# Enhanced Normalized Metered Energy Consumption Analysis with Rapid Interventions

## Final Report

**ET23SWE0049**

Prepared by:

**Brett Close,** TRC

**Bilsay Varcin,** TRC

**Niranjana Sreejayan,** TRC

**Aoyu Zou,** UC Berkeley

**Paul Raftery,** UC Berkeley

September 15, 2025

Image: Façade of a residential multifamily building.

## Acknowledgements

The team thanks Paul Raftery and Aoyu Zou at the University of California, Berkeley, for their valuable insights and support throughout this study—particularly for conducting the device-level analysis. Their contributions enhanced the rigor of our analysis and strengthened our approach. The team also offers our sincere gratitude to Caton Mande at the University of California, Davis, whose expertise in device-level integration and control systems played a crucial role in the implementation and interpretation of the data. Their combined efforts were integral to the success of this research project. The team also thanks the California Energy Commission's Electric Program Investment Charge (EPIC) program for its support of this work.

# Executive Summary

## Purpose and Scope

This study evaluated two methods for estimating energy savings from heat pump water heaters, using data from 25 homes in two multifamily buildings located in San Jose, California, and Woodland, California. The first method, traditional normalized metered energy consumption analysis, compares long pre- and post-intervention periods using utility billing data. The second method, enhanced normalized metered energy consumption with rapid interventions, applies randomized on/off control intervals to estimate savings more quickly. The intervention system used forecasts of energy cost, comfort, and greenhouse gas emissions to decide when to slightly preheat the tank. For example, if it anticipated high electricity prices soon and low emissions around midday, it raised the temperature to avoid running the heat pump during expensive or high-emission times while keeping comfort as the top priority. Our objective was to determine whether enhanced normalized metered energy consumption with rapid interventions can produce savings estimates comparable to those from traditional normalized metered energy consumption. Many programs cannot support 12 months of pre- and post-intervention data collection due to schedule or budget constraints. Enhanced normalized metered energy consumption with rapid interventions addresses this by mixing baseline and intervention data throughout the study period, allowing models to adapt in near-real time to changes in behavior or external conditions. Because both traditional normalized metered energy consumption and enhanced normalized metered energy consumption with rapid interventions are estimates of an unknown quantity, device-level data analysis is compared to the actual heat pump hot water heater usage.

## Methodology

The project team collected three rounds of Pacific Gas & Electric utility billing data, completed in June 2025. The intervention ran from November 2024 to August 2025, alternating control settings on randomized schedules with balanced 10-week blocks. Each block included a similar number of baseline and intervention days for every day of the week—for example, five Mondays with the intervention control strategy and five Mondays on baseline before moving on to the next block. Each 24-hour period was executed following a control change to prevent carryover effects. We used regression models to estimate hourly energy use, controlling for heating and cooling demands as well as daily hot water usage. The project team analyzed enhanced normalized metered energy consumption with rapid interventions only for the post-intervention period, which included days when the control was active (treatment) and days when it was inactive (baseline). The team compared traditional normalized metered energy consumption data between pre-intervention baseline days and post-intervention treatment days. Any post-period days that returned to baseline were dropped to keep a clean before-after comparison. The team fit a fixed effects panel model to control for time invariant differences between units.

## Key Findings

### Peak Hours

When focusing on peak hours, no method produced statistically significant savings in San Jose, although the device-level measurement and verification model showed a small effect that approached significance. All methods produced statistically significant savings in Woodland. Across both locations, the device-level measurement and verification model consistently delivered strong statistical performance, while both enhanced normalized metered energy consumption with rapid interventions and traditional normalized metered energy consumption methods produced marginal or insignificant results. Enhanced normalized metered energy consumption with rapid interventions showed smaller standard errors in San Jose and similar precision to traditional normalized metered energy consumption in Woodland, supporting its improved precision.

### All Hours of the Day

When the team examined all hours of the day, it was clear that the device-level model continued to perform best. It produced statistically significant negative savings in San Jose and across all sites combined. Neither normalized metered energy consumption method showed significant savings anywhere in the all-hours analysis. Despite this, formal tests indicated that neither method differed significantly from the device-level model: enhanced normalized metered energy consumption with rapid interventions consistently produced results closest to those from the device-level approach and, in most cases, produced lower-variance estimates that allow for greater precision.

## Implications

Enhanced normalized metered energy consumption with rapid interventions provided results as accurate as traditional normalized metered energy consumption while offering greater precision and notable practical advantages. Enhanced normalized metered energy consumption with rapid interventions allows for faster evaluation and verified savings without requiring long baseline periods. It works well when equipment supports rapid control switching. For efficiency programs needing timely, accurate verification, enhanced normalized metered energy consumption with rapid interventions offers a viable alternative to traditional normalized metered energy consumption. Enhanced normalized metered energy consumption with rapid interventions also adapts continuously to changing conditions. For example, it adjusts to factors like tenant turnover, seasonal changes, or shifts in usage patterns. This makes it more flexible than traditional normalized metered energy consumption, which assumes no external changes between long pre- and post-periods. Its shorter data collection window reduces the risk of external changes over time. Detecting savings using any normalized metered energy consumption method based on utility meter data remains more challenging compared to device-level analysis when overall savings are small, even with enhanced normalized metered energy consumption with rapid interventions. This is because utility billing data can be noisy, especially when the expected savings are minimal. Since enhanced normalized metered energy consumption with rapid interventions relies on controlled interventions, available on/off control capabilities are essential.

## Recommendations

The project team recommends using enhanced normalized metered energy consumption with rapid interventions for evaluations when rapid control switching is technically possible and when faster results are desired. The project team will continue to validate the method across larger samples and varying control technologies to confirm its accuracy and applicability in broader contexts.

## Abbreviations and Acronyms

| Acronym | Meaning |
|---|---|
| CEC | California Energy Commission |
| EPIC | Electric Program Investment Charge |
| E-NMEC RI | Enhanced normalized metered energy consumption with rapid interventions |
| HPLF | Heat pump load flexibility |
| HPWH | Heat pump water heater |
| kWh | Kilowatt-hour |
| M&V | Measurement and verification |
| MPC | Model predictive control |
| NMEC | Normalized metered energy consumption |
| PG&E | Pacific Gas & Electric |

# Contents



## Tables



## Figures

# Introduction

This report examines the differences between two methods for estimating energy savings using utility billing data and device-level meter data, comparing conventional normalized metered energy consumption (NMEC) analysis (California Public Utilities Commission 2020) with an alternative approach, enhanced NMEC with rapid interventions (E-NMEC RI). The project team sought to understand whether E-NMEC RI could successfully estimate project savings and outperform traditional NMEC within a shorter timeframe.

The report includes:

- An overview of the NMEC and E-NMEC RI methods
- A summary of how the two methods were applied to the same set of sites
- A comparison of savings results from utility billing data and device-level data
- A discussion of the results

# Background

NMEC is a widely used approach for estimating energy savings that relies on actual utility meter data—e.g., electricity or gas use—instead of modeled or deemed savings, offering higher confidence in measured impacts. Using utility billing data and weather information, NMEC compares energy use before and after an energy efficiency intervention. Traditional NMEC often requires long periods of data, such as one year before and one year after the intervention, to produce reliable results. States like California have formally adopted NMEC into their evaluation protocols, which helps make the evaluation process more transparent and consistent.

The E-NMEC RI method shortens this process. Instead of using long pre- and post-periods, E-NMEC RI turns the intervention on and off in short, randomized intervals. This creates multiple shorter treatment and control periods, all during the post-installation period. The method uses regression models to estimate savings quickly and may reduce issues caused by customer turnover, seasonal effects, or other changes.

This study relied on data from a related (in-progress) project—Optimizing Heat Pump Load Flexibility (HPLF) for Cost, Comfort, and Carbon Emissions—conducted by the University of California, Davis, with TRC as a subcontractor under the California Energy Commission (CEC) Electric Power Investment Charge (EPIC) funding. The CEC Optimizing HPLF project deployed different control strategies for heat pump water heaters (HPWHs) in 25 dwelling units in 2 multifamily buildings in San Jose, California, and Woodland, California. The project team used the same homes in this study and compared results from device-level meter data with results from utility billing data, applying both the traditional NMEC and E-NMEC RI methods and comparing them to saving estimates based on device-level usage rather than whole-home utility meter data.

## Objectives

The main goals of this study were to:

- Test whether the E-NMEC RI method could successfully estimate energy savings comparable to those from conventional NMEC
- Compare energy savings results from E-NMEC RI and traditional NMEC using utility billing data with the results from device-level meter data analysis
- Evaluate how well the methods captured the impact of load shift interventions during the peak hours of 4:00 p.m. to 9:00 p.m.
- Assess the uncertainty of the savings estimates using both data types across all methods

## Methodology and Approach

### Implementing the Analysis Framework

The project team set up the necessary infrastructure to support three types of analysis: device-level analysis using HPWH usage data from meters installed on the HPWHs, conventional NMEC using utility meter data, and E-NMEC RI using utility meter data. All three approaches use energy consumption data to estimate savings from HPWH control strategies; the team treats the device-level results as the benchmark and compares the two NMEC approaches to the device-level analysis.

The project team recruited tenants from the CEC Optimizing HPLF project, who each signed a consent form allowing the team to access and download their Pacific Gas and Electric (PG&E) utility billing data. Participants received $100 for each utility data download, in addition to a participation stipend through the CEC Optimizing HPLF project. The project team completed all three rounds of utility data collection—with the final download conducted at the end of June—and used the November and February data to build our analytical framework and perform early checks to confirm that the data behaved as expected.

To support the E-NMEC RI method, the research team from the related CEC Optimizing HPLF project installed control devices on the HPWHs. These devices receive signals through the Attune platform, which switches the water heaters between standard and intervention control modes on a randomized schedule.[1] The same platform collects detailed device-level usage data, which forms the basis of the device-level analysis and serve as a reference for comparison with utility billing-based methods.

### Executing Measurement and Verification Activities

The second activity involved executing measurement and verification (M&V) activities across all selected sites. The CEC project team configured advanced load management systems on the HPWHs and installed device-level meters to track energy and hot water flow. This involved setting time-based

[1] Attune is an Internet-of-Things-based monitoring and control platform that provides real-time visibility and automated control of building assets, including heating equipment, via wireless sensors and cloud dashboards. https://www.attuneiot.com/

control strategies to shift water heating loads away from peak demand periods while ensuring consistent water temperature for users. The M&V activities tracked both energy usage and peak load shifting over time, measuring the system's response to the intervention.

The M&V method used a randomized schedule to switch between the two control strategies. Each apartment followed a separate schedule, and the study used a three-day sampling interval. For example, one unit might have received three days of intervention followed by three days of baseline, just by chance. Another unit could have received nine consecutive days of intervention before switching back to baseline. The team removed days from the analysis when the control strategy did not operate on consecutive days. These days could show carryover effects, where the previous day's control setting affected the next day's performance.

The schedule began on September 26, 2024, and continued until August 28, 2025. The team restarted the schedule on November 18, 2024, after resolving an initial issue where set-point commands failed due to a server space error, as well as a separate cloud-based issue. The team divided each schedule into 10-week blocks, and at the end of each block, the schedule included a roughly equal number of baseline and intervention days for each hour of the day. This timing helped maintain balanced sampling and reduced the chance of detecting false effects.

For E-NMEC RI, the baseline referred to days after the start of the randomization when the model predictive control (MPC) controller was off, whereas the traditional NMEC baseline referred to the period before randomization started. For both methods, the project team defined intervention as days after the start of randomization, i.e., when the MPC controller was on.

## Data Collection

The project team initially collected utility data from 11 homes in Woodland and 12 homes in San Jose. This data included energy usage and cost for both baseline and intervention periods for all sites. However, by the third round of data collection, two residents had moved out and data could not be obtained from two additional homes. After formatting and examining for data completeness, the project team used 12 sites—4 in San Jose, 8 in Woodland—in the traditional NMEC analysis, and 14 sites—5 in San Jose, 9 in Woodland—in the E-NMEC RI analysis. This included baseline and intervention-period data from both utility meters and device-level sources. Depending on the method, the team selected homes where there were no significant gaps in recorded periods. E-NMEC RI required complete data once randomization began, whereas traditional NMEC required comprehensive data before randomization, as well as during randomization, when the MPC controller was on.

The team used utility billing data downloaded from PG&E for both traditional NMEC and E-NMEC RI, allowing for comparison across treatment and non-treatment periods using meter-level energy use.

The research team also collected device-level meter data through the Attune communication platform. Attune integrated with sensors and meters installed at each dwelling, and the system transmitted data continuously to a secure cloud server using a cellular modem. This detailed dataset supported precise evaluation of how the HPWHs responded to the control strategy, which the team used in the device-level analysis.

## Outdoor Weather Conditions

Outdoor temperature was not expected to significantly affect HPWH performance in this study. However, outdoor weather still played an important role in energy modeling, as seasonal shifts in outdoor temperature affected both hot water usage patterns and the temperature of the incoming cold-water supply. These factors influenced energy and water demand and therefore were controlled for in the NMEC regression analysis to accurately isolate the effects of the control strategy. Although average conditions were similar between the two study sites, the sampled temperatures showed a narrower and colder distribution compared to typical annual patterns. The HPWHs in Woodland were installed in exterior locations, while the HPWHs in San Jose were placed in interior closets, although it is unclear how this difference in location affected energy savings.

## Differences Between E-NMEC RI, Traditional NMEC, and Device-Level M&V

This study compared two savings estimation approaches based on utility meter data, or NMEC, to an ideal scenario of site-level M&V with device-level data.

For the two NMEC approaches, the team estimated energy savings using a regression model based on utility meter data. The project team considered outdoor air temperature, hour of the day, and the strategy sampled as the independent variables. The team also included outdoor air temperature in the regression to account for how outdoor air temperature affected hot water usage, which allowed the model—known as a panel fixed effects model—to separate the effect of the control intervention from weather-related consumption changes. This type of regression implicitly controlled for all time-invariant characteristics within a specific site, such as insulation level or the number of residents, while estimating average slope coefficients. Both models took the form:

$$EnergyUse_{it} = \alpha_i + RI_{it} * \beta_{RI} + \delta_{HOD,t} * \beta_{HOD} + HDH65_{it} * \beta_{HDH65} + CDH65_{it} * \beta_{CDH65} + \varepsilon_{it}$$

Where:

- $i$ indexes the specific site (apartment)
- $t$ indexes the hour
- $EnergyUse_{it}$ is hourly energy usage (kwh) for unit $i$ in time $t$
- $\alpha_i$ is a site-specific constant that incorporates the effect of all time-invariant characteristics
- $RI_{it}$ is 1 for time $t$ when unit $i$ has the intervention and 0 for otherwise
- $\delta_{HOD,t}$ is a vector of indicators for the hours of the day (HOD) that are equal to 1 when time $t$ is equal to hour HOD and 0 otherwise
- $HDH65_{it}$ is heating degree hours for unit $i$ in time $t$ with a changepoint of 65 degrees
- $CDH65_{it}$ is cooling degree hours for unit $i$ in time $t$ with a changepoint of 65 degrees
- Each $\beta$ is a slope coefficient
- $\varepsilon_{it}$ is an idiosyncratic error term for unit $i$ in time $t$

The difference between the E-NMEC RI and the traditional NMEC is in the time periods $t$, which are included in the model data.

- For E-NMEC RI, the model included only time periods $t$ after the start of the intervention, with the value of $RI_{it}$ alternating between 0 and 1 depending on the randomization.

- For traditional NMEC, the model included a year a pre-period data with $RI_{it}$ equal to 0, and intermittent post period data where $RI_{it}$ was equal to 1, with the time periods where $RI_{it}$ was equal to 0 omitted from the model.

The project team compared the results from these models to a device-level M&V model that used the metered usage of the HPWH. The device-level M&V model used daily data with two separate models run for each site—one model for days with the intervention and one model for days without the intervention—rather than hourly data with a single model for all sites, and took the form:

$$EnergyUse_t = \beta_0 + \delta_{DOW,t} * \beta_{DOW} + Outdoor_t * \beta_O + Water_t * \beta_W + GPM_t * \beta_{GMP} + \varepsilon_t$$

Where:

- $t$ indexes the day
- $EnergyUse_t$ is daily energy usage (kwh) on day $t$
- $\delta_{DOW,t}$ is a vector of indicators for the days of the week (DOW) that are equal to 1 when day $t$ is equal to day DOW and 0 otherwise
- $Outdoor_t$ is the average daily outdoor air temperature on day $t$
- $Water_t$ is the average daily hot water supply temperature on day $t$
- $GPM_t$ is the average hot water draw on day $t$
- Each $\beta$ is a slope coefficient
- $\varepsilon_t$ is an idiosyncratic error term for unit *i* in time $t$

Here, the team presents both overall daily results and results focusing on the peak hours of 4:00 p.m. to 9:00 p.m. For the two NMEC models, the team simply dropped the hours outside the peak period. For the device-level M&V model, the team ran a model with daily data but aggregated the variables only for certain hours—4:00 p.m. to 9:00 p.m. for $EnergyUse_t$ and $Outdoor_t$,and 2:00 p.m. to 9:00 p.m. for $Water_t$ and $GPM_t$—to account for time delays between water draw and water heater energy use.

Savings for the two NMEC models were based on the coefficient $\beta_{RI}$. This coefficient equaled the average hourly difference in energy usage between time periods with and without the intervention, across all sites in the model. Average daily savings were then equal to $-24*\beta_{RI}$, and average daily peak savings were equal to $-5*\beta_{RI}$. Savings for the device-level M&V model were based on the difference in predicted energy usage between the baseline and intervention model. Savings for each site were estimated by predicting usage over the whole analysis period by using either the baseline or intervention model and then calculating the difference per day. The average energy savings were then the average of the site-level savings.

For our comparison, the team used the device-level M&V results for the reference estimate, as these were based on device-level usage itself. The project team then compared the results of the two NMEC approaches.

# Findings

## Overview

Across both the San Jose and Woodland sites, the team observed energy-use reductions during peak hours under the control strategy. Woodland showed consistently stronger and statistically significant peak-period savings, while San Jose results were more variable and not statistically significant. These outcomes suggest that household water usage patterns may influence savings detection.

Cost analysis revealed minimal differences between baseline and intervention periods, mainly due to modest equipment loads and high uncertainty at the whole-home level. However, when focusing on high-usage households and units with strong device-level responses, the team detected clearer utility-level signals.

Most importantly, our cross-method statistical comparison found no significant difference between E-NMEC RI and traditional NMEC in terms of estimated energy savings; both were generally within the statistical uncertainty of the device-level result. In most cases, the E-NMEC RI had improved precision and a smaller deviation from the device-level results. This suggests that E-NMEC RI provides comparable—if not improved—accuracy, while offering operational advantages like shorter evaluation windows and improved validity due to the randomization of the intervention. Therefore, E-NMEC RI may be a viable substitute for traditional NMEC when device-level control capabilities are available.

## Results

In this section, the project team analyzes how traditional NMEC and E-NMEC RI perform using utility billing data and compares them to device-level meter data analysis. The team evaluates whether the E-NMEC RI method, which toggles control strategies on and off frequently, can match the savings estimates from the device-level analysis, and compares the results to the traditional population NMEC approach. The team also assesses both methods for consistency in measured savings and confidence intervals.

### Heat Pump Water Heater Usage

This section describes the supply hot water temperature distribution and daily hot water usage for baseline and intervention days. For most units, the supply temperature was slightly lower during intervention days than on baseline days. However, the difference was small and not statistically significant. Daily average hot water usage was slightly higher during intervention days for most units as well.

### SUPPLY WATER TEMPERATURE

The team analyzed hot water temperature data based on a set point of 110°F. Water heaters typically provide water at around 110ºF to 120°F to balance comfort and safety, so the team defined a hot water run-out event as any instance where the supply temperature fell below 110°F. To focus on actual hot water draws, the team excluded periods with no hot water demand and ignored the first minute of each draw. That initial minute often showed ambient water temperature before hot water arrived.

Figure 1 shows the distribution of supply water temperature minute readings for units during baseline and intervention days. Supply temperatures were slightly lower on intervention days for most units, but the difference was small and not statistically significant. Run-out events occurred occasionally, happening at a similar rate in both baseline and intervention periods. That is, the bottom whiskers of the box and whisker plot illustrate that the minimum temperature reading is generally lower, but the overall number of events is similar. This analysis confirms that the intervention did not lead to a higher frequency of hot-water run-out events.

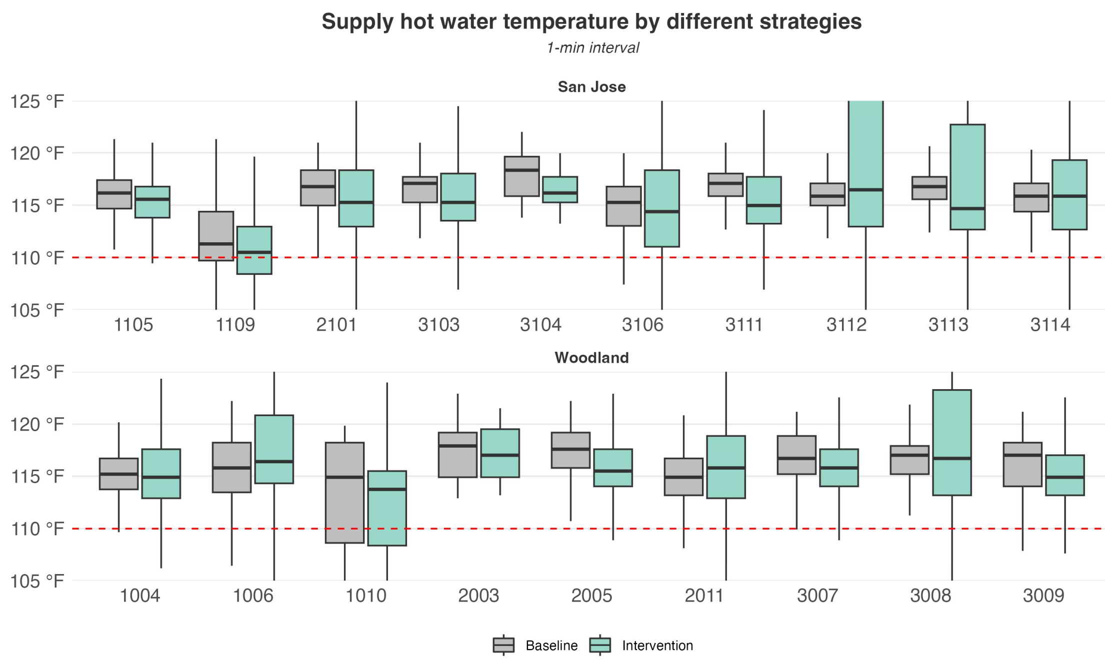


**Figure 1: Distribution of supply hot water temperature when hot water was used for each unit between sampled baseline days and intervention days.**

### HOT WATER FLOW

The project team analyzed hot water usage based on flow meter data collected before and after the mixing valve, reviewing average daily usage and hourly flow patterns to compare baseline and intervention days. Figure 2 shows the average daily hot water usage for units measured from the tank flow meter, i.e., before the mixing valve, and the calculated flow after mixing, i.e., after the mixing valve. Our analysis found that most units used slightly more hot water on intervention days. This increase aligns with the slightly lower supply temperatures observed during those days, which may have prompted occupants to draw more hot water.

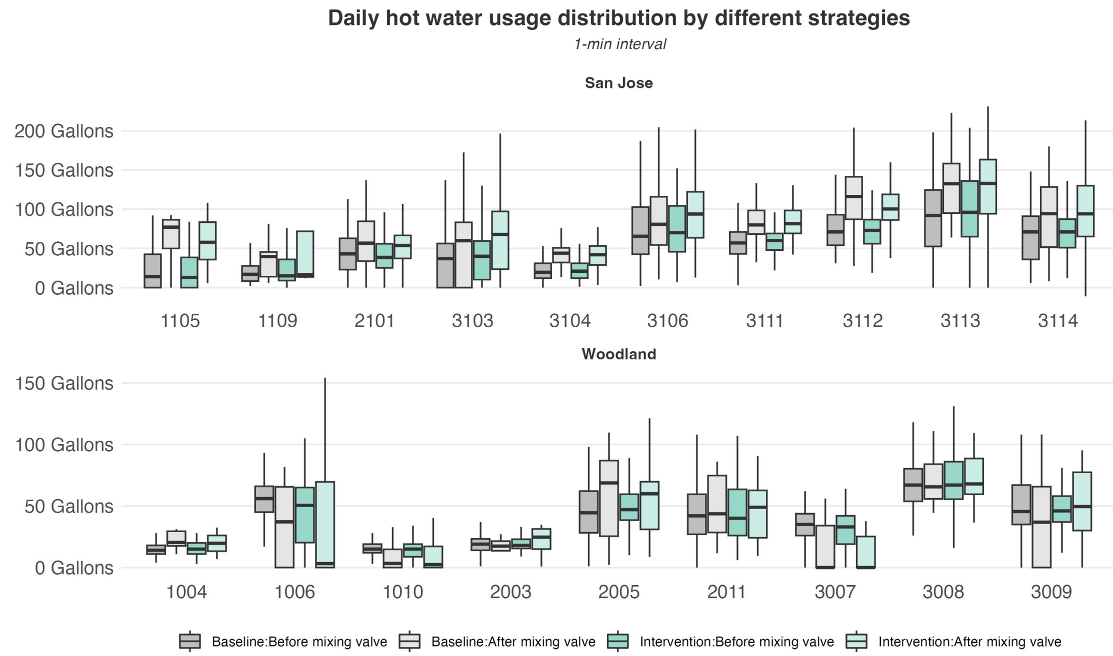


**Figure 2: Daily average hot water usage by different strategies.**

The differences in flow rates were small and varied significantly across units, which was likely due to individual habits and random sampling variation. Overall, the team did not find any consistent or meaningful pattern that would suggest the intervention had strongly influenced hot water consumption.

## Peak Window Device Energy Usage

The team analyzed how energy usage measured at the HPWH changed under different control strategies, focusing on the peak window of 4:00 p.m. to 9:00 p.m. In our analysis, units in San Jose showed some energy savings during that window, but the change was not statistically significant. This aligns with the increase in power the team observed during the intervention periods. In contrast, units in Woodland showed significant energy reductions. Since overall daily usage remained similar between baseline and intervention days, this suggests that the control strategy successfully reduced energy consumption during the peak window.

When examining the relationship between hot water usage and savings, the team found that energy savings tended to decrease above a daily usage threshold of about 50 gallons. Based on this finding, the project team included daily hot water usage as a control variable in our regression models to improve accuracy.

## Cost Analysis

Through the MPC cost function framework—where a blended objective function consists of energy cost, marginal emissions, and comfort violation—the team analyzed differences in cost through energy bills and emissions between the baseline and intervention periods. The project team calculated daily energy costs using power consumption and time-of-use tariff rates, with peak hours at $0.50 per kWh and off-peak hours at $0.36 per kWh.

The project team's analysis showed that the intervention had no meaningful effect on cost, with daily energy costs remaining similar between baseline and intervention periods. Even though our methods detected shifts in load timing, these changes did not translate into cost reductions. The primary challenge lies in measurement uncertainty, especially for utilities where equipment usage is modest and the intervention effects are small. For households with higher usage and more load shifting, more significant cost impacts may be detectable.

### Peak Window Site Energy Usage

The team also assessed how much site-level energy savings were possible during the 4:00 p.m. to 9:00 p.m. peak hours, when electricity demand and rates are often highest.[2] HPWHs account for roughly 20 percent of overall household electricity use, so we examined whether shifting their operation during this window would produce detectable savings when viewed at the whole-house utility meter.

Woodland units produced a measurable reduction in usage during the peak hours, but San Jose units showed smaller estimated savings. One challenge associated with detecting meaningful savings at the site level is the larger uncertainty in the measurements, especially when the units use less hot water and savings are even less noticeable at the submeter level. This comparison could be improved by limiting the sites included in the analysis to only those with substantial savings based on the device-level M&V. However, detecting meaningful savings at that level remains challenging due to higher measurement uncertainty, especially for households with low usage or small net intervention effects.

### Cross-Method Comparison of Savings Estimates

Because the study's control strategy focused on cost savings, the team expected the primary impact to be on reductions in usage during peak hours when energy was most expensive. When focusing on the peak hours, the team found no statistically significant energy savings for San Jose for any of the methods, though the device-level M&V model had a small estimate that was nearly significant. All models found statistically significant energy savings for Woodland. Across all sites in both San Jose and Woodland, the device-level M&V model showed statistically significant results, and the two NMEC models showed results that were nearly—but not quite—significant at the 10 percent level. Note that for San Jose and the overall results, the standard error of the savings is substantially smaller for E-NMEC RI than for traditional NMEC, and the standard errors are quite similar for Woodland. This is consistent with E-NMEC RI offering improved precision over traditional NMEC. Results for each model and site are shown in Table 1 below.

**Table 1: Normalized daily peak (4:00 p.m. to 9:00 p.m.) energy savings results across methods.**

| Location | Method | Average Daily Savings (kWh) | Standard Error | P-value |
|---|---|---|---|---|
| San Jose | Device-level site-specific M&V | 0.09 | 0.05 | 0.05 |

[2] Device-level results are discussed above in Peak Window Device Energy Usage.

| Location | Method | Average Daily Savings (kWh) | Standard Error | P-value |
|---|---|---|---|---|
| San Jose | E-NMEC RI | 0.02 | 0.1 | 0.87 |
| San Jose | Traditional NMEC | 0.74 | 0.71 | 0.37 |
| Woodland | Device-level site-specific M&V | 0.14 | 0.04 | 0.00 |
| Woodland | E-NMEC RI | 0.27 | 0.2 | 0.17 |
| Woodland | Traditional NMEC | 0.52 | 0.13 | 0.01 |
| All locations | Device-level Site-specific M&V | 0.08 | 0.02 | 0.00 |
| All locations | E-NMEC RI | 0.18 | 0.14 | 0.23 |
| All locations | Traditional NMEC | 0.47 | 0.28 | 0.13 |

The central analysis for this study focused on how well the methods agreed. In all cases, the E-NMEC RI results were closer to the device-level M&V results than the traditional NMEC results. To make a formal comparison, the team conducted a two-sided test of the null hypothesis where the NMEC result equaled the device-level M&V result. In all cases, the team failed to reject the null hypothesis for the E-NMEC RI result, whereas the team rejected the null hypothesis for one of the traditional NMEC approaches in Woodland. Additionally, all p-values were larger for the E-NMEC RI comparisons, indicating that the differences were likely due to random variation rather than a true difference. These results indicate that, for this study, the E-NMEC RI model outperformed the traditional NMEC model, as shown below in Table 2.

**Table 2: Statistical comparison of normalized daily peak (4:00 p.m. to 9:00 p.m.) energy savings between E-NMEC-RI and traditional NMEC.**

| Location | Comparison Model | Difference in Savings (kWh) | Standard Error | P-value |
|---|---|---|---|---|
| San Jose | E-NMEC RI | −0.03 | 0.12 | 0.81 |
| San Jose | Traditional NMEC | 0.69 | 0.71 | 0.33 |
| Woodland | E-NMEC RI | 0.15 | 0.20 | 0.45 |
| Woodland | Traditional NMEC | 0.40 | 0.14 | 0.01 |

| Location | Comparison Model | Difference in Savings (kWh) | Standard Error | P-value |
|---|---|---|---|---|
| All locations | E-NMEC RI | 0.10 | 0.14 | 0.48 |
| All locations | Traditional NMEC | 0.39 | 0.28 | 0.16 |

The team also estimated the overall savings impacts using all hours of the day. When examining results across all hours, the San Jose site showed statistically significant negative savings only for the device-level M&V model, with usage higher rather than lower on days with the control strategy. Both E-NMEC RI and traditional NMEC models showed high uncertainty and no statistical significance. For Woodland, the device-level M&V showed statistically significant negative savings of p=0.07, but neither NMEC method yielded statistically significant results. This same pattern was evident in the results across all sites, as shown below in Table 3.

**Table 3: Normalized daily energy savings results across methods (all hours).**

| Location | Method | Average Daily Savings (kWh) | Standard Error | P-value |
|---|---|---|---|---|
| San Jose | Device-level site-specific M&V | −0.26 | 0.08 | 0.00 |
| San Jose | E-NMEC RI | −0.40 | 0.66 | 0.59 |
| San Jose | Traditional NMEC | 0.78 | 1.15 | 0.55 |
| Woodland | Device-level site-specific M&V | −0.11 | 0.08 | 0.07 |
| Woodland | E-NMEC RI | 0.27 | 0.52 | 0.63 |
| Woodland | Traditional NMEC | −0.38 | 0.88 | 0.69 |
| All locations | Device-level site-specific M&V | −0.19 | 0.06 | 0.00 |
| All locations | E-NMEC RI | 0.035 | 0.41 | 0.93 |
| All locations | Traditional NMEC | 0.15 | 0.73 | 0.84 |

In every comparison, the E-NMEC RI results were closer to the device-level M&V results than the traditional NMEC results. To test this formally, the team conducted a two-sided test of the null hypothesis. In all cases, the team could not reject the null hypothesis when comparing E-NMEC RI or traditional NMEC to device-level M&V. The relative p-values vary, largely due to the substantially higher uncertainty in the traditional NMEC results. Table 4 presents the full comparison.

**Table 4: Statistical comparison of normalized daily energy savings results across methods (all hours) between E-NMEC-RI and traditional NMEC.**

| Location | Comparison Model | Difference in Savings (kWh) | Standard Error | P-value |
|---|---|---|---|---|
| San Jose | E-NMEC RI | -0.14 | 0.66 | 0.83 |
| San Jose | Traditional NMEC | 1.04 | 1.16 | 0.37 |
| Woodland | E-NMEC RI | 0.38 | 0.53 | 0.47 |
| Woodland | Traditional NMEC | -0.26 | 0.88 | 0.77 |
| All locations | E-NMEC RI | 0.23 | 0.41 | 0.58 |
| All locations | Traditional NMEC | 0.34 | 0.73 | 0.64 |

## Recommendations

The project team recommends the use of E-NMEC RI for evaluations when rapid control switching is technically possible and faster results are desired. Its randomized intervention framework enabled the team to estimate savings more quickly and, in some cases, with better precision compared to traditional NMEC. The team also recommends continued validation of E-NMEC RI across larger and more diverse deployments, including different technologies and customer settings, to strengthen confidence in its accuracy and scalability.

# References

California Public Utilities Commission. 2020. *Rulebook for programs and projects based on normalized metered energy consumption.* Vol. (Version 2.0). CA.