APPENDIX D

REPORT

Duke Ohio 2016 Power Manager Evaluation

March 2017

Josh Bode, M.P.P.
Candice Potter, M.S.
Ankit Jain, M.P.P.
Greg Sidorov, M.S.
Trevor Cummings, M.S.
Rachel Flaherman, B.S.

Abstract

This study analyzes the impact of Duke Energy Ohio's Power Manager program on electricity demand for a range of weather conditions and dispatch hours under normal and emergency operation protocols. Power Manager is a voluntary demand response program that provides incentives to residential customers who allow Duke Energy to reduce the use of their central air conditioner's outdoor compressor and fan on summer days with high energy usage. The impacts were evaluated using a randomized control trial design. Each customer was randomly assigned to one of six groups at the start of the summer—a primary group with 75% of the population (approximately 30,000 customers) and five research groups, each with 5% of the population (2,000 customers per group). During each event, a control group of approximately 2,000 households was withheld to establish the baseline absent activation of Power Manager. In addition, as part of the evaluation, Nexant collected air conditioner end use data for 95 units in 89 homes, conducted a visual inspection of 103 load control devices, surveyed participants immediately after an event and control day, and interviewed program staff and implementers.

Acknowledgements

The study required careful collaboration with the Duke Energy Ohio's evaluation and operations team, MadDash, Inc., Nexant field engineers, and Nexant's survey data collection lab. In specific, the inputs from Duke Energy's team—Rich Philip, John Kappesser, Marjan Salek, Rose Stoeckle, Danielle Maple and Regina Harris—were critical to proper implementation of the randomized control trial and the analysis. Their comments and edits are reflected throughout the report. Marjan deserves special mention because she took on the critical and monumental task of addressing individual devices to the randomly assigned groups. Dr. Michael Sullivan and Dr. Jon Cook provided critical input to the design of the study and the sample size simulations. A special thanks to Mad Dash, Inc. whose staff implemented the installation of air conditioner end use data loggers and inspected load control devices. Nexant field engineers were critical in retrieving end use data loggers and downloading the data. The Nexant survey data collection team led the recruitment of the end use sample, coordinated scheduling between field staff and customers, implemented the survey data collection, and coordinated the retrieval of data loggers.

Figures

Tables

1 Executive Summary

This report presents the results of the 2016 Power Manager impact and process evaluations for the Duke Energy Ohio territory. Power Manager is a voluntary demand response program that provides incentives to residential customers who allow Duke Energy to reduce the use of their central air conditioner's outdoor compressor and fan on summer days with high energy usage. During normal events, the signal to load control devices to reduce air conditioner use is phased in over the first half hour and the reduction is sustained through the remainder of the event and phased out over the half hour immediately after the event. During emergency operations, all devices are instructed to instantaneously shed loads and deliver larger demand reductions (75% cycling).

1.1 Impact Evaluation Key Findings

The impact evaluation is based on a randomized control trial. Each customer who had an addressable load control device at the start of the summer was randomly assigned to one of six groups—a primary group with 75% of the population and five research groups, each with 5% of the population. During each event, a control group of approximately 2,000 households was withheld to provide an estimate of energy load profiles absent activation of Power Manager. [1] In addition, as part of the evaluation, Nexant collected air conditioner end use data for 95 units in 89 homes and conducted a visual inspection of 103 load control devices. While Nexant also analyzed impacts for customers with end-use data loggers via regression methods, the randomized control trial results are the primary evaluation results. They are more precise, require no modeling, and rely on what is indisputably the best evaluation method.

During the summer of 2016, between 48,105 and 48,178 air conditioner units were actively participating in Power Manager and had load control devices. The average household had 1.06 load control devices installed.

Figure 1-1 summarizes the demand reductions for the 2016 general population and emergency test curtailment events as a function of weather. Table 1-1 summarizes the reductions attained during each event in 2016, as estimated using the randomized control trial. The July 21, 2016 event included a side-by-side test of demand reduction under different dispatch hours during which 75% of customers were dispatched for the 3:30pm to 4:00pm event and four research groups were dispatched at different times. The July 25, 2016 event included side-by-side tests of emergency and normal operations in order to estimate the incremental demand reductions due to emergency operations.

A few key findings are worth highlighting:

- Demand reductions were 0.79 kW per household for the average general population event.
- Peak day impacts under normal operations were 1.01 kW per household on July 25, 2016, when max temperatures reached 93˚F.

[1] A total of 45,152 customers with 47,675 load control devices – all customers enrolled in Power Manager in spring of 2016 - were randomly assigned to the six groups. The analysis was implemented with smart meter data for 39,804 customers. The difference in counts is due the combination of customer turnover (primarily due to moving) and the ability of Duke Energy Ohio to extract the smart meter data.

- Emergency operations produced larger impacts than normal operations, 1.49 kW vs. 1.05 kW per household for the same hour on the hottest day in 2016, when the daily maximum temperature reached 93˚F. Reductions from emergency operations exceeded those from normal operations by 41.9%.
- The magnitude of impacts varies slightly by dispatch window. Demand reductions ranged from 0.67 to 0.84 kW per household on July 21, when different randomly assigned groups were dispatched at different times. As a percentage of loads, the demand reductions varied less, ranging from 21.0% to 23.4%, suggesting that most of the differences by event window are a function of the underlying amount of air conditioner load.
- Demand reductions grow larger in magnitude when temperatures are hotter and resources are needed most.
- The difference in impacts between customers who signed up for the lower and higher load control options was minimal.
- There is no evidence that customers compensate for air conditioner curtailments by increasing other end uses—whole building impacts are no different than end use impacts.
- The randomized control trial using smart meter data produced highly precise estimates, enabled side-by-sides tests, and should be the primary evaluation method.

Figure 1-1: Demand Reduction by Load Control Option as a Function of Weather

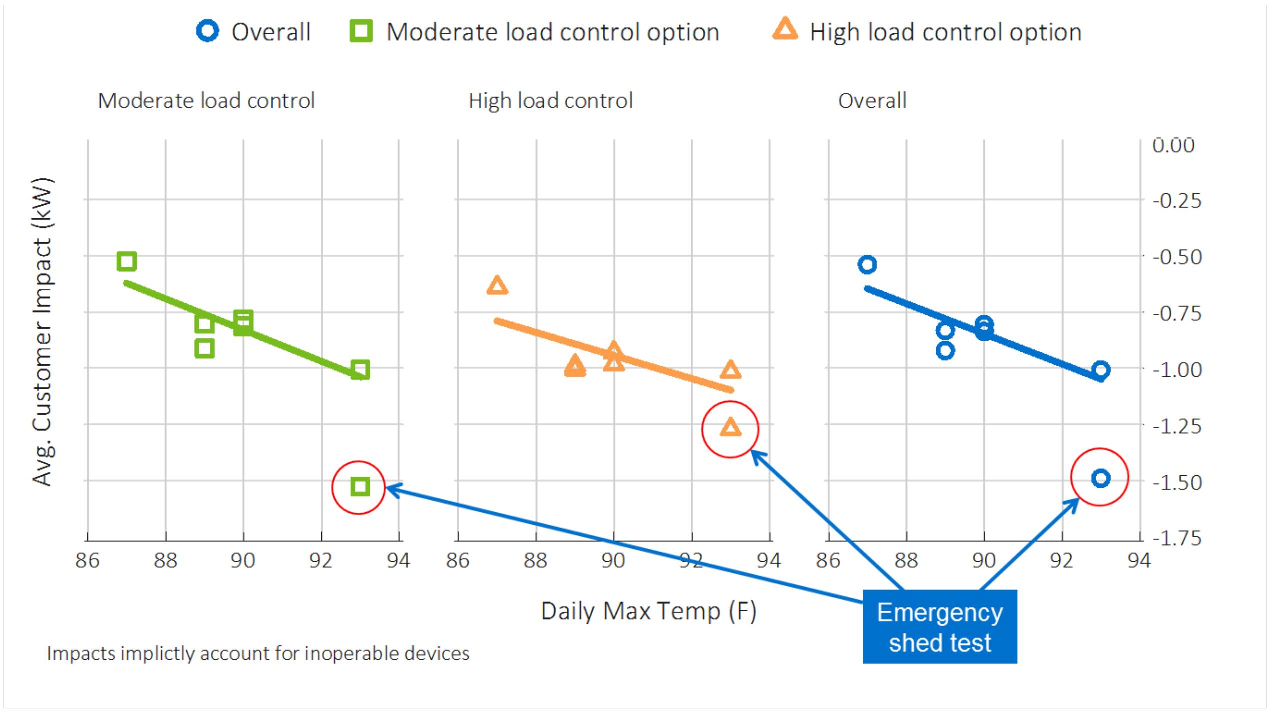


Table 1-1: Randomized Control Trial Demand Reductions for Individual Events

Event Date	Start Time	End Time	Load without DR	Impact	Std. Error	90% Confidence Interval		% Impact	90% Confidence interval		Daily Max	Avg. Daily Temp
						Lower bound	Upper bound		Lower Bound	Upper Bound		
7/21/2016	11:30 AM	2:00 PM	2.84	-0.67	0.05	-0.59	-0.74	-23.4%	-20.7%	-26.2%	90	80.3
	1:30 PM	4:00 PM	3.24	-0.73	0.05	-0.65	-0.80	-22.4%	-20.0%	-24.8%		
	3:30 PM	6:00 PM	3.59	-0.84	0.04	-0.78	-0.90	-23.3%	-21.6%	-25.0%		
	5:30 PM	7:00 PM	3.64	-0.82	0.05	-0.74	-0.91	-22.6%	-20.4%	-24.9%		
	6:30 PM	8:00 PM	3.50	-0.74	0.05	-0.65	-0.82	-21.0%	-18.7%	-23.4%		
7/22/2016	2:30 PM	5:00 PM	2.87	-0.54	0.03	-0.50	-0.58	-18.9%	-17.4%	-20.3%	87	79.9
7/25/2016	3:30 PM	6:00 PM	3.86	-1.01	0.04	-0.95	-1.07	-26.1%	-24.5%	-27.7%	93	83.0
	4:00 PM	5:00 PM	3.82	-1.49	0.05	-1.41	-1.57	-39.0%	-36.9%	-41.1%		
8/25/2016	3:30 PM	6:00 PM	3.52	-0.81	0.04	-0.75	-0.87	-23.0%	-21.3%	-24.6%	90	81.7
8/29/2016	3:30 PM	6:00 PM	3.39	-0.83	0.04	-0.77	-0.89	-24.6%	-22.8%	-26.3%	89	78.8
9/7/2016	3:30 PM	6:00 PM	3.52	-0.92	0.04	-0.86	-0.98	-26.2%	-24.6%	-27.9%	89	78.9
Average General Population Event			**3.42**	**-0.79**	**0.02**	**-0.77**	**-0.82**	**-23.2%**	**-22.4%**	**-24.0%**	**89**	**80.4**

1.2 Device Operability and Site Performance

A significant opportunity to improve load control programs is by identifying nonperforming devices or sites. These can be due to broken or disconnected control devices or because some devices fail to receive control event paging signals. They also can occur because of broken air conditioner units or because some customers do not use their air conditioners during event hours. As part of the evaluation, Nexant undertook two distinct initiatives to identify underperforming sites. First, a field study was implemented during which 103 devices were inspected and tested, with the goal of quantifying the share of inoperable devices. This estimate, however, does not factor in paging network communication failures or sites that do not have their air conditioner on during event hours. Second, Nexant used smart meter data in conjunction with data analytics to identify sites that underperform or do not deliver demand reductions and delivered to Duke Energy Ohio (DEO) a dataset with 39,627 customers, identifying which sites were underperforming and likely had missing or failing devices, paging network issues, or did not use air conditioning during afternoon hours on hotter days.

Key findings from the device operability and site performance analysis include:

- Based on field tests, 95 out of the 103 (92.2%) devices tested were operational, with a 90% confidence interval of ±4.34%.

- Most sites with inoperable devices have multiple failures.
- Not all customers demonstrated a load reduction pattern during events. The event day load profiles suggest that 6,956 of the 39,627 (17.6%) sites analyzed did not exhibit a demand reduction pattern. This can be due to failing or missing devices, paging network issues, or lack of air conditioner loads.
- Efforts to inspect paging network strength and repair devices are missing or failing should focus on larger customers. They are less prone to misdiagnosis and more cost effective.

1.3 Process Evaluation Key Findings

The process evaluation was designed to inform efforts to continuously improve programs by identifying program strengths and weaknesses, opportunities to improve program operations, program adjustments likely to increase overall effectiveness, and sources of satisfaction or dissatisfaction among participating customers. The process evaluation consisted of interviews with key program managers and implementers, post-event surveys implemented immediately after events, and control day surveys implemented on days with similar temperature but when customer's air conditioners were not established.

Key findings from the process evaluation include:

- 121 Power Manager participants were interviewed within 24 hours of the July 21 event, which had a high temperature of 91°F with a heat index of 95°F.
- 92 Power Manager participants were interviewed during a hot nonevent day (a control day), July 14, which had a high of 88°F with a heat index of 92°F. The control day surveys were used to establish a baseline for comfort, event awareness, and other key metrics.
- A strong majority of all respondents, 75%, reported that they are familiar with the Power Manager program.
- Only 13% of respondents on the event day reported that their homes were uncomfortable, while all of them experienced a load control event that afternoon. By comparison, 7% of Power Manager customers surveyed on a hot nonevent day reported they felt uncomfortably hot. While more respondents of the post-event survey stated that their home was uncomfortable that day than respondents of the nonevent survey (13% vs. 7%, respectively), the difference is not statistically significant and the difference in reported thermal discomfort cannot be ascribed to the Power Manager event.
- Over three quarters of participants would recommend the Power Manager program to others.
- The Power Manager staff and vendors are customer focused and undertake a number of activities both during the load control season and afterward to ensure that participants are satisfied with their Power Manager program experience.

2 Introduction

This report presents the results the 2016 Power Manager impact and process evaluation for the Duke Energy Ohio (DEO) territory. Power Manager is a voluntary demand response program that provides incentives to residential customers who allow DEO to reduce the use of their central air conditioner's outdoor compressor and fan on summer days with high energy usage.

Because DEO has full deployment of smart meters and access to Power Manager customers' interval data, the impact evaluation is based on a randomized control trial that randomly assigned customers to six different groups. During each event, at least one of the groups was withheld to serve as a control group and provide an estimate of customer's energy profiles absent activation of Power Manager. The randomized control trial was employed during normal Power Manager operations and during specific tests designed to address key research questions. In addition, DEO launched a field study to collect air conditioner end use data, assess device operability, and assess if customers were compensating for the air conditioner curtailments by increasing utilization of fans or other end uses.

The process evaluation was designed to inform efforts to continuously improve the program by identifying program strengths and weaknesses, opportunities to improve program operations, program adjustments likely to increase overall effectiveness, and sources of satisfaction or dissatisfaction among participating customers. The process evaluation consisted of interviews with key program managers and implementers, post-event surveys implemented immediately after events, and control-day surveys implemented on days with similar temperature but when customer's air conditioners were not controlled by the Power Manager program.

2.1 Key Research Questions

The study data collection and analysis activities were designed to address the main impact evaluation and process evaluation research questions.

Impact Evaluation Research Questions

- What were the demand reductions achieved during each event called in 2016?
- Did impacts vary for customers in moderate (1.0 kW) and high (1.5 kW) load control options?
- Were impacts at the whole building level (net) different from AC end use demand reductions (gross)?
- Do impacts vary based on the hours of dispatch and/or weather conditions? If so, how?
- What is the device failure rate?

Process Evaluation Research Questions

- What is the extent to which participants are aware of events, bill credits, and other key program features?
- What is the participant experience during events?
- What are the motivations and potential barriers for participation?
- What are the processes associated with operations and program delivery?
- What are program strengths and areas for potential improvement?

2.2 Program Description

Power Manager is a voluntary demand response program that provides incentives to residential customers who allow DEO to reduce their central air conditioner's outdoor compressor and fans on summer days with high energy usage. All Power Manager participants have a load cycling switch device installed on at least one outdoor unit of qualifying air conditioners. The device enables the customer's air conditioner to be cycled off and on to reduce load when a Power Manager event is called. DEO initiates events by sending a signal to all participating devices through a corporate paging network. The signals instruct the switch devices to cycle the air conditioning system on and off, reducing the run time of the unit during events.

The program participates in the energy and capacity markets of the PJM market, but DEO generally limits participation in the energy market to days when the wholesale price exceeds $65/MWh. Duke regularly bids Power Manager into the capacity market, which means that the program must be available for PJM emergency events. Absent an emergency, the DEO operations team schedules and calls events for local emergency, economic, or testing reasons.

Power Manager events typically occur between May and September in DEO territory. Participants receive financial incentives for their participation that depend on the amount of load control they experience during an event. At enrollment, Power Manager customers elect one of two load control options that are available—moderate or high load control. Approximately 85% of Power Manager devices in DEO are enrolled in the normal option and the remaining 15% are enrolled in the higher load control option.[2] The payments received by participants include a one-time installation credit of $25 for the moderate load control option ($35 for high load control) plus bill credits for each cycling event that occurs. The minimum bill credit for 2016 participation was $5 for customers enrolled in the moderate option and $8 for customers enrolled in the high load control option.

Starting in 2017, DEO will begin using a new cycling algorithm known as *true cycle algorithm*. The algorithm uses learning days to estimate the run time (or duty cycle) of air conditioners as a function of hour of day and temperature at each specific site and aims to curtail use by a specified amount. Events in 2016, however, were based on the target cycle algorithm.

2.3 Participant Characteristics

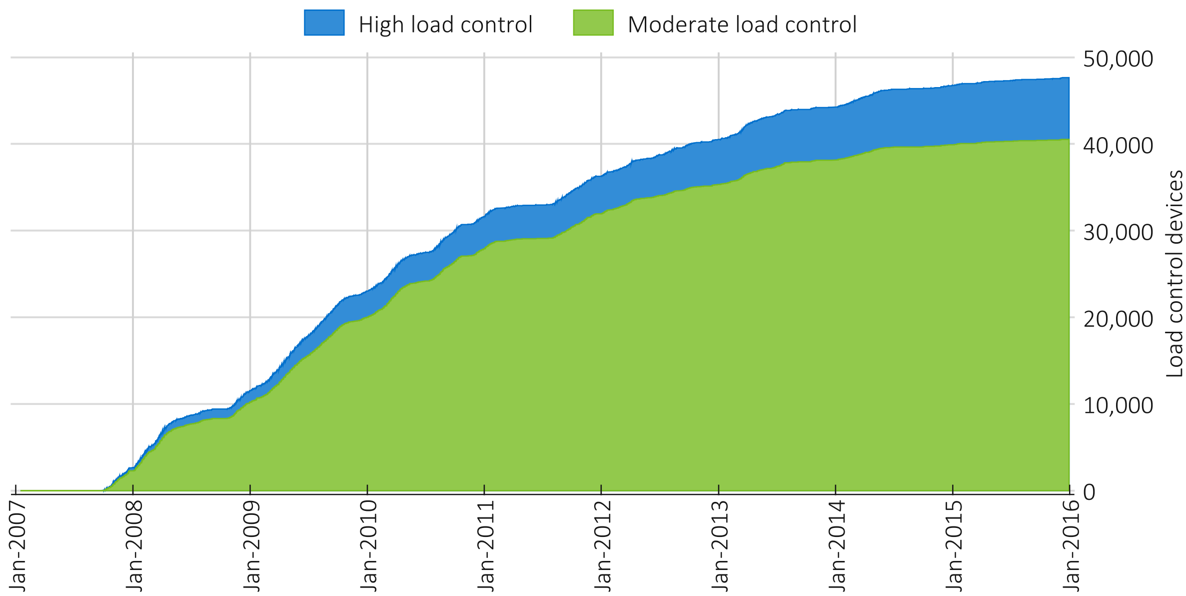
The Duke Energy Ohio service territory is in the Southern portion of Ohio and centered in the Cincinnati area. By the end of summer 2016, slightly over 48,105 air conditioner units were part of Power Manager. Of those units, 14.9% enrolled in the higher load control option. On average, customers enroll 1.06 air conditioner units per site.

DEO serves approximately 760,000 residential customers. To enroll on Power Manager, customers must be in DEO territory, own their single family home, and have a functional central air conditioning unit with

[2] Customers who ask to de-enroll are offered a low load control option to minimize attrition. Less than 1/15th of one percent of devices are enrolled in the low load control option.

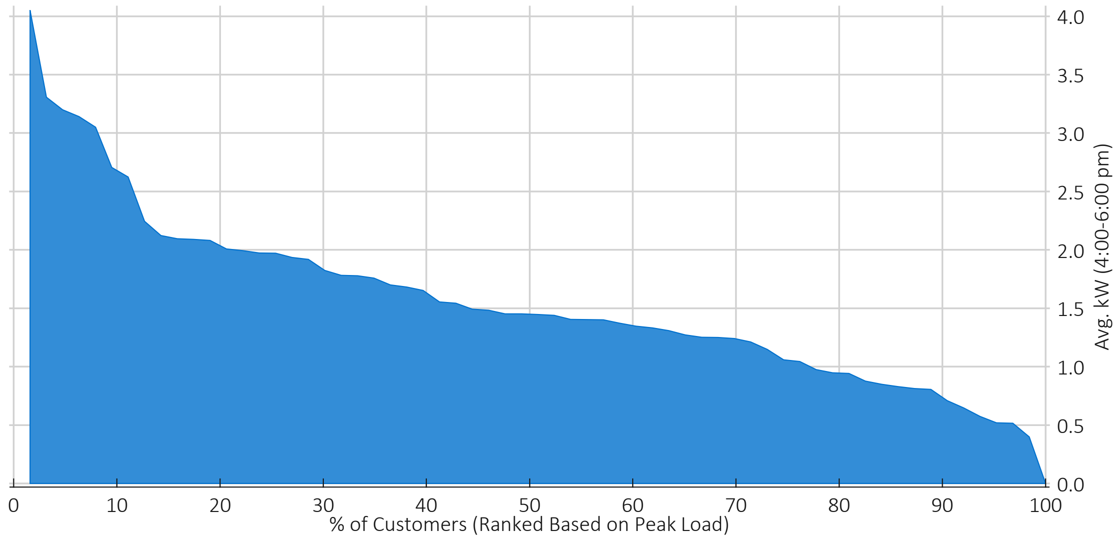
an outside compressor. Based on the program rules and a residential appliance saturation survey Duke Energy implemented in 2013, approximately 54.7% of customers meet the eligibility criteria.[3] To date, DEO has enrolled approximately 10.9% of eligible customers. Figure 2-1 visualizes enrollment in Power Manager over time.

Figure 2-1: Power Manager Participation Over Time



3 71.4% of residential customer in the territory own single family homes and, of those, 76.6% have central air conditioners. The estimate does not include heat pumps.

Figure 2-2: Distribution of Air Conditioner Peak Period Loads

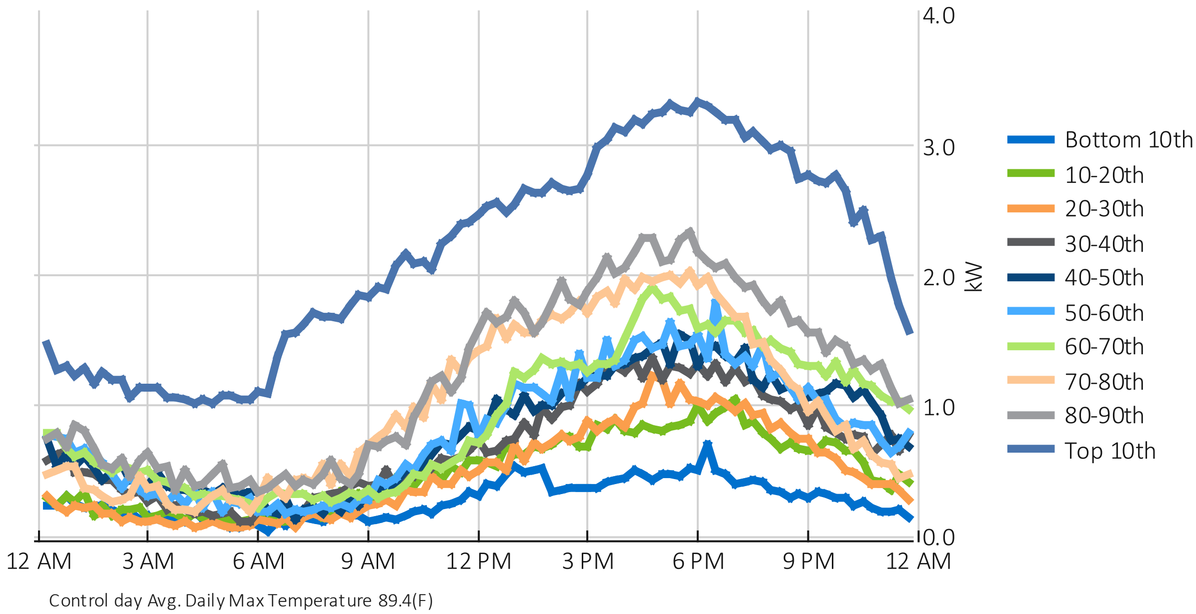


One of the advantages of end use data collection is the ability to assess whether customers use their air conditioner units during key hours on hotter days. By design, the events were not called on all of the hottest days, enabling us to assess air conditioner use absent load curtailment events. A total of 19 nonevent days were identified with daily maximum temperatures exceeding 88˚F and averaging 89.4˚F (vs. 89.7˚F for actual events). For simplicity, these days are referred to as control days.

Figure 2-2 shows the distribution of air conditioner demand across customers on hot nonevent days. We isolated the 4 to 6pm period because it aligns with the time period for most Power Manager events. Air conditioner use by Power Manager participants varied substantially, reflecting different occupancy schedules, comfort preferences, and thermostat use and settings. Roughly 40% of air conditioner loads exceeded 1.5 kW. As with any program, some customers who enrolled use little or no central air conditioners during late afternoon hours on hotter days. They are, in essence, free riders. The bulk of the costs for recruitment, equipment, and installation have already been sunk for these customers and, as a result, removing these customers may not improve cost effectiveness substantially. However, given the availability of smart meter data, we recommend assessing nonparticipant afternoon loads on hotter days prior to marketing in order to target customers who are cost effective to enroll.

Figure 2-3 provides additional detail and shows the hourly air conditioner end use loads for different customer groups. The customers were classified into 10 equally sized groups, known as deciles, based on their air conditioner use during hot nonevent days. Each line represents the hourly air conditioner loads for the average customer in each decile.

Figure 2-3: Air Conditioner End-use Hourly Loads by Size Decile



2.4 2016 Event Characteristics

In 2016, DEO dispatched Power Manager six times in addition to the PJM test event. The general population events occurred between 3:30 and 6:00pm, except for the July 22, 2016 event, which started at 2:30 and lasted until 5:00pm. DEO bids Power Manager resources into the PJM market during those time periods. The PJM event was prescheduled well in advance and happened to land on a cooler day with a daily maximum temperature of 78˚F. During a PJM event, Power Manager customer loads needed to be less than the peak load contribution (PLC) minus the magnitude of DR resources bid into the capacity market.

Table 2-1: 2016 Event Operations and Characteristics

Event Date	Start Time	End Time	Daily Max (˚F)	Type of Event	# of Devices	Devices dispatched	Control group	Notes
7/21/2016	3:30 PM	6:00 PM	90	General Population (GP)	48,178	36,134	2,409	Group 3 held back as control
	11:30 AM	2:00 PM				2,409	2,409	Group 1 dispatched
	1:30 PM	4:00 PM				2,409	2,409	Group 2 dispatched
7/21/2016	3:30 PM	6:00 PM	90	Research	48,178	36,134	2,409	Group 0 dispatched
	5:30 PM	7:00 PM				2,409	2,409	Group 4 dispatched
	6:30 PM	8:00 PM				2,409	2,409	Group 5 dispatched
7/22/2016	2:30 PM	5:00 PM	87	GP Event	48,178	43,360	2,409	Groups 1 and 2 held back

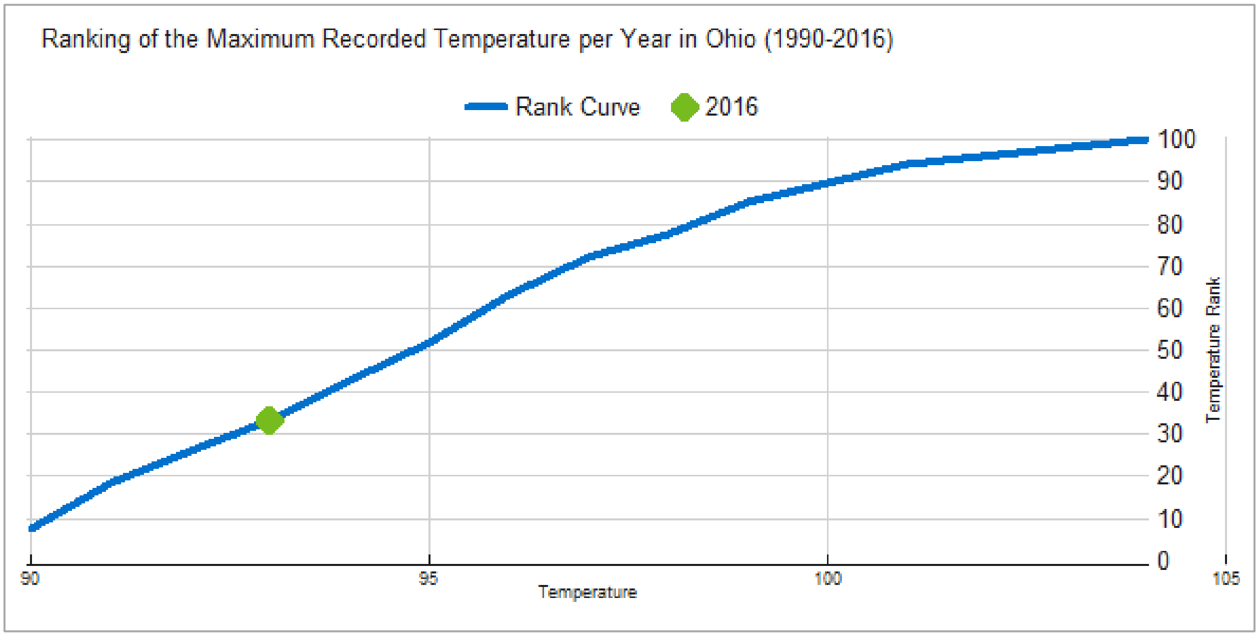
7/25/2016	3:30 PM	6:00 PM	93	GP Event + Emergency shed	48,178	43,360	2,409	Group 2 held back
8/25/2016	3:30 PM	6:00 PM	90	GP Event	48,128	45,722	2,406	Group 4 held back
8/29/2016	3:30 PM	6:00 PM	89	GP Event	48,128	45,722	2,406	Group 5 held back
9/1/2016	4:00 PM	5:00 PM	78	PJM System Test Event	48,108	48,108	0	All customers dispatched
9/7/2016	3:30 PM	6:00 PM	89	GP Event	48,105	45,700	2,405	Group 1 held back

In addition, DEO overlaid two research experiments alongside the general population events on July 21 and July 25. On July 21, DEO implemented a side-by-side test of five groups to assess if and how demand reductions varied for different dispatch periods. On July 25, a research group was dispatched using emergency shed operations side-by-side with a control group and a group that experienced normal operations. The objective was to assess how the magnitude of the emergency shed compares to traditional operations. Emergency operations reflect the full demand reduction capability of the program, but are employed judiciously.

With the exception of emergency shed tests, the control of the air conditioner units is phased in, at random, over the first 30 minutes. Likewise, at the end of an event, instructions to resume normal operations are gradually sent to individual air conditioners. The demand reductions reported in this study are for the time period when units' full load reduction were achieved—that is, the phase in and phase out periods are excluded since they do not reflect the demand reduction capability.

In comparison to the past 25 years, 2016 was a relatively cool year in DEO territory. Figure 2-4 shows how the maximum temperature in 2016 compares to historical annual maximum temperatures. Overall, nearly 70% of historical years experienced hotter temperatures.

Figure 2-4: Comparison of 2016 Maximum Temperature to Historical Years



3 Methodology and Data Sources

This section details the study design, data sources, sample sizes, and analysis protocols for both the impact and process evaluations. For clarity, we provide details about the methodology separately for the impact and process evaluation.

3.1 Impact Evaluation Methodology

The 2016 Power Manager evaluation included four main activities designed to meet the research objectives. The primary evaluation results are based on a randomized control trial, which included the entire Power Manager participant population in Ohio. The additional data collection and analysis were supplemental and designed to address specific research questions. Because of this, the focus of the methodology discussion is on the randomized control trial design and analysis.

Table 3-1: Summary of Impact Evaluation Components

Evaluation Component	Description
Randomized control trial using smart meters	▪ Primary evaluation results ▪ Population with addressable devices (39,804) randomly assigned to six groups, one with 75% of population and five research groups, each with 5% of the population ▪ During events, at least one group withheld to serve as a control group and establish the baseline ▪ Comparison of means between treatment and control
Air conditioner end use meter sample	▪ Data loggers installed on 95 devices at 89 households, 61 households used for analysis[4] ▪ Spot measurements of voltage, amps, kW, and connected load conducted at 54 sites ▪ Used to compare end use to whole building demand reductions and assess if customers compensated for air conditioner curtailments ▪ Used nonevent days to infer the baseline ▪ Regression model selected based on out of sample testing of multiple models
Whole building data for customers with end use metered air conditioners	▪ Whole house meter installed for same household with air conditioner end use data loggers ▪ Used to compare end use to whole building demand reductions and assess if customers compensated for air conditioner curtailments ▪ Used nonevent days to infer the baseline ▪ Regression model selected based on out of sample testing of multiple models
Device operability inspections and analysis	▪ Field inspection of 103 devices, 8 (7.8%) were inoperable ▪ Event day shape analysis for all customers to identify devices that are and are not curtailing loads during events

[4] Some device loggers either did not record data for the full summer or did not download data. Expected losses are around 25% but were higher in Duke Ohio

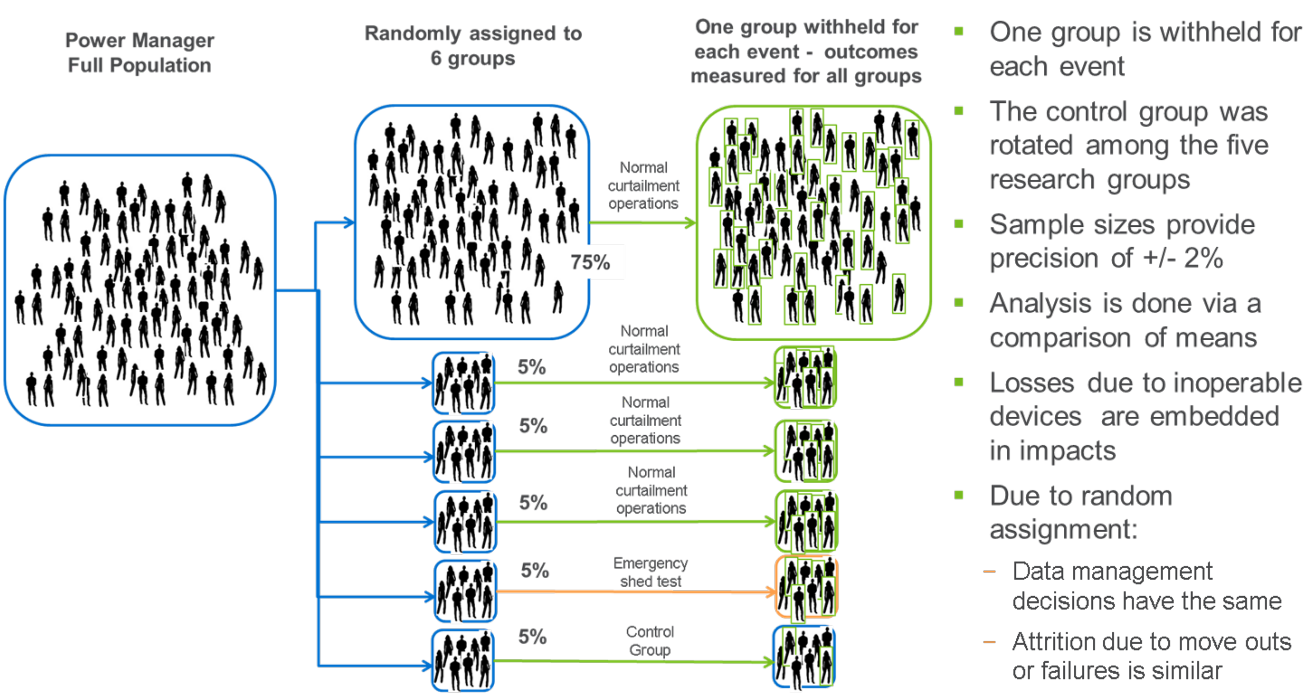
3.2 Randomized Control Trial Design and Analysis

Randomized control trials are well recognized as the gold standard for obtaining accurate impact estimates and have several advantages over other methods:

- They require fewer assumptions than engineering-based calculations;
- They allow for simpler modeling procedures that are effectively immune to any kind of model specification error; and
- They are guaranteed to produce accurate and precise impact estimates with proper randomization and large sample sizes.

The RCT design randomly separated the DEO Power Manager population into two groups—treatment and control—for each event day. On an event day, all load control devices in the treatment group were activated, while none of the devices in the control group were activated. Because of random assignment, the only systematic difference between the two groups is that one set of customers was curtailed and the other group was not. During research events, distinct operation strategies were employed to enable side-by-side testing, but in all instances a control group was withheld. Figure 3-1 shows the conceptual framework of the random assignment.

Figure 3-1: Randomized Control Trial Design



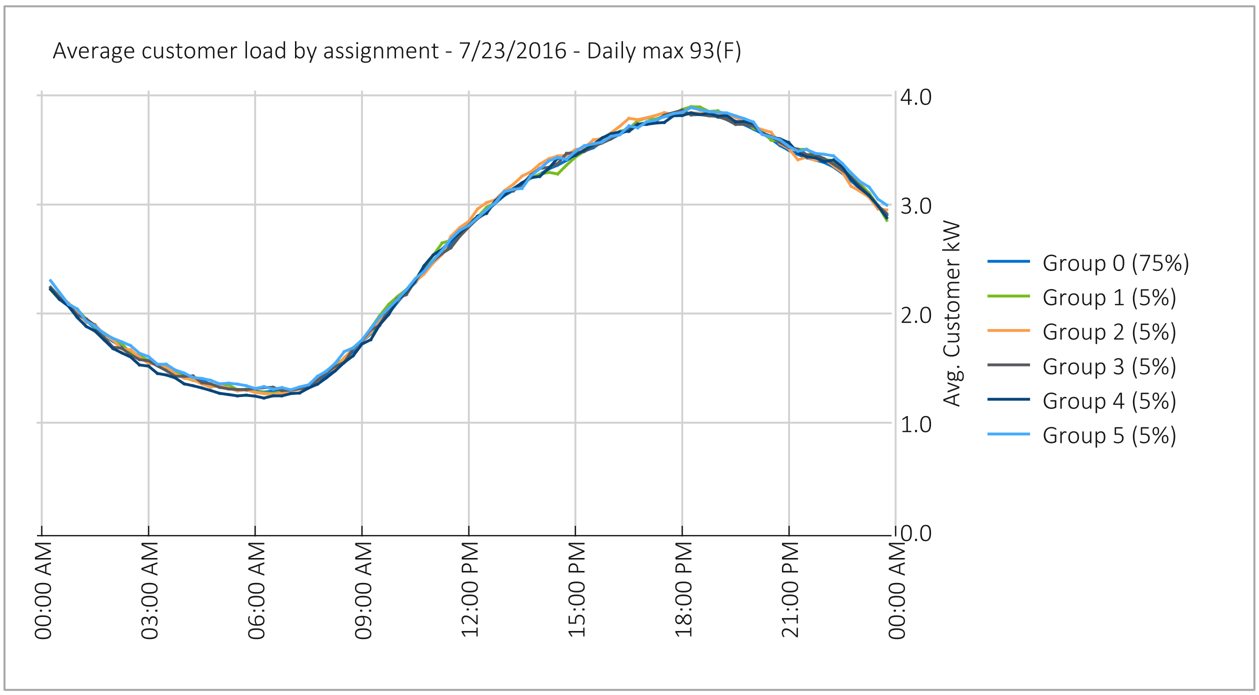
The Power Manager participant population with addressable load control devices was randomly assigned into six distinct groups prior the 2016 summer based on the last two digits of the device serial number.[5]

[5] Some households have multiple load control devices. In these instances the homes were randomly assigned and all devices are home were in the same group.

At the beginning of the summer, the main general population group includes 75% of participants – approximately 30,000 participants. The remaining five research groups each include 5% of participants, or roughly 2,000 customers each. Before implementation, Nexant conducted simulation based power analysis using smart meter data for load control participants and concluded the sample sizes were sufficient to provide a ±2% Margin of Error with 90% confidence. The purpose of creating six distinctive randomly assigned groups was twofold. First, it allowed side-by-side testing of cycling strategies, event start times, or other operation aspects to help optimize the program. Second, it also allowed DEO to alternate the control group, increasing fairness but also helping avoid exhausting individual customers by dispatching them too often solely for research purposes.

To ensure the randomization was properly implemented, the loads for each of the six groups were compared to each other on all days when none of the groups experienced an event. Figure 3-2 shows hourly loads for each group on the hottest, nonevent day (July 23). The customer loads are nearly identical, which provides strong evidence that the assignment of devices into the six different groups was indeed random. It also reflects the precision of control group as a method for estimating the counterfactual.

Figure 3-2: Validation of Random Assignment and Precision — Loads on the Hottest Nonevent Day



For each event, one of the five research groups was withheld to serve as a control group and establish the counterfactual or baseline—the electricity load patterns in the absence of curtailment. Within the experimental framework of an RCT, the average usage for control group customers provides an unbiased estimate of what the average usage for treatment customers would have been if an event had not been called. Because of this, estimating the load impacts for an event requires simply calculating the difference in loads between the treatment and control groups during each 15 minute interval, including the event period and hours following the event when snapback can occur. The demand reductions reflect net impacts and account for customer use of fans to compensate for curtailment of air conditioners, device failures, and paging network communication issues.

The standard error, used to calculate the confidence bands, is calculated using the formula shown in Equation 1.

Equation 1: Standard Error Calculations for Randomized control trial

$$Std.\,Error\ of\ Difference\ between\ Means_i = \sqrt{\frac{sd_c^2}{n_c} + \frac{sd_t^2}{n_c}}$$

Where *sd* is the stand deviation, *n* is the sample size, *t* and *c* are the treatment and control groups respectively, and *i* refers to individual time intervals

3.3 Analysis Protocol for End Use Metered Customers

As noted earlier, the DEO study also included end use metering for a sample of 95 air conditioner units at 89 households. The main purpose was to asses if whole house demand reductions matched end use demand reductions or if customers were compensating for air conditioner curtailments by increasing the use of fans or other equipment.

Nexant used regression analysis to model the relationship between weather and demand on nonevent days in order to establish what customer energy use patterns would have been absent curtailments—known as the counterfactual. This approach works because the intervention—air conditioner curtailments—is introduced on some days and not on others, making it possible to observe load patterns with and without demand reductions. The repeated, or ON/OFF pattern, enables the evaluator to assess whether the outcome—electricity use—rises or falls with the presence or absence of event dispatch instructions. This approach hinges on having comparable nonevent days. When all of the hottest days are event days, the counterfactual is based on extrapolating trends beyond the range of nonevent temperatures, producing less accurate and less unreliable impact estimates for the hottest days. By design, DEO avoided dispatching Power Manager resources on all of the hottest days.

Figure 3-3 illustrates the underlying concept using actual DEO end use load data. The blue circles reflect the individual nonevent weekdays and the orange line shows the trend between peak hour loads and weather. The green X's show the load during event days. The regression modeling calculates the demand reduction as the difference between the estimated loads, absent air conditioners and actual loads during event days. The example below is simplified for illustration purposes. In practice, regression modeling typically includes other explanatory variables besides weather such as day of week effects and seasonal, or month effects.

A key question every evaluator must address is how to decide which model produces the most accurate and precise counterfactual. In many instances, multiple counterfactuals are plausible but provide different estimated demand reductions. The model selection was based on testing 10 distinct model specifications and employing a systematic approach to identify the most accurate and precise estimation method, described in Appendix A.

Figure 3-3: Illustration of Within-subject Regression Models with 2016 Duke Energy Ohio Data

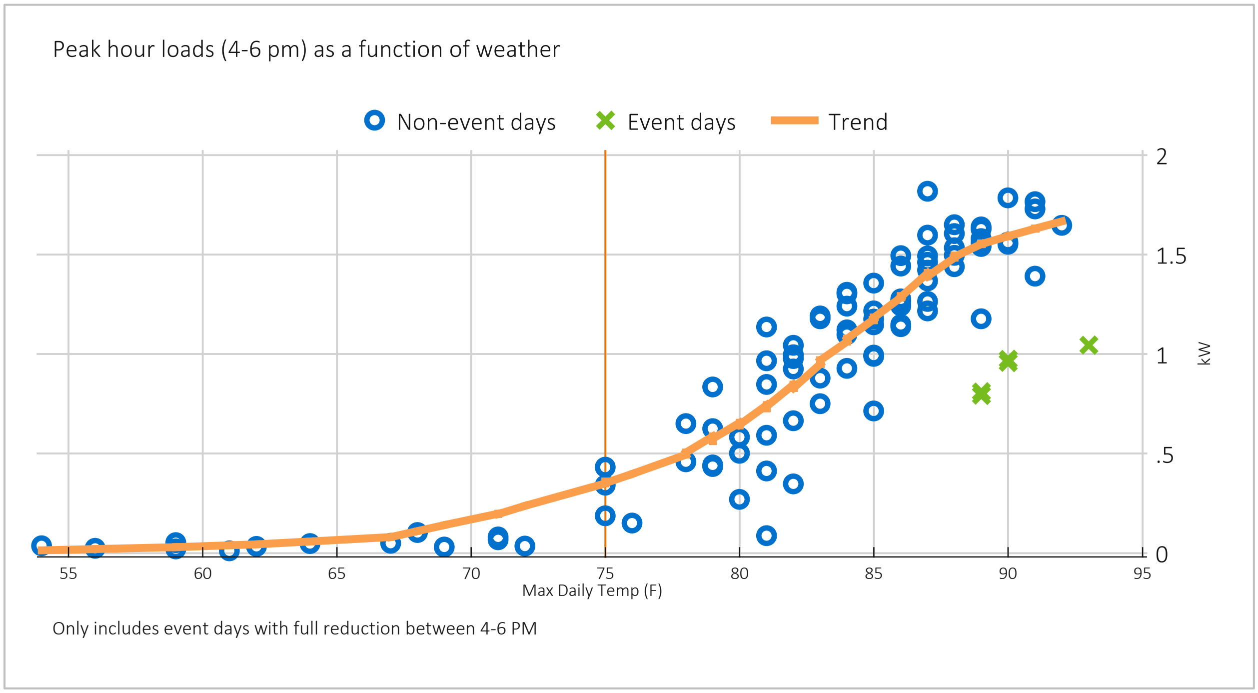
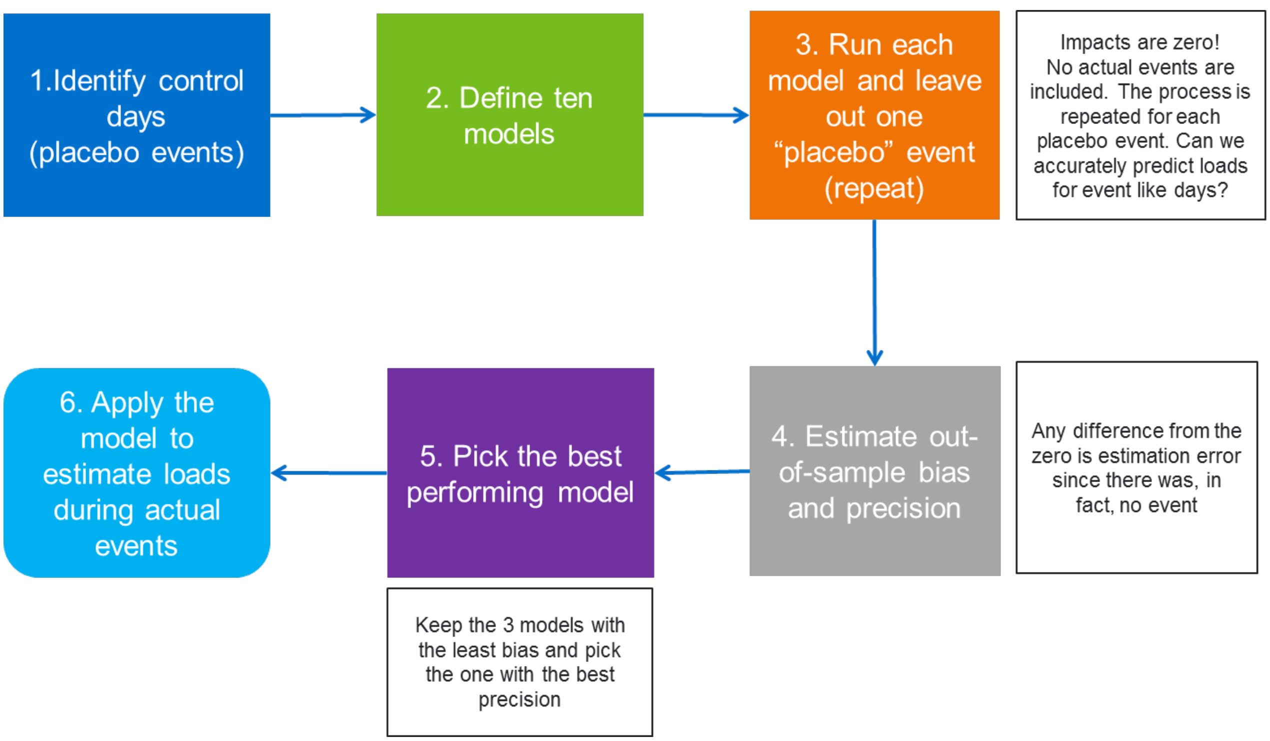


Figure 3-4: Model Out-of-Sample Validation and Selection



The process for model selection relied on out-of-sample placebo tests. First, the model specifications are defined. The models tested are summarized in Appendix A. Second, hotter days when air conditioners were not curtailed are defined as placebo days. Because load control devices were not activated during those days, the impacts are by definition zero. Any estimated impact by models is in fact due to model error. Third, we run each of the 10 models using nonevent data but leave out a single placebo event. The regression model is used to predict electricity use on the withheld placebo event day—an out-of-sample prediction. We repeated the process for each placebo event and record the actual and predicted loads for each placebo event. For DEO, a total of 20 placebo days were employed. Fourth, the out-of-sample predictions are compared to actual electricity use observed on that day, which is used to calculate metrics for bias and precision. Fifth, the best model is identified by first narrowing the candidate models to the three with least bias and then selecting the model with the highest precision. Finally, the best performing model is used to estimate the counterfactual for actual event days.

Table 3-2 summarizes metrics for bias and precision.[6] Table 3-3 summarizes the results for each model tested. Bias metrics measure the tendency of different approaches to over or under predict and are measured over multiple days. The mean percent error describes the relative magnitude and direction of the bias. A negative value indicates a tendency to under predict and a positive value indicates a tendency to over predict. This tendency is best measured using multiple days. The precision metrics describe the magnitude of errors for individual events days and are always positive. The closer they are to zero, the more precise the results. The mean percentage error was used to narrow down to the three models with the least bias. The CV(RMSE) metric was used to identify the most precise and final model among the remaining candidates. The best performing model (#7) incorporated both the temperature during the time period and the heat buildup in six hours immediately prior.

Table 3-2: Definition of Bias and Precision Metrics

Type of Metric	Metric	Description	Mathematical Expression
Bias	Average Error	Absolute error, on average	$AE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)$
	Mean Percentage Error (MPE)	Indicates the percentage by which the measurement, on average, over or underestimates the true demand reduction.	$MPE = \frac{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)}{\bar{y}}$
Precision	Root mean squared error	Measures how close the results are to the actual answer in absolute terms, penalizes large errors more heavily	$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$
	CV(RMSE)	Measures the relative magnitude of errors across event days, regardless of positive or negative direction. It can be though us as the typical percent error, but with heavy penalties for large errors.	$CV(RMSE) = \frac{RMSE}{\bar{y}}$

[6] Bias is also referred to as accuracy. Precision is sometimes called goodness-of-fit.

Table 3-3: Out of Sample Bias and Precision Metrics for Each Model Tested

Model	Variables	End-Use Bias: Avg. Error	End-Use Bias: Mean Percent Error	End-Use Precision: Root mean square error	End-Use Precision: Normalized RMSE	Whole building Bias: Avg. Error	Whole building Bias: Mean Percent Error	Whole building Precision: Root mean square error	Whole building Precision: Normalized RMSE
1	- Pre-event load (11 am to 1 pm) - Cooling degree hours (Base 70F) - Day of week and month	-0.01	-0.7%	0.15	10.7%	-0.02	-0.6%	0.23	7.7%
2	- Pre-event load (11 am to 1 pm) - Cooling degree days (Base 65F) - Day of week and month	-0.02	-1.3%	0.16	11.7%	-0.01	-0.2%	0.24	7.9%
3	- Pre-event load (11 am to 1 pm) - Maximum temperature for day - Day of week and month	-0.01	-0.7%	0.17	12.2%	0.00	-0.1%	0.24	7.8%
4	- Pre-event load (11 am to 1 pm) - Avg. temperate in prior 24 hours - Day of week and month	-0.03	-2.2%	0.17	12.4%	-0.03	-1.1%	0.24	8.0%
5	- Pre-event load (11 am to 1 pm) - CDH and CDD - Day of week and month	-0.01	-0.6%	0.15	10.8%	-0.01	-0.2%	0.24	7.8%
6	- Pre-event load (11 am to 1 pm) - Avg. temperate in prior 24 hours and current CDH - Day of week and month	-0.01	-0.7%	0.15	10.9%	-0.01	-0.4%	0.24	7.8%
7	- Pre-event load (11 am to 1 pm) - Average CDH in prior 6 hours and current CDH - Day of week and month	0.00	0.0%	0.14	10.4%	0.00	0.1%	0.23	7.5%
8	- Pre-event load (11 am to 1 pm) - Average CDH in prior 12 hours and current CDH - Day of week and month	0.00	-0.3%	0.15	10.7%	0.00	0.0%	0.23	7.7%
9	- Pre-event load (11 am to 1 pm) - Average CDH in prior 18 hours and current CDH - Day of week and month	-0.01	-0.6%	0.15	10.7%	-0.01	-0.2%	0.24	7.7%
10	- Pre-event load (11 am to 1 pm) - Average CDH in prior 24 hours and current CDH - Day of week and month	-0.01	-0.7%	0.15	10.9%	-0.01	-0.4%	0.24	7.8%

3.4 Device Operability Testing Protocols

As part of the study, Nexant was responsible for all fieldwork related to recruiting customers for end use data collection and installation and collection of data loggers. The customers were recruited from a random sample of the Power Manager participant population. Prior to installing data loggers on air conditioners, Nexant tested whether load control devices were functional.

The inspection consisted of:

- Onsite spot measurements of the kW, voltage, amperage, and power factor;
- Information about the AC unit;
- Inspection of the load control device for presence, proper installation, physical condition, and operability; and
- Inspection of the load control device connection wires, including presence, physical condition, and whether the connection was secure.

End use data loggers were only installed on air conditioner units with functional load control devices. In total, 89 out of the 95 (93.7%) devices inspected had functional load control devices.

3.5 Process Evaluation Methodology

The process evaluation included four main activities:

- A survey of Power Manager participants in the 24 hours immediately following an event;
- A survey of Power Manager participants on a hot, nonevent day (a control day). By design, the survey mirrored the event day survey and served to establish the baseline response, absent curtailments, for customer responses about comfort, awareness, and other program features;
- Interviews with program managers and implementers; and
- A review of the data files, enrollment, and operation processes.

Table 3-4 lists the overarching objectives and the related process evaluation activity. Data collection included a mix of interviews and surveys designed to obtain information sufficient to understand the experience of Duke Energy staff, implementation staff, and participating customers. Surveys included both post-event and nonevent data collection to enable comparison of participant responses with regard to comfort level, other cooling strategies, and the extent to which these experiences are attributed to Power Manager. Based on power analysis, Nexant concluded that the sample of 68 participants during the event and the hot nonevent days was sufficient to deliver 90/10 precision. By design, the goal was to collect 50% of the responses via telephone and 50% online. In practice, the survey targets were exceeded by a substantial amount because the response to telephone surveys was quick, with most phone surveys completed within three hours immediately after the event. Online response, on the other hand, tended to occur late at night or early in the morning.

Table 3-4: Process Evaluation Research Objectives and Data Sources

Process Evaluation Objective	Interviews with Key Contacts	Post-event Survey	Nonevent Survey
Assess customer awareness		✓	✓
Understand participant experience	✓	✓	✓
Identify potential barriers to participation		✓	✓
Document current program processes associated with recruitment, enrollment, and curtailment	✓		
Identify program strengths and potential improvements	✓	✓	✓

4 Randomized Control Trial Results

The goals of this study include understanding the load impacts associated with the Power Manager program under a variety of conditions. General population event dates were selected to understand the available load reduction capacity under a variety of temperature conditions during normal operations, while emergency shed events demonstrated the available capacity for short-duration events during extreme conditions. In addition, one test day was used to understand how load reduction capacity varied as a function of dispatch window by signaling different customer groups at different times of day. This section presents the results for these event days. A comparison of load impacts by dispatch option (moderate versus high load control) is also presented.

4.1 Overall Program Results

The load impact estimates derived from the randomized control trial analysis for the general population events, as well as the emergency shed test that occurred side-by-side with normal operation of the program on July 25, are presented in Table 4-1. Results for the July 25 emergency event are presented as a separate event from the general population event. The load impacts presented here, along with the accompanying confidence intervals, are the average changes in load during the indicated dispatch windows, excluding the first 30 minutes of dispatch for the normal operation events since this is the time period when devices are phased-in at random.

Table 4-1: Randomized Control Trial per Customer Impacts

Event Date	Start Time	End Time	Load without DR	Impact	Std. Error	90% Confidence Interval		% Impact	90% Confidence interval		Daily Max	Avg. Daily Temp
						Lower bound	Upper bound		Lower Bound	Upper Bound		
7/21/2016	11:30 AM	2:00 PM	2.84	-0.67	0.05	-0.59	-0.74	-23.4%	-20.7%	-26.2%	90	80.3
	1:30 PM	4:00 PM	3.24	-0.73	0.05	-0.65	-0.80	-22.4%	-20.0%	-24.8%		
	3:30 PM	6:00 PM	3.59	-0.84	0.04	-0.78	-0.90	-23.3%	-21.6%	-25.0%		
	5:30 PM	7:00 PM	3.64	-0.82	0.05	-0.74	-0.91	-22.6%	-20.4%	-24.9%		
	6:30 PM	8:00 PM	3.50	-0.74	0.05	-0.65	-0.82	-21.0%	-18.7%	-23.4%		
7/22/2016	2:30 PM	5:00 PM	2.87	-0.54	0.03	-0.50	-0.58	-18.9%	-17.4%	-20.3%	87	79.9
7/25/2016	3:30 PM	6:00 PM	3.86	-1.01	0.04	-0.95	-1.07	-26.1%	-24.5%	-27.7%	93	83.0
	4:00 PM	5:00 PM	3.82	-1.49	0.05	-1.41	-1.57	-39.0%	-36.9%	-41.1%		
8/25/2016	3:30 PM	6:00 PM	3.52	-0.81	0.04	-0.75	-0.87	-23.0%	-21.3%	-24.6%	90	81.7
8/29/2016	3:30 PM	6:00 PM	3.39	-0.83	0.04	-0.77	-0.89	-24.6%	-22.8%	-26.3%	89	78.8
9/7/2016	3:30 PM	6:00 PM	3.52	-0.92	0.04	-0.86	-0.98	-26.2%	-24.6%	-27.9%	89	78.9
Average General Population Event			**3.42**	**-0.79**	**0.02**	**-0.77**	**-0.82**	**-23.2%**	**-22.4%**	**-24.0%**	**89**	**80.4**

Overall load impacts for the average customer in the test group ranged between 0.54 kW and 1.01 kW during normal operations. The wide range in impacts was primarily driven by differences in temperature during the event days, as will be discussed in more detail in a subsequent section of this report. By design, events were called under different weather conditions and for different operating hours to better estimate the demand reduction capability under different conditions. The emergency shed event had a much higher load impact of 1.49 kW.

Except for the PJM test, at most, 95% of the sites were dispatched since at least 5% of the population was withheld to serve as a control group and establish the baseline. Had all resources been dispatched under normal operation on July 25, the hottest event day, the program would have delivered 46.1 MW. If instead, all resources had been dispatched using emergency operations, reduction would have been 68.0 MW, despite a relatively cool weather year.

Since all of the analysis included customers with inoperable devices, the results implicitly take device inoperability into account. Because we used random assignment, each of the test groups accurately represent the percentage of customers with inoperable devices among the entire population and the estimated load impacts are appropriately de-rated by the nonworking devices included in the test groups.

These same impacts are shown graphically in Figure 4-1, along with the average customer load profiles for the test and control groups. Compared to the control group load profile, there is a clear drop in test group load during the dispatch period, along with a small snapback in energy usage immediately after the events. Note that, based on the control group load profiles, there is more load available for reduction on hotter days.

Figure 4-1: Load Profiles of Average Test and Control Group Customers on General Population Event Days

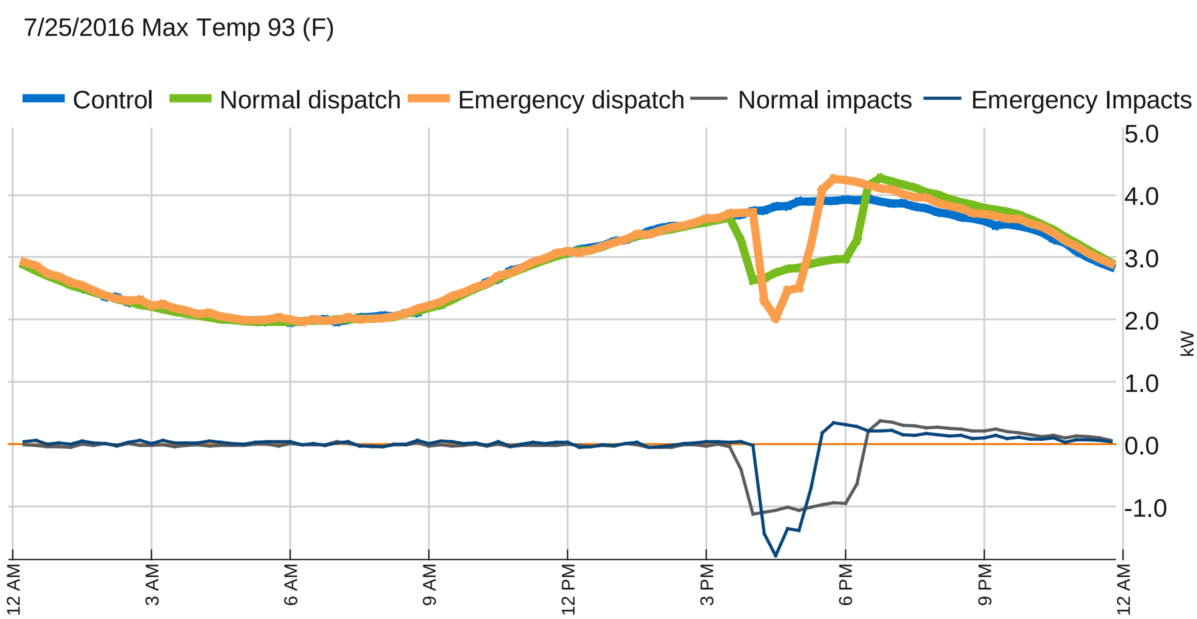
4.2 Normal Operations Versus Emergency Shed Test

Impacts for the July 25 event are presented in Figure 4-2 for both normal and emergency operations. As shown in the graph, the group that was dispatched via normal operations had a 30 minute period (3:30 to 4pm) during which devices were phased in randomly, whereas all of the devices in the emergency shed test group were dispatched simultaneously at the start of the 4pm event and instructed to implement 75% cycling (AC unit is off ¾ of each hour). As a result, the magnitude of the overall load reduction was much greater for customers in the emergency shed group.

Emergency operations produced larger impacts than normal operations, 1.49 kW vs. 1.05 kW per household for the common dispatch hour from 4 to 5pm (average load reduction for normal operations during the entire two hour event window was 1.01 kW). Reductions from emergency operations exceeded those from normal operations by 41.9%.

The emergency shed event ended at 5pm, after which time the load for this dispatch group returned to nearly the same level as the control group, with some additional snapback. The normal operation group continued to show steady load drop until the end of its dispatch window at 6pm.

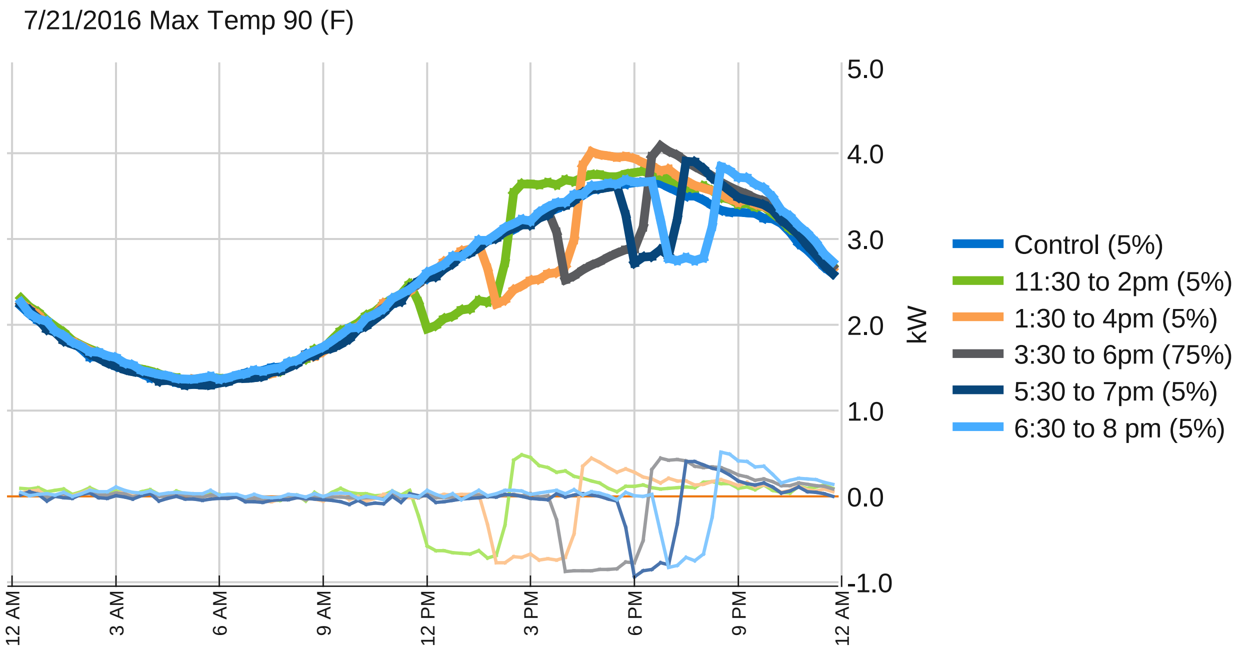
Figure 4-2: Load Profiles for Emergency and Normal Operations on July 25 Event



4.3 Impacts by Dispatch Period

Load profiles for the various test groups for the July 21 cascading event test are presented in Figure 4-3, along with the load profile for the control group. The plot shows the load reduction and accompanying snapback associated with each group's dispatch, as compared to the control group. As can be seen from the plot and from the prior table, there were slight differences in the estimated load impacts with larger per customer impacts occurring in the late afternoon hour, up until the last event which began at 7pm (excluding the 30 minute ramp-in period at the beginning of the event). Impacts during all dispatch windows were fairly steady throughout the events. While the magnitude of impacts varied slightly by dispatch window (between 0.67 and 0.84 kW per household), the percent load reduction was actually quite similar for each group. As a percentage of loads, the demand reductions varied less, ranging from 21.0% to 23.4%, suggesting that most of the differences by event window are a function of the underlying amount of air conditioner load.

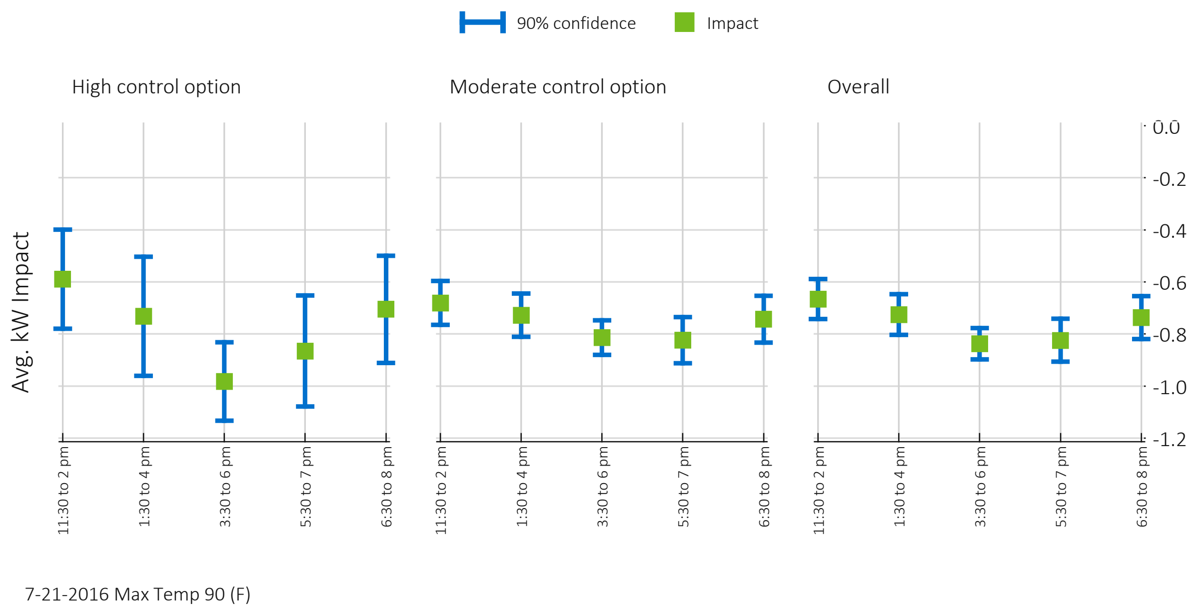
Figure 4-3: Load Profiles for July 21 Dispatch Window Test



The point estimates for the load impacts, along with the 90% confidence intervals, for each test group is presented in Figure 4-4. The results are broken down by program option (moderate versus high load control), as well as for program participants in general. Note that the width of the confidence intervals are largely driven by the sample sizes, and thus the confidence intervals for the higher load control option customers are much wider because only 15% of customers sign up for it and, as a result, treatment and control group sample sizes were smaller.

In all cases, the load impacts show the same pattern with average load reduction increasing for later dispatch windows until the last event. However, the difference in impacts is not great enough to rule out the possibility that it could be explained by estimation error, as indicated by the overlapping confidence intervals for the various dispatch windows.

Figure 4-4: Point Estimates and Confidence Intervals for July 21 Cascading Events

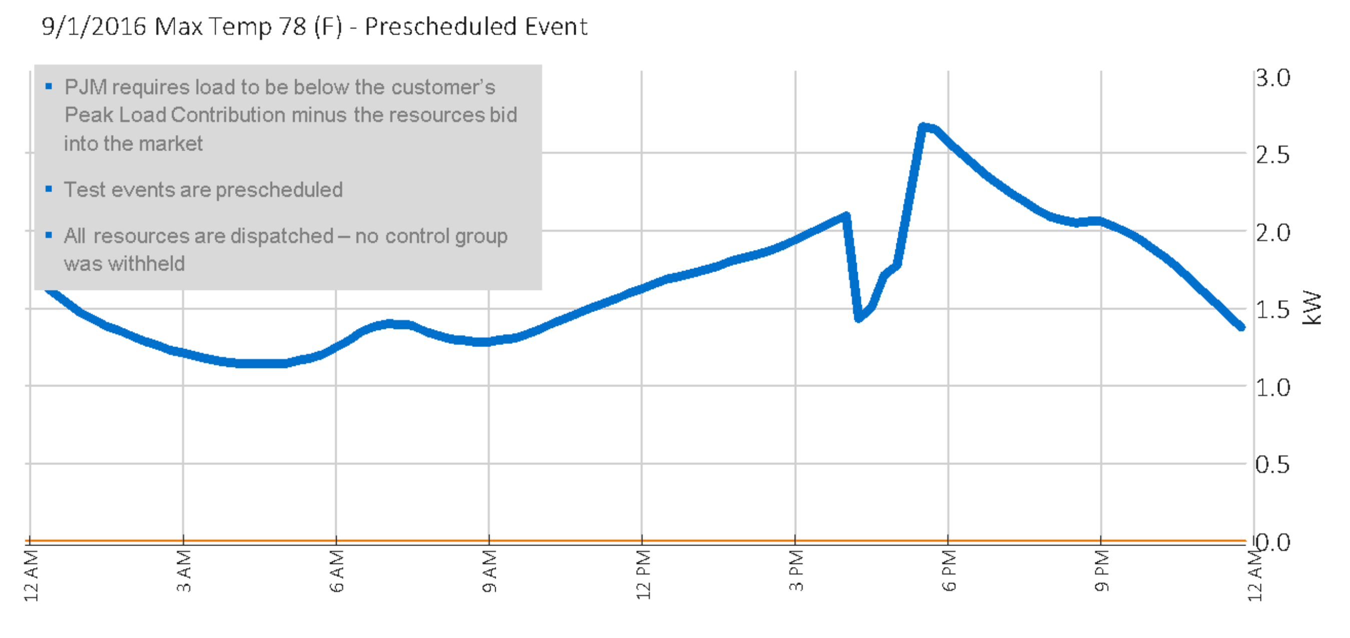


4.4 PJM System Test

In addition to the general population events, the cascading event, and the emergency load shed event, DEO dispatched all resources on September 1 for a pre-scheduled PJM test event. Because no customers were withheld for a control group, it was not possible to calculate impacts for this event using the randomized control trial method. However, the load profile for the event in Figure 4-5 includes a clear load drop. Because the test occurred on a fairly cool day (maximum temperature of only 78°F), there was relatively little load to drop. As such, the load shed appears to be of much smaller magnitude than those observed on hotter days.

For PJM events, the objective is to keep loads below a specific threshold. The threshold is based on the aggregate peak load contribution of participants under 1-in-2 weather year planning conditions minus the amount of resources bid into PJM. When events are called on cooler days, less reduction is needed to maintain loads below the threshold. For DEO Power Manager participants, the peak load contribution value is 3.19 kW per household—loads were kept substantially below the target threshold during the PJM test event.

Figure 4-5: PJM System Event



4.5 Weather Sensitivity of AC Load and Demand Reductions

The load reduction capacity of Power Manager is dependent on weather conditions, as shown in Figure 4-6. The plot shows the estimated average customer impact for each event as a function of daily maximum temperature. There is a clear correlation between higher temperatures and greater load reduction capacity, with the greatest load reductions occurring on the hottest day. Both emergency and normal operation impacts are displayed on this plot for that day, with the greater magnitude impacts attributable to the emergency operations customers.

While the weather correlation is clear, the question remains: How much of the bigger reduction capacity is due to larger air conditioners loads versus larger demand reductions? Both percent reduction and air conditioner loads grow with hotter temperatures. The whole house reductions were 18.9% on the coolest event day (87°F) and 26.1% on the hottest day (93°F). Figure 4-7 shows the weather sensitivity of whole house load for the average customer in Power Manager. All nonevent weekdays with a daily high above 70°F were classified into two degree temperature bins. The plot shows how the loads vary by hour as temperatures grow hotter.

The key finding is simple. Demand reductions grow larger in magnitude when temperatures are hotter and resources are needed most. Because peak loads are driven by central air conditioner use, the magnitude of air conditioner loads available for curtailment grows in parallel with the need for resources. Not only are air conditioner loads higher, but the program performs at its best when it is hotter.

Figure 4-6: Weather Sensitivity of Load Reduction based on Randomized Control Trial Analysis

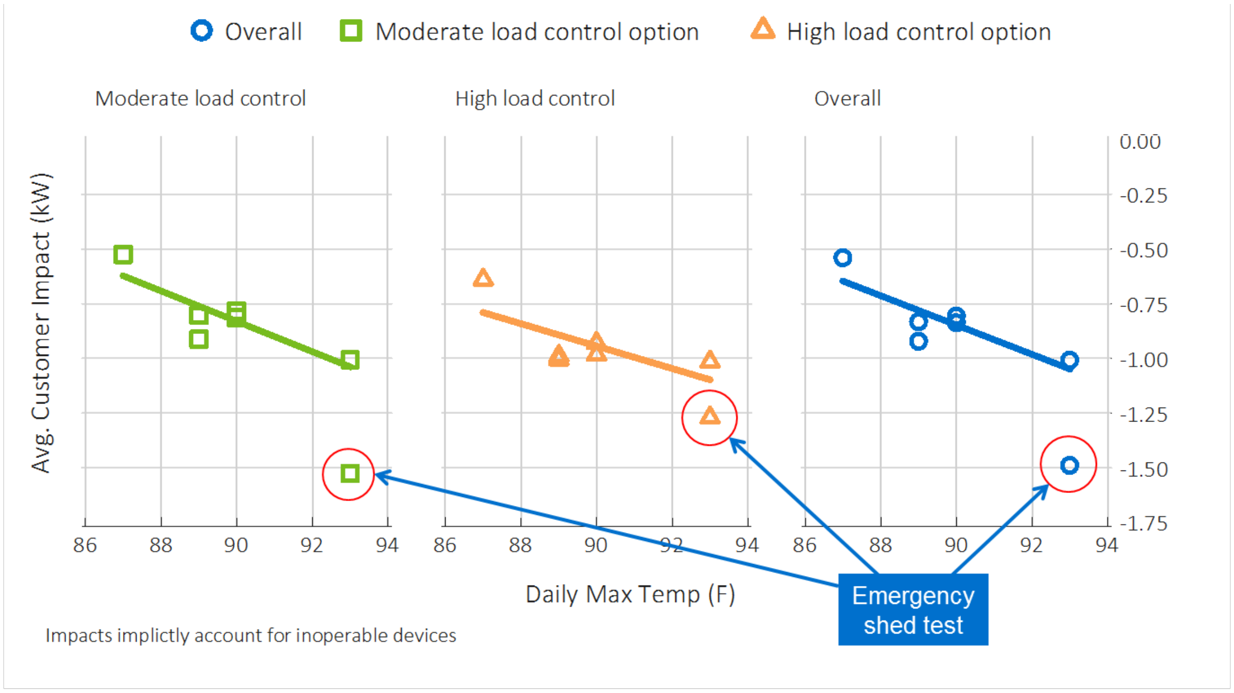
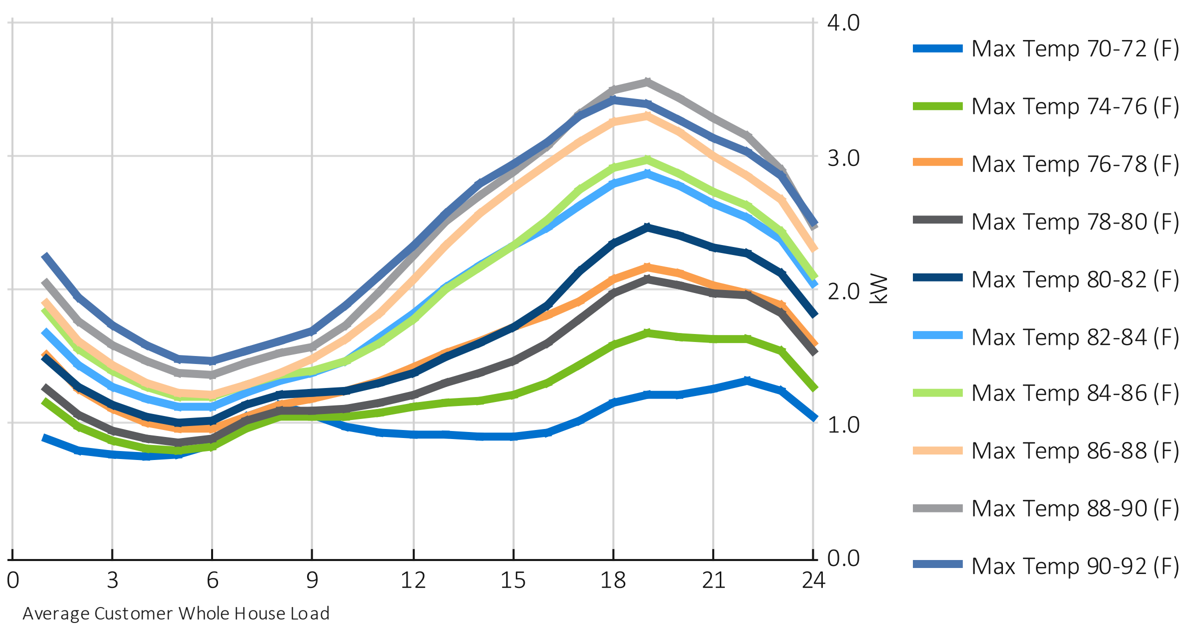


Figure 4-7: Weather Sensitivity of Average Customer Loads



4.6 Impacts by Customer Load Control Option

Figure 4-8 compares the load impact estimates for customers enrolled in the moderate versus high load control option, along with the 90% confidence intervals for each event. In general, point estimates for load reduction are greater for high load control option customers on any given event day. However, because there were relatively fewer customers in the high load control option subgroup, the confidence intervals for these point estimates are quite wide. While the differences are statistically significant for a few of the days, for many days the difference in load impact estimates falls within the range of uncertainty.

The difference in impacts between customers who signed up for the lower and higher load control options was minimal. Because customers self-select into the option, a key question is whether customers who use less air conditioning tend to sign up for the higher load control options. Figure 4-8 compares the control group loads and reductions for customers on the moderate and high load control options for the hottest day in 2016, July 25. Customers who signed up for lower load control option do have larger loads, but the difference is small and does not explain why the higher load control option did not deliver larger load reductions. The remaining explanation of the lack of the difference is the implementation of the cycling algorithm.

Starting in 2017, DEO will begin using a new cycling algorithm but retain the option for customers to sign up for the lower and higher load control options. A key question is whether the transition to the new algorithm will lead to clear differences in the magnitude of demand reductions between customers who elect the lower and higher load control options.

Figure 4-8: Comparison of Load Impact Results by Control Option for all Events

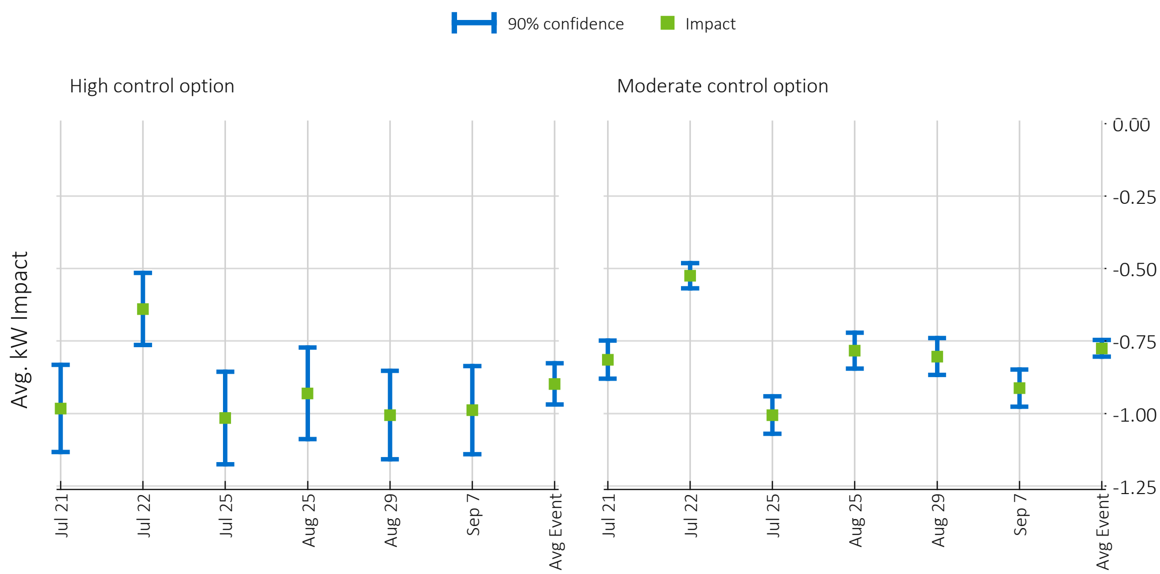
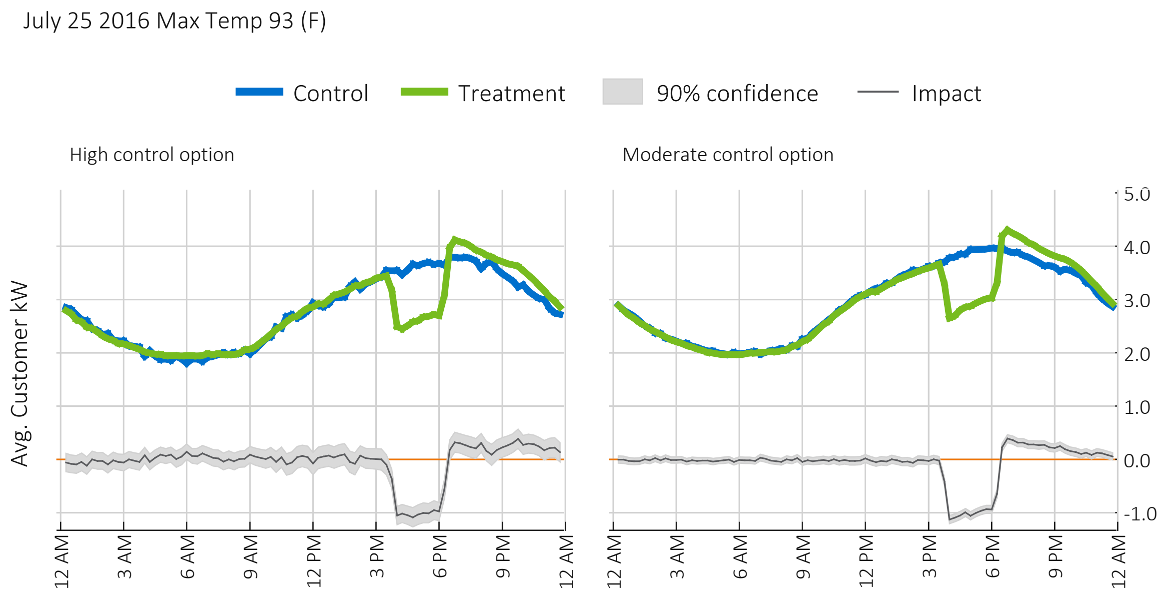


Figure 4-9: Comparison of Hourly Loads by Control Option on Peak Day



4.7 Impacts by Customer Size

As noted earlier, air conditioner use by Power Manager participants varies substantially, reflecting different occupancy schedules, home sizes, comfort preferences, and thermostat use and settings. Table 4-2 shows the Power Manager demand impacts for customers of different sizes. For the comparison, customers were classified into 10 equally sized groups, known as deciles, based on their 4 to 6pm electricity use during hot nonevent days.

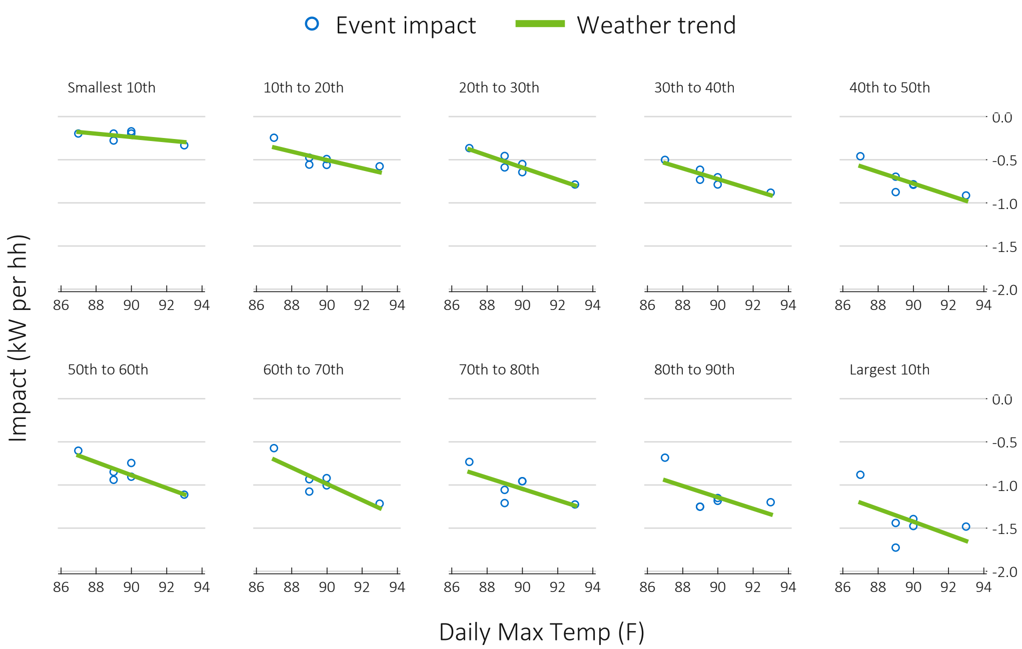
Customers with larger loads delivered larger demand impacts. Customers among the smallest 10th on average reduced demand by 0.23 kW per household; while customers among the largest group on average reduced demand by 1.40 kW. While the pattern of larger impacts among larger customers is clear, the reductions as a percent of whole house loads are very similar for nearly all groups, except the smallest ones. Within each size group, customers delivered larger demand impacts with hotter weather, as summarized in Figure 4-10.

Table 4-2: Average Event Impacts by Customer Size

Size Group	Load without DR	Impact	Std. error	90% Confidence Interval		% Impact	90% Confidence interval	
				Lower bound	Upper bound		Lower Bound	Upper Bound
Smallest 10th	1.44	-0.23	0.03	-0.18	-0.28	-16.0%	-12.7%	-19.4%
10th to 20th	2.26	-0.49	0.02	-0.45	-0.53	-21.5%	-19.7%	-23.3%
20th to 30th	2.61	-0.57	0.02	-0.53	-0.61	-21.8%	-20.2%	-23.3%
30th to 40th	2.93	-0.71	0.03	-0.66	-0.75	-24.1%	-22.6%	-25.5%
40th to 50th	3.15	-0.76	0.03	-0.71	-0.80	-24.0%	-22.7%	-25.3%
50th to 60th	3.46	-0.86	0.03	-0.82	-0.91	-24.9%	-23.6%	-26.2%
60th to 70th	3.78	-0.96	0.03	-0.91	-1.00	-25.3%	-24.0%	-26.5%
70th to 80th	4.13	-1.02	0.03	-0.97	-1.07	-24.8%	-23.6%	-26.0%
80th to 90th	4.66	-1.12	0.03	-1.07	-1.17	-24.1%	-23.0%	-25.2%
Largest 10th	6.00	-1.40	0.05	-1.32	-1.48	-23.4%	-22.1%	-24.7%

*Impact are only for general population events

Figure 4-10: Customers within Each Size Group Deliver Larger Impacts with Hotter Weather



4.8 Key Findings

A few key findings are worth highlighting:

- Demand reductions were 0.79 kW per household for the average general population event.
- Peak day impacts under normal operations averaged 1.01 kW per household over the course of the two hour dispatch window on July 25, 2016, when the daily maximum temperature was 93˚F.

- Emergency operations produced larger impacts than normal operations, 1.49 kW vs. 1.05 kW per household for the same hour on the hottest day in 2016. Reductions from emergency operations exceeded those from normal operations by 41.9%.
- The magnitude of impacts varied slightly by dispatch window in absolute terms, but not so much as a percentage of available load. Demand reductions ranged from 0.67 to 0.84 kW per household on July 21, with larger impacts generally occurring later in the day. As a percentage of loads, the demand reductions varied less, ranging from 21.0% to 23.4%, suggesting that most of the differences by event window are a function of the underlying amount of air conditioner load.
- Demand reductions grow larger in magnitude when temperatures are hotter and resources are needed most.
- The difference in impacts between customers who signed up for the lower and higher load control options was within the range of uncertainty.

5 Whole Building Versus End Use Impacts

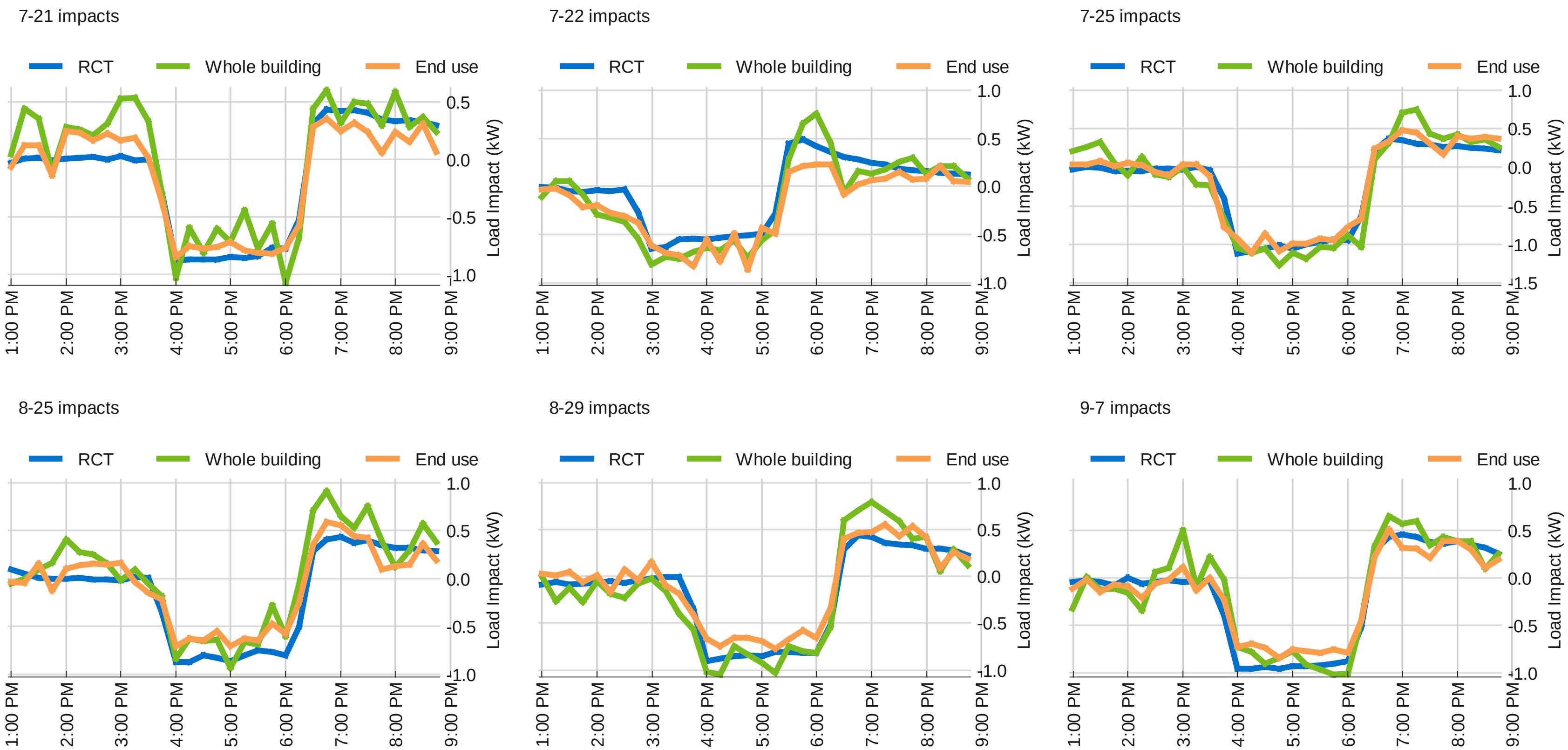
Along with randomized control trial analysis, the Power Manager program in DEO's territory was evaluated using within-subjects regression of load data collected from a sample of program participants. This analysis was applied to end use data collected from customers' AC units, as well as to the whole building Smart Meter data for the same group of customers. The same regression model was applied to both sets of data to ensure consistency in the analysis and allow for a valid comparison between the two sets of results.

The purpose for this was to compare whether whole building impacts would predict similar impacts to those derived from the end use data. Any deviation between the two would imply that customers were offsetting load reductions through other end uses. However, the study found that both evaluation methods produced similar impact estimates.

5.1 Comparison of Load Impacts by Method

Hourly load impact estimates for general population events derived from regression analysis of end use and whole building data for the same group of customers, as well as results derived from randomized control trial analysis, are compared in Figure 5-1. Hourly load impacts are similar for all analysis methods presented, though the two sets of results produced by regression analysis exhibit considerably more noise, likely due to the relatively small sample size and the inherent uncertainty associated with modeling. However, the overall magnitude of the load impacts was essentially the same regardless of data source and analysis method.

Figure 5-1: Comparison of Load Profiles by Analysis Method

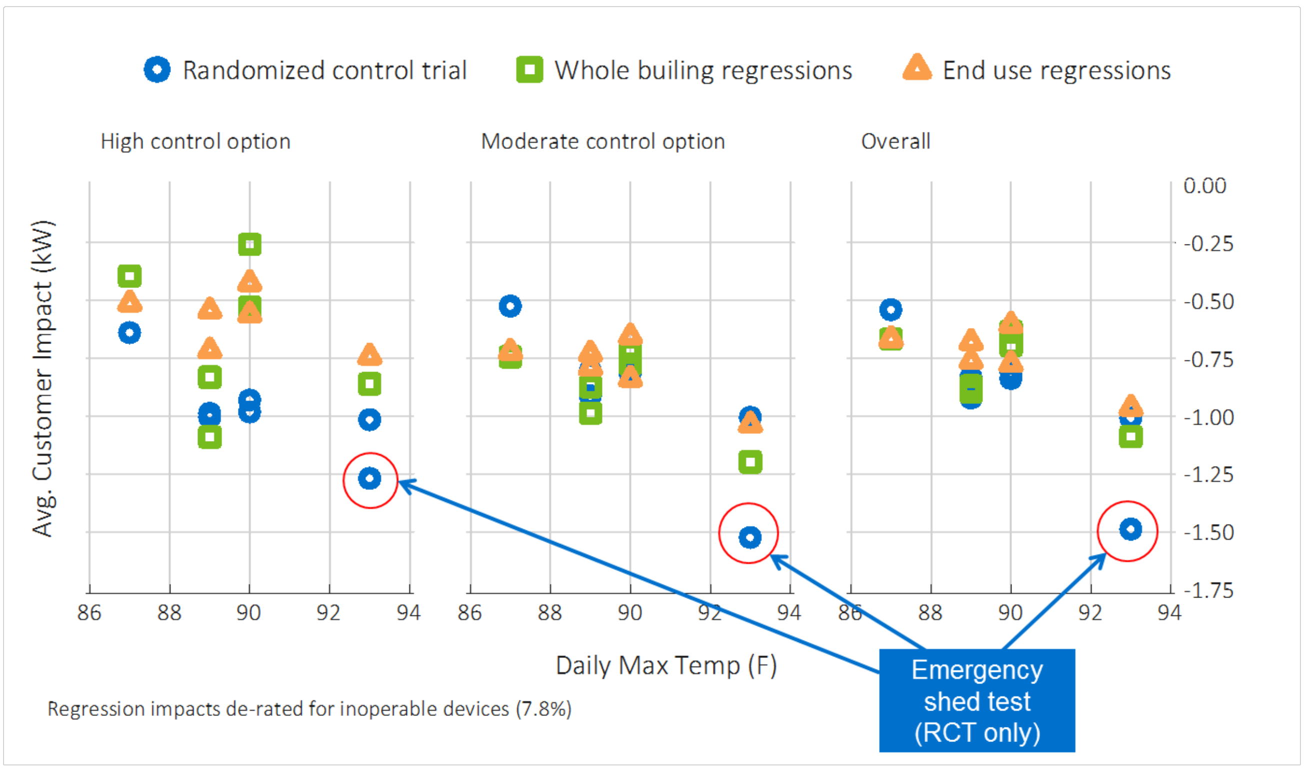


Whole building and end use regression impacts de-rated for inoperable devices

The average load impacts during the event windows are plotted against maximum daily temperature in Figure 5-2. The results are broken down by control option (as well as overall average impacts across both groups), and are differentiated for each analysis method. The estimated load impacts are similar to one another at each temperature level, though the randomized control trial results also include the impacts for the emergency shed event, which was not evaluated by regression analysis due to a lack of customers in the emergency dispatch group among the customers in the end use sample.

The same load impacts are estimated by three different analysis methods, providing a high degree of confidence in the results. Furthermore, the fact that similar load impacts are predicted by whole building and end use data indicates that Power Manager customers are not offsetting AC load reductions by increasing usage of other end uses. Furthermore, the weather sensitivity implications produced by the randomized control trial results are confirmed by the end use and whole building regression analysis.

Figure 5-2: Load Impacts vs. Temperature for Each Analysis Method



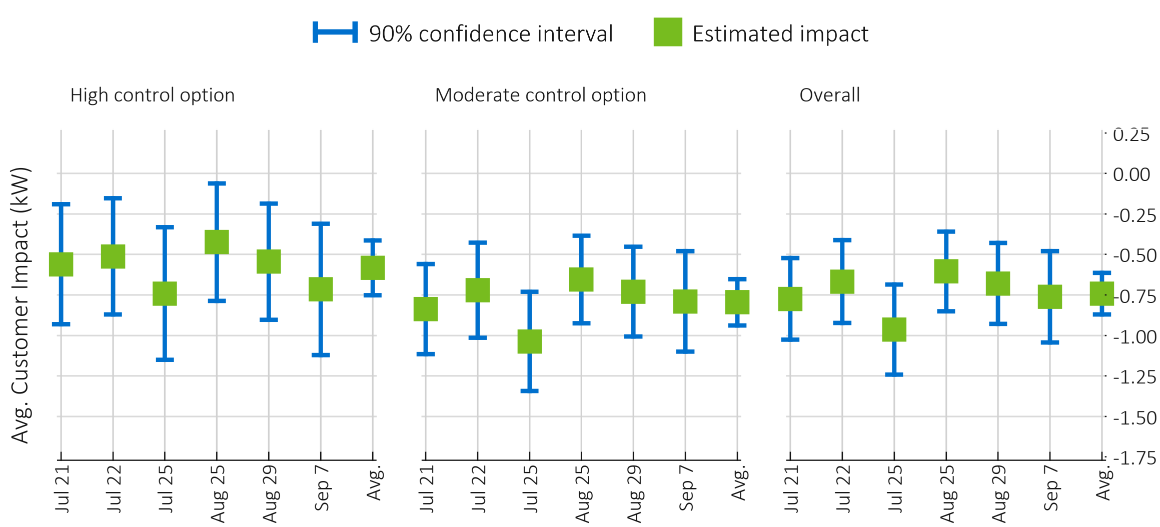
5.2 Air Conditioner End Use Metered Customer Load Impacts

Details of the ex post results derived from regression analysis of end use data are provided in Table 5-1. As with the randomized control trial results, these load impacts are the average values for load reduction during the event windows indicated, minus the initial 30 minute phase-in period. Overall, load impacts are similar to the results predicted by randomized control trial analysis and display the same weather sensitivity presented previously. Average impact over all of the general population events is 0.74 kW/customer, as compared to the randomized control trial average of 0.79 kW/customer. As noted earlier, the confidence intervals are somewhat larger for this analysis compared to the randomized control trial analysis due to the substantially smaller sample size involved.

Table 5-1: End Use Load Impacts Based on Regression Analysis

Event Date	Start Time	End Time	Load without DR	Impact	Std. Error	90% Confidence Interval		% Impact	90% Confidence interval		Daily Max	Avg. Daily Temp
						Lower Bound	Upper Bound		Lower Bound	Upper Bound		
7/21/2016	3:30 PM	6:00 PM	1.83	-0.77	0.15	-0.52	-1.03	-42.4%	-28.6%	-56.2%	90	80.3
7/22/2016	2:30 PM	5:00 PM	1.43	-0.67	0.16	-0.41	-0.92	-46.7%	-28.9%	-64.6%	87	79.9
7/25/2016	3:30 PM	6:00 PM	2.07	-0.96	0.17	-0.69	-1.24	-46.6%	-33.2%	-60.0%	93	83.0
8/25/2016	3:30 PM	6:00 PM	1.59	-0.60	0.15	-0.36	-0.85	-38.0%	-22.5%	-53.5%	90	81.7
8/29/2016	3:30 PM	6:00 PM	1.54	-0.68	0.15	-0.43	-0.93	-44.2%	-27.9%	-60.4%	89	78.8
9/7/2016	3:30 PM	6:00 PM	1.63	-0.76	0.17	-0.48	-1.04	-46.9%	-29.5%	-64.2%	89	78.9
Average General Population Event			**1.68**	**-0.74**	**0.08**	**-0.61**	**-0.87**	**-44.2%**	**-36.5%**	**-51.8%**	**90**	**80.4**

Figure 5-3: Load Impacts and Confidence Intervals for End Use by Option



Sample consists of 61 customers
48 on the moderate option, 12 on the high control option, and 1 on the low control option

These results are also presented in Figure 5-3, along with a breakdown by control option. Comparing the results for customers enrolled in the moderate and high load control option, the results are close enough to one another for each event day that they can be explained by estimation error using this analysis method. Any differences that exist between the two control options were not picked up due to the small sample size.

5.3 Whole Building Impacts for Customers with Metered Air Conditioners

Similar to the end use derived results, the details of the load impacts estimated by regression modeling of whole building data is presented in Table 5-2. The results are similar enough to the end use results that the difference could be attributed to estimation error. The whole building analysis estimates an average impact of 0.81 kW/customer across the events, versus 0.74 kW/customer as predicted by the end use data. This indicates that customers are not offsetting AC load reductions through other end uses.

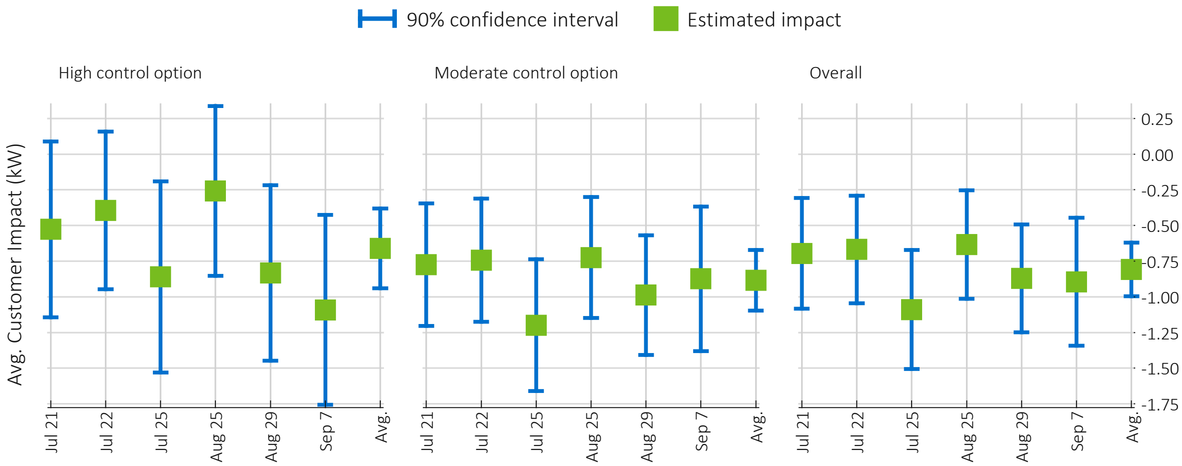
Table 5-2: Whole Building Load Impacts Based on Regression Analysis

Event Date	Start Time	End Time	Load without DR	Impact	Std. Error	90% Confidence Interval		% Impact	90% Confidence interval		Daily Max	Avg. Daily Temp
						Lower Bound	Upper Bound		Lower Bound	Upper Bound		
7/21/2016	3:30 PM	6:00 PM	3.68	-0.70	0.24	-0.31	-1.08	-18.9%	-8.4%	-29.5%	90	80.3
7/22/2016	2:30 PM	5:00 PM	3.18	-0.67	0.23	-0.29	-1.04	-21.0%	-9.1%	-32.8%	87	79.9
7/25/2016	3:30 PM	6:00 PM	4.04	-1.09	0.25	-0.67	-1.51	-26.9%	-16.6%	-37.3%	93	83.0
8/25/2016	3:30 PM	6:00 PM	3.24	-0.63	0.23	-0.25	-1.01	-19.5%	-7.8%	-31.3%	90	81.7
8/29/2016	3:30 PM	6:00 PM	3.34	-0.87	0.23	-0.49	-1.25	-26.1%	-14.8%	-37.4%	89	78.8
9/7/2016	3:30 PM	6:00 PM	3.42	-0.89	0.27	-0.45	-1.34	-26.2%	-13.0%	-39.3%	89	78.9
Average General Population Event			3.48	-0.81	0.11	-0.62	-1.00	-23.2%	-17.8%	-28.6%	90	80.4

The point estimates and confidence intervals for the load impacts are presented in Figure 5-4, with the results also broken down by control option. Compared to the end use results, there is considerably more uncertainty for the estimates derived from whole building data even though the samples are identical. This is due to the additional noise introduced by the additional end uses that are measured by whole building data. However, all impact estimates are statistically significant.

Comparing the results for moderate and high load control customers, there is simply too much uncertainty around the point estimates to draw any conclusions on how the control options compare using this analysis technique.

Figure 5-4: Load Impacts and Confidence Intervals for Whole Building by Option



Households are the same as the same as those with AC end use measurement

5.4 Key Findings

A few key findings are worth highlighting:

- There is no evidence that customers compensate for air conditioner curtailments by increasing other end uses—whole building impacts are no different than end use impacts.
- Findings were consistent across analysis methods, providing a high degree of confidence in the results.
- Regression analyses produced similar results to the randomized control trial analysis, but were much less precise.

6 Device Operability and Site Level Performance

A significant problem in load control programs is nonperforming devices or sites. These can be due to broken or disconnected control devices or because some devices fail to receive control event paging signals. They also can occur because of broken air conditioner units or because some customers do not use their air conditioners during event hours. Due to the significant cost of direct verification of device operation, it is not financially feasible to blindly send service technicians to every property to check device operation. Up until recently, with no way to identify broken devices, it has just been easier, and more cost effective, to recruit new customers. If DEO is able to remotely identify sites that underperform due to broken or missing devices, or because of paging network communication failures, it could increase the aggregate impacts of the program without as much cost as new customer acquisition.

Using 15 minute interval data from DEO's air conditioning cycling load control program, Power Manager, Nexant undertook the task of creating methods to identify probable broken or missing devices. Our effort involved two main steps:

- A field study designed to physically test whether load control devices were functional. The main purpose of this study component was to quantify the share of inoperable devices. This estimate, however, does not factor in paging network communication failures or sites that do not have their air conditioner on during event hours. As we discuss later, the incidence rate is one of the critical components that affects the precision of efforts to identify broken or missing devices.
- Use of data analytics to develop methods that identify sites that underperform or do not deliver demand reductions. A device that is not functional does not reduce air conditioner demand over multiple events.

The field study was implemented in tandem with the installation of air conditioner data loggers and serves to quantify the device failure base rate. While data analytics was used to identify underperforming sites, a separate verification test to determine the precision of the diagnosis has not yet been implemented. Our expectation is that using whole building smart meter data to identify nonperforming or missing devices will lead to substantial improvements over blindly sending technicians to assess if devices are performing. These efforts, however, are most precise if they are restricted to households that clearly use air conditioners during hotter weather conditions. These customers also provide the most value since they use air conditioners during peaking conditions. Diagnosis of nonperforming devices is less accurate when it is applied to sites with low or no air conditioner use during peak hours of hotter days.

6.1 Device Operability Field Test

As part of the study, Nexant was responsible for all fieldwork related to recruiting customers for end use data collection and installation and collection of data loggers. The customers were recruited from a random sample of the Power Manager participant population. Prior to installing data loggers on air conditioners, Nexant tested whether load control devices were functional. The inspection consisted of:

- Onsite spot measurements of the kW, voltage, amperage, and power factor;
- Information about the AC unit;
- Inspection of the load control device for presence, proper installation, physical condition, and operability; and

- Inspection of the load control device connection wires, including presence, physical condition, and whether the connection was secure.

End use data loggers were only installed on air conditioner units with functional load control devices. Based on field tests, 95 out of the 103 (92.2%) devices are operable, with a 90% confidence interval of ±4.34%. The estimate represents a lower bound for performance improvement because it does not account for devices that do not perform due to paging network issues or because the air conditioner is not in use during afternoon peak hours on hotter days.

Table 6-1: Device Operability Field Study Results

Metric	Value
Devices inspected	103
Inoperable devices	8
Operable devices (i.e., loggers installed)	95
Device failure rate	7.8%

6.2 Use of Smart Meter Data to Identify Underperforming Sites

DEO smart meters collect residential whole building data for each 15 minute interval. To identify underperforming sites, it is necessary to assess whether air conditioner units are on when load control events are called and whether or not the devices lead to reduction in the air conditioner load control demand. There are two related challenges for doing so with whole building data: air conditioner use varies substantially across households and the footprint of air conditioner use is often not clearly identifiable with hourly data on individual days. Before detailing the method used to identify nonperforming sites and the results, it is useful to understand some of the fundamental challenges in diagnosis.

6.2.1 Fundamentals of Diagnosis

The accuracy of any diagnosis depends on the answer to three questions:

- Are failures common? Technically, this is the incidence rate or base rate of failures. Based on the field inspection it is approximately 7.8%.
- How well does the test identify failures when there is indeed a failure? This is technically referred to as *sensitivity*.
- How often does the test incorrectly diagnose a failure when none occurred? Technically, this is the *false positive rate*. The inverse of the false positive rate is known as *specificity*.

In describing diagnostic tests, it is common to focus on how well the test identifies failures when there are indeed failures—*sensitivity*. It is also possible, however, for the diagnosis to misclassify devices that are, in fact, operable as failing.[7] The right question is: what fraction of all devices classified as failing is indeed failing? The answer to this question is known as *precision*.[8] Diagnosis is inherently difficult when failures are uncommon. When that occurs, a well-designed diagnostic test with high sensitivity and high specificity may still over diagnose.

One way to improve the precision of diagnosis is to apply the test only to populations where the diagnostic test is known to perform well or populations that are expected to have higher failure rates (e.g., older devices). The risk of misdiagnosis is highest among customers who rarely use their air conditioner during peaking conditions and who should not be targeted for reactivation in the first place. Customers who use their air conditioners during peaking conditions should be less prone to misdiagnosis. These customers are also more valuable and cost effective to reactivate. Older devices are also expected to have higher failure rates and, as a result, a lower rate of misdiagnosis.

The main takeaway is that using whole building interval data to identify underperforming devices can be very successful under the right settings. The method should not be applied blindly to all sites but should ideally focus on customers with higher air conditioner use and/or sites with older control devices.

6.2.2 Classification Algorithm

Devices that are functional and receiving the paging signal reduce demand or notch the load shape during the load control events. Devices that are nonperforming do not alter the load shape. There were four main components to the algorithm:

- Use of whole building load shape over multiple events where the event start and end times are standardized;
- Load drop when the event begins—air conditioner units are phased in over time;
- Snapback immediately after control of the air conditioner is released; and
- A high correlation between temperature and loads during hotter nonevent days.

[7] For example, assume 10,000 devices, out of which 500 are failing (5%). A test that correctly identifies 95% of failing devices (475) is highly sensitive. However, if that same test misclassifies 5% of the 9,500 devices that are operable (475), it is not very precise. In total, 950 devices will have been identified as failing, but only 475 (50%) are correctly classified. Another 25 failing devices are missed entirely.

[8] Precision = True Positives / (True Positives + False Positives)

Figure 6-1: Process for Developing Algorithm to Detect Underperforming Sites

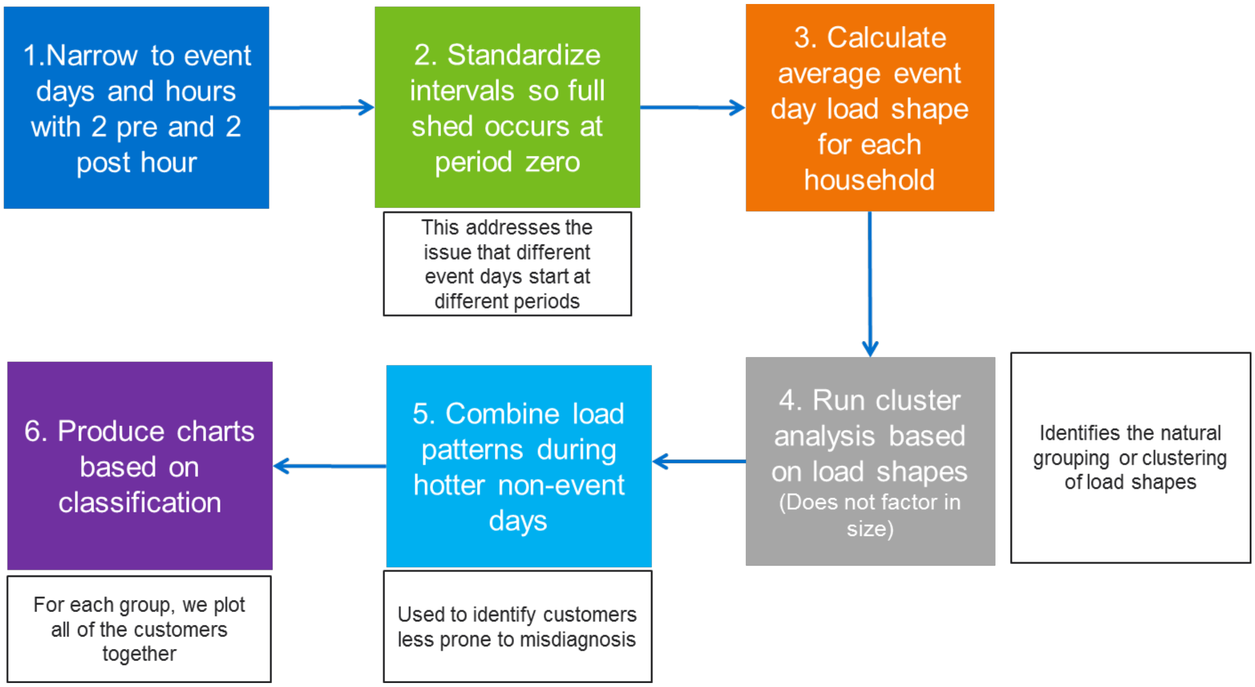


Figure 6-1 describes the process used to develop the algorithm. First, the data was narrowed to six hours for each general population event day—two hours immediately preceding the event, two hours of actual full reduction, and two post event hours. Next, the time intervals were standardized so full reduction capability occurred at time zero. This was necessary to account for slight differences in dispatch times.[9] Third, the average load shape by time interval was calculated for each household in Power Manager with smart meter data. We focused on the average event day load shape across multiple events to reduce the risk of misdiagnosis. A device that does not respond to multiple events is easier to identify than a device that does not respond to a single event. Over 99% of the 39,627 sites analyzed experienced 2 or more events and 73% of sites experienced 6 events.

The core of the algorithm is the fourth step—the use of cluster analysis on event day load shape data. Cluster analysis is an exploratory data analysis technique used to classify customers into natural groupings (or clusters) based on a specific set of observable characteristics. The customers within a cluster are similar to each other based on the observed variables and, at the same time, differences between the groups are maximized. To isolate load shapes from customer size, the loads for each time period were normalized as percentage consumption over the six hour period—that is, for each customer the normalized event day load shape added up to 100%. Each customer was then assigned to one of six load shapes through cluster analysis. Customers with distinct load drops during events and snapback after the

[9] For example, if an event started at 3:30pm, it attained full reduction by 4pm because load control devices are phased in randomly; thus 4pm was considered time zero. If the event instead started at 2:30, 3pm was considered time zero after accounting for the load control phase in.

event were grouped together; while customers who did not experience a clear load drop were grouped with similar customers.

The final step was to combine the load shape classification with the customer size data in order to avoid misdiagnosis. Customers who do not use their air conditioner much during peak hours (4 to 6pm) on hotter days are more prone to miscategorization. The lack of air conditioner use can be mixed with a lack of event response.

6.2.3 Results

Figure 6-2 illustrates the six prototypical load shapes produced by the cluster analysis. The shape for customers in groups 2, 3, and 4 suggests a distinct load drop. Customers in groups 1 and 5 have smaller but still distinct load drop shapes. The shape for group 6 suggests no load reduction took place for these customers during events despite the automation. This could be due to missing or failing devices, paging network gaps, or lack of air conditioner loads.

Figure 6-2: Prototypical Event Day Load Shapes (Cluster Analysis)

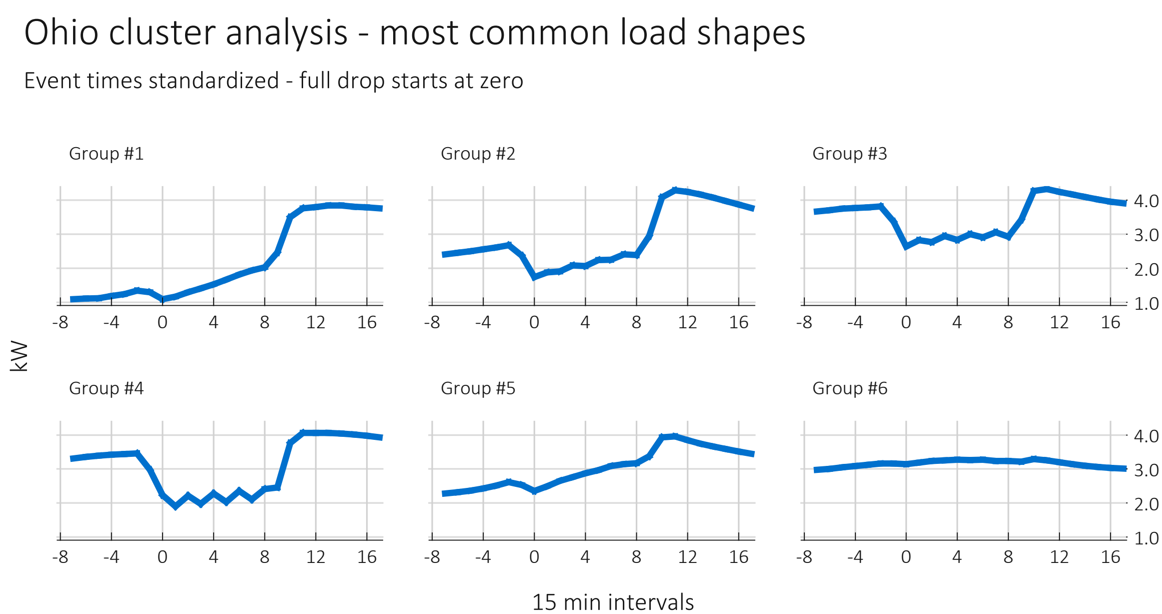
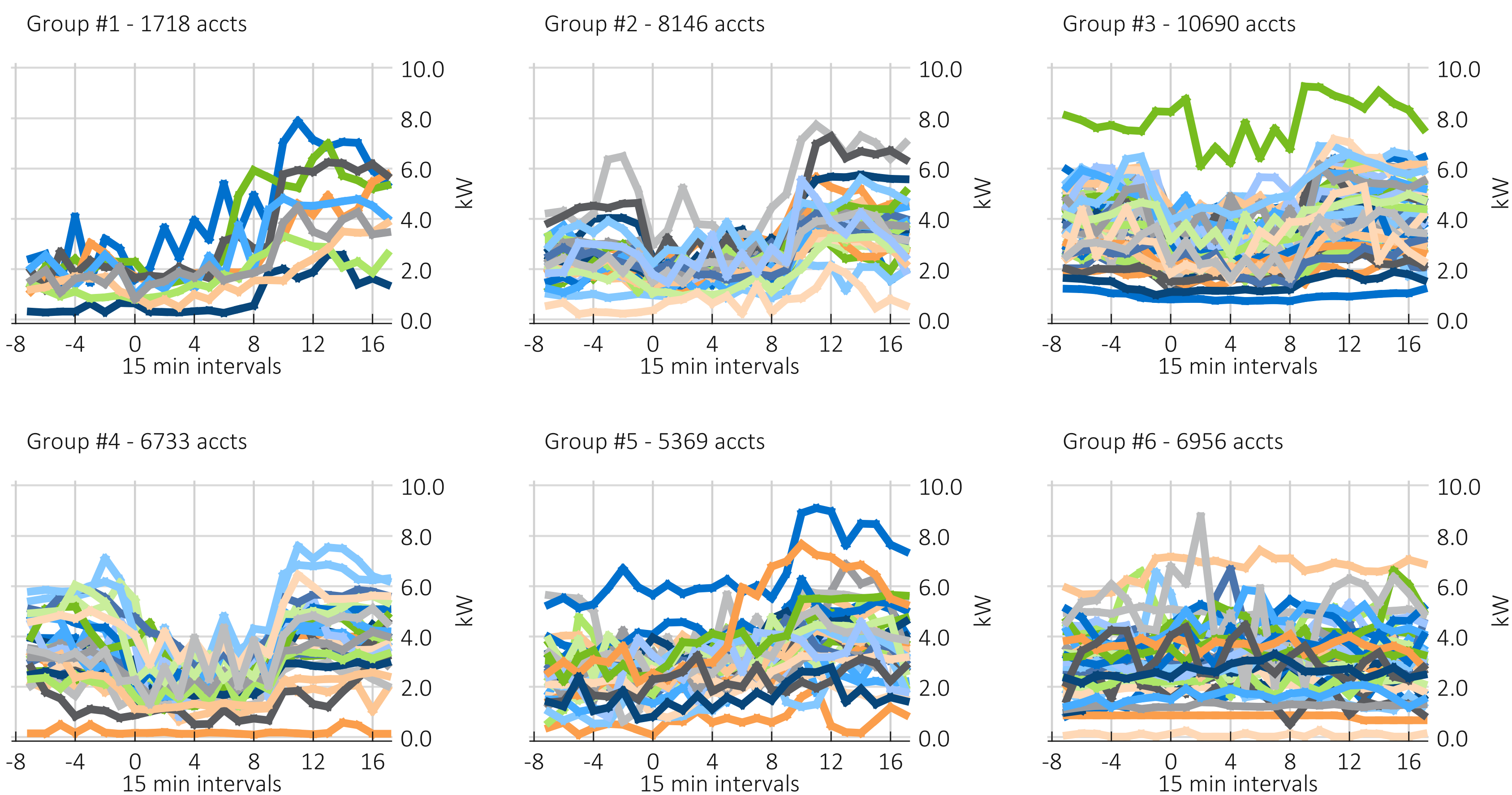


Figure 6-2 visualizes the categorization for a random sample of 200 sites. The customers in each group follow the prototypical shapes but sometimes differ in size due to the fact that the algorithm isolated shapes. In total, 6,956 of the 39,627 sites analyzed (17.6%) did not exhibit a demand reduction pattern and another 13.5% were assigned to group 5, which delivered smaller percent load reductions. Customers may exhibit little or no reduction pattern due to inoperable devices (7.8%), paging network communication failures, or because those customers rarely use air conditioning during event hours. It is important to separate performance from weather sensitivity and customer size. Smaller customers may be underperformers due to the lack of air conditioners and are less cost effective, even with a functional device. Thus, we recommend focusing direct verification efforts on larger customers.

Figure 6-3: Ohio Cluster Analysis – Individual Sites Classified Based on Event Day Load Shapes



Each line is a household during the average event day - Random sample of 200
Event times standardized - full drop starts at zero

6.3 Key Findings

Key findings from the investigation into device operability, include:

- End use data loggers were only installed on air conditioner units with functional load control devices.
- Based on field tests, 95 out of the 103 (92.2%)devices tested are operable, with a 90% confidence interval of ±4.34%.
- Most sites with inoperable devices have multiple failures.
- The event day load profiles suggest that 6,956 of the 39,627 sites analyzed, 17.6% did not exhibit a demand reduction pattern. This can be due to failing or missing devices, paging network issues, or lack of air conditioner loads.
- Efforts to inspect paging network strength and verify if devices are missing or failing should focus on larger customers. They are less prone to misdiagnosis and more cost effective.

7 Process Evaluation

Process evaluation, particularly when combined with the insight obtained from impact evaluation, informs efforts to continuously improve programs by identifying program strengths and weaknesses, opportunities to improve program operations, program adjustments likely to increase overall effectiveness, and sources of satisfaction or dissatisfaction among participating customers. The primary objectives for the process component of the evaluation include:

- Assessing the extent to which participants are aware of events, bill credits, and other key program features;
- Understanding the participant experience during events: comfort, occupancy, thermostat adjustments, and strategies employed to mitigate heat;
- Identifying motivations and potential barriers for participation, including expectations, sources of confusion or concern, intention to stay enrolled, and likelihood of recommending the program to others;
- Documenting the operations, recruitment, enrollment, outreach, notification, and curtailment activities associated with program delivery; and
- Identifying program strengths and potential areas for improvement.

7.1 Survey Disposition

Nexant developed a survey for customers participating in the Power Manager program that was deployed immediately following a Power Manager event. The survey was administered via phone and email to maximize response rates in the 24 hour window directly following a Power Manager event. The post-event survey addressed the following topics:

- Awareness of the specific event day;
- Any actions that increased household comfort during a Power Manager event. Do participants report changing AC settings, using other equipment (including window units, portable units, or ceiling fans) to mitigate heat buildup? Were participants home during the event? Are they usually home during that time period?
- Satisfaction with the Power Manager program and bill credits earned;
- Expectations and motivations for enrolling. What did participants expect to gain from enrollment? To what extent are they motivated to earn incentive payments versus altruistic motivations such as helping to address electricity shortfalls during periods of high peak demand and/or reducing the environmental effects of energy production?
- Do participants expect to remain enrolled in the program in future years?

In addition to the post-event survey, a nonevent survey was also deployed immediately following a hot, nonevent day. This nonevent day survey was nearly identical to the post-event survey to facilitate comparison with the results of the event day survey, with only references to specific event awareness removed. Both the event and nonevent surveys were administered to Power Manager participants, providing a treatment and control day structure to apply to the data collected. Since event awareness and thermal comfort are primary areas of inquiry for the survey, the control data (from the nonevent surveys) provides the opportunity to net out any propensity for thermal discomfort or belief that a Power

Manager event is occurring that would naturally happen on any hot day of the summer. In this way, it is possible to evaluate whether statistically significant differences in event awareness and reports of thermal discomfort exist between customers who actually experience a Power Manager event and customers who do not.

The survey was completed by 121 customers on an event day (the *treatment group*) and 92 customers on a hot nonevent day (the *control group*). The overall response rate was 6%. All surveys were conducted on the day of the event or the nonevent. The plan was to survey about 50% of people by phone and 50% by email, but on the event day we could not reach 50% by phone; we note that telephone response rates in Ohio for the event day survey were likely impacted by the Republican National Convention that was taking place that day in Cleveland. The distribution of phone calls and emails, with response rates, is shown in Table 7-1. All responses in this section summarizing survey results have been weighted to reflect the survey design for 50% of completions by phone and email each.

The temperature on the event day was a high of 91°F with a heat index of 95°F, which was somewhat higher than the temperature on the nonevent day, which was a high of 88°F with a heat index of 92°F. Table 7-1 outlines the treatment and control group survey dispositions.

Table 7-1: Survey Disposition

Total Responses	Group Size	Date	Temp	Phone/ Email Distribution	Response Rate
213 Responses	121 Event Day	Thursday, July 21	high 91°F (heat index 95° F)	36% Phone	7%
				64% Email	5%
	92 Control day	Thursday, July 14	high 88°F (heat index 92°F)	49% Phone	11%
				51% Email	5%

Nearly all survey respondents—95% of both groups—own their residence. More than half of households surveyed have two or fewer residents, but 26% of treatment and 14% of control households have four or more residents. There was no apparent systematic difference in the age of respondents between the treatment and control groups. The mean age of respondents is 58 years and the most commonly reported level of education was a bachelor's degree: 29% of respondents said that they graduated from college. Nearly as many (28%) have a graduate or professional degree and 20% graduated from high school but not college.

7.2 Program and Event Awareness

The customer surveys were designed with the key objective of evaluating participants' awareness of Power Manager events, but a few questions were also included to gauge participants' general awareness of the program and its key features. Every respondent who was contacted to complete the survey was a Power Manager participant at the time of the survey, and a strong majority of the respondents, 75%, reported that they are in fact familiar with the Power Manager program. However, participants are not as aware of a key feature of the program, the option to decline to participate in a Power Manager event on a specific day; only 52% of respondents reported that they are aware of that option. Respondents also reported on whether or not they had seen Power Manager event credits on their bill. Relatively few respondents affirmed that they have seen credits on their bill: 13% of respondents reported that they have seen a credit, while 57% reported that they had not, and the balance of respondents, 31%, reported that they did not know. It is possible that due to the timing of the survey, which was midseason, these customers did not receive credits in 2016 at the time of the survey. Duke Energy also screened the list of customers who said they did not receive bill credits to make sure errors were not made; no customers in fact did not receive a bill credit when they should have.

All three of these questions were asked of both the treatment group and the control group. That is, the questions were asked of a group of customers that had experienced a Power Manager event that day and a group of customers who had not. It would not be expected that there would be significant differences in these questions addressing program awareness between these groups. Indeed, the responses to these three questions do not significantly differ across treatment and control customers.

Both the opt-out feature and bill credits are designed to be program features that enhance customer satisfaction with the program; with only about half of respondents aware of the event opt-out option and less than 15% of respondents recalling receiving a bill credit, an opportunity exists to improve participants' awareness of these customer-friendly program features.

Every Power Manager participant that was randomly selected to receive the post-event survey, i.e., the treatment group, experienced an actual Power Manager event that day, Thursday, July 21. A total of 121 customers completed the post-event survey. Only 13% of the treatment group respondents reported that their homes were uncomfortable that day, while all of them experienced a load control event that afternoon. As a program with no pre-event notification, a decrease in thermal comfort in the home is the key factor for assessing event awareness. In Ohio, with only 13% of respondents stating that they were uncomfortable the day of the event, event awareness by that measure is quite low. However, it could also be that a number of those respondents would say that their home was uncomfortably hot at times on any hot day of the year, regardless of whether or not the Power Manager program had a load control event. To control for this possibility, another randomly selected group of Power Manager participants were also surveyed on a hot day when a Power Manager event did not occur, Thursday, July 14. A total of 7% of respondents reported that their home was uncomfortable on this nonevent day. While more respondents of the post-event survey stated that their home was uncomfortable that day than respondents of the nonevent survey (13% vs. 7%, respectively), the difference is not statistically significant. Put simply, the increase in reported thermal discomfort cannot be ascribed to the Power Manager event. It is small enough that it could reasonably have occurred by chance. The response frequencies are tabulated in Table 7-2.

Table 7-2: Was there any time today when the temperature in your home was uncomfortable? Response Frequencies Weighted by Mode, Nt = 121 and Nc = 92

Response	Treatment	Control
Yes	13%	7%
No	71%	81%
Don't know	16%	11%

Of those relatively few customers (17 post-event and 7 nonevent survey respondents) who reported that they were uncomfortable at some time during the day of the survey, all but one reported that they started feeling thermal discomfort between the hours of noon and 6pm.[10] Asked when the period of thermal discomfort in their home ended, there was a slight shift in responses towards later in the day, where all but two respondents reported that their home stopped feeling uncomfortable during the period of 5 to 11pm.[11]

These customers who reported thermal discomfort were also asked to rate their discomfort using a five-point scale, where 1 represents "not at all uncomfortable" and 5 represents "very uncomfortable." Frequencies of the responses are summarized in Figure 7-1, which shows an unexpected result: the distribution of responses tends toward the upper end of the discomfort spectrum for the control customers. It would be expected that treatment customers would be more uncomfortable both because they experienced a load control event and because the treatment day was hotter than the control day.[12] The statistical test for the difference in distribution is called the chi-squared test, and it shows that the difference between the two groups is significant at the 90% confidence level. However, it is important to be cautious about statistical significance in this case, since only seven control customers answered the question—one control respondent rated their discomfort at level one, and the other six rated their discomfort at three, four, or five.

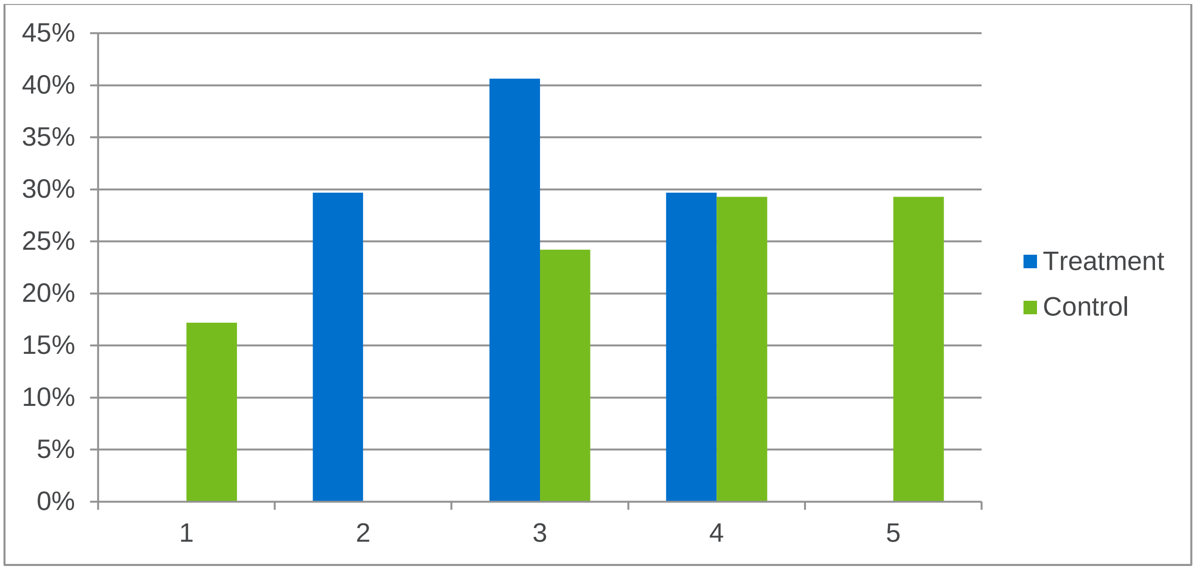
10 The one respondent reported feeling uncomfortable starting at 11am.

11 The two respondents reported feeling uncomfortable ending at 11am and 11pm.

12 The event day had a high of 91°F. The nonevent day had a high of 88°F.

Figure 7-1: Please rate your discomfort using a scale of one to five, where one means "not at all uncomfortable" and five means "very uncomfortable."

Response Frequencies Weighted by Mode, Nt = 17 and Nc = 7



Those respondents who reported that their homes had been uncomfortably hot that day were asked to state in their own words what they think caused the discomfort. The most commonly reported rationale is that the discomfort in their home was due to the weather being hot; 45% of these 24 respondents gave that reason. The next most common reason was a range of responses grouped into the category "other" because they have no defining characteristic. The responses range from "Power went out" to "I had a new installation this winter, they put in new siding, and the AC runs all day." The third most reported reason given for thermal discomfort was that the air conditioning unit was not on, where 17% of respondents gave that reason. Notably, ascribing thermal discomfort to Duke Energy controlling the air conditioner was only the fourth most common response: only 11% of respondents gave that reason. Table 7-3 summarizes the responses given to this survey question, across treatment and control customers and altogether. The totals may not add up to 100% because respondents could cite more than one reason. Also, none of the differences between treatment and control are statistically significant, which is not unexpected given the small sample size.

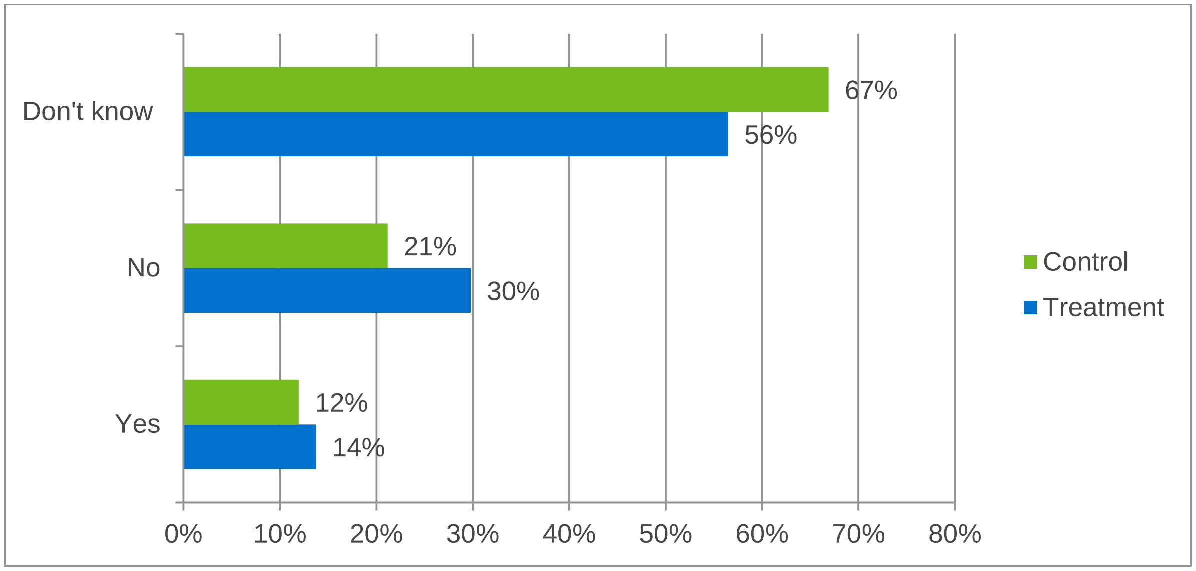
Table 7-3: What do you think caused the temperature to be uncomfortable?

Response Frequencies Weighted by Mode, N_t = 17 and N_c = 7

Reason	Treatment	Control	All
Air conditioner unit was not on	19%	12%	17%
Duke Energy was controlling air conditioner	11%	12%	11%
It was a very hot day	46%	41%	45%
Other	30%	17%	26%
Air conditioner doesn't work properly	0%	17%	5%

All survey respondents were also asked directly whether or not they thought a Power Manager event had been called in the past few days. The most common response was “don’t know,” where 56% of treatment customers and 67% of control customers stated that they didn’t know if there was a Power Manager event in the past few days. The prevalence of “don’t know” responses here is not surprising in light of the fact that Duke Energy does not actively notify participants of load control events. Figure 7-2 presents response frequencies for treatment and control respondents; the differences between treatment and control responses to this question were not statistically significant. Across all respondents together, 61% did not know if there was a Power Manager event recently, 13% thought that there was an event recently, and 26% did not think that there was an event recently.

Figure 7-2: Do you think a Power Manager event occurred in the past few days?
Response Frequencies Weighted by Mode, N_t = 121 and N_c = 92



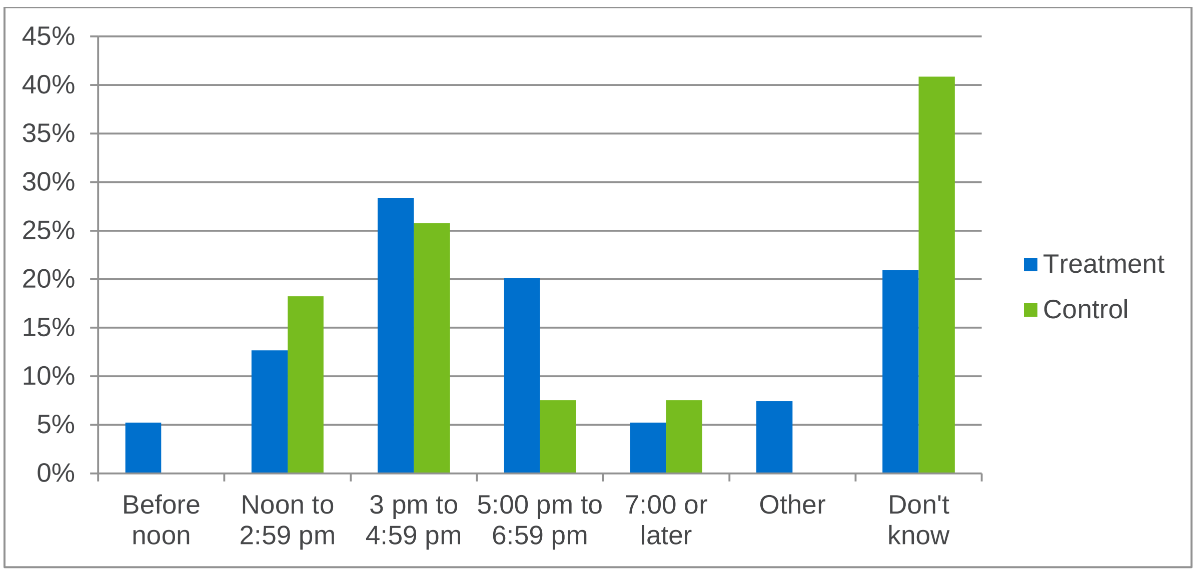
The relatively few respondents (17 treatment and 12 control) who thought there was a Power Manager event recently were asked a few questions about the event(s) that they perceived to have happened. First, when asked what day they thought the event occurred on, only 41% of the treatment customers correctly stated that there was an event that day; for comparison, 15% of control customers said there was an event day that day, and the difference between treatment and control customers identifying “today” as a Power Manager event day was not significant. Thus, we can’t conclude from this survey that actually experiencing a Power Manager event makes a customer any more likely to correctly identify when a Power Manager event takes place. These customers were also asked to describe how they determined that a Power Manager event was occurring, and the responses are summarized in Table 7-4. The most common response, given by 60% of respondents, is that they concluded an event was occurring because the temperature inside their home went up. The next most commonly reported rationale was because it was hot outside (17% of respondents giving this reason) and the third most common response was that they did not hear the air conditioning running the way they normally do, with 14% of respondents stating that reason. There were no statistically significant differences between the response patterns of treatment customers and control customers for this question.

Table 7-4: How did you determine that an event was occurring?
Response Frequencies Weighted by Mode, $N_t = 17$ and $N_c = 12$

Reason	Treatment	Control	All
It was a hot day outside - I knew from the temperature outside	18%	15%	17%
It got warmer inside - the inside temperature went up	59%	59%	59%
Did not hear the air conditioner running like I knew it should	18%	8%	14%
Some other way	7%	8%	7%
Don't know	10%	18%	14%

Respondents who thought there was a Power Manager event recently were also asked what time they thought the event occurred and whether or not they were home at that time. More than half of both the treatment group and control group customers said that they first noticed the event during the period of noon to 7pm; differences in the response pattern between the treatment and control groups are not statistically significant. One hundred percent of these respondents affirmed that they were home at the time they thought the event was occurring.

Figure 7-3: About what time did you first notice this event?
Response Frequencies Weighted by Mode, $N_t = 17$ and $N_c = 12$



7.3 Program Experience

Aside from occasional program communications to program participants, the primary way that Duke Energy customers experience the Power Manager program is during load control events. A large majority of survey respondents, 83%, stated that there is normally someone home between the hours of noon to 6pm on weekdays. Similarly large proportions of respondents also report that they are frequent users of their air conditioning systems. Table 7-5 shows the percentage of respondents who reported that they use their air conditioners every day for four different time period and day type combinations. Generally, between 80% and 90% of Power Manager survey respondents reported using their air conditioners every day, considering both weekdays and weekends, during both the afternoon and the evening. Statistically significant differences in response patterns were not observed here.

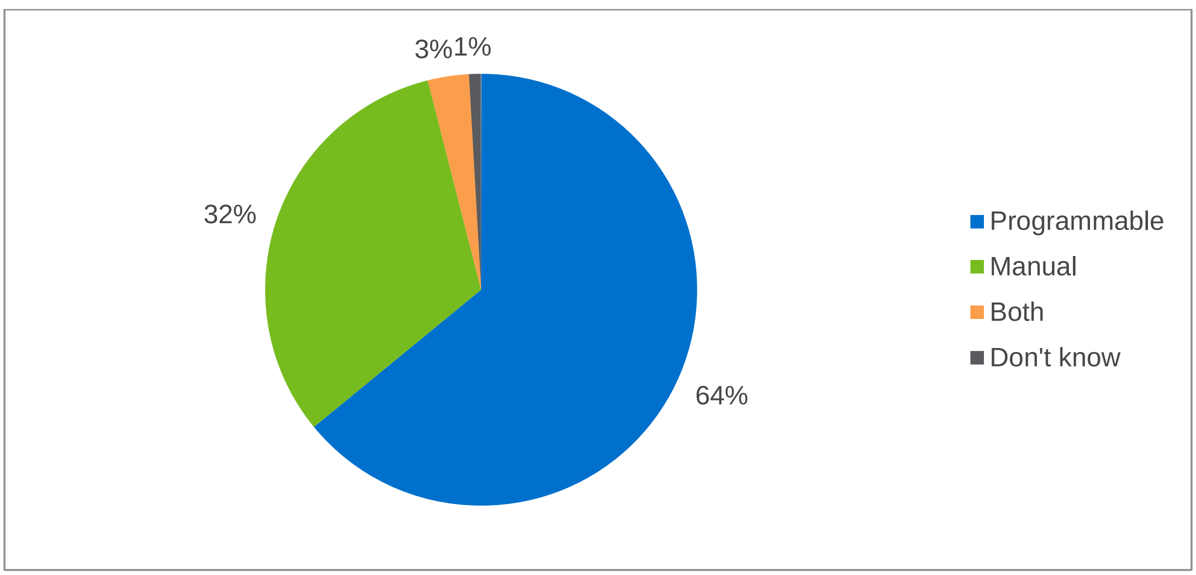
These survey responses confirm that Power Manager participants are in fact largely at home and using their air conditioners during the times that the program is likely to be launched when the need arises to use the program resource. As such, monitoring participant comfort levels is confirmed to be an important evaluation activity so that thermal comfort can be maintained at high enough levels to retain customer participation.

Table 7-5: How frequently do you or someone else in your household use your air conditioning system? Response Frequencies Weighted by Mode, N_t = 121 and N_c = 92

Day and Time	% of Treatment Responding "every day"	% of Control Responding "every day"
...weekday afternoons (12-6 PM)	86%	83%
...weekend afternoons (12-6 PM)	86%	90%
...weekday evenings (6 PM-12 AM)	85%	83%
...weekend evenings (6 PM-12 AM)	86%	86%

In addition to occupancy patterns and frequency of air conditioning usage, Power Manager participants' experience with the program is affected by how they operate their air conditioning systems. Beginning with the type of thermostat(s) installed in the home, survey responses show that there is a mix of both manual and programmable thermostats installed in the homes of Power Manager participants. Figure 6-4 summarizes the types of thermostat(s) that survey respondents reported. More than half, 64%, have a programmable thermostat, while 32% of respondents say that they have a manual thermostat.

Figure 7-4: What type of thermostat(s) do you have?
Response Frequencies Weighted by Mode, N_t = 121 and N_c = 92



Among the customers that have a programmable thermostat, 35% reported using the programmability feature to allow the thermostat to cool to different temperatures at different times, and a further 44% of customers set their thermostat at a constant temperature, representing 78% of respondents. Among customers without programmable thermostats, 53% say that they keep their thermostat set at a constant temperature. This relatively high incidence of using a thermostat setpoint should encourage thermal comfort associated with events. If during the course of an event, the home's internal temperature rises by one or two degrees, when the event is over, the thermostat will reliably detect the higher temperature and automatically cool the home to the desired temperature, without relying on the customer to feel uncomfortable first and manually turn the air conditioning on themselves. These reported air conditioning usage behaviors are supportive of the earlier finding that, on the whole, Power Manager participants are not aware of events when they occur.

In a similar vein, we asked customers who reported that they thought there was a Power Manager event recently whether or not they took any actions as a result of the perceived event. One only customer (of 17 who said that they thought there was a Power Manager event) said they did something different because of the event. This customer reported that they did stay home (they didn't leave because of the event) and that they used a "homemade" fan. Responses to these questions also provide consistent evidence that Power Manager events are not disruptive to participants and do not result in an increase in using other appliances for cooling that also use electricity.

7.4 Motivation and Potential Barriers for Program Participation

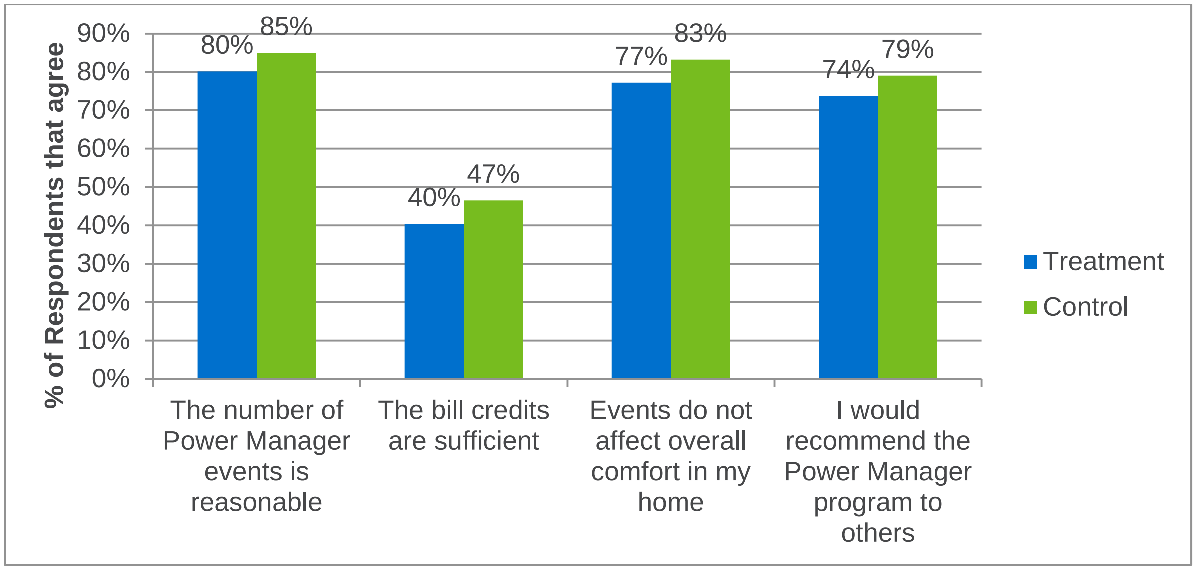
Respondents were provided with a list of possible reasons for enrolling and asked which reason was most important to them, and the survey responses reveal that Power Manager participants are motivated to be a part of the program for a diverse set of interests. The most frequently reported motivation are the bill credits; with 33% of respondents citing this as their most important motivator. The second-highest motivator is helping the environment—nearly as many respondents cited this reason as cited the first highest reason; 27% of respondents said helping the environment was the most important reason for enrolling. The remaining 40% of respondents were nearly equally split in among the remaining possible answers, including "don't know." Between 12% and 17% of respondents reported that they are participating in Power Manager because they want to "do their part for Duke Energy Ohio," or avoid electric service interruptions, or that they don't know what motivated them. Table 7-6 summarizes the survey responses. Differences in response patterns between treatment and control customers are not statistically significant.

Table 7-6: Which of the following reasons was most important to you when enrolling? Response Frequencies Weighted by Mode, N_t = 121 and N_c = 92

Reason	Treatment	Control	All
Earning a credit on my bill	30%	36%	33%
Doing my part for DEO	10%	14%	12%
Helping the environment	27%	27%	27%
Avoiding electrical service interruptions	12%	11%	12%
Don't know	20%	13%	17%

Customers were asked to rate, on a scale of 1 to 5, their agreement with various positive statements about Power Manager. Customers widely agreed that they would recommend the Power Manager program to others; that Power Manager events do not affect the overall comfort in their home; and that the number of Power Manager events is reasonable. Over 70% of both treatment and control customers agreed with those statements. But only 40% of treatment customers and 47% of control customers agree that the bill credits are sufficient. The distribution of responses for those who answered each question is shown in Figure 7-5.

Figure 7-5: How would you rate the following statements about Power Manager?
Response Frequencies Weighted by Mode, N_t = 121 and N_c = 92



The survey concluded with an opportunity for customers to provide free form suggestions on how they think the Power Manager program might be improved. Only 39% of respondents (83 of 213) offered suggestions. Among those offering suggestions for improvement, there were three common requests. The first, mentioned by 18 of 83 people, reflected a desire for more bill credits. The second, also mentioned by 18 people, expressed desire for feedback after an event. Most of these customers requested a more prominent notification of their bill credits. They say they didn't know an event occurred, and they would like to know if an event occurred and how much they earned in credit:

- "Periodically, reminders of credits applied. I forget to look for them."
- "I hadn't noticed anything on the bill, so highlight the credits."
- "If it would tell you on the bill how many times they had an event at the end of the summer."
- "Make the credits more obvious by notifying by email when the credit is issued in addition to showing on the bill."

The third most common comment, reported by 17 people, is the desire for warning before an event occurs or during the event:

- "Send me a text when an event occurs so I don't think my air conditioner is broken."

- "Alert consumers to when the event is happening."
- "Some people might not be home during the day, but that doesn't mean that their beloved animals aren't. I'd be upset if I found out that Duke turned my air conditioning down without a notice or text/email."
- "It would be nice to have a schedule of the occurrences to prepare for them."

Other comments centered on the AC cycling pattern or praise for the program:

- "Do not leave it off as long as it does, shorter intervals."
- "None, look into putting more in areas where there is time for high call and automatically be able to cycle."
- "Works fine for me."
- "Could do it more often."
- "The Power Manager program is so efficient that I do not know these events occur!"

Table 7-7 summarizes categorizations of the freeform responses. Suggestions categorized as "other" include requests to lower the overall cost of electricity, thanking the interviewer for the phone call, and concern for specific family members. Many respondents gave more than one comment, and often they gave one comment that fit into a category and one that did not. Since the answers often fit into multiple categories, the percentages add up to more than 100%.

Table 7-7: What suggestions do you have to make the Power Manager program work better for you? Response Frequencies Weighted by Mode, $N_t = 49$ and $N_c = 34$

Statement	Treatment	Control	All
Other	39%	31%	36%
I want more credits	19%	24%	21%
I want more feedback	15%	29%	21%
I want more notification	19%	18%	19%
It's a good program	10%	2%	7%
Change the cycling strategy	5%	7%	6%

Responses were positive when participants were asked to rate the likelihood of staying enrolled in Power Manager, with the large majority of respondents saying that they intend to stay in the program. Fully 71% of treatment and 74% of control respondents said they would "very likely" remain enrolled. Responses are tabulated in Table 7-8. Those customers, five respondents in total, who said they were not at all likely to stay enrolled were asked why, gave five disparate answers, which are tabulated in Table 7-9.

Table 7-8: How likely is it that you will stay enrolled in Power Manager? Would you stay…?
Response Frequencies Weighted by Mode, N_t = 121 and N_c = 92

Response	Treatment	Control	All
Not at all likely	4%	0%	22%
Somewhat likely	14%	13%	14%
Very likely	71%	74%	72%
Don't know	11%	13%	12%

Table 7-9: Why are you not at all likely to stay enrolled in Power Manager?
Response Frequencies Weighted by Mode, N_t = 5 and N_c = 0

Response	Group
It doesn't save me enough to warrant being a part of it	Treatment
Moved to a new house and installed geothermal	Treatment
I have already withdrawn from the program [because of the discomfort associated occasionally when we had guests]	Treatment
Because it is not working for my household, and I wasn't even aware I was enrolled in it	Treatment
Discontinued at recommendation of my AC supplier	Treatment

7.5 Interview Findings

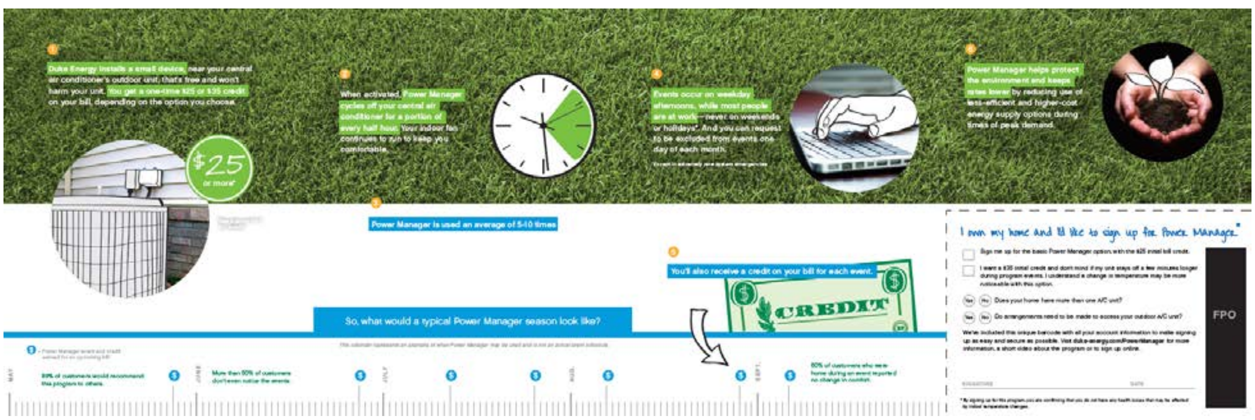
Power Manager is a mature demand side resource that is actively used in the course of operating Duke Energy Ohio's electric system. The demand savings delivered by Power Manager are made possible through the teamwork of internal and external stakeholders that manage the program's budget and goals, communicate with participants, maintain the Yukon event dispatch software, and interact with the customer at every stage of the program lifecycle, from enrollment, to device installation, to device removal. Three primary stakeholder groups, the Duke Energy program management team, Eaton Power Systems, and GoodCents, work together to deliver Power Manager to DEO customers. Nexant interviewed seven individuals from these organizations. Overall, through the course of our conversations, we observe that Power Manager maintains a customer focused orientation and is currently engaged in a number of initiatives to improve program operations and customer service. The remainder of this section will describe the Power Manager offering at DEO and what Duke Energy's activities are to bring in new program participants and support annual enrollment goals. A description of Duke Energy's activities to maintain Power Manager as a reliable system resource follows, which is followed in turn by an outline of work that continues after each load control season concludes to ensure Power Manager's continued success. This section concludes with a review of the activities that are planned or currently underway to further improve program operations and participating customer experience.

7.5.1 Program Offer and Enrollment Goals

Work to recruit new Duke Energy Ohio participants into Power Manager is concentrated in spring, just prior to the load control season. DEO's enrollment goal for 2016 was 1,040 participants, which was reached. The majority of recruitment into Power Manager takes place through outbound calling, fulfilled by the third party call center provider, CustomerLink. In some years, there are also direct mail and email recruitment campaigns initiated and managed by Duke Energy. For DEO, year round recruitment campaigns aren't necessary; the relatively small enrollment goals are achievable by targeted campaigns each spring.

As an outbound call center, CustomerLink is prepared to address common questions or concerns that DEO customers who are not familiar with the program may have, in addition to describing the basic features of the program, many of which are friendly to the program participants. Outbound callers are ready to speak to the fact that Duke Energy's customer research has shown that 85% of customers who are home during an event don't notice it, that there are generally only 5 to 7 events each summer, and that events typically end by 6pm, which is when many customers are just coming home from work. Another participant friendly aspect of the program is that air conditioning units enrolled in the program are cycled rather than completely curtailed.[13] Power Manager is also not called on weekends or weekday holidays. The load control devices used by the program, switches that directly control the air conditioner's compressor, are a proven technology that does no harm to the customer's air conditioner or the home's electric distribution system. Further, Duke Energy Ohio customers have the ability to opt-out of one Power Manager event per month without any penalty with respect to their incentive. Figure 7-6 provides an example of recent Power Manager marketing collateral used in the DEO jurisdiction.

Figure 7-6: Excerpt from Power Manager Direct Mail Marketing Collateral



The Duke Energy Ohio program offer provides for two different cycling levels—a moderate and high load control—determined by how much load shed the switches will yield during events through cycling (by cycling the air conditioner compressor's operation more or less during any given event hour).[14] Customers are encouraged to enroll in Power Manager through a one time sign up incentive, provided as a bill credit on their Duke Energy bill: $20 for the moderate option, and $35 for the high load control

[13] Unless a load control event is called as a result of a system emergency. In that case air conditioning units would experience full load shed. Emergency Power Manager events are extremely rare.

[14] There is also a low load control option, however it is not marketed. The low load control option is offered only to customers who are considering exiting the program in order to improve retention.

option. Further, Duke Energy Ohio Power Manager participants receive small annual payments that depend on the number and duration of events called during the summer. The annual payments are typically around $5 for moderate load control and $8 for the high load control option. In the realm of electric utility air conditioning cycling programs, these financial participation incentives are relatively small. Therefore, Duke Energy (and CustomerLink) emphasize messaging around community and environmental benefits to generate customer interest in the program. The program offer, which centers on the use of the outdoor switch, rather than an indoor programmable communicating thermostat, is found generally to be most successful with customer segments that are attracted to “set-it-and-forget-it” arrangements and those customers who would prefer not to have a service provider enter the home. Duke Energy has found that these preferences are correlated with older, higher income, and higher education demographics.

GoodCents is a third party provider that manages Power Manager customer care, handling participants’ inquiries about the programs or requests for customer service, and also all fieldwork. Power Manager fieldwork ranges from scheduling and routing load control device installations, training and managing a staff of device installers, responding to any device service calls, and responding to customer requests to remove load control devices. GoodCents reports that most new device installations are handled within 30 days of the customer’s enrollment, and that most customers don’t request installation appointments to work around pets or access issues. As a result, most installation appointments can be fulfilled using cost-effective routing and scheduling. GoodCents also manages and staffs all quality assurance inspections and fieldwork.

7.5.2 Power Manager Program Operation and Maintenance

In terms of maintaining Power Manager as a reliable system resource for the Duke Energy Ohio system operators, Eaton Power System plays an important role as a resource to assist Duke Energy program staff keep the Yukon software system running smoothly, managing firmware issues that can arise, addressing the switches, training GoodCents’ switch installers, and monitoring and managing the program’s communications links, which are provided through paging networks. An annual all-hands Spring Training event hosted by Duke Energy brings all the Power Manager program stakeholders together to discuss the upcoming load control season’s work. Also particular to 2016, a large scale quality assurance audit effort of load control switches was undertaken and staffed by GoodCents.

When it’s time to start calling events during the summer load control season, there is no proactive customer notification for each event. However, customers may call a toll-free number to get updates on the status of whether or not Duke Energy plans to call or has called a Power Manager event. At Duke Energy Ohio, program managers must decide when load control events will be called by 10:30am on a day-ahead basis. This day-ahead dispatch plan supports the program’s bid into the PJM market. Here, both strike prices and weather factor in to the decision to call load control events. The event calling team involves staff in system operations and fuels in addition to demand response operations. However, overall, demand response operations staff maintain control of the decision to call nonemergency events. Power Manager is viewed as an important resource for the Duke Energy Ohio system that depends on the participating customers’ willingness to remain enrolled. Therefore, all events are called with a view towards whether or not it will be a detriment to the experience of the participants. Considerations taken in this area are the number of events that have already been called during the current summer, or, during heat spells, during that week. Demand response operations staff also consider other finer points that lie outside of the program rules that are indicative of customers' willingness to continue to participate in the program; for example, whether or not Power Manager event hours have frequently gone late into the afternoon.

7.5.3 Program Monitoring and Postseason Program Maintenance

Duke Energy undertakes a number of activities both during the load control season and afterward to ensure that participants are satisfied with their Power Manager program experience and that the program is on track to provide an excellent customer experience going forward.

GoodCents, as the third party contractor that manages Power Manager customer contacts, has service level agreements in place with Duke Energy that outline service benchmarks, with both penalties for nonperformance and opportunities for incentives when benchmarks are exceeded. There are specific benchmarks in place to ensure that, during event days in particular, customer calls coming into GoodCents are handled quickly, efficiently, and that accurate information is provided to the customers calling in. Additionally, Duke Energy program managers monitor the number of calls coming in to the toll-free notification line, in addition to the number of calls coming into the GoodCents call center to detect any emerging issues associated with the program experience. Device removal requests are also tracked for this purpose.

Duke Energy uses seasonal reminder or thank you cards that are sometimes sent before the load control season, or sometimes after, to provide Power Manager tips for having a comfortable experience with the program. These cards are also sometimes used to recognize the program's megawatt contributions to reducing system load that year. Duke Energy's jurisdictions in the Midwest, including Ohio, typically have not used these mailings in the past. However, DEO plans to employ one of these program mailings this year to communicate upcoming program changes.

In 2016, DEO program managers are also leveraging smart meter interval data to identify Power Manager participants that may have broken or removed load control switches. Another effort that is currently underway to improve program performance is to look for missing switches in the homes of Power Manager participants who have recently received a rebate for a new heating ventilation and air conditioning system (HVAC); when these new systems are installed, the Duke Energy load control switch is usually left disconnected from the new system.

7.5.4 Upcoming Program Changes and Initiatives

Duke Energy is also engaged in a number of initiatives to change the program offering to make it more attractive to customers and to improve program performance. Emergency load control for customers on the moderate cycling program option will be moved from 75% to 66% in recognition that these customers selected moderate cycling due to the fact that they are at home and have more air conditioning usage during hot afternoons than those on the higher cycling option. The availability of event notification on the Duke Energy Ohio website will be evaluated, with a goal of making it easier for customers to access information about Power Manager events. Finally, Duke Energy is also engaged in a study to identify and change out certain models of older switches that are known to have high failure rates.

7.6 Key Findings

Key findings from the process evaluation include:

- 121 Power Manager participants were interviewed within 24 hours of the July 21 event, which had a high temperature of 91°F with a heat index of 95°F.
- 92 Power Manager participants were interviewed during a hot nonevent day (a control day), July 14, which had a high of 88°F with a heat index of 92°F. The control day surveys were used to establish a baseline for comfort, event awareness, and other key metrics.
- A strong majority of all respondents, 75%, reported that they are familiar with the Power Manager program.
- Only 13% of respondents on the event day reported that their homes were uncomfortable, while all of them experienced a load control event that afternoon. By comparison, 7% of Power Manager customers surveyed on a hot nonevent day reported they felt uncomfortably hot. While more respondents of the post-event survey stated that their home was uncomfortable that day than respondents of the nonevent survey (13% vs. 7%, respectively), the difference is not statistically significant and the difference in reported thermal discomfort cannot be ascribed to the Power Manager event.
- Over three quarters of participants would recommend the Power Manager program to others.
- The Power Manager staff and vendors are customer focused and undertake a number of activities both during the load control season and afterward to ensure that participants are satisfied with their Power Manager program experience.

Appendix A Regression Models Tested

All regression models were performed and the average customer loads throughout the summer using 15 minute interval data. The same sample of customers was analyzed using whole house interval and air conditioner end use data. The analysis only included days when maximum temperature exceeded 75˚F.

For the individual event day impacts (ex post), the regression equation took the general form of Equation 1, which will be estimated using a dataset made up of hourly observations of the average load in the M&V sample. Equation 2 describes the model used to estimate average event impacts for the general population events. The average event impacts were estimated separately to account for the effect of repeated treatments on confidence intervals.

Equation 1 and Equation 3 represents a within-subjects approach in which the observations on nonevent days are used to predict the counterfactual load for Power Manager customers on event days. A few points are noteworthy. The models were run separately for each 15 minute interval (equivalent to a fully interacted model) to account for occupancy patterns and produce different weather coefficients and constants. The only component that varied across the 10 models tested was how the weather variables were specified. Table A-1 shows the weather variables and explains the underlying concept for each model tested. To improve precision, same-day loads for the pre-event hours of 11am to 1pm were included to capture any differences between event and nonevent days that are not reflected in the model. The pre-event same day load variable functions as a same-day adjustment and is included because customers are not notified of the event in advance.

Equation 2: Ex Post Regression Model Individual Events

$$kW_{t,i} = a_i + \sum_{j=1}^{J} b_{i,j}\text{event}_{t,j} + c \cdot preeventkW_t + d_i \cdot weather_{i,t} + \sum_{k=1}^{7} e_{i,k}\text{dayofweek}_{i,k} + \sum_{l=5}^{10} f_{i,l}\, month_t + \varepsilon_{i,t}$$

Equation 3: Ex Post Regression Model Average Event (General Population Events)

$$kW_{t,i} = a_i + b_i\text{avgevent}_t + c \cdot preeventkW_t + d_i \cdot weather_{i,t} + \sum_{k=1}^{7} b_{i,k}\text{dayofweek}_{i,k} + \sum_{l=5}^{10} f_{i,l}\, month_t + \varepsilon_{i,t}$$

Where:

a	Is the constant or intercept
$b_{i,j}$	Represents the treatment effect of Power Manager during each interval, i, and each event day, j.
c-f	Are other model coefficients
i, k, l	i, k *and* l are indicators that represent individual 15 minute intervals (96 in a day), days of the week, and months of the year
t	Represents each date in the analysis dataset
$event$	Is a binary variable indicating whether Power Manager was dispatched on that day
$preeventKW$	Represents the same-day loads for the pre-event hours of 11am to 1pm. The variable functions as a same-day adjustment and is included because customers are not notified of the event in advance.
$weather$	10 different ways to specify weather were tested. Those are detailed in Table A-1.
$dayofweek$	Are a set of mutually exclusive binary variables to capture day of week effects
$month$	Are a set of mutually exclusive binary variables to capture monthly or seasonal effects
ε	Represents the error term

Table A-1: Weather Variables by Model Tested

Model	Weather variables	Concept
1	Cooling Degree Hour Base 70˚F (CDH)	The same hour temperature drives electricity use but air conditioner loads are only linear when temperatures are above 70˚F
2	Cooling Degree Day Base 65˚F (CDD)	The overall daily average temperature drives electricity use but air conditioner loads are only linear when average daily temperatures exceed 65˚F
3	Daily Maximum Temperature	The daily maximum temperature drives air conditioner electricity use
4	Average temperature over the 24 hours immediately prior	Heat buildup over the 24 hours immediately prior to time period drives electricity use
5	CDH and CDD	Both the daily average temperatures and same hour temperatures drive air conditioner electricity use
6	Same hour CDH and average temperature over the 24 hours immediately prior	Air conditioner use if influenced both by the temperature during that hour and by average temperature over the 24 hours immediately prior
7	Same hour CDH and average CDH over the 6 hours immediately prior	Air conditioner use if influenced both by the temperature during that hour and by heat buildup, as measured by CDH, over the 6 hours immediately prior
8	Same hour CDH and average CDH over the 12 hours immediately prior	Air conditioner use if influenced both by the temperature during that hour and by heat buildup, as measured by CDH, over the 12 hours immediately prior
9	Same hour CDH and average CDH over the 18 hours immediately prior	Air conditioner use if influenced both by the temperature during that hour and by heat buildup, as measured by CDH, over the 18 hours immediately prior
10	Same hour CDH and average CDH over the 24 hours immediately prior	Air conditioner use if influenced both by the temperature during that hour and by heat buildup, as measured by CDH, over the 24 hours immediately prior