Advancing Multifamily Decarbonization Through Field Demonstrations of Emerging Heat Pump and Heat Recovery Technologies

Neil Donnelly, New Ecology, Inc., Sankhanil Goswami, New Ecology, Inc.
John Beaumont, New Ecology, Inc., Steve McCarthy, New Ecology, Inc.

ABSTRACT

Manufacturers of new technologies offer the promise of carbon reduction and progress towards programmatic goals. Vetting new technologies outside of a laboratory environment is critical to widespread adoption by utility-sponsored energy efficiency programs.

This paper presents the design and early implementation of four multifamily field demonstration projects that evaluate key challenges, installation considerations, and pathways to scale emerging electrification technologies. Each project represents a distinct strategy for decarbonizing thermal systems in existing buildings.

Project A compares central CO_2 heat pump water heater installations at two sites to evaluate pre-installation load measurement for system sizing and characterize operational performance through continuous monitoring.

Project B investigates drain-water heat recovery, integrated with water-source heat pumps to capture wastewater energy for domestic water heating, exploring control logic, system configuration, and detailed flow-metering for design optimization.

Project C involves installation of saddle-style heat pumps across multiple apartment units, examining installation logistics, time-study methods, and tenant comfort monitoring to evaluate pathways for large-scale adoption.

Project D demonstrates replacement of a large central chiller with an air-to-water heat pump system operating in parallel with existing boilers under a hybrid control strategy, focusing on design coordination, equipment sizing, and integration of monitoring infrastructure.

Together, these projects highlight the critical role of field-based demonstrations in bridging design intent and operational performance, establishing groundwork for scalable, data-driven decarbonization in multifamily buildings.

Introduction

Manufacturers of emerging building energy technologies offer compelling promises of carbon reduction, efficiency gains, energy savings, and tenant comfort. Yet the pathway from new product to reliable, replicable, and scalable systems for deployment is littered with failure points that consistently undermine these promises, limiting the reach and speed of adoption of technologies that can help meet energy reduction goals. Utility-sponsored energy efficiency programs face pressure to adopt new technologies, but the gap between laboratory ratings and real-world operation creates substantial risk for building owners investing capital and for tenants whose comfort depends on system reliability.

As emerging technologies progress from product development toward field deployment, whether delivered as packaged units or site built engineered systems, they encounter a recurring sequence of failure points that can degrade performance well below design intent (Roth et al.

2024). Technologies and buildings are often matched arbitrarily: an owner or management company discovers a technology and brings it to a utility efficiency program, instead of technologies being brought to the table and then matched with an ideal building for demonstration. Poor assumptions about building loads lead to faulty system designs mismatched between system and loads. Inexperienced contractors, unfamiliar with emerging equipment, produce installations that deviate from specifications, deviations that go unnoticed until commissioning or much later. Commissioning personnel, equally unfamiliar with what "right" looks like for these systems, overlook construction errors and configure controls incorrectly. By the time the system is operational, compounding errors have degraded performance well below design intent, and the ability to rectify these errors becomes more limited and costly. Even when systems initially meet goals, the question remains: will those goals be maintained over time without continuous visibility into system operation, and how replicable are the results to a wider range of buildings?

This pattern matters because success in decarbonization programs is too often measured by installation counts rather than verified performance and satisfied owners and tenants. Many reported savings are claimed or deemed not measured. Experience from previous monitoring efforts (Sklarsky, J., et al. 2025) has shown that even seemingly accurate savings estimates frequently fail to hold up under scrutiny or beyond the first year. When programs scale technologies without clearly documented best practices, and based on flawed assumptions compounded across design, installation, and commissioning, the result is a portfolio of underperforming systems and eroded confidence in specific technologies and electrification as a viable pathway.

A typical project follows a standard sequence: define goals, detail scope, design the system, bid on the project, construct and install, and commissioning. At each stage, failure points can compound (Figure 1). This paper argues that rigorous field demonstration, grounded in site selection, pre-installation measurement, construction and commissioning QC, and continuous performance monitoring, can interrupt this failure cascade at multiple points and create the keys to success for replicating the measure. Measuring existing conditions before system design

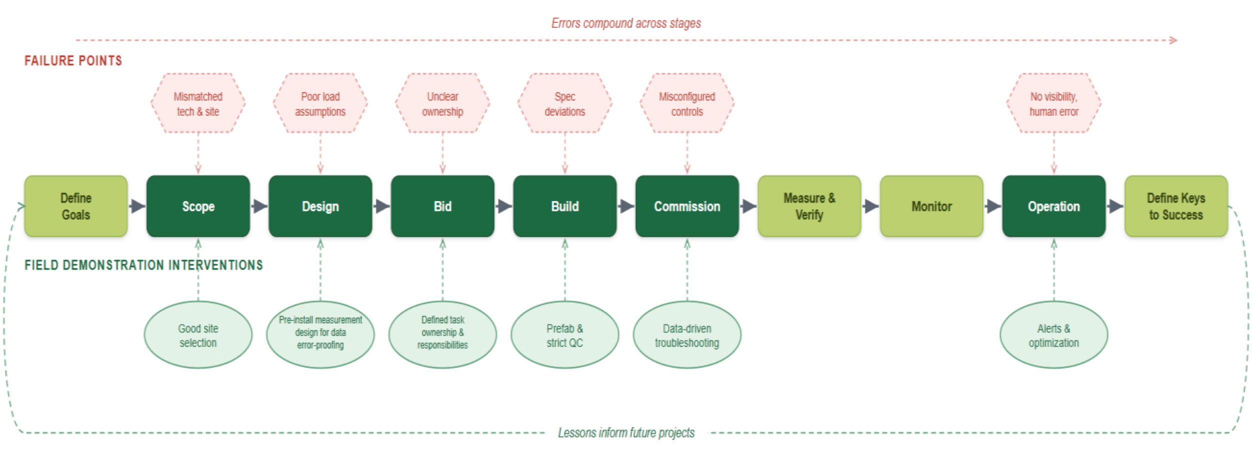

Figure 1. Project roadmap and failure points

addresses the root cause: poor assumptions. Data-driven site selection identifies buildings were emerging technologies are most likely to succeed. Detailed monitoring infrastructure enables commissioning personnel and the project team to diagnose problems that would otherwise go undetected, pushing the project beyond verification to optimization.

A successful project provides its intended service reliably and at the expected efficiency, with failures that can be readily diagnosed and fixed, which requires the operational visibility most installations lack. This paper presents four multifamily field demonstration projects, each a distinct strategy for decarbonizing mechanical systems in existing cold-climate buildings, that together show how measurement, verification, and continuous monitoring interrupt the failure cascade and lay groundwork for scalable, evidence-based decarbonization.

Project A: Central Heat Pump Water Heaters

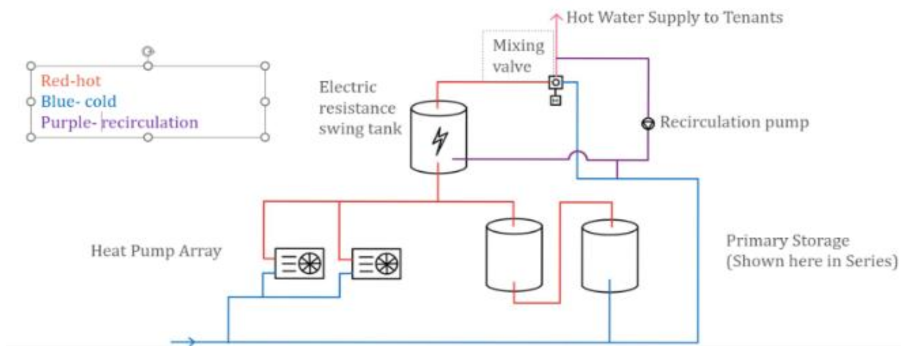

Figure 2. Simplified central CO_2 heat pump water heater system

Project Context

This project compares early implementations of central heat pump water heater (CHPWH) installations at two sites to evaluate the importance of baseline and ongoing data collection for system design and operation, defining keys to success and characterization of operational performance. The comparison illustrates how lessons from early installations can improve outcomes at subsequent sites when continuous monitoring enables iterative learning. A line diagram of the installed CHPWH is shown in Figure 2.

Location W is a high rise of 132 units of senior housing originally served by an electric resistance domestic hot water system and replaced with an array of 10 Sanco heat pumps (combined rated output of 154,000 Btu/hour) and 1,000 gallons of primary storage in series, and a swing tank. Recirculation is served through two loops, serving floors 4-10 and floors 1-3.

Location S is a 2-story 98 unit senior housing development originally served by condensing gas-fired water heaters and storage. The CHPWH system was 12 Sanco heat pumps (combined rated output of 184,800 Btu/hour) and 1,000 gallons of primary storage in series, and a swing tank.

Pre-measurement of domestic hot water loads and recirculation loads was completed at Location S to inform system sizing and to evaluate recirculation losses. This analysis was prompted by observations at Location W, where high recirculation losses caused a disproportionate share of electricity consumption to occur at the swing tank.

Pre-measurement data suggested that both total heat pump capacity and storage volume were oversized by a factor of two. However, because the project as scoped by the manufacturer's representative had passed the cost-effectiveness evaluation required by the utility efficiency program, the decision was made to proceed with the full design. The rationale for this decision centered on the modular nature of the system: an array of 12 heat pumps with oversized storage was expected to operate efficiently while providing additional benefits, including greater load-shifting capability and increased resilience during equipment malfunctions due to the larger storage volume.

Proceeding with the manufacturer's full design also aimed to reduce potential disputes later in the project. By deferring to the manufacturer's sizing, the project team sought to limit finger-pointing among the plumber, manufacturer's representative, and manufacturer when commissioning and operational issues inevitably arose.

Significant piping and control issues arose at the first installation site, Location W, which required multiple follow-up visits by the plumber during early startup and commissioning. Disputes between the manufacturer, manufacturer's representative, and plumber regarding discrepancies between the design documents and as-built conditions informed a set of lessons for overseeing the physical design and construction of the second system at Location S. Based on this experience, the project team applied close quality control at Location S, with explicit attention to pipe sizing, valve orientations, reverse return piping, and the location of each system connection. This strict QC enabled the plumbers work to be completed and signed off as correct without additional callbacks, enabling commissioning to proceed with the planned data, and identify a series of controls issues that could have been blamed on inaccurate construction.

Data Collection and Analysis Plan

System data streams for monitoring and optimization were collected as follows:

- Ultrasonic flowmeters and temperature kits to measure the domestic hot water demand, recirculation losses, and the isolated energy flows from the heat pump array
- Circuit level monitoring of the heat pump array, swing tank and controls
- System temperature monitoring through the central controller at Location S. This capability was then made remotely accessible after the first year of monitoring and optimization to support more detailed observation and optimization.

The analysis for these systems was based on the calculation of a daily system coefficient of performance (COP) from measured loads (DHW and recirculation) and input energy (electrical use of HP & swing tank) as defined in the advanced water heating specification. Specific analysis into the division of energy used by the heat pump array and the swing tanks were developed as part of the optimization once elevated swing tank operation was observed at Location W.

Designing data collection from the outset of the project enabled remote monitoring of electricity consumption at the heat pumps and swing tank, as well as temperature data from the central controller. The accessibility of this data, combined with the intent to use it for commissioning verification, exposed failures and suboptimal control decisions that had resulted in primary storage being unheated, overheated, or fully depleted when heat pumps entered lockout. Once temperature data was integrated into the remotely viewable dashboard, the feedback cycle accelerated, allowing issues to be identified daily rather than over the course of weeks. Initial results from system optimization are shown in Figure 3; these results represent a midpoint in the optimization process before remote temperature data was made available for the current round of optimization targeting heat pump staging and load shifting. Time of use billing is not in place for these buildings, so load shifting is being pursued to shift heat pump run-time into warmer parts of the day to boost the daily system COP. The oversizing of the heat pump array and primary storage volume both support this load shifting capability.

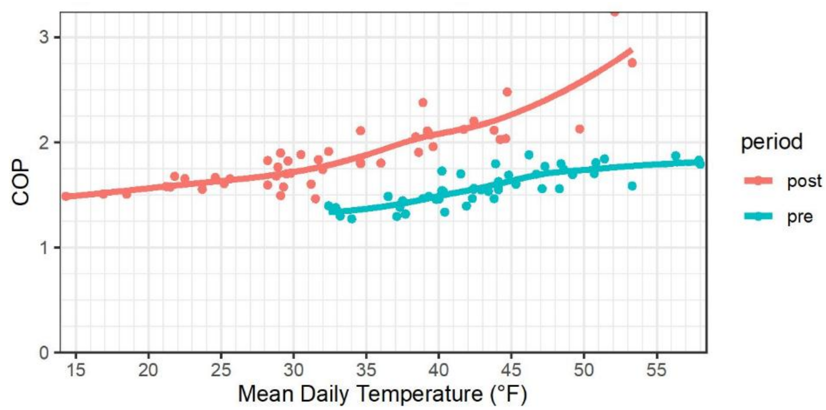

Figure 3. System COP vs outdoor air temperature

Lessons Learned

System commissioning at Location S was completed by the manufacturer's representative and observed by the project team. The project team intentionally refrained from intervening during commissioning to identify what failures would emerge without close oversight and a manufacturer's representative completing commissioning while still being very new to the system. This approach exposed a lack of coordination and unclear ownership of controls implementation and setpoints.

Controls implementation failures occurred at three levels: the central controller was not correctly wired to the heat pumps, the heat pumps were not configured to accept commands from the central controller, and the central controller was not configured to issue commands to the heat pumps. The persistence of these failures through commissioning highlighted breakdowns in responsibility, coordination, and verification at multiple points in the project delivery chain

Control setpoints were also identified as a failure point. Following commissioning, there had been no discussion of efficiency-oriented setpoints for heat pump staging or swing tank operation, making clear two additional gaps in responsibility for setpoint coordination (and education) and ownership of a defined efficiency goal to be held to.

Following commissioning, the swing tank setpoint was left at 135°F. This tank has a fixed differential of 8°F, meaning the 10 kW electric resistance element would activate whenever the swing tank temperature dropped below 127°F. With the heat pump array self-modulating to maintain a target outlet temperature of 145°F, the setpoint difference between primary storage (maintained at approximately 140°F) and the swing tank setpoint of 135°F created conditions for significant and unnecessary resistance element operation, as shown by the swing tank energy usage (blue) in Figure 4, regardless of how much hot water was being used in the building.

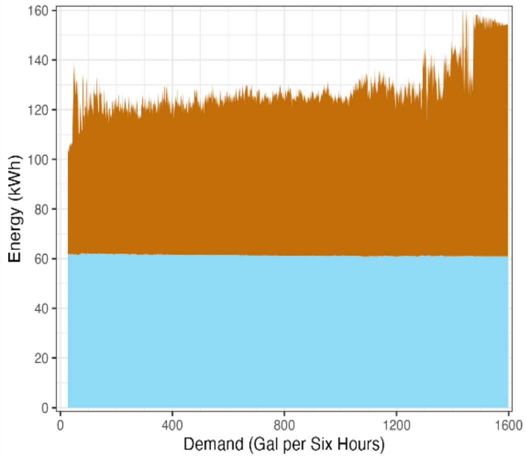

Figure 4. Energy use breakdown, with swing tank setpoint of 135°F

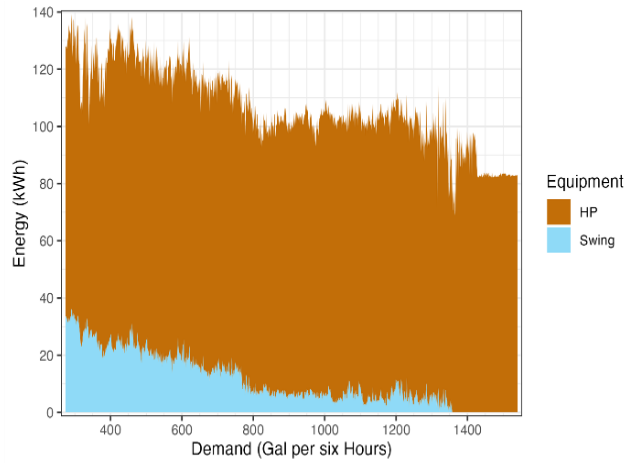

Figure 5. Energy use breakdown, with swing tank setpoint of 128°F

These errors in setpoint coordination, particularly around swing tanks, are endemic to systems with swing tank designs, and presents a significant loss in COP and increased operational costs for many systems that can be easily fixed. A lack of understanding of how swing tanks are intended to operate, and how simple setpoint coordination can improve performance, represents a significant but easily solvable issue.

The swing tank setpoint was identified right away by the team and corrected after confirming that the manufacturer's representative considered commissioning complete and following a brief monitoring period to quantify the impact of the misconfiguration. Left uncorrected, this issue was significantly increasing system electricity consumption. The remedy required a single setpoint adjustment: reducing the swing tank setpoint to 128°F so that the heating element activates only when tank temperature drops below 120°F and the energy usage is shown in Figure 5. This change reduced daily swing tank electricity use by 71%. Designing for visibility into system operation and energy usage enabled this issue to be identified right away and have the impact quantified. In this project, this immediate demonstration of the value of monitoring kept the project team at the table to push on to deeper issues of setpoint optimization.

Project B: Drain Water Heat Recovery

Technology Overview

This project investigates the efficiency and cost-effectiveness of multiple drain-water heat recovery (DWHR) system designs within a multifamily building portfolio.

Drain-water heat recovery technology is relatively simple and well established, particularly for vertically oriented units. In Europe, larger centralized horizontal systems have also demonstrated strong cost-effectiveness, even in regions with very low energy prices.

The four levels of drain-water heat recovery discussed in this project are summarized below. Level 1 and Level 2 systems were implemented as part of the current project scope:

Level 1: Passive heat recovery: A fully passive design where heat from wastewater in the drain is transferred directly to incoming cold potable water through a heat exchanger. Level 1 systems were installed in both vertical and horizontal configurations, primarily to temper incoming cold-water supplies located near vertical drain stacks or along sections of horizontal drain piping. The pre-heated water is used as makeup to the domestic hot water system or to temper all water entering the building.

Level 2: Active heat recovery with recirculation: Level 2 systems utilize longer horizontal prefabricated drain-water heat recovery units with factory-built fittings. A temperature-controlled circulation pump, small storage buffer, and temperature controls are added to the potable water side as shown in Figure 6 and Figure 7. The buffer, a large-diameter copper pipe section, is installed either on the main cold-water feed or the DHW makeup line, depending on the building's plumbing layout.

Figure 6. Installed Level-2 Drain Water Heat Recovery System

Figure 7. Interior View of Level-2 Drain Water Heat Recovery System

The circulation pump allows the system to harvest additional heat even when wastewater discharge and domestic water demand are not synchronized. When the control system detects sufficient thermal energy available at the heat exchanger, it increases pump flow to extract and store heat in the buffer. The pre-heated water is used as makeup to the domestic hot water system or to temper all water entering the building.

Level 3: Drain-source heat pump for domestic hot water: The incoming cold-water loop is replaced with a closed-loop heat pump circuit. Heat is transferred from the drain water to this loop via the drain-water heat recovery unit. Water-source heat pumps then extract heat from the closed-loop circuit to heat domestic hot water.

Level 4: Drain-source heat pump for heating, cooling, and domestic hot water: Additional water-source heat pumps are added and networked together to serve multiple building loads. The system can simultaneously support heating, cooling, and domestic hot water, with a central controller managing the distribution of recovered energy between loads.

Four of the twelve buildings in this project received prefabricated "Level 2" drain-water heat recovery systems, among the first installed in North America. Two of the largest were equipped with advanced monitoring infrastructure to support detailed performance evaluation.

These two larger installations were also intentionally oversized to be "heat-pump-ready," allowing for potential future upgrades to Level 3 or Level 4 drain-source heat pump systems as additional funding becomes available.

Site Selection

The buildings in this clustered portfolio were initially identified during routine building audits with many sharing a key characteristic: large, open basements with exposed sanitary drain piping. Their proximity to one another was itself a selection factor, as having multiple sites within close range allowed the manufacturer and designer to provide hand-on oversight throughout the construction and commissioning, which was critical given that these were the first Level 2 installations in North America.

Of the twelve buildings included in the project, four were selected as the best candidates for Level 2 drain-water heat recovery systems. These buildings had plumbing configurations that allowed the sanitary drains to be consolidated and re-piped into a single location before exiting the building. Key selection criteria included:

(i) The ability to combine all sanitary drain lines into a single centralized drainpipe

(ii) Adequate drainpipe size to accommodate total building wastewater flow

(iii) Proximity of the drain location to the building's main domestic water supply and domestic hot water system, as applicable to the proposed design.

These conditions made it possible to install centralized heat recovery systems capable of capturing heat from the majority of wastewater, leaving the building using a prefabricated, easy-to-install system. The remaining eight buildings received Level 1 systems matched to their individual plumbing layouts. The intentional variation in Level 2 configurations across the four sites, that is pre-tempering of all water entering the building vs pre-tempering just the makeup water to the DHW system, was designed to build a diverse dataset and develop the analysis tools and decision framework needed to scale these systems to future buildings.

Data Collection and Analysis Plan

Advanced monitoring systems were incorporated into the design to measure system performance, verify recovered energy, support optimization of thermal controls, and inform laboratory testing and modeling with actual performance and load data.

The system is a heat exchanger with buffer capacity on the potable water side. Temperature and flow measurements on the potable side allow detailed analysis of energy recovered, however the measurement of the drain side flow and temperatures required more expensive equipment and were not able to be included in this project to have fully defined energy flows on both sides of the heat exchanger. The variety of DWHR implementations (DHW makeup only, all-building flow) enabled building a diverse dataset, driving the analysis tools needed to characterize each design, and feeding a decision tree for selecting and scaling these systems in future buildings

Due to the rapidly changing flow and temperatures in this potable system, direct immersion RTDs (resistance temperature detectors) were used for faster response time.

By integrating monitoring requirements into the design and prefabrication process, the project team achieved efficient installation and rapid deployment of data collection systems. This allowed system performance data to be available quickly for commissioning and early analysis of heat recovery performance.

Analysis is currently underway to develop savings calculations and models to represent the different implementation methods (level 1/2, DHW pre-temper, whole building tempering) and to incorporate the longer term trends of seasonality of water temperatures, building temperatures and consumption patterns, with the initial rounds of analysis showing savings of more than 30% of daily domestic hot water energy use.

Lessons Learned

The pre-fabricated central horizontal systems are the first of their kind in North America, and thus require more feedback into system design and iteration than a demonstration project working out keys to success for scaling adoption. This puts a different focus on the commissioning and optimization phases and targets different outputs to provide feedback to the system and control design. This is being seen with larger buffer volumes in the next generation of designs, allowing more efficient heat transfer with colder potable water entering the heat exchanger, and being able to store more energy to further decouple the available drain energy from the instantaneous potable demand.

The following incident demonstrates how continuous monitoring directly supports building operations by catching failures that would otherwise go undetected outlined in the project roadmap (Figure 1). In one case, someone unknowingly closed a valve that cut potable flow through the heat exchanger. Because of the Level 2 design, water still reached the building, so no heat recovery was happening, but no tenant ever felt it. The team spotted the drop in the standard monitoring dashboard and resolved it over the phone with on-site maintenance. This failure mode recurs, a well-meaning contractor or maintenance person closing a valve while addressing something else, and unlike a heating or domestic hot water fault it produces no complaint to flag it. Pre-tempering systems like this therefore need automated alerts, since the fallback otherwise is noticing an unexplained utility-bill increase months later. This is a great example of specifically identifying significant operational failure modes which will not result in complaints and setting up automated guard-rails.

These Level 2 systems occupy an important position within a phased decarbonization pathway. They provide a standalone early-adoption measure that reduces domestic hot water loads with only a circulation pump as a moving part, while also being sized and configured to support future Level 3 or Level 4 implementation. As a result, the systems are being evaluated both as completed Level 2 installations and as measurement platforms for future drain-source heat pump applications.

The value of this approach is that the project is already collecting the kind of information that would otherwise be needed later, after a more complex system is considered. By measuring actual wastewater heat recovery potential, testing control strategies, and identifying operational issues now, the project can reduce the risk of poor assumptions or design mistakes in future phases. This gives the building a stronger starting point for upgrades, with limited added cost in the current project, and supports a decarbonization pathway that can move forward in steps rather than relying on a single deep energy retrofit

Project C: Saddle-Style Heat Pumps in Public Housing

Figure 8. Interior View of Installed Saddle Style Heat Pump

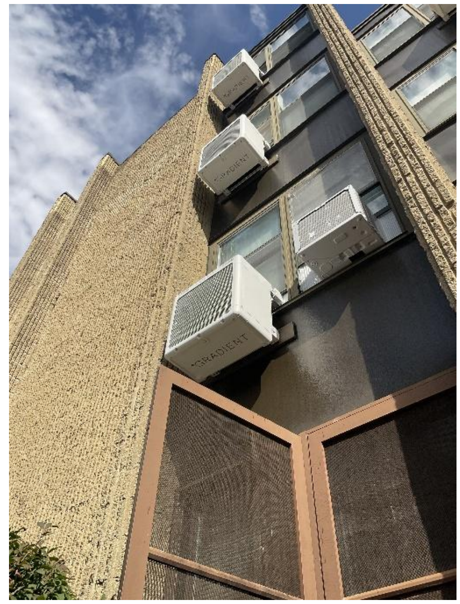

Figure 9. Exterior View of Installed Saddle Style Heat Pumps

Project Context

This project installed saddle-style heat pumps in apartment living rooms. The study examines installation logistics, coordination of unit-level controls in a partial displacement project, and tenant comfort monitoring to evaluate practical pathways for large-scale adoption in public housing settings, where split incentives, tenant coordination, and building constraints create unique implementation challenges.

The Boston Housing Authority (BHA) supported the New York City Housing Authority (NYCHA) in soliciting the development of a new saddle-style retrofit heat pump. These units have been demonstrated in New York City(NYSERDA 2023) and are now being deployed in supporting cities. BHA selected a building from their portfolio with existing electric resistance baseboard heat for the pilot installation.

The project installed one saddle-style heat pump per apartment across a mix of studio, one-bedroom, and two-bedroom units. Existing electric baseboards were retained with new data-connected thermostats added to manage their operation. Heat pumps were installed in living rooms only, as bedroom windows were not large enough to accommodate the units (Figure 8, Figure 9).

Implementing this as a partial displacement of existing electric resistance heat introduced competing controls vying for setpoint dominance creating a challenge which amplified across 100 apartments, each presenting an opportunity for dueling controls between heat pump and baseboard thermostats. Connectivity posed an additional challenge: the four-story solid concrete building had no existing central Wi-Fi, complicating data stream networking and requiring follow-up troubleshooting visits even after early-stage signal assessments.

Installation logistics for this project were shaped by both the novelty of the technology and the scale of the deployment. Introducing a new system to a local market inflates installation pricing due to contractor unfamiliarity. The repetitive apartment-by-apartment nature of this project turned this into an advantage where time studies tracked discrete installation tasks across units, building installer proficiency and a documented benchmark for future projects to encourage realistic contractor bidding.

Commissioning with Distributed Partial Displacement

During commissioning of data streams, the thermostats and heat pump switches were closely reviewed. Using the remote thermostat dashboard to identify faults for follow-up, the share of thermostats reporting to the central hub increased from 72% to 95%, with the remaining 5% identified as having pre-existing power issues that are currently being fixed by BHA's electricians.

This process exemplifies the data pipeline stages of the project roadmap (Figure 1), and highlights where commissioning failures commonly occur. Coordination of installation, data connection, configuration, and commissioning each required distinct effort and team ownership to produce functional data streams. Data connection involved planning and installing the infrastructure to route monitoring streams back to the project team for remote access. Configuration required implementing settings on individual components including heat pumps, smart switches, and thermostats to ensure connectivity-capable equipment was actually communicating with the network. Thermostats were installed but not configured during the same visit, requiring follow-up appointments and additional tenant disruption to enable data collection and remote setpoint control. Clearly defining each step and the responsible party at the outset will improve project budgeting and clarify when each data stream is ready for handoff in future projects.

With installation of the heat pump systems and the new connected thermostats for electric baseboard, these two competing controls were initially installed without setpoint limits to allow time for identifying and troubleshooting installations. The team's visibility into apartment-level systems at this point was limited to space temperature information and tenant thermostat setpoints available through the connected thermostats, along with electrical usage data from the smart plugs installed with the heat pumps. When most of the project team considered the project complete with heat pumps installed and operational, there was still no coordination in place between the two control systems. This put all control in tenants' hands to independently set thermostat setpoints and manage their use of the heat pumps. Initial adoption and use of the heat pumps was below 10% at this stage.

Heat pump adoption was determined through analysis of thermostat and smart-plug data to determine the dominant method of heating, with apartments sorted into 'heat pump', 'baseboard', 'mixed' or 'none' based on the total usage of each system for each day. Quantifying this low heat pump adoption early on with direct measurement, rather than months later with utility bill analysis allowed targeted follow-up through thermostat limit roll-outs and expanded tenant education strategies.

Thermostat limits were coordinated with tenant education sessions, and one-pagers on operating the heat pumps were updated to simplify and focus the teaching-points on the very basics of use. Once these updated materials were translated, distributed and multiple education sessions had been held, the online thermostat dashboard was used to decrease setpoint limits available for electric baseboard. Limits were reduced from 85F down to 70°F in living rooms, 72°F in primary bedrooms, and 74°F in secondary bedrooms as an initial reduction to evaluate savings, tenant experience, and the ability of the heat pump to condition adjacent bedroom spaces.

With apartments sorted into their dominant heating system on a daily basis, and the ability to sort apartments with very low (non-compressor) power draws on the heat pump, apartments which were baseboard dominant could be further split into ‘ready for HP’ or ‘not ready’. ‘Ready for HP’ apartments showed some power draw at the HP, signifying that the unit was powered on, and in most cases was in heating mode, but with a heating setpoint below the baseboard thermostat heating setpoint. These heat pumps were ready to be pushed into the dominant category in most cases with more tenant education about setpoints and limiting the baseboard setpoint limits further.

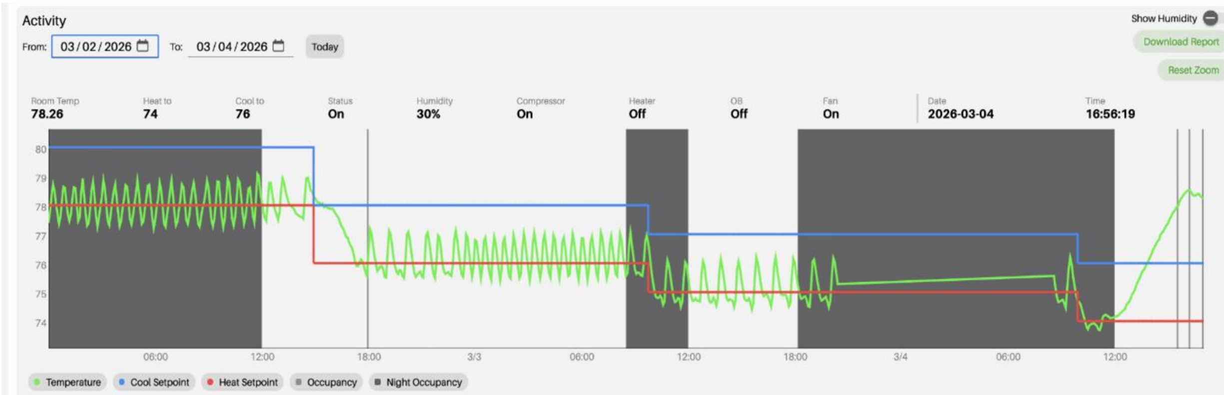

Figure 10. Thermostat data, showing max setpoint reductions and heat pump adoption

Figure 10 shows three days from the baseboard thermostat dashboard during the initial setpoint reductions. As the baseboard limit (red) is lowered each day, the space temperature (green) follows it down in the cycling sawtooth of baseboard operation, until the limit drops below the heat pump setpoint and the heat pump takes over. The space temperature then rises above any previous baseboard setpoint, so the tenant stays comfortable on less energy.

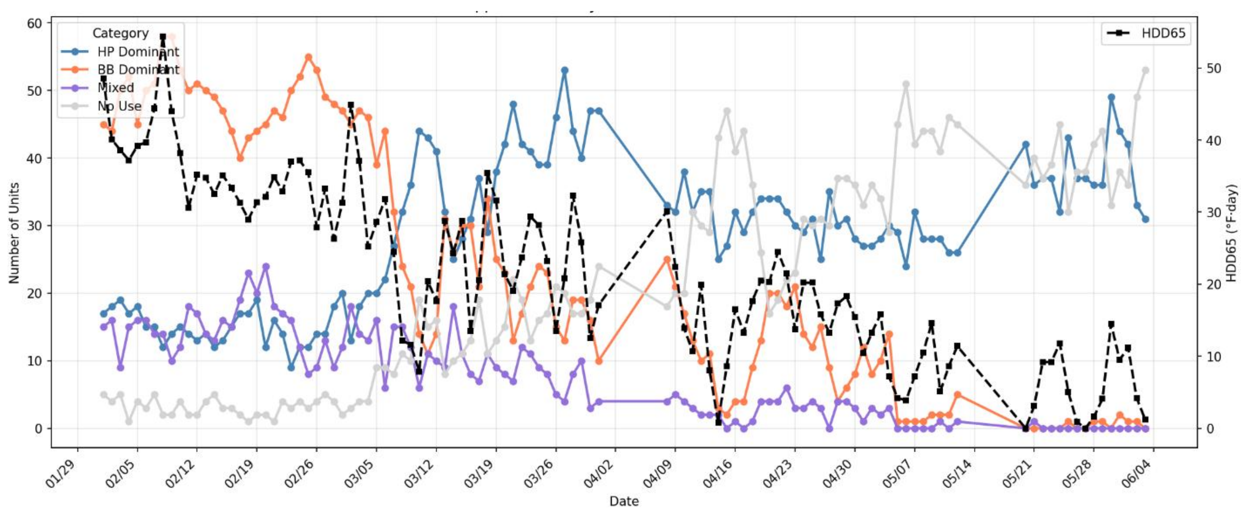

Figure 11. Unit-Level heating behavior over time: heat pump vs baseboard dominance with heating degree days(HDD)

By keeping the heat pump's maximum setpoint at 85°F while capping the baseboard lower, the strategy makes the heat pump the only way to warm a room further, shifting tenant preference toward it. The first reduction pushed heat pump adoption above 40% by early March (Figure 11) but left setpoints able to compete below 70°F. After clearing remaining heat pump issues at a few apartments, the living room baseboard was limited to 65°F with a 20°F differential, effectively disabling it unless the room drops below 45°F. Implemented April 15, this drove a further jump in heat-pump-dominant operation. The same approach extends to the bedrooms at the end of the required heating season (September 15 to May 31 in Massachusetts), disabling the baseboard for the summer for the first time and preventing simultaneous heating and cooling.

Data Collection and Analysis Plan

System data streams for monitoring and optimization were collected as follows:

- Heat pump power usage through smart plugs via a dedicated mesh network to three hubs distributed between floors although the heat pump units themselves have not yet been networked to allow remote fleet management and visibility.
- Connected thermostats communicate via a dedicated mesh network to a single central hub, providing space temperatures, setpoints, baseboard run-time and remote control
- Circuit level monitoring at two selected apartments to measure baseboard power draws and to help true up smart plug measurements using e-gauge.

Figure 12 shows the resultant daily energy consumption curves for the building when normalized and extrapolated from the subset of apartments operating as heat pump or baseboard dominant throughout the monitoring period so far. Due to the slow adoption of heat pumps, high variability in individual apartment heating set points and distributed setpoint controls, visualization at to the building level cuts through the individual noise of apartment level data to evaluate the measure's savings at the building level.

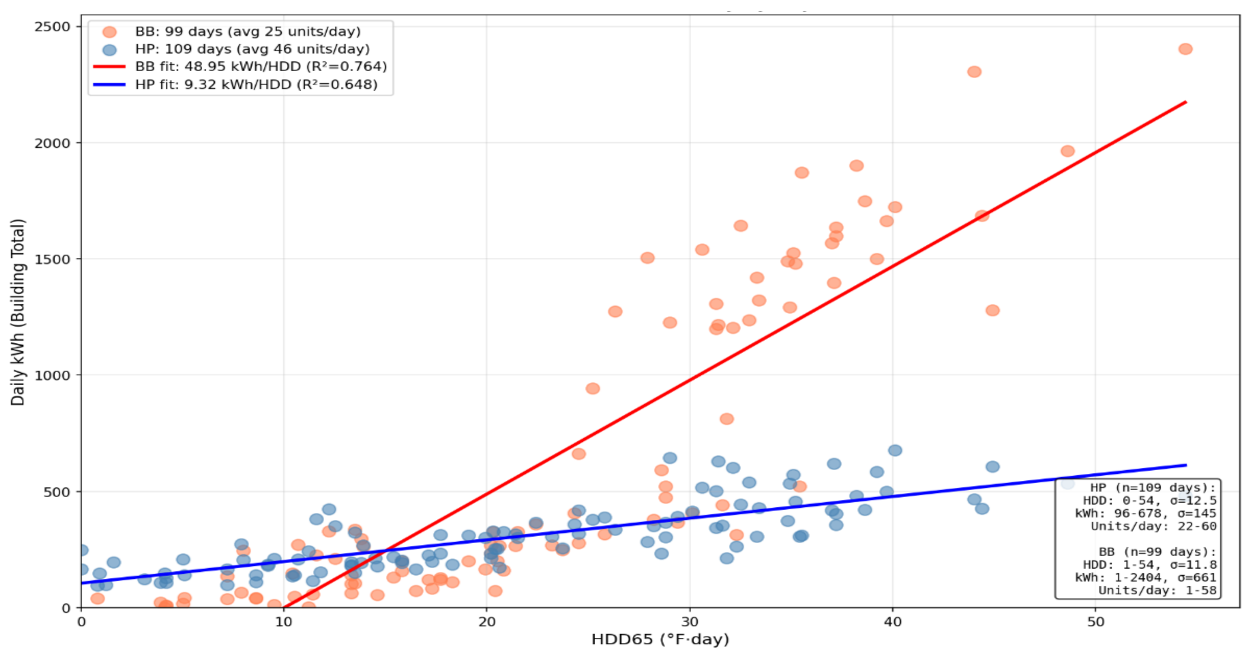

Figure 12. Building-Level heating demand vs. heating degree days: baseboard and heat hump comparison

Lessons Learned

In our previous experience (Roth, K., and M. Kromer. 2024) with centrally controlled heating and hot water systems monitored and optimized years into their operation, more than 80% of buildings were identified as able to have setpoints reduced without impacting tenant comfort, and more than 50% were identified as serving one load (heating) with a heating water setpoint intended for a different load (domestic hot water). These setpoint coordination are a significant failure point, exemplifying the need for error-proofing systems and controls as much as possible, to ensure ongoing system performance. This issue is compounded within this project with the elements of partial displacement and distributed controls.

Partial displacement projects bring control setpoint and coordination to the highest level of importance, with an elevated opportunity for failure, as the coordination of multiple systems would fall outside the typical oversight of contractors and equipment reps focused on their own scope, and not used to working within a partial displacement solution. The final action of disabling baseboard use in the living rooms by implementing a 20°F setpoint differential was required to error-proof the system coordination.

Distributed controls, baseboard thermostats and onboard heat pump controls (both of these controls being new to tenants as well), elevate the potential failure of partial displacement to near certainty. Tenant Education was a core focus after the project was installed, with a series of in-person presentations and materials prepared and translated, coordinated with the building wide thermostat limiting setpoints. In the end, tenant education did not significantly impact the adoption and use of heat pumps but served to build trust and reduce the pushback when heating setpoints were limited and disabled.

Data commissioning was a failure point at the design and bid stages, where responsibility for configuring the thermostats and smart plugs and delivering a working dashboard was never clearly assigned, forcing follow-up visits to finish what handoff should have covered.

Project D: Air-to-Water Heat Pump Integration

Project Context

This project demonstrates the replacement of a large central chiller at the end of its useful life with an air-to-water heat pump system that provides full cooling capacity while also displacing the existing gas-fired boiler plant for heating above a defined switchover temperature. This partial-displacement approach enables significant electrification of building heating loads while retaining the existing boiler plant as backup capacity during extreme cold conditions.

The building is currently served by modulating, condensing gas-fired boilers for heating and a rooftop chiller for cooling. The hydronic distribution system operates as a two-pipe configuration with seasonal changeover between heating and cooling modes. Fan-coil units serve as terminal equipment throughout the building. The recently installed high-efficiency boilers, the dual-temperature distribution well suited to heat pump integration, and the aging chiller together made the building a strong candidate for a replicable replacement model.

The two heating plants were designed to operate in series. Return water from the building first passes through a heat exchanger served by the heat pumps, then reaches the boilers through closely placed tees downstream. The heat exchanger is sized for full building flow, and an automated bypass valve routes flow around it when building switches to boiler operation. The three boilers are piped primary-secondary, each with a dedicated pump tied into the building loop through closely spaced tees. This lets the heat pumps carry moderate-weather heating while the boilers assume the load once outdoor temperature drops below the switchover threshold.

Site Selection

The building was selected from facilities within the funding utility's shared gas and electric service territory, then screened on system characteristics and electrical capacity. Its dual-temperature hydronic system with fan-coil units suited air-to-water heat pumps, connected thermostats already supplied interior temperature data for monitoring at no added cost or tenant disruption, and the existing electrical service could support the heat pumps without an upgrade. The team also considered a second heat-only pilot in a building with fin-tube radiation, but most candidates lacking a chiller would have needed costly electrical service upgrades, which made that option unattractive. The screening and the load measurement below illustrate the site selection and pre-installation measurement interventions in Figure 1.

Data Collection and Analysis Plan

Prior to equipment selection, the team directly measured the building's heating loads with ultrasonic BTU meters. These measurements established actual heating demand, informed heat pump sizing, and supported energy savings projections, fixing the capacity needed to displace gas heating cost-effectively down to the switchover temperature.

A comprehensive monitoring system was then installed to verify performance and energy impacts. Permanent electromagnetic flow meters serve the building hydronic loop, the heat pump loop, and the boiler manifold inlet, each paired with supply and return RTDs, with an additional RTD downstream of the heat pump heat exchanger and upstream of the boiler tees for redundancy. Electrical monitoring covers the heat pumps on both sides of the 208-to-480-volt step-up transformer, isolating heat pump power from transformer losses, along with the VFD building circulators, the perimeter pumps serving stairwells and common areas, and the heat pump loop pumps. Modbus connections to the heat pumps and an outdoor air temperature sensor mounted at the equipment complete the dataset. Together these streams resolve the thermal contribution of the heat pumps and boilers, electrical consumption, and control optimization.

Lessons Learned

Project D demonstrates the front-end of the Figure 1 roadmap. Choosing the building with data-driven site selection and measuring its loads directly prevented the two failures that begin most cascades, a mismatch between technology and building and design based on poor load assumptions. The project is currently delayed due to the discovery of deteriorated structural roof beams beneath the planned heat pump mounting location. Structural repairs are required before heat pump equipment can be installed.

While this delay has postponed installation, it has created an opportunity to collect additional space conditioning load data using the monitoring equipment already in place. This equipment will remain operational once the heat pumps are installed. Extended monitoring during the current heating season will further refine load profiles and improve understanding of building performance prior to system commissioning.

Discussion

Field demonstration matters because the cost-effectiveness models that approve these technologies rest on assumptions a real building rarely honors. Across the four projects the same failures recur, mismatched site, poor load assumptions, unclear controls ownership, and setpoints misconfigured at handoff, which is why Figure 1 frames them as a single roadmap rather than four separate stories.

Two practices address the early stages. Measurement must precede design, since the load data that sized the heat pumps in Project D and flagged the twofold oversizing of water heater storage in Project A cannot be recovered afterward. Verification then has to be designed into the system rather than bolted on, as the prefabricated heat recovery units in Project B and the monitoring specified from the outset in Project A both shows.

Implementation fails most often at ownership and commissioning. Partial displacement in Project C put coordination of distributed setpoints at the center of the work, and Project A passed mis-wired controls because responsibility was split across the plumber, representative, and manufacturer. Continuous monitoring is what catches these faults, shortening the feedback cycle from weeks to a day and surfacing the swing tank misconfiguration whose correction cut daily resistance use by 71%.

Conclusion

Field demonstration turns predictable failure points into solvable ones. The shared lesson across these projects is that measuring before design, instrumenting for verification, and monitoring continuously belong in standard practice, not as optional add-ons. Adopting them lets utilities and housing authorities judge programs by verified performance rather than installation counts.

References

NYSERDA (New York State Energy Research and Development Authority). 2023. "Window Heat Pump Units Installed in Public Housing as Part of Clean Heat Challenge." nyserda.ny.gov/About/Newsroom/2023-Announcements/2023-09-20-Governor-Hochul-Announces-Installation-Of-Window-Heat-Pumps-For-New-York-City.

Roth, K., and M. Kromer. 2024. "Using Connected Boiler Data to Accurately Quantify Overheating and Energy Savings from Outdoor Air Reset (OAR) Curve Changes in Multifamily Buildings." In *Proceedings of the 2024 ACEEE Summer Study on Energy Efficiency in Buildings*. Washington, DC: ACEEE.

Sklarsky, J., K. Roth, M. Kromer, N. Donnelly, and S. Goswami. 2025. *Optimizing Hydronic Heating for* Comfort *and Performance in Multifamily Housing*. Boston, MA: New Ecology. www.osti.gov/biblio/2499544.

Wehbi, Z., R. Taher, J. Faraj, M. Ramadan, C. Castelain, and M. Khaled. 2022. "A Short Review of Recent Studies on Wastewater Heat Recovery Systems: Types and Applications." *Energy Reports* 8 (9): 896–907.