Comfort First, Flexibility Next: Grid-Interactive Water Heating That Puts People Front and Center

Caton Mande, UC Davis Western Cooling Efficiency Center
Christy Green, UC Davis Western Cooling Efficiency Center
Loren dela Rosa, UC Davis Chemical Engineering
Matt Ellis, UC Davis Chemical Engineering

ABSTRACT

Comfort must come first in the transition to grid-interactive water heating. This paper presents laboratory and field validation of a cloud-based supervisory model predictive control (MPC) for hybrid heat pump water heaters (HPWHs) that prioritizes hot-water availability while co-optimizing electricity cost and greenhouse-gas (GHG) emissions.

Laboratory testing established an upper bound on performance, with ~15% utility-cost reduction, up to 61% marginal-emissions CO_2 reduction, and ~79% reduction in 4–9 p.m. peak runtime relative to manufacturer rule-based control (RBC).

A field demonstration across 17 low-income multifamily households in two Northern California climate zones (CZ4, CZ12) used a randomized rapid measurement and verification design to isolate impacts of weather and occupant behavior. The supervisory MPC reduced peak-period energy use by 31% in CZ12 ($p=0.01$) and 20% in CZ4 ($p=0.08$). Field cost impacts were modest: CZ4 experienced a small, non-significant increase driven by winter tariffs, while CZ12 peak savings were offset by off-peak increases. Emissions decreased by an average of 3% (CZ4) and 9% (CZ12).

Average COP remained ~2.3 across modes, with average delivered water temperature maintained relative to baseline. Comfort, defined here as uninterrupted access to hot water, was likewise maintained: 15 of 17 households reported that the MPC did not impede their hot water access, and 16 of 17 said they would recommend the technology overall.

Results demonstrate comfort-first supervisory MPC delivers peak reduction and emissions benefits in real-world deployments, with performance shaped by installation constraints and demand variability—key considerations for equitable electrification as tariffs and load flexibility programs evolve.

Introduction

Residential water heating is the second biggest energy end use in U.S. homes (U.S. Energy Information Administration 2020). As electrification and decarbonization efforts accelerate, heat pump water heaters (HPWHs) are positioned to be a key enabling technology because of substantial efficiency gains compared with electric resistance units. Additionally, the water in the tank can be used as thermal storage to support grid interactivity without the need for major modification.

However, HPWHs are commonly deployed as 240V hybrid units that pair a heat pump with electric resistance backup, and the substantial efficiency gains noted above are mainly realized when the heat pump does the heating rather than the resistance elements. Like conventional electric resistance water heaters, HPWHs are typically operated with a static temperature setpoint: the unit simply reheats whenever the water temperature falls below the control threshold, regardless of how much time is available to recover. This can trigger faster but less efficient resistance elements more often than necessary, undermining the very efficiency gains that motivate HPWH adoption. It also frequently occurs during expensive, carbon-intensive evening hours when people return home. This problem is not unique to HPWHs, but solving it is especially valuable for them: smarter control can keep the heat pump doing more of the work while also shifting that work to cheaper, cleaner hours.

While these opportunities have motivated substantial interest in grid-interactive water heating, the success of flexible control strategies ultimately depends on maintaining the service that residents expect. Occupants do not directly experience avoided emissions, lower system peaks, or improved load flexibility; they experience whether hot water is available when they need it. Even occasional comfort shortfalls can disproportionately influence perceptions of technology performance and willingness to participate in flexible-load programs. As utilities and policymakers increasingly look to residential water heating as a grid resource, demonstrating that flexibility can be delivered without compromising occupant comfort is becoming as important as demonstrating technical performance.

Supervisory model predictive control (MPC) offers a promising approach: by anticipating hot-water demand, electricity prices, and grid emissions, the controller can proactively charge the tank, shifting operation away from costly periods while maintaining hot-water comfort. This work advances that approach through a comfort-first supervisory controller that balances cost and emissions via a single blended "effective rate," while heavily penalizing comfort violations to ensure that comfort remains the primary objective. The controller interfaces with the water heater through its existing app, requiring no additional hardware, and preserves built-in safety mechanisms through the native rule-based controller (RBC).

In previous work, the authors developed this supervisory MPC approach across three phases (dela Rosa et al. 2023). A second phase addressed practical implementation challenges including: model mismatch, controller tuning, demand forecasting, and resistance-heater management, which are necessary for reliable real-world deployment (Dela Rosa et al. 2024; dela Rosa et al. 2025). Building on this foundation, laboratory testing validated the supervisory MPC on a 65-gallon HPWH using perfect forecasting knowledge for hot water demand, energy costs, and marginal GHG emissions. These tests showed the potential for ~15% utility-cost reduction, up to 61% marginal-emissions CO_2 reduction, and ~79% peak-period runtime reduction relative to the native RBC (Mande and Meyers 2024).

This paper focuses on the final phase of the overall project: a field demonstration of the supervisory MPC in 17 low-income multifamily homes, split between two properties: one in California climate zone 4 (CZ4) and the other in CZ12. The demonstration evaluated peak load shifting, emissions reduction, and occupant experience under the comfort-first supervisory MPC.

Real-world deployment introduces uncertainty in occupant behavior and exogenous variables that may reduce the reliability of deterministic scheduling approaches. To address these limitations, a simulation study was conducted using a stochastic MPC framework applied to data collected in the field study. This simulation demonstrates how the deterministic MPC deployed in the field may be refined for future applications to improve robustness to uncertain conditions.

Background and Prior Work

Existing grid-interactive water heating programs rely on CTA-2045, which allows utilities or third-party aggregators to send simple "load-up" or "shed" commands to water heaters through the CTA-2045 certified port or the manufacturer's application programming interface. While CTA-2045 itself can work on a local area network, in practice the "load-up" or "shed" command must still reach the household from an external party, which today typically means an internet connection. While robust and widely deployed, RBC is inherently reactive and lacks the adaptability needed to optimize for dynamic grid signals or stochastic occupant water draw patterns (dela Rosa et al. 2023).

MPC offers a proactive alternative by solving a constrained optimization problem over a receding horizon. MPC leverages a dynamic thermal model of the water tank alongside forecasts of future hot water demand, electricity prices, and grid carbon intensity to determine optimal heating schedules (Reyes Premer et al. 2026; dela Rosa et al. 2023). This forward-looking approach enables controllers to anticipate rather than merely reacting to changes in system conditions.

Research in predictive HPWH control has evolved from simple cost minimization in simulations to practical, multi-objective implementations addressing several key challenges:

A central challenge is balancing model accuracy with computational efficiency. Approaches range from simple one-node lumped models that enable real-time control and scalability (dela Rosa et al. 2025; Buechler et al. 2025; Reyes Premer et al. 2026) to multi-node stratified models that better capture vertical temperature layers during water draws (dela Rosa et al. 2023). Recent field studies have successfully deployed second-order linear models to improve comfort in 120V units without imposing high computational overhead (Reyes Premer et al. 2026).

Accurate prediction of occupant water use patterns is essential but difficult due to the stochastic nature of residential hot water draws. Researchers have proposed ensemble machine learning models combining Random Forest, XGBoost, and other techniques, as well as quantile-based heuristics to reduce the frequency of comfort violations where hot water is depleted (Reyes Premer et al. 2026).

For hybrid HPWHs, preventing frequent activation of inefficient electric resistance backup elements is critical to realizing efficiency gains. Researchers have developed logic-based constraints that restrict resistance heating to genuine hot water depletion events, ensuring the heat pump performs the majority of the heating work (dela Rosa et al. 2025).

This work builds on prior laboratory validation (Mande and Meyers 2024) by demonstrating the supervisory MPC approach in a real-world field setting with 17 low-income multifamily homes across two California climate zones. The field demonstration evaluates the controller's ability to achieve peak load shifting and emissions reduction while maintaining occupant comfort under actual operating conditions with imperfect forecasts and diverse usage patterns. This represents a critical step in validating the comfort-first supervisory MPC can deliver reliable performance across varied households without requiring specialized hardware installation.

Methods

Field Demonstration

The field demonstration took place at two low-income multifamily properties in California climate zone 4 and 12, referred to hereafter as the CZ4 and CZ12 Buildings. These sites were selected to measure the performance of the supervisory MPC, referred to hereafter as the Advanced Water Heating Control (AWHC), performance in different climate zones and building configurations.

The CZ4 Building, located in San Jose, contains 70 apartments with a mix of one- to four-bedroom layouts. Twenty-eight units are reserved for formerly unhoused residents, and four more are occupied through county housing-authority referrals. The property serves a linguistically and culturally diverse population of families, individuals, couples, and seniors, many of whom have previously experienced housing insecurity. Each apartment is equipped with a 65- or 80-gallon HPWH located in an unvented interior utility closet with a louvered door and an undercut at the base.

The CZ12 Building in Woodland includes 39 apartments, ranging from one- to three-bedrooms. Residents are predominantly low-income agricultural workers, including multi-generational families and seniors. These apartments also use 65- or 80-gallon HPWHs but the systems are located in an exterior utility closet and are ducted to ambient outdoor air.

To support equitable participation, the broader research team at the University of California, Davis (UC Davis) prepared bilingual (English/Spanish) informational materials describing the study's objectives, expectations, and potential benefits. UC Davis staff coordinated with on-site management to host community meetings where residents could learn about the project, ask questions, and voice concerns. These outreach events—held between March and April 2022 at the CZ4 Building and between September and October 2021 at the CZ12 Building—helped establish rapport and encourage enrollment. Although all households were eligible, recruitment emphasized a representative distribution of unit sizes (one- through three-bedroom). Participation was voluntary and interested residents completed an informed-consent process outlining study procedures and participant rights.

By May 2022, a total of 26 households across the two properties had joined the study. Over the subsequent three years, participant numbers declined more than anticipated, primarily due to resident turnover associated with pandemic-related disruptions and evolving eligibility for subsidized housing. Two additional households withdrew for other reasons, one citing concerns about increased electricity costs and another noting that participation requirements were inconvenient. Supplemental recruitment in 2024 helped boost enrollment, and by the start of the AWHC monitoring period, 17 households remained active, representing household sizes from one to seven occupants.

Control and Equipment Performance

To isolate the influence of the AWHC from confounding factors such as weather conditions and resident hot-water use patterns, the field demonstration employed a randomized rapid-M&V design, based on methods developed by (Raftery et al. 2024), in which baseline and AWHC performance were interleaved within the same observation period rather than measured sequentially. This approach ensured that both control approaches were exposed to similar weather conditions and resident behavior patterns over the course of the study. Figure 1 confirms this for weather, showing that average outdoor air temperatures and their distributions were nearly identical between the two approaches for each apartment. Hot water use patterns were not independently verified in this way, but the interleaved design makes it unlikely that resident behavior differed systematically between conditions.

The control method for each HPWH alternated between its native RBC – operating at the default manufacturer settings (eco mode and a water setpoint of 120 °F [48.89°C]) – and the AWHC, following a randomized three-day block schedule. To prevent thermal carryover from biasing results, data from the first day after each control transition were omitted from subsequent analysis ("Drop Day") (Figure 2).

Each water heater was equipped with 1-minute-resolution instrumentation capturing key operating variables—outlet and inlet water temperatures, mixed water temperature, evaporator inlet conditions, flow rate, and electrical power draw (Figure 3). Sensor details are summarized in Table 1. Baseline monitoring began in December 2021 for the CZ12 Building and in May 2022 for the CZ4 Building. AWHC operation began in September 2024, with a follow-up control-algorithm refinement implemented in November 2024.

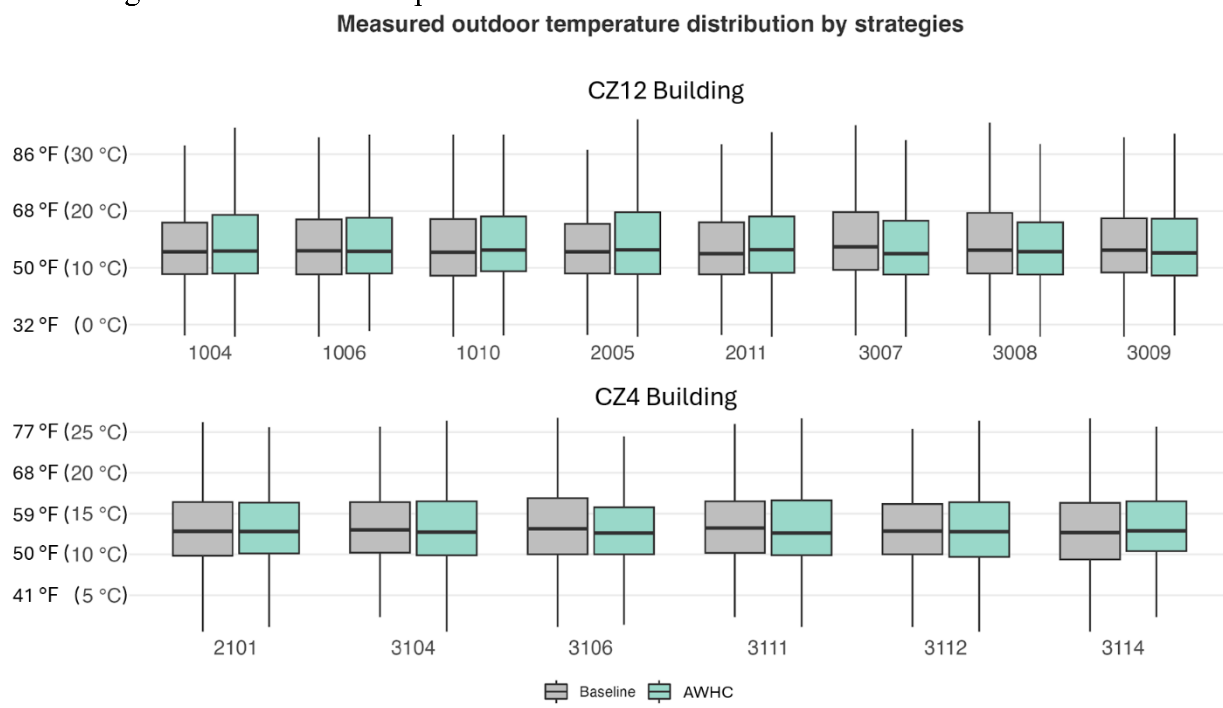

Figure 1: Distribution of outdoor air temperature measured while each control method was active.

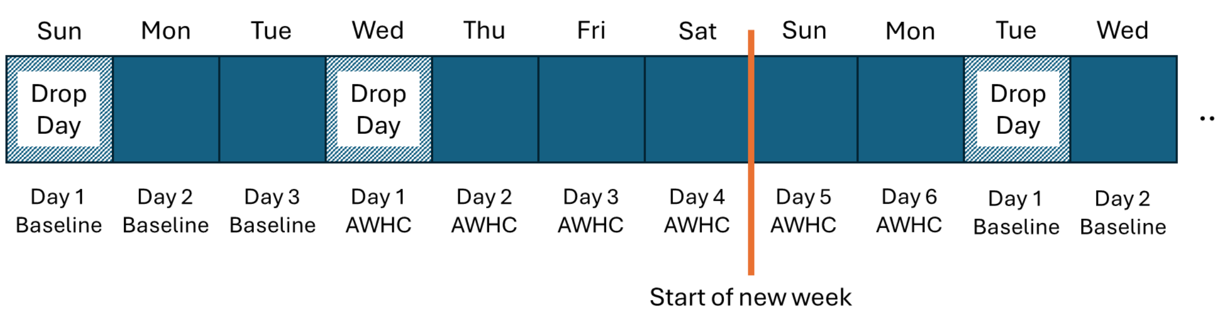

Figure 2: Randomized 11-day switchback schedule with two transitions between baseline and AWHC. Alternative randomized sequences, including all-baseline or all-AWHC periods, would also be considered valid.

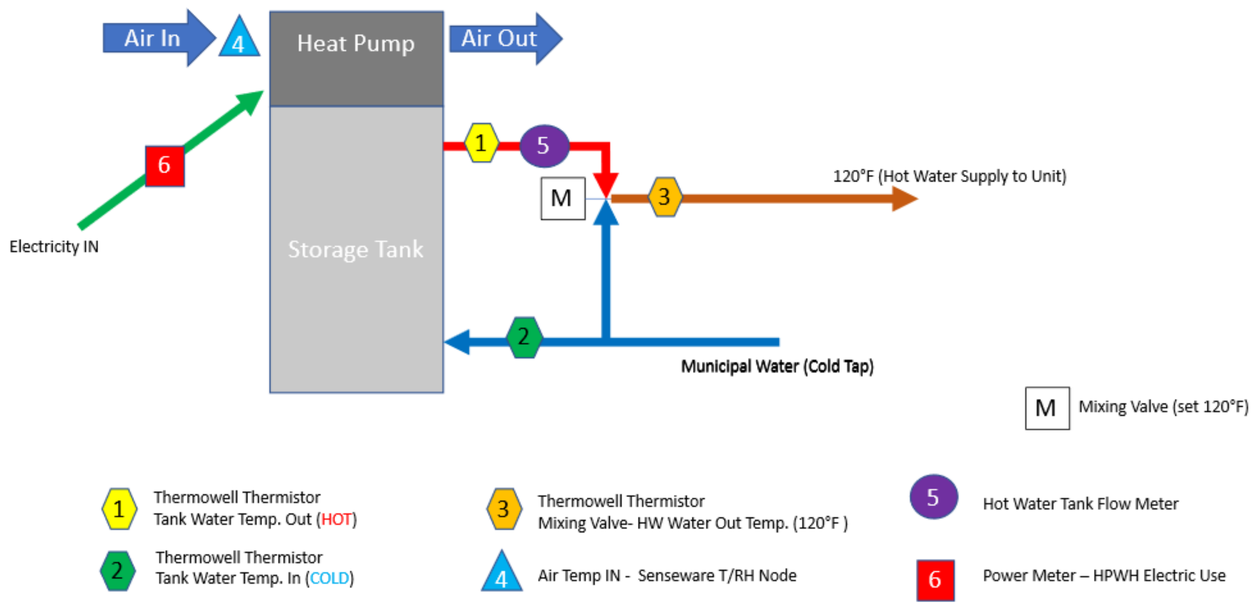

Figure 3: Data collection schematic diagram for sensors used for AWHC field demonstration. Sensor details are provided in Table 1.

Table 1: Summary of measurements for each HPWH in field demonstration. Sensor number corresponds to Figure 3.

Sensor #	**Sensor type**	**Measured Parameter**	**Unit**
1	Thermistor	HPWH hot water outlet temperature	°F
2	Thermistor	HPWH cold water inlet temperature	°F
3	Thermistor	Mixing valve mixed water output temperature	°F
4	T/RH sensor	HPWH evaporator inlet air temperature	°F
4	T/RH sensor	HPWH evaporator inlet air relative humidity	%
5	Flow meter	HPWH hot water outlet flow	GPM
6	Power meter	HPWH electrical power/energy	kW/kWh

Forecasts of Exogenous Inputs

The AWHC required forecasts of four exogenous inputs—electricity price, weather, marginal GHG emissions, and hot water demand—to proactively control the HPWHs. The electricity price forecast was based on Pacific Gas & Electric's (PGE's) time-of-use rates, which included a peak period from 4 to 9 PM, with all other hours classified as off-peak. Existing forecast providers were used for both weather and marginal GHG emissions. OpenWeather was used as the local weather forecast provider based on the latitude and longitude of the demonstration sites. WattTime was used for marginal GHG emissions based on the grid balancing authority, which was the same for both sites.

Unlike electricity prices, weather, and marginal GHG emissions, residential hot water demand lacks an established third-party forecasting service. This challenge is well documented in the literature, as hot water consumption is highly stochastic and strongly influenced by occupant behavior, making it difficult to predict using conventional time-based approaches.

To address this challenge, the AWHC forecasting method leverages the observation that the sequence and relative timing of water-use events are often more consistent than the events themselves. For example, while the exact timing of cooking or showering may vary from day to day, the interval between these activities tends to remain relatively stable. Rather than predicting demand at fixed times, the method identifies historical days exhibiting similar intra-day usage patterns and uses them to inform future demand.

The developed forecasting algorithm employs a stacked kernel approach in which autocorrelation is used as the similarity metric. The previous 24 hours of water-use data are resampled at multiple temporal resolutions (15 minutes, 1, 2, 4, 6, 8, and 12 hours), creating a set of kernels that represent the same period at different levels of aggregation. Each kernel is then compared against candidate historical windows spanning from 7 to 50 days before the forecast date, with all historical data resampled using the same temporal windows. The lower bound of 7 days was selected to preserve day-of-week matching, while the upper bound of 50 days was identified during tuning for the demonstration sites. Although the source data are recorded at 1-minute intervals, they are aggregated to 15-minute resolution to ensure candidate windows align on common boundaries.

After calculating autocorrelation scores across all kernels and candidate windows, the 10 highest-scoring matches are selected, and their average demand profile is used to forecast the next hour of hot water use. A full 24-hour forecast is generated iteratively: each newly forecasted hour is appended to the most recent 23 hours of observed data to create an updated 24-hour template, which is then used to predict the subsequent hour. This sliding-window process continues until a complete 24-hour demand forecast is produced for use by the AWHC.

Occupant Experience

Household surveys were developed to capture both baseline and post-retrofit conditions, with an emphasis on hot-water availability and comfort, patterns of hot-water and HVAC use, thermostat practices, and perceptions of energy costs. Social science researchers at UC Davis designed the surveys to align with project goals, prioritizing closed-ended questions to reduce respondent burden while incorporating selective open-ended items for additional context. All survey materials and implementation procedures underwent review and approval by the UC Davis Institutional Review Board (IRBNet ID: 1951687-3) and were deployed using the Qualtrics platform.

Baseline data collection took place in September 2023 and again in May 2024 to document household characteristics and usage behaviors. Because field deployment of the AWHC technology was postponed after the first survey round, a second baseline survey was administered to ensure current household information prior to activation of the control system. Post-retrofit surveys were conducted in January, February, and October 2025 to assess residents' experiences with the installation process and their interactions with the load-flexible control features once operational.

A trained UC Davis researcher administered all surveys by telephone using computer-assisted interviewing in either English or Spanish, depending on participant preference. Responses were entered directly into Qualtrics and stored in a secure project database. Each interview lasted approximately 15–25 minutes and was completed by one adult household member. Participants received a $100 gift card for each completed survey as compensation for their time. The resulting dataset was examined using a combination of quantitative and qualitative analytic techniques to characterize household experiences before and after implementation of the advanced control strategies.

Results

Control and Equipment Performance

The effectiveness of shifting electrical consumption outside the 4–9 p.m. peak-demand window varied notably between the two study sites. Figure 4 shows the household-level average peak-period savings differed by climate zone. In the CZ12 Building, the AWHC resulted in a statistically significant reduction, with a mean savings of 31% per day in peak-hour energy use ($p = 0.01$). In contrast, the CZ4 Building achieved a smaller, statistically insignificant reduction, with a mean savings of 20% per day ($p = 0.08$). The limited flexibility observed at the CZ4 site was largely associated with higher hot-water demand and constrained airflow in the indoor utility closets, which restricted HPWH performance during load-shifting events.

Overall bill costs under PG&E's winter and summer time-of-use tariffs, CZ4 Building participants experienced a mean increase of $0.10 per day during the winter months and a mean decrease of $0.10 per day during the summer months. Although the magnitudes offset one another across tariff seasons, the winter tariff's longer duration — active for 8 of 12 months — meant that AWHC produced a net annual cost increase at this site; however, this effect was not statistically significant. In the CZ12 Building, on-peak savings were offset by corresponding increases in off-peak consumption for both the winter and summer tariff seasons and were not statistically significant.

Carbon-emission impacts during the monitoring period were modest at both sites, with the CZ4 Building achieving a mean daily reduction of 3% and the CZ12 Building achieving 9%. The limited magnitude of these reductions is attributed to two compounding factors: marginal emissions rates on the Northern California grid follow an inverse time-of-use profile — lowest midday when solar generation suppresses grid emissions and highest during morning and evening hours — and the monitoring dataset was weighted toward winter months, when shorter daylight hours compress the low-emissions midday window and increase the share of high-emissions hours that cannot be displaced through load shift.

Across both properties and control modes, average COP remained stable at approximately 2.3, indicating that overall system efficiency was similar with and without AWHC. However, during AWHC operation at the CZ4 site, HPWHs relied slightly more on electric resistance elements, contributing to an average daily power increase of 19 W per unit. This pattern suggests that, in some cases, the control strategy prioritized maintaining hot-water availability during peak-shifting periods by less efficient electric resistance modes, an issue compounded by the limited airflow and longer runtimes inherent to the indoor-closet installations.

Figure 5, shows that delivered hot-water temperatures generally remained above 110 °F (43.3 °C). Nonetheless, several units experienced longer durations below this threshold during AWHC operation, and average time spent under 105 °F (40.6 °C) increased modestly. While these deviations remained within acceptable bounds for most households, the results indicate that AWHC occasionally amplified pre-existing performance constraints in undersized or airflow-limited systems without leading to outright service failures.

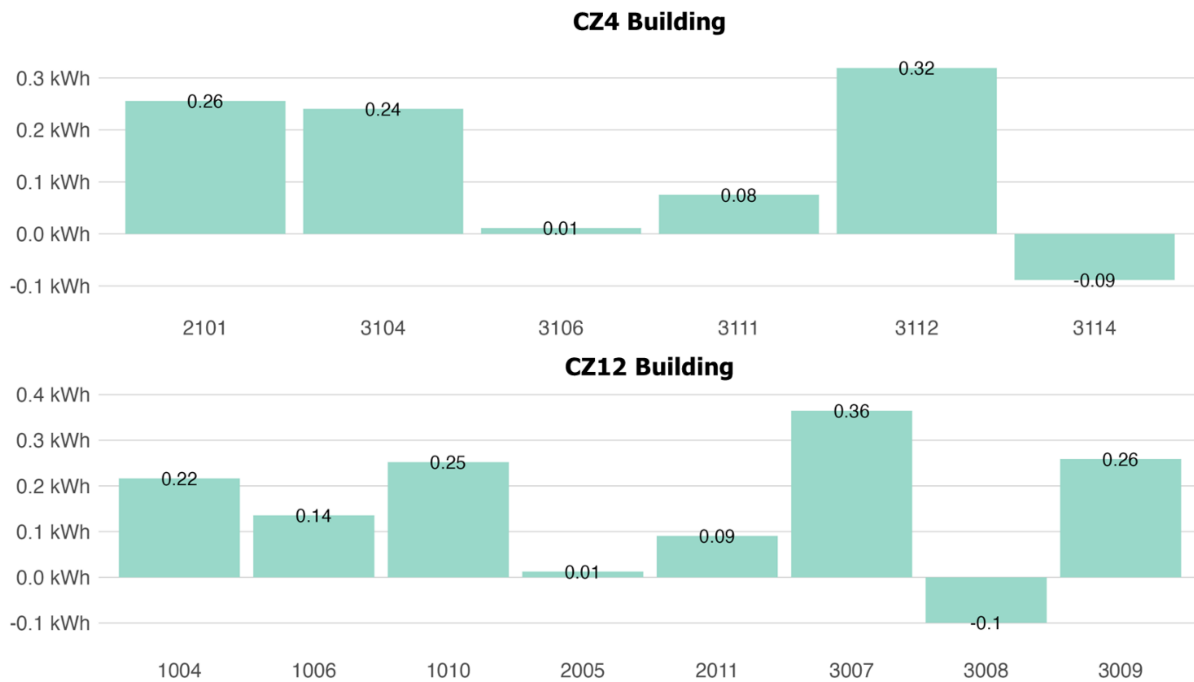

Figure 4: AWHC mean peak energy savings (4 PM to 9 PM) based on data collected between November 2024 and July 2025.

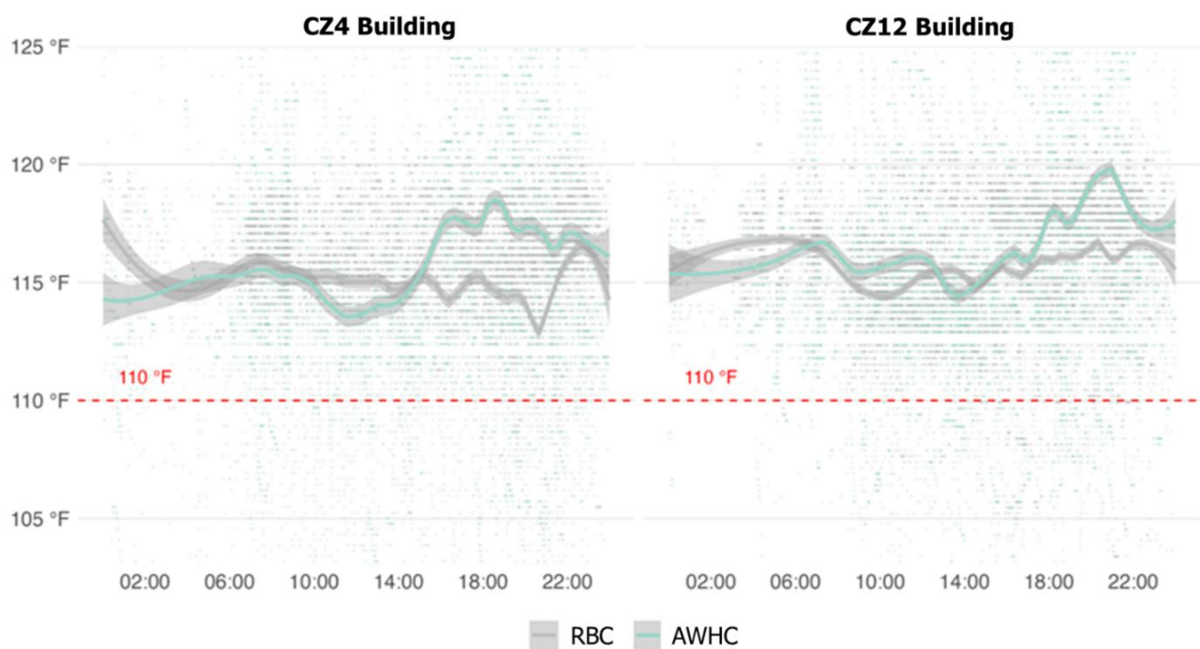

Figure 5: HPWH supply hot water temperature for RBC and AWHC over the course of a day.

Occupant Experience

Survey responses revealed broadly favorable attitudes toward the advanced water-heating controls. Among the 17 households that completed the survey, 16 reported they would be very likely to recommend the technology to others. On a more granular level, nearly two-thirds rated their satisfaction with how long hot water remained available as high (Figure 6), suggesting that while access was rarely interrupted outright, the duration of supply left some room for improvement in high-demand households. This is consistent with reports from several high-use households of occasional inconveniences, such as slower recovery to preferred temperatures or insufficient supply during back-to-back showers.

Perceptions of economic impact were more varied. Roughly 7 of respondents were uncertain about changes in utility costs—largely due to overlapping rate increases during the study period—while about one-quarter reported experiencing lower bills. Most participants (n=15) stated that the AWHC did not impede their ability to access adequate hot water for managing health-related needs, including conditions such as arthritis or dermatological issues. A small number of households, however, reported that reduced maximum temperatures occasionally required them to boil supplemental water for specific medical or hygiene practices.

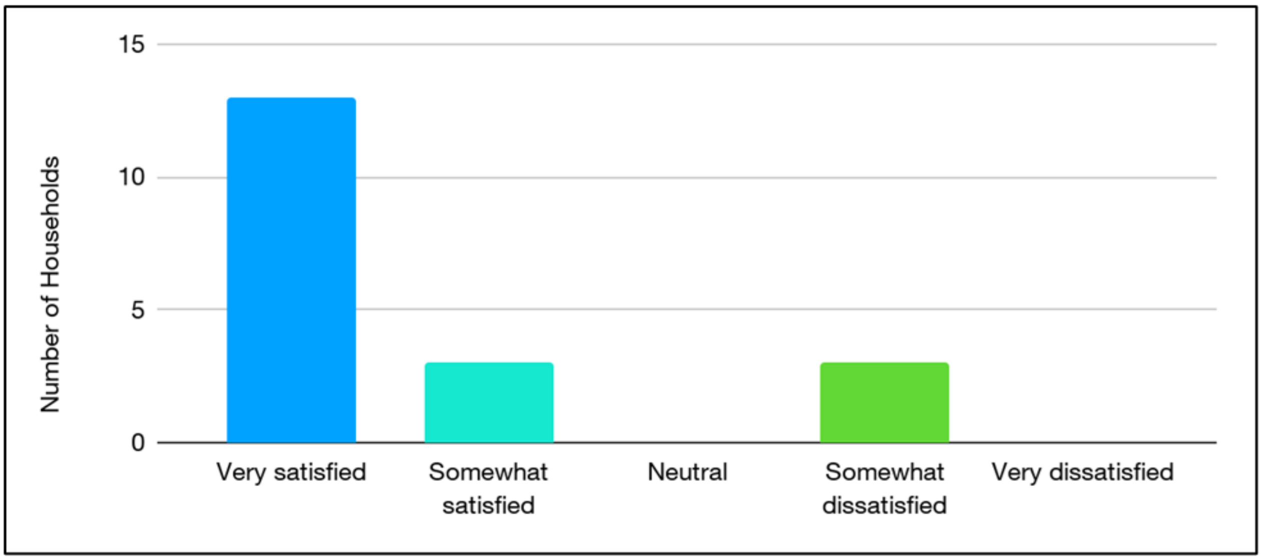

Figure 6: User survey results for the duration hot water availability under AWHC

Simulation of Chance-Constrained Stochastic MPC

Real-world environments introduce uncertainty through variations in occupant behavior and exogenous conditions. Uncertain conditions may reduce the reliability of deterministic scheduling approaches, such as the one deployed in this study. To address these limitations, a chance-constrained stochastic MPC framework was developed and evaluated in a simulation environment using historical data captured during the field deployment.

The forecast used in the stochastic MPC is a modified version of the deterministic forecasts that selects a dynamic k number of forecasts, where k is the number of forecasts with a Pearson correlation coefficient > 0.7. Each of these forecasts is then used to develop an optimization scenario for the MPC. The MPC cost function becomes:

$$\min_u \sum_{i=1}^{K} \pi_i \sum_{n=0}^{N-1} p_{eff,n} \tilde{P}_n \Delta t$$

where π_i is the probability associated with a given scenario.

Maintaining occupant hot water availability is handled through a probabilistic comfort constraint. Instead of requiring the minimum temperature requirement to be satisfied for every possible demand scenario, which can result in overly conservative operation, the controller limits the expected severity of temperature violations across all scenarios. A temperature shortfall variable is introduced:

$$\delta_{i,n} \geq T_{min} - T_{i,n}$$
$$\delta_{i,n} \geq 0$$

where $\delta_{i,n}$ represents the amount by which the tank temperature falls below the minimum acceptable temperature T_{min}. The probability-weighted expected temperature shortfall is constrained according to:

$$\sum_{i=1}^{K} \pi_i \delta_{i,n} \leq \epsilon T_{tol} + \sigma_n$$

where π_i is the probability of scenario i, ϵ is the allowable violation probability, T_{tol} is the temperature tolerance used to scale the acceptable risk, and σ_n is a non-negative slack variable. This constraint allows low-probability or small temperature violations while preventing large, expected comfort impacts. The slack variable allows the optimization problem to remain feasible during extreme demand events but is penalized to discourage violations.

To prevent excessively high tank temperatures during load-up events, an upper temperature constraint is applied to each scenario:

$$T_{i,n} \leq T_{max}$$
$$\epsilon_{i,n}^{+} \geq 0$$

where $\epsilon_{i,n}^{+}$ is the magnitude of any upper temperature violation. This allows small constraint violations when necessary while penalizing them in the objective function.

A terminal constraint is also introduced to ensure that the water heater does not end the prediction horizon with insufficient stored thermal energy. The terminal temperature shortfall is defined as:

$$\delta_s^f \geq T_f - T_{i,N}$$

where T_f is the required terminal temperature and N is the final time interval in the prediction horizon. The probability-weighted terminal shortfall is constrained by:

$$\sum_{i=1}^{K} \pi_i \delta_i^f \leq \epsilon_f T_{tol} + \sigma_f$$

where ϵ_f is the allowable terminal violation probability and σ_f is the terminal slack variable. The terminal constraint prevents the MPC from minimizing short-term energy cost by allowing excessive depletion of stored hot water near the end of the horizon, which negatively impacts performance in the next horizon interval.

HPWH load flexibility using the stochastic MPC was simulated over one week at site 3008, which was a site with high hourly DHW consumption and high variability in draw patterns. Performance was compared against the deterministic MPC approach simulated over the same period. Application of the stochastic MPC yielded a 34% cost reduction, which was greater than the 19% cost reduction achieved in the deterministic MPC simulation. However, peak load reduction between the two simulations was similar, with the stochastic MPC achieving a peak load reduction of 45% and the deterministic MPC achieving a peak load reduction of 49%. These results indicate that both control strategies were effective at shifting HPWH operation away from high-cost periods, while the primary benefit of the stochastic MPC was improved management of uncertainty in future DHW demand. By considering multiple possible demand trajectories, the stochastic MPC was able to proactively maintain stored thermal energy using the higher-efficiency heat pump element, reducing reliance on electric resistance heating during unexpected high-demand events. As a result, the stochastic MPC achieved greater overall energy and cost savings while maintaining comparable load-shifting performance.

These initial simulation results build on findings from the field demonstration by exploring opportunities to improve MPC performance under the high-demand and variable operating conditions where deterministic approaches may demonstrate limitations. These results represent an initial evaluation based on a one-week simulation at a single site. Additional simulations across longer time periods and a larger sample of DHW usage profiles are needed to validate the performance benefits of the stochastic MPC approach and determine the generalizability of these findings.

Discussion

The field results demonstrate that a comfort-first supervisory MPC approach can deliver meaningful peak-load reductions in real-world low-income multifamily settings, but performance is strongly mediated by installation context and demand characteristics. The statistically significant 31% reduction in 4–9 p.m. energy use in CZ12 highlights the value of pairing hybrid HPWHs with predictive preheating under favorable airflow and moderate demand patterns. In contrast, the smaller and statistically marginal 20% reduction observed in CZ4 reflects the operational constraints inherent to indoor-closet installations, where limited airflow results in longer runtimes and more frequent use of electric-resistance heating during load-shifting periods.

Despite these constraints, comfort outcomes were generally strong. Most households reported satisfaction with hot-water duration and did not attribute comfort losses to the AWHC. The few cases reporting temperature shortfalls or recovery delays aligned with objectively observed conditions—high simultaneous demand or borderline-sized systems—indicating that the control strategy largely preserved hot-water availability within the bounds of existing equipment capabilities.

Across both sites, average COP remained stable at 2.3, suggesting that supervisory MPC can shift load without degrading whole-system efficiency. However, the modest rise in resistance-element usage in CZ4 shows that MPC must account for installation limitations to avoid unintended trade-offs. Future refinements may incorporate airflow diagnostics or adaptive resistance-avoidance logic to maintain both comfort and efficiency.

To investigate the potential value of incorporating uncertainty directly into the control formulation, a stochastic MPC approach was evaluated using field-derived DHW demand profiles. Initial simulation results from a high-demand, high-variability site showed that stochastic MPC achieved greater operating cost reduction than deterministic MPC while maintaining comparable peak-period load shifting. The improved performance was primarily associated with increased use of efficient heat pump operation and reduced reliance on electric resistance heating. By considering multiple probable demand trajectories rather than a single expected forecast, stochastic MPC may proactively manage stored thermal energy in response to uncertain occupant behavior. These preliminary results suggest that probabilistic control approaches may help address the high-demand edge cases observed during field deployment; however, additional simulations across more households, seasons, and operating conditions are needed to validate these findings.

Importantly, acceptability among low-income households was high. Transparency, bilingual materials, and structured touchpoints with the research team appear to have contributed to trust, as evidenced by a 90% "very likely to recommend" rate. These findings reinforce the importance of centering occupant experience—especially in communities disproportionately affected by energy insecurity—when developing load-flexible residential technologies.

Overall, the field demonstration validates that predictive, API-based supervisory control can scale across existing HPWH fleets, but achieving equitable performance requires nuanced treatment of local constraints, equipment diversity, and occupant needs.

Limitations and Future Work

The study involved 18 households across two properties, limiting generalizability across the broader California multifamily stock or single-family homes. While the focus on low-income communities is critical for ensuring equitable access to load-flexible technologies, future demonstrations across a wider range of household types, climates, and usage patterns are needed to better characterize performance variability and scalability.

Performance differences between the field sites demonstrated the importance of installation characteristics and occupant demand patterns. Airflow-restricted closets in CZ4 meaningfully constrained performance and introduced more resistance-element usage. Future deployments should include airflow screening protocols or automated diagnostics that alert operators to suboptimal installation conditions.

The single-node thermal model provides speed and robustness but may under-represent stratification dynamics under high-draw conditions. Incorporating adaptive stratification weighting, simplified hybrid observers, or data-driven demand-sequence recognition may better capture tank availability for large or irregular draw profiles. Additionally, the deterministic MPC approach deployed in the field relies on a single forecast of future DHW demand, which does not explicitly represent uncertainty in occupant behavior. Initial simulations using a stochastic MPC framework with probabilistic demand forecasts and chance constraints demonstrated the potential to improve operation in high-demand, high-variability households. However, these results were based on a limited simulation period and should be evaluated across additional households, seasons, and operating conditions to determine whether stochastic control provides consistent improvements over deterministic MPC

Resident perceptions of bill impacts were mixed due to concurrent rate restructuring by utilities. Evaluating AWHC under real-time pricing, critical-peak pricing, or more dynamic tariffs may provide clearer insights into economic benefits and trade-offs.

Finally, while this study focused on local comfort, cost, and emissions metrics, future work should evaluate portfolio-level participation in aggregated demand-flexibility programs or virtual power plants (VPPs), where coordinated HPWH fleets could provide dispatchable load reductions.

Conclusions

This work demonstrates that comfort-first, cloud-supervisory MPC can effectively shift HPWH operation away from peak hours while maintaining occupant satisfaction in low-income multifamily homes. The control strategy reduced peak-period energy use by 31% in CZ12 and 20% in CZ4, with minimal comfort impacts for the majority of households. Variability between the two sites underscores the influence of installation characteristics and household demand, suggesting that airflow conditions and system sizing must be explicitly considered in future deployments.

Laboratory results showed strong potential for cost and emissions benefits, and field results confirm that these gains can be realized in practice—particularly where equipment is not constrained by airflow or high coincident demand. High acceptance among residents indicates that supervisory MPC, when combined with proactive communication and equitable program design, can support broader adoption of grid-interactive heat pump technologies.

As California moves toward more dynamic tariffs and increasing reliance on flexible load, the comfort-first approach validated here offers a scalable, retrofit-friendly pathway for integrating HPWH fleets into demand-response and decarbonization strategies while prioritizing the needs of households least able to bear the risks of poor comfort or bill volatility.

Acknowledgements

We gratefully acknowledge support from the California Energy Commission for funding this work under project number EPC-19-015. We also gratefully acknowledge Dr. Sarah Outcault and Cinthia Y. Perea for managing relationships with the demonstration sites and participants, including participant recruitment, Spanish translation, and interviews capturing participant experience. We thank Dove Feng, Marian Goebes, and Daniel Hacking of TRC for managing the M&V infrastructure, and Paul Raftery and Aoyu Zou of UC Berkeley for their collaboration on the rapid-M&V approach and analysis of results.

AI-assisted tools used: Claude, Research Rabbit.
Used for: Copyediting, summarizing, and searching
Extent: light editing and citation research
Human review: The authors reviewed and verified all content and citations.

References

Buechler, Elizabeth, Aaron Goldin, and Ram Rajagopal. 2025. "Designing Model Predictive Control Strategies for Grid-Interactive Water Heaters for Load Shifting Applications." *Applied Energy* 382 (March): 125149. https://doi.org/10.1016/j.apenergy.2024.125149.

dela Rosa, Loren, Caton Mande, and Matthew J. Ellis. 2024. "Impact of Model Mismatch on MPC Performance for Heat Pump Water Heaters." *2024 American Control Conference (ACC)*, July, 5364–69. https://doi.org/10.23919/ACC60939.2024.10644183.

Mande, Caton, and Fred Meyers. 2024. *Optimizing Cost and Carbon Footprint: Laboratory Testing of Model Predictive Control for Smart Management of Heat Pump Water Heaters*.

Raftery, Paul, Aoyu Zou, Thomas Parkinson, and Geoff Hancock. 2024. “Reliably Estimating the Impact of an Active Control Strategy in a Building.” *Journal of Building Engineering* 98 (December): 111134. https://doi.org/10.1016/j.jobe.2024.111134.

Reyes Premer, Levi D., Elias N. Pergantis, Leo Semmelmann, Davide Ziviani, and Kevin J. Kircher. 2026. “Model Predictive Control Lowers Barriers to Adoption of Heat-Pump Water Heaters: A Field Study.” *Energy Conversion and Management* 348 (January): 120723. https://doi.org/10.1016/j.enconman.2025.120723.

dela Rosa, Loren, Caton Mande, and Matthew J. Ellis. 2023. “Supervisory Multi-Objective Economic Model Predictive Control for Heat Pump Water Heaters for Cost and Carbon Optimization.” *Proceedings of the ASHRAE Winter Conference*, February, 517–25. https://doi.org/10.63044/w23del60.

dela Rosa, Loren, Caton Mande, and Matthew J. Ellis. 2025. “Practical Strategies for Managing Resistance Heating in Heat Pump Water Heater Predictive Control.” *Chemical Engineering Research and Design* 215 (March): 180–92. https://doi.org/10.1016/j.cherd.2025.01.024.

U.S. Energy Information Administration. 2020. “Residential Energy Consumption Survey.” https://www.eia.gov/consumption/residential/data/2020/index.php?view=consumption.