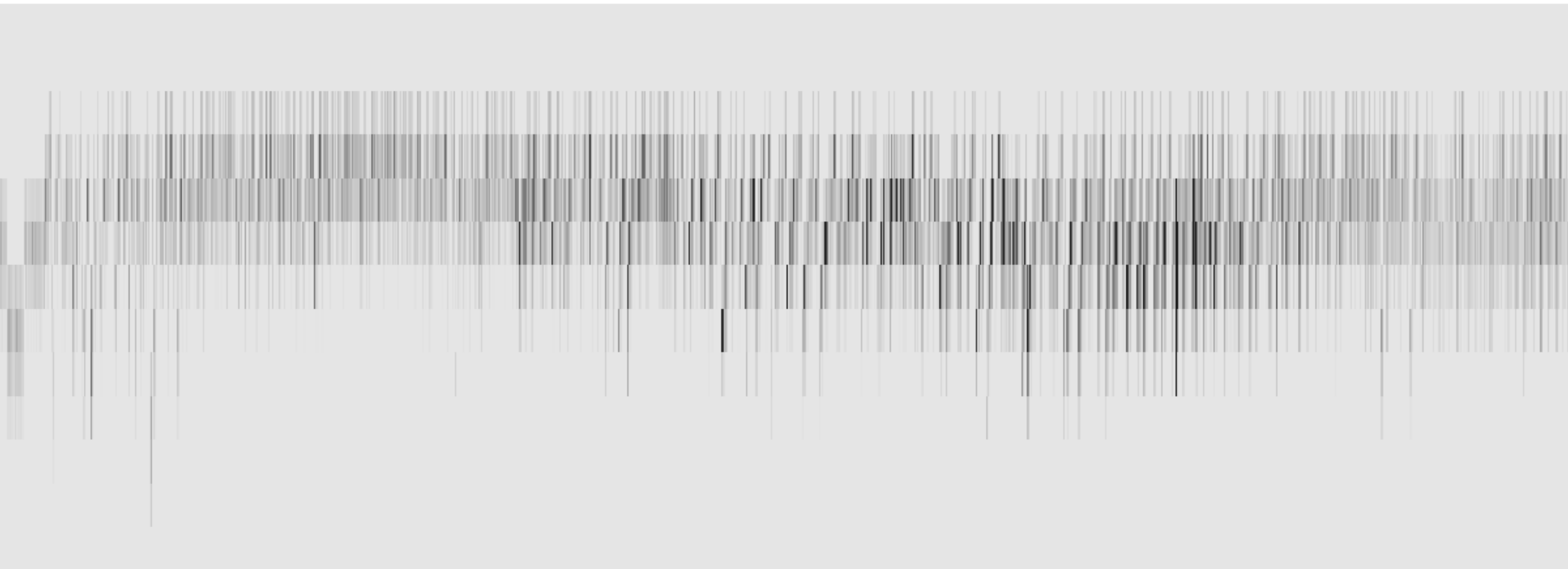
2018 Load Impact Evaluation for OhmConnect's DR Resource



October 21, 2019

Prepared for OhmConnect

CALMAC Study ID CPU0199.01

Prepared by
Sam Borgeson
Phil Price
Convergence Data Analytics, LLC

Convergence
Data Analytics

Cover image: Drawing from the pool of all 8,279 individual ex post event estimations that inform this report, each column represents from bottom to top, the computed impacts (darker is higher) between noon and midnight. The events are ordered from left to right by event start times.

Contents

Figures

Tables

Evaluation Context

Despite its adherence to the DR Load Impact Protocols (LIPs), this evaluation is not being done for a conventional utility-run DR program. OhmConnect intends to provide Resource Adequacy (RA) capacity to load serving entities (LSEs) outside of the Commission's Demand Response Auction Mechanism (DRAM) pilot, and understands the application of the LIPs to be necessary to ensure compliance with the requirements for qualifying capacity outlined in D.19-06-026. Specifically, D.19-06-026 stipulates that, beginning with the 2020 RA compliance year, "[a]ll demand response resources, whether third-party or investor-owned utility-managed, shall receive qualifying capacity values based on application of the load impact protocols." (Ordering Paragraph 18). Convergence Data Analytics, LLC has been hired to perform a DR Load Impact Protocols compliant evaluation of OhmConnect's DR resource and this report is the result of those efforts.

Because this is the first of its kind application of the Load Impact Protocols to a third-party DR resource, this report evaluates OhmConnect's DR resource using the DR Load Impact Protocols and serves as an existence proof that such an evaluation is feasible. However, we have encountered several circumstances that are likely to be of interest to the evaluation community. In particular:

- Third parties were not a topic of discussion when the DR Load Impact Protocols were written. The application of the protocols to third party resources place requirements on third parties and evaluators that are typically placed on IOU staff.
- Third parties will often not have access to the full set of customer account, historical meter, and non-participant data that utilities do.
- Third parties may call orders of magnitude more events with more widely varying participation and dispatch conditions than IOU-run DR programs.
- Third parties may be undergoing rapid growth with significant changes baked into their enrollment forecasts.
- Third parties will tend to have trade secrets or other sensitive or protected information they are shielding from their competitors and such information may not be available for public scrutiny.
- Third parties are unlikely to be as firmly established within the regulatory framework as IOUs and will tend to both fall into the gray area of regulation more often and require more regulatory support than established IOU evaluators.
- There is not consensus among stakeholders as to how third parties should proceed under D.19-06-026 and the rules and decision governing Resource Adequacy (RA) to produce their RA numbers. This is an area that will require further clarification from the CPUC.

These topics are addressed in more detail in a discussion section that the end of this report.

Executive Summary

This report presents results of a load impact analysis of OhmConnect's aggregated residential behavioral DR capabilities. OhmConnect is a demand response (DR) aggregation company that offers residential customers a system of rewards and encouragements to shift consumption away from "OhmHour" events. OhmConnect has "gamified" DR. Customers accrue points for OhmHour participation. The points are accrued by customers based on the difference between their metered OhmHour consumption vs. their rolling average baseline consumption and are redeemable for cash and/or prizes[1]. Customers that save more deeply and consistently are rewarded with more points per event. There is also a system of tokens that can be used in the token marketplace for prizes or for special cards that give customers temporary "power ups" that help customers protect their savings streaks and earn more points[2]. Customers are tracked at the sub-LAP level, differentiated into tiers based on past performance, and encouraged to configure devices (i.e. smart plugs and communicating thermostats) that OhmConnect can dispatch through direct load control.

During 2018, the year this report is based on, OhmConnect was primarily bidding their events into the DR Auction Mechanism (DRAM). In a process that is performed once a week, the total pool of OhmConnect customers was dynamically split into sub-groups that form resources qualified by CAISO to participate in the DRAM. For this reason, OhmHours have been responsive to periods of grid stress as reflected by prices in the spot market for electricity. Specifically, OhmHours tend to be called during periods of high locational marginal prices (LMP). However, OhmHours are called year-round. OhmConnect called at least one event on each of 182 days in 2018.

In this document we present:

1. Ex post load impact estimates for OhmConnect events for the year 2018 (PY2018)
2. Ex ante forecast of OhmConnect events for program years 2020-2030
3. A deeper dive into issues of particular relevance to the future of OhmConnect evaluations and forecasts

Within these analyses, we examined impacts across geography and by customer segments. This included findings for:

- *Sub-LAPs.* The natural geography for DRAM resources is the sub-LAP. Therefore, sub-LAPs are the primary geography used to disaggregate event impacts in this report.
- *Local capacity areas (LCAs).* There are ten California Independent System Operator (CAISO) LCAs[3] in California, spanning a great deal of geographic/climatic variability: **PG&E territory**: Greater Bay, Greater Fresno, Humboldt, Kern, North Coast / North Bay, Sierra, Stockton; **SCE territory**: Big Creek / Ventura, LA Basin; **SDG&E territory:** San Diego. The LCA analysis provides insights on the magnitude of available capacity from events in each geographic area.
- *Tiers.* OhmConnect places customers into performance tiers including Silver, Gold, and Platinum, based on their event performance (depth and consistency) over time. Tiers are relevant to customers because they are markers of status and point rewards are higher for the higher tiers.
- *Mega events.* When savings are particularly important, OhmConnect changes messaging and increases point rewards for "mega events". Events of this type can be evaluated and compared to regular events.
- *Direct load control devices.* Through subsidies, encouragement, and point/tier rewards, OhmConnect customers are encouraged to deploy one or more direct load control devices in their homes. These devices include smart plugs and communicating thermostats that are automatically dispatched by OhmConnect during events. Customers with devices can be expected to save more deeply and reliably than those without.

All of the above categories inform our understanding of ex post results and are explanatory variables in our ex ante model. (It is trained using sub-LAPs, not LCAs, but they directly correspond with one another).

[1] See https://www.ohmconnect.com/faq/en/articles/1056874-how-do-i-use-my-points for the rules on redeeming the points and appendix A for a more detailed description provided by the company.
[2] See https://intercom.help/stealthohmconnect/en/articles/1056869-tokens for more information on how tokens are used.
[3] See http://www.caiso.com/informed/Pages/StakeholderProcesses/LocalCapacityRequirementsProcess.aspx for more details on the CAISO local capacity requirements process.

The OhmHour events examined for this report lasted from 1 to 2 hours, with timing that falls within the RA window, which extends from 4 pm to 9pm. In the 2018 sample, just 0.25% of the events started before , the RA window, and none extended after it. OhmConnect is very active in calling events. In total, data from 726 events, spanning 182 days of 2018, was provided by OhmConnect to CDA for evaluation. Ex ante results are based on hypothetical event performance across the full RA window.

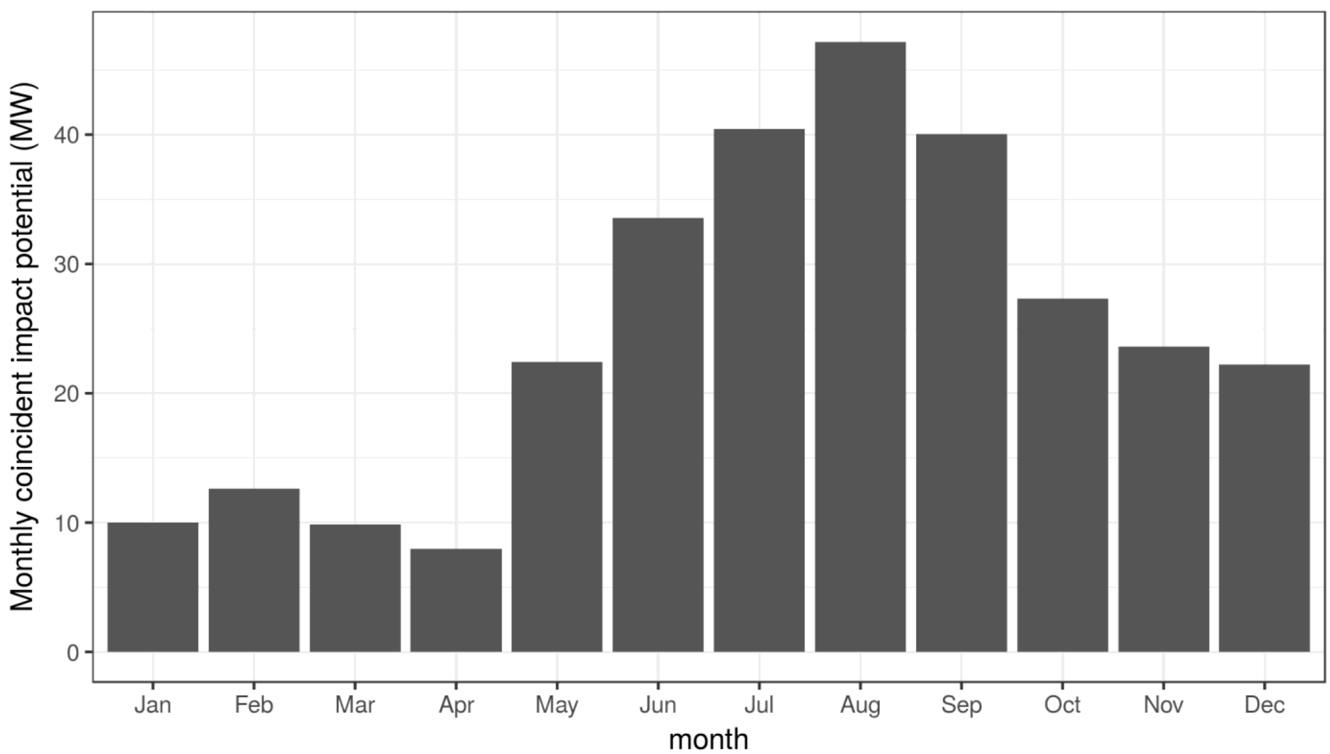
Ex Post Load Impacts

Raw events vs. monthly summaries

OhmConnect called 726 events on 182 days of 2018. It is impractical to represent each individual event in all the tables and figures of this evaluation. Therefore, we have created a summary view of event results by month. In the "monthly roll-up" of events, a single unbiased average of per-participant impacts, aggregate impacts, reference loads, outside temperature, and count of participants across all events with the same start and end times for each month are created.

PY2018 Aggregate Impacts: The aggregate ex post program load impact for a typical event in the summer (June-September) of 2018 was 1.1 MW for just 2,142 customers, but this is misleading because OhmConnect calls different slices of the customer base across rapid-fire events that cumulatively lead to the dispatch of the majority of customers each month. We call the product of the per-participant impacts and the count of unique customers dispatched in a month the "monthly coincident impact". Figure 1 presents monthly coincident impacts for each month in PY2018. The monthly coincident impact for August was 47.2 MW from 70,984 customers.

Figure 1: Ex post monthly coincident aggregate impact by event month



PY2018 Per-Participant Impacts: The average per-participant impact in 2018 was 0.51 kW with significant variation across sub-LAP, month of year, and customer segments. Figure 2 presents the average per-participant impacts for all customer sub-groups defined by device, tier, LCA, and mega event. Among these, device ownership and tiers are the most highly differentiated. Non-device owners show average impacts of 0.55 kW while device owners average 1.29 kW. The Silver (and below) tier achieves an average impact of just 0.44 kW, Gold provides 0.69 kW, and Platinum (and above) deliver the greatest per-participant impacts of any single sub-group with 1.57 kW. We also note that (1) mega events did not appear to out-perform conventional events on average in 2018 and (2) there was stand-out performance in the hot climates of Fresno and Sierra and under-performance in the cooler coastal weather of the Bay Area, Humboldt, and even San Diego.

Figure 2. Comparison of typical per-participant impacts (average of June-September) for several categories of customer attributes

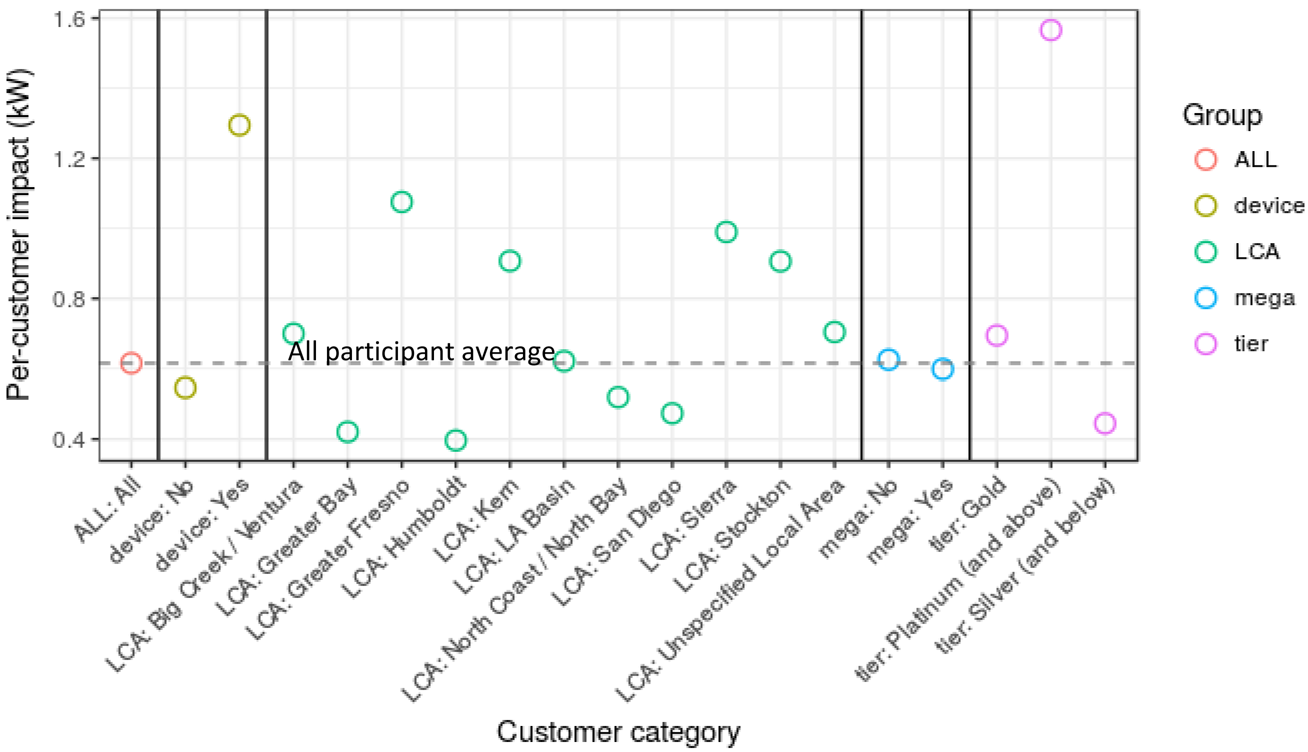
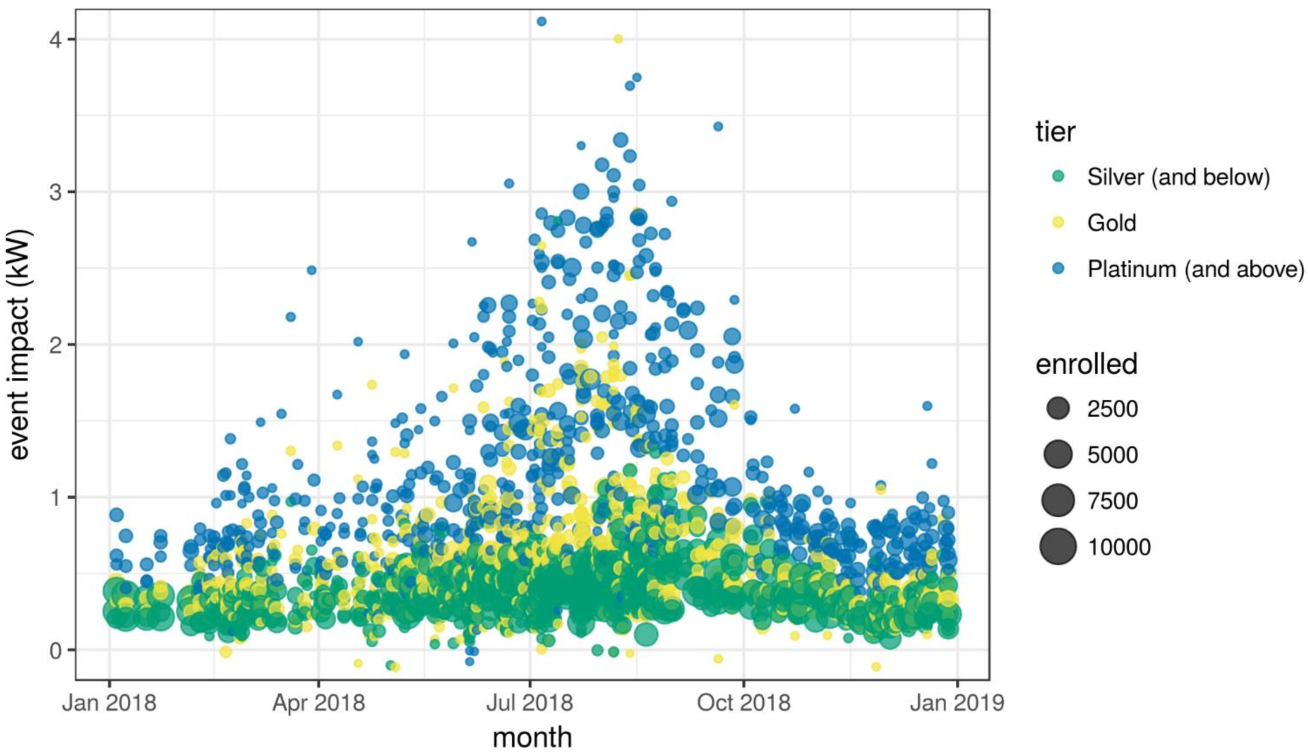


Figure 3 illustrates typical per-participant event average impacts for all events called in 2018 (i.e., each circle represents an event) for each customer tier. This figure shows the extraordinary number of events as well as the clear differentiation between customer tiers and the correlation between summer weather and elevated per-participant performance. For reference, platinum tier participants average 3.5x the impact of Silver (and below) tier participants.

Figure 3: Per-participant average impacts by tier for each event in 2018



Comparison with PY2017 and PY2019: Table 1 offers a comparison of ex post results drawn from 2017, 2018, and 2019 and from winter (Feb) and summer (Aug). OhmConnect has been growing the number of unique customers called per month. There were 16,000 unique event participants in August of 2017, just shy of 71,000 a year later, and 80,000 in August of 2019. In August, higher per-participant reference loads are associated with the two years with higher average temperatures. Despite having lower temperatures and reference loads, both per-participant impacts and impacts as a % of reference loads were highest in 2018, but the greater average number of event participants in 2019 leads to aggregate impacts per-event much higher in 2019 than other years, however, the monthly coincident aggregate impacts for 2018 and 2019 are comparable.

Table 1. 2018 ex post comparison with 2017 ex post and 2019 (YTD) ex post

month	year	# of events	average # of participants	# monthly unique participants	total # participants	average temperature (F)	reference load (kW)	per-participant impact (kW)	aggregate reference load (MW)	aggregate impact (MW)	impact %	coincident reference load (MW)	coincident aggregate impact (MW)
Feb	2017	6	1530.83	7707	9185	51.75	1.09	0.31	1.68	0.47	27.94	8.44	2.36
	2018	47	1524.40	39999	71647	54.56	1.13	0.32	1.73	0.48	27.78	45.39	12.61
	2019	51	2685.59	65517	136965	50.54	1.31	0.37	3.52	1.00	28.53	85.76	24.47
Aug	2017	15	1844.67	16672	27670	**82.15**	**2.08**	0.58	3.84	1.06	27.64	34.72	9.60
	2018	119	2250.86	70984	268075	78.30	1.89	**0.66**	4.26	1.50	**35.10**	134.46	**47.19**
	2019	22	10976.41	80220	241526	**81.40**	**2.10**	0.61	23.07	6.70	29.05	168.61	**48.99**

Ex Ante Load Impacts

As opposed to 'ex post' analyses, which quantify what has happened in the past, 'ex ante' forecasts attempt to quantify the future. Ex ante load impact forecasts predict the load impact that would occur in standard event times and conditions for a total of four hypothetical weather-years. For this report, forecasts are made for the peak day of each month for two standard weather years – corresponding to conditions that are expected to lead to peak electric load in either one out of every two years, or one out of every ten years on average. There is a slight difference in the conditions that cause peak load for the California Independent System Operator (CAISO) and for Investor-Owned Utilities (IOU). Combining the two peaking conditions with the two types of weather year leads to four sets of standard weather data.

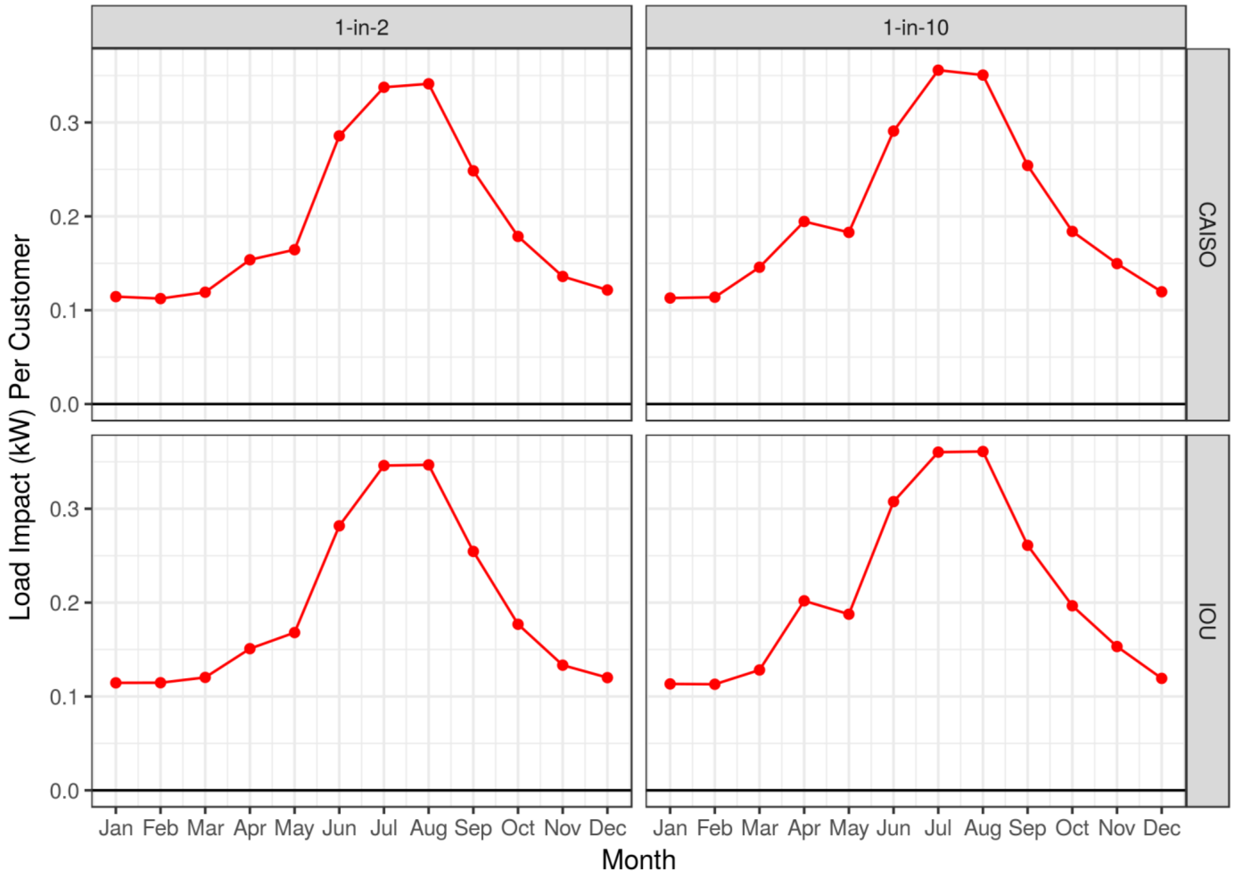
Ex ante predictions are made by statistical models that predict the load impact per participant for different categories of customers (such as customers with and without electric devices that are automatically controlled to reduce load during a demand response event). The models were fit to data from calendar year 2018 and used to predict the load impact per participant for the standard weather conditions mentioned above, for the projected mix of customers by category that is forecast for future years. The predicted load impact per participant multiplied by the projected number of participants equals the projected aggregate load impact.

For reasons discussed in the body of the report, the number of participants that is used for the ex ante predictions is much lower than OhmConnect's projected customer enrollment: after accounting for strategic exclusions and opt-outs, only about 25-40% of OhmConnect's customers are projected to contribute load shed during a standard event.

A major product of the ex ante analysis is a 'table generator' – a spreadsheet – that allows the user to select a forecast year and a set of weather conditions and see estimates of load impact by hour, in each sub-LAP separately or the total of all sub-LAPS.

Figure 4 shows the predicted load impact per OhmConnect customer for each of the four sets of weather conditions described above, for the resource adequacy (RA) time window that runs from 4-9 p.m. For these four hypothetical weather-years the highest load impact per customer occurs in July and August, at about 0.35 kW per customer; the lowest occurs in December through March, at about 0.12 kW per customer.

Figure 4. Predicted load impact per customer, by month, for four standard sets of weather conditions.



OhmConnect expects to grow quickly over the next few years (see the section on enrollment forecasts in the body of the report). The projected growth in enrollment varies by sub-LAP according to where OhmConnect intends to market their program most heavily combined with where they think it will be most popular. The result is a projected rapid increase in aggregate load impact.

Figure 5 shows the predicted aggregate load impact for 2019; this is the product of OhmConnect predicted enrollment in each month times the predicted load shed per customer in each month, again for the standard weather conditions. The maximum is around 41 MW statewide, in August. Whereas the predicted load impact per customer (Figure 4) is about the same in January as in December, this is not true of the aggregate because the number of projected event participants increases substantially over the course of the year.

Figure 5. Predicted aggregate load shed by month, for four standard sets of weather conditions.

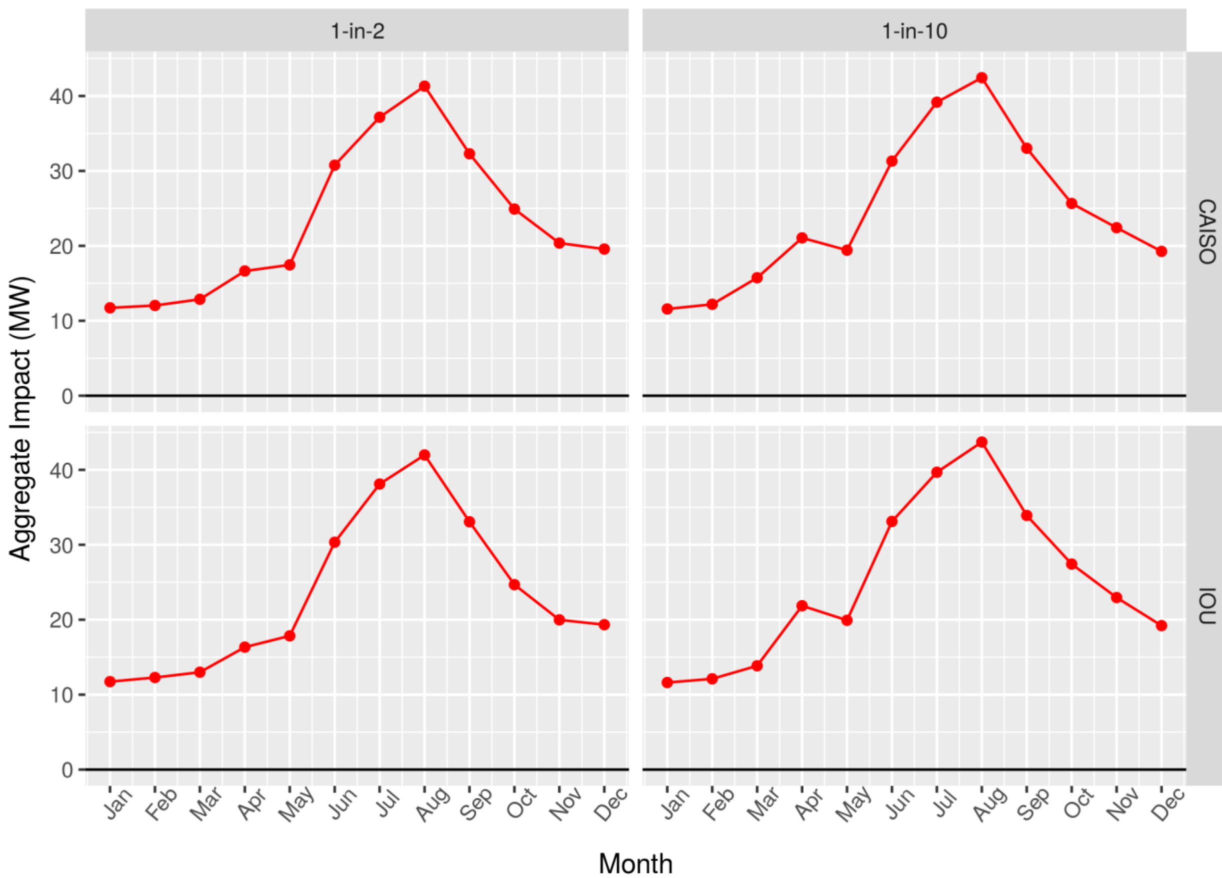


Figure 6 shows the same sort of information, this time for the next several years. The projected year-over-year increase in customers (and thus event participants) leads to a very rapid year-over-year increase in load impact.

Figure 6. Forecast aggregate load impact, by month and year. Points for 2022 are hidden behind 2023.

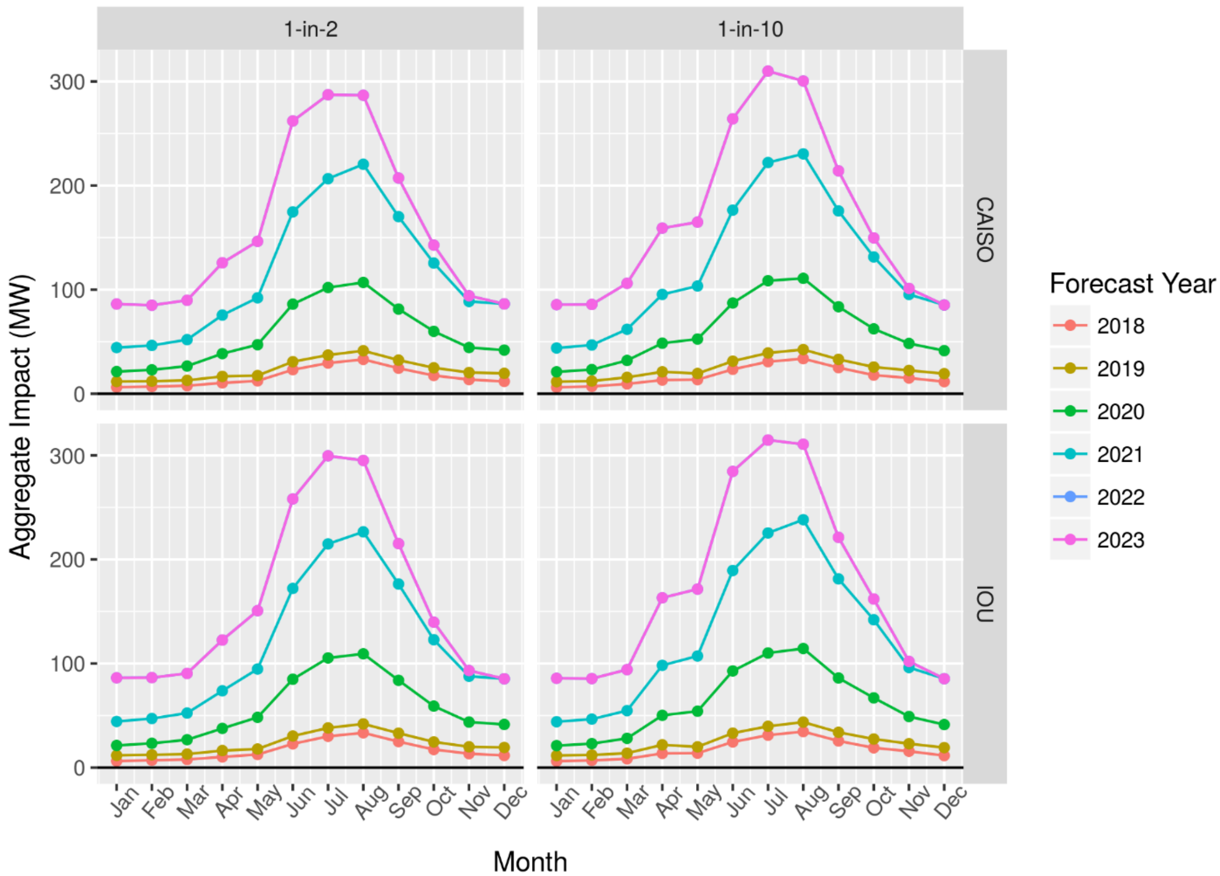


Table 2 shows the forecast load impact by year for the August 1-in-2 monthly CAISO peak day. The very large projected increase over the next two years is mostly due to a projected large increase in enrollment, especially in sub-LAPS with high load shed per participant; secondarily, the fraction of customers who provide direct load control is forecast to continue to increase, and these customers provide more load shed than others. The fastest growth is projected to be in sub-LAPs that experience high temperatures, so the average event temperature experienced by OhmConnect's customers increases from year to year. Table 3 shows the predictions for IOU peak days rather than CAISO.

Table 2. Forecast Aggregate Load Impact for August CAISO 1-in-2 day.

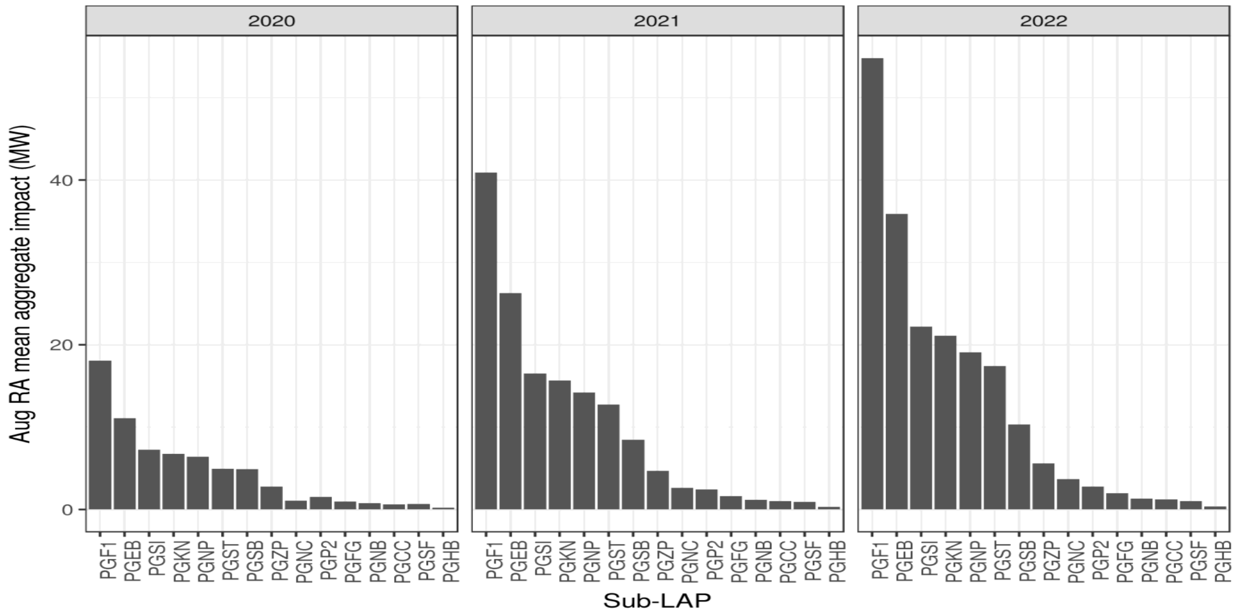
Year	Temp (F)	Aggregate Impact (MW)
2018	80.58	32.97
2019	81.37	41.31
2020	82.86	106.93
2021	83.99	220.43
2022	84.48	286.72
2023	84.48	286.72

Table 3. Forecast Aggregate Load Impact for August IOU 1-in-2 day.

Year	Temp (F)	Aggregate Impact (MW)
2018	82.14	33.40
2019	83.19	41.98
2020	85.12	109.29
2021	86.45	226.50
2022	87.04	295.09
2023	87.04	295.09

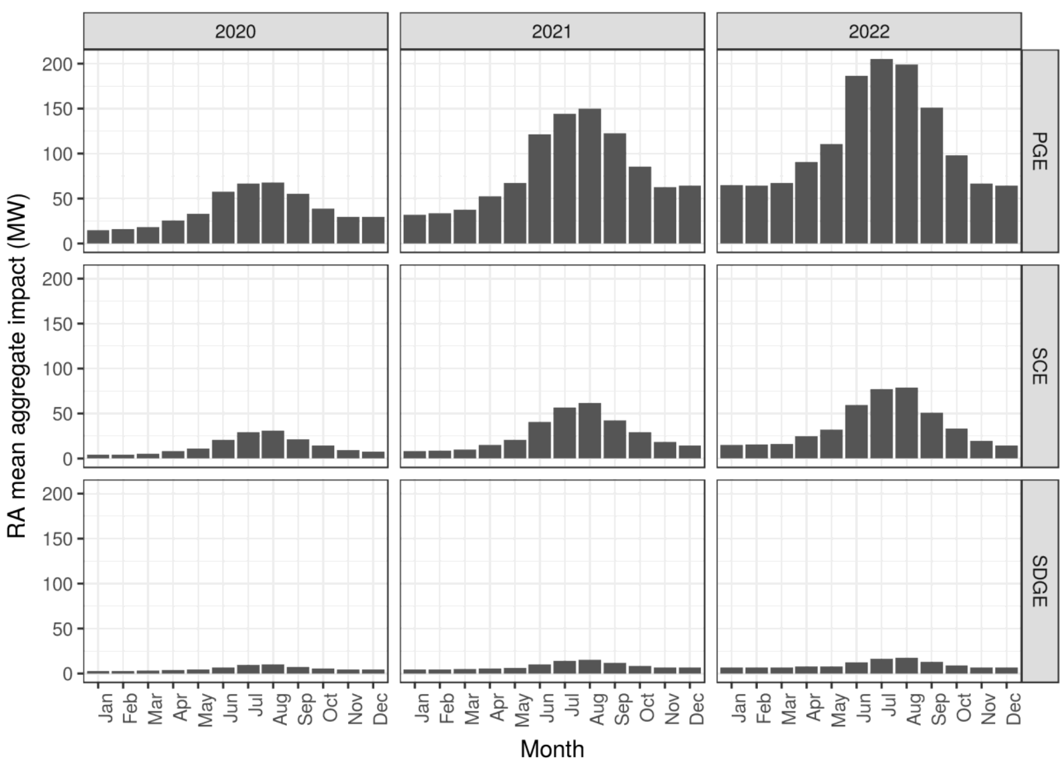
The ex ante predictions were made for each sub-LAP - there are sub-LAP-specific terms in the models, as illustrated in the appendix on modeling – but sub-LAPs can be mapped onto Local Capacity Areas to make forecasts at the LCA level. Figure 7 is an example: it shows the predicted aggregate impact for August, for the IOU 1-in-2 weather year.

Figure 7. Forecast aggregate impact by sub-LAP, for August of different years.



Aggregate impacts by month are shown in Figure 8, separately for PG&E, SCE, and SDG&E.

Figure 8. Forecast aggregate impact by month, separately by utility.



Conclusions and Recommendations

Based on our evaluation of the 2018 dispatch of OhmConnect's behavioral DR resource, we provide the following recommendations:

Recommendation	Description
Investigate more robust baselining and comparison group methodologies for estimating event impacts	OhmConnect's DR events have short durations, so comparison day baselines with same day adjustments (SDAs) can be used for estimating impacts for the handful of hours around event windows. Joining a long list of evaluators, we have documented that raw 10 in 10 baselines would not have been serviceable without SDAs. But more importantly, baselines can't establish that the event savings were caused by the event as convincingly as randomized or even synthetic controls. For a pioneering behavioral program to convince its skeptics that its results are legitimate, adopting the most rigorous methods would go a long way.
Emphasize device deployment	Customers with direct load control devices save more deeply and more reliably than purely behavioral customers. Emphasizing device deployment in the future will make savings more reliable and help customers accrue rewards with less effort. And on the grid side, automated DR is a higher quality resource.
Recruit where customers perform	We have documented significant variation in event impacts by location (sub-LAP). If a company like OhmConnect can optimize recruitment to emphasize performance when and where the grid needs it most, all will benefit.
Work to clarify the application of LIPs to third parties	Determining how to apply the DR Load Impact Protocols (LIPs, written in 2008 to describe the methods and process for evaluating ratepayer funded DR programs with an emphasis on long term planning) to OhmConnect's Third Party resource, which is intended to be sold to non-IOU Load Serving Entities for RA compliance, has been challenging at times. This report and the other evaluation assets should help guide progress towards setting standards that can be extended into the future, but care should be taken to achieve greater clarity around what exactly is required of future evaluations of this resource.
Standardize how the unique aspect of OhmConnect's resource are tracked and presented in future reports	One of the challenges of this evaluation process was the number of events that OhmConnect calls. The OhmConnect event dispatch platform consists of emails and text messages, backed by a centralized "gamified" website. OhmConnect called 726 events on 182 days in 2018. This is up to two orders of magnitude more than traditional DR programs call. It presents challenges in how to standardize results, how to share them via table generators, etc. We recommend that care is taken when evaluating this type of program, and that over time, the method used by CDA continues to be improved.

Introduction to the 2018 OhmConnect Resource

OhmConnect is a free service that rewards customers for reducing their demand when requested via text message or email. Once a user signs up and connects their utility data to the OhmConnect platform, they receive a notification about once per week to save energy. These demand response events, which vary between one and four hours, are called *#OhmHours*. If a user saves energy compared to their baseline, they are rewarded through prizes, cash, or other incentives. Rewards are proportional to the amount of energy saved and OhmConnect aggregates the event impacts into a "behavioral DR" resource and sells them into various grid-facing markets and programs[4].

OhmConnect is continually recruiting new customers and refining its methods and strategies for dispatching events. The data set used for this evaluation consists of event data from approximately 100,000 customers called during 2018, along with data from approximately 45,000 called during 2017. In addition, while we focused our evaluation on 2018, 2019 data through August was also available to the evaluation team. OhmConnect called slightly more customers by September of 2019 than were called during all of 2018. **Overall, therefore, this evaluation is based on all RA window dispatched event data from 2018, backed by all 2017 event data, as well as 2019 data through August of 2019.**

OhmConnect calls events through email and text messaging notifications, backed by a centralized website that implements a set of "gamified" rewards for event participants. OhmConnect has the technical capacity to call any combination of customers at any time and they are very active in calling events. Over the course of 2018, there were events called on each of 182 days of the year, and typically more than one event was called on event days.

In total, OhmConnect called 726 events, most of which lasted only one hour. The large majority of these events began between 6 pm and 8 pm. Figure 9 illustrates the count of events called for each month of 2018. Table 4 below it tabulates the number of events starting at each hour of the day. For the purposes of this report, and in alignment with OhmConnect's internal practice, events are considered distinct if they involve a unique combination of date, utility, and event characteristics (e.g., standard events versus "mega events" described below, and event start and end times).

[4] Because of the unique nature of OhmConnect's offering, CDA has requested a full summary of OhmConnect's DR model directly from them. Their description can be found in Appendix A of this report.

Figure 9. Event counts for each month of 2018, but the number of hours the event lasted (aka 'duration')

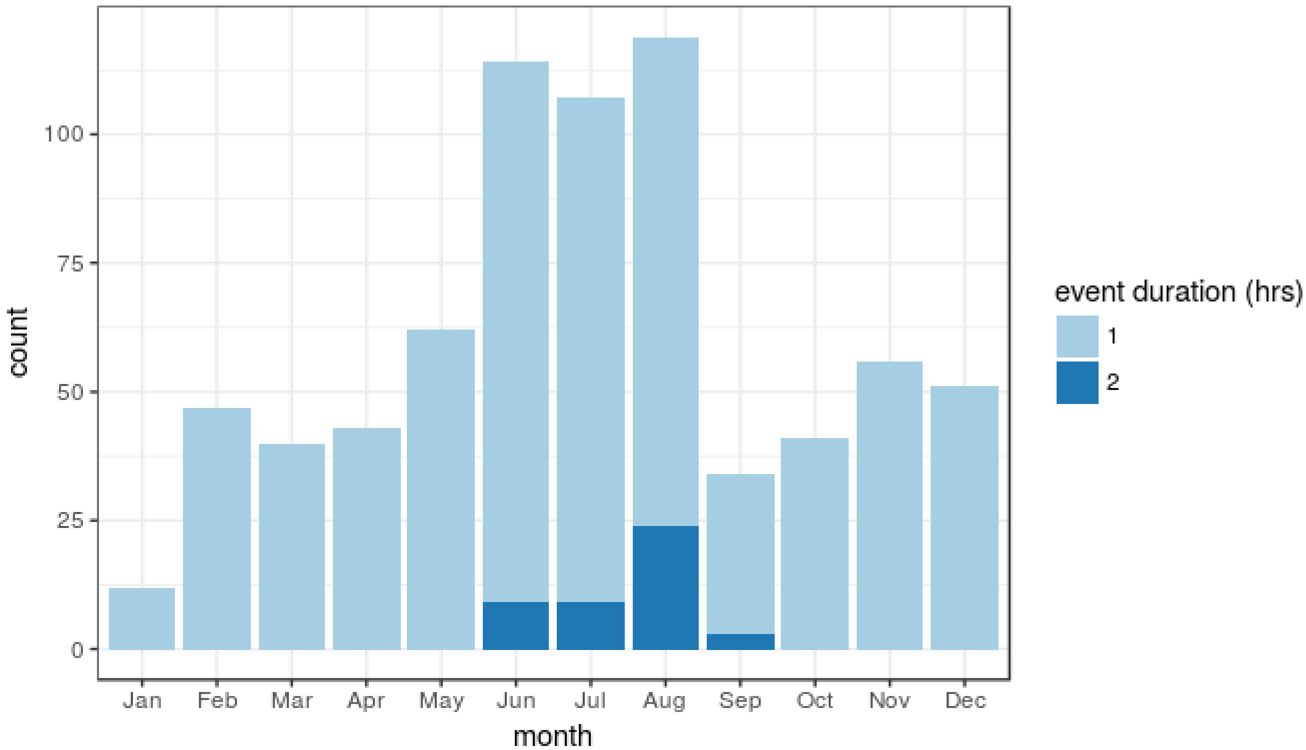
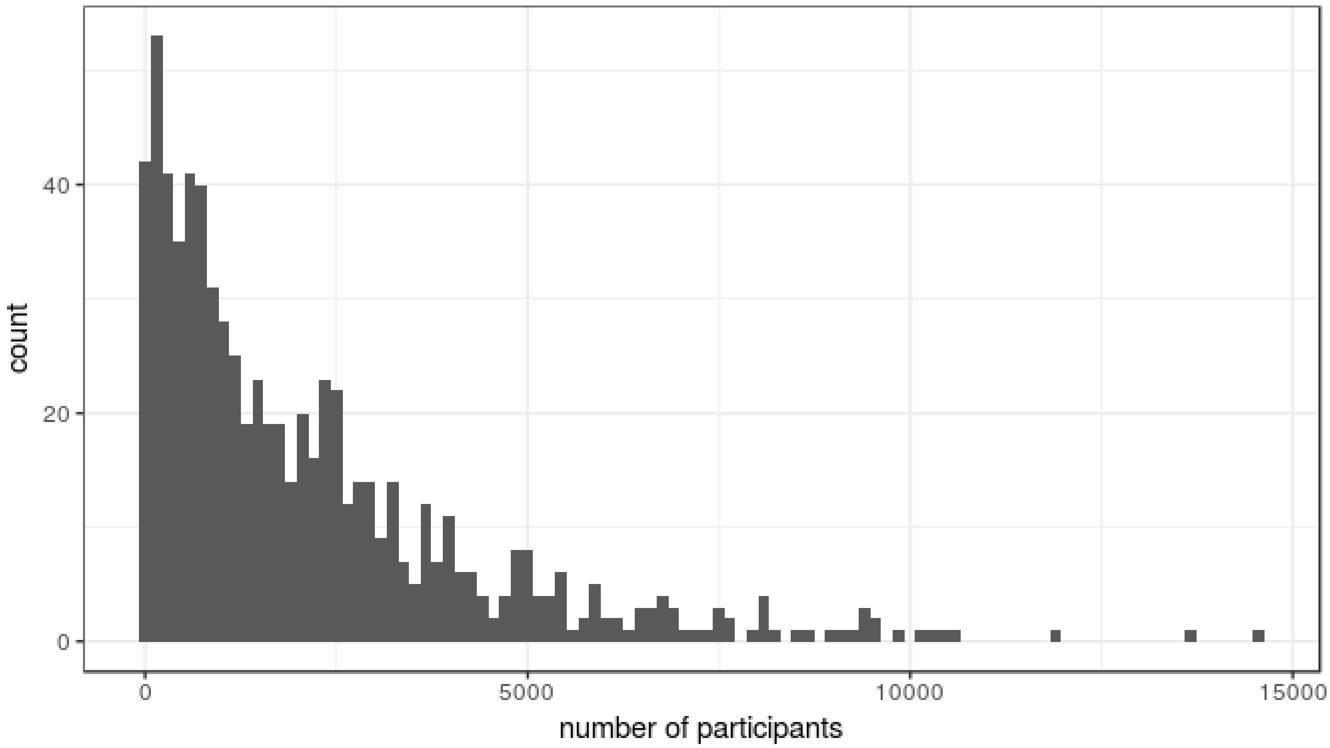


Table 4. Event counts by start time

start	count
2 pm	1
3 pm	1
4 pm	26
5 pm	59
6 pm	182
7 pm	249
8 pm	208

Events are called with a wide-ranging number of participants. Figure 10 presents a histogram of the number of customers participating in all events from 2018. The average is 2,142, but 50% of events are below 1,453: the distribution has a long tail up toward higher counts.

Figure 10. Histogram of the number of customers participating in OhmConnect’s 2018 events



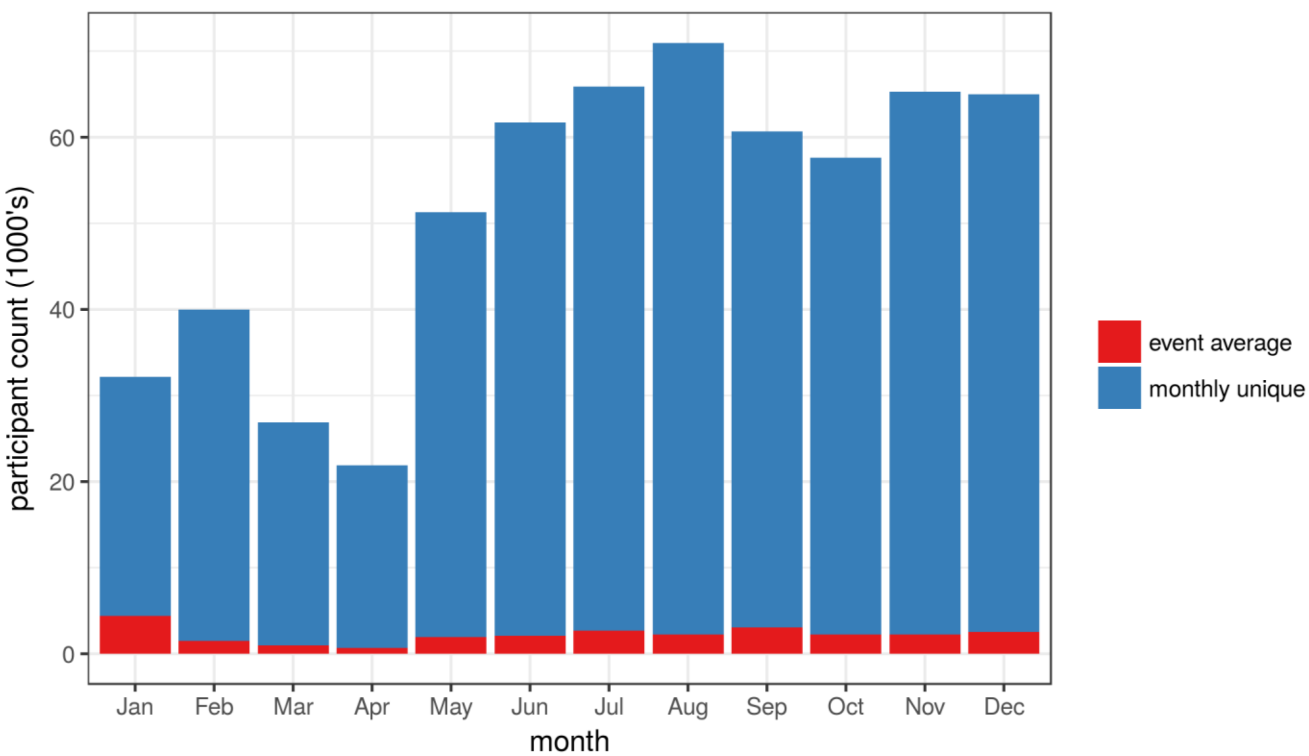
In addition to the variability in the number of customers called for each event, there was also a lot of variability in which groups of customers were called. The average event calls just 2% of active users, but over the course of a month, tens of thousands of customers are called. Figure 11 illustrates the average number of customers called per event in each month and the count of *unique customers* called in each month.

While OhmConnect’s approach to calling events is foundational to their program model, it presents a puzzle for the evaluation of the resource. Typical DR program events call a significant fraction of their customer base for every event. They may be resolved by geography or other relevant factors, but they are rarely called in such small fractions of their total enrolled population. Thus, it is possible to assess the full dispatchable capacity of traditional DR in terms of their aggregate impact during single events. The aggregate impact of an OhmConnect event, by comparison, is a small fraction of their dispatchable resource. The result is that individual OhmConnect events may not be representative of the full population (or the potential for that population).

Monthly coincident impacts and loads

To address this dilemma, we report two different versions of ex post aggregate results in sections of this report. We report the per-event **aggregate impact and aggregate reference loads**, but we also report the monthly **"coincident impact" and "coincident reference loads"** by multiplying monthly average per-participant impacts and reference loads by the count of unique customers dispatched during each month. The idea behind those numbers is to give a sense of the magnitude of saving that would be achievable if OhmConnect called events with all of their monthly dispatched participants. Because those numbers are based on unique participants, rather than the total count of called participants (i.e. including some more than once), it is a more realistic view of resource potential.

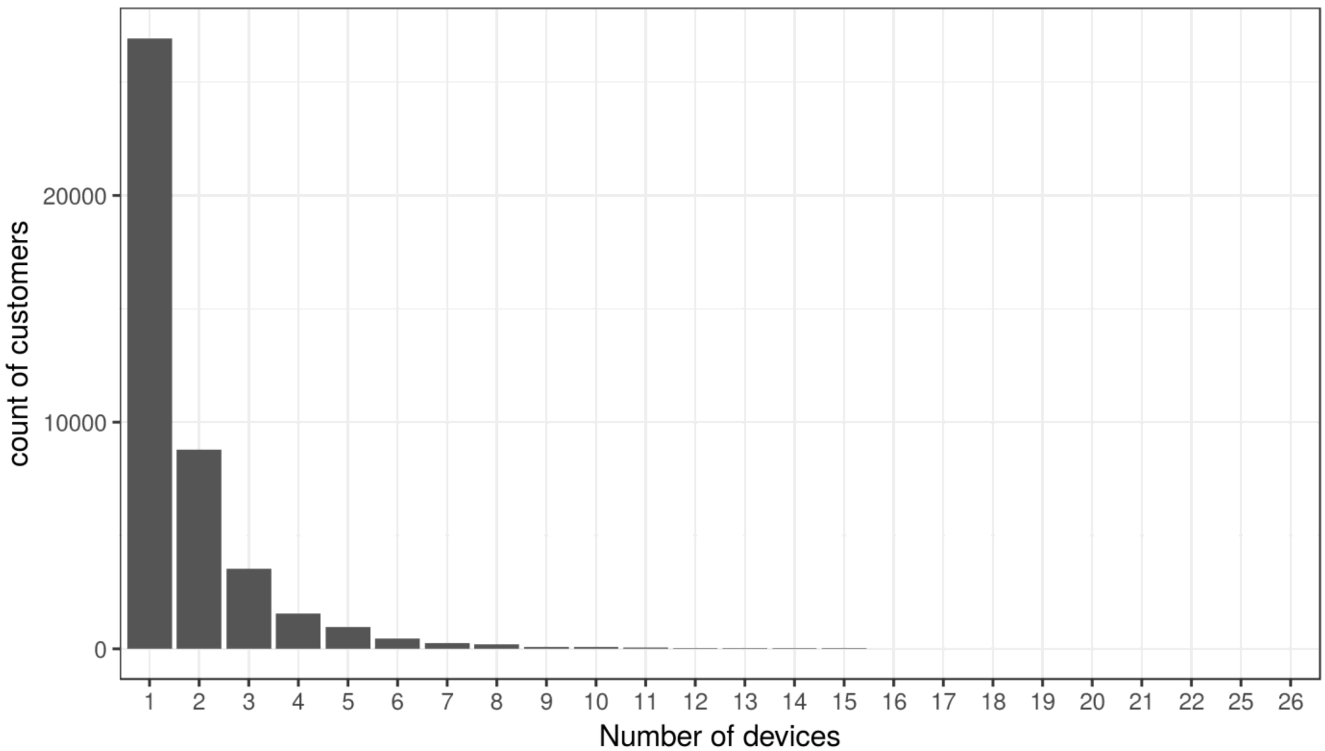
Figure 11. Event average participant count vs. the number of unique customers called during each month.



Devices

OhmConnect encourages its customers to install devices in the form of smart plugs and communicating thermostats that can be configured to be automatically triggered during an OhmHour as a form of Automated Demand Response. The high degree of automation OhmConnect is able to achieve with devices is thanks to their two-way communication infrastructure that monitors all registered devices for their “heartbeat” and records a real-time indication of whether each customer participating in an event had a device that was currently connected as a part of their event data. Due to the resulting level of event response automation, these devices tend to increase the magnitude and reliability of event impacts. In 2018, approximately 36,000 devices were dispatched across 100,000 customers. Figure 6 illustrates the count of households with one or more device. While the majority of device-owning households have just one, there is a significant portion of the population that has several devices in use.

Figure 12. Histogram of the count of customers with one or more device (smart plugs or thermostats) in their homes.



Tiers

OhmConnect has “gamified” DR responses by offering behavioral incentives for customers to respond more deeply and reliably to events. One of the key parts of the incentive system is the concept of earned customer “tiers”, which are ordered as Bronze, Silver, Gold, Platinum, Diamond according to past performance. There are just a handful of customers below Silver or above Platinum tiers. Therefore, we consolidate the tiers for this report into Silver (and below), Gold, and Platinum (and above). During 2018, approximately 53% of event participants were Silver (or below), 31% were Gold, and 15% were Platinum (and above).

As might be expected, the highly motivated customers who out-perform their peers are also more likely to have adopted direct load control devices. Table 5 summarizes the % of customers in each tier with at least one device and the proportion of all devices that can be found within that tier. The Platinum customers make up just 15% of customers, but comprise 31% of customers with devices because nearly 1 in 4 of them has at least one device.

Table 5. Proportion of tier members with devices and proportion of devices in each tier (note percentages rounded to nearest whole number may not add to 100%)

tier	% of customers in tier	% of tier with devices	% of devices in tier
Silver (and below)	53	9	39
Gold	31	12	30
Platinum (and above)	15	24	31

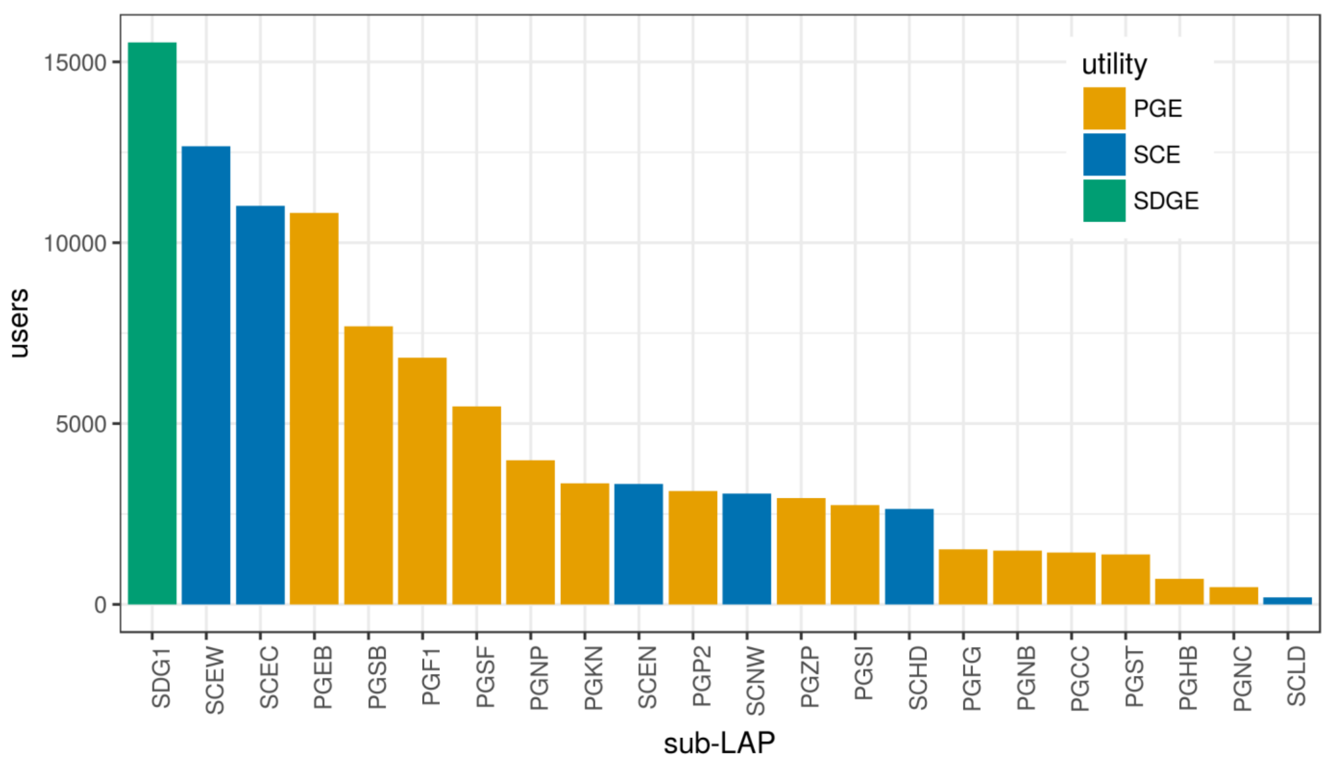
Mega events

For "mega events", OhmConnect increases the rewards and modifies their messaging to customers. Events are considered to be the dispatch of a customer group unique to a utility, date, time range, and mega status. In 2018, Around 13% of events were mega events.

Sub-LAPs

OhmConnect calls events when the grid is stressed. The most important events come when Locational Marginal Prices (LMPs) are high. For this reason, the natural geography for grouping customers is the sub-LAP. In total, there are 22 sub-LAPs. Figure 13 illustrates the number of active customers from each sub-LAP, where "active" indicates that the customer was dispatched at least once in 2018. In 2018, these dispatches included 353 events in PG&E territory, 264 for SCE, and 109 for SDG&E.

Figure 13. The count of customers active during 2018 in each of California's sub-LAPs, colored by utility territory



LCA

The CAISO recently re-drew sub-LAPs to fall cleanly within LCAs. Now that sub-LAPs and LCAs have been harmonized, each sub-LAP belongs to at most one LCA. Although sub-LAPs are the primary geography for OhmConnect's business, this report presents some results in terms of LCAs. There are 10 LCAs, as well as a group of customers that do not fall into any LCAs.

Table 6. Count of customers active in 2018 by all combinations of LCA, device ownership, and tier

	Device			Tier		
LCA	**All**	**Yes**	**No**	**Silver (and below)**	**Gold**	**Platinum (and above)**
Big Creek / Ventura	12916	495	7742	10432	1588	896
Greater Bay	55067	4302	37254	43660	7198	4209
Greater Fresno	14369	472	10519	12033	1414	922
Humboldt	1688	79	1487	1213	269	206
Kern	6882	201	4574	5723	772	387
LA Basin	46972	1893	26971	38692	5295	2985
North Coast / North Bay	7554	450	5645	5861	1105	588
San Diego	27712	1405	16580	22041	3763	1908
Sierra	5061	306	3962	4174	537	350
Stockton	2677	103	1719	2231	271	175
Unspecified Local Area	21510	773	14956	17737	2393	1380

Key concepts and decisions

This section marshals in one place evaluation-relevant details of OhmConnect's resource and the circumstances of this evaluation that have required careful thought and/or contributed substantially to the outcomes or interpretations presented in this report. The information and caveats presented in this section are necessary to understand our methods and results.

Baselines vs. control groups

The results in this evaluation are based on the use of a 10 in 10 baseline with a same day adjustment. We have been asked why we did not use non-participant controls instead. The answer is twofold. First, as a non-utility, OhmConnect did not have access to utility account data for any of their customers or any data at all on non-customers for us to request. Further, OhmConnect customers can request that OhmConnect deletes their historical data at any time and finally, OhmConnect only stores event-day consumption and 10 in 10 baseline data in their production data systems. They may or may not have been able to compile sufficient data for us to use synthetic controls in the time available and we may or may not have been able to match controls in a manner that we believed to be unbiased, but we were not confident that this report could be filed on time with so many uncertainties in the timing and availability of workable data. Second, for short duration events (OhmConnect called just 1 or 2 hour events in 2018), baseline methods can produce reasonable results.

OhmConnect does not randomly select controls for each event, so the use of a control group would require the construction of a synthetic control group after the fact. To enable synthetic controls, an evaluator must make unverifiable assumptions about what makes customers similar and apply them as the matching criteria (outside of special cases, it cannot be proven that these assumptions produce unbiased controls).

In the DR evaluation world, similarity is typically determined by a mix of customer attributes and non-event day load characteristics. OhmConnect has set the expectation with their customers that their data will only be collected and used for event evaluation. And their version of event evaluation has always involved 5 in 10 or 10 in 10 baselines (meter data from a few days before each event gathered for just the event participants). Even if we were to accept the baseline days as the comparison days, we still couldn't have that data for non-event participants unless OhmConnect went out and retro-actively harvested it for all candidate controls from utility green button systems. This was not something we deemed feasible in the time available nor did we judge it necessary, given the caveats associated with synthetic controls.

We do typically use synthetic control groups in our DR evaluation work, but the resulting comparison loads still "drift" compared to event loads and often require a same day adjustment heading into each event. At that point, many of the benefits of true controls have been lost as the "time series" constraints of the same day adjustment come to exert significant influence over the evaluated impacts. In truth, for a typical DR program dispatched under 20 times a season, evaluators will inspect all event outcomes and "hand tune" their approach to both matching and the same day adjustments until the results satisfy intuition, relationship to past results, and diagnostic heuristics on expected outcome. Whether evaluators are using controls or comparison days, there is a lot of judgement required in that process, which is why evaluation isn't just a mechanical process without human input. In our judgement baselines/time series approaches, with same day adjustments, can produce reasonable estimates for short duration events.

Having said all that, we allow that our estimates could be either over or under the unobserved ground truth outcomes. We are recommending that OhmConnect look into dispatching at least a subset of their events with true randomized controls and that they look into what it would take to have more non-participant meter data on hand for control matching in the future.

5-hour ex-ante events

By definition DR is meant to be a temporary reduction or shift in loads. The reality is that very few DR resources have been designed to be dispatched for 5 hours in a row. DR through thermostat set point adjustments, for example, only suppresses AC loads until the building hits the new set point and starts cooling again. Behavioral DR, which relies on the active participation of customers is also subject to "fatigue" during long events. For these reasons, it is unusual to see 5-hour DR events in any program. However, the Load Impact Protocols require ex ante estimates for all 5 hours of the RA window. This poses a question for evaluators: are the ex ante numbers just individual time-of-day predictions for event performance (i.e. for each of the 5 hours) or are they really intended to model hypothetical 5-hour sustained events? In practice, CDA observes that evaluators often choose the former – each hour is estimated separately. This is partially due to lack of data from long events, but also due to their interpretation of what "capacity" means in the context of DR resources and the lack of guidance from the LIPs on the topic (the LIPs merely call out resource adequacy as an "other potential application"). In the present case, we have data from mostly one-hour events and we have estimated each RA hour separately to produce the ex ante numbers for this evaluation.

Imputed solar customers

OhmConnect does not have access to utility customer account data related to net metering or other indicators of PV ownership. However, customers with on-site PV generation can exhibit loads that are highly variable (for example on a partly cloudy day), dip below zero, etc. This makes PV customers a likely source of uncertainty or non-representative results. In response, OhmConnect has flagged all customers with loads that dip below zero and probable PV customers. In 2018, there were approximately 4,500 customers flagged as having PV and 102,100 without. CDA has separately tracked those customers for this evaluation, excluding them entirely from ex post per-participant impact calculations, participant counts, and other summaries of customer counts found in this report. However, ex ante models do include an indicator based on such customers and they are included (weighted by their small population) in ex ante predictions. Our assumption, as manifest in the enrolled customer forecast is that the PV customers will remain the same portion of each combination of sub-LAP and customer tier in the population as they were in 2018 over the course of the ex ante forecast window, so they change with sub-LAP and tier enrollment forecasts.

Defining inclusion fraction, participation, etc.

Among all of their enrolled customers, OhmConnect has historically excluded customers experiencing unusually low consumption (i.e. event day consumption lower than the average of the 10 preceding non-event days) from being dispatched as part of events. The ratio of the number of customers not excluded in this manner (aka **"included customers"**) to the full pool of eligible customers is the **"event inclusion fraction"** (or percentage). For reference, inclusion fractions can often dip below 50%. Among the included customers, there are some who opt-out of individual events or events with certain characteristics (like time of day) and there are others who have de-enrolled. Included customers who nevertheless did not participate in an event are called **"non-participants"**. The remaining customers who are called for an event and do participate are called the **"event participants"**. All ex post results are presented in terms of their participant counts, with average impacts as **"load impact per-participant"**. By convention, the ex post table generators use the misleading label of "enrolled" for participant count. We have adopted that convention despite the fact that the numbers presented are the count of event participants.

The inclusion fraction and opt-out rates are most relevant to the calculation of ex ante aggregate impacts. The enrollment forecast provides the top line enrollment numbers, but as described above, event participation is moderated by the inclusion fraction and then further reduced by customer opt-outs, de-enrollment, and other forms of non-participation. We have used historical event data from 2017 through August of 2019to calculate seasonal inclusion fractions (see the ex ante methods for more details) and we have used 2018 event data to calculate non-participation rates by month, tier, and device ownership. The ex ante aggregate impact calculations start with the enrollment forecast and make adjustments using the empirically derived inclusion fraction and then non-participation rates to calculate a count of forecasted participants. Aggregate ex ante impacts are the model-predicted per-participant impacts multiplied by the forecasted participants.

However, there is one final wrinkle to these adjustments. By convention, the ex ante table generator presents its results in terms of enrollment counts, not participation counts, so the aggregate impacts (calculated as described above) are divided by the forecasted enrollment counts to produce impact per enrolled customer, which become the numbers presented as the "Average per Enrolled Customer" in the table generator. In the rest of the report, our numbers are typically presented as average per event participant or are labeled otherwise.

OhmConnect plans to loosen their customer inclusion rules for future events, so they are very likely to see contributions from historically excluded customers to their participant counts and aggregate impacts. Nevertheless, *CDA decided to conservatively exclude them entirely from ex ante participant counts and aggregates until estimates of their event impacts can be made*. This all but guarantees that the reported ex ante aggregate numbers are lower than OhmConnect can achieve through including a greater fraction of customers in the future (all else being equal), but we had no empirical basis for assigning numerical values to their expected impacts.

Note to DRAFT report readers: CDA mis-understood the nature of the excluded customers at the time of compiling the draft report and table generators, so draft ex ante aggregate estimates were not pro-rated by the inclusion fraction. Final ex ante aggregate impacts have been corrected.

Monthly event "roll-ups"

CDA identified and estimated ex post impacts for 726 discrete OhmConnect DR events in 2018, with the permutations of customer attributes and locations bringing the total number of discrete event day evaluations to 8,279. There are practical reporting issues raised by having such a large number of events and event results to report on. To address this CDA performed **"monthly roll-ups"** on events to produce representative monthly summaries of all event activity. For all events called for the same time of day during the same month in the same utility territory and sharing the same set of customer attributes, the monthly roll-up computes the average number of participants, the total counts of unique participants, the total count of all participants (i.e. with users counted more than once if they are in more than one event), and the participant-count-weighted average reference loads, impacts, and outside temperatures, with associated errors and aggregate values. The month of June, for example, had 114 events, which were summarized by 18 "monthly events." This is because there were 6 different time ranges of events called in the PG&E territory, 7 different time ranges in SCE's territory, and 5 in SDGE's territory. These "monthly events" summarize all the events that occurred is a more manageable format and are the data products that are fed into the ex post table generator.

Per-event vs. monthly "coincident impacts and loads"

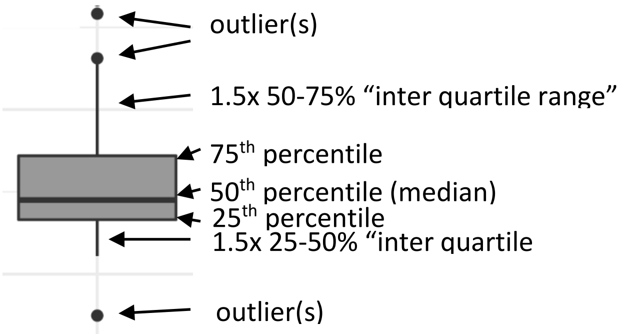
As specified in the Load Impact Protocols, ex post results are presented with aggregate impacts derived from event participation. However, the aggregate impact of an OhmConnect event is a small fraction of their dispatchable resource, which may not be representative of the full population (or the potential for that population). Therefore, the per-event aggregate impacts provide little intuition for magnitude of OhmConnect's resource. However, over the course of a month, OhmConnect dispatches tens of thousands of unique customers. To aid interpretation of the results, CDA calculates and reports the unique participation for monthly roll-up events, and the "coincident impact" and "coincident loads", which we define as the per-participant average event impact for the month multiplied by the number of unique participants called each month.

The "coincident" in these terms refers to the coincident dispatch of customers. In other words, what if they had all been called at the same time? We recognize that "coincident" can be used in contexts similar to this to describe activity that occurs at the same time as the grid peak demand. **Our "coincident impact" should not be misconstrued to imply that the impacts were delivered at the time of the grid peak.**

Box and whisker plots

A frequently used visual tool in this report is a "box and whisker plot". This type of plot visually summarizes the full distribution of data it represents by providing visual depictions of the 25th, 50th (aka median), 75th percentiles of the data values as well as whisker extensions that run up to either the min/max value or 1.5x the "inter-quartile range" beyond their starting point, with values beyond the whiskers plotted as "outlier" points. These are a good alternative to mean values with error bars because they more accurately reflect the distribution of data in cases where it is not well fit by a normal distribution or is not symmetrical about the mean and they also preserve the actual max and min values. They are a good alternative to plotting all the data because the human eye can falter trying to accurately break down a cloud of points into percentiles. Where we deem it appropriate, we have also included labeled mean or weighted mean values in our box and whisker plots.

Figure 14. Box and whisker plot example.



Key Research Questions and Study Methods

Key Research Questions

The research:

1. Estimates the ex post load impacts for the OhmConnect resource for PY2018 (the most recent full year)
2. Estimates the ex ante load impacts for the OhmConnect resource for years 2020-2029
3. Looks at device, customer tier, mega event, sub-LAP, and LCA effects
4. Looks at the impacts of weather and time of year

Challenges to our analysis included:

- **Tight evaluation timeline:** The ruling that mandates the use of the DR Load Impact Protocols for OhmConnect’s resource was finalized just a few months prior to the RA QC filing deadline at the end of October (D.19-06-026 Ordering Paragraph 18). Every aspect of this evaluation was shaped by the need to complete in weeks an effort that typically takes months.
- **Large number of events:** The large volume of events called by OhmConnect presents logistical challenges related to managing and processing so much data and practical challenges related to how the results should be presented and summarized.
- **Limited non-event-day data:** OhmConnect is not a utility but they have been granted access to their customers’ meter data via Green Button. However, it is not their practice to keep non-event-day data in their data systems. OhmConnect’s data systems have been optimized for computing just-in-time individual event results on a per-participant basis using 10 in 10 or 5 in 10 baselines. For this reason, potential control matching or baselining schemes involving non-event-day data were not feasible.

Ex Post Impact Analysis Methods

For the **ex post analysis**, we estimated load impacts and reference loads for participants on event days compared to the same day adjusted top 10 in 10 non-event day baselines. The analysis produced estimates of hourly reference loads and event average impacts, with errors, for all participants. We also estimated load impacts for mega events and sub-groups determined by device ownership, customer tier, sub-LAP, and LCA.

More specifically, we estimated average hourly reference loads and load impacts (both with uncertainties) and tabulated participant counts (called enrollment for historical purposes in the ex post table generator) and weighted temperatures for each hour of each event day for every customer sub-group modeled for or reported in this report.

The steps to evaluate each event are as follows:

(1) Meter readings from the event day are “actual event load” per customer.

(2) For each event day, the previous ten non-holiday weekdays that were not demand response event days are the "comparison days."
(3) For each hour of the day, calculate the mean load over the ten comparison days. This is the unadjusted "raw baseline load" for that hour.
(4) Compute mean actual event load and baseline load, with uncertainties, across all event participants (and then do the same for all examined sub-groups of customers participating in each event).
(5) Determine a "same-day adjustment" that trues up the difference between baseline and actual load during non-event hours[5].
(6) The "adjusted baseline load" for a given hour is the unadjusted baseline modified by the same-day adjustment (discussed below).
(7) Impacts are computed as the difference between adjusted baseline and actual load during event hours.

Actual event and raw baseline loads

To compute event impacts, we leveraged OhmConnect's access to 24 hours of "actual" hourly meter data and "Top 10 of 10" "raw baseline" calculations for each individual customer for each event day. The actual and raw baseline loads were averaged across all event participants into event average actual and raw baseline load shapes, with hourly standard deviations representing the variability across the population of participants.

If the baselines track actual event day loads closely (during non-event hours), event impacts could be calculated as the simple difference between the actual loads and the raw baseline. However, baselines are, by definition, captured on non-event days while event days are characterized by unusually hot weather and unusual grid conditions. The result is that even when aggregated across all event participants, the raw baseline is not a good proxy for what the actual loads would have been absent an event. This conclusion comes primarily from observing the poor match between non-event hours across the two, where the raw baseline is typically lower than the non-event-hour actual loads and ramps more slowly as well.

A good mental image is of baseline loads with similar non-AC end uses compared to actual event day loads. The "load shape" of air conditioning ramps into the mid-day and peaks in the afternoon, unless it is determined by occupancy, in which case it may peak during the evening. Event days tend to be hotter than their baseline days and the divergence between the raw baseline and the event actual loads tend to exhibit a pattern consistent with elevated event-day AC loads.

Same day adjustment

Same day adjustment (SDA): additive and multiplicative

When there is a mismatch between the available baseline and event actual data, it is common to apply a "same day adjustment" that trues up the difference between the adjusted baseline and the actual loads in the hours just prior to the event. The simplest same day adjustments involve calculating the mean baseline and actual loads from 4 hours to 1 hour before the event start. For an additive adjustment, the difference between these two can be added to the entire raw baseline, producing a fixed offset that causes the loads just prior to the event to overlap.

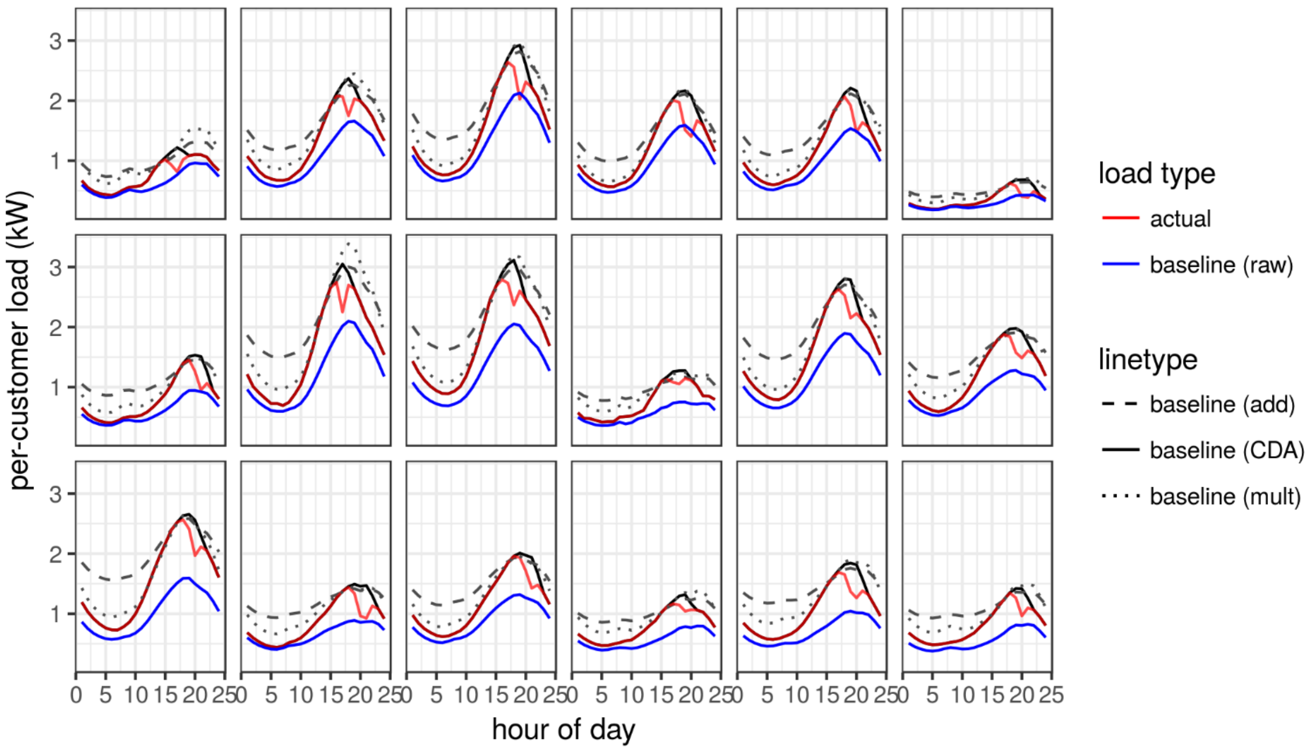
[5] The multiplicative baseline adjustment values used in the DRAM, described in the Southern California Edison document "10-Day Average Baseline and 'Day-Of' Adjustment" (document NR-285-V4-0418), are available to us, but we have found those corrections inadequately control for the difference between comparison and event day consumption.

For a multiplicative adjustment, the raw baseline can be multiplied by the ratio of the pre-event loads, one again ensuring that the load just prior to the event are well matched (ratio of 1) while scaling all other hours by the same amount.

Imperfections in SDAs

Figure 15 illustrates baselines with additive (add), multiplicative (mult) SDA, and the approach used for this evaluation and explained below (CDA or "dynamic adjustment"). The closer the baseline is to the non-event-period actual consumption, the better a baseline it is.

Figure 15: Event-wide mean raw baseline and actual load with 3 same day adjustment options (first 18 monthly summary events from July, where each event differs by utility and event timing).



For this evaluation, we considered and rejected both additive and multiplicative adjustments. Due to the lower ramping rates of the raw baselines, the *actual loads grew faster than the additive adjusted baselines into the critical event period*. A baseline that under-estimates loads produces under-estimates of impacts, so the concern about the additive adjustment was that it would under-estimate event impacts. For both "mult" and "add", pre-event baseline loads are consistently above the actuals, and the adjusted baselines typically continue higher than the actuals after the events. These characteristics make the adjustments poor choices for estimating 24 hour impact curves.

The multiplicative adjustment might have done better under ideal conditions, but because it is based on a ratio, it is highly sensitive to loads that are near zero and the data set has no reliable indicator of solar customers. The resulting ratios could balloon into the hundreds and those customers would exert a very oversized and clearly incorrect influence over the results. In the DRAM, the CAISO uses a multiplicative same day adjustment, but has a cap on the ratio. This approach introduces an arbitrary cutoff in the data and we found that even with the cutoff, the customers close to that cutoff value can be over-represented in the results, likely resulting in over-estimates of savings.

Dynamic adjustment SDA

In the end, we opted for a “dynamic” same day adjustment that enforces the following properties:

(1) it is an additive adjustment with different values for each hour of the day whose properties are derived from the difference between the actual event loads and the raw baseline loads.

(2) For the pre-event period, the adjustment is equal to the difference between the raw baseline and the actual event loads.

(3) For the second hour before the event, the adjustment continues at half its hourly rate of change (i.e. slope) from the preceding hours. In other words, if the gap between the baseline and the actual was growing that trend is extended with half the slope. If the gap was closing, that trend is extended. If there is no gap, there is no adjustment.

(4) For the hour just before the event, the slope of the adjustment is cut in half again.

(5) The rate of change in the adjustment is zero during the event.

(6) The adjustment ramps down in a linear fashion to become equal to the difference between the raw baseline and actual loads once again two hours after the event.

(7) Thereafter the adjustment is equal to the difference between the actual and baseline load until the end of the day, so that the estimated impact is zero for these hours.

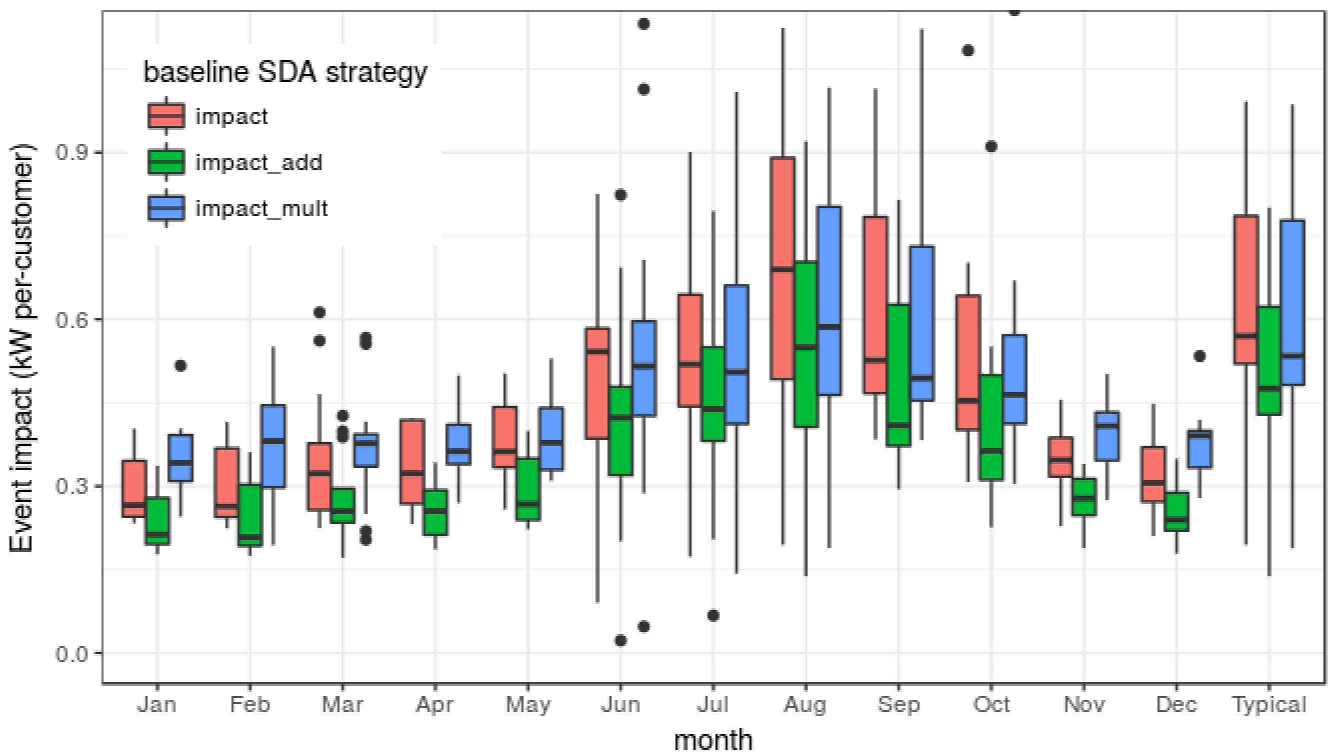
In effect the dynamic SDA performs an interpolation of the difference between the raw baseline and actual loads, anchored to the hours prior to and after each event. This approach is similar to the additive adjustment, but it allows the adjustment to take on different values in each hour just before, during, and after events. This approach takes advantage of the time-series nature of loads and the fact that the majority of events lasted just an hour and none lasted longer than two hours to anchor a tightly constrained interpolation of the additive difference to conditions *both before and after* each event.

Validation of SDA approach

Figure 16 compares monthly and typical per-participant impact estimates generated using the three discussed SDA approaches. As would be expected if baseline days ramp more slowly than event days, the additive adjustment (green) comes in consistently below the other two. The multiplicative adjustment (blue) and the "dynamic" SDA we applied (red) have similar outcomes during summer months, but the multiplicative one makes systematically higher estimates during winter months, raising concerns that it could be exaggerating winter baseline ramps compared to winter actual loads.

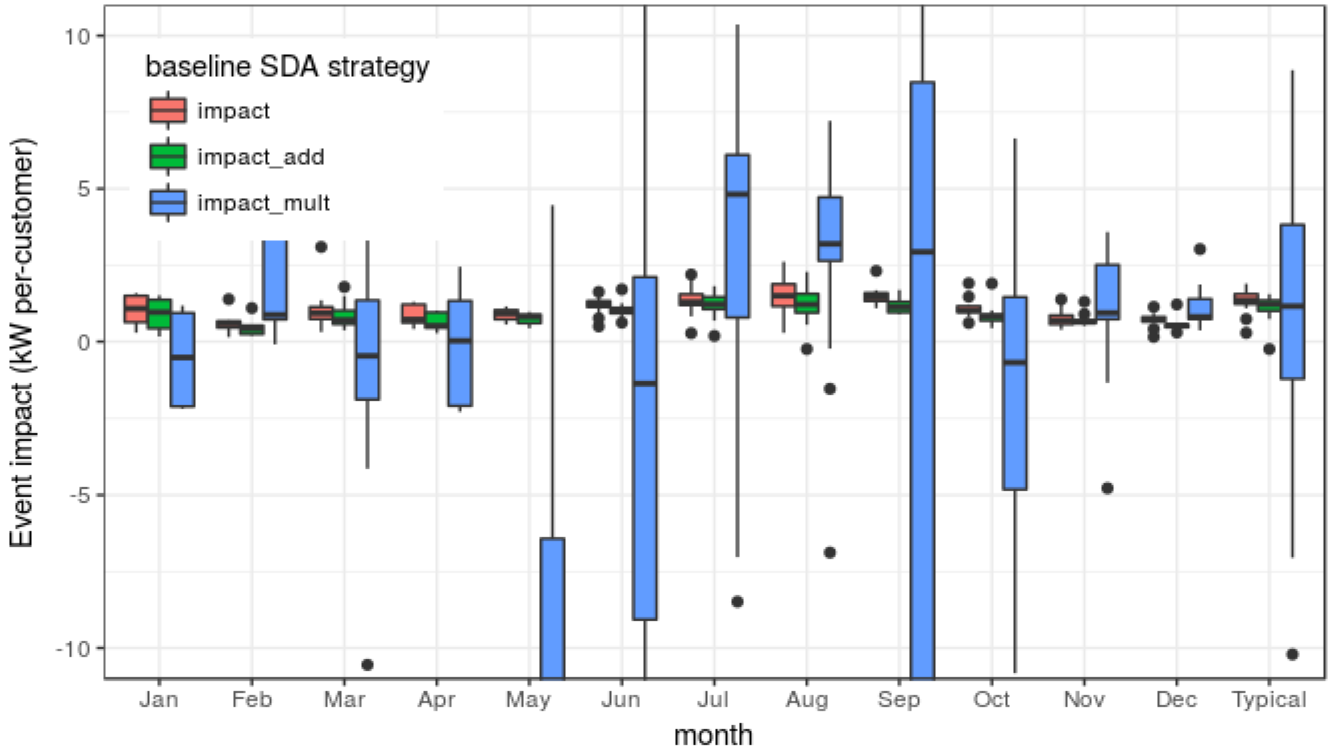
The general correspondence between all three SDA strategies suggests that they are all extracting similar information from the underlying signal and appears to lower the stakes of the choice. However, it is important to keep in mind that neither "add" nor "mult" match actual loads well during non-event hours and would create spurious morning and night "impacts" during those hours if not further constrained.

Figure 16. Comparison of monthly roll-up events for each of the considered same day adjustment strategies



It is also important to note that the behavior of the "mult" SDA is non-linear near zero load (i.e. the ratio that it is based on could involve dividing by numbers close to zero). Figure 17 presents a comparison of the same methods applied this time to customers known to have PV systems. It is clear that "mult" is completely inappropriate for such situations and is producing highly erratic impact estimates.

Figure 17. Event impacts for each SDA for customers with PV



Estimating errors

To quantify the error of impact estimates, CDA computed the underlying variability of actual event and baseline loads that would occur when taking random sub-samples of the participants. These quantities are known as the standard error in the hourly means of the loads. The resulting "jitter" in the outcomes they characterize can be understood as uncertainty in impacts caused by the distribution of event performance across customers.

For each hour of the day the standard error of the impact is computed as follows,

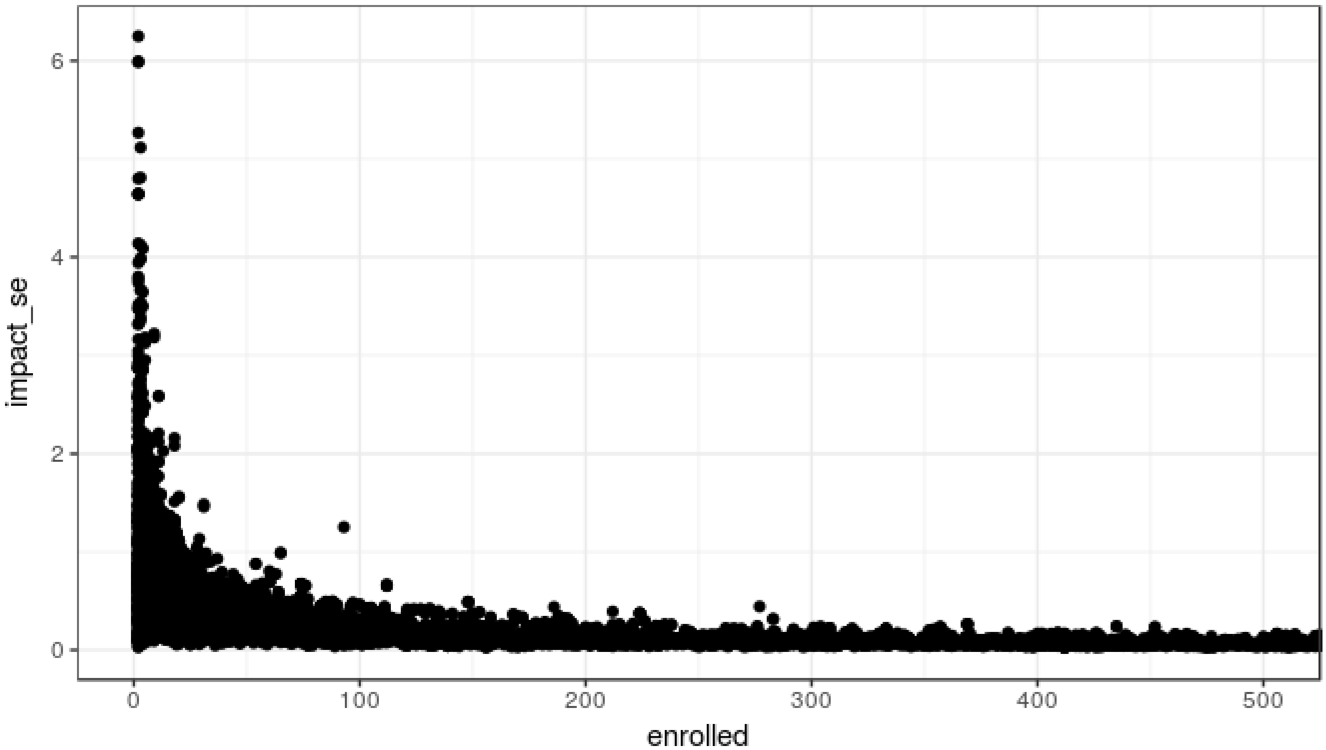
```
impact_se = sqrt( (sd_baseline_kwh^2)/n + (sd_actual_kwh^2)/n )
```

where n is the number of customers participating in the event, sd_baseline_kwh is the standard deviation of baseline loads across participants and sd_actual_kwh is the standard deviation of actual event loads across participants.

Validity assessment of the study findings

Figure 18 illustrates the relationship between the event participant count and the estimated impact standard errors of event hours. The data comes from results from all participants and all permutations of sub-groups participating in each event (i.e. multiple points per-event). For small groups of participants, errors can grow quite large, but standard errors have largely settled into a small range by the time we are looking at events with 100 participants, which is the cutoff for reporting out in the ex post table generator.

Figure 18. Impact standard error from event hours vs. count of event participants.



For events with at least 100 participants, Table 7 characterizes the percentage of all error estimates from event hours that meet a series of criteria. For example, just under 3% of standard errors are over 250W. In less than ½ of 1% of cases, the standard errors are greater than the impact, and in just over 2% of the cases the standard errors are greater than half of the impacts, which was our stated goal from the evaluation plan.

Table 7. Characteristics of impact standard errors

criteria	% of event hrs
se > 50w	59.15
se > 100w	25.73
se > 250w	2.97
se > impact	0.49
se > impact/2	2.13

Ex Ante Load Impact Analysis Methods

As opposed to 'ex post' analyses, which quantify what has happened in the past, 'ex ante' forecasts attempt to quantify the future. Ex ante load impact forecasts predict the load impact that would occur in standard event times and conditions for a total of four hypothetical weather-years. For this report, forecasts are made for the peak day of each month for two standard weather years – corresponding to conditions that are expected to lead to peak electric load in either one out of every two years, or one out of every ten years on average. There is a slight difference in the conditions that cause peak load for California Independent System Operator (CAISO) and for Investor-Owned Utilities (IOU). Combining the two peaking conditions with the two types of weather year leads to four sets of standard weather data.

Unless explicitly stated, all forecasts discussed in this report are made for events in the Resource Adequacy (RA) time window that runs from 4 p.m. through 9 p.m.

Ex ante predictions are made by statistical models that predict the load impact per participant for different categories of customers (such as customers with and without electric devices that are automatically controlled to reduce load during a demand response event). The models were fit to data from calendar year 2018 and used to predict the load impact per participant for the standard weather conditions mentioned above, for the projected mix of customers by category that is forecast for future years. The predicted load impact per participant multiplied by the projected number of participants equals the projected aggregate load impact.

In addition to being used to predict the load for standard weather conditions, the ex ante models can be used to predict the load shed in other conditions. In this report we discuss fitting models to weather and enrollment data from 2018 and using them to predict the load impact for 2019 events through August, which allows us to check the performance of the models.

A major product of the ex ante analysis is a 'table generator' – a spreadsheet – that allows the user to select a forecast year and a set of weather conditions and see estimates of load impact by hour, in each sub-LAP separately or the total of all sub-LAPs.

Ex Ante Methods

Given the ex post input data, there are five major steps to producing the ex ante forecasts. These are outlined below, followed by technical details.

1. Fit statistical models that predict the ex post load impact for each event participant as a function of available explanatory variables.
    a. Customers are divided into categories based on (a) whether they have a device (a communicating thermostat or smart plug) that OhmConnect directly controls to trigger an event, (b) the performance tier OhmConnect has assigned to each customer, and (c) whether they have solar generation. A different model is fit for each customer category, for both event hours and non-event hours, and separate models are fit for 'summer' months (May through October) and 'non-summer' (November through April).
    b. Models were fit only to data from events in the 4 p.m – 9 p.m. RA window.
    c. Customer categories and model variables, were chosen based on:
        i. Substantive knowledge (e.g. air conditioning and heating loads are affected by outdoor air temperature);

ii. Data availability (e.g. there were few hour 17 ex post events, so these events were lumped with hour 18 events for analysis, using an indicator variable described below);
iii. Goodness-of-fit (e.g. an indicator variable was added for the summer shoulder months of May and October because without this variable the load impact in those months tended to be over-predicted compared to ex post results).

d. Explanatory variables for the event-hour models were:
   i. Indicator variable for the sub-LAP.
   ii. Indicator variable for whether the sub-LAP routinely experiences summer temperatures over 90 F (and is therefore expected to have a large fraction of customers with air conditioning). This was used in conjunction with the cooling-degrees variable (discussed below) to allow the relationship between load impact and temperature to differ between these sub-LAPs and cooler sub-LAPs.
   iii. Indicator variable for whether the event is an OhmConnect 'mega' event;
   iv. Indicator variable for whether the hour is 17 or 18 (i.e. from 4 – 6 p.m.);
   v. Indicator variable for whether the hour is 21 (from 8-9 p.m.);
   vi. Indicator variable for whether the month is a summer vacation month (July or August);
   vii. Indicator variable for whether the month is a summer 'shoulder' month (May or October);
   viii. A heating- or cooling-degree variable: In summer this is temp75 = max(0, temperature – 75 F), in winter it is Tunder60 = min(0, 60 F – temperature).

e. Conceptually similar models (slightly simpler) were fit to data from the non-event hours, as well as the baseline data.
f. All models were linear regression models.

2. Obtain a forecast of customer counts, by customer category, in future years, as well as forecasts of how many customers will participate in the standard ex ante events.
   a. OhmConnect provided a forecast of the total number of customers they anticipate in each future month, by sub-LAP and tier, as well as the total number of customers with controllable devices in each sub-LAP. These numbers were mapped to the four customer categories by apportioning devices among categories within a sublap and tier to match the proportions observed in 2018.
3. Obtain the standard weather conditions for which to make forecasts. These are the standard CAISO and IOU 1-in-10 and 1-in-2 weather years.
4. Apply the statistical models to predict the load impact per participant, in each customer category, for the forecast customer counts and weather conditions.
5. Aggregate the predictions across customer categories to generate predictions of the program-wide load impact per participant; multiply by the forecasted number of participants to generate the predicted aggregate load impact; and divide the aggregate load impact by the total number of OhmConnect customers to determine the predicted load impact per customer.

As discussed earlier in this document, all of the modeling above was performed only for customers included by OhmConnect in events, i.e. whose event-day load exceeded their 10-of-10 baseline in the hour prior to the OhmConnect event, and who did not opt-out of the event. In the summer months this applied to about half of their customers; in other months the fraction was lower. In the ex ante predictions we assume that the other customers – the majority – contribute no load impact at all. In each month, the number of eligible OhmConnect customers is multiplied by the empirical "inclusion fraction" who meet the baseline-exceedance criterion in the relevant season, and then multiplied by the fraction of the remaining customers who are expected to participate in the event (based on empirical opt-out data from 2018, tabulated by customer category). The resulting number of forecasted participants is multiplied by the load impact per participant from the model discussed above.

OhmConnect provided data on each customer whose actual load exceeded their 10-of-10 baseline load in the hour preceding the event, for all hours of every day the customer met that criterion and was eligible to participate in an OhmConnect event. Unless otherwise specified, all discussion below refers to the subset of OhmConnect customers who met this baseline exceedance criterion.

Customers were divided into categories defined by a combination of: sub-LAP; whether they have solar generation; whether they have given OhmConnect direct control of one or more electrical devices (such as a thermostat or switch); and what 'tier' they have been assigned. (Tier is based on OhmConnect's assessment of how well the customer performs in the context of the program). All customers within a given category and who were eligible for a particular Demand Response event are grouped, and summary statistics are calculated for the group for that event, along with some event-specific information. Table 8, below, shows a small portion of the available information.

"n" is the number of customer-hours in the category described by that row, and the outdoor temperature is the participant-weighted mean temperature experienced by all of the customers described by that row, as determined from the weather station closest to each customer.

About 4% of the temperatures are unknown due to missing weather station data, but these are all from rows with low n, so in all they represent less than 0.1% of all customer-hours. For these cases we imputed the temperature to be the mean temperature of all customers in the sub-LAP for which a temperature was available.

The resulting data file for 2018 has information on 726 OhmConnect events that included at least one hour between 4 p.m. and 9 p.m., further broken down by customer attributes (sub-LAP, whether the customer provides direct load control, etc.), totaling over 18,000 rows of ex post data input into the ex ante models.

Table 8. Example of aggregated data that are inputs to the ex ante models.

Event ID	sub-LAP	Date	Event hours	Outdoor temp (F)	Direct load control?	Solar?	Participating?	n	Load Impact per Participant (kW)
712	PGCC	2/15/18	20-20	50.6	No	No	Yes	215	1.10
712	PGCC	2/5/18	20-20	50.6	Yes	No	Yes	71	1.42
712	PGCC	2/5/18	20-20	50.6	No	Yes	Yes	9	0.74

A customer category is a combination of (Tier, Direct Load Control, Solar), yielding 3 x 2 x 2 = 12 different customer categories. Furthermore, a different model is fit for summer and for winter, in each category. Thus there are 24 combinations of (season x customer category). Within each of these customer categories there are separate models to predict the load impact during event hours, the load impact during non-event hours, and the baseline load.

Each of the models is a linear regression model that predicts load impact from the explanatory variables, including both 'indicator variables' (dummy variables) such as sub-LAP and continuous variables such as the number of degrees (F) above 75 F. See the ex ante appendix for details. Predictions are made only for 'participants', i.e. customers who did not opt out of the event, are thought to have received an event notification, and are still enrolled in OhmConnect's program (customers who de-enroll are marked as non-participants until their de-enrollment propagates through OhmConnect's weekly resource allocation process).

Ex Ante Weather

IOU ex-ante weather data.

The evaluation team received monthly-peak and typical event day weather data from three IOUs - PG&E, SDG&E and SCE. The data included hourly temperature data for weather stations spanning the IOU territory for both the IOU and CAISO monthly peak demand days under 1-in-2 and 1-in-10 weather conditions and under typical event conditions*. In addition to the weather data files, SCE and SDG&E also provided a file that mapped each weather station to one or mode zip codes.

(* See *Statewide Demand Response Ex Ante Weather Conditions*, *Prepared for San Diego Gas and Electric Co., Pacific Gas and Electric Co. and Southern California Edison Co. January 30, 2015, Nexant, Inc.*)

Summary of data received from each of the IOUs:

IOU	Weather Stations	Temperature Records
PG&E	25	60,000
SCE	25	60,000
SDG&E	10	30,000

Mapping ex-ante weather to OhmConnect customers

The IOU ex ante weather data was specific to the weather geography relevant to each utility. To use the weather data for OhmConnect's ex ante forecast, each customer had to be assigned to a weather station/geography. Lacking a weather station-to-zip-code map for the PG&E territory, OhmConnect used the weather station's approximate latitude and longitude to associate the nearest weather station to each OhmConnect customer's known location.

For SCE and SDG&E, OhmConnect mapped the weather station level data to sub-LAP-level data using the following steps:

1. For SCE and SDG&E: Associate a weather station to an OhmConnect customer by mapping the customer's zip code to the associated station in the data provided by the IOU.
   For PG&E: Associate a weather station to an OhmConnect customer based on the shortest straight line distance between the customer's service address and the approximate coordinates of the weather station.
2. Count the number of enrolled customers in each sub-LAP served by each associated weather station.
3. For each sub-LAP, calculate a percentage weight for each weather station based on the percentage of customers in the sub-LAP associated with that station.

4. Calculate a weighted-average of the temperature for a given month and hour at the sub-LAP-level using the weights calculated in step 3.

The following table illustrates how weather station-level temperature is normalized to a sub-LAP-level temperature. The data used in the example is based on customers in the SCHD sub-LAP (within SCE territory), for month 8, hour 20, weather-year 1-in-2 and a day type of IOU MONTHLY SYSTEM PEAK DAY.

Table 9. Example of sub-LAP-level weather calculation.

Weather Station	sub-LAP	Temperature (t)	Customers (n)	Percentage of Customers (w)	Weighted Temperature (T = w*t)
101	SCHD	73.00	45	45/848 = 5.31%	73.00*5.31% = 3.87
141	SCHD	92.72	32	32/848 = 3.77%	92.72*3.77% = 3.50
191	SCHD	99.90	134	134/848 = 15.80%	99.90*15.80%=15.79
192	SCHD	99.63	296	296/848 = 34.91%	99.63*34.91% = 34.77
194	SCHD	92.88	339	339/848 = 39.98%	92.88*39.98% = 37.13
195	SCHD	91.75	2	2/848 = 0.24%	91.75*0.24* = 0.22
sub-LAP-level weighted temperature:3.87+3.50+15.79+34.77+37.13+0.22 =					**95.28**

Adjusting for Incomplete Participation

Inclusion Fraction

As described earlier, OhmConnect excludes customers whose pre-event load is below their 10-of-10 baseline load from events. As a conservative estimate of their program in the future, we assume any customer whose pre-event load is below their 10-of-10 baseline load, if included, would contribute no load impact. The fraction of customers who do meet this baseline exceedance criterion is called the 'inclusion fraction'. The ex ante participation forecasts require a prediction of what the inclusion fraction would be for each of the standard ex ante event days. For this we used the average empirical inclusion fraction over years 2016-2018, by season, as tabulated in Table 10. Empirical Inclusion Fraction from 2016 through 2018. For example, for all future weather years we assume the inclusion fraction is 0.40 (40%) in March, April, and May. Below we will explain how this number is used in the ex ante predictions.

Table 10. Empirical Inclusion Fraction from 2016 through 2018.

season	month	Enrolled customers	participants	Inclusion fraction	Seasonal inclusion fraction
Fall	Sep	320533	120918	0.38	0.39
Fall	Oct	442110	171762	0.39	0.39
Fall	Nov	492064	191706	0.39	0.39
Spring	Mar	393369	144246	0.37	0.40
Spring	Apr	322832	135728	0.42	0.40
Spring	May	637205	268084	0.42	0.40
summer	Jun	986673	519951	0.53	0.48
summer	Jul	938069	507132	0.54	0.48
summer	Aug	1325440	537647	0.41	0.48
winter	Jan	255197	90336	0.35	0.39
winter	Feb	548004	217797	0.40	0.39
winter	Dec	565456	227216	0.40	0.39

Participation Fraction

Roughly 25% of the remaining customers – those who satisfy the baseline exceedance criterion – are known not to participate in the event for various reasons. Most of these have unsubscribed from OhmConnect's event notifications; some of them have recently de-enrolled from OhmConnect's events at the time of the event but this change has not yet propagated through the system; and a small fraction have used OhmConnect's website to opt out of the event.

As shown in Table 11, In 2018 the participation fraction was very high for customers who have a device that OhmConnect can directly control, and is lower for other customers. Customers whose device status is unknown (mostly from the early part of 2018) have a participation fraction that is in between.

The ex ante models were fit to data from customers who participated in the events, and thus predict the load shed per *participant.* To convert those numbers to the load shed per *customer*, the load shed per participant is multiplied by the participation fraction, i.e. the number of participants per customer: (load shed/customer) = (load shed per participant) x (participants / included customer) x (included customers/total enrollment)

Table 11. Participation as a fraction of the number of customers, by tier and device category, in 2018.

Tier	Has Devices In Event?	Participation Fraction
Silver	No	0.70
Silver	Yes	0.96
Silver	Unknown	0.82
Gold	No	0.86
Gold	Yes	0.96
Gold	Unknown	0.90
Platinum	No	0.87
Platinum	Yes	0.98
Platinum	Unknown	0.92

How Inclusion Fraction and Participation Fraction are Used

The ex ante aggregate load impact for a given customer category in a given month is given by:
Aggregate Load Impact = (OhmConnect total customers in the category) x (seasonal inclusion fraction) x (participation fraction for the tier) x (predicted load impact per customer in the category)

Ex Post Results

Monthly roll-up of events

CDA identified and estimated ex post impacts for 726 discrete OhmConnect DR events in 2018, with the permutations of customer attributes and locations bringing the total number of discrete event day evaluations to 8,279. There are practical reporting issues raised by having such a large number of events and event results to report on. To address this CDA performed "monthly roll-ups" on events to produce representative monthly summaries of all event activity. For all events called for the same time of day during the same month in the same utility territory and sharing the same set of customer attributes, the monthly roll-up computes the average number of participants, the total counts of unique participants, the total count of all participants (i.e. with users counted more than once if they are in more than one event), and the participant-count-weighted average reference loads, impacts, and outside temperatures, with associated errors and aggregate values. The month of June, for example, had 114 events, which were summarized by 18 monthly events. This is because there were 6 different time ranges of events called in the PG&E territory, 7 different time ranges in SCE's territory, and 5 in SDGE's territory. These "monthly events" summarize all the events that occurred in a more manageable format and are the data products that are fed into the ex post table generator.

As discussed in the "Coincident impacts and loads" section of the introduction to the OhmConnect resource, we can optionally report "coincident reference loads" and "coincident impact" for rolled up events, where those are the aggregate values that are calculated by multiplying the count of the unique participants in a month by that month's weighted average per-participant impact.

Table 12 below takes an additional roll-up pass on the monthly events (using the same calculation logic), producing a single representative rolled up ex post summary event for each month of the year. In other words, the events that were separated by utility and event timing have been aggregated to single monthly representative events.

Table 12. Summary of 2018 events, conditions, enrollment, and impacts.

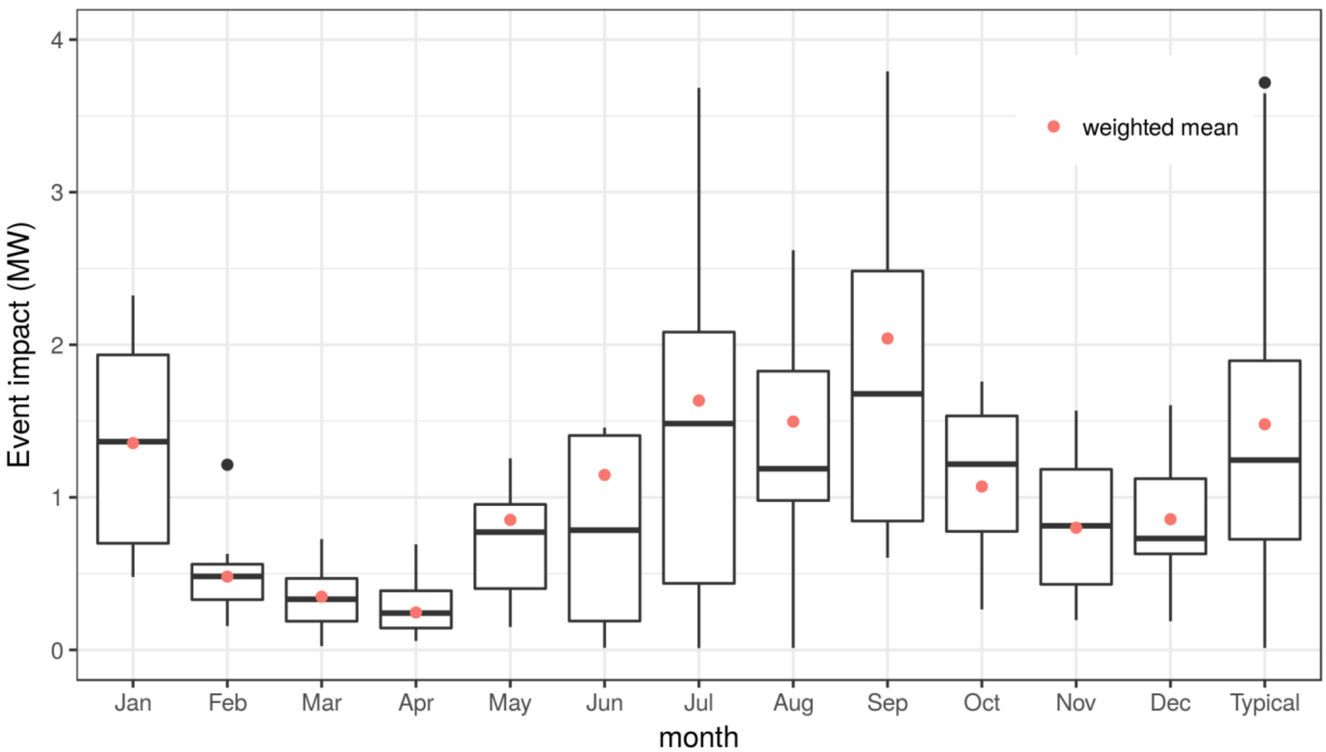
month	# of events	average # of participants	# monthly unique participants	total # participants	average temperature (F)	reference load (kW)	per-participant impact (kW)	aggregate reference load (MW)	aggregate impact (MW)	impact %	coincident reference load (MW)	coincident aggregate impact (MW)
Jan	12	**4,362.92**	32,135	52,355	60.17	1.04	0.31	4.56	1.36	29.74	33.57	9.98
Feb	47	1,524.40	39,999	71,647	54.56	1.13	0.32	1.73	0.48	27.78	45.39	12.61
Mar	40	952.15	26,887	38,086	58.84	1.13	0.37	1.08	0.35	32.32	30.47	9.85
Apr	43	676.02	21,916	29,069	60.99	1.17	0.36	0.79	0.25	30.97	25.69	7.96
May	62	1,948.27	51,288	120,793	64.75	1.30	0.44	2.53	0.85	33.73	66.50	22.43
Jun	**114**	2,109.18	61,744	240,697	73.94	1.68	0.54	3.55	1.15	32.29	103.99	33.58
Jul	107	2,664.41	65,906	**285,392**	**81.55**	**2.17**	0.61	**5.77**	1.63	28.31	**142.80**	40.42
Aug	119	2,250.86	**70,984**	268,075	78.30	1.89	**0.66**	4.26	1.50	35.10	134.46	**47.19**
Sep	34	3,090.88	60,644	105,090	76.50	1.87	**0.66**	5.77	**2.04**	35.37	113.22	40.05
Oct	41	2,259.44	57,633	92,637	67.50	1.29	0.47	2.91	1.07	**36.75**	74.32	27.32
Nov	56	2,214.48	65,282	124,011	60.67	1.18	0.36	2.60	0.80	30.77	76.73	23.61
Dec	51	2,506.61	65,015	127,837	52.81	1.32	0.34	3.30	0.86	25.93	85.71	22.22

For this summary one can observe the number of actual events called in each month, with more activity during the summer than winter. The majority of months have average event participation counts around 2,000 customers, and 8 of the 12 months have total unique participant counts between 50,000 and 70,000 customers. Per-participant reference loads are strongly correlated with outside temperatures, and peak in the hottest months of the year. Per-participant impacts roughly track reference loads, and also peak in the summer. Note that impact as a percentage of reference loads does not have a strong seasonal pattern. Event activity is lower during winter months, but the set of fall and winter months toward the end of the year shared similar average participant counts and unique monthly participants as the summer months. Thus the coincident aggregate loads and impacts mainly correlate with their average per-participant analogs.

Aggregate Ex Post Summary

Figure 19 depicts the monthly and typical (typical is the average for June-September) monthly roll-up event aggregate impacts. The ex post aggregate load impact on a typical 2018 event day was 1.48 MW, with the largest aggregate load shed (2.04 MW) occurring during the month of September, the month with the highest average event participation.

Figure 19. 2018 event average ex post aggregate load shed (overall, by month)



As discussed previously, monthly coincident impact (unique participants multiplied into average per-participant impact) is a better metric of resource potential than per-event aggregate performance, so Figure 20 presents the monthly coincident aggregate impacts. The monthly coincident impacts peak in August, which tied with September for the highest per-participant impacts, but called the highest monthly total of unique customers.

Figure 20. 2018 monthly coincident impact ex post load shed (overall, by month)

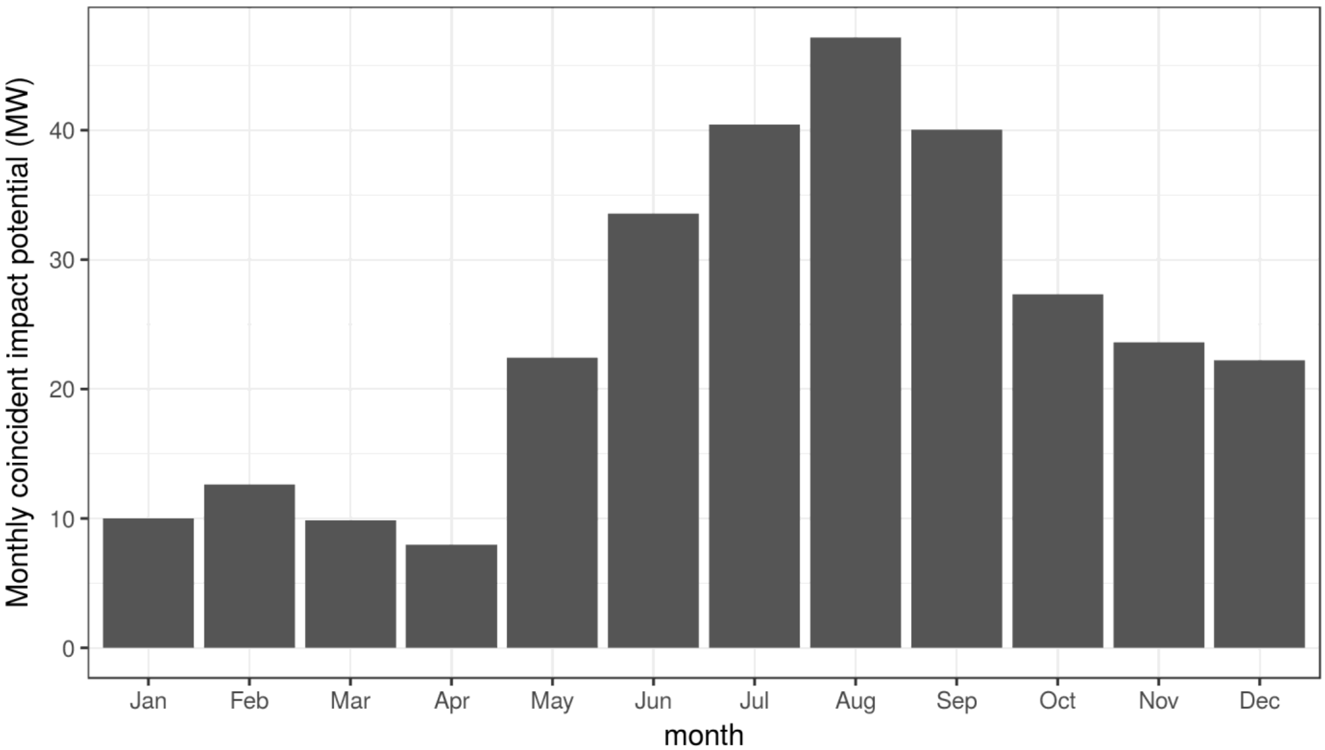


Figure 21 presents the monthly per-event aggregate impacts again, but this time disaggregated by customer tier. Silver customers dramatically outnumber Gold and Platinum customers, so they dominate aggregate impacts despite low per-participant impact. Note that in January, the average event participant count was significantly higher than other months. See the determinants of aggregate results section for more information on this.

Figure 21. Event average aggregate load impacts by customer sub-groups

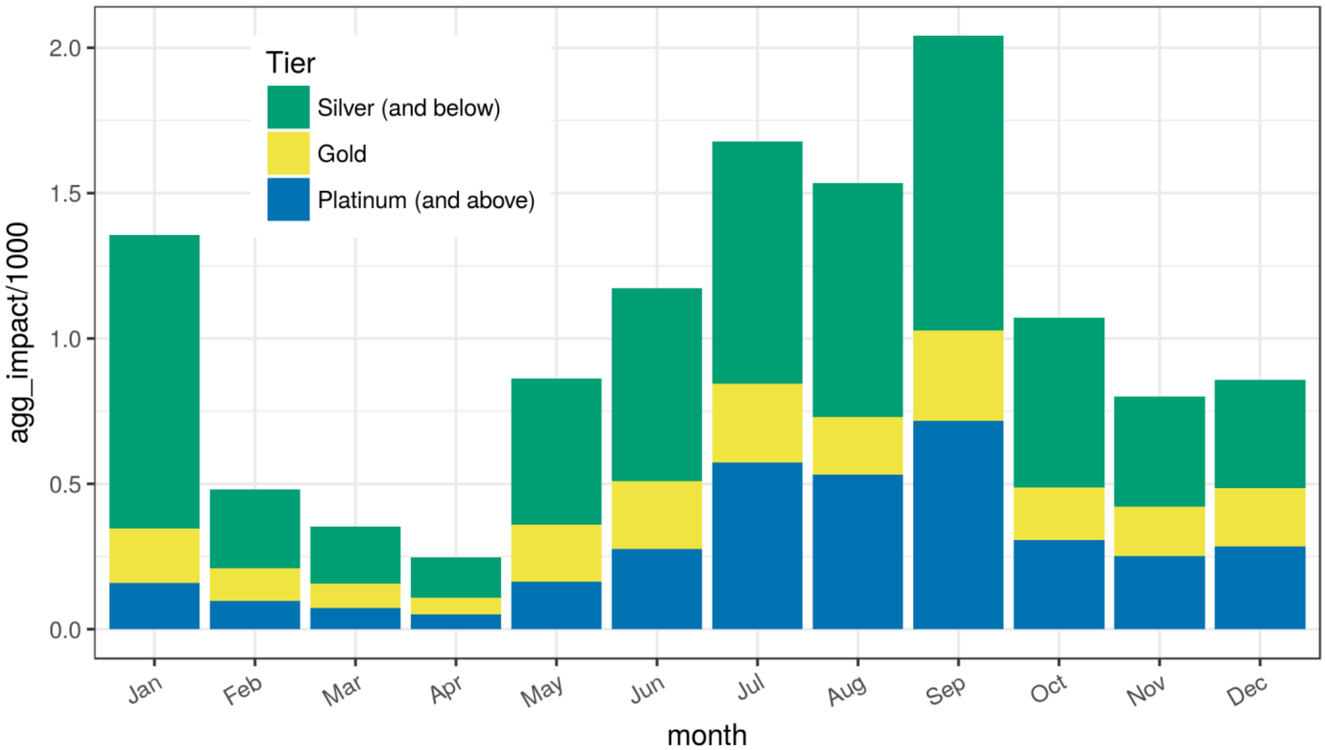
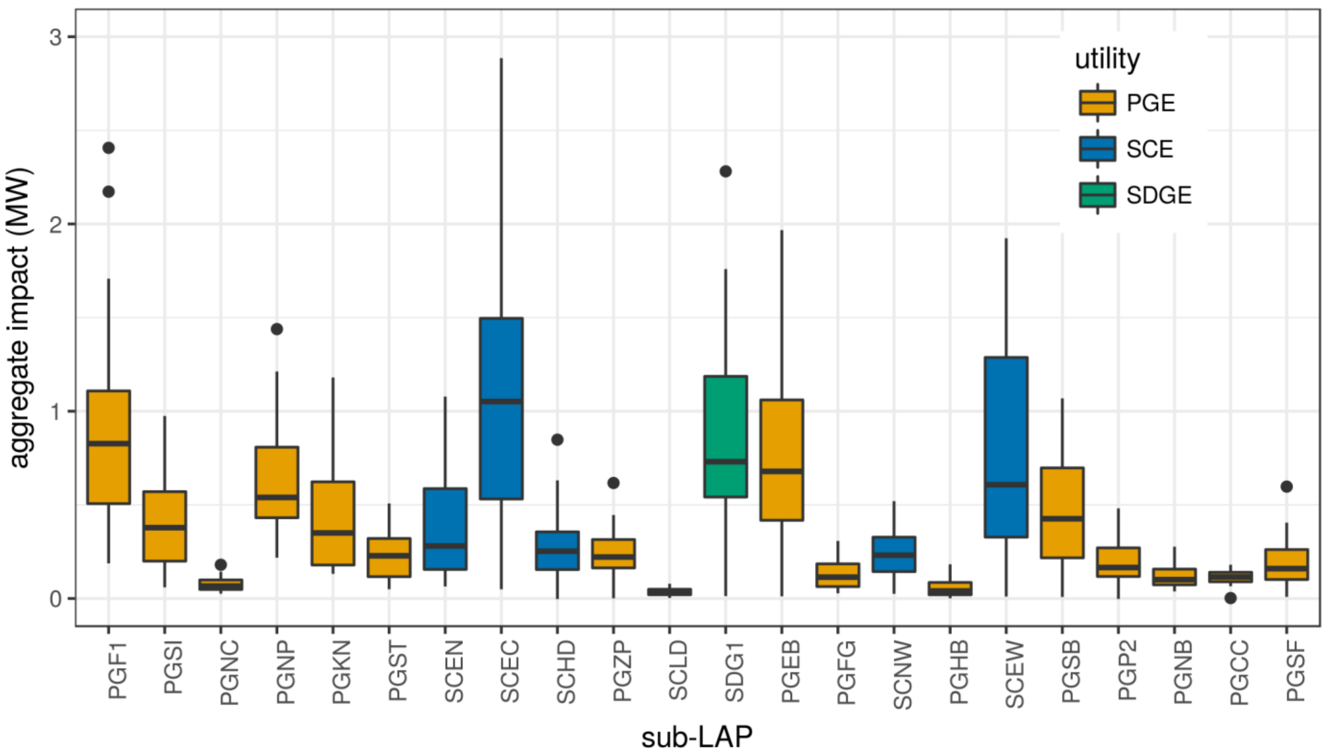


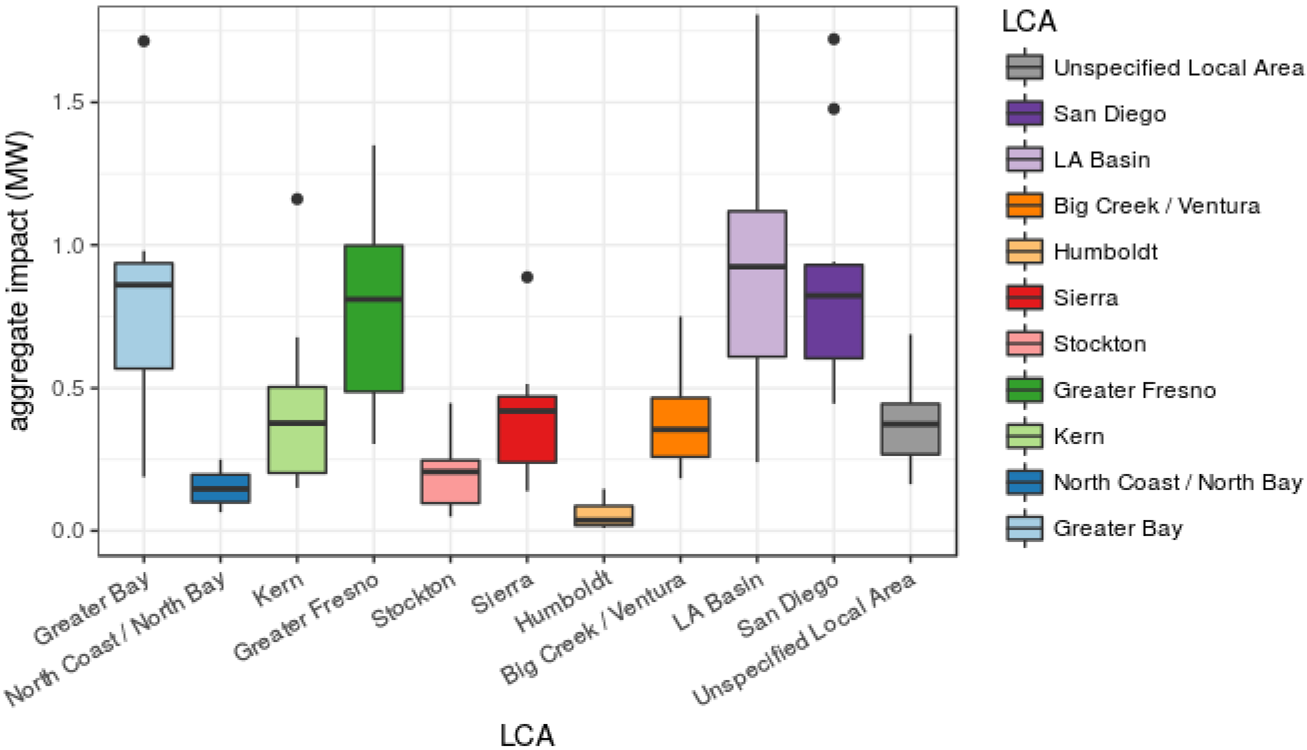
Figure 22 and Figure 23 present box and whisker summaries of aggregate impacts for all events grouped by sub-LAP and LCA respectively. The sub-LAPs are ordered for highest to lowest average impacts and colored by the utility whose territory they fall within. The four sub-LAPs with the greatest aggregate impacts are also the top four sub-LAPs by enrollment.

Figure 22. Aggregate impact across all events grouped by sub-LAP and colored by utility. The order from left to right is in descending average per-participant impacts.



Among the LCAs, the Greater Bay and LA Basin have the top two counts participating customers which overcomes low per-participant impacts placing them toward the top of aggregate impacts. By comparison, Fresno has less than 1/3 the count of participating customers than the Greater Bay does, yet its high per-participant impacts deliver comparable aggregate impacts to the most populous ones.

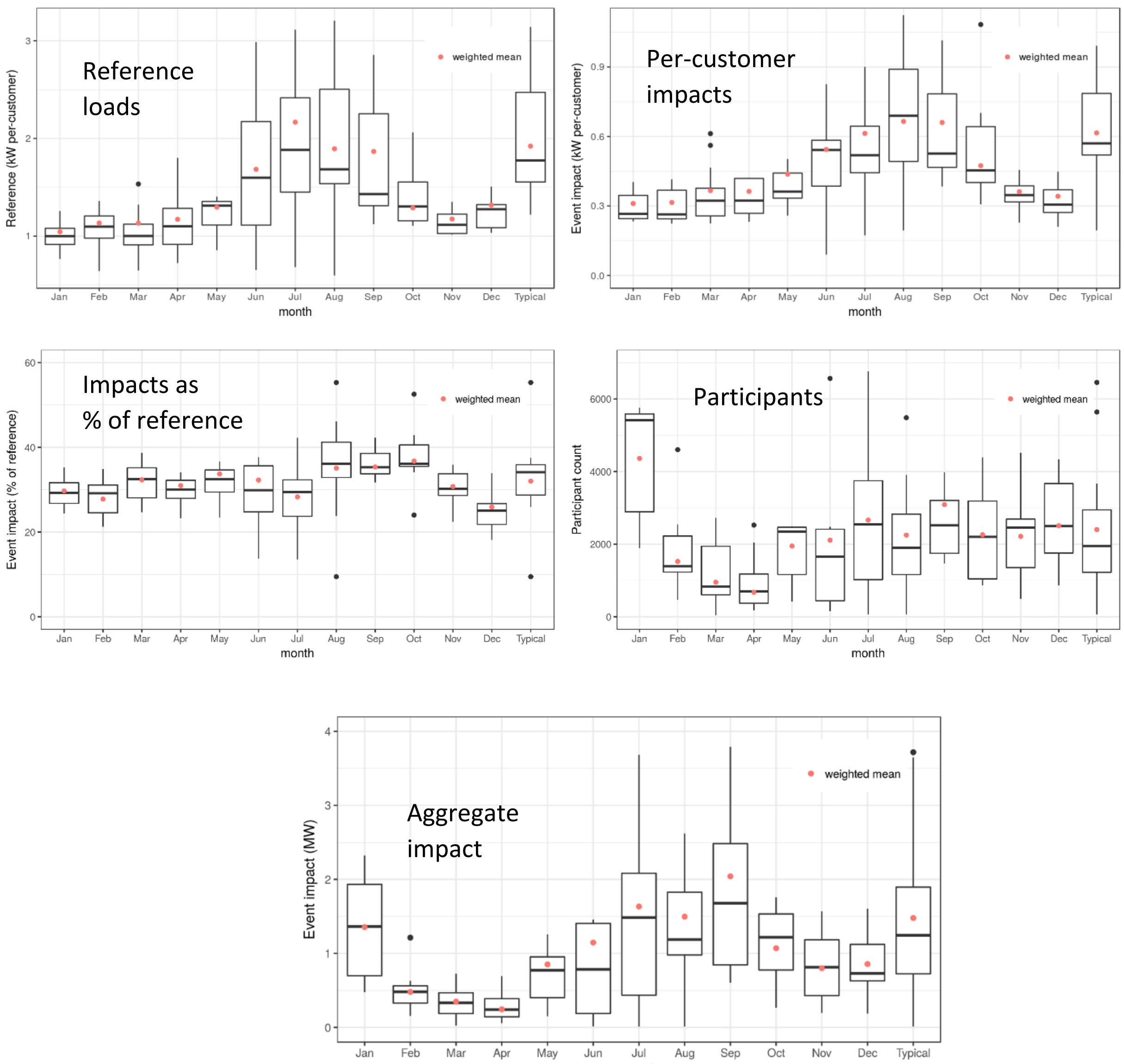
Figure 23. Aggregate impact across all event grouped and colored by LCA.



Determinants of aggregate results

Aggregate impacts are population weighted per-participant impacts. Per-participant impacts are a function of reference loads. Reference loads tend to be driven by seasonal factors. Figure 24 presents monthly summaries of per-participant reference loads, per-participant impacts, and average event enrollment, followed by the aggregate impacts. For this comparison, it can be verified that reference loads are highly seasonal and the per-participant impacts track reference loads and are therefore seasonal as well. A look at the impacts expressed as a percentage of reference loads reveals a fairly stable fraction of reference loads being saved throughout the year. The month with the greatest impact as a % of reference loads is actually October, which has undistinguished per-participant impacts. Finally, the enrollment per event numbers reveal a flat or gradually increasing trend throughout the year, with anomalously high event sizes in January. Each of these determinants leaves its imprint on the final aggregate results.

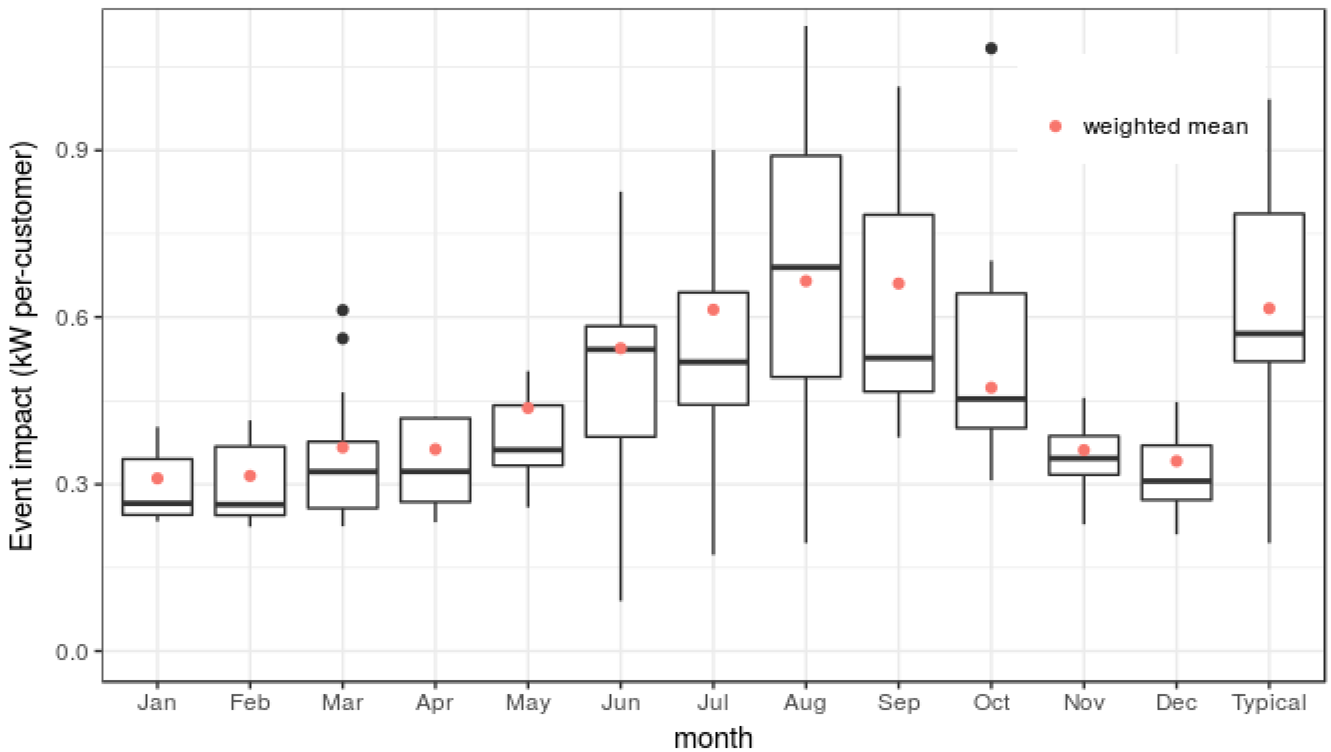
Figure 24. Monthly average reference loads, per-participant impacts, impacts as a percentage of reference loads, and enrollment counts for each month and typical (June-September) months, including the weighted mean for each.



Average Ex Post Load Impacts by Event

Figure 25 summarizes the monthly summary events per-participant impacts for all events called in 2018. There is a clear seasonal pattern to the results. The highest monthly weighted average per-participant impact is 0.66 kW during both August and September. The typical values is 0.62 kW. January and February come in the lowest at 0.31 kW and 0.32 kW respectively.

Figure 25. Monthly and typical per-participant average event impacts.



The figures in the next section explore the determinants of per-participant event performance more deeply. Figure 26 depicts the average impacts per-participant (y-axis) for every event called in 2018 vs. the population weighted outside temperature during the events (x-axis), with dot sizes roughly corresponding to participant counts. The figure gives a visual sense of the number of events called overall and confirms a rough correlation between impacts and outside temperature, with plenty of variability caused by other factors. It is unusual to see DR events from cold weather conditions and there is a possible hint of the impacts increasing gradually at the lowest temperatures (OhmConnect customers may be shedding heating loads in the winter). Dot sizes suggest that the hot weather/summer period include the events with the greatest participation.

Figure 26. Mean 2018 ex post event load shed (overall, by date and LCA)

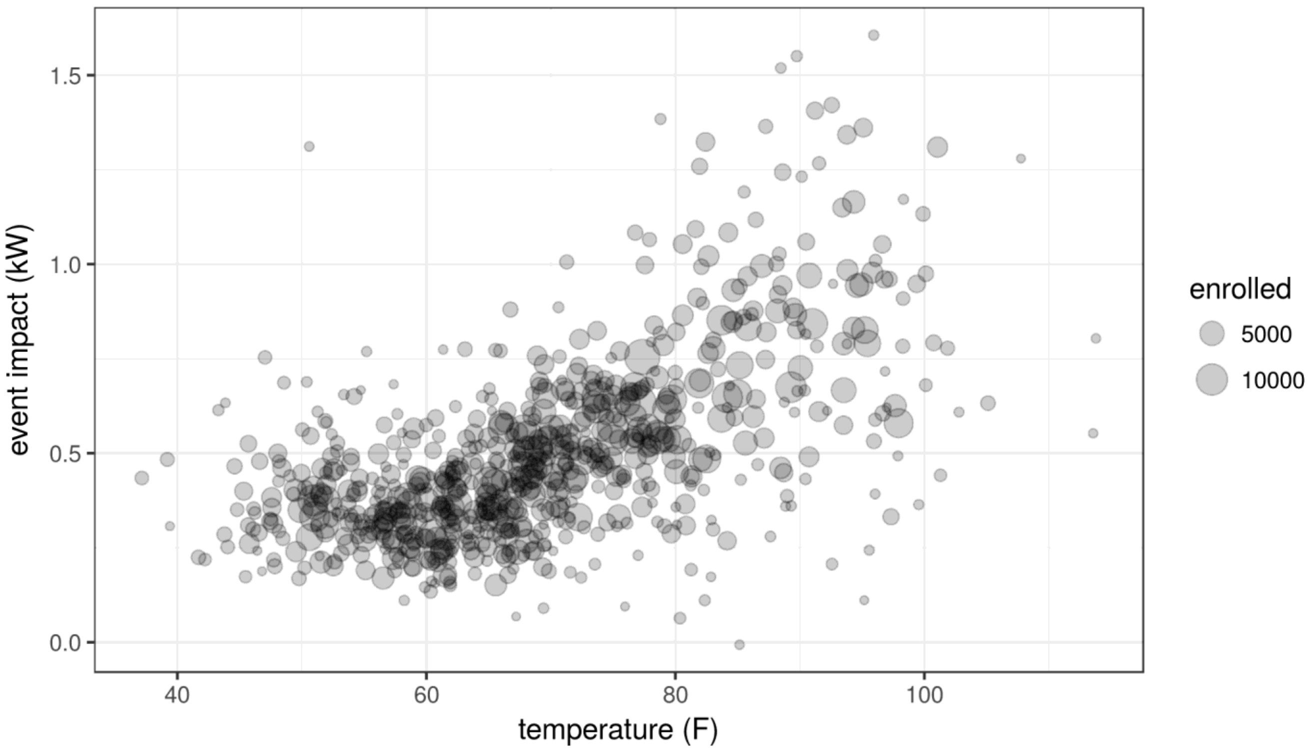
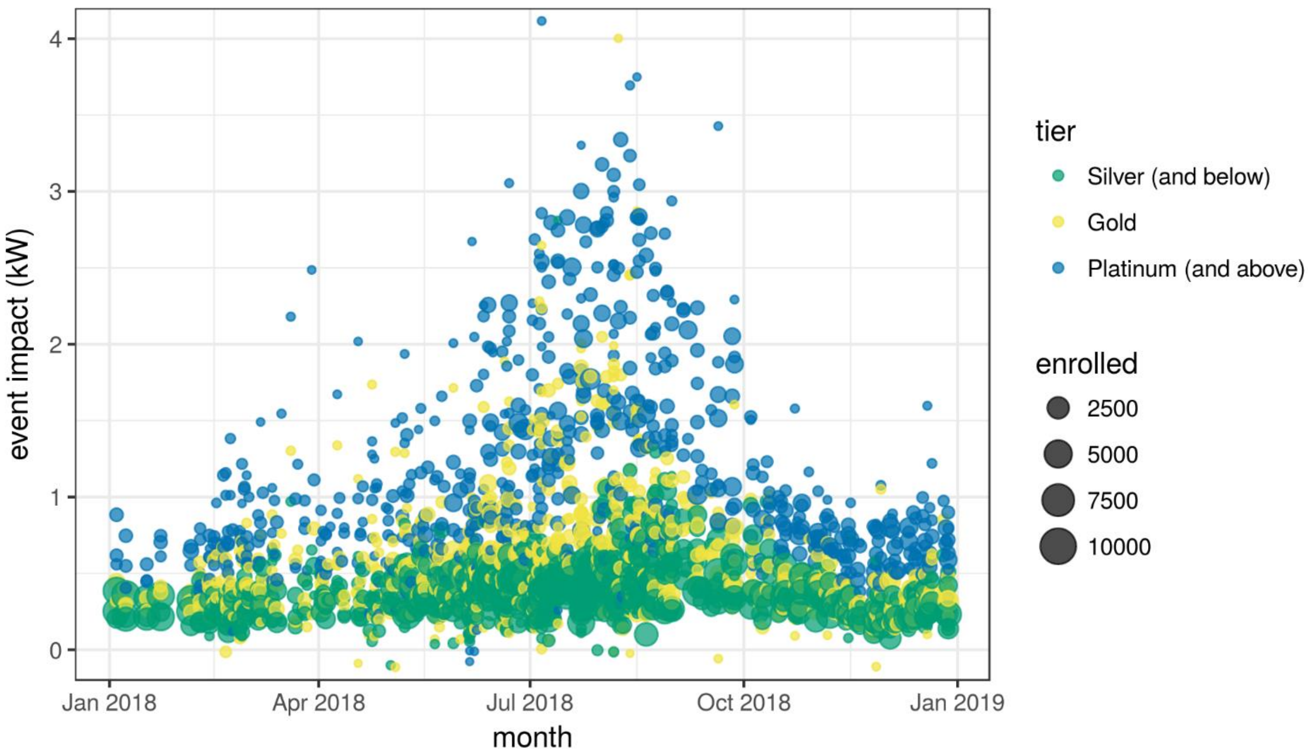


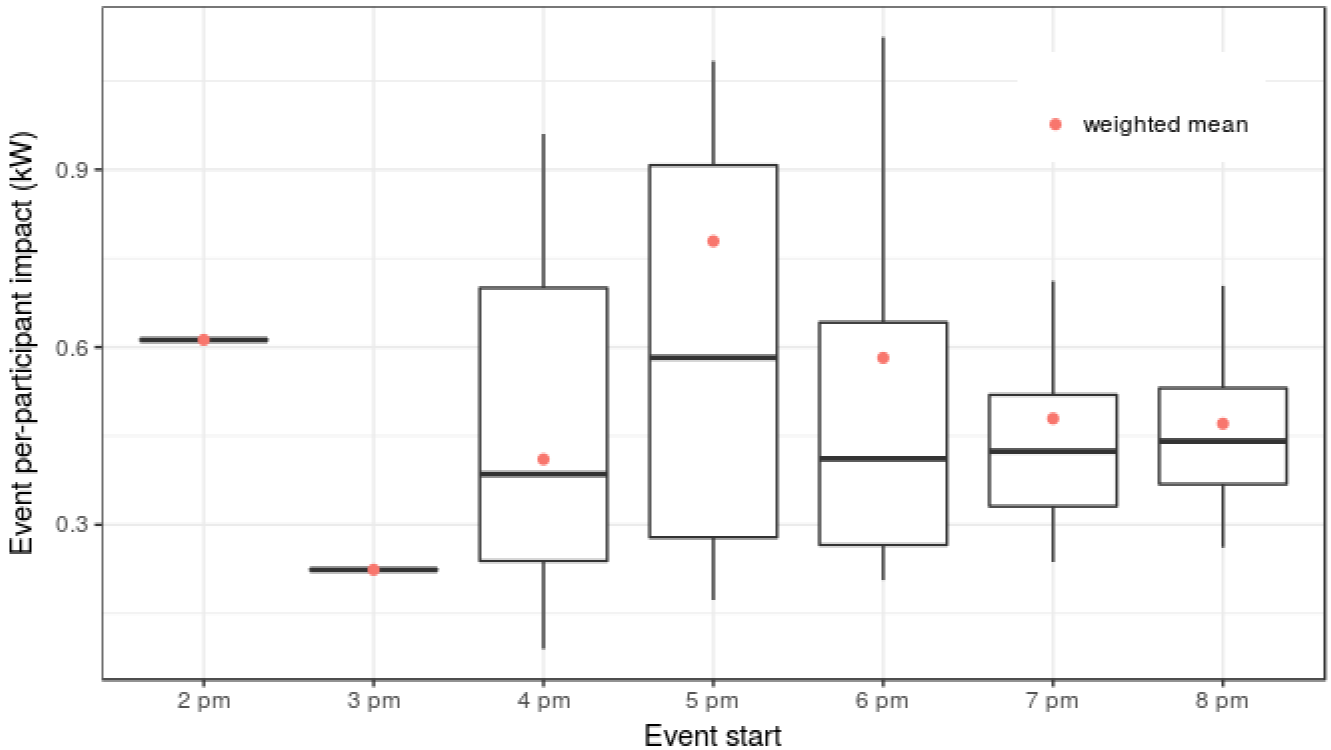
Figure 27 depicts average per-participant impacts (y-axis) for all the events called in 2018 as a time series (x-axis), with dot colors corresponding to customer tier and sizes related to participation numbers. This plot underscores the tendency for Platinum (and above) tier customers to out-perform their lower tier peers, but also how that effect is modulated by outside conditions. This will be explored more deeply in the section on tiers below.

Figure 27. Time series of all evaluated per-participant event impacts, grouped by customer tier.



Event timing can impact outcomes for a lot of reasons. First and foremost, behavioral DR relies upon occupancy for its non-automated saving. Also, the amount of cooling loads online on event days will tend to vary by time of day. Figure 28 Provides a view of the range of monthly summary per-participant event impacts for each hour events start in. The extent of the boxes include all seasonal variability. Events with start time of 5pm exhibit higher variability but also higher mean impacts than other events. There was just one event each starting at 2 pm and 3 pm.

Figure 28. Event outcomes by time of day.



Load Impacts by Customer Sub-group

This section presents load impacts for device ownership, mega events, customer tier, sub-LAP, and LCA. Table 13 summarizes typical (i.e. the aggregation of results from Jun-Sep) outcomes for all the relevant sub-group categories.

Table 13. Typical outcomes for all evaluated customer sub-groups

Group	Category	# Typical unique participants	Average temp. (F)	per-participant ref. load (kW)	Per-participant impact (kW)	Coincident reference load (MW)	Coincident aggregate impact (MW)	Impact % (of ref)
ALL	All	85,088	77.95	1.92	0.62	163.48	52.38	32.04
device	No	78,939	78.14	1.91	0.55	150.73	43.12	28.61
	Yes	9,566	76.20	2.04	1.29	19.50	12.39	63.53
mega	No	84,103	76.95	1.88	0.63	158.44	52.60	33.20
	Yes	72,714	79.66	1.98	0.60	144.34	43.54	30.17
tier	Silver (and below)	75,787	78.61	2.00	0.44	151.52	33.72	22.25
	Gold	34,035	74.24	1.51	0.69	51.30	23.64	46.08
	Platinum (and above)	17,413	78.38	1.93	1.57	33.61	27.27	81.15
slap	PGF1	5,433	94.43	3.09	1.08	16.78	5.84	34.81
	PGSI	2,325	86.20	2.70	0.99	6.29	2.30	36.62
	PGNC	398	84.33	2.33	0.76	0.93	0.30	32.57
	PGNP	3,337	88.07	2.61	0.93	8.71	3.12	35.79
	PGKN	2,712	94.98	3.06	0.91	8.29	2.46	29.67
	PGST	1,146	88.25	2.81	0.91	3.22	1.04	32.22
	SCEN	2,839	90.78	3.04	0.86	8.63	2.43	28.21
	SCEC	9,451	85.55	2.76	0.77	26.09	7.25	27.80
	SCHD	2,262	98.81	2.80	0.69	6.34	1.57	24.78
	PGZP	2,474	77.35	1.91	0.60	4.73	1.49	31.41
	SCLD	166	102.07	3.30	0.73	0.55	0.12	21.99
	SDG1	11,797	74.19	1.42	0.47	16.71	5.59	33.45
	PGEB	9,117	70.71	1.34	0.51	12.22	4.69	38.39
	PGFG	1,271	69.76	1.36	0.53	1.73	0.68	39.14
	SCNW	2,555	73.58	1.66	0.52	4.24	1.32	31.13
	PGHB	623	56.34	1.04	0.40	0.65	0.25	37.94
	SCEW	10,755	73.72	1.48	0.46	15.90	5.00	31.43
	PGSB	6,708	68.21	1.13	0.45	7.60	3.01	39.56
	PGP2	2,685	67.16	1.18	0.44	3.17	1.18	37.15
	PGNB	1,298	68.29	1.23	0.43	1.60	0.56	34.85
	PGCC	1,242	60.61	0.97	0.34	1.20	0.42	35.23
	PGSF	4,588	61.98	0.60	0.18	2.74	0.83	30.45
LCA	Big Creek / Ventura	5,391	82.82	2.40	0.70	12.94	3.77	29.14
	Greater Bay	23,071	67.86	1.11	0.42	25.71	9.70	37.75
	Greater Fresno	5,433	94.43	3.09	1.08	16.78	5.84	34.81
	Humboldt	623	56.34	1.04	0.40	0.65	0.25	37.94
	Kern	2,712	94.98	3.06	0.91	8.29	2.46	29.67
	LA Basin	20,361	79.56	2.15	0.62	43.76	12.65	28.91
	North Coast / North Bay	2,967	71.15	1.44	0.52	4.27	1.54	36.10
	San Diego	11,797	74.19	1.42	0.47	16.71	5.59	33.45

Sierra	**2,325**	**86.20**	**2.70**	**0.99**	**6.29**	**2.30**	**36.62**
Stockton	**1,146**	**88.25**	**2.81**	**0.91**	**3.22**	**1.04**	**32.22**
Unspecified Local Area	**9,313**	**85.05**	**2.28**	**0.70**	**21.26**	**6.56**	**30.86**

Device: Customers with devices have typical impacts that average 1.29 kW, while those without have impacts averaging 0.55 kW. This is one of the highest gaps between customer groups and underscores the reliable contribution of direct load control to DR resources.

Mega events: There is little difference in typical average outcomes between mega events, with higher rewards and stronger messaging, and normal events (0.63 kW compared to 0.60 kW).

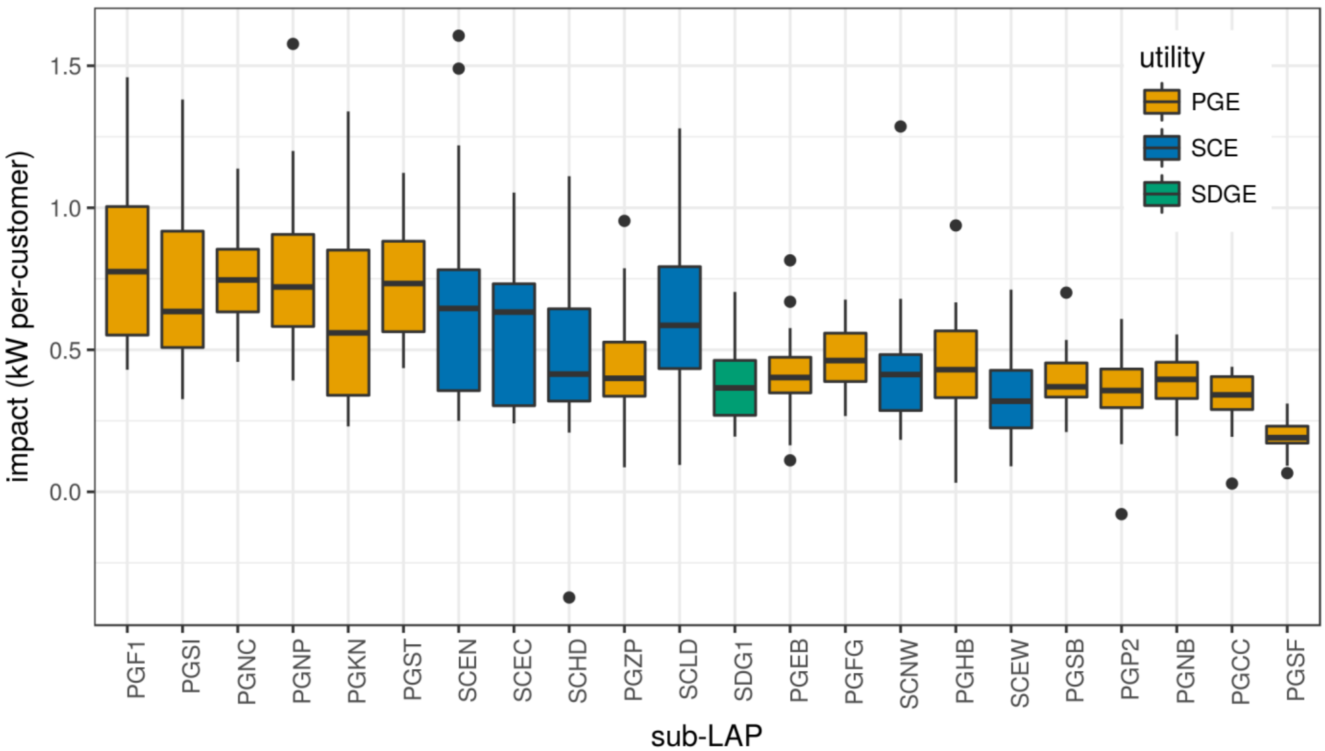
Customer tiers: Along with device ownership, tiers show the strongest differentiation within a category. Silver (and below) registers an average of 0.44 kW, while Gold is 0.69 kW and Platinum (and above) is 1.57 kW, the single highest per-participant average of any sub-group.

Sub-LAP and LCA: Among the geographic groups, Fresno (as the PGF1 sub-LAP and the Greater Fresno LCA) out performs with an average per-participant impact of 1.08 kW.

Sub-LAP-specific Results

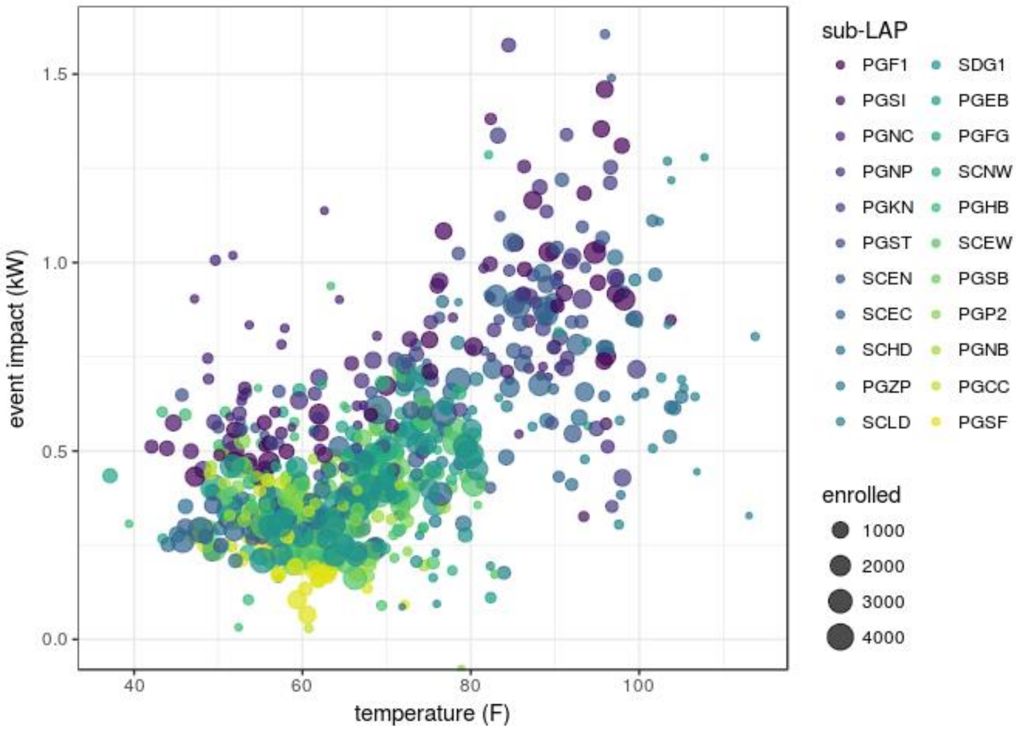
Figure 29 depicts the per-participant impacts for all events grouped by sub-LAP as box and whisker ranges. The sub-LAPs are ordered from left to right by descending average performance. The Fresno and Sierra sub-LAPs (PGF1 and PGSI) have the highest savings and San Francisco and Contra Costa counties (PGSF and PGCC) have the lowest. Virtually all of the lower performers are coastal and the high performers are inland, underscoring the role that climate plays in DR potential.

Figure 29. Per-participant event impact for all events grouped by sub-LAP in descending order of sub-LAP mean impacts and colored by utility.



Taking another look at sub-LAP performance, Figure 30 presents a scatter plot of per-participant impacts (y axis) vs. outside temperature (x axis) for all monthly roll-up events. Even in cooler weather, PGF1 tends to stick near the top of the pack.

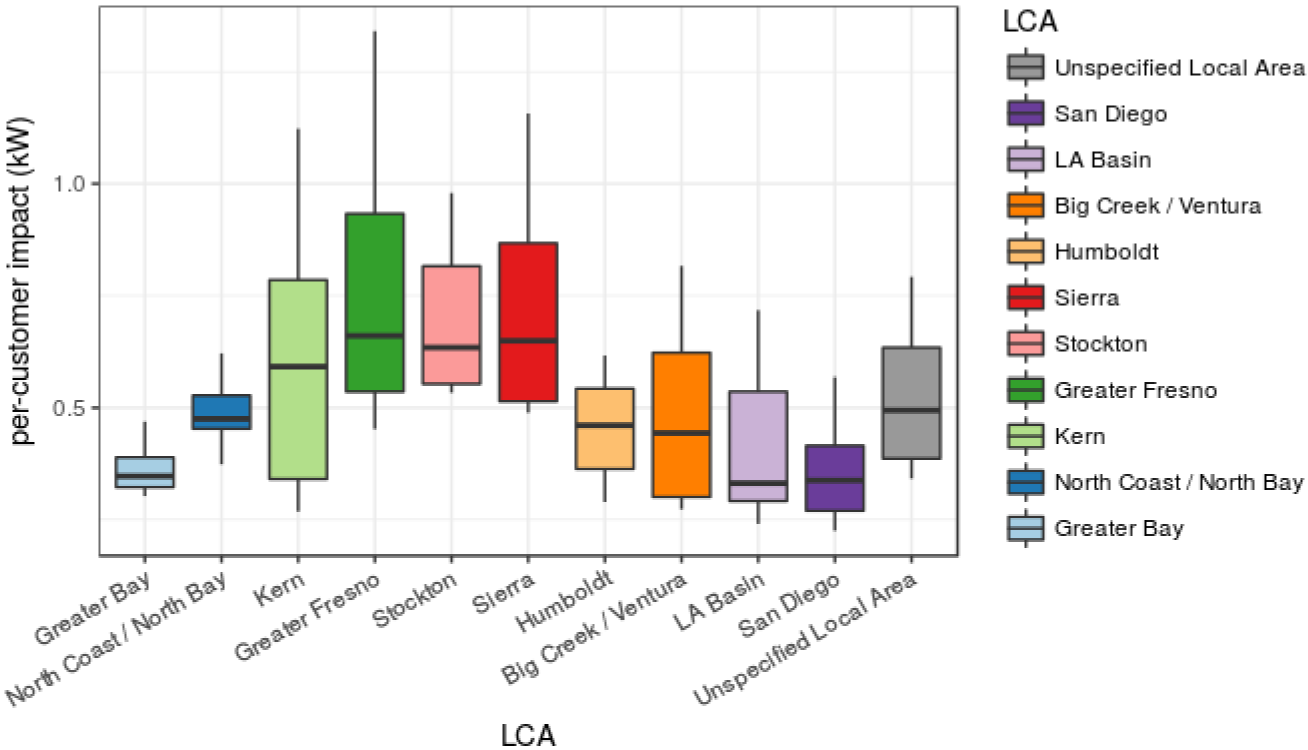
Figure 30. Scatter plot of monthly summary event average impacts vs. outside temperature, colored by sub-LAP



LCA-specific Results

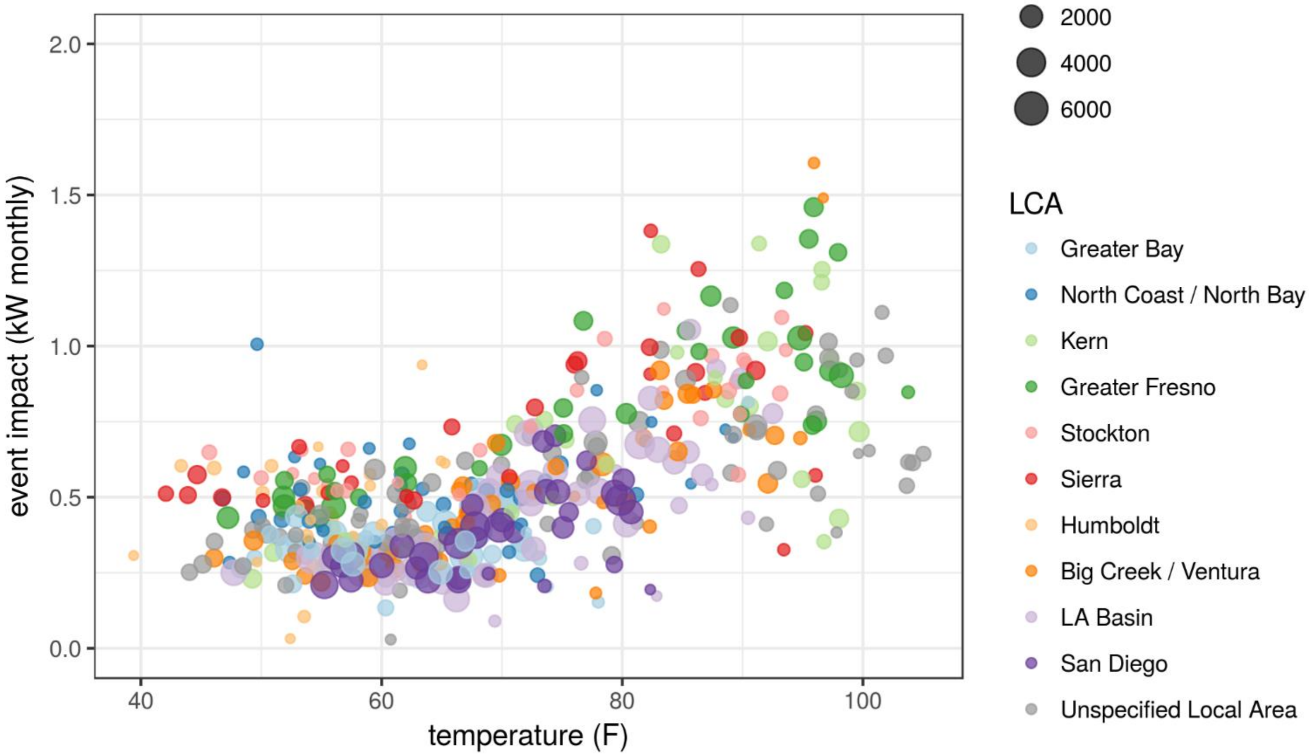
Figure 31 illustrates the average per-participant impact for all events grouped into box and whisker plots and colored by LCA. Because they are supersets of sub-LAPs, the results for LCA mirror the sub-LAP results. However, some sub-LAPs are not in LCAs and it can be verified that the per-participant performance in those locations is notably higher than what is found in coastal LCAs.

Figure 31. Per-participant impacts for all events, groped by LCA



In Figure 32, we see the relationship between per-participant event impact, temperature, and LCA, based on monthly event roll-ups. The San Diego and LA Basin LCAs tend to have the greatest number of customers in their events, but have modest temperatures and per-participant impacts. This plot reveals that Sierra tends to perform about as well as Fresno under similar temperatures, but it is not as frequently in the absolute upper range of temperatures.

Figure 32. Scatter plot of monthly average impacts vs. outside temperature, colored by LCA



Customer Tier Results

Customer tiers are the customer categories with the greatest divergence in per-participant impacts. Figure 33 illustrates the per-participant average impacts from all events for each of the tiers. The mean, spread, and maximum values for each tier increase in order from Silver (and below) to Gold to Platinum (and above).

Figure 33. Comparison of all event outcomes grouped by tier

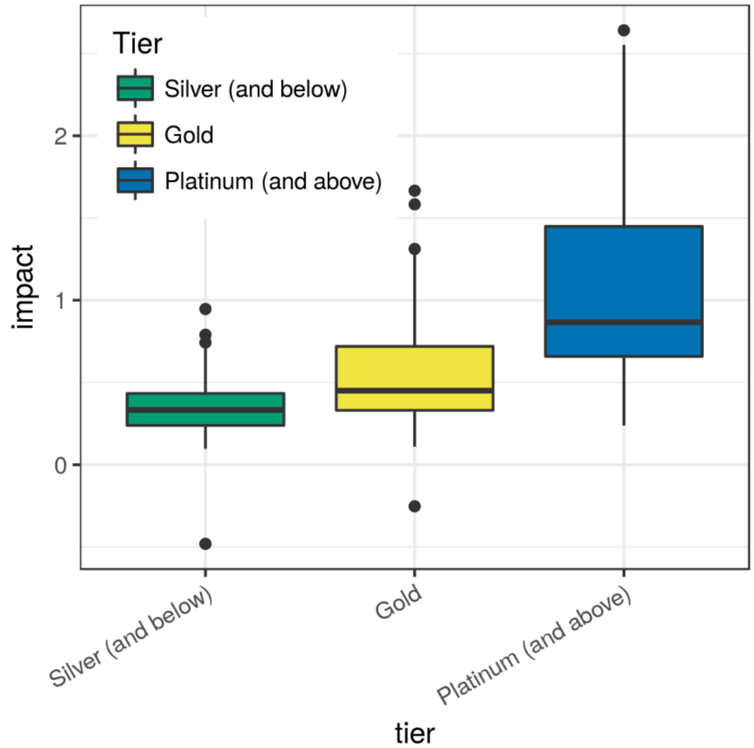


Figure 34 provides box and whisker plots for each tier for each month. While the Platinum tier always out-performs, this effect is amplified during summer months, strongly suggesting that Platinum customers are more aggressive with turning back their cooling loads.

Figure 34. Monthly summaries of all event outcomes, grouped and colored by tier.

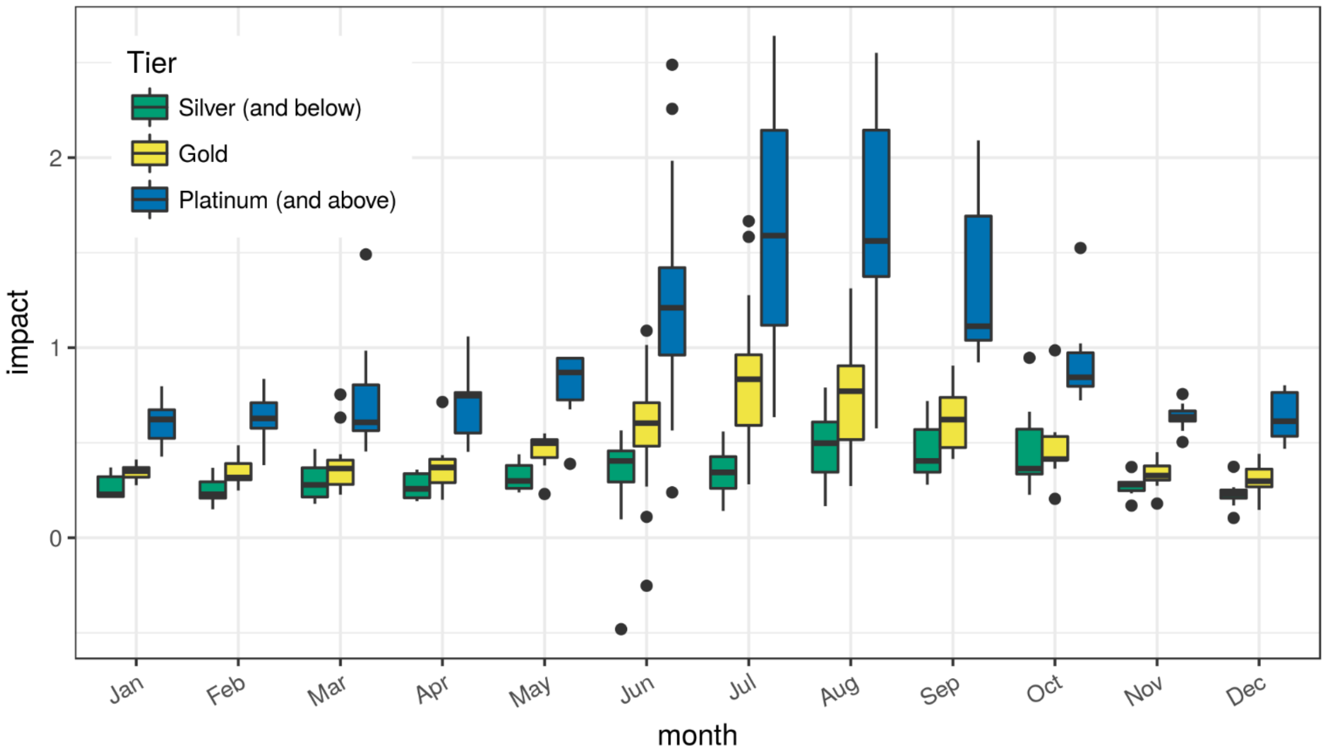
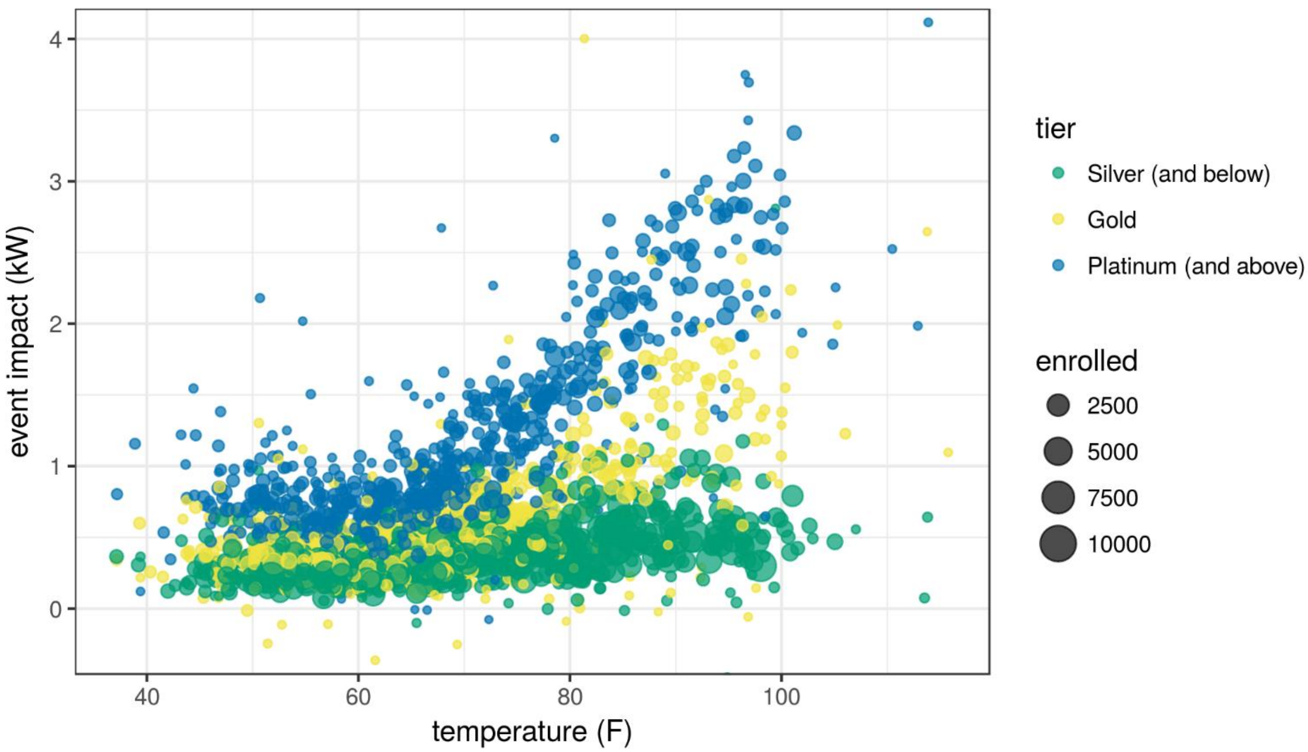


Figure 35 underscores the relationship between tier performance and temperature with a scatter plot based on all events from 2018. The results are crystal clear: Platinum customers tend to perform well at any temperature compared to their peers, but their defining feature is that they have impacts that are more reliably and strongly correlated with outside temperatures. Indeed, as we will see in the “device” section, they have a high incidence of automated devices that assist them with such tasks and there is a similar pattern in the performance of all customers with controllable devices.

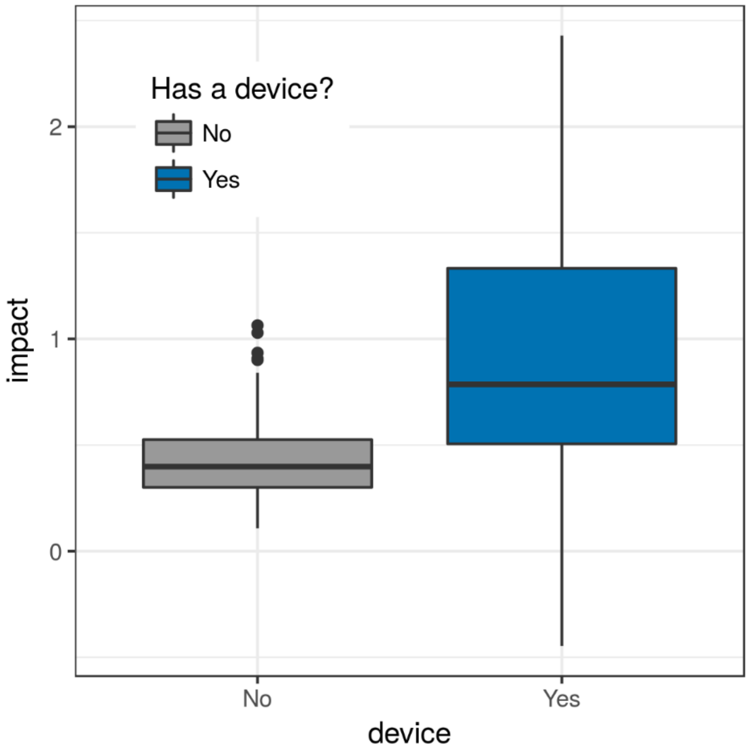
Figure 35. Scatter plots of average impact vs. outside temperature for all events, grouped and colored by tier.



Device Ownership Status Results

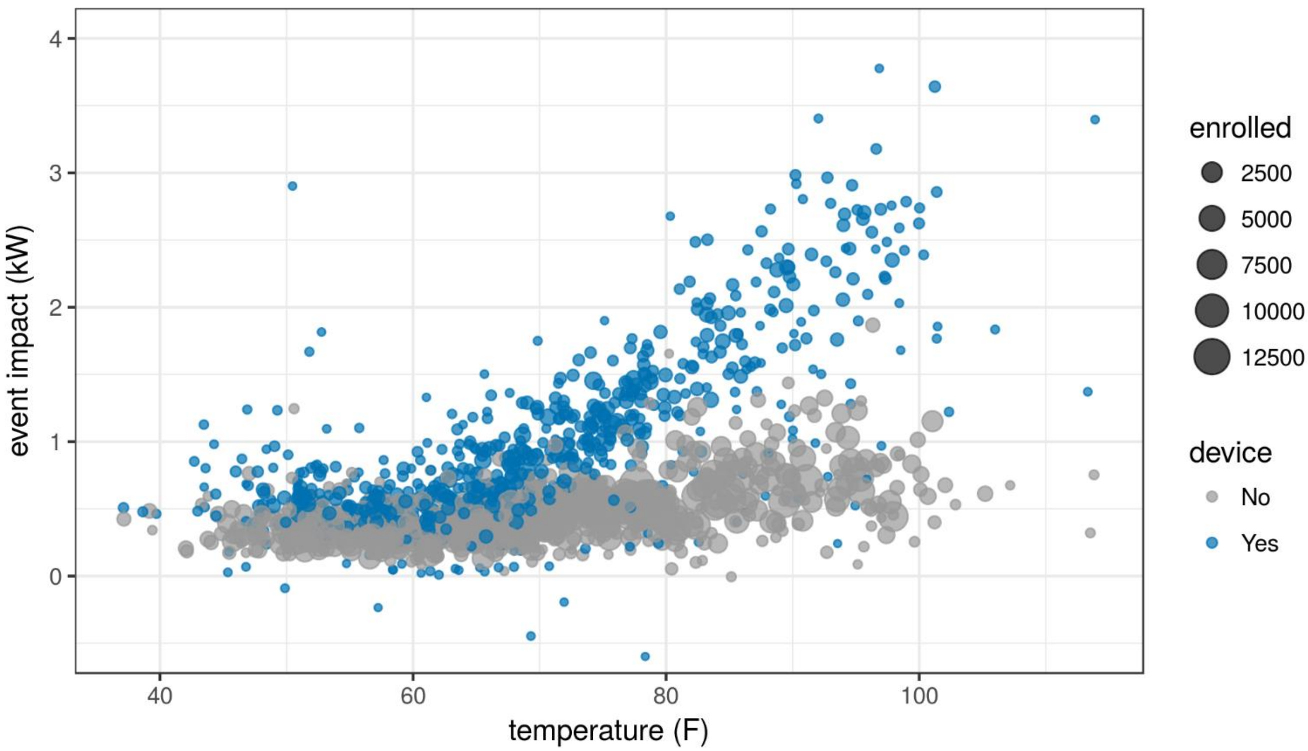
With deterministic dispatch and a tendency to be connected to larger loads, direct load control tends to be the most reliable form of DR. OhmConnect has prioritized getting devices into the hands of their customers lately and their emphasis is backed up by Figure 36 illustrating per-participant event impacts for all events in 2018 in households with and without devices.

Figure 36. Comparison of all event outcomes, grouped by device ownership.



The heart of the role devices can play in DR is revealed in Figure 37, which depicts the per-participant impacts vs. outside temperature for all events in 2018 for customers with (some form of direct load control) and without devices (purely behavioral). The performance of those with devices is higher than those without under all conditions, but there is much more overlap in cooler weather. The key factor in the divergence between the two groups is the strong correlation between impacts for device owners and outside temperature (i.e. the devices are very likely controlling cooling loads). The weak correlation between impacts and temperature for non-device "behavior only" households support the hypothesis that people are turning off or deferring similar types of loads regardless of weather. Interestingly, there is similar coverage of both groups across the full range of temperatures, indicating that there is not a strong climate pattern to device ownership (although that might be something to encourage in the future).

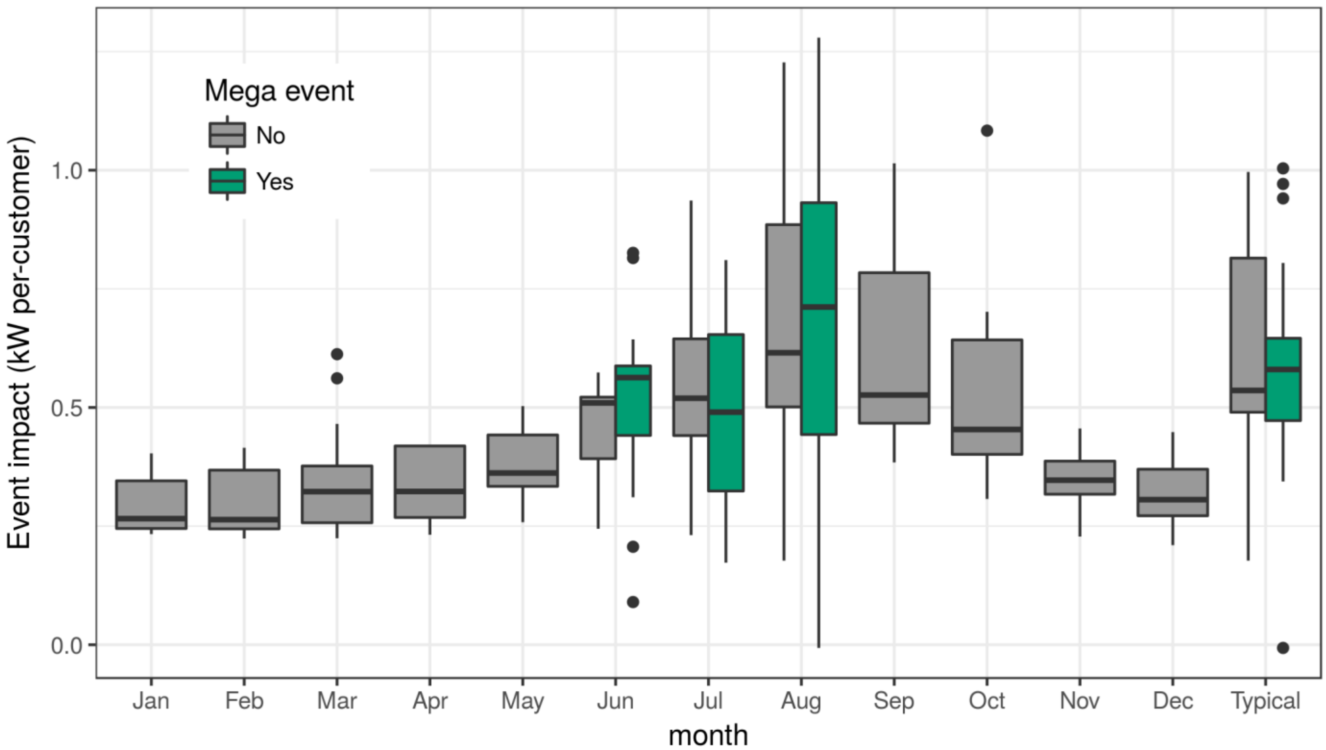
Figure 37. Scatter plot of all event impacts vs. outside temperature, grouped and colored by device ownership.



Mega Event Results

The results of our analysis of mega events suggests that they are not producing impacts systematically higher than normal events. This is illustrated by Figure 38. It is a mystery to CDA why the additional emphasis is not producing measurable differences, but there is a school of thought that studies EE and DR in terms of thresholds of motivation required for action. It is possible that OhmConnect customers, who have self-selected to be enrolled in the first place have already crossed the threshold of motivation required to participate (or not) and are already committed whatever actions they will take on event days. Furthermore, we've already explored the extent to which the top performing customer groups rely on automated device controls to achieve their greatest savings and the extent to which device driven outcomes are temperature dependent. The portion of savings that come from non-device actions is much smaller in summer than in the winter months. There is also a lot of nuance associated with the ways that marketing can lead to specific outcomes. It may be that OhmConnect has not yet dialed in the most salient ways to present mega events to their customers.

Figure 38. Monthly per-participant event impacts grouped and colored by mega event status.



Current Ex Post to Prior Ex Post

Table 14 presents the single monthly ex post event roll-ups for all events in January, February, July, and August for 2017, 2018, and 2019. This table supports the comparison of outcomes from one year to the next. For this summary, one can see that early 2017 was still early in the development of OhmConnect's resource. Jan and Feb of that year saw fewer than 10,000 unique (or total!) participants. Depending on the month, either 2018 or 2019 have the best per-participant performance, which tracks reference loads in every case, except when Aug 2018 per-participant impacts are higher than Aug 2019, despite having lower reference loads.

Table 14. Ex post results summary for comparison across 2017, 2018, and 2019

month	year	# of events	average # of participants	# monthly unique participants	total # participants	average temperature (F)	reference load (kW)	per-participant impact (kW)	aggregate reference load (MW)	aggregate impact (MW)	impact %	coincident reference load (MW)	coincident aggregate impact (MW)
Jan	2017	18	434.78	3820	7826	54.84	1.10	0.22	0.48	0.09	19.49	4.22	0.82
	2018	12	4362.92	**32135**	52355	60.17	1.04	0.31	4.56	1.36	29.74	33.57	**9.98**
	2019	18	1675.28	26487	30155	54.67	**1.33**	**0.36**	2.22	0.60	26.80	**35.15**	9.42
Feb	2017	6	1530.83	7707	9185	51.75	1.09	0.31	1.68	0.47	27.94	8.44	2.36
	2018	47	1524.40	39999	71647	54.56	1.13	0.32	1.73	0.48	27.78	45.39	12.61
	2019	51	2685.59	**65517**	136965	50.54	**1.31**	**0.37**	3.52	1.00	28.53	**85.76**	**24.47**
Jul	2017	27	2321.15	19391	62671	80.91	1.98	0.56	4.59	1.31	28.47	38.35	10.92
	2018	107	2664.41	65906	285392	81.55	**2.17**	**0.61**	5.77	1.63	28.31	**142.80**	**40.42**
	2019	14	11273.43	**69903**	157857	78.67	1.99	0.53	22.39	6.02	26.90	138.83	37.35
Aug	2017	15	1844.67	16672	27670	82.15	2.08	0.58	3.84	1.06	27.64	34.72	9.60
	2018	119	2250.86	70984	268075	78.30	1.89	**0.66**	4.26	1.50	35.10	134.46	47.19
	2019	22	10976.41	**80220**	241526	81.40	**2.10**	0.61	23.07	6.70	29.05	**168.61**	**48.99**

Current Ex Post to Prior Ex Ante

This report represents the first time OhmConnect's DR resource has been evaluated using the DR Load Impact Protocols. For this reason, there is no prior ex ante to compare to. See the ex ante results section for a comparison of ex post to the ex ante model based on 2018 events.

Ex-Ante Results

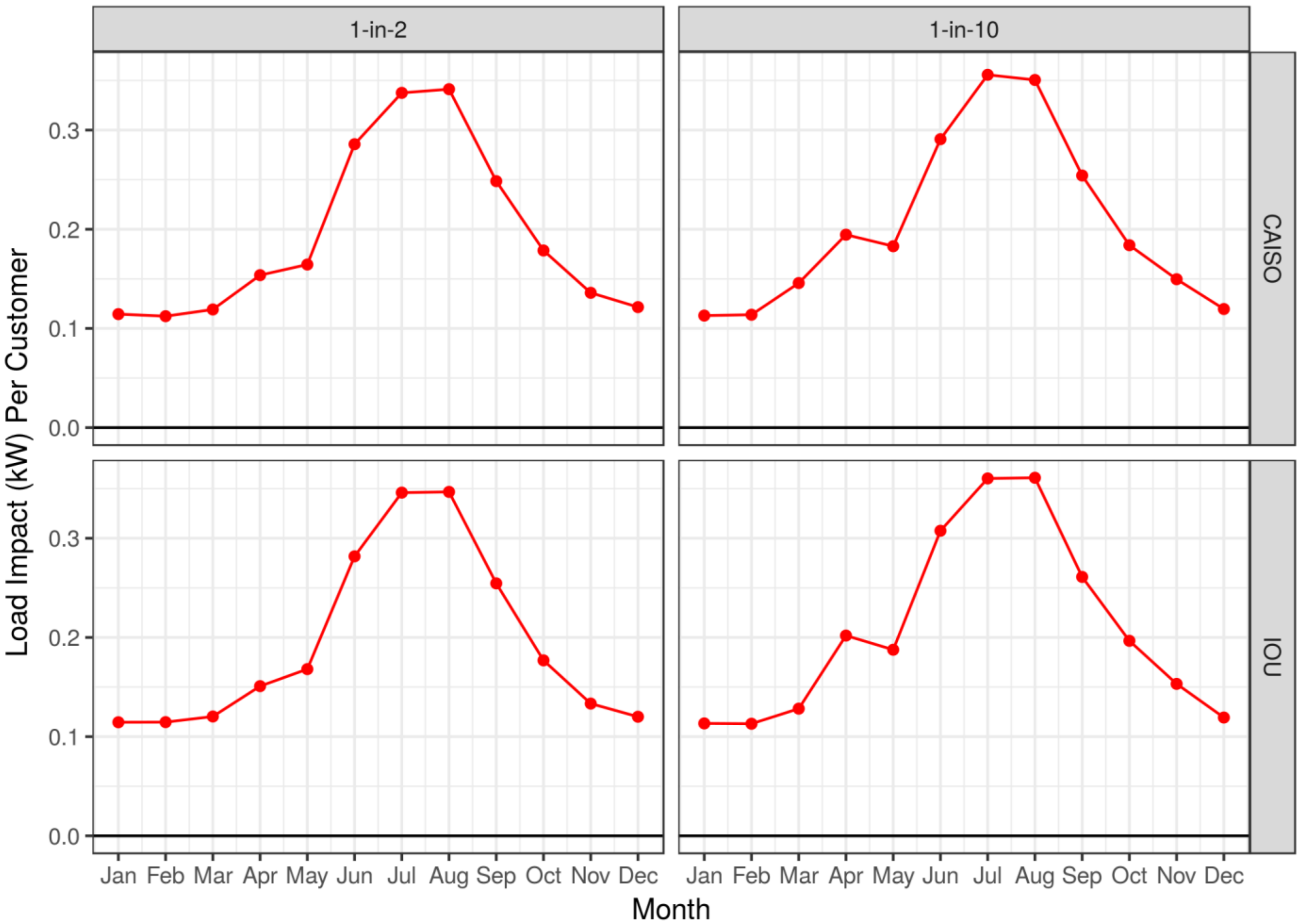
Ex ante load impacts are based on forecasts for standard event times and conditions. Forecasts are made for two standard weather years – that is, two sets of monthly peak temperatures – corresponding to conditions that are expected to lead to peak electric load in one out of every two years and one out of every ten years on average. There is a slight difference in the conditions that cause peak load for the statewide California Independent System Operator (CAISO) and Investor-Owned Utilities (IOU) territories. Accounting for the two peaking conditions with the two weather years leads to four sets of standard weather data per territory.

Unless explicitly stated, all forecasts discussed in this report are made for events in the Resource Adequacy (RA) time window that runs from 4 p.m. through 9 p.m.

Predictions are made by statistical models that predict the load impact per participant for different categories of customers (such as customers with and without electric devices that are automatically controlled to reduce load during a demand response event). The models were fit to data from calendar year 2018 and used to predict the load shed per participant for the standard weather conditions mentioned above, for the projected mix of customers by category that is forecast for future years. The models, their coefficient fits, and their diagnostic information are described in an appendix to this report.

Figure 39 shows the predicted load impact per customer for each of the four sets of weather conditions described above (recall that ex post is presented as per-participant). All four sets of conditions produce rather similar predictions. For these four hypothetical weather-years the highest load impact per customer occurs in July and August, at about 0.35 kW per customer; the lowest occurs in December through March, at about 0.12 kW per customer.

Figure 39. Predicted load impact per customer, by month, for four standard sets of weather conditions.



OhmConnect expects to grow quickly over the next few years (see the section on enrollment forecasts, below). The projected growth in enrollment varies by sub-LAP according to where OhmConnect intends to market their program most heavily and by customer attributes OhmConnect plans to target their recruitment around. For example, they are currently focused on getting more of their customers to adopt devices. The result is a projected rapid increase in load impact.

Figure 40 shows the predicted aggregate load impact for 2019; this is the product of OhmConnect predicted enrollment in each month times the predicted load shed per customer in each month, again for the standard weather conditions. The maximum is around 41 MW statewide, in August. Whereas the predicted load impact per customer (Figure 4) is about the same in January as in December, this is not true of the aggregate because the number of projected event participants increases substantially over the course of the year.

Figure 40. Predicted aggregate load shed by month, for four standard sets of weather conditions.

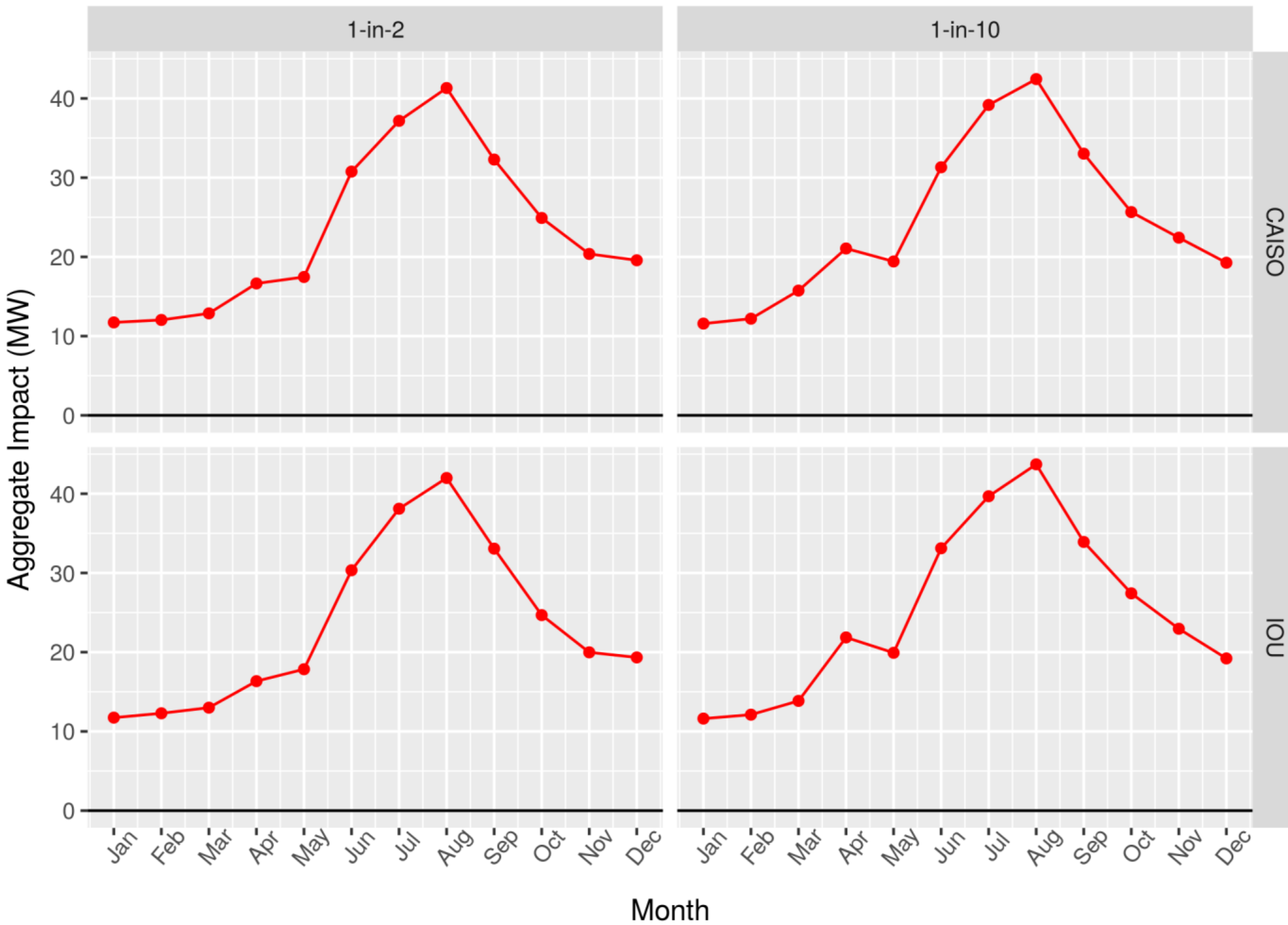


Figure 41 shows the same sort of information, this time for the next several years. The projected year-over-year increase in customers (and thus event participants) leads to a very rapid year-over-year increase in load impact.

Figure 41. Forecast aggregate load impact, by month and year. Points for 2022 are hidden behind 2023.

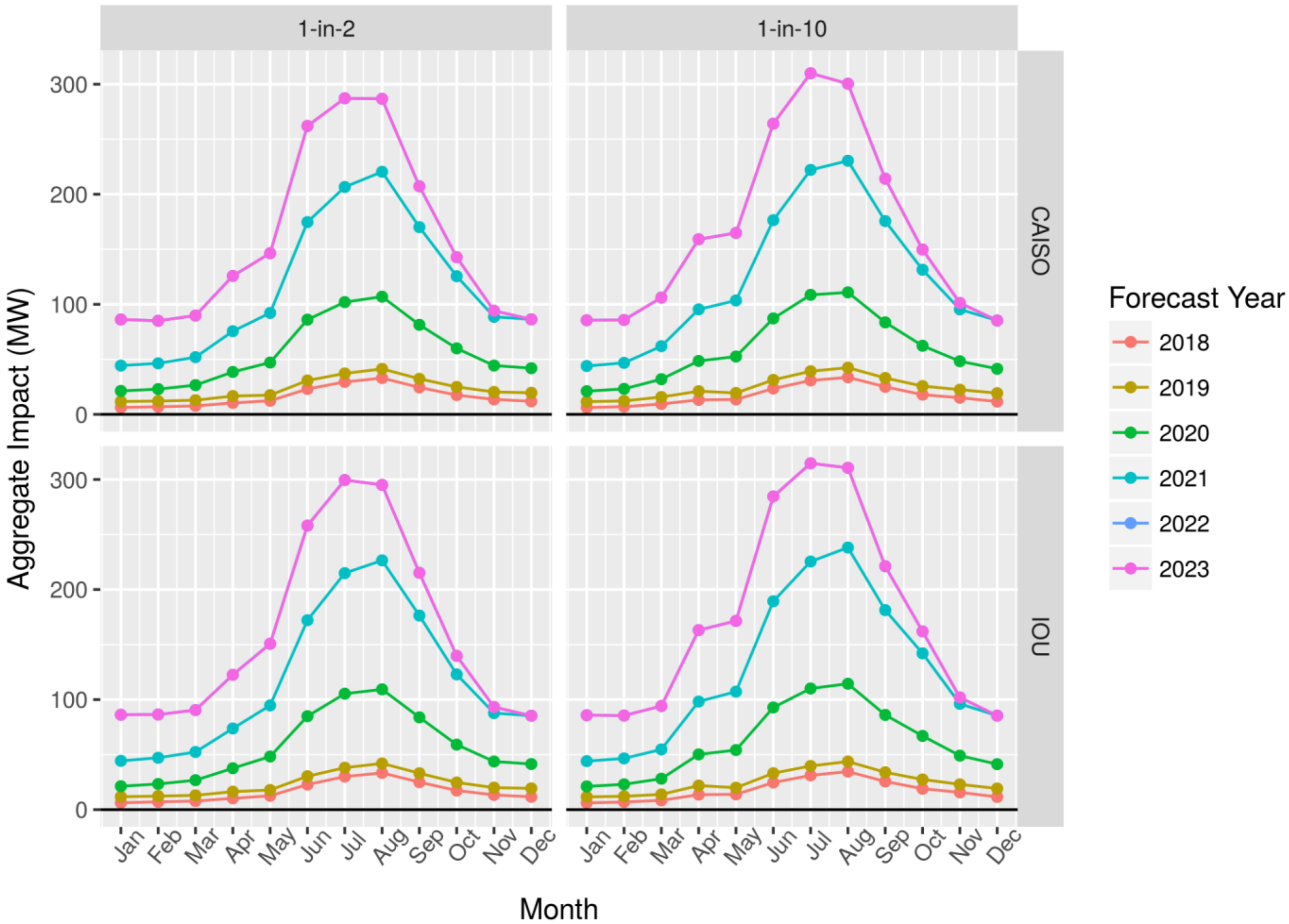


Table 15 shows the forecast load impact by year for the August 1-in-2 monthly CAISO peak day. The very large projected increase over the next two years is mostly due to a projected increase in enrollment, especially in sub-LAPs with high load shed per participant; secondarily, the fraction of customers who provide direct load control devices is forecast to continue to increase, and these customers provide more load shed than others.

Table 16 shows the August aggregate results for IOU peaking conditions rather than CAISO peaking conditions.

Table 15. Forecast Aggregate Load Impact for August CAISO 1-in-2 day.

Year	Temp (F)	Aggregate Impact (MW)
2018	80.58	32.97
2019	81.37	41.31
2020	82.86	106.93
2021	83.99	220.43
2022	84.48	286.72
2023	84.48	286.72

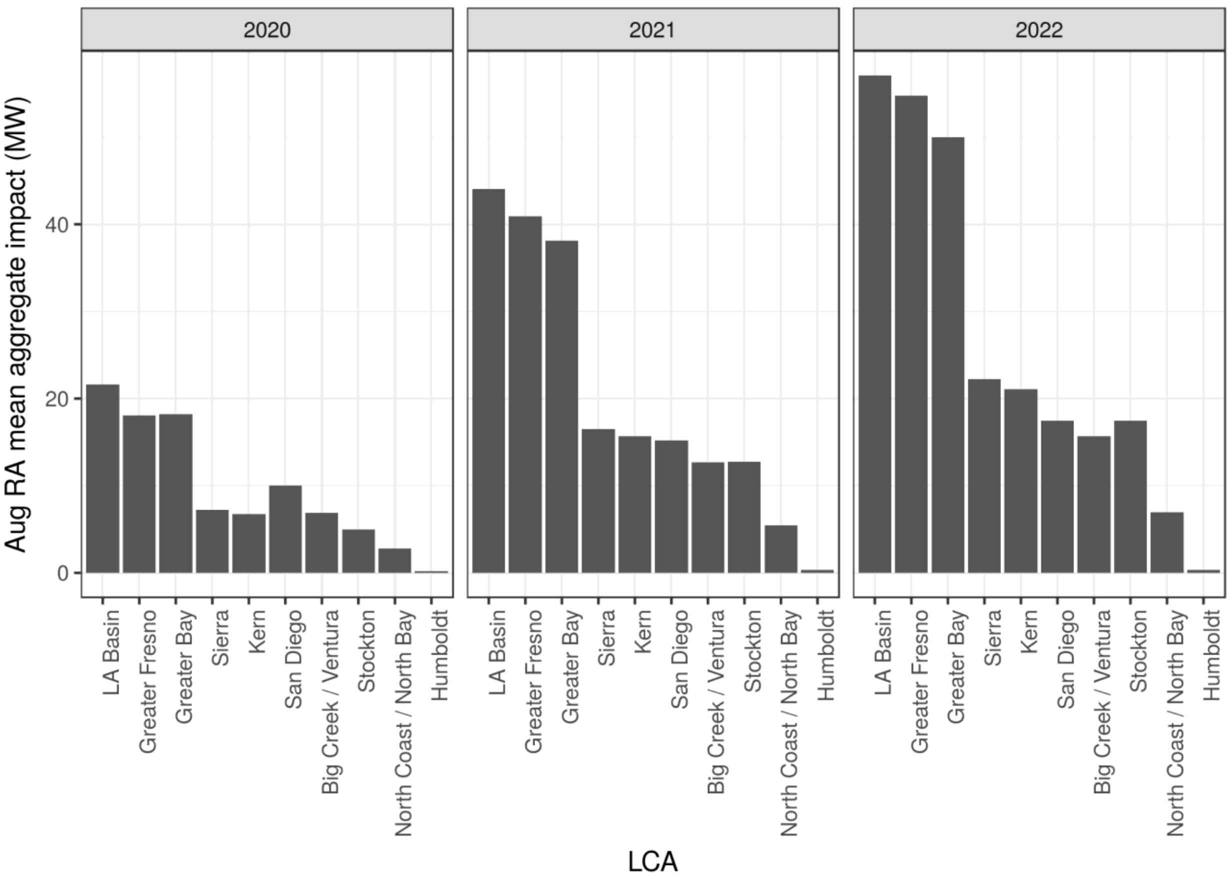
Table 16: Forecast Aggregate Load Impact for August IOU 1-in-2 day.

Year	Temp (F)	Aggregate Impact (MW)
2018	82.14	33.40
2019	83.19	41.98
2020	85.12	109.29
2021	86.45	226.50
2022	87.04	295.09
2023	87.04	295.09

This report is being written in October 2019, and data for the first eight months of the year are available for comparing to predictions. However, the events that were called in 2019 do not correspond to the standard weather conditions, so to evaluate the accuracy of the statistical model we instead compare predictions to observations for the outdoor temperatures at which the events occurred, and so should not be compared to the tables above. Instead, we make this comparison in the “Future ex post to ex ante” section below.

The ex ante predictions were made for each sub-LAP - there are sub-LAP-specific terms in the models, as illustrated in the appendix on modeling, and the number of customers in each customer category varies by sub-LAP – but sub-LAPs can be mapped onto Local Capacity Areas to make forecasts at the LCA level. Figure 42 is an example: it shows the predicted aggregate impact for August, for the IOU 1-in-2 weather year.

Figure 42. Forecast aggregate impact by Loacal Capacity Area, mean over RA hours, for August of different years.



Aggregate impacts by month are shown in Figure 43, separately for PG&E, SCE, and SDG&E and by year.

Figure 43. Forecast aggregate impact by month (average over RA hours), separately by utility.

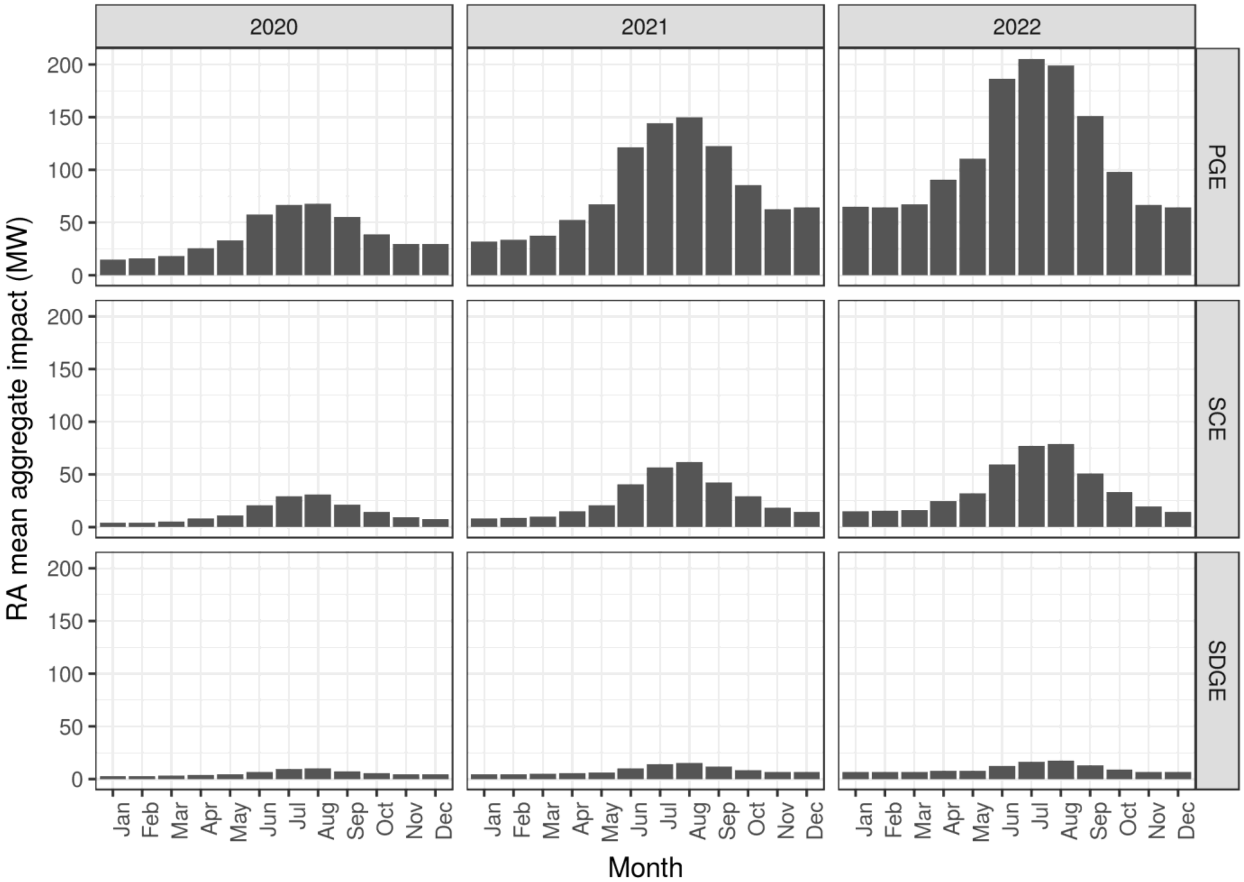
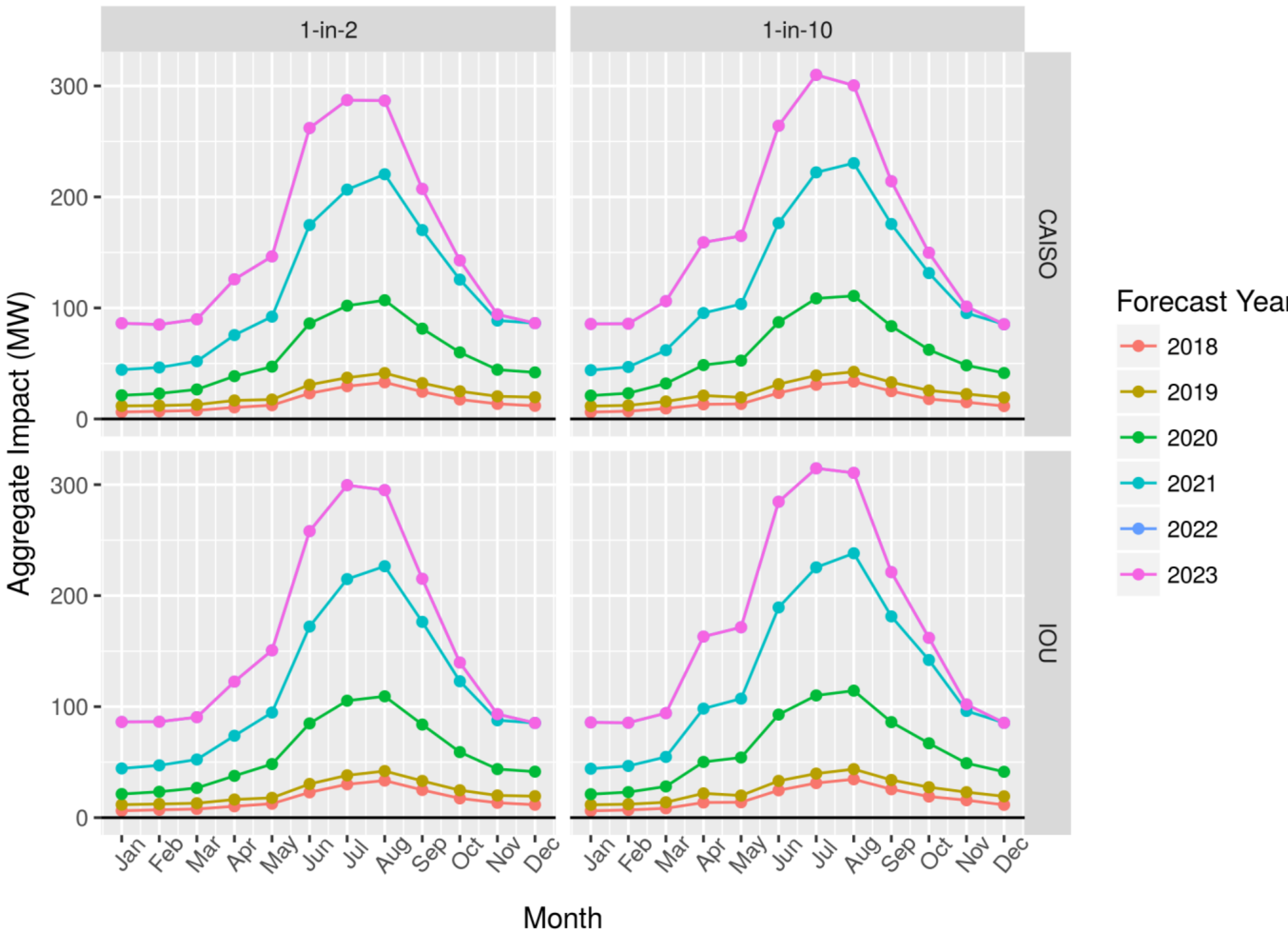


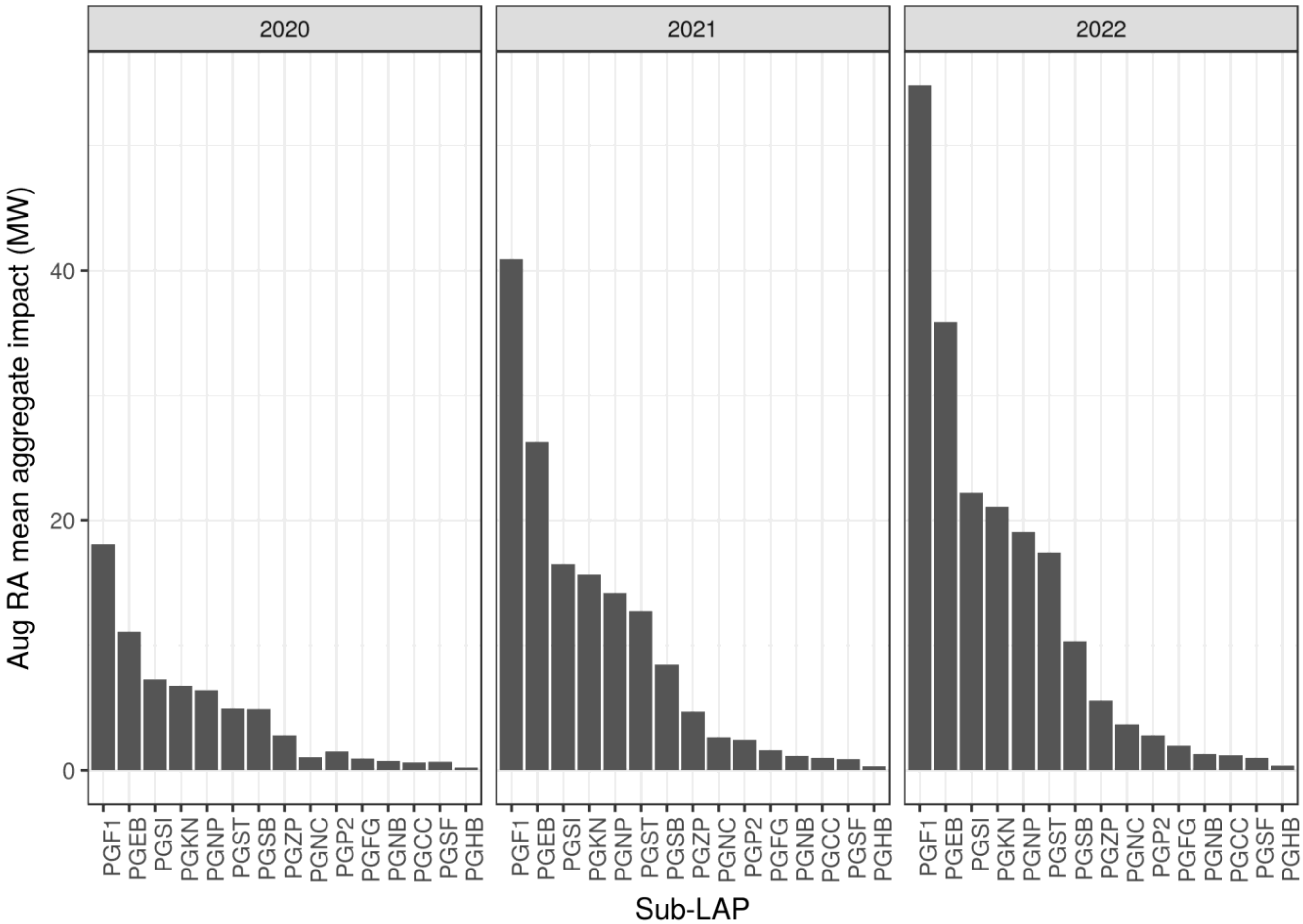
Figure 44 shows the aggregate load impact by month and year for each of the four standard sets of weather conditions. By 2023 the forecast load shed for a July or August event in a typical year is around 300 MW, and in a hot year (the 1-in-10 year) it is even higher.

Figure 44. Forecast aggregate load impact, by month and year. Points for 2022 are hidden behind 2023.



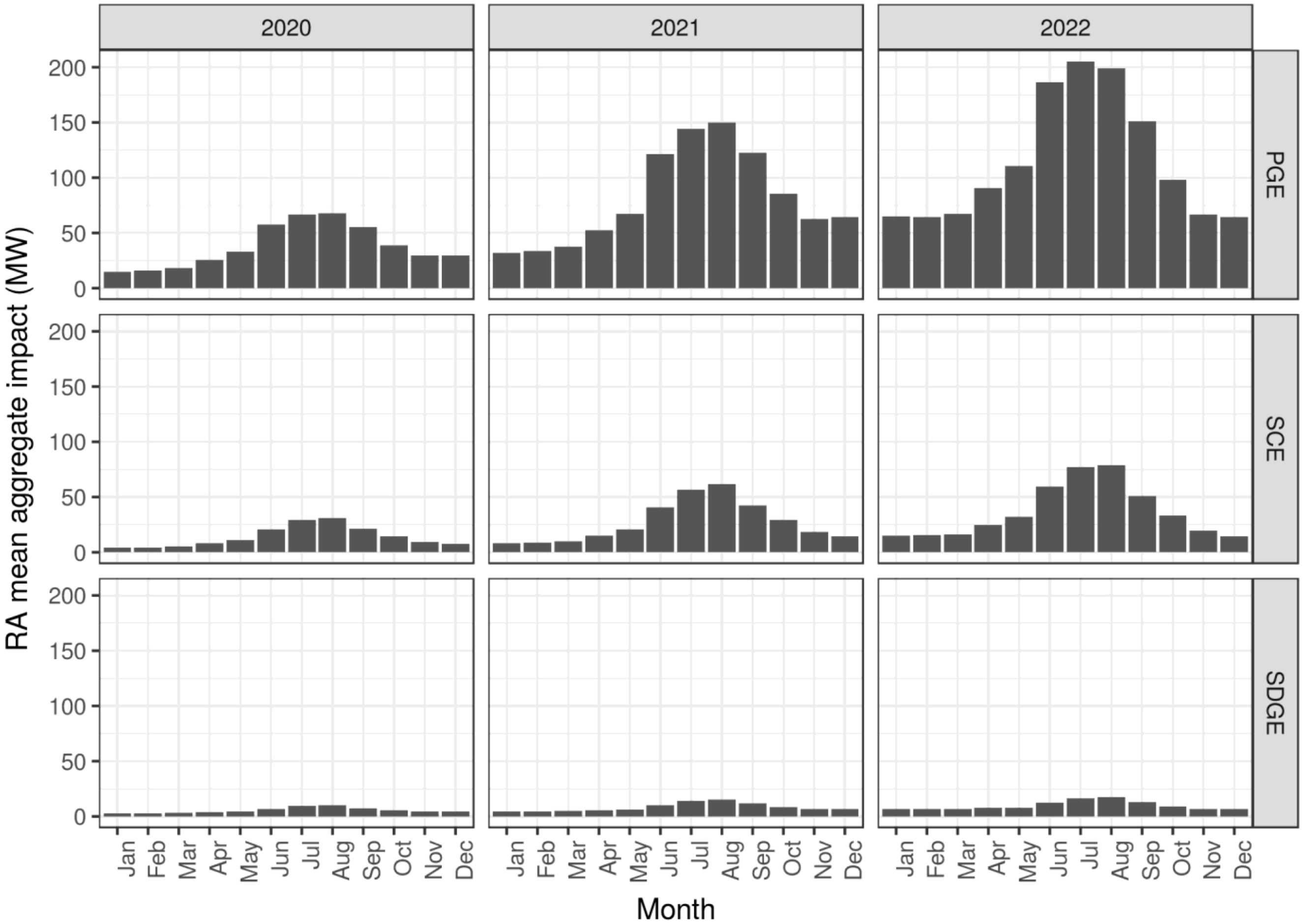
The ex ante models are fit at the sub-LAP level and predictions are made for each sub-LAP - there are sub-LAP-specific terms in the models, as illustrated in the appendix on modeling – but sub-LAPs can be mapped onto Local Capacity Areas to make forecasts at the LCA level. Figure 45 is an example: it shows the predicted aggregate impact for August, for the IOU 1-in-2 weather year.

Figure 45. Forecast aggregate impact by Local Capacity Area, for August of different years. Average over RA hours.



Aggregate impacts by month are shown in Figure 46, separately for PG&E, SCE, and SDG&E.

Figure 46. Forecast aggregate impact by month, separately by utility. Average over RA hours.



Current Ex Post Compared to Current Ex Ante

Comparing the ex post data to the predictions from the ex ante models for the same year helps confirm that the models perform adequately.

Table 17 summarizes the OhmConnect events that took place between 4-9 p.m. in 2018, averaged by month, and compares the observed load shed to the predictions. ‘Mean Impact’ refers to the mean impact *per customer* for both ex post and ex ante (rather than per-participant).

Table 17. Comparison of current ex post impacts to current ex ante load impact predictions.

Month	Ex Post Mean Impact (kW)	Ex Ante Mean Impact (kW)	Reference Load (kW)	% Load Impact
January	0.34	0.33	1.10	31.10
February	0.32	0.31	1.17	27.24
March	0.34	0.33	1.09	31.64
April	0.36	0.37	1.14	31.36
May	0.44	0.46	1.34	32.68
June	0.64	0.64	2.00	31.82
July	0.61	0.66	2.04	29.90
August	0.67	0.66	1.82	36.53
September	0.66	0.64	1.81	36.56
October	0.50	0.50	1.36	36.98
November	0.39	0.41	1.23	31.45
December	0.35	0.36	1.35	26.18

Future Ex Post Compared to Ex Ante

The conventional way of comparing current ex post to future ex ante is to scale the ex post results to account for projected enrollment differences between past and future and then compare the scaled results. That approach is undesirable for this program because OhmConnect calls so many events, with widely varying numbers of customers, on many different days, with very different numbers of events called at different times of day and in different months. Instead, we compare the 'future ex post' to the future ex ante predictions. 'Future ex post' seems to be an oxymoron: obviously one cannot look back on the future. But in this case 2018, the last calendar year for which we have complete data, is 'current', but this report is being written in October 2019 and data through August 2019 are available, so we are able to compare the ex ante predicted load impact with actual data through August 2019 in Table 18. As with comparing 2018 ex post to 2018 ex ante, we use the 2018 ex ante model with input temperatures that applied to the actual 2019 events, and then roll up the results to the month.

Table 18. Comparison of ex post impacts for the first eight months of 2019 to the predictions from the 2018 ex ante model applied to actual 2019 temperatures and enrollment.

Month	Ex Post Mean Impact (kW)	Ex Ante Mean Impact (kW)	Reference Load (kW)	% Load Impact
January	0.35	0.34	1.33	26.36
February	0.37	0.33	1.29	28.37
March	0.33	0.36	1.16	28.69
April	0.27	0.42	1.01	27.08
May	0.36	0.42	1.07	33.66
June	0.54	0.52	1.96	27.75
July	0.54	0.60	1.97	27.39
August	0.61	0.54	2.11	28.92

Enrollment Forecast

To make ex ante load forecasts, a forecast for the number of customers by sub-LAP and customer category is needed. CDA used forecasts provided by OhmConnect for this purpose. OhmConnect has already seen significant growth in 2019 and projects an increasing number of customers for the remaining months of 2019 and for the next several years. The number is assumed to increase until the end of 2021 and then remain flat.

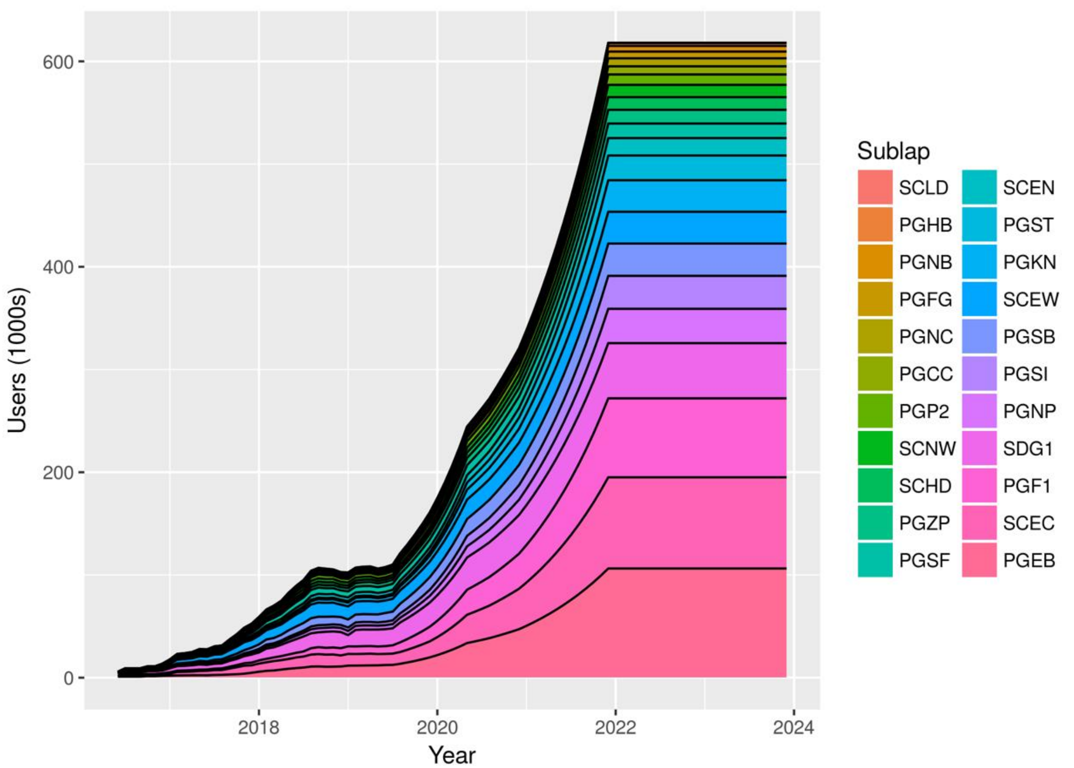
Enrollment Forecast Rationale

CDA did not produce the enrollment forecast used for the ex ante modeling. The forecast was provided by OhmConnect based on their internal growth models used for budgeting and planning. They have experienced rapid growth historically and offered a written explanation for their forecast in their own words, which we have included as Appendix B of this report.

Effects of Forecasted Enrollment

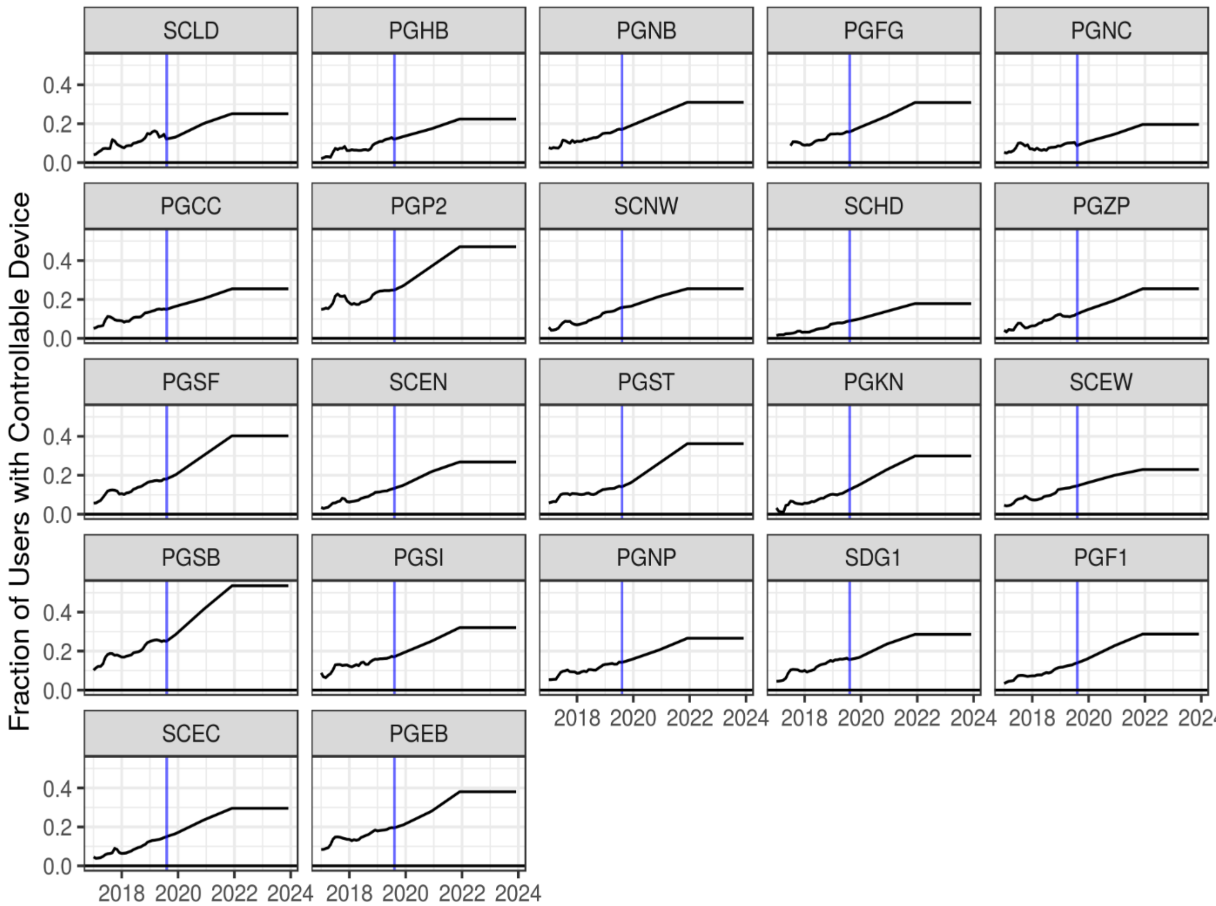
OhmConnect projects the mix of customers among sub-LAPs to change substantially. The projected growth in enrollment varies by sub-LAP according to where OhmConnect intends to market their program most heavily combined with where they think it will be most popular. The enrollment forecasts show that OhmConnect intends to prioritize recruitment from sub-LAPs that provide the most load impact per customer. Figure 47 shows the projected number of customers in each sub-LAP as a function of time.

Figure 47. Projected enrollment by sub-LAP.



OhmConnect also expects to improve the load shed per customer, for customers within a given sub-LAP, by increasing the number of customers who allow OhmConnect to enable direct load control and by marketing to customers who are willing to take more effective actions during a demand response event. In the forecasts these show up as a larger increase in the number of customers in the 'Gold' and 'Platinum' tiers than in the 'Silver' tier, and an overall increase in the number of customers with controllable devices. The panels of Figure 48 are in order by the projected number of customers (highest at top left) to match Figure 47.

Figure 48. Projected fraction of customers with direct load control devices, by sub-LAP



The effects within a sub-LAP are expected to be modest in all but a few sub-LAPs, as shown in Figure 49 below. This figure shows the predicted load impact per customer in each sub-LAP, for the IOU 1-in-2 weather year, in each year through 2023. Changes over time are caused by an increase in controllable devices, and to some extent to improved recruiting of higher-performing customers. Panels of the plot are in order of increasing 2020 enrollment from upper left through lower right, and are labeled by sub-LAP.

Figure 49. Predicted load impact per customer by month, in each sub-LAP, for several years.

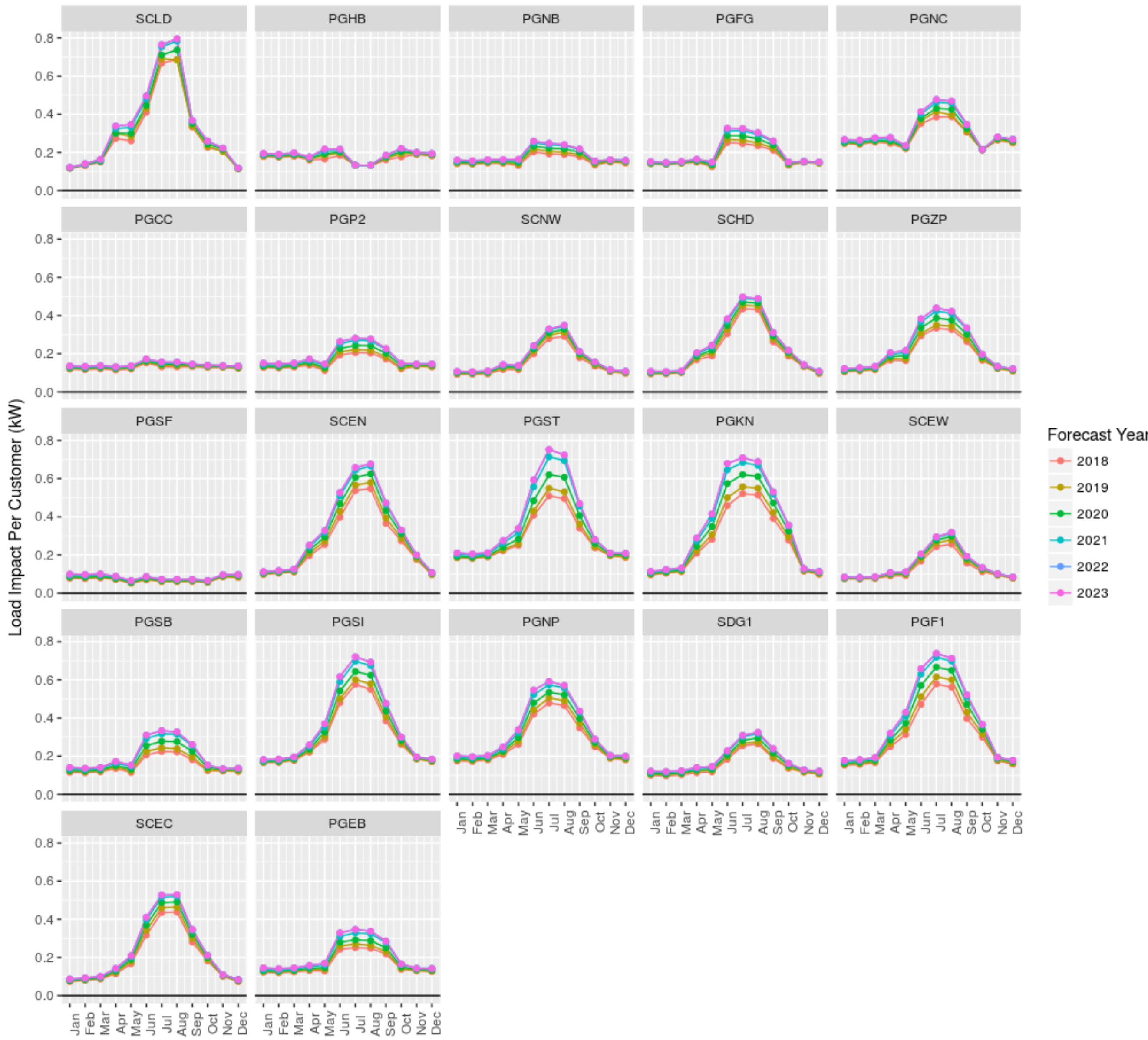
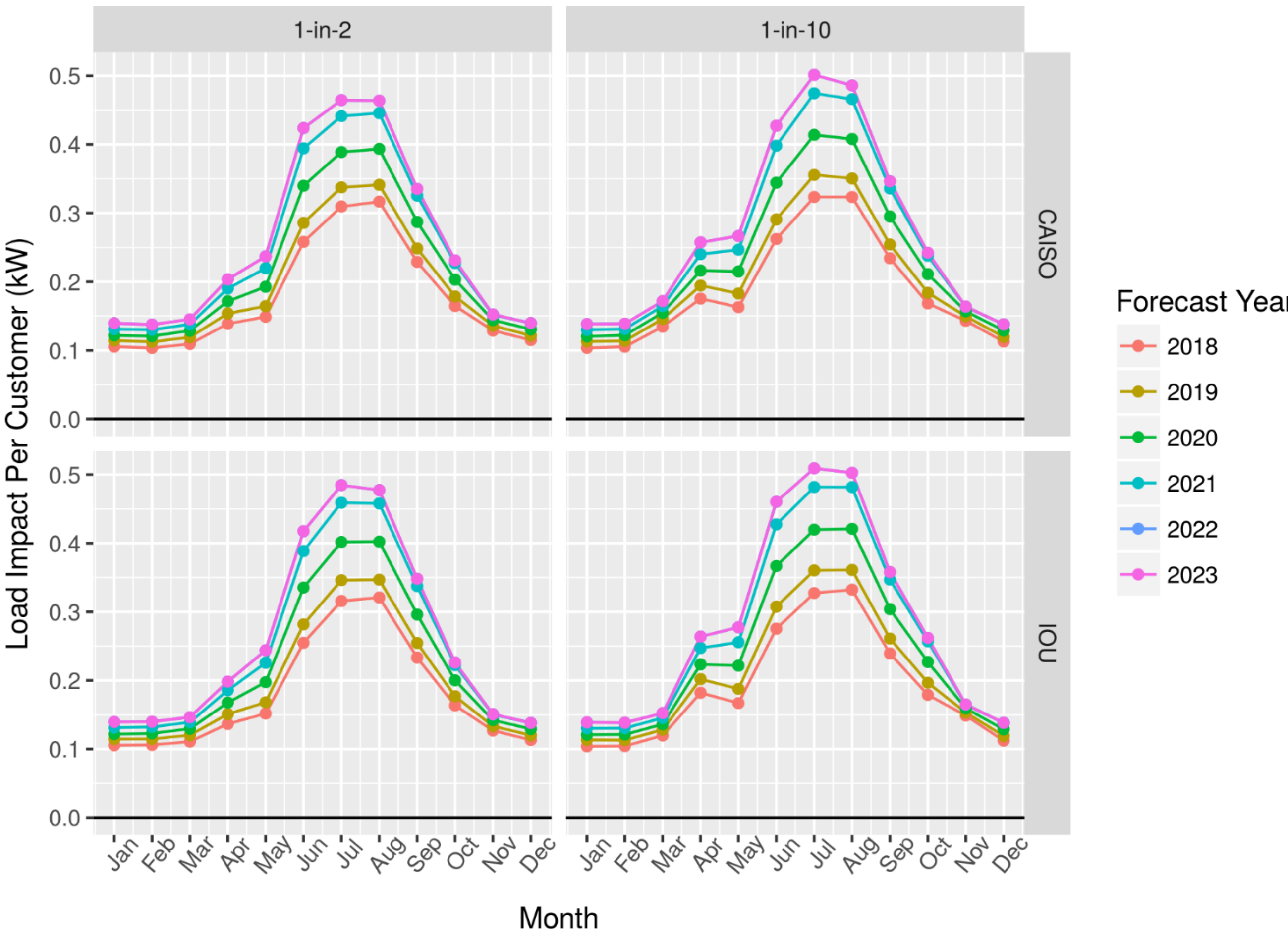


Figure 50 shows the predicted load impact per customer by month through 2023. The projections for 2022 are identical to those for 2023, so the plotted points for 2023 are hidden behind those for 2022. Almost all of the year-to-year change in load impact per customer is due to the fact that a higher percentage of future customers are expected to come from sub-LAPs where customers provide high load impact and the ownership of load control devices is expected to increase, rather than improvements within each sub-LAP.

Figure 50. Predicted load impact per customer, by month and year. Points for 2022 are hidden behind 2023.



The large projected increase in total customer count combines with the projected improvement in load impact per customer to lead to a very large increase in the projected program-wide aggregate load impact. Figure 51 shows the aggregate load impact by month and year for each of the four standard sets of weather conditions. By 2023 the forecast load impact for a July or August event in a typical year is around 300 MW, and in a hot year it is even higher.

Figure 51. Forecast aggregate load impact, by month and year. Points for 2022 are hidden behind 2023.

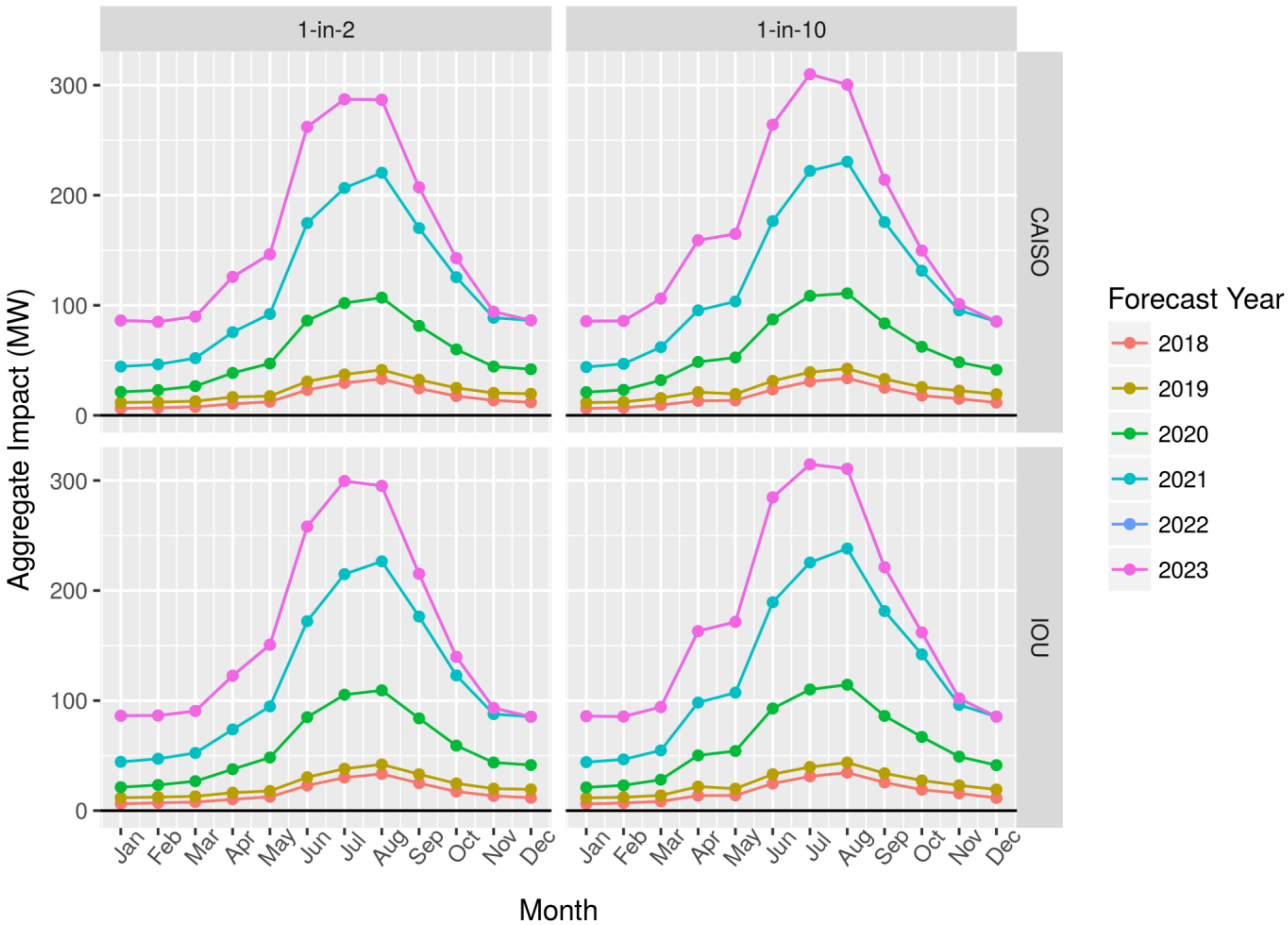


Table 19 shows the forecast load impact by year for the August CAISO 1-in-2 monthly peak day. The projected increase over the next two years is due mostly to a projected increase in enrollment, especially in sub-LAPs with high load shed per participant. Table 20 shows the forecast load impact by year for the August IOU 1-in-2 monthly peak days.

Table 19. Forecast Aggregate Load Impact for August CAISO 1-in-2 day.

Year	Temp (F)	Aggregate Impact (MW)
2018	80.58	32.97
2019	81.37	41.31
2020	82.86	106.93
2021	83.99	220.43
2022	84.48	286.72
2023	84.48	286.72

Table 20: Forecast Aggregate Load Impact for August IOU 1-in-2 day.

Year	Temp (F)	Aggregate Impact (MW)
2018	82.14	30.06
2019	83.19	38.11
2020	85.12	105.36
2021	86.45	214.89
2022	87.04	299.51
2023	87.04	299.51

Potential sources of error in the results.

Baseline with SDA: risk of over-estimation

Any evaluator is concerned with where and how they might be wrong. In the case of this evaluation, we are comfortable enough with the 10 in 10 baseline method, in part due to its history of use, to produce results using it with same day adjustments. At the same time, we can clearly see that the unadjusted 10 in 10 baseline tends to (a) exhibit loads that are consistently lower than actuals prior to many events and (b) exhibit load ramping in the period prior to events that is less steep than event actual loads.

Baselines were consistently lower than event day consumption for customers who were included in the analysis. This left a lot of work to the same day adjustments (SDA). We found that a multiplicative SDA was (a) unstable for solar customers and (b) had a long tail of implausibly large multipliers that biased results upwards. On the other hand, we found that ramp rates using the additive SDA would likely be underestimated. Our resulting "dynamic additive SDA" attempted to address these issues and tends to fall between results measured with the multiplicative and additive SDA.

In other words, our work represents an improvement relative to the existing simple multiplicative or additive SDAs, but if both of those produce biased results, then our SDA probably does too.

Indeed, we can see some evidence of "savings" among the closest thing we have to controls: non-participants. However, they are a small and self-selected group with lots of problems related to representativeness and sample size.

Inclusion fraction: likely underestimation

The approach we have taken where we exclude the historical fraction of customers excluded from OhmConnect events when computing forecasted participation credits those customers with zero savings. It is likely that this group has positive average impacts, but we were unable to estimate those values and unwilling to credit them without evidence.

Conclusions and Recommendations

Behavioral DR resources are tricky to cultivate and understand. All aspects of human behavior are subtle, and superficially similar ideas can have very different outcomes determined by the nuances of their implementation. All too often, EE and DR potential is expressed in terms of simple/rational economic utility decision making, but most people do not pay enough to care about their energy or most do not understand what drives their consumption well enough to have the incentives or tools required to make purely utilitarian energy choices. What distinguishes OhmConnect's DR resource from others is their gamified approach to behavioral incentives, their very active schedule of events called in every month of the year, and their support of "bring-your-own" direct load control devices, with statuses tracked in real-time.

Overall, we conclude that OhmConnect does deliver significant DR resources on a per-participant and aggregate basis and they are able to do so more often (726 events in 2018) and under a wider range of conditions (every month of the year in hot and cold weather) than most DR programs attempt to do. In our evaluation, we have uncovered patterns consistent with what can be seen in other DR programs:

- Per-participant impacts tend to reflect per-participant reference loads, which tend to be driven by outside temperatures.
- Automated demand response, managed through communicating devices deployed in the home is more reliable and tends to produce deeper savings than purely behavioral DR, but the magnitude of the effect

is temperature sensitive, suggesting that cooling loads are the largest controlled loads (in terms of energy content).

- There are lower overall impacts, significant variability, and relatively low temperature sensitivity to be found in purely behavioral DR. This suggest that customers are (somewhat unreliably) controlling or deferring similar loads under most conditions.
- The low impact of mega events further underscores how weak the connection can be between motivating messages and altered behaviors and the extent to which direct control of large cooling loads can overwhelm more subtle event outcome influences.

Based on our evaluation of the 2018 OhmConnect resource, we provide the following recommendations:

Continue to promote the adoption of devices. Habits are hard to form and ad-hoc behavior is hard to control. The single biggest lever behavioral scientists and marketer have to shape energy consumption is influencing decisions about energy consuming/saving products. OhmConnect should continue to apply their marketing and IT know-how to growing the portion of their customers that rely on automated DR devices. As we've seen, the savings will be larger and more predictable.

Continue to recruit customers in areas and categories with high DR potential as revealed by their per-participant load impacts. Segment-specific results provide insights on the best individual targets and we can clearly see that customers in Fresno and Sierra areas are out-performing others, customers with devices are out-performing others, and customers with higher reference loads are out-performing others. Future recruitment efforts should target populations with more to offer, provided they are in locations where DR is of value to the grid.

Study the difference between smart plugs and thermostats for load control. OhmConnect has a unique resource in that they have a pool of smart plugs as well as thermostats under their control. Future work should study the differences in outcomes enabled by these different categories of devices.

Look to future diversification of device controls. The adoption of electric cars and the push to electrify space and water heating suggest that the future is bright for direct control of electric loads in the service of DR. These trends most likely represent the future of automated DR resources and stand to break the weather dependence that today's automated controls tend to have. Greater Bay customers don't use much air conditioning, but their charging EVs could be sleeping giants.

Look into randomized controls. Hands down, the best method for evaluating a resource with as much nuance and uncertainty as behavioral DR is through randomized controlled trials, where event participants and their controls are assigned at random from a pool of candidates. OhmConnect's IT infrastructure is already used to call a widely varying array of events and assign customers in and out of resource groups. Those same capabilities could be applied toward ensuring that each event (or a representative sample of events) have randomized controls. Since OhmConnect does not tend to call all of their customers at once, they need not reduce their resource potential to identify controls.

Call all customers on some days. A major leap for this evaluation is taking the results from many small ex post events and aggregating them into a resource level forecast. It would be useful for future evaluators to have empirical data on how OhmConnect customers do when a large fraction of all enrolled customers is called at the same time.

Find some way to standardize the practice around reporting on so many events. OhmConnect really does stand out in the context of other DR providers for the quantity of events they call and the range of conditions over which they call them. In many ways, the tools and systems of the Load Impact Protocols have been challenged by evaluating so many events in one year. The precedents set here by CDA should be extended or modified until a consistent approach to reporting that serves all stakeholders can be found.

Questions for Future Evaluation of Third-party DR Resources

- Third parties were not a topic of discussion when the DR Load Impact Protocols were written. The application of the protocols to third party resources place requirements on third parties and evaluators that are typically placed on IOU staff.
    - Does DRMEC need to clarify any methodological distinctions it would make when evaluations are for third parties? Do third parties need representation on DRMEC?
    - What party needs to be on the relevant service lists to notify stakeholders of draft and final documents and facilitate comment periods?
    - Should there be specific requirements for the first time a third-party resource is evaluated vs. subsequent times?
    - What level of access to third party data should the CPUC be granted?
- Third parties will often not have access to the full set of customer account, historical meter, and non-participant data that utilities do.
    - Should utilities be required to provide supporting data (i.e. customer account attributes) to evaluators when third parties do not have access to them?
    - How should data restrictions be factored into the review of evaluations?
- Third parties may call orders of magnitude more events with more widely varying participation and dispatch conditions than IOU-run DR programs.
    - What is the correct way to establish the full resource potential from such events?
    - Should third parties be required to dispatch "full resource" events for evaluation purposes?
- Third parties may be undergoing rapid growth with significant changes baked into their enrollment forecasts.
    - How should forecasts be evaluated? And by whom?
- Third parties will tend to have trade secrets or other sensitive or protected information they are shielding from their competitors and such information may not be available for public scrutiny.
    - How should the limits of what can be disclosed be factored into the reporting requirements?
- Third parties are unlikely to be as firmly established within the regulatory framework as IOUs and will tend to both fall into the gray area of regulation more often and require more regulatory support than established IOU evaluators.
    - Are there changes to roles and process that should be considered in the context of third-party filings? For example, should evaluators be required to be more closely involved with DRMEC or interacting with CPUC staff in a manner similar to IOU evaluation staff?
- Per comments received and conflicting advice, there is not consensus among stakeholders as to how third parties should proceed under D.19-06-026 and the rules and decision governing Resource Adequacy (RA) to produce their RA numbers.
    - This is an area that will require further clarification from the DRMEC and CPUC.
- Many DR resources cannot sustain their full capacity for 5 hours.
    - Are ex ante results supposed to model a single hypothetical event (with fatigue) or the capacity of the resource if called on each of the RA window hours? The current practice suggests near consensus on the latter.

Appendix A: About OhmConnect

CDA notes: The content in this section is provided for context on the resource being evaluated. OhmConnect wrote it all and CDA has not evaluated any claims made or altered it in any way.

Quick Overview

OhmConnect is a free service that rewards customers for saving energy when it matters most to the grid and the environment. Once a user signs up and connects their utility, they receive a notification about once per week to save energy. These demand response events, which vary between one and four hours, are called *#OhmHours*. If a user saves energy compared to their baseline, they are rewarded through prizes, cash, or other incentives. Rewards are proportional to the amount of energy saved.

Introduction

OhmConnect is a residential demand response (DR) platform with a simple message to its users: *Save Energy. Get Paid*. OhmConnect's objective is to empower individuals to take charge of their energy use, enabling users to participate in wholesale demand response and receive financial compensation for doing so. OhmConnect meets these goals through a software platform (Figure 1) that manages an individual's energy to reduce demand through behavioral and/or automated responses to DR events, called #OhmHours.

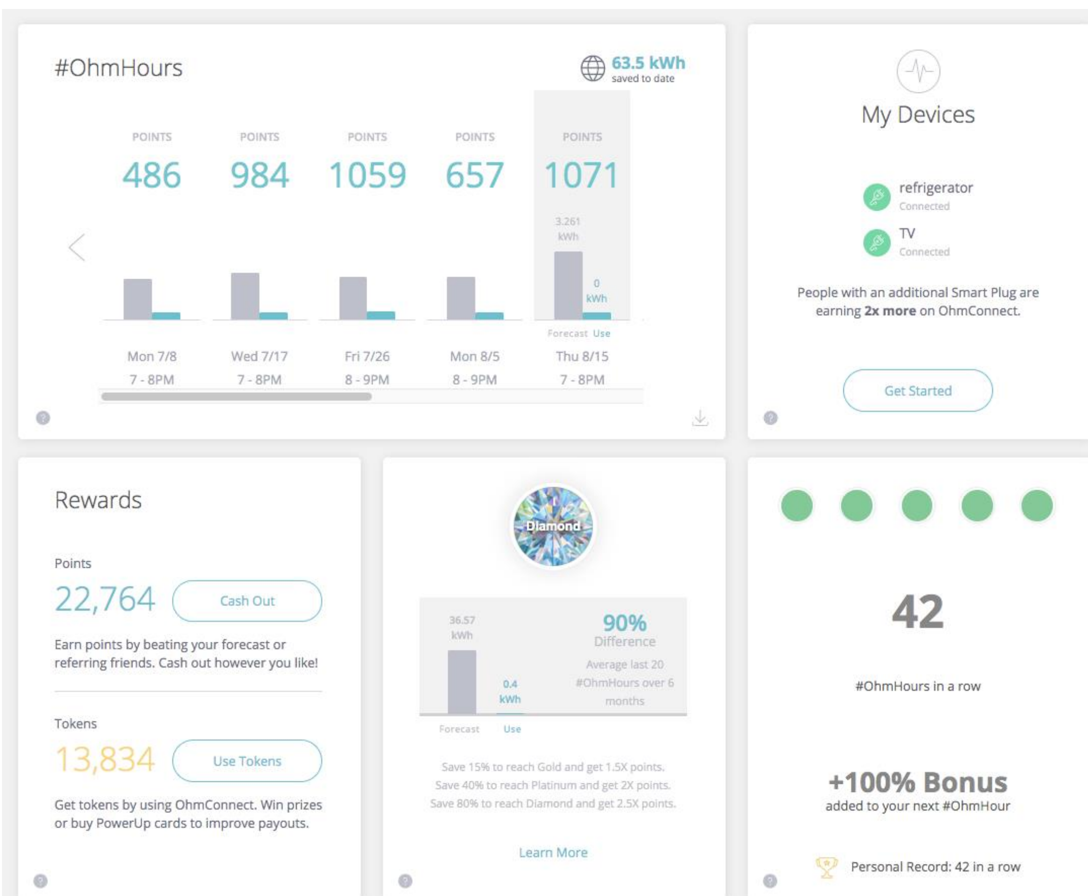


Figure 1. OhmConnect's dashboard, gamifying the demand response experience.

Automated versus Behavioral Demand Response

Users participate in #OhmHours behaviorally, by shutting off devices inside the home, and automatically, through devices controlled by OhmConnect. Although anyone can participate in OhmConnect without a device, empowering users to purchase and connect smart devices in their homes increases the overall value the users receive during #OhmHours by enabling more frequent and prolonged response.

OhmConnect allows users to connect devices, smart plugs and thermostats, to their account during sign-up and any time thereafter. When an #OhmHour occurs, OhmConnect sends automated signals to all connected devices to reduce household energy consumption. These devices are either shut off entirely or their settings are altered (e.g., a thermostat's temperature is increased) remotely and without any manual customer intervention. Users with and without connected devices can also manually shut off or delay use of other appliances that cannot be controlled remotely. OhmConnect is committed to increasing the penetration of automated demand response controls among its users and is continuously experimenting with the rebate structure, cadence, and marketing to increase automation uptake within the OhmConnect platform.

About Demand Response Events, "#OhmHours"

OhmConnect calls DR events about once a week throughout the year. Events are typically called between 4pm and 9pm and are one or two hours long, although they can last as long as four hours.

How OhmConnect Dispatches Users

OhmConnect notifies its users of an upcoming event via email, text message, or both, per user preference. Users are typically notified a full day before the #OhmHour is scheduled (once OhmConnect received its Day Ahead award) and again approximately 30 minutes before the #OhmHour starts. To earn money during an event, a user must use less energy than their forecast. The forecast is shared with dispatched users via the day-ahead and day-of notifications, as well as via their dashboard (Figure 2). This forecast is calculated using the 10-in-10 baseline. Once OhmConnect receives the relevant interval meter data, it will populate the OhmConnect Forecast widget with the user's actual performance (actual usage) during the event.

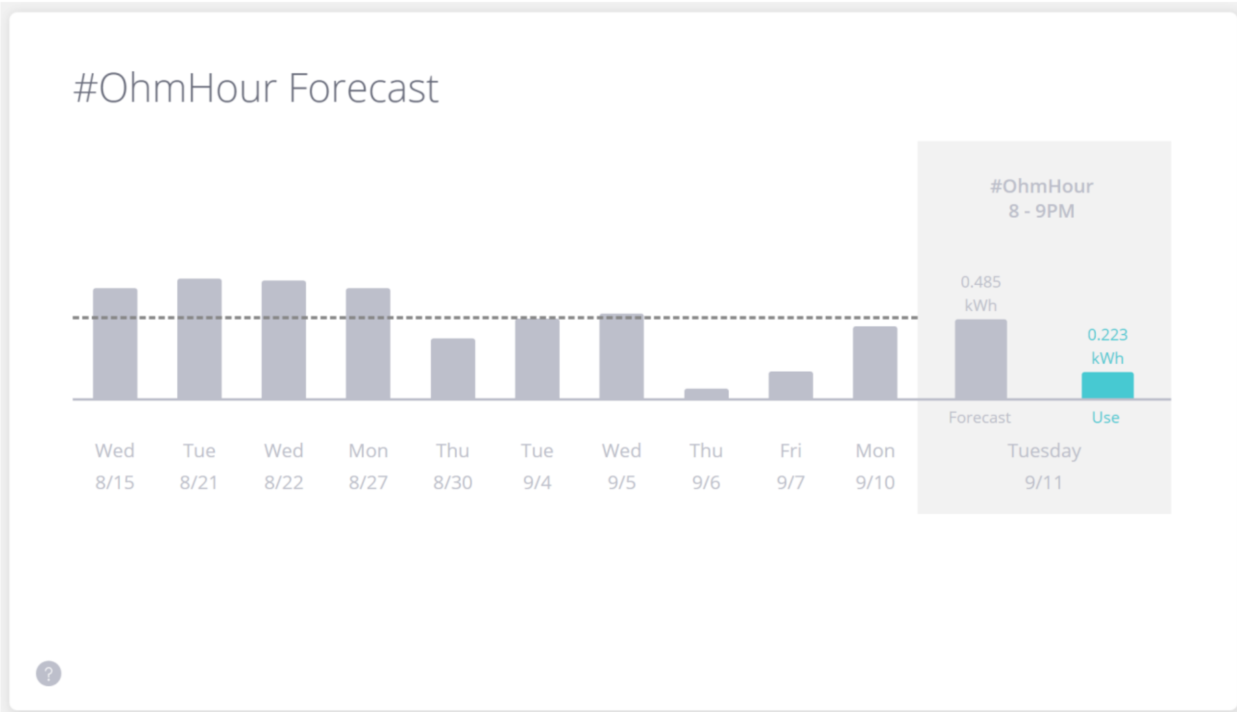


Figure 2. A user can see their energy usage forecast prior to each event.

How OhmConnect Incentivizes Performance

If the user beats their forecast, they are rewarded. The size of the reward is not fixed; it depends on many factors, including the season, the prevailing price of electricity, and whether the event is MEGA (more on MEGA below).

A user is rewarded for each #OhmHour in which they beat their forecast. OhmConnect's primary mechanism of rewards is direct payments via the point system. Users are rewarded for each kilowatt hour (kWh) they reduce against their forecast during an event. If they use *more* energy than their forecast, they lose points. In addition to the energy reduction payments, users that have demonstrated high engagement with OhmConnect can earn bonus points during events. Examples of these types of bonuses include additional points for reaching certain status level within the platform or performing well on multiple #OhmHours in a row. These are discussed at length below, "Rewarding Deep Engagement".

Energy reduction payments are always given on a scale, i.e., the more a user saves, the more they are able to earn. Some users will earn as much as $1000s for energy saving events[6], while many others earn as little as $30. Over the past year, OhmConnect has paid out millions of dollars to users through direct cash payments (via Paypal) and prizes (such as a trip to Disneyland or Iceland).

[6] OhmConnect's largest direct reward winner in the past year earned $3,500; they also gave out a $100,000 grand prize to a single user.

Users can request a payout by redeeming their points once they reach $10. Each point is equivalent to $0.01. Options to redeem points for rewards includes a PayPal cashout, Amazon and Target gift cards, and application of the dollar value to purchases in the OhmConnect store. Users can also donate their points to a charity, or send them to another OhmConnect user (Figure 3, right).

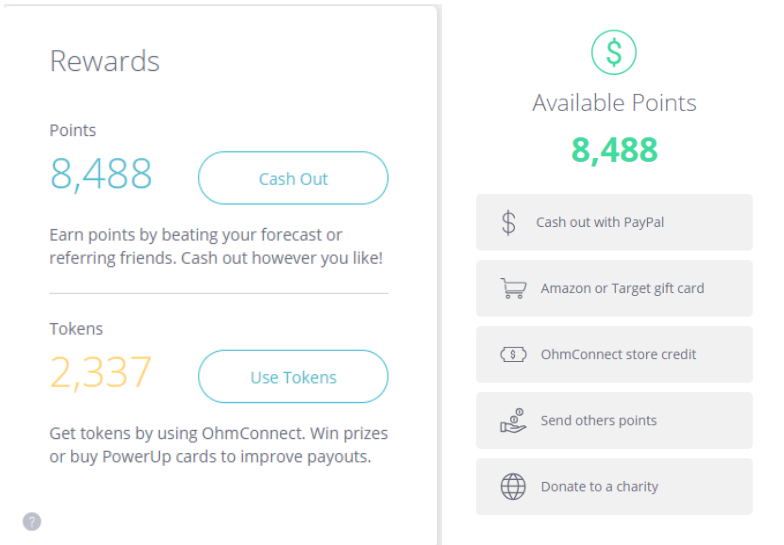


Figure 3. OhmConnect rewards its users using points.

How OhmConnect Rewards Deeper Engagement

OhmConnect motivates users and boosts engagement by using gamification methods, such as streaks and status levels, both of which add a bonus to a user's next #OhmHour earnings. Streaks keep track of how many #OhmHours in a row a user has successfully reduced their energy consumption; status levels are awarded based on a user's average energy reductions during #OhmHours. Below are two examples of gamification methods, but OhmConnect has experimented with dozens of different methods, and the methods that prove to be most successful change over time.

- Streaks: When OhmConnect users beat their forecast during DR events, they add to their streak. For example, positive performance in 5 events in a row leads to a streak of "5". A streak allows users to receive additional points with each successful sequential #OhmHour. Customers receive an extra 2.5% bonus points for each #OhmHour in a row they consume less than their #OhmHour forecast, up to 100% or a streak of 40. In the example provided in Figure 4 (left), below, this specific OhmConnect users has a streak of 69, giving him or her the maximum 100% bonus on their #OhmHour base points. Negative performance during an event results in both negative points and a reset of the streak value to 0.

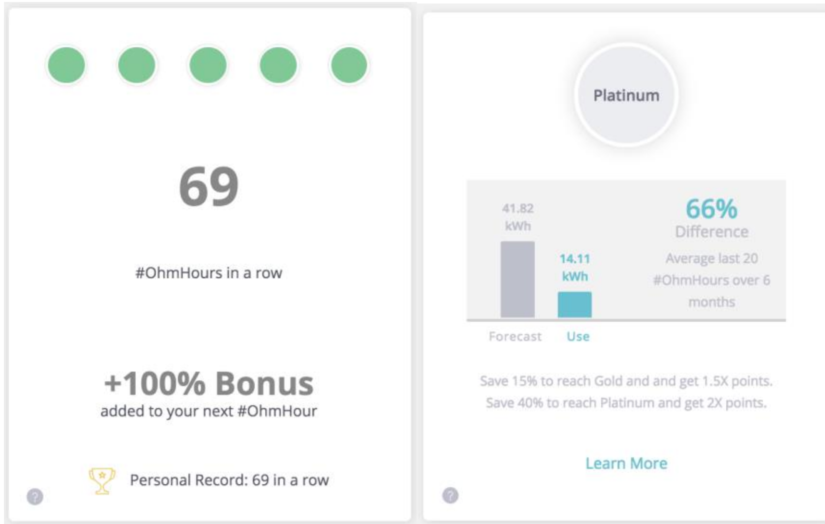


Figure 4. Streaks and Status Widgets

- Tier Status Levels: OhmConnect users receive a different status—Silver, Gold, Platinum, or Diamond—based on how much energy they've saved, on average, during their #OhmHours (Figure 4, right). Users are bumped up to Silver as soon as they share their utility energy usage data with OhmConnect, and they reach Gold, Platinum, and Diamond when they save a set average percentage of energy from their baseline over the previous 20 #OhmHours. Gold, Platinum, and Diamond users have to save an average of 15%, 40%, and 80% or greater, respectively.

Similar to streaks, the tiered status level results in a points multiplier during #OhmHours. Gold, Platinum and Diamond users receive a 50%, 100%, and 150% bonus, respectively, on the base points they've earned during an event.

MEGA Events

To incentivize user performance during times of greatest grid need, OhmConnect increases its rewards and modifies its messaging around "mega events". A MEGA #OhmHour goal is tougher to beat than a normal forecast. However, if beat, the payouts are higher.

The MEGA goal is calculated for each participating user by applying the 10-in-10 baseline *plus* a large discount (such as 75%). So, if a user's #OhmHour baseline is 1kWh, their MEGA goal could be 0.25 kWh. If the customer uses less energy during the event than their MEGA goal, they are paid 3X the points they would otherwise receive for a regular #OhmHour. They are also entered into a weekly sweepstakes for a variety of prizes. If the user *does not* beat their MEGA goal, but *does* use less energy than their normal baseline, they are awarded points according to the non-MEGA scale. They are also not entered into the weekly sweepstakes. In 2018, approximately 160 of 1250 events were MEGA events.

Appendix B: OhmConnect enrollment forecast rationale

CDA note: OhmConnect is a start-up and considers its plans for the future, including recruitment targets and strategies related to their growth forecast to be proprietary information. CDA requested as much support as OhmCounnect could provide in public for their enrollment forecast and the remainder of text in this section is copied verbatim from their response. The text was written from their perspective, so in this appendix "we" and "our" refer to OhmConnect.

OhmConnect has created a set of growth projections for use in the Load Impact Protocols (LIP) report. These growth projections are conservative compared to historical growth. Leveraging our team's deep expertise in consumer engagement and acquisition, we have averaged 12% month over month (MoM) growth during periods when we are investing in marketing. We expect that growth to continue, but have *conservatively* projected our growth to be half of our prior average (at 6.0%) for the purposes of the LIPs.

Background

Our team is filled with senior leaders and executives from across the consumer space, including LinkedIn, Capital One, Zynga, and Facebook. The DNA of the team is focused on different acronyms than the energy sector is used to: LTR (Lifetime Revenue), CAC (Customer Acquisition Cost), DAU (Daily Active Users), and NPS (Net Promoter Score). The majority of the team, including two of the three founders, have consumer background and are outside of the energy space. Cadir Lee, our President, was the CTO of Zynga, a mobile gaming company that acquired hundreds of millions of customers, with high engagement; Zynga is known for popular games such as Words with Friends and Farmville.

OhmConnect is a consumer-first company that takes pride in having a deep relationship with our users. This strong, active relationship has driven the majority of our growth. Over 50% of our user growth has been driven by Organic or Referral based channels, which are scalable as they are not based on digital marketing spend.

Growth Projections

In periods with digital marketing spend, we average 12% MoM growth. We are happy to provide additional supporting data but would need to do so confidentially as growth marketing is extremely sensitive and a key trade secret for OhmConnect.

We do not always invest in additional digital marketing spend depending on a variety of factors. For example, in 2017, our growth far exceeded what we had contracted; in 2018-19, SCE did not pay us for contract delivery for 10 months; in 2019, PGE did not pay us for 5 months during their bankruptcy. Our user base represents ~1% of California households. Our testing in specific cities suggests that we could easily grow 10-20x before reaching any saturation point, equating to 4-5 years based on the LIP projections before hitting that amount.

Recent testing from the summer of 2019 indicates we could grow even faster than 12% MoM as we expand our marketing channels. We have seen promising results from adding traditional advertising to our mix of digital marketing. As consumer companies generate more brand awareness, they often see faster growth which would also accelerate our growth in the near term.

From our team's experience at other consumer companies, growth often accelerates as awareness combined with a great consumer experience leads to adoption from demographics beyond what was originally targeted.

Our conservative model of 6% MoM growth is strongly backed up by our history to date, lack of saturation and ongoing marketing results. Anything less than 6% would be counter to all evidence.

Appendix C: Ex Ante Models

Customers are divided into categories based on:

1. 'Tier', which is an assessment OhmConnect makes concerning the mean load shed of a customer and their reliability. For purposes of this study, the tiers were compressed to Silver, Gold, and Platinum; they have a small number of Bronze and Diamond customers but these were lumped into Silver and Platinum, respectively.
2. Whether or not the customer has devices that OhmConnect can directly call to decrease load.
3. Whether or not the customer has solar generation.

Within each of these customer categories, three models are fit: one to predict the load impact during event hours, one to predict the load impact during non-event hours, and one to predict the baseline. The load impact during event hours is the main quantity of interest so that is what we discuss in this appendix. We will explain the model by example.

Table 21 includes the coefficients for a linear model to predict the load impact per participant. in a given event hour. Consider customers in the 'Silver' tier who do not have solar generation but do have at least one electric device that OhmConnect can control The table shows the coefficients that apply to that set of customers. To illustrate, consider sub-LAP PGSI in a July OhmConnect event, in hour 19 (that is, between 6 and 7 p.m.) when the temperature in the sub-LAP is 85 F.

From the first set of coefficients, prefixed 'slap', we find slapPGSI, which has a coefficient estimate of 1.14 kW.

Then move down to the temp75 coefficient. This parameter multiplies a variable called temp75 that is equal to (Temperature – 75 F) for temperature greater than 75 F, and is zero otherwise. The coefficient is estimated to be 0.03 so for this 85 F event this term contributes 0.03 * (85 – 75) = 0.3 kW.

Assume this is not a 'mega' event, so the 'is_mega' coefficient does not come into play.

We are considering hour 19, so neither then 'hr1718' nor the 'hr21' coefficient has an effect.

Moving down: PGSI is a 'hot' sub-LAP – one in which most customers have air conditioners – but for this customer category the coefficient estimate for the interaction between hot_slap and temp75 is 0 (to two decimal places) so this term contributes 0 kW..

Finally, we move down to the PGSI:july_aug term, which contributes an additional 0.13 kWh.

Adding all of the above together, the prediction for this group of customers in this event at this hour is a load impact of 0.89 + 0.3 + 0.0 + 0.13 = 1.32 kW.

Each customer category has a different set of coefficients.

Model-measurement agreement is illustrated below the table.

Table 21. Coefficient estimates and uncertainties for the event-hour model for Silver customers who do not have solar generation and do not have devices controlled by OhmConnect.

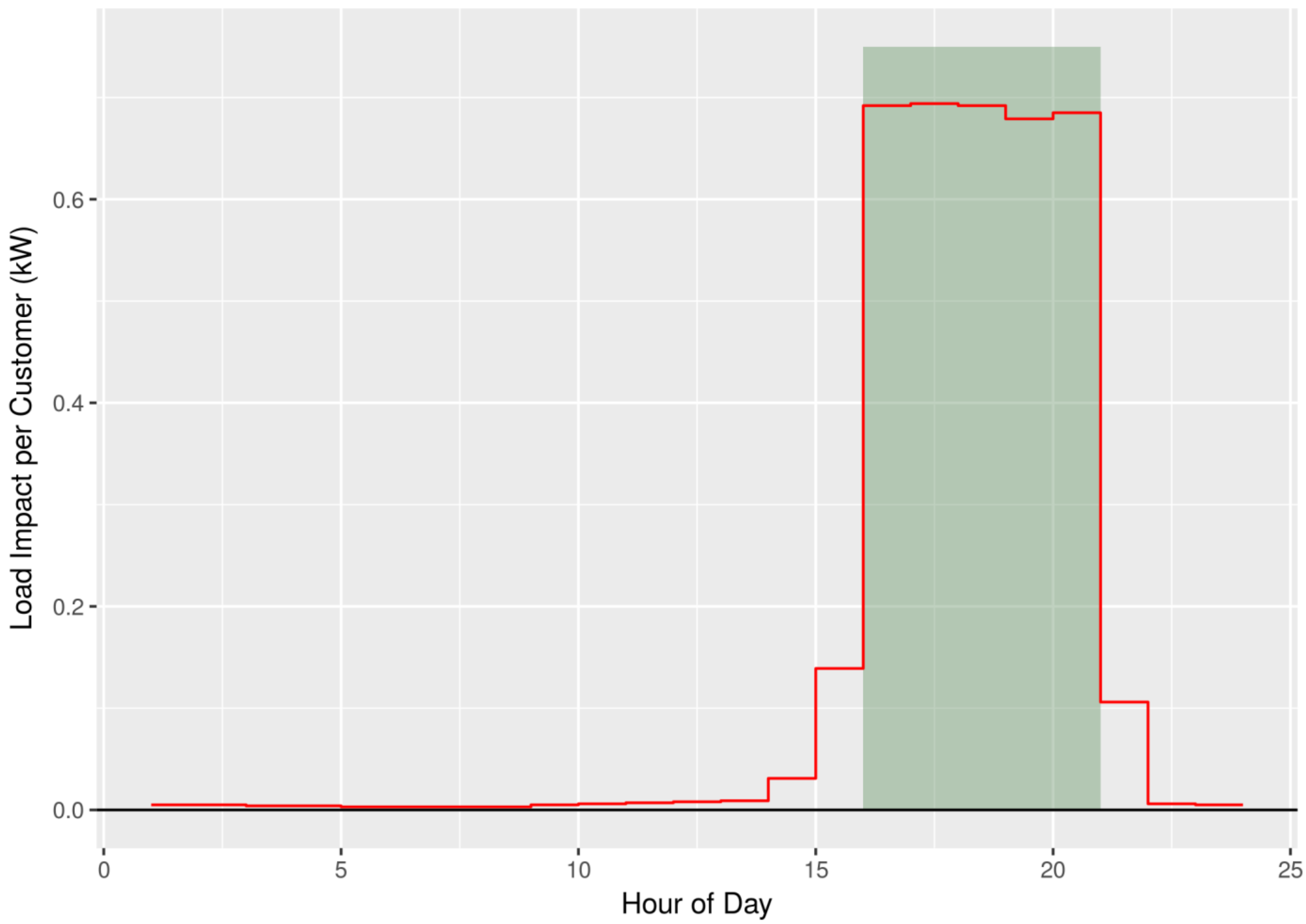
Parameter	Coefficient Estimate	Standard Error	t-statistic	p.value

slapPGCC	0.32	0.14	2.20	0.03
slapPGEB	0.67	0.05	14.54	0.00
slapPGF1	1.38	0.11	13.17	0.00
slapPGFG	0.67	0.13	5.07	0.00
slapPGHB	0.78	0.35	2.21	0.03
slapPGKN	1.16	0.14	8.50	0.00
slapPGNB	0.49	0.13	3.64	0.00
slapPGNC	0.45	0.45	1.00	0.32
slapPGNP	0.96	0.10	9.74	0.00
slapPGP2	0.50	0.07	7.26	0.00
slapPGSB	0.54	0.04	12.30	0.00
slapPGSF	0.16	0.07	2.39	0.02
slapPGSI	1.14	0.10	11.50	0.00
slapPGST	1.19	0.14	8.36	0.00
slapPGZP	0.88	0.11	8.21	0.00
slapSCEC	1.03	0.05	19.59	0.00
slapSCEN	1.29	0.10	13.15	0.00
slapSCEW	0.55	0.05	10.70	0.00
slapSCHD	0.35	0.14	2.44	0.02
slapSCLD	0.57	0.28	2.00	0.05
slapSCNW	0.60	0.10	6.31	0.00
slapSDG1	0.51	0.04	12.14	0.00
temp75	0.03	0.00	6.42	0.00
is_mega	0.00	0.02	0.16	0.87
hr1718	-0.01	0.05	-0.23	0.82
hr21	0.04	0.02	2.17	0.03
temp75:hot_slap	0.00	0.01	-0.66	0.51
slapPGCC:jul_augTRUE	-0.08	0.19	-0.43	0.67
slapPGEB:jul_augTRUE	-0.06	0.06	-1.04	0.30
slapPGF1:jul_augTRUE	-0.26	0.11	-2.51	0.01
slapPGFG:jul_augTRUE	-0.32	0.17	-1.83	0.07
slapPGHB:jul_augTRUE	-0.70	0.44	-1.59	0.11
slapPGKN:jul_augTRUE	-0.16	0.16	-1.00	0.32
slapPGNB:jul_augTRUE	-0.07	0.16	-0.46	0.64
slapPGNC:jul_augTRUE	0.17	0.55	0.31	0.76
slapPGNP:jul_augTRUE	-0.14	0.13	-1.15	0.25
slapPGP2:jul_augTRUE	0.06	0.08	0.72	0.47
slapPGSB:jul_augTRUE	-0.02	0.05	-0.43	0.66

slapPGSF:jul_augTRUE	0.00	0.08	0.01	0.99
slapPGSI:jul_augTRUE	0.13	0.12	1.07	0.29
slapPGST:jul_augTRUE	0.42	0.20	2.12	0.03
slapPGZP:jul_augTRUE	0.24	0.14	1.69	0.09
slapSCEC:jul_augTRUE	0.06	0.07	0.86	0.39
slapSCEN:jul_augTRUE	-0.05	0.13	-0.39	0.69
slapSCEW:jul_augTRUE	0.10	0.06	1.58	0.11
slapSCHD:jul_augTRUE	0.05	0.17	0.27	0.79
slapSCLD:jul_augTRUE	0.64	0.43	1.49	0.14
slapSCNW:jul_augTRUE	0.02	0.12	0.16	0.87
slapSDG1:jul_augTRUE	0.19	0.05	3.57	0.00
slapPGCC:may_octTRUE	0.22	0.24	0.94	0.34
slapPGEB:may_octTRUE	-0.26	0.07	-3.48	0.00
slapPGF1:may_octTRUE	-0.44	0.16	-2.81	0.01
slapPGFG:may_octTRUE	-0.16	0.21	-0.75	0.45
slapPGHB:may_octTRUE	0.06	0.49	0.13	0.90
slapPGKN:may_octTRUE	-0.46	0.25	-1.82	0.07
slapPGNB:may_octTRUE	-0.18	0.21	-0.90	0.37
slapPGNC:may_octTRUE	-0.63	0.65	-0.96	0.34
slapPGNP:may_octTRUE	-0.19	0.17	-1.09	0.27
slapPGP2:may_octTRUE	-0.13	0.11	-1.12	0.26
slapPGSB:may_octTRUE	-0.20	0.08	-2.58	0.01
slapPGSF:may_octTRUE	0.06	0.11	0.58	0.56
slapPGSI:may_octTRUE	-0.51	0.16	-3.18	0.00
slapPGST:may_octTRUE	-0.18	0.27	-0.66	0.51
slapPGZP:may_octTRUE	-0.28	0.19	-1.50	0.13
slapSCEC:may_octTRUE	-0.43	0.09	-4.62	0.00
slapSCEN:may_octTRUE	-0.70	0.18	-3.92	0.00
slapSCEW:may_octTRUE	-0.23	0.09	-2.58	0.01
slapSCHD:may_octTRUE	0.32	0.25	1.30	0.19
slapSCLD:may_octTRUE	-0.87	0.58	-1.51	0.13
slapSCNW:may_octTRUE	-0.35	0.16	-2.23	0.03
slapSDG1:may_octTRUE	-0.10	0.07	-1.53	0.13

As mentioned above, predictions are made not just for the event hours but also for the load impact in other hours. Figure 52 shows an example. The load impact estimates outside the event hours depend somewhat on the details of the same-day adjustment (discussed in the Ex Post section of this report). We see evidence of some load impact in the hour or so before the event, which is expected. Less expected is that there does not appear to be a 'rebound' effect: to the extent that people reduce their electric load by turning up their thermostat during cooling season, one would expect the air conditioning to have to work harder than usual once they set their thermostat back to its normal level, but we do not see this effect in the data. The certainty in this conclusion is low, however.

Figure 52. Hourly prediction of load impact per participant for sub-LAP SCLD, for the 1-in-2 CAISO weather year.



Model-Measurement Comparison

The statistical models are fit to each combination of sub-LAP and customer category, and are fit to every event and customer category. For each event the actual mean load impact can be calculated and compared to the prediction. These will disagree for two reasons: (1) due to factors including finite sample sizes and the variability in the same-day adjustment, the observed load impact for a given set of customers will not equal the true load impact. And (2) even if the observed impact were equal to the true impact, the model could not fit the data perfectly because the model itself is imperfect.

Figure 53. Mean observed load impact for each combination of sub-LAP and customer category, for 2018, versus the prediction from the 2018 ex ante model. Only groups with more than 100 customer-hours of participation are shown.

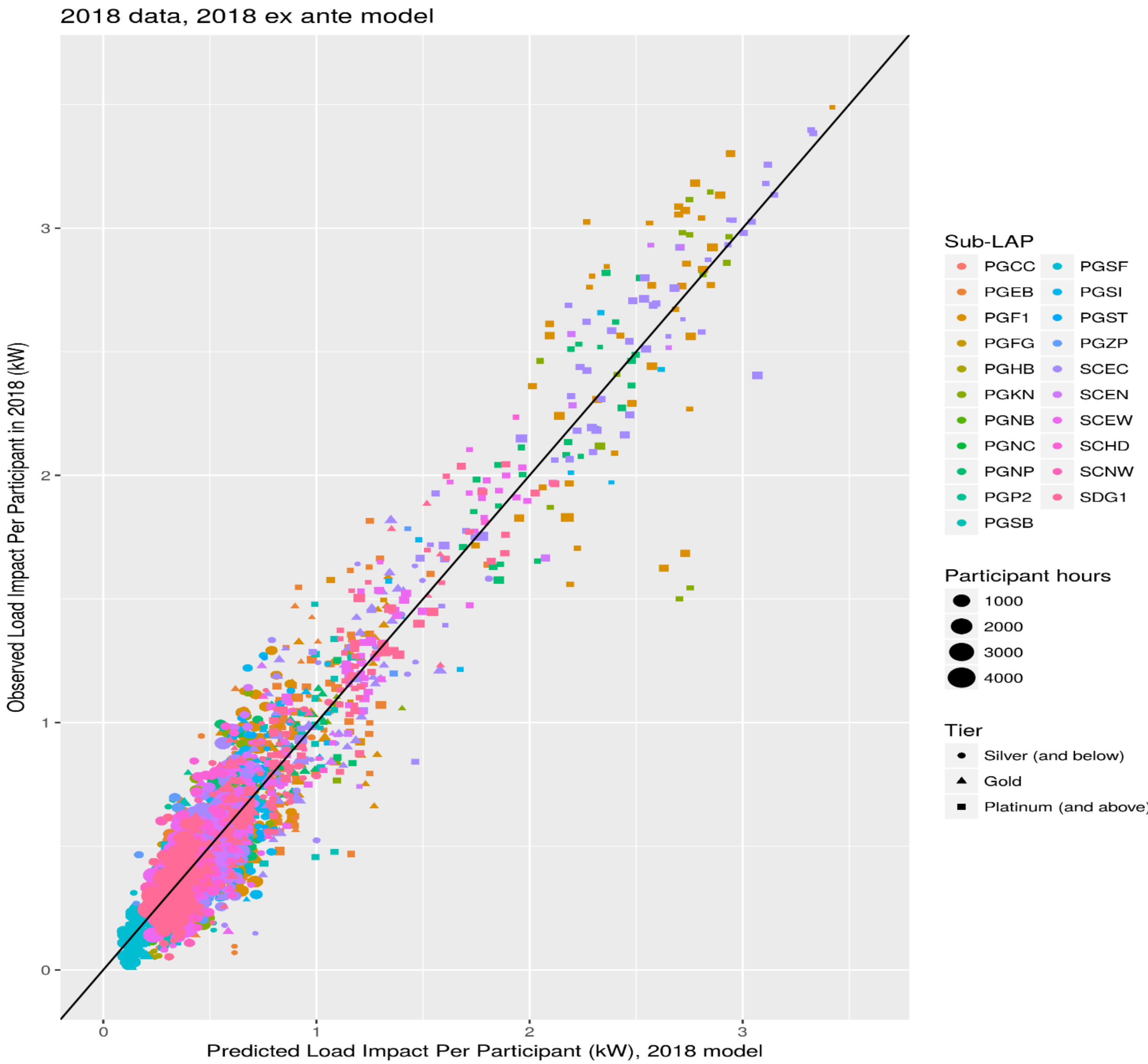
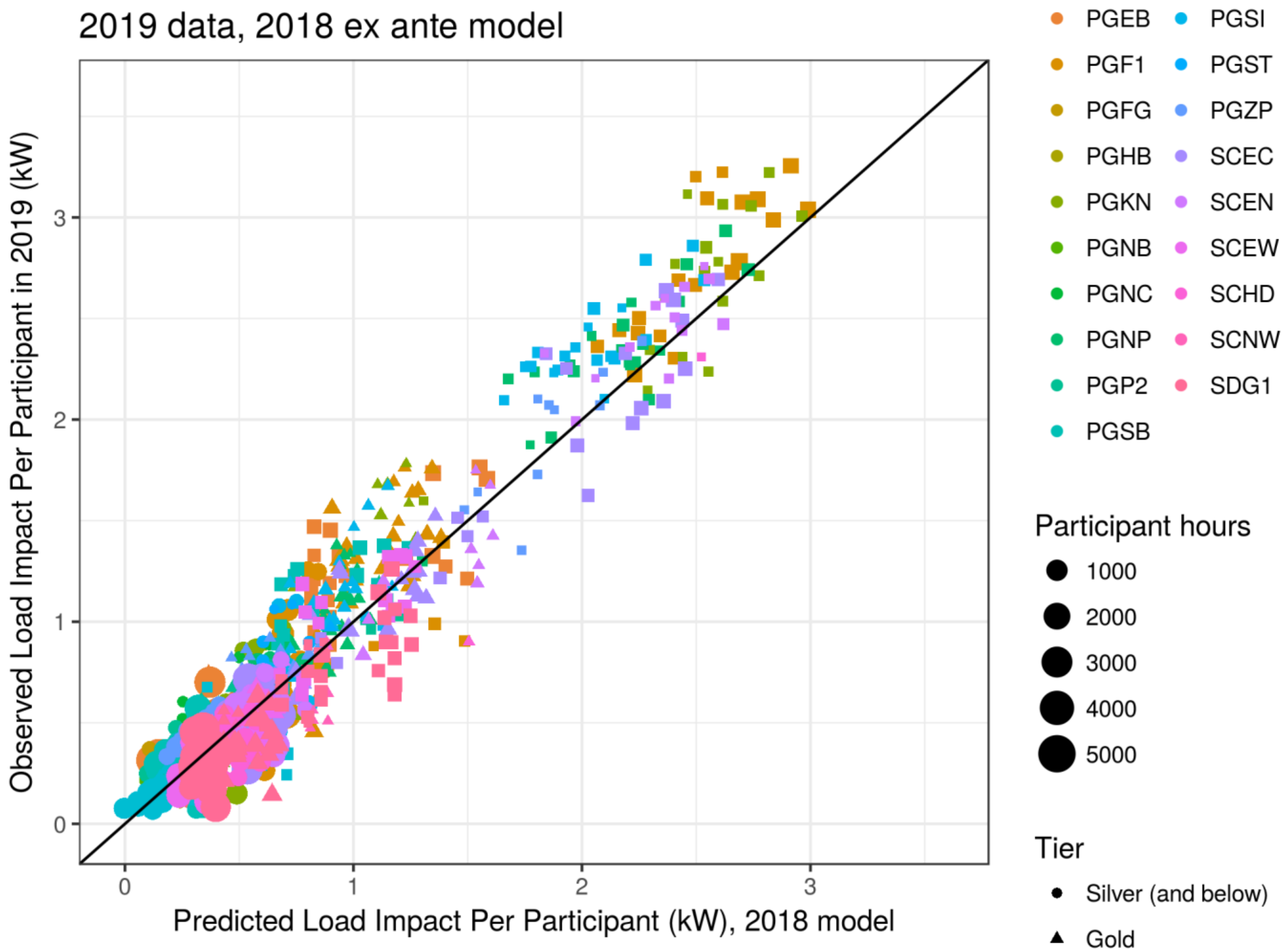


Figure 54. Mean observed load impact for each combination of sub-LAP and customer category, for 2019, versus the prediction from the 2018 ex ante model. Only groups with more than 100 customer-hours of participation are shown.



Appendix D: Sub-LAP and LCA geography

The following images illustrate the geography of California's LCAs (aka Local Reliability Areas on the map) and sub-LAPs. There are two maps of sub-LAPs because PG&E has published a detailed and current map of the sub-LAPs in their territory, while there is not a similar level of detail in the available public map from CAISO for the whole state.

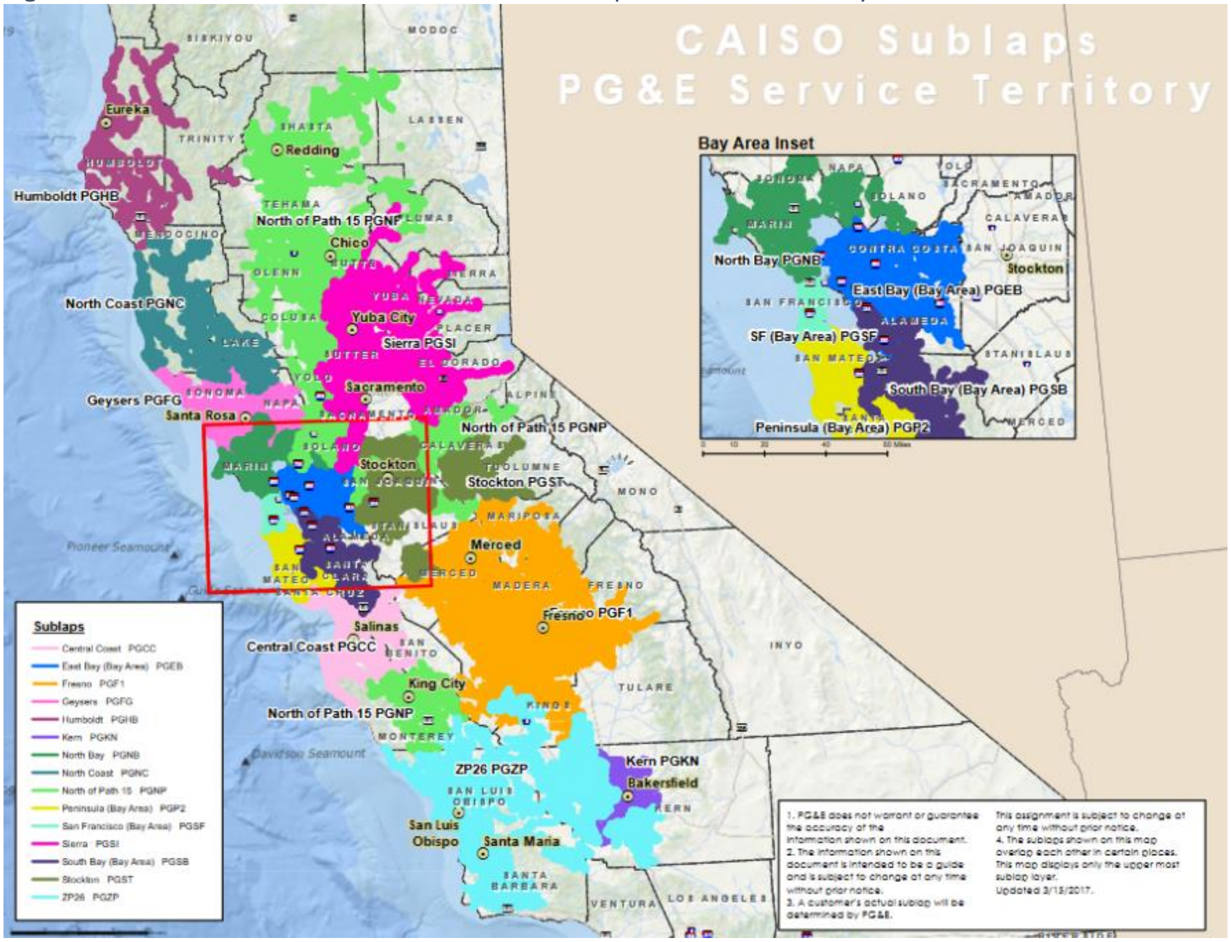
Figure 55. Map of California's LCAs (aka LRAs)

Original at: https://ww2.energy.ca.gov/maps/reliability/Local_Reliability_Areas.pdf

Figure 56. The best statewide map of sub-LAPS CDA could locate from CAISO (it is out of date in PG&E's territory – see below)

Original: https://www.caiso.com/Documents/PDR_RDRRParticipationOverviewPresentation.pdf

Figure 57. A More detailed and current sub-LAP map for PG&E's territory.



Original: https://www.pge.com/pge_global/common/pdfs/save-energy-money/energy-management-programs/demand-response-programs/2018-demand-response/2018-demand-response-auction-mechanism/PGE-Sub-Lap-Map-201703.pdf