Demand Side Analytics
DATA DRIVEN RESEARCH AND INSIGHTS

CALMAC ID: ESP0001.01
FINAL REPORT

Program Year 2024 Load Impact Evaluation of Enersponse's Demand Response Resources

Prepared for Enersponse

By Demand Side Analytics, LLC

April 1, 2025

ACKNOWLEDGEMENTS

Demand Side Analytics Team

- Josh Bode, M.P.P.
- Parker Gauthier, M.S.
- David Pojunas, M.S.

Enersponse Team

- James McPhail
- Steven Gerard
- James Muraca
- Edwin Pham
- Grace Park-Bradbury

ABSTRACT

This report provides an evaluation of Enersponse's demand response (DR) performance in California for the 2024 program year, covering activity from October 2023 through September 2024. Enersponse is a Demand Response Provider (DRP) that works with commercial, industrial, and agricultural customers to temporarily reduce electricity use when the power grid is stressed. In 2024, Enersponse managed more than 300 MW of customer load across over 3,800 sites located in the Pacific Gas and Electric (PG&E), Southern California Edison (SCE), and San Diego Gas & Electric (SDG&E) service territories. These resources were enrolled in statewide DR programs such as the Demand Response Auction Mechanism (DRAM), Demand Side Grid Support (DSGS), and Resource Adequacy (RA), and were dispatched by the California Independent System Operator (CAISO) based on real-time grid needs.

The purpose of this evaluation is to estimate the actual energy savings delivered during demand response events in 2024 (known as ex-post impacts), examine how performance varied across customer types, programs, and locations, and estimate the load reduction potential of Enersponse's resources under typical planning conditions (referred to as ex-ante impacts). The evaluation also aims to understand when and where demand reductions were most effective and to inform expectations for future DR capability.

In 2024, Enersponse resources delivered an average of 34.3 MW of demand reduction per event hour. This value represents an aggregation of average event-level performance across all dispatched sites, accounting for the fact that not all sites were called at once and that dispatch conditions varied across events. The scale of load reduction differed by region and customer sector. While over half of the participating sites were retail and supermarkets, most of the actual load reduction came from fewer than 1,000 agricultural and pumping sites. The variability in grid dispatch was significant: some events involved only a small number of sites, while others—such as the event on August 27—had widespread participation, with over 2,600 sites dispatched and more than 28 MW of reduction achieved.

To estimate the expected load reduction potential under utility planning scenarios, the study also produced ex-ante impact values. Under August "1-in-2" weather conditions and assuming full dispatch of available resources, Enersponse's portfolio was estimated to be capable of providing 32.8 MW of load reduction for 2024. With a conservative 10% annual growth rate in customer enrollment and resource development, this potential is projected to rise to 39.8 MW by 2026.

Several improvements were added in 2024 to increase the accuracy of impact estimates, including the utilization of more detailed weather data such as temperature, rainfall, and reservoir levels, as well as utility-wide load patterns by customer class. A competitive modeling process was used to select the best-performing method for each individual site.

TABLE OF CONTENTS

FIGURES

TABLES

1 EXECUTIVE SUMMARY

Enersponse has been a Demand Response Provider (DRP) serving customers across North America since 2016. This document presents Enersponse's load impact evaluation for their California demand response resources from October 2023 through September 2024, also known as program year 2024 (PY2024). Enersponse was awarded 30 MW of Net Qualifying Capacity (NQC) for October 2023-June 2024, and 40 MW of NQC for July 2024-September 2024. For PY2025, the CPUC awarded Enersponse 10 MW of NQC. Enersponse is seeking **39.8 MW** NQC for 2026, as summarized in Table 1.

Table 1: Ex-Ante Projections for Qualifying Capacity (2026) Under August 1-in-2 Utility Weather Conditions

Utility Name	Number of Customers	MWs
SCE	2,672	28.7
PG&E	1,481	9.2
SDG&E	361	1.9
TOTAL	**4,514**	**39.8**

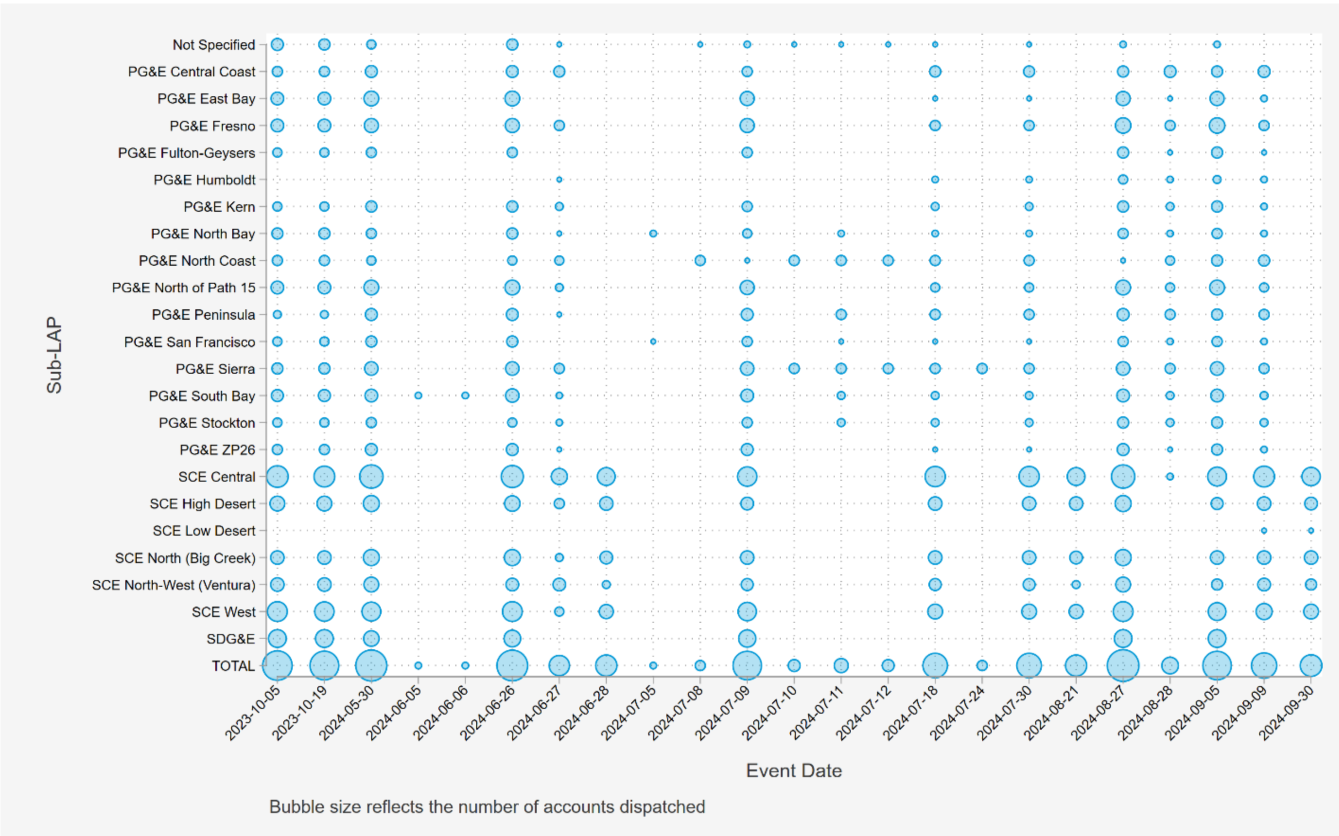
In PY2024, the Enersponse California portfolio included over 300 MW of load, delivering over **34 MW** of load reduction for the average event hour from over 3,800 sites located inside Pacific Gas and Electric (PG&E), Southern California Edison (SCE), or San Diego Gas and Electric (SDG&E) service territory. In 2024, Enersponse actively enrolled its resources in Demand Response Auction Mechanism (DRAM), Demand Side Grid Support (DSGS), and Resource Adequacy and delivered demand reductions in response to CAISO dispatch instructions. The 2026 NQC proposal is based on the actual performance of resources in PY 2024 events and assumes modest growth of 10% per year through 2026.

The primary objectives of this evaluation are to:

1. Estimate the hourly ex-post load impacts.
2. Identify how impacts vary by segment (program, utility, sector, etc.).
3. Identify when and where load impacts are concentrated.
4. Estimate the ex-ante load reduction capability under planning conditions and assess how well these align with ex-post results and prior ex-ante impacts.

At a high level, ex-post impacts estimate the historical demand reductions delivered under the conditions they were called upon by the California System Operator (CAISO), which dispatches resources at different grid locations, known as sub-LAPs, on different days and times depending on the grid conditions and the bids into the market. In 2024, all Enersponse demand response resources were dispatched multiple times. However, CAISO did not dispatch all of the Enersponse resources all at once at any point in time in 2024. Even when resources were dispatched on the same day, the event window varied by program and location due to the CAISO instructions. In addition, the industry mix of the resources dispatched makes substantial difference in assessing performance. Figure 1 highlights the variability in dispatch, displaying the number of Enersponse sites dispatched for each event by sub-LAP. Many events had very low site counts, such as the events on June 5th and July 5th. Of all the events, the event on August 27th had the most complete dispatch, with a total of 2,639 sites participating. Sites are a useful but inadequate proxy for the share of resources dispatched: while Enersponse has over 2,000 supermarket and retail sites combined (over 50% of sites), the majority of the demand reduction is delivered by under 1,000 agriculture and pumping resources.

Figure 1: Dispatch variability by Event Date and Sub-LAP



Because not all resources were called on each event, the load reductions varied substantially by date. Table 2 provides a comprehensive view of Enersponse PY 2024 resources by CAISO sub-LAP grid areas. To produce the table, the average reduction delivered over PY2024 event hours was calculated for each site and then aggregated by sub-LAP. This approach provides an estimate of the performance for the average event hour while accounting for the reality that the sites were not all dispatched simultaneously. Overall, the Enersponse resources delivered **34.3 MW** of resources for the average event hour in 2024.

Table 2: Average Hourly Impacts by Sub-LAP Grid Area[1]

Utility	Sub-LAP	Site Counts	Average Event Temperature (°F)	Per-Site Reference Load (kW)	% Impact	Per-Site Impact (kW)	Aggregate Impact (MW)
PG&E	Unclassified	[redacted]					
	PG&E Central Coast	95	69.28	85.11	36.5%	31.10	2.95
	PG&E East Bay	139	84.35	94.70	6.3%	5.94	0.83
	PG&E Fresno	173	98.56	67.83	2.8%	1.90	0.33
	PG&E Fulton-Geysers	[redacted]					
	PG&E Humboldt	[redacted]					
	PG&E Kern	55	97.52	60.53	6.0%	3.65	0.20
	PG&E North Bay	65	83.94	78.48	7.5%	5.89	0.38
	PG&E North Coast	66	72.73	18.29	-2.5%	-0.46	-0.03
	PG&E North of Path 15	151	94.09	98.30	5.4%	5.31	0.80
	PG&E Peninsula	77	76.62	84.86	5.9%	4.98	0.38
	PG&E San Francisco	46	70.58	176.42	3.4%	6.02	0.28
	PG&E Sierra	109	94.39	80.56	7.0%	5.62	0.61
	PG&E South Bay	88	83.04	99.24	4.6%	4.61	0.41
	PG&E Stockton	[redacted]					
	PG&E ZP26	55	84.52	85.23	1.6%	1.40	0.08
SCE	Unclassified	[redacted]					
	SCE Central	988	92.50	91.11	22.0%	20.04	19.80
	SCE High Desert	199	93.83	44.41	10.0%	4.46	0.89
	SCE Low Desert	[redacted]					
	SCE North (Big Creek)	218	93.20	85.96	13.1%	11.23	2.45
	SCE North-West (Ventura)	184	77.53	63.67	4.1%	2.60	0.48
	SCE West	620	76.99	74.32	3.3%	2.44	1.52
	Unclassified	[redacted]					
SDG&E	SDG&E	298	78.40	56.24	6.2%	3.48	1.04
TOTAL		**3,811**	**84.32**	**80.70**	**11.2%**	**9.01**	**34.33**

[1] Due to variability in CAISO dispatch, the average reduction delivered over PY2024 event hours was calculated for each site, including all hours the sites were dispatched. The site level impacts were then aggregated by sub-LAP.

Table 3 shows the ex-post impacts for each event day but is less useful for understanding the resource capability due to fact that not all resources were dispatched on the same day and event window varied by program and location on event days. Of all the events, the event on August 27th had the most complete dispatch with 2,639 sites participating and 28.2 MW of load reduction. However, the August 27 events did not include dispatch of Enersponse DSGS resources.

Table 3: Average Hourly Impacts by Utility and Event Day[1]

Date	PGE Impact (MW)	PGE Sites	SCE Impact (MW)	SCE Sites	SDGE Impact (MW)	SDGE Sites
10/5/2023	-0.13	461	8.04	1,291	1.10	247
10/19/2023	1.48	461	5.75	1,175	1.07	247
5/30/2024	2.21	818	17.68	1,569	0.26	151
6/5/2024	[redacted]		0.00	0	0.00	0
6/6/2024	[redacted]		0.00	0	0.00	0
6/26/2024	5.47	859	16.60	1,379	0.61	206
6/27/2024	1.07	135	2.38	294	0.00	0
6/28/2024	[redacted]		16.04	514	0.00	0
7/5/2024	[redacted]		0.00	0	0.00	0
7/8/2024	[redacted]		0.00	0	0.00	0
7/9/2024	5.34	693	1.81	872	0.48	233
7/10/2024	[redacted]		0.00	0	0.00	0
7/11/2024	0.18	53	0.00	0	0.00	0
7/12/2024	[redacted]		0.00	0	0.00	0
7/18/2024	2.73	184	22.38	789	0.00	0
7/24/2024	[redacted]		0.00	0	0.00	0
7/30/2024	3.09	184	19.08	788	0.00	0
8/21/2024	0.00	0	17.98	510	0.00	0
8/27/2024	5.97	800	21.00	1,570	1.27	269
8/28/2024	3.71	191	[redacted]		0.00	0
9/5/2024	4.65	810	2.93	722	1.57	271
9/9/2024	3.19	215	20.16	870	0.00	0
9/30/2024	0.00	0	14.19	592	0.00	0

[1] Due to variability in CAISO dispatch, the average reduction per event involved two steps: a) For each site, estimate the average impacts delivered per event date across the hours dispatched, and b) sum the impacts per event date for the sites dispatched by CAISO.

The magnitude and performance of DR resources under planning conditions are referred to as ex-ante impacts and are informed by performance during historical events. Ex-ante impacts estimate the load reduction capability under planning conditions, assuming full dispatch of all resources, and standardizing the hours of dispatch. Table 3 summarizes the Enersponse ex-ante impacts, or DR capability, broken out by utility. The 2024 ex-ante value reflects the resources Enersponse had in place in August 2024. The estimates for future years assume 10% annual growth in resources until 2031, with a 3% growth rate thereafter. Rather than project large amount of growth, Enersponse has recently adopted a practice of grounding ex-ante impacts on resources it currently controls and forecasting very modest year-on-year growth. As result, the projections are lower than in prior evaluation reports.

Table 4: Aggregate Ex-Ante Impacts (MW) – August Utility 1-in-2 Weather (Worst Day)

Forecast Year	SCE 1-in-2	PG&E 1-in-2	SDG&E 1-in-2	Total
2024	23.69	7.58	1.53	32.80
2025	26.06	8.34	1.69	36.09
2026	28.66	9.17	1.86	39.70
2027	31.53	10.09	2.04	43.66
2028	34.68	11.10	2.24	48.03
2029	38.15	12.21	2.47	52.83
2030	41.97	13.43	2.72	58.12
2031	46.16	14.77	2.99	63.93
2032	47.55	15.22	3.08	65.85
2033	48.97	15.67	3.17	67.82
2034	50.44	16.14	3.27	69.86
2035	51.96	16.63	3.37	71.96

There were several key findings resulting from the analysis of Enersponse's PY2024 resources. Table 5 presents these findings, broken out into four major categories.

Table 5: Key Findings

Category	Key Findings
Historical Performance	▪ Enersponse resources enrolled by September 30, 2024, delivered 34.3 MW for the average event hour in 2024. ▪ When adjusted for planning conditions, the resource capability was 32.8 MW, slightly lower than the projections made for the portfolio level in 2023, and similar to the NQC awarded to Enersponse for PY2024.
Performance Characteristics	▪ The Pumping & Agricultural sectors led in load reductions, especially during the late summer. Retail and supermarkets contributed moderate reductions, while other sectors had a minimal impact. ▪ Event impacts were typically consistent across event hours with minimal pre- and post- event increases in consumption, and often persist beyond the event dispatch hours.
Operations and Analysis	▪ DSA recommends that Enersponse continues the practice of grounding ex-ante impacts on resources it currently controls and to forecast modest year-on-year growth (e.g. 10%) ▪ While not directly part of the evaluation, DSA identified a gap between the NQC awarded to Enersponse in PY2024 and the nominations to CAISO. The nominations to CAISO are being updated to match the load reduction capability identified in the evaluation. ▪ The PY2024 analysis included substantial data and modeling updates that improved the accuracy of the impact estimates, including model tournaments to identify the best model for site, and inclusion of additional data sources such as rainfall, reservoir levels, and utility wide aggregate load profiles by rate class.
Forward Projections	▪ Projected enrollment growth of 10% per year is expected to increase aggregate impacts from 32.8 MW in 2024 to over 39.7 MW by 2026. ▪ Enersponse resources that were historically enrolled in DRAM will need to shift to other options such as CBP, DSGS, or resource adequacy contracts with other parties (i.e., utilities, CCA's, or other aggregators)

2 INTRODUCTION

This report presents the results of Enersponse's demand response (DR) participation for the 2023-2024 season. As part of its regulatory obligations to the California Public Utilities Commission (CPUC), Enersponse engaged Demand Side Analytics (DSA) to conduct an independent, third-party evaluation of its DR portfolio. The evaluation was designed to provide a comprehensive assessment of observed demand response impacts during the October 1, 2023, to September 30, 2024, DR season (PY2024) while also examining expected impacts under planning conditions to forecast potential load reductions. Enersponse engages in three key DR pathways in California: The Demand Response Auction Mechanism (DRAM), Resource Adequacy (RA), and the Demand Side Grid Support (DSGS) program. These programs provide financial incentives for customers to reduce electricity usage during periods of high system demand or grid emergencies. Table 6 provides a brief overview of each program.

Table 6: Program Descriptions

Program	Description
DRAM	A competitive procurement mechanism where third-party aggregators bid DR capacity into the CAISO wholesale market. DR resources are treated like generation assets and dispatched as needed.
RA	A regulatory framework ensuring that utilities secure sufficient capacity to meet forecasted demand. DR resources can qualify as RA capacity and must be available for dispatch during peak periods.
DSGS	A DR initiative designed to provide load reductions during critical grid events, particularly extreme weather or emergencies. Participants respond to CAISO-directed dispatch events.

Enersponse participates in each DR pathway with a distinct mix of customer types, with pumping stations, supermarkets, and retail establishments among the most prevalent participants. Moreover, Enersponse's resources are geographically diverse, with participants under the jurisdictions of the three California investor-owned utilities (IOUs) – Pacific Gas & Electric (PG&E), Southern California Edison (SCE), and San Diego Gas & Electric (SDG&E). This assortment in customer composition resulted in varying climate conditions and weather sensitivity levels. Some customer segments exhibit strong correlations between temperature and energy consumption, whereas others demonstrate minimal weather sensitivity, requiring tailored analytical approaches to assess demand response impacts. In addition, pumping and agricultural customer loads and load reduction capability are influenced by California water reservoir levels and rainfall, with more pumping occurring in dryer years.

To ensure robust impact estimation, DSA implemented a site-level model tournament, systematically evaluating multiple methodologies to identify the most reliable reference load estimation techniques. The selected site-level models enabled the measurement of individual-level impacts, which were subsequently aggregated to utility, Sub-Load Aggregation Point (sub-LAP), sector, and program levels.

For this evaluation, DSA requested and analyzed multiple datasets from Enersponse, including:

- Advanced Metering Infrastructure (AMI) data sourced from each utility, providing consumption records for participants.
- Service agreement number (SAN)-level characteristics, categorizing participants by industry, location, and operational attributes to enable segmentation analysis.
- Event dispatch schedules detailing the timing, duration, and participation criteria for each DR event for each program.

2.1 KEY RESEARCH QUESTIONS

For clarity, this evaluation's key research questions can be found in Table 7.

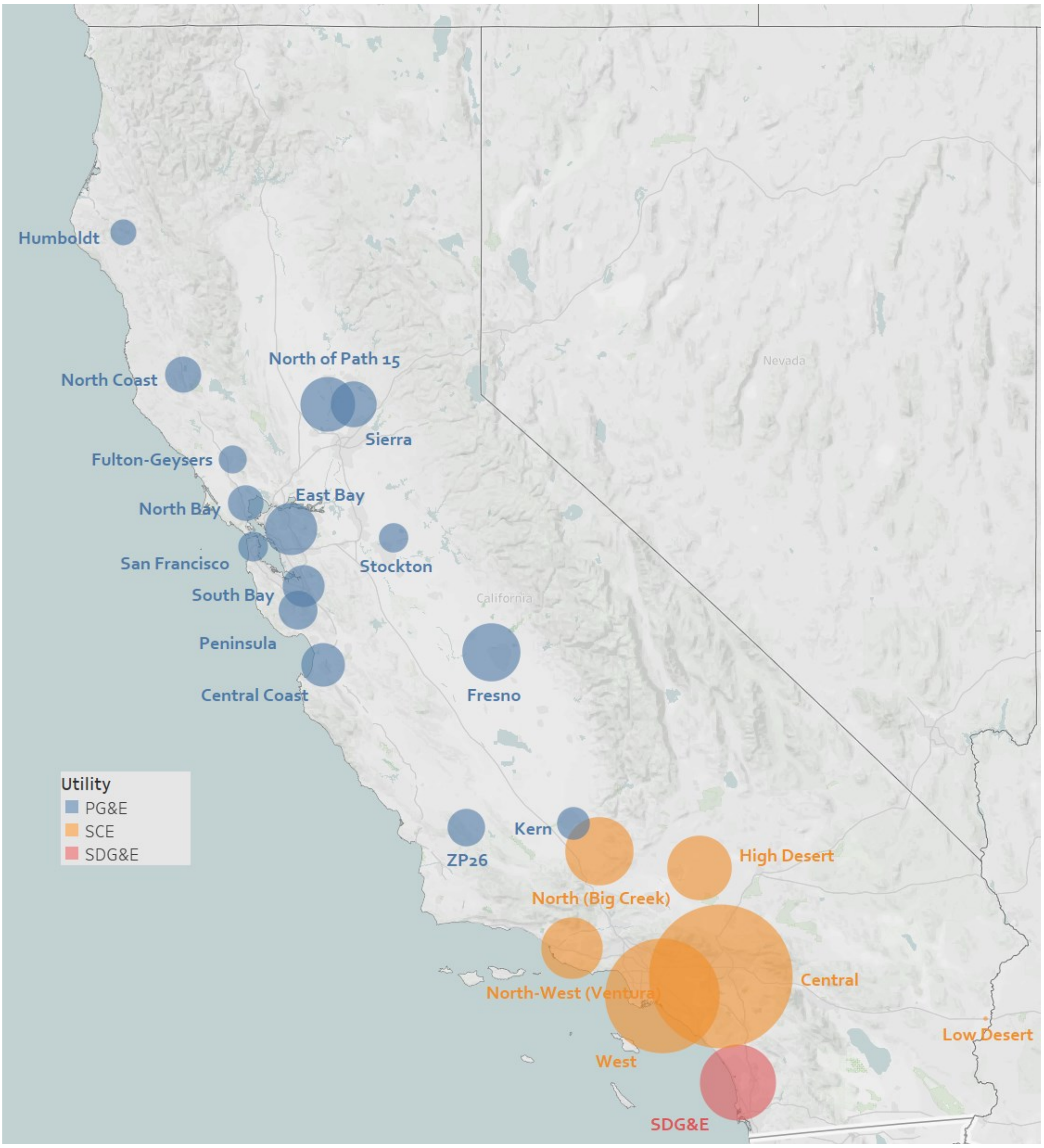
Table 7: Research Questions

	Research Question
1	What were the load impacts from Enersponse's DR resources observed during PY2024?
2	How do impacts vary by segment (program, utility, sector, etc.)?
3	When and where are load impacts concentrated; what sites are the key drivers of the aggregate impacts?
4	What is the ex-ante load reduction capability under planning conditions? How well do these align with ex-post results and prior ex-ante forecasts?

2.2 PARTICIPANT CHARACTERISTICS

During the PY2024 season, Enersponse had 4,131 unique service agreements participate as part of their DR resources. Figure 2 maps these customers across the sub-LAPs of California's three IOUs, with the highest concentration in SCE's territory, followed by PG&E's, and a smaller presence in SDG&E's service area. This geographic distribution highlights that different weather patterns were experienced by sites, influencing electricity consumption and responsiveness to DR events. California's climate varies significantly across regions, ranging from the cooler coastal zones of northern California to the arid and warmer areas of the Central Valley and inland Southern California. Participants in PG&E's territory span multiple climate zones, including cooler coastal areas in the Bay Area and hotter inland regions such as the Central Valley. In contrast, customers in SCE's and SDG&E's territory are more concentrated in southern California, which has warm and dry climate conditions.

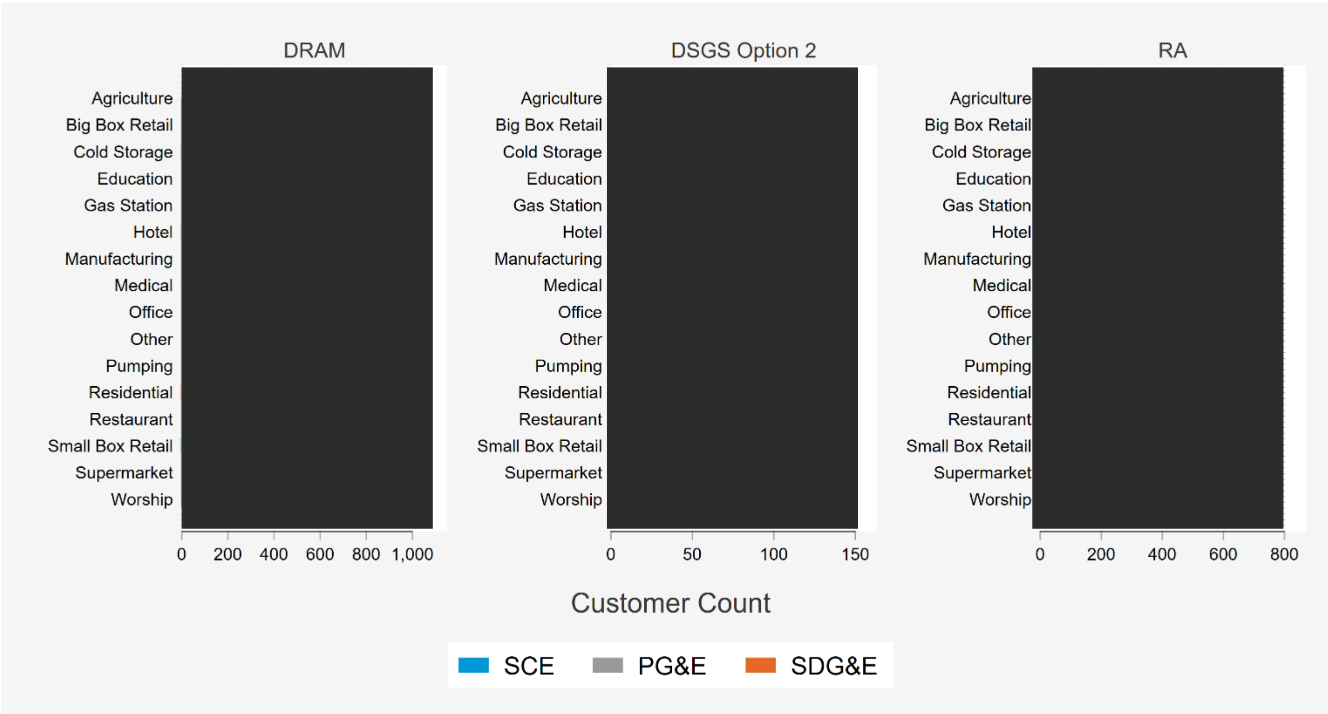
Figure 2: Geographic Distribution of Customers



**Bubble size reflects the number of sites located in the highlighted area*

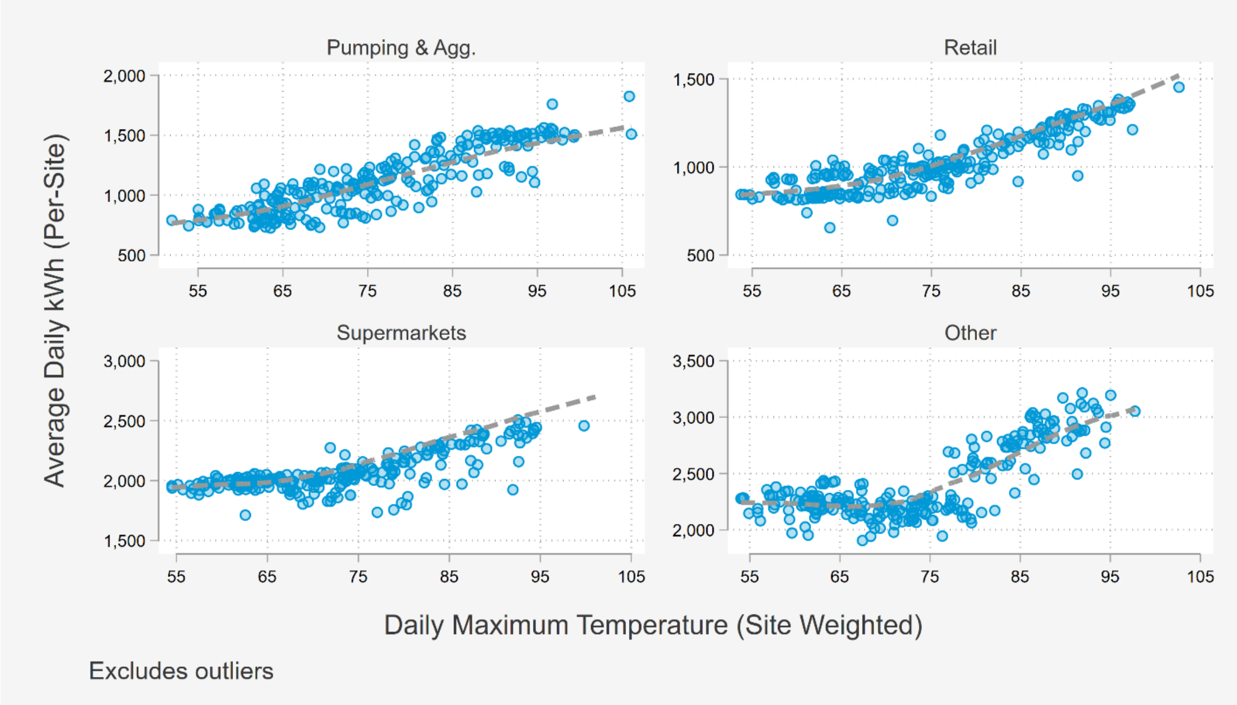
Figure 3 highlights the diversity of customers across the different DR programs, showing how participation varies by industry, utility, and program. While supermarkets, retail, and pumping customers make up a significant share of participants, each program attracts a distinct mix of industries. Some programs see more concentration in specific sectors, while others feature broader representation across industries. Participation also varies by utility territory, with certain programs dominated by customers from a single utility while others draw from multiple regions.

Figure 3: Counts by Sector, Program, and Utility



The diversity in customer participation underscores the importance of understanding how temperature and water levels influence electricity consumption. Figure 4 illustrates the average relationship between daily consumption (kWh) and maximum daily temperature by sector. While all industries exhibit some degree of temperature sensitivity, the strength of this relationship differs, reflecting the complex relationship between environmental conditions and demand. Subsequent sections of this analysis explore this relationship further.

Figure 4: Weather Sensitivity Per Site by Sector



Demand Side Analytics
DATA DRIVEN RESEARCH AND INSIGHTS

2.3 SYSTEM PEAKING CONDITIONS

Figure 5 illustrates how California's peak loads exhibit a significant concentration within a limited number of hours, highlighting the importance of targeted DR interventions. This plot arranges system demand in descending order, with the highest load hours appearing first, visualizing the skewed distribution of electricity consumption. Throughout the 2023-2024 analysis period, the CAISO system load rarely exceeded 40,000 MW, highlighting the fact that demand mostly occurs at levels well below peak conditions. However, during extreme weather events and periods of heightened electricity usage, demand can surge substantially. The highest system load recorded during the analysis period occurred on September 5th, reaching 47,759 MW.

Figure 5: CAISO System Load Duration Curves

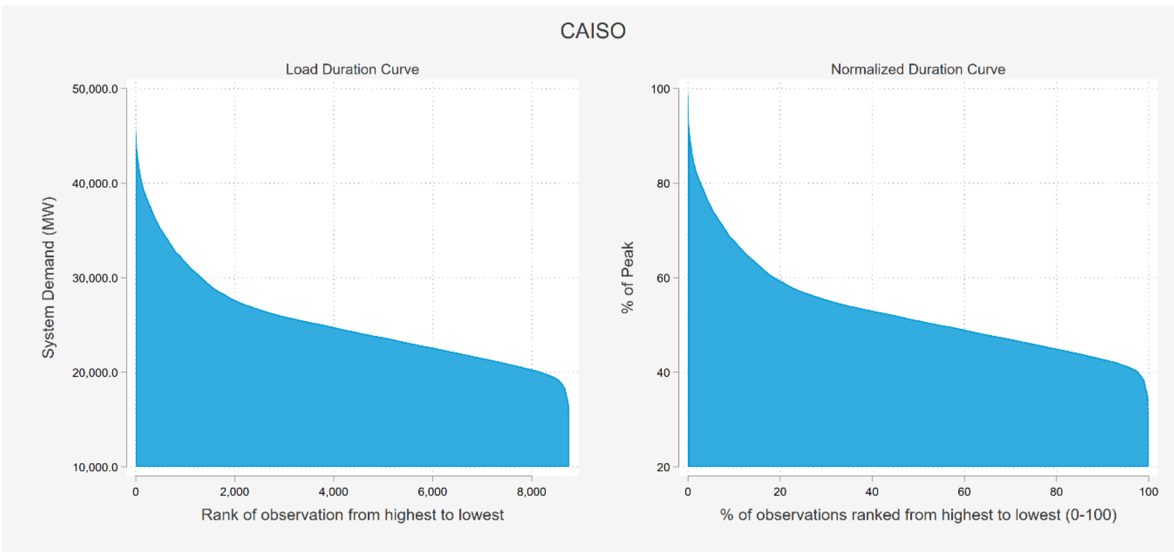
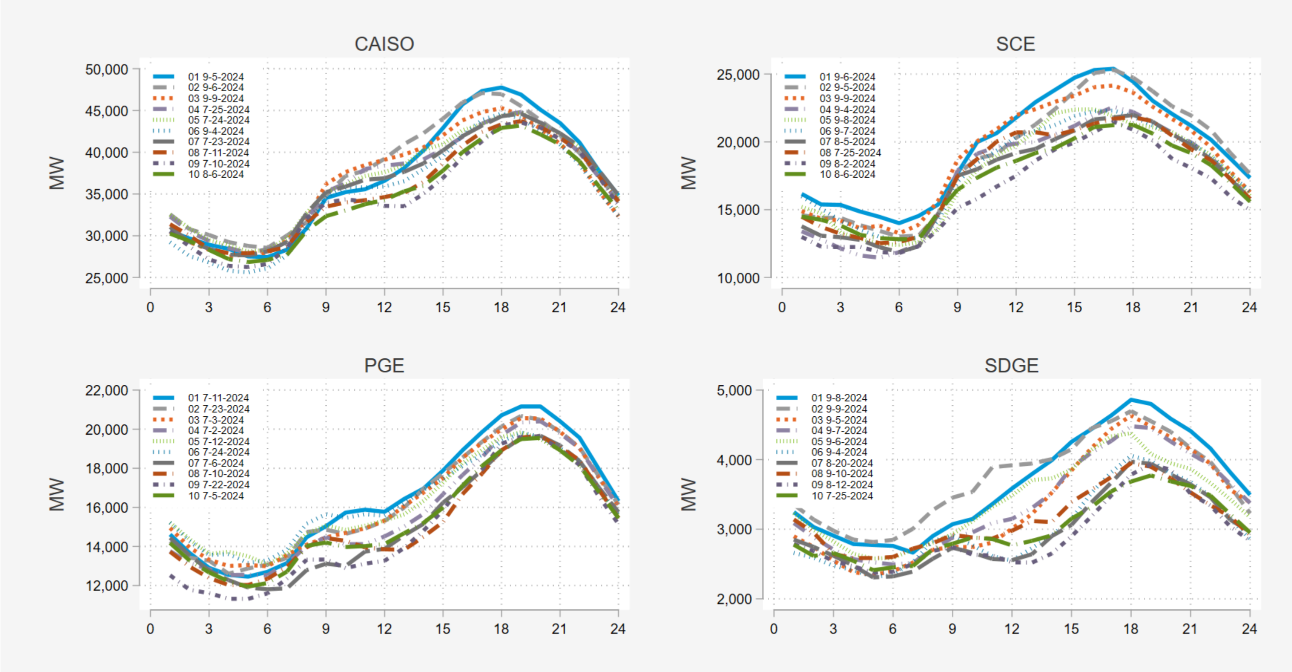


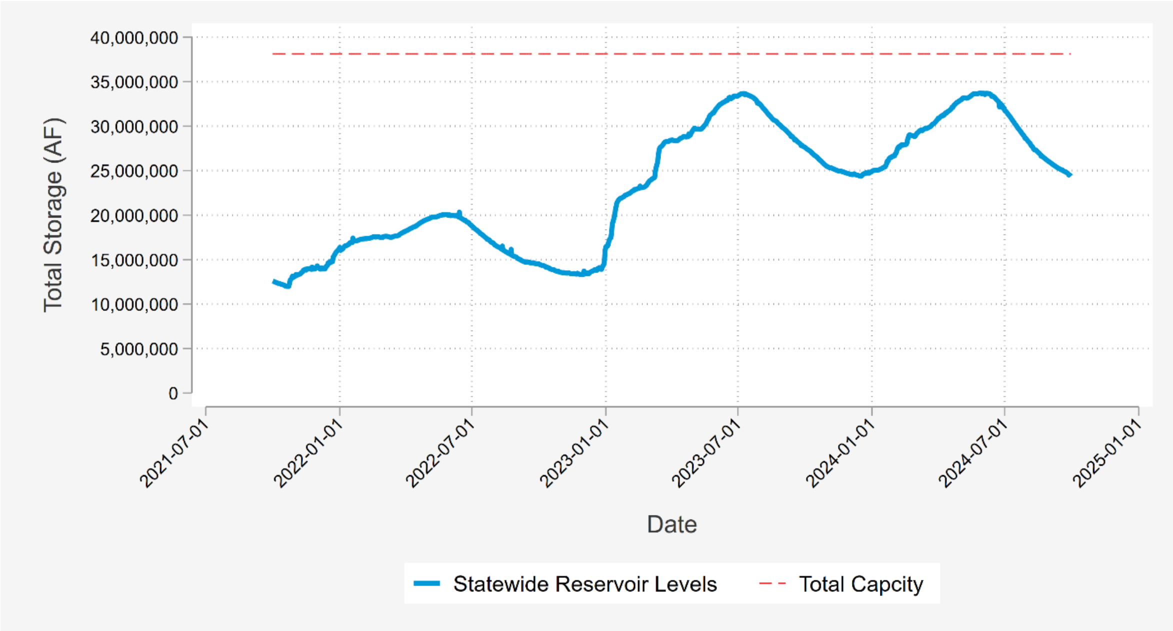
Figure 6 compares the ten days with the highest loads for each utility to the days with the largest overall system demand in 2024. While there is significant overlap between these sets of days, some variations exist due to localized demand patterns and regional weather conditions. Many of these high-demand days were selected as DR events.

Figure 6: Top Ten Load Days by Utility



Since a large portion of Enersponse's customer base and demand reductions comes from pumping customers, DSA investigated the relationships between customer loads, reservoir levels, and precipitation levels. Figure 7 displays the statewide reservoir water levels in California since the summer of 2021. Generally, statewide water levels have remained high since 2023. Higher water levels are generally associated with lower pumping loads.

Figure 7: Statewide Reservoir Levels over Time



2.4 EVENT DISPATCH

Event days and dispatch durations varied by program and location in PY2024. Overall, participation was greatest in the DRAM and RA program, while DSGS participation was low early in the season but increased for later events. DRAM events were uniformly scheduled, each lasting two hours and occurring at the same time each event day. RA events were typically four hours long, with variation in start and end times throughout the season. DSGS events had the greatest variability, with dispatch durations ranging from one to five hours. For additional context, the tables below provide a detailed summary of event dispatch and the conditions experienced by participants in PY2024, broken out by program.

Table 8: DRAM Event Conditions

Program	Event Times	Dispatch Length	Sites Dispatched	Average Event Temperature (°F)	Statewide Reservoir Levels
DRAM	10/05/2023 16:00 to 18:00	2 Hours	708	86.21	72%
	10/19/2023 16:00 to 18:00	2 Hours	708	82.12	69%
	05/30/2024 16:00 to 18:00	2 Hours	1,368	82.31	88%
	06/26/2024 16:00 to 18:00	2 Hours	1,570	82.48	85%
	06/28/2024 16:00 to 18:00	2 Hours	514	86.08	85%
	07/09/2024 16:00 to 18:00	2 Hours	924	85.07	81%
	07/18/2024 16:00 to 18:00	2 Hours	540	89.87	79%
	07/30/2024 16:00 to 18:00	2 Hours	540	88.35	75%
	08/21/2024 16:00 to 18:00	2 Hours	511	91.15	70%
	08/27/2024 16:00 to 18:00	2 Hours	1,577	87.84	69%
	09/05/2024 16:00 to 18:00	2 Hours	1,043	92.64	67%
	09/09/2024 16:00 to 18:00	2 Hours	620	97.12	67%
	09/30/2024 16:00 to 18:00	2 Hours	592	86.18	64%

Table 9: DSGS Event Conditions

Program	Event Times	Dispatch Length	Sites Dispatched	Average Event Temperature (°F)	Statewide Reservoir Levels
DSGS Option 2	06/05/2024 16:00 to 17:00	1 Hours	4	86.72	88%
	06/06/2024 16:00 to 17:00	1 Hours	4	72.63	88%
	06/27/2024 16:00 to 18:00	2 Hours	141	79.85	85%
	07/05/2024 17:00 to 19:00	2 Hours	4	78.55	82%
	07/08/2024 18:00 to 20:00	2 Hours	25	58.46	82%
	07/10/2024 18:00 to 20:00	2 Hours	23	97.63	81%
	07/10/2024 17:00 to 20:00	3 Hours	25	59.15	81%
	07/11/2024 19:00 to 20:00	1 Hours	18	87.80	81%
	07/11/2024 18:00 to 20:00	2 Hours	23	102.60	81%
	07/11/2024 17:00 to 20:00	3 Hours	25	70.85	81%
	07/11/2024 16:00 to 21:00	5 Hours	25	59.40	81%
	07/12/2024 19:00 to 20:00	1 Hours	23	95.70	80%
	07/12/2024 16:00 to 20:00	4 Hours	25	59.93	80%
	07/18/2024 17:00 to 19:00	2 Hours	190	79.32	79%
	07/24/2024 19:00 to 20:00	1 Hours	23	91.75	77%
	07/30/2024 17:00 to 19:00	2 Hours	190	76.90	75%
	08/28/2024 16:00 to 19:00	3 Hours	197	79.79	69%
	09/05/2024 18:00 to 20:00	2 Hours	42	70.50	67%
	09/09/2024 16:00 to 18:00	2 Hours	221	80.75	67%

Table 10: RA Event Conditions

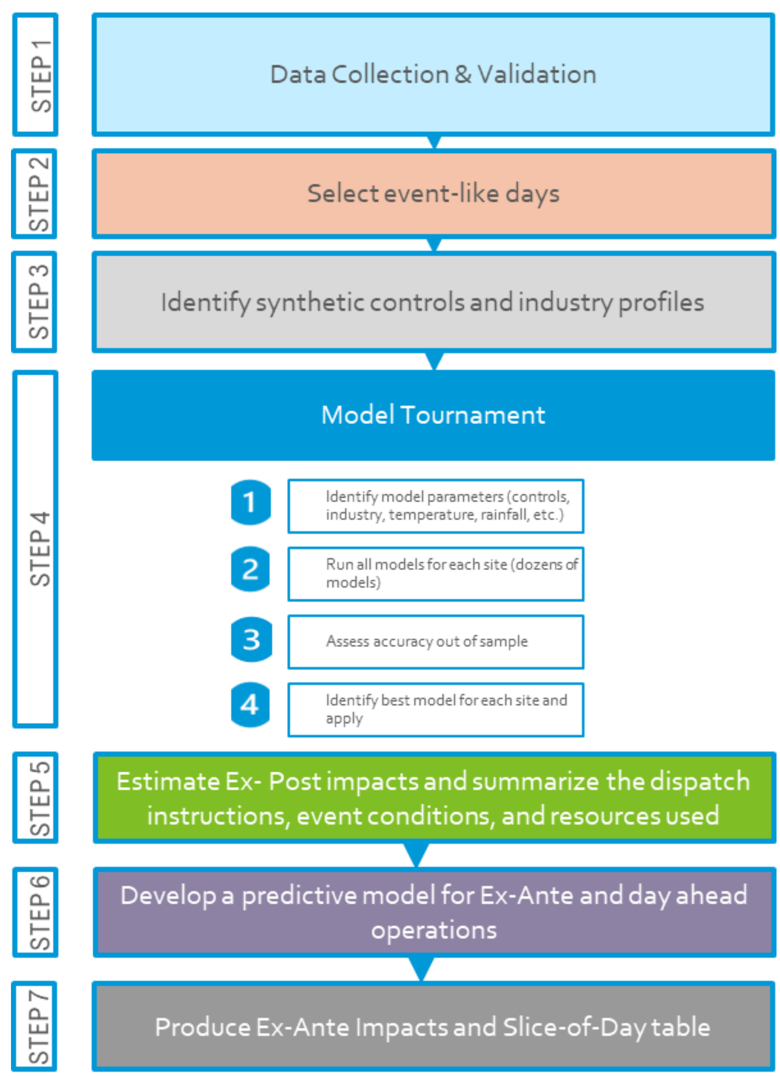
Program	Event Times	Dispatch Length	Sites Dispatched	Average Event Temperature (°F)	Statewide Reservoir Levels
RA	10/05/2023 16:00 to 20:00	4 Hours	1,291	83.68	72%
	10/19/2023 16:00 to 20:00	4 Hours	1,175	79.79	69%
	05/30/2024 16:00 to 20:00	4 Hours	1,170	74.48	88%
	06/26/2024 16:00 to 20:00	4 Hours	874	81.99	85%
	06/27/2024 16:00 to 20:00	4 Hours	288	82.64	85%
	07/09/2024 16:00 to 20:00	4 Hours	874	86.76	81%
	07/18/2024 17:00 to 21:00	4 Hours	243	82.82	79%
	07/30/2024 17:00 to 21:00	4 Hours	242	81.71	75%
	08/27/2024 16:00 to 20:00	4 Hours	849	82.16	69%
	08/27/2024 16:00 to 19:00	3 Hours	286	85.50	69%
	09/05/2024 16:00 to 20:00	4 Hours	718	96.66	67%
	09/09/2024 16:00 to 20:00	4 Hours	244	95.01	67%

3 METHODOLOGY

The primary challenge of an impact evaluation is the need to accurately detect changes in demand while systematically eliminating plausible alternative explanations for those changes, including random chance. The core question is: did the dispatch of DR resources cause a decrease in hourly demand, or can the difference be explained by other factors? To estimate demand reductions, it is necessary to estimate what demand patterns would have been in the absence of the intervention – this is called the counterfactual or reference load.

Broadly, there are two main approaches to establish the counterfactual. The first approach relies on non-event days to developed a model that accurately estimates load patterns in the absence of events. A unique characteristic of demand response resources is that the primary intervention – demand response dispatch – is introduced on some days and not on others, making it possible to observe energy use patterns with and without demand reductions. The second approach is to create a control group, either by random assignment or matching. A control group consists of customers who are similar to participants, experienced the same event day conditions, but are not dispatched during events (or were not transitioned to time varying pricing). Ideally, control and participant groups should have similar energy usage patterns when the intervention is not in place and the only systematic difference is that the treatment group is receives DR dispatch instructions while the control group does not. The best methods typically use both non-event day data and control groups.

Figure 8: Analysis Overview



The methodological approach DSA used follows a structured process, as illustrated in Figure 8. The approach is centered on a model tournament to identify the most accurate regression model for each site, which is then used to estimate the load impacts. The main steps included:

1. **Data Collection and Validation**. The analysis relies on interval data, customer characteristic profiles, event dispatch instructions, and local weather data. In addition, DSA expanded the list of explanatory variables to include rainfall, water reservoir levels, publicly available PG&E granular industry profiles of actual loads, and actual historical rate profiles by utility and rate class. The data was carefully vetted and synthesized into a comprehensive dataset containing the key features necessary for the evaluation. DSA used comprehensive pre-analysis checklist to the data quality and analysis data structure (Appendix A).
2. **Select event-like days to be used for the model tournament.** In order to assess accuracy, it is critical to know that predictions reflect observed conditions. Event-like days were chosen based on their similarity to actual event days in terms of temperature and aggregate loads. A subset of event like days were withheld from the modeling and used to assess the accuracy of the models tested (and out-of-sample test).
3. **Identify controls and industry profiles for each site.** The California utilities produce actual historical hourly loads (8760) for the average customer in each rate class. In addition, PG&E produces publicly available historical hourly profiles (8760) by rate class, climate region, and industry. These profiles are used as non-equivalent control groups to help account changes to demand that are not due to weather or the DR dispatch instructions.
4. **Run the model tournament and identify the best model for each site.** The tournament is used to identify the most accurate model for each participating site. The model tournament included 17 regression models, each with a different combination of explanatory variables (temperature, load profiles, rainfall, snowpack, etc.). Each model was estimated three times, withholding a different group of event-like days (testing days) to assess accuracy. The out-of-sample predictions were compared to the actual loads for the event-like days to develop performance metrics and identify the best performing models. For each site, the models were narrowed down to models with bias below +/-1% (or top 3 with the least bias), and then the model with the smallest prediction errors across (as measured by RRMSE) was selected.
5. **Estimate ex-post impacts using the best model for each site.** The most accurate model for each site was used to estimate the impacts on event dispatch days and hours. The objective was to produce the most accurate estimates of the delivered demand reductions under the conditions the resources were called. In 2024, CAISO did not dispatch all resources at the same time for all grid areas (sub-LAPs). The event days and event hours varied by site and grid location and in all instances a subset of the Enersponse resources were dispatched.
6. **Use historical event performance to develop predictive models for ex-ante planning and operations.** The predictive models are designed to estimate load impact under planning conditions and for future operations. They enable estimates of the resource capability for different weather conditions, event start times, and dispatch durations. A distinct predictive model was estimated for each sub-lap and industry group using the 2024 load impact estimates for each site. We did not include multiple years because of substantial growth in Enersponse resources, the large number of 2024 events, and the change in the third-party evaluator.

7. **Produce the ex-ante tables and slice of day tables for resource adequacy and NQC.** By design, the ex-ante impacts reflect the resource capability under planning conditions, assuming all resources are dispatched. The ex-ante forecasts include a 10% year-on-year growth in resources until 2031 and a 3% growth rate thereafter, assuming a similar mix of customer and similar performance as what was observed historically.

Appendix B includes additional details regarding the model tested and performance metrics. Appendix C details the Ex-ante models and the performance of the predictive models.

3.1 EX-POST IMPACT ESTIMATION

The evaluation of ex-post impacts follows a methodical approach. By analyzing participant consumption on non-event days, we can calculate the electricity usage that would have normally been observed absent any program dispatch. During successful DR events, the drop in energy usage should be obvious during the event window. Table 11 provides an overview of the process used by DSA to estimate the ex-post impacts.

Table 11: Ex-Post Analysis Approach

Methodology Component	Description
1. Population or sample analyzed	The analysis considers the full population of participants active on the event days. 3,811 of 4,131 total participants had complete interval data. The population analyzed only includes these customers
2. Data included in the analysis	The analysis utilized utility AMI data, CalMAC temperature data, California Data Exchange Center (CDEC) water data, utility load profiles, and Enersponse provided customer characteristics and event information
3. Model selection	Each individual participant was assigned a model identified based on out-of-sample metrics for bias and fit. The process relies on identifying proxy days to be used for out-of-sample testing. The models were developed using the training data and applied, out-of-sample, to the testing data using a cross-validation approach. For each model specified, we produce standard metrics for bias and goodness of fit. The best model is identified by narrowing the candidate models to the three with the least bias and then selecting the model with the highest precision.
4. Segmentation of impact results	The results were segmented by: ▪ Program ▪ Local Capacity Area ▪ Sub-LAP ▪ Sector ▪ Utility

3.2 EX-ANTE IMPACT ESTIMATION

While the ex-post analysis quantifies the actual impact of past DR events, ex-ante estimation provides a forward-looking assessment of potential load reductions under future dispatch conditions. To support operational planning, a predictive panel model was developed to estimate DR impacts at both the program level and within CAISO's sub-LAPs. Unlike the ex-post methodology, which relies on site-level modeling, the ex-ante approach aggregates customer data to the sub-LAP level prior to estimation. Table 12 provides an overview of the ex-ante analysis.

Table 12 – Ex-Ante Analysis Approach

Methodology Component	Description
1. Historical Performance Used	PY2024 Ex-Post impacts and consumption data
2. Process for producing reference loads	Key steps included: ▪ Aggregate data to combinations of sub-LAP and sector groupings ▪ Estimate the relationship between non-event day consumption and temperature using a temperature spline model, absorbing individual customer fixed-effects ▪ Predict reference loads for ex-ante conditions
3. Process for producing ex-ante impacts	Key steps included: ▪ Aggregate event performance to combinations of sub-LAP and sector groupings ▪ Identify candidate models to be used for making ex-ante predictions ▪ Use cross-validation training and testing methodologies to estimate the out-of-sample performance of each model ▪ Predict impacts for ex-ante conditions
4. Accounting for enrollment growth	It is expected that the customer mix will grow in future years. A growth rate of 10% was used to forecast future participation up until 2031, and 3% thereafter.

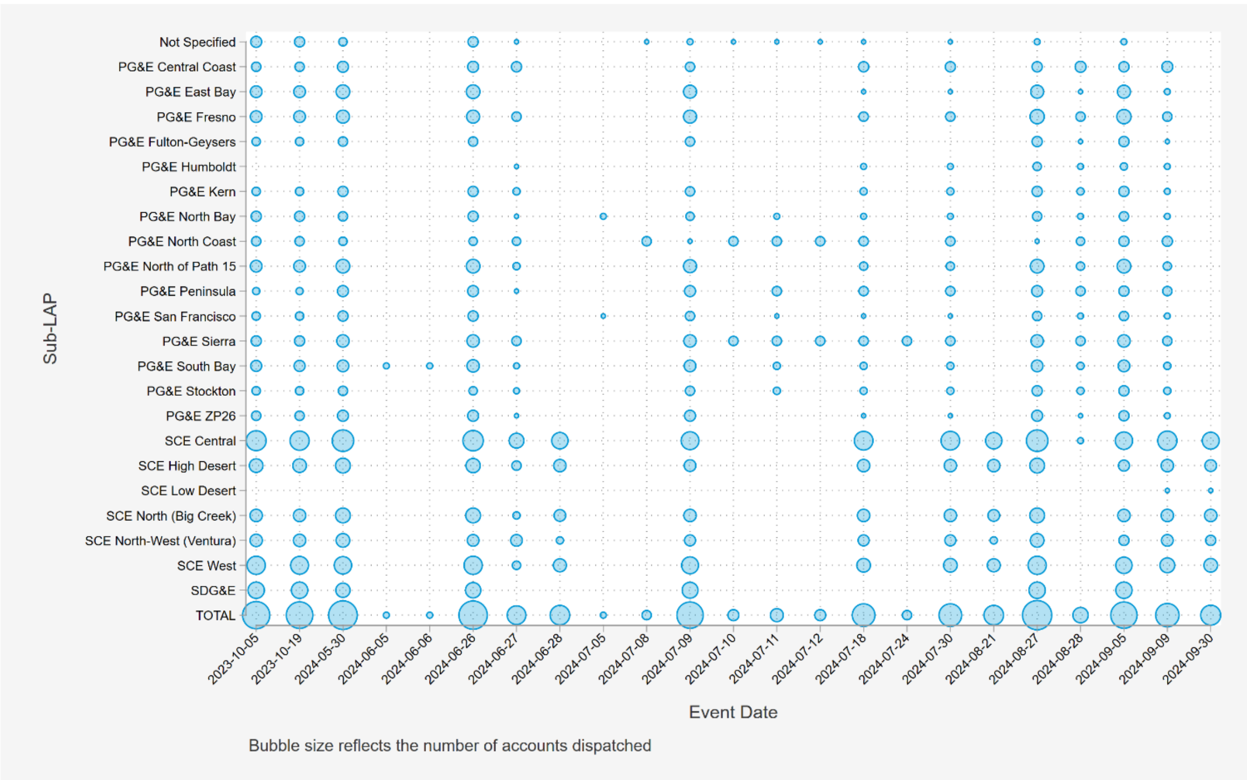
4 EX-POST EVENT BASED LOAD IMPACTS

This section examines the load impacts of DR events occurring in the analysis period, focusing on both aggregate and program-level performance.

4.1 PORTFOLIO RESULTS

When analyzing the impacts observed over the PY2024 season, it is crucial to consider the variation in resources dispatched for each event. Figure 9 highlights this variability, displaying the number of Enersponse sites dispatched for each event by sub-LAP. Many events had very low site counts, such as the events on June 5th and July 5th. Of all the events, the event on August 27th had the most complete dispatch, with a total of 2,639 sites participating. Even when resources were dispatched during the same hours, the event window varied by program and location.

Figure 9: Dispatch variability by Event Date and Sub-LAP



Because not all resources were called on each event, the load reductions varied substantially by date. Table 13 shows

Demand Side Analytics
DATA DRIVEN RESEARCH AND INSIGHTS

Table 2 a comprehensive view of Enersponse PY 2024 resources by CAISO sub-LAP grid areas. To produce the table, the average reduction delivered over PY2024 event hours was calculated for each site and then aggregated by sub-LAP. This approach provides an estimate of the performance for the average event hour while accounting for the reality that the sites were not all dispatched simultaneously. Overall, the Enersponse resources delivered 34.3 MW of resources for the average event hour in 2024.

Table 13: Average Hourly Impacts by Sub-LAP Grid Area[1]

Utility	Sub-LAP	Site Counts	Average Event Temperature (°F)	Per-Site Reference Load (kW)	% Impact	Per-Site Impact (kW)	Aggregate Impact (MW)
PG&E	Unclassified	[redacted]					
	PG&E Central Coast	95	69.28	85.11	36.5%	31.10	2.95
	PG&E East Bay	139	84.35	94.70	6.3%	5.94	0.83
	PG&E Fresno	173	98.56	67.83	2.8%	1.90	0.33
	PG&E Fulton-Geysers	[redacted]					
	PG&E Humboldt	[redacted]					
	PG&E Kern	55	97.52	60.53	6.0%	3.65	0.20
	PG&E North Bay	65	83.94	78.48	7.5%	5.89	0.38
	PG&E North Coast	66	72.73	18.29	-2.5%	-0.46	-0.03
	PG&E North of Path 15	151	94.09	98.30	5.4%	5.31	0.80
	PG&E Peninsula	77	76.62	84.86	5.9%	4.98	0.38
	PG&E San Francisco	[redacted]					
	PG&E Sierra	109	94.39	80.56	7.0%	5.62	0.61
	PG&E South Bay	88	83.04	99.24	4.6%	4.61	0.41
	PG&E Stockton	[redacted]					
	PG&E ZP26	55	84.52	85.23	1.6%	1.40	0.08
SCE	Unclassified	[redacted]					
	SCE Central	988	92.50	91.11	22.0%	20.04	19.80
	SCE High Desert	199	93.83	44.41	10.0%	4.46	0.89
	SCE Low Desert	[redacted]					
	SCE North (Big Creek)	218	93.20	85.96	13.1%	11.23	2.45
	SCE North-West (Ventura)	184	77.53	63.67	4.1%	2.60	0.48
	SCE West	620	76.99	74.32	3.3%	2.44	1.52
	Unclassified	[redacted]					
SDG&E	SDG&E	298	78.40	56.24	6.2%	3.48	1.04
TOTAL		**3,811**	**84.32**	**80.70**	**11.2%**	**9.01**	**34.33**

[1] Due to variability in CAISO dispatch, the average reduction delivered over PY2024 event hours was calculated for each site, including all hours the sites was dispatched. Next, the site level impacts were aggregated by sub-LAP.

Table 3 shows the ex-post impacts for each event day, but is less useful for understanding the resource capability due to fact that not all resources were dispatched on the same day and event window varied by program and location on event days. Of all the events, the event on August 27th had the most complete dispatch with 2,639 sites participating and 28.2 MW of load reduction. However, the August 27 events did not include dispatch of Enersponse DSGS resources.

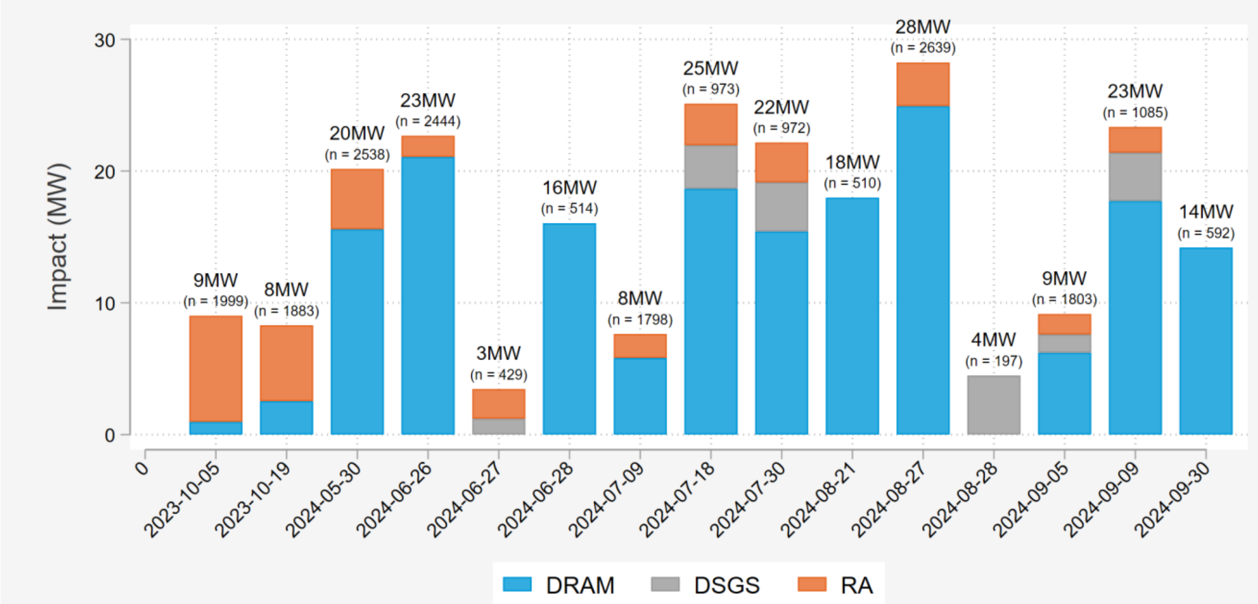
Table 14: Average Hourly Impacts by Utility and Event Day[1]

Date	PGE Impact (MW)	PGE Sites	SCE Impact (MW)	SCE Sites	SDGE Impact (MW)	SDGE Sites
10/5/2023	-0.13	461	8.04	1,291	1.10	247
10/19/2023	1.48	461	5.75	1,175	1.07	247
5/30/2024	2.21	818	17.68	1,569	0.26	151
6/5/2024	[redacted]		0.00	0	0.00	0
6/6/2024	[redacted]		0.00	0	0.00	0
6/26/2024	5.47	859	16.60	1,379	0.61	206
6/27/2024	1.07	135	2.38	294	0.00	0
6/28/2024	0.00	0	16.04	514	0.00	0
7/5/2024	[redacted]		0.00	0	0.00	0
7/8/2024	[redacted]		0.00	0	0.00	0
7/9/2024	5.34	693	1.81	872	0.48	233
7/10/2024	[redacted]		0.00	0	0.00	0
7/11/2024	0.18	53	0.00	0	0.00	0
7/12/2024	0.02	31	0.00	0	0.00	0
7/18/2024	2.73	184	22.38	789	0.00	0
7/24/2024	[redacted]		0.00	0	0.00	0
7/30/2024	3.09	184	19.08	788	0.00	0
8/21/2024	0.00	0	17.98	510	0.00	0
8/27/2024	5.97	800	21.00	1,570	1.27	269
8/28/2024	3.71	191	[redacted]		0.00	0
9/5/2024	4.65	810	2.93	722	1.57	271
9/9/2024	3.19	215	20.16	870	0.00	0
9/30/2024	0.00	0	14.19	592	0.00	0

[1] Due to variability in CAISO dispatch, the average reduction per event involved two steps: a) Estimate average impacts delivered per event date across hour dispatched for each site, and b) Sum the impacts per event date for all sites dispatched by CAISO.

Figure 10 similarly presents the averaged observed impacts during event hours for events where at least 100 sites were dispatched, aggregating results by program. These aggregated reductions ranged from as low as 3 MW on June 27 to a peak of 28 MW on August 27, with multiple events exceeding 20 MW. Nearly all the variation was due to partial dispatch of Enersponse resources rather than variability in performance. DRAM consistently contributed the largest share of reductions, while RA and DSGS provided additional but smaller reductions. It is important to note that RA and DSGS events tended to be longer in length than DRAM events, spreading impacts over 3 to 4 hours as opposed to 2 hours.

Figure 10: Average Hourly Portfolio Level Impacts by Date



*Includes events with greater than 100 sites dispatched
*Total impacts and site counts represent the average observed values during event hours, aggregated across all programs

4.2 DRAM

Overall Performance

Table 15 displays the average hourly event performance for DRAM participants by event. DRAM events delivered varying levels of load reductions throughout the season, largely influenced by dispatch levels. The number of sites dispatched ranged from full participation in some events to as low as 32%, with peak deployment occurring on June 26 and August 27 when over 1,500 sites participated. Load reductions ranged from 0.97 MW (2.5%) to 24.94 MW (16.4%), with the highest percentage reductions occurring in events with lower dispatch rates, such as August 21 (30.4%) and June 28 (28.6%). Temperature played a significant role, with the highest site-weighted temperature (97°F on September

Demand Side Analytics
DATA DRIVEN RESEARCH AND INSIGHTS

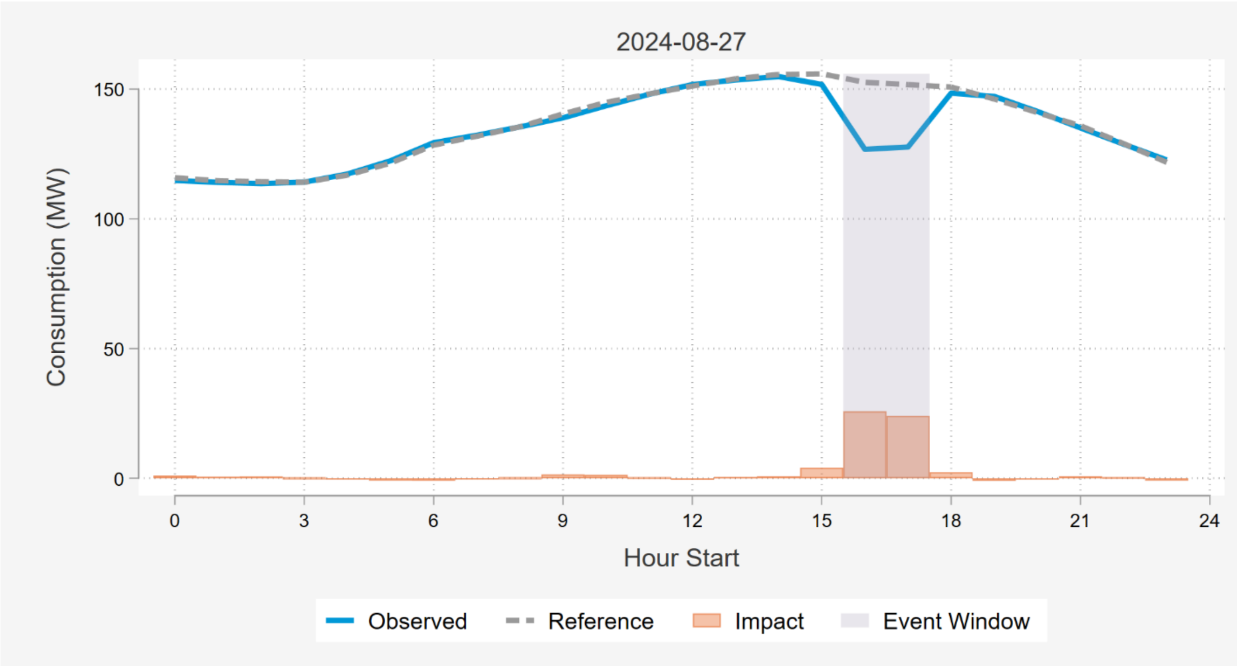
9) coinciding with a substantial 18.1% load reduction. However, response varied across events, indicating that factors beyond temperature also influenced outcomes.

Table 15 - Overall Event Performance - DRAM

Event	Total Sites Dispatched	% of Customers Dispatched	Reference load (MW)	Load w/ DR (MW)	Load reduction (MW)	% Load reduction	Avg temp (F, site weighted)	Std. error
10/05/2023 16:00 to 18:00	708	100%	38.58	37.61	0.97	2.5%	86	0.260
10/19/2023 16:00 to 18:00	708	100%	37.21	34.66	2.54	6.8%	82	0.256
05/30/2024 16:00 to 18:00	1,368	93%	118.23	102.64	15.59	13.2%	82	0.637
06/26/2024 16:00 to 18:00	1,570	98%	144.14	123.06	21.08	14.6%	82	0.643
06/28/2024 16:00 to 18:00	514	32%	56.16	40.11	16.04	28.6%	86	0.608
07/09/2024 16:00 to 18:00	924	62%	84.26	78.44	5.82	6.9%	85	0.208
07/18/2024 16:00 to 18:00	540	36%	69.94	51.27	18.66	26.7%	90	0.611
07/30/2024 16:00 to 18:00	540	36%	67.78	52.37	15.42	22.7%	88	0.608
08/21/2024 16:00 to 18:00	511	32%	59.41	41.36	18.05	30.4%	91	0.617
08/27/2024 16:00 to 18:00	1,577	98%	152.24	127.30	24.94	16.4%	88	0.650
09/05/2024 16:00 to 18:00	1,043	62%	97.64	91.43	6.21	6.4%	93	0.256
09/09/2024 16:00 to 18:00	620	37%	98.12	80.38	17.74	18.1%	97	0.927
09/30/2024 16:00 to 18:00	592	35%	82.09	67.89	14.20	17.3%	86	0.909

Figure 11 illustrates an example event from August 27, 2024, showcasing the observed consumption compared to the reference load during the event window. The event led to a clear reduction in electricity usage, with consumption dropping notably below the reference load during the dispatch period. The shaded region marks the event window, where demand response efforts resulted in a measurable reduction, as shown by the orange impact bars. Pre- and post-event consumption trends suggest that participants maintained typical load patterns outside the event period, with minimal evidence of significant pre-cooling or post-event rebound. Additionally, there is little decay of impacts over the event period.

Figure 11: Example Event Day – DRAM 8/27/2024



Performance by Sub-LAP

Table 16 displays the average individual event performance for DRAM events by utility. The sites included in this table are sites that were a part of DRAM at the end of the PY2024 season. Generally, the largest percent impacts were in the SCE's Central and Northern jurisdictions.

Table 16: Average Event Performance by Sub-LAP – DRAM

Utility	Sub-LAP	Site Counts	Average Event Temperature (°F)	Per-Site Reference Load (kW)	% Impact	Per-Site Impact (kW)	Aggregate Impact (MW)
PG&E	Unclassified	[redacted]					
	PG&E Central Coast	50	70.7	68.5	4.1%	2.83	0.14
	PG&E East Bay	136	84.6	95.6	6.4%	6.08	0.83
	PG&E Fresno	142	98.7	47.8	4.0%	1.90	0.27
	PG&E Fulton-Geysers	[redacted]					
	PG&E Humboldt	[redacted]					
	PG&E Kern	[redacted]					
	PG&E North Bay	62	84.2	75.3	8.1%	6.10	0.38
	PG&E North Coast	[redacted]					
	PG&E North of Path 15	132	94.0	76.2	7.9%	6.02	0.80

Utility	Sub-LAP	Site Counts	Average Event Temperature (°F)	Per-Site Reference Load (kW)	% Impact	Per-Site Impact (kW)	Aggregate Impact (MW)
	PG&E Peninsula	52	79.8	119.3	5.8%	6.90	0.36
	PG&E San Francisco	[redacted]					
	PG&E Sierra	83	93.0	77.7	7.5%	5.82	0.48
	PG&E South Bay	79	83.5	98.8	4.9%	4.82	0.38
	PG&E Stockton	37	94.0	114.3	4.7%	5.36	0.20
	PG&E ZP26	52	84.4	81.8	0.8%	0.67	0.04
SCE	SCE Central	311	93.1	146.8	34.2%	50.13	15.59
	SCE High Desert	89	94.8	56.1	13.0%	7.26	0.65
	SCE Low Desert	[redacted]					
	SCE North (Big Creek)	84	90.3	129.7	20.6%	26.75	2.25
	SCE North-West (Ventura)	[redacted]					
	SCE West	153	76.6	145.9	4.2%	6.15	0.94
SDG&E	Unclassified	[redacted]					
	SDG&E	294	78.3	56.6	6.2%	3.49	1.03
TOTAL	**TOTAL**	**2,001**	**84.7**	**95.8**	**13.3%**	**12.71**	**25.43**

Performance by Sector

Figure 12 categorizes the dispatched sites into industry-defined sectors based on classifications provided by the Enersponse, grouping similar customer types to assess sector-level DR impacts. The Pumping & Ag. sector, which includes pumping and agricultural customers, consistently delivered the largest share of load reductions, particularly in late summer when demand was highest. The Retail sector, comprising both small-box and big-box retail stores, also played a significant role, with notable contributions during high-demand events such as August 27, which recorded the highest total reduction of 24.9 MW. Supermarkets are represented as a standalone category and exhibited steady but moderate reductions across events, while the Other category accounted for a very small share of total reductions.

Figure 12: Average Hourly Impacts by Sector - DRAM

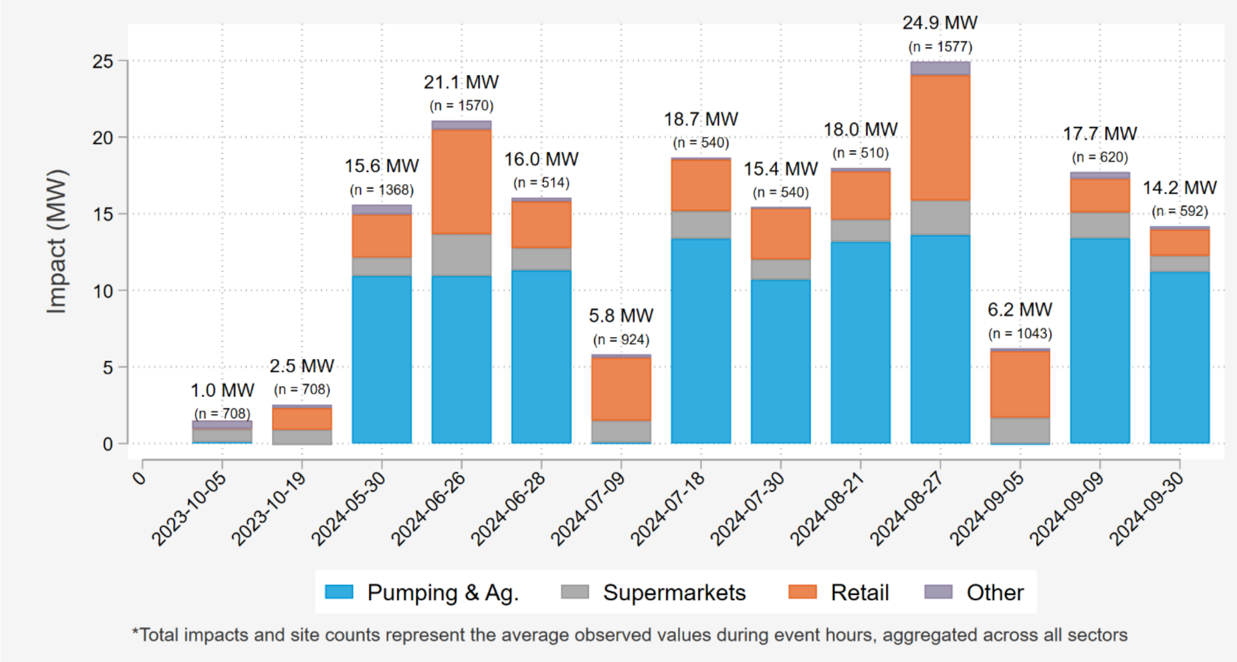


Table 17 provides a more detailed look into the average event performance by sector. Pumping customers made up the bulk of impacts with large participation counts and high percent impacts. Performance was also driven by supermarkets with very high participation counts, however, their percent impacts were lower.

Table 17: Average Event Performance by Sector - DRAM

Sector	Site Counts	Average Event Temp. (°F)	Per-Site Reference Load (kW)	% Impact	Per-Site Impact (kW)	Aggregate Impact (MW)
Other	217	89.6	103.7	3%	3.24	0.70
Pumping & Ag.	362	86.1	42.3	63%	26.82	9.71
Retail	369	87.3	74.7	10%	7.71	2.84
Supermarkets	1,053	84.4	96.5	6%	5.55	5.84

4.3 RA

Overall Performance

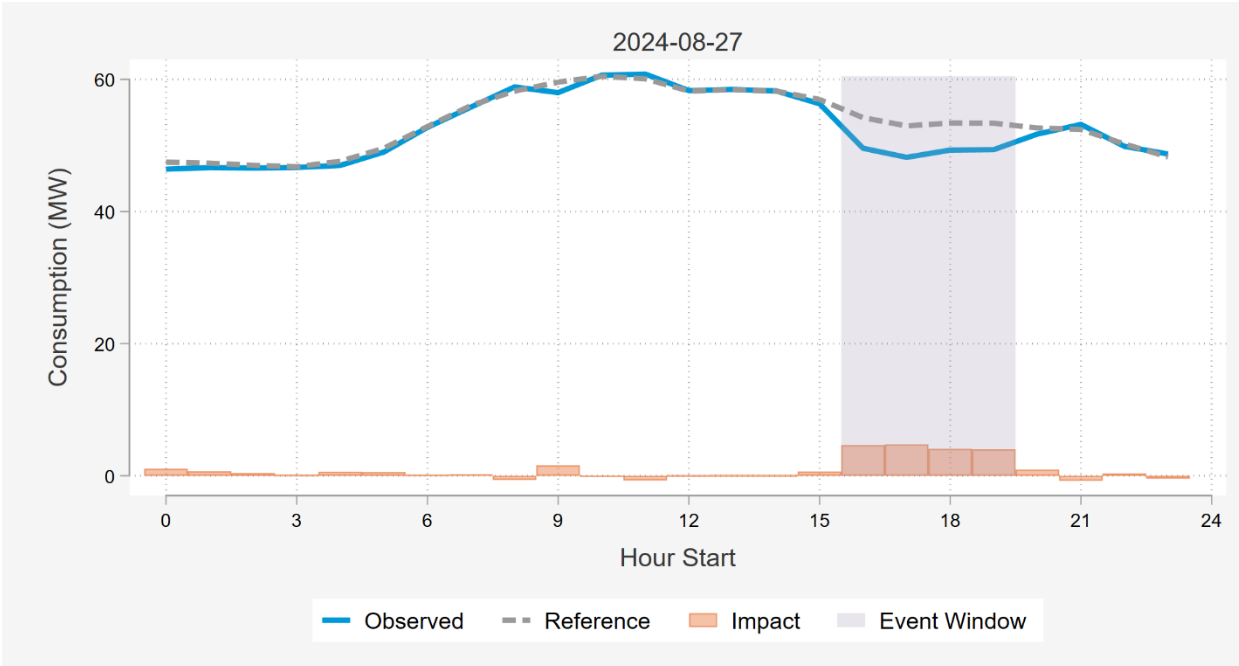
RA events delivered varying levels of load reductions. While some events dispatched a high percentage of available resources, others, particularly in mid-summer, had much lower participation, with as little as 18% on July 18 and July 30. Interestingly, events with lower dispatch levels often achieved higher percentage reductions, with July 18 (21.5%) and August 27 (22.1%) showing the strongest per-site responses. Absolute reductions peaked at 8.05 MW on October 5, while the lowest observed impact was 0.90 MW on August 27, despite it being one of the more responsive events on a percentage basis. Temperature alone did not determine load reductions, as some high-heat events, such as September 5 (97°F), produced only a 4.4% reduction, whereas more moderate temperatures often coincided with stronger demand response.

Table 18:Overall Event Performance - RA

Event	Total Sites Dispatched	% of Customers Dispatched	Reference load (MW)	Load w/ DR (MW)	Load reduction (MW)	% Load reduction	Avg temp (F, site weighted)	Std. error
10/05/2023 16:00 to 20:00	1,291	86%	110.80	102.75	8.05	7.3%	84	0.452
10/19/2023 16:00 to 20:00	1,175	78%	88.10	82.35	5.75	6.5%	80	0.435
05/30/2024 16:00 to 20:00	1,170	83%	46.46	41.89	4.57	9.8%	74	0.430
06/26/2024 16:00 to 20:00	874	62%	38.92	37.32	1.60	4.1%	82	0.253
06/27/2024 16:00 to 20:00	288	20%	14.48	12.24	2.24	15.5%	83	0.352
07/09/2024 16:00 to 20:00	874	64%	41.35	39.54	1.81	4.4%	87	0.254
07/18/2024 17:00 to 21:00	243	18%	14.63	11.49	3.14	21.5%	83	0.347
07/30/2024 17:00 to 21:00	242	18%	14.81	11.79	3.02	20.4%	82	0.343
08/27/2024 16:00 to 20:00	849	70%	38.56	37.66	0.90	2.3%	82	0.261
08/27/2024 16:00 to 19:00	286	24%	14.62	11.39	3.23	22.1%	86	0.391
09/05/2024 16:00 to 20:00	718	62%	35.64	34.09	1.56	4.4%	97	0.255
09/09/2024 16:00 to 20:00	244	21%	14.88	12.92	1.95	13.1%	95	0.355

Figure 13 depicts an event from August 27, 2024, showing a clear reduction in electricity consumption during the dispatch period. Load declines notably during the event window compared to the expected baseline, indicating successful curtailment. The reduction is sustained throughout the event, with minimal signs of pre-event adjustments or post-event rebound effects. While overall consumption follows typical daily patterns, the observed drop during the event highlights effective demand response performance.

Figure 13:Example Event Day – RA 8/27/2024



Performance by Sub-LAP

Table 19 illustrates the average individual event performance by utility. The site counts include participants who were a part of RA at the end of the PY2024 season. The percent impacts were generally lower for RA than DRAM, but the sites in the SCE Central sub-LAP again exhibited the largest relative impacts.

Table 19:Average Event Performance by Sub-LAP – RA

Utility	Sub-LAP	Site Counts	Average Event Temperature (°F)	Per-Site Reference Load (kW)	% Impact	Per-Site Impact (kW)	Aggregate Impact (MW)
PG&E	PG&E Fresno	[redacted]					
SCE	Unclassified	[redacted]					
	SCE Central	671	92.2	64.0	8.3%	5.31	3.57
	SCE High Desert	110	93.1	35.0	6.3%	2.19	0.24
	SCE North (Big Creek)	134	95.0	58.5	2.6%	1.50	0.20
	SCE North-West (Ventura)	149	77.1	40.6	5.3%	2.16	0.32
	SCE West	467	77.1	50.9	2.4%	1.23	0.57
SDG&E	SDG&E	[redacted]					
TOTAL	**TOTAL**	**1,573**	**84.4**	**57.9**	**5.2%**	**3.02**	**4.75**

PERFORMANCE BY SECTOR

RA event impacts were largely driven by the Pumping & Ag. sectors, which contributed the most load reductions across all events, like trends observed in DRAM. Retail and supermarkets made smaller but notable contributions, particularly in early-season events such as October 19 and May 30, where they accounted for a meaningful portion of total reductions. The Other category again played a more limited role. These results highlight the concentration of RA event impacts within agricultural and industrial sectors.

Figure 14- Average Hourly Impacts by Sector - RA

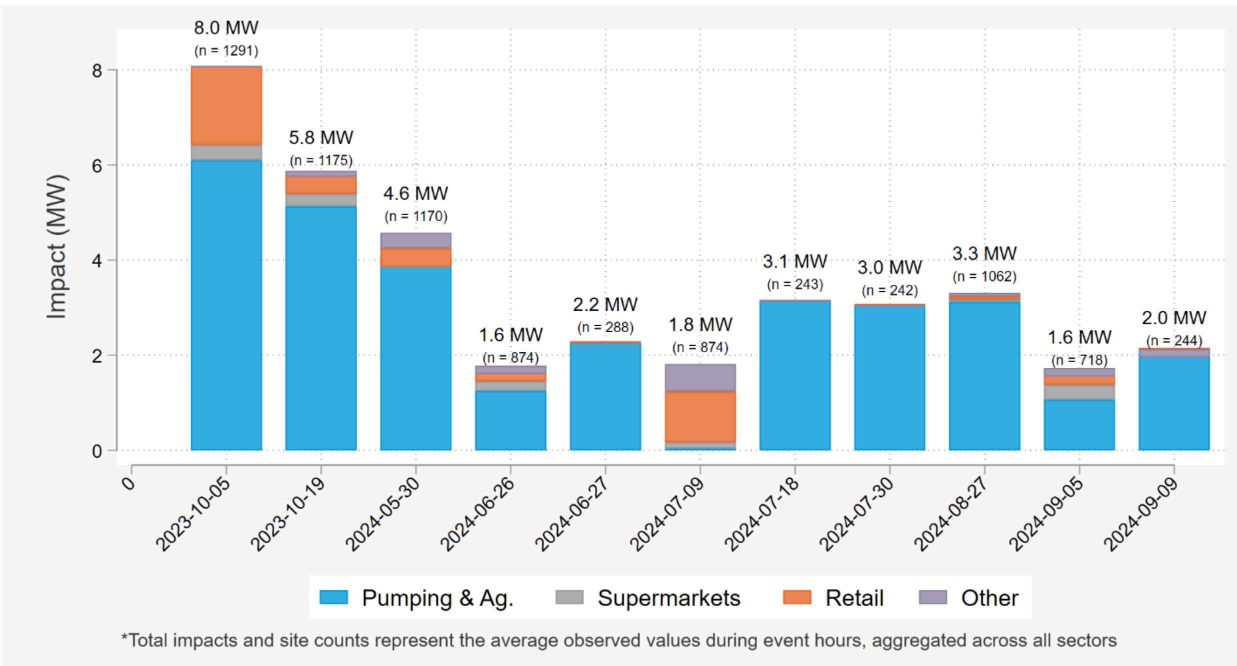


Table 20 provides a more detailed look into the average event performance by sector. Similar to DRAM, Pumping customers typically drove higher impact, albeit not as high as DRAM participants.

Table 20:Average Event Performance by Sector - RA

Sector	Site Counts	Average Event Temp. (°F)	Per-Site Reference Load (kW)	% Impact	Per-Site Impact (kW)	Aggregate Impact (MW)
Other	184	85.2	105.3	0%	-0.04	-0.01
Pumping & Ag.	506	88.0	44.3	13%	5.94	3.00
Retail	491	87.0	60.7	2%	1.44	0.71
Supermarkets	392	86.2	53.1	6%	3.27	1.28

4.4 DSGS

Overall Performance

DSGS performance varied significantly across the season, with some events achieving substantial load reductions while others had very low dispatch rates. High-dispatch events, such as August 28 (88% dispatched, 4.48 MW reduction) and July 30 (90% dispatched, 3.73 MW reduction), demonstrated strong overall performance, with reductions exceeding 20% of the reference load. Similarly, July 18 (22.8% reduction) and September 9 (16.1% reduction) showed effective curtailment. In contrast, events with low dispatch levels and minimal participant engagement often resulted in negligible reductions, with some instances of increased consumption relative to the reference load, such as July 10 (-18.8%) and July 5 (-2.8%).

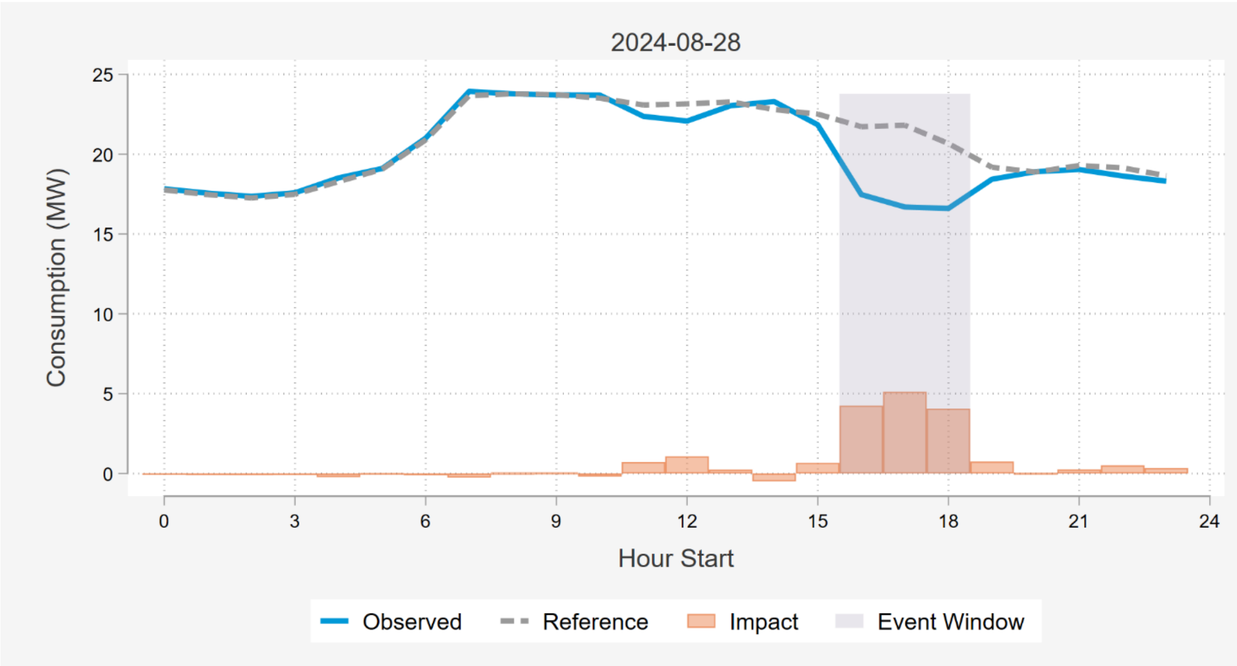
Table 21: Overall Event Performance - DSGS[1]

Event	Total Sites Dispatched	% of Customers Dispatched	Reference load (MW)	Load w/ DR (MW)	Load reduction (MW)	% Load reduction	Avg temp (F, site weighted)	Std. error
06/27/2024 16:00 to 18:00	141	90%	10.30	9.09	1.21	11.8%	80	0.345
07/18/2024 17:00 to 19:00	190	90%	14.48	11.18	3.30	22.8%	79	0.419
07/30/2024 17:00 to 19:00	190	90%	13.97	10.24	3.73	26.7%	77	0.415
08/28/2024 16:00 to 19:00	197	88%	21.40	16.92	4.48	20.9%	80	0.356
[redacted]								

Figure 15 illustrates the demand response performance for the August 28, 2024 event, demonstrating a reduction in electricity consumption during the dispatch period. During the event, load dropped significantly below the reference baseline, indicating effective curtailment. The reduction was sustained throughout the dispatch period, with no immediate rebound effect once the event ended.

[1] This table includes events with more than 100 sites dispatched.

Figure 15:Example Event Day – DSGS 8/28/2024



Performance by Sub-LAP

Table 22 shows the average individual event performance by utility. Sites counts include customers that were part of the DSGS program at the end of the PY2024 season. DSGS impacts were primarily concentrated in PG&E's territory, with customers in the Central Coast providing the bulk of reductions.

Table 22:Average Event Performance by Utility – DSGS

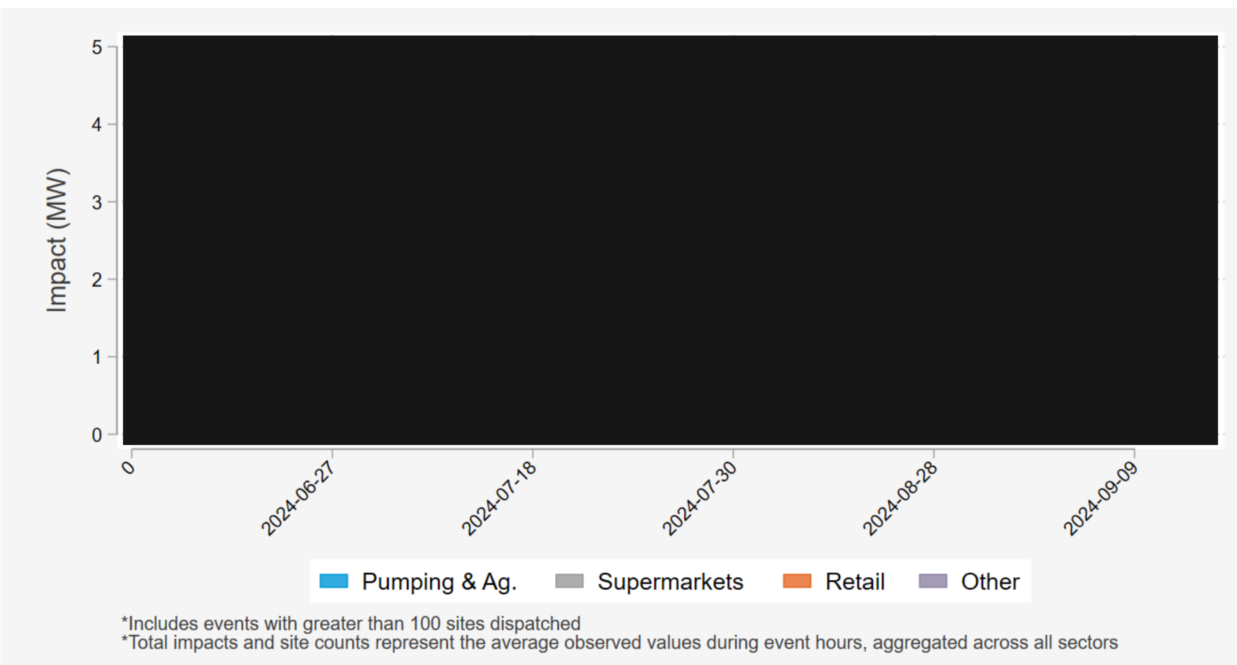
Utility	Sub-LAP	Site Counts	Average Event Temperature (°F)	Per-Site Reference Load (kW)	% Impact	Per-Site Impact (kW)	Aggregate Impact (MW)
PG&E	[redacted]						

Utility	Sub-LAP	Site Counts	Average Event Temperature (°F)	Per-Site Reference Load (kW)	% Impact	Per-Site Impact (kW)	Aggregate Impact (MW)
	[redacted]						
SCE	[redacted]						
SDG&E	[redacted]						
TOTAL		237	80.6	104.5	16.7%	17.49	4.15

Performance by Sector

The sector-level analysis of DSGS load impacts reveals that the Other category provided the majority of reductions across most events. Figure 16 depicts how this grouping consistently contributed the largest share of impacts. The Pumping & Agriculture sector also played a role, though its contributions varied across events, with more notable reductions on July 30 and September 9. Retail and supermarkets provided minimal or no measurable reductions with limited participation. Early-season events generally exhibited lower total reductions, with some instances of negligible impacts, highlighting lower performance when fewer sites were dispatched.

Figure 16: Average Hourly Impacts by Sector - DSGS



Finally, Table 23 shows the average event performance by sector. In contrast to the other programs, Agricultural and Cold Storage customers made up the bulk of impacts when they participated in events.

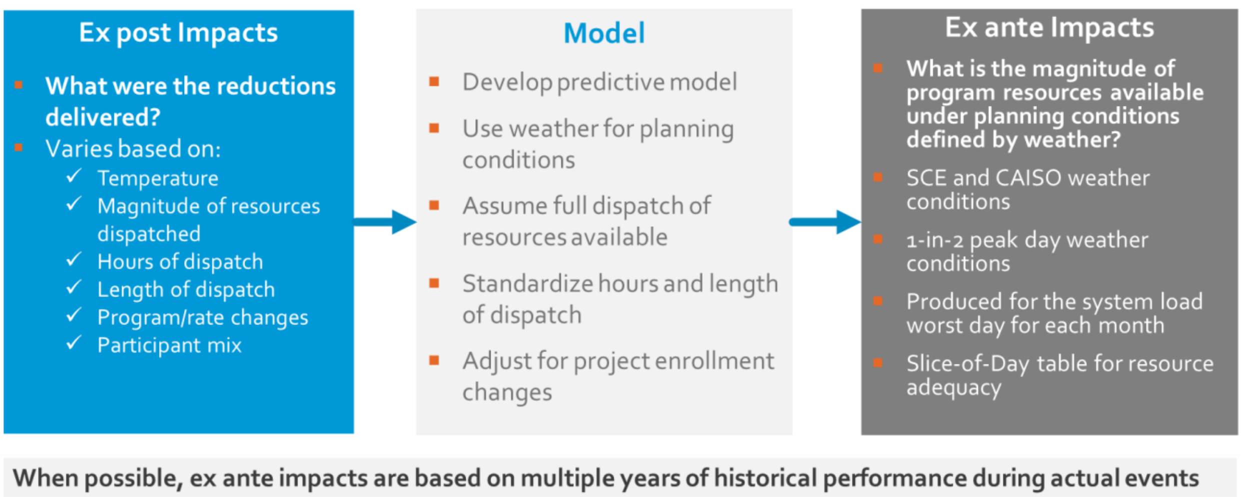
Table 23: Average Event Performance by Sector - DSGS

Sector	Site Counts	Average Event Temp. (°F)	Per-Site Reference Load (kW)	% Impact	Per-Site Impact (kW)	Aggregate Impact (MW)
Other	137	83.8	154.5	10%	16.16	2.21
Pumping & Ag.	100	77.9	36.0	53%	19.22	1.92

5 EX ANTE IMPACTS FOR PLANNING AND OPERATIONS

The electric grid is designed to maintain reliability under peaking conditions when temperatures are typically hottest. Thus, the magnitude and performance of DR resources under peaking conditions used for planning is critical for understanding the degree they can offset other resources, such as peaking gas power plant. Load impacts under planning conditions are referred to as ex-ante impacts and are informed by performance during historical events. They are an estimate of the load reduction capability that align with peak day weather, standardized hours, and length of dispatch. In essence, they are weather normalized impact values that are helpful when forecasting a resource's load reduction potential for different forecasts. This section explores the ex-ante analysis conducted by DSA.

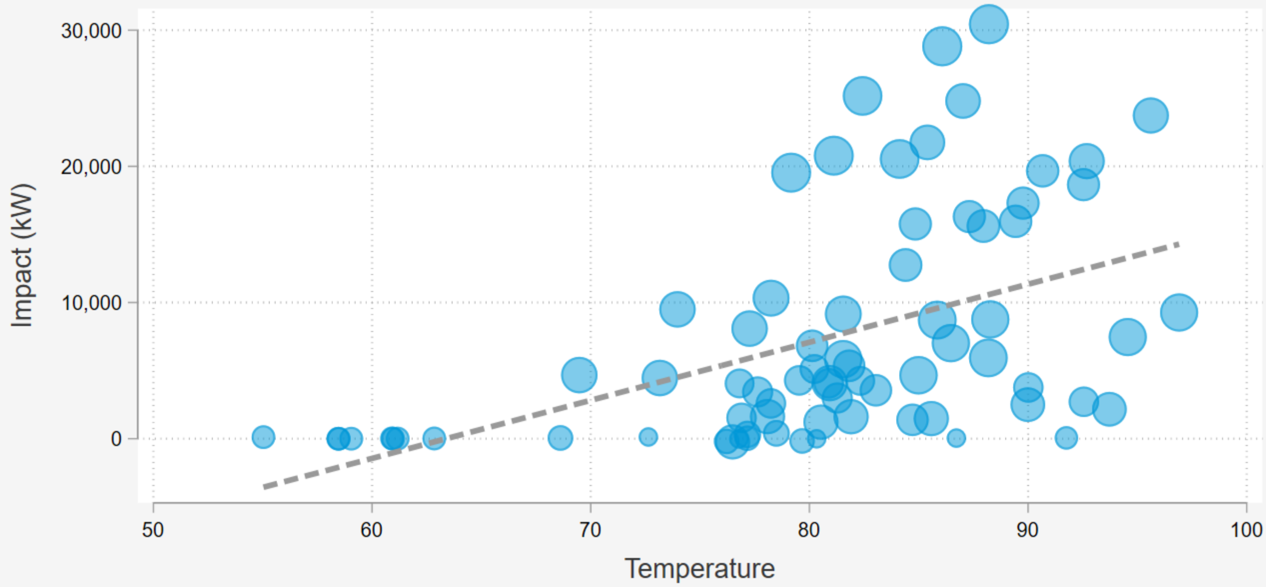
Figure 17: Modeling for planning conditions (Ex-ante Model)



5.1 MODELING EX POST PERFORMANCE AS A FUNCTION OF WEATHER

To develop ex-ante impacts, DSA aimed to understand the weather sensitivity of impacts. Figure 18 shows how the hourly ex-post impacts tended to correlate with higher temperatures. The approach used by DSA to estimate ex-ante impacts parallels the development of the reference loads under ex-post methodologies. However, one of the primary challenges lies in the uncertainty surrounding the relationship between impacts and non-temperature-related factors, such as reservoir levels and industry-specific dynamic load profiles, during the forecasted period. Consequently, the modeling approach needed to estimate reference loads and impact magnitudes without relying on variables unavailable for ex-ante forecasting.

Figure 18:Correlation Between Hourly Impacts and Temperature



Each bubble represents performance for an event hour with all dispatch sites aggregated.
Individual hours have variation in event windows and customer mix.
Bubble size indicates the number of sites dispatched

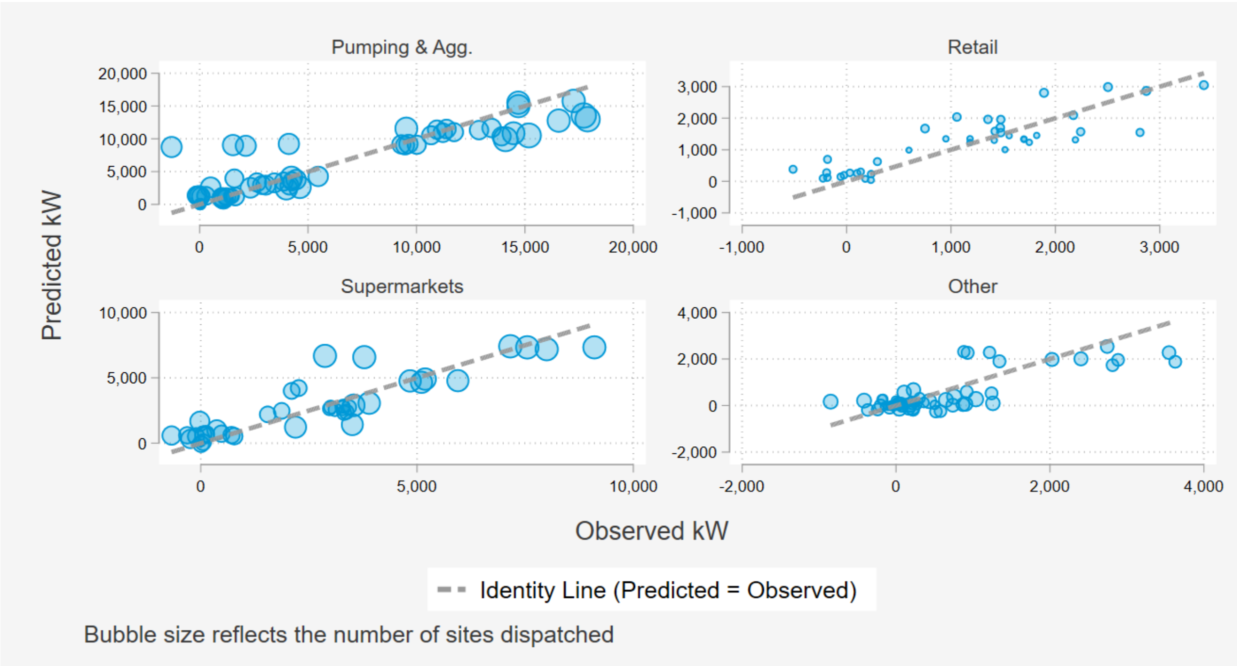
DSA addressed this challenge by estimating reference loads and impacts at the industry and sub-LAP levels, employing an individual fixed-effects model to absorb the variation associated with each site. This approach offered the advantage of incorporating the weather sensitivity specific to each industry's geographic location while controlling for individual variation by capturing the average usage for each customer. The more detailed analysis of the ex-ante impact estimation can be found in Appendix C.

The industry groupings used in the analysis included four categories:

1. Pumping and Agricultural Customers: This group encompasses customers primarily engaged in water pumping for agricultural and irrigation purposes.
2. Retail: Includes both small-box and big-box retail establishments, this category captures a wide range of commercial retail operations.
3. Supermarkets: Focuses specifically on grocery stores and supermarkets, which have distinct load profiles due to refrigeration and lighting demands.
4. Other: Comprising all remaining customers that do not fit into the other three categories, representing a diverse set of industries.

Figure 19 presents the average out-of-sample predictions of impacts for each industry group. The predicted impacts generally aligned closely with the observed values, reinforcing the robustness of the modeling approach and supporting its applicability for estimating impacts under planning conditions.

Figure 19: Observed vs Predicted Hourly Impacts by Industry Group



5.2 EX-ANTE LOAD IMPACT PROJECTIONS

Rather than project large amount of growth, Enersponse has adopted a practice of grounding ex-ante impacts on resources it currently controls, and forecasting very modest year-on-year growth. As result, the projections are lower than in prior evaluation reports.

Enersponse is projecting a 10% year-over-year increase of resource capability until 2031, after which the resources are assumed to grow at a rate of 3% per year. The customer counts at the end of the 2024 season totaled to 3,730 sites. Since customers can be switched between programs, program specific breakouts were not constructed. The assumed event window for dispatch is between 4:00 p.m. and 9:00 p.m., with resources capable of being dispatched for more than four hours. Table 24 depicts the 1-in-2 ex-ante impacts during the August Worst Day for each utility. With the steady increase in participation over time, the aggregate impacts are generally expected to increase over the next ten years.

Table 24: Aggregate Ex-Ante Impacts (MW) – August Worst Day – Utility 1-in-2 Weather

Forecast Year	SCE 1-in-2	PG&E 1-in-2	SDG&E 1-in-2	Total
2024	23.69	7.58	1.53	32.80
2025	26.06	8.34	1.69	36.09
2026	28.66	9.17	1.86	39.70
2027	31.53	10.09	2.04	43.66
2028	34.68	11.10	2.24	48.03
2029	38.15	12.21	2.47	52.83
2030	41.97	13.43	2.72	58.12
2031	46.16	14.77	2.99	63.93
2032	47.55	15.22	3.08	65.85
2033	48.97	15.67	3.17	67.82
2034	50.44	16.14	3.27	69.86
2035	51.96	16.63	3.37	71.96

Table 25 displays the projected impacts by month for 2025 highlighting the seasonality of the impacts. Generally, the summer months provide the largest impacts, paired with higher temperatures and drier conditions. Notably, October saw low impacts for SCE, highlighting their makeup of pumping and agricultural customers which see lower demand during harvest seasons.

Table 25: Aggregate Ex-Ante Impacts (MW) by Monthly Worst Day in 2026

Day Type	SCE 1-in-2	PG&E 1-in-2	SDG&E 1-in-2	Total
January Worst Day	13.62	4.80	-0.24	18.18
February Worst Day	13.33	4.91	-0.17	18.07
March Worst Day	12.25	3.86	-0.06	16.05
April Worst Day	17.16	6.37	0.40	23.93
May Worst Day	20.38	7.52	0.38	28.28
June Worst Day	24.43	10.49	0.83	35.75
July Worst Day	28.47	10.89	1.43	40.79
August Worst Day	28.66	9.17	1.86	39.70
September Worst Day	26.53	8.98	1.58	37.09
October Worst Day	21.98	8.57	1.17	31.73
November Worst Day	19.64	5.85	0.47	25.97
December Worst Day	16.63	5.36	-0.05	21.94

5.3 RESOURCE ADEQUACY SLICE-OF-DAY TABLE

Table 26 presents the slice-of-day analysis for all participants using utility weather data under 1-in-2 worst day planning conditions. In general, impacts tend to peak at the start of events across all seasons. Pre- and post-event impacts are estimated to be minimal. Additionally, while impacts remain relatively stable throughout event hours, they do show a slight decline over time.

Table 26: 2026 Ex-Ante Slice-of-Day Table –All Participants, 1-in-2 Weather Conditions

Units: MW

Hour Ending	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
15	-10.14	-10.47	0.00	0.00	0.00	-0.56	1.78	0.46	-0.98	-4.60	-6.28	-8.03
16	-1.64	-1.89	-9.69	-5.65	-2.39	7.67	10.39	8.18	7.60	4.12	2.40	0.57
17	21.61	21.08	-1.13	2.37	5.45	39.09	44.89	43.66	41.08	35.41	30.05	26.09
18	21.42	21.07	19.54	28.30	32.56	38.61	43.39	41.05	38.58	34.92	29.33	24.95
19	18.69	18.53	17.78	25.92	30.21	36.27	40.81	39.92	37.73	32.67	26.41	21.92
20	15.10	15.38	15.04	22.78	26.72	32.90	37.56	37.25	34.94	28.75	22.51	18.69
21	14.08	14.26	13.99	21.46	25.73	31.84	37.23	36.57	33.09	26.87	21.51	18.03
22	-2.32	-2.61	13.89	21.19	26.14	6.84	8.79	7.92	6.49	3.36	1.38	-0.39
23	-9.24	-9.82	-2.37	0.92	4.24	0.06	1.73	0.96	-0.36	-3.31	-5.50	-7.31
24	0.00	0.00	-9.98	-6.56	-3.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Demand reductions are positive (Blue)

Load increases are negative (Orange)

*Table is presented in prevailing time. Tables with all months in PST are available in the table generator

5.4 COMPARISON TO 2024 EX-POST IMPACTS

Enersponse was awarded 30 MW of Net Qualifying Capacity (NQC) for October 2023-June 2024, and 40 MW of NQC for July 2024-September 2024. For June 2025-September 2025, the CPUC awarded Enersponse 10 MW of NQC. By contrast, Enersponse delivered 34 MW of resources for the average event hour in PY2024.

In 2024, all Enersponse demand response resources were dispatched multiple times. However, CAISO did not dispatch all of the Enersponse resources all at once at any point in time in 2024. Even when resources were dispatched on the same day, the event window varied by program and location due to the CAISO instructions. The lack of a coordinated dispatch of all resources at the same time makes a direct comparison between the impacts delivered and the resource capability under planning conditions challenging.

Table 27 compares the ex-ante forecast for the August Worst Day grid planning conditions with the average event hour, and the ex-post impacts for August 27, 2024—the day with the highest number of sites dispatched during the summer of 2024. Although no DSGS sites were dispatched on August 27th and enrollment have grown since, the reductions delivered, 28.26 MW, are comparable to the ex-ante demand reduction capability estimated for planning conditions, 32.80 MW. In addition, Enersponse resources delivered 34.33 MW for the average event hours in 2024, slightly more than the 32.80 MW estimated for planning conditions. The average event hour metric is designed to measure performance across all event hours and sites while accounting for the reality that the sites were not all dispatched simultaneously.

Table 27: Comparison to 2024 Ex Post

Day Type	# Dispatched	Event Hour Avg Temp	Avg Cust Ref (kW)	% Impact	Avg Cust Impact (kW)	Agg Impact (MW)
Ex-Ante: 2024 August Worst Day 1-in-2 (4:00 - 9:00PM)	3,730	86.2	78.85	10.0%	8.79	32.80
Ex-Post: 8/27/2024 (4:00-6:00PM)	2,639	85.14	76.39	14.0%	10.7	28.26
Ex-Post Average Event Hour [1]	3,811	84.32	80.70	11.2%	9.01	34.33

[1] Due to variability in CAISO dispatch, the average reduction delivered over PY2024 event hours was calculated for each site, including all hours the sites was dispatched. Next, the site level impacts were aggregated to produce the load reduction for the average event hour.

5.5 COMPARISON TO 2023 EX-ANTE VALUES

Table 28 displays a comparison between the ex-ante impacts produced this year to the ex-ante impacts forecast for PY2024 in the 2023 report. Generally, the Enersponse portfolio's load impact capability in 2024, **32.8 MW** aligns closely with the **33.52 MW** projected in the PY2023 evaluation (98% of projected). However, the mix of customers evolved substantially, with larger number of sites and different geographic footprint than initially projected.

Table 28: Comparison of 2024 to 2023 Ex-Ante Estimates for PY2024 (August 1-in-2)

Utility	Metric	2023 Report	2024 Report
PG&E	Impact (MW)	0.75	7.58
	Sites	656	1,224
	Reference Load (MW)	31.3	114.82
	% Reduction	2.40%	6.60%
SCE	Impact (MW)	32.38	23.69
	Sites	1,799	2,208
	Reference Load (MW)	255.0	189.12
	% Reduction	12.70%	12.52%
SDG&E	Impact (MW)	0.39	1.53
	Sites	166.0	298
	Reference Load (MW)	6	22.96
	% Reduction	6.30%	6.68%
Portfolio	**Impact (MW)**	**33.52**	**32.80**
	Sites	**2,621**	**3,730**
	Reference Load (MW)	**292.4**	**326.90**
	% Reduction	**11.5%**	**10.03%**

6 KEY FINDINGS AND RECOMMENDATIONS

The Enersponse DR portfolio demonstrated effective load reductions across multiple DR pathways during the 2023-2024 season. These reductions varied by program, utility, and sector, highlighting the program's potential to support grid stability during peak demand periods. Below are DSA's key findings and recommendations from this analysis:

Table 29: Key Findings

Category	Key Findings
Historical Performance	▪ Enersponse resources enrolled by September 30, 2024, delivered 34.3 MW for the average event hour in 2024. ▪ When adjusted for planning conditions, the resource capability was 32.8 MW, slightly lower than the projections made for the portfolio level in 2023, and similar to the NQC awarded to Enersponse for PY2024.
Performance Characteristics	▪ The Pumping & Agricultural sectors led in load reductions, especially during the late summer. Retail and supermarkets contributed moderate reductions, while the "Other" category had a minimal impact. ▪ Event impacts were typically consistent across event hours with minimal pre- and post- event increases in consumption, and often persist beyond the event dispatch hours.
Operations and Analysis	▪ DSA recommends that Enersponse continues the practice of grounding ex-ante impacts on resources it currently controls and forecast modest year-on-year growth (e.g. 10%) ▪ While not directly part of the evaluation, DSA identified a gap between the NQC resources awarded to Enersponse in PY2024 and nomination to CAISO. The nominations to CAISO are being updated to match the load reduction capability identified in the evaluation. ▪ The PY2024 included substantial data and modeling updates that improved the accuracy of the impact estimates, including model tournaments to identify the best model for site, and inclusion of additional data sources such as rainfall, reservoir levels, and utility wide aggregate load profiles by rate class.
Forward Projections	▪ Projected enrollment growth of 10% per year is expected to increase aggregate impacts from 32.8 MW in 2024 to over 39.7 MW by 2026. ▪ Enersponse resources that were historically enrolled in DRAM will need to shift to other options such as CBP, DSGS, or resource adequacy contracts with other parties (i.e., utilities, CCA's, or other aggregators)

Table 30: Recommendations

Category	
Nomination Alignment	▪ Align nominations to CAISO with the load reduction capability identified in the evaluation. ▪ Produce nomination forecasts that reflect the resources in place and their weather sensitivity
Operational Planning	▪ Schedule at least one test event per year when all resources are dispatched at the same time for the same event hours to demonstrate the full capability of Enersponse resources. CAISO dispatch varies substantially, making it difficult demonstrate the full resource capability without a coordinate test. ▪ Enersponse may want to consider developing an evaluation operation dispatch plan to introduce intentional variation in event duration, weather, and start times. The goal would be to provide data to better inform performance under planning conditions.
Ex-ante Weather	▪ Work to incorporate hydro conditions, including reservoir levels, and rainfall into ex-ante planning conditions.
Customer Targeting	▪ Targeted recruitment to sites that deliver larger percent reductions, which can be estimated more precisely. Generally, pumping and retail customers drove impacts across the Enersponse portfolio. Further recruitment in these sector may beneficial for achieving higher capacity reductions.

APPENDIX A – PRE-ANALYSIS DATA VALIDATION CHECKLIST

Category	Question	
Interval or monthly data	1	Is the data in the right units? Is it kW or kWh? Did we keep the right units in aggregating data (e.g. 15 minute to hourly data)?
	2	Is the data adjusted for daylight savings time? Does the adjustment match was done for other files (e.g., weather, events)?
	3	Did we properly deal with inflow and outflow from customers with behind the meter generation? Did we properly deal with multi-phase customers that have multiple channels?
	4	Do we have gaps in time? Are we missing data for some customers? Are we missing data for some months or periods? Are there a lot of large customers?
	5	Did we check for outliers or highly influential customers?
	6	Is the data at the right unit of analysis? Are we analyzing data at the service point, premise, or customer level?
	7	Did we check if it is the same customers at the same site? If not, does it matter?
	8	Is the dataset unbalanced? If so, by how much and do we need to do anything about it?
Customer characteristics	9	Do we have all the relevant participant data? Are some customers missing? Do we have extra customers?
	10	Do we have information regarding dual enrollment in other programs? Were these customers affected by other interventions?
	11	Do we have all the relevant characteristics? Or are we missing some to either link to other files? Or characteristics that we use to summarize results?
	12	Do we have a master characteristics file that pull together all the relevant customer information (in case we need to use it later)?
	13	Are we using the right customer size categories?
	14	Are we taking into account program start and end dates?
Treatment or Event data	15	Did the utility/implementer provide us with the correct event/treatment information?
	16	Are there treatment periods where no effect is apparent?
	17	Are there non-treatment periods that reflect a reduction in loads? If so, what is the explanation?
	18	Did we include information about events from overlapping programs? Was it properly coded?
	19	Do all customers get dispatched at the same time or do we have a wide range of dispatch/implementation practices? If the later, have we properly coded when different customers were dispatched/implemented?
	20	Is the treatment variable properly specified?
Weather data	21	Is the daylight saving time adjustment in the weather data consistent with the interval and event data?

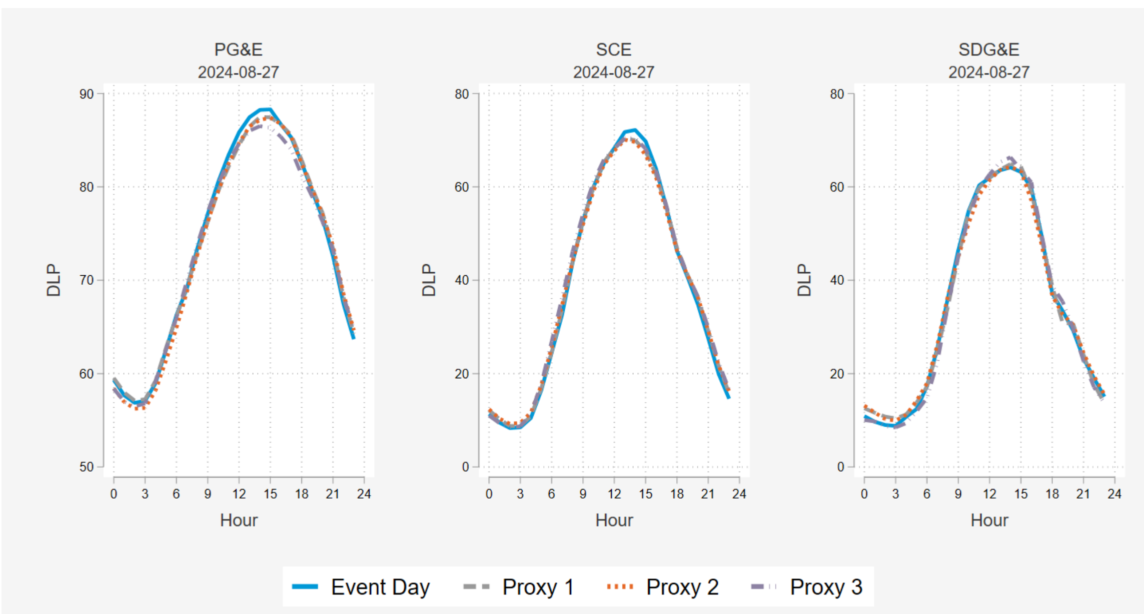
Category		Question
	22	Is weather data in the right time zone (some weather data is provided in UTC)?
Control groups	23	If the control group is based on random assignment, did we run checks to ensure there are indeed no systematic differences?
	24	Was the matching done using usage or load data from the **pre-treatment** period?
	25	How many control candidates per participant do we have?
	26	Are the control group candidates true control group candidates? Or do they experience other interventions (e.g., enrolled in other DR programs)?
	27	Are all the variables used for matching populated (if applicable)? Or are some variables missing for participants or a substantial share of control candidates?
	28	Do weights need to be applied to the control group? Why? Are any of those weights overly large?
	29	Is the participant sample representative of the participant population? Did we lose customers because some variables were blank? Did we lose customers because a suitable match couldn't be found? How influential is this?
	30	Does the number of participants vary substantially over the course of the analysis period?
	31	Do some or all participants have hard eligibility screens (e.g., must have AC to be eligible)? Were those screens applied to the control group candidates?
	32	How did we deal with variables with extreme distributions (where some values can be extremely influential)?
Pre-analysis checks	33	Did we save the analysis dataset file? Was it checked? Are variables populated for everyone?
	34	Did we produce program levels graphs using raw load/usage data to check for any odd patterns?
	35	Do the loads change when we expected them to? Is the treatment effect apparent visually?
	36	How weather sensitive are the customers? How volatile is their load?
	37	If it is an event based program, are there comparable non-event days?
	38	If it the treatment stays on once it is in place, does the effect occur shortly after the treatment is put in place? Or does the treatment precede the effect?

APPENDIX B – EX-POST MODEL TESTING AND PERFORMANCE METRICS

To estimate ex-post load impacts, a model tournament was conducted to determine the most accurate counterfactual model for each individual customer site. This approach ensures that demand reductions attributed to DR participation are not confounded by external factors such as temperature variations or random fluctuations. The model tournament evaluates different modeling approaches and selects the one that provides the most reliable counterfactual estimate for each site.

A critical component of this methodology was the use of proxy days for out-of-sample validation of model accuracy. Proxy days are non-event days that share similar load characteristics with actual DR event days, making them suitable for assessing how well different counterfactual models perform in predicting demand under event-like conditions. Selecting appropriate proxy days was essential to ensuring that the evaluation models were tested against realistic demand patterns rather than average or non-representative conditions. Because there is no event on theses event-like days, the load impact is known to be zero, allowing us to assess model accuracy.

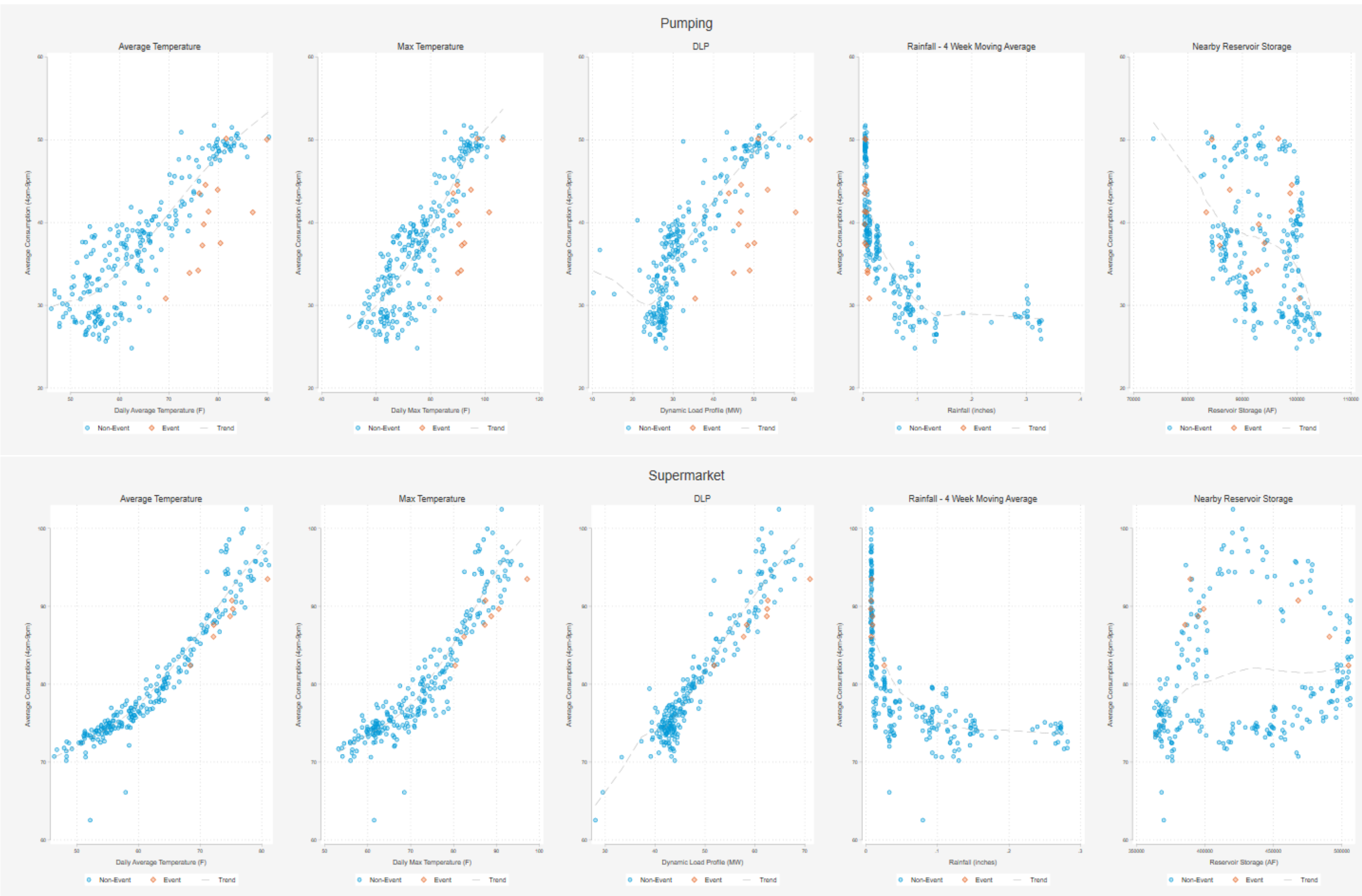
Figure 20: Comparison of Event Day Dynamic Load Profiles to Proxy Days Used for Out-of-Sample Testing



For each utility, three days were selected for each event based on the hourly consumption profile of the most common rate class within that utility's customer base. From this pool, ten proxy days were selected at random for each utility to be used for out-of-sample testing. These proxy days were then randomly divided into three groups. A cross-validation procedure was applied in which one group was left out while the remaining two were included in the training data. This process was iterated three times to ensure that each proxy day received an out-of-sample prediction. The iterative approach allowed some of the proxy days to be included in the model training while also ensuring that extreme, event-like days were being accounted for in model training.

Another key step in the model tournament involved identifying key predictive variables, including historical demand patterns, industry classification, temperature, precipitation, and other exogenous factors that influence electricity consumption. Once the relevant variables were defined, multiple candidate models were estimated for each customer site.

Figure 21: Example of Correlations Between Demand and Key Predictors by Sector



Demand Side Analytics
DATA DRIVEN RESEARCH AND INSIGHTS

Table 31: Summary of Individual Regressions Tested for Accuracy

Explanatory Variables	Models Tested for Accuracy															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Hourly temperature 4 part spline	X	X	X	X	X											
Average daily temperature 4 part spline						X	X	X	X	X						
Mean 17 (Average temp from 12:00 AM - 5:00 PM)											X	X	X	X	X	X
Day of week indicator variables	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Dynamic load profile for corresponding utility and rate class	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
Morning load (12:00 AM to 12:00 PM)		X	X	X	X		X	X	X	X		X	X	X	X	
Solar indicator x solar irradiance			X	X	X			X	X	X			X	X	X	
4 week moving average of rainfall				X	X				X	X				X	X	
CA statewide water reservoir levels					X					X					X	
4 week same weekay same hour average load (lagged load skips event days and non-eligible days)																

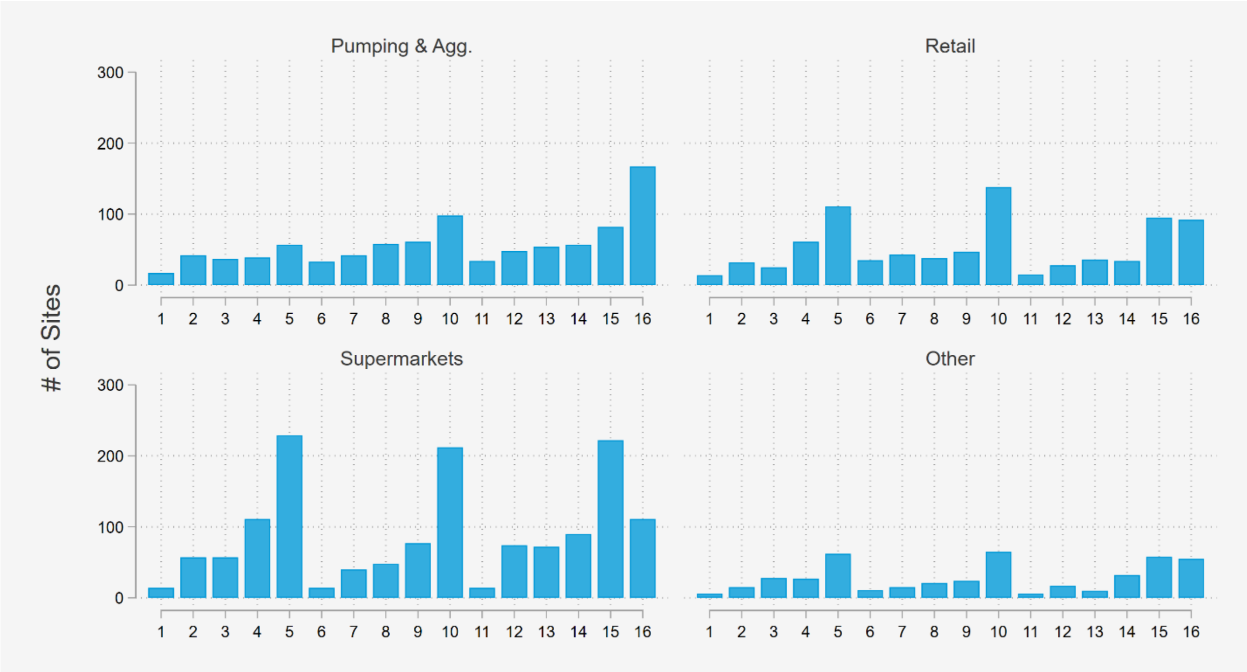
Each model was evaluated using the out-of-sample proxy days to assess predictive accuracy. Two primary metrics guided model selection: percentage bias, which quantifies the tendency to over- or under-predict consumption during event-like days and relative root mean squared error (RRMSE), which measures the noise between actual and predicted loads. To identify the best model for each site:

- The candidates were narrowed to unbiased models defined as the models with bias under +/-1% or the three models with lowest bias.
- Then, the model with the lowest RRMSE among the unbiased models was selected as the counterfactual model.

By leveraging a model tournament approach and incorporating proxy days into the validation process, the evaluation framework ensured that the most accurate counterfactual is applied at the site level. The selected models were used to estimate site-specific DR reductions, which were then aggregated to generate program-wide impact estimates.

Figure 22 summarizes the distribution of the best model by sector. For most sectors, the model that include weather plus all the additional explanatory variables was selected most often. The exception is pumping and agriculture, where the model that did not include weather was selected most often based on the model accuracy tournament.

Figure 22: Summary of Best Model Distribution by Sector Group



While best models were identified for each individual level, a more critical question is the performance for the portfolio as whole. The portfolio performance was assessed by aggregating the predicted and actual values for the out-of-sample proxy event days. This was done for each model and using the best model for each individual site.

Table 32 summarized the portfolio level performance metrics using the out-of-sample proxy event days. Overall, picking the best model for each site lead to results that do not over or underestimate the loads (% bias) and have the highest precision (MAPE, RRMSE, and R-squared). Figure 23 plots the predicted and actual values on the out-of-sample proxy event days. It also includes the identify line (y=x) to help assess if the predicted and actual values deviates. Figure 24 compares the hourly predicted and actual loads for all the out-of-sample proxy event days. At a portfolio level, the model predictions are precise, accurate, an unbiased.

Table 32: Portfolio Level Model Tournament Performance Metrics on Out-of-Sample Proxy Event Days

Model Number	Average Actual Load (kW)	Average Predicted (kW)	Average Error	% Bias	Absolute % Bias	MAPE	RRMSE	R Squared
Best Model	114,421	114,408	-13.54	-0.01%	0.01%	0.77%	0.012	99.96%
1	114,421	114,650	228.54	0.20%	0.20%	3.44%	0.037	99.59%
2	114,421	114,447	25.73	0.02%	0.02%	0.85%	0.013	99.95%
3	114,421	114,448	26.39	0.02%	0.02%	0.86%	0.013	99.95%
4	114,421	114,432	10.80	0.01%	0.01%	0.85%	0.013	99.95%
5	114,421	114,406	-15.27	-0.01%	0.01%	0.80%	0.012	99.96%
6	114,421	114,809	388.06	0.34%	0.34%	3.51%	0.040	99.53%
7	114,421	114,427	5.89	0.01%	0.01%	0.91%	0.014	99.94%
8	114,421	114,428	7.14	0.01%	0.01%	0.90%	0.014	99.94%
9	114,421	114,405	-15.91	-0.01%	0.01%	0.90%	0.014	99.94%
10	114,421	114,377	-44.51	-0.04%	0.04%	0.85%	0.014	99.95%
11	114,421	115,046	625.23	0.55%	0.55%	3.60%	0.038	99.57%
12	114,421	114,464	43.00	0.04%	0.04%	0.89%	0.014	99.94%
13	114,421	114,465	44.21	0.04%	0.04%	0.89%	0.014	99.94%
14	114,421	114,444	23.07	0.02%	0.02%	0.90%	0.014	99.94%
15	114,421	114,451	29.70	0.03%	0.03%	0.87%	0.015	99.93%
16	114,421	114,460	39.12	0.03%	0.03%	0.98%	0.016	99.93%

All Metrics were computed using out-of-sample prediction on proxy event days

Figure 23: Portfolio Level Comparison of Predicted Versus Actuals on Out-of-Sample Proxy Event Days

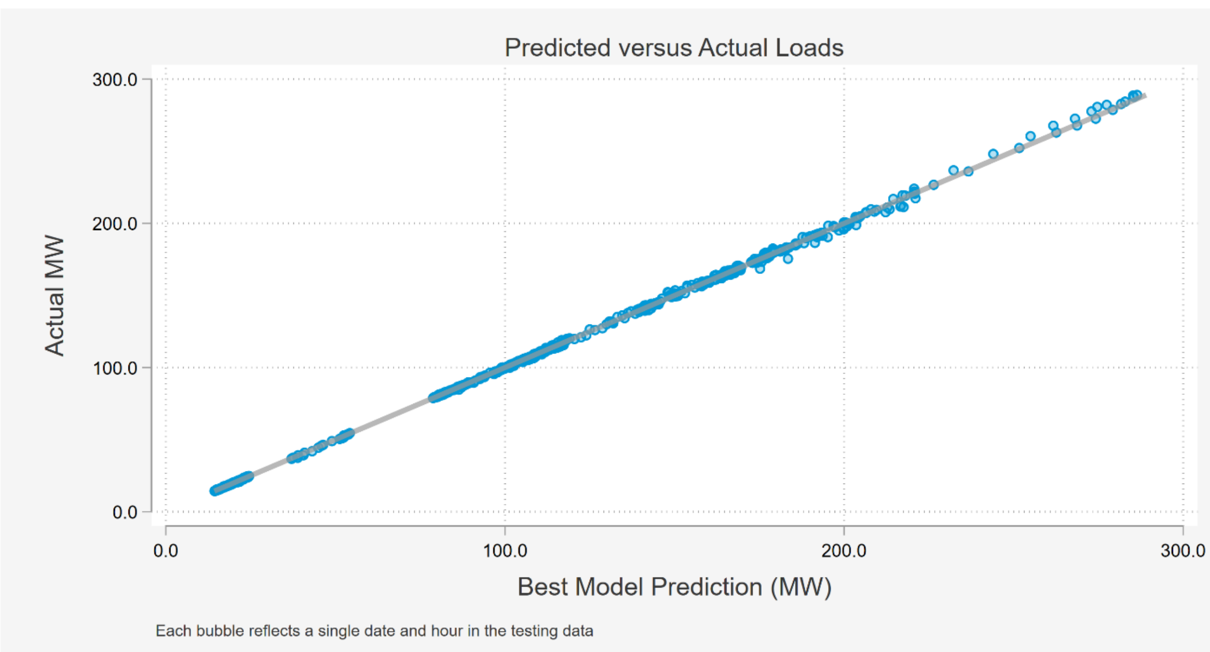


Figure 24: Portfolio Level Hourly Predicted Versus Actuals on Out-of-Sample Proxy Event Days

APPENDIX C – EX-ANTE MODEL TESTING AND PERFORMANCE METRICS

DSA's first step in the ex-ante analysis involved estimating reference loads using a fixed-effects modeling approach with a temperature spline specification. The model included dummy variables for temporal effects to account for seasonal and day-of-week variations in consumption. This modeling framework was applied separately for each hour of the day, each sub-LAP, and across different industry groupings.

The reference load model was specified as follows:

Equation 1: Reference Load Regression Specification

$$kWh_{i,t}^{g,s,h} \;=\; \beta_0^{g,s,h} + \sum_{j=1}^{4}(\beta_j{}^{g,s,h} * \; TempSpline_{j,i,t}) + \sum_{m=1}^{12}(\delta_m^{g,s,h} * \; D_{month,m,t}) + \gamma^{g,s,h} D_{weekday,t} + \alpha_i + \; \varepsilon_{i,t}^{g,s,h}$$

Where:

Term	Description
$\boldsymbol{kWh_{i,t}^{g,s,h}}$	The consumption for customer i, at time t, within the industry group g, sub-LAP s, and hour h
$\boldsymbol{\beta_0^{g,s,h}}$	Intercept, specific to the unique combinations of industry, sub-LAP, and hour
$\boldsymbol{TempSpline_{j,i,t}}$	The temperature spline terms based on the average daily temperature (with knots at 55-, 60-, and 65-degrees Fahrenheit) for customer i at time t. The associated coefficients $\beta_j{}^{g,s,h}$ captures how load responds to temperature for each industry, sub-LAP, and hour combination
$\boldsymbol{D_{month,m,t}}$	Monthly dummy variable. $\delta_m^{g,s,h}$ allows the effects to vary across each combination of industry, sub-LAP, and hour
$\boldsymbol{D_{weekday,t}}$	Weekday vs weekend dummy variable. $\gamma^{g,s,h}$ allows the effects to vary across each combination of industry, sub-LAP, and hour
$\boldsymbol{\alpha_i}$	Individual (site-level) fixed effects, capturing the time-invariant differences in baseline consumption for each customer i
$\boldsymbol{\varepsilon_{i,t}^{g,s,h}}$	Error term

DSA then estimated impacts using a model tournament methodology similar to the approach employed in the ex-post analysis. The best models were selected for each combination of industry grouping and sub-LAP. These models also included individual fixed effects, enabling the predictions to reflect the average impact observed for each individual site. Equation 2 shows an example regression specification, while Table 33 shows the different models tested.

Equation 2 –Example Impact Regression Specification

$$Imapct_{i,t}^{g,s,h} = \beta_0^{g,s,h} + \beta_1^{g,s,h} * \widehat{ReferencekWh}_{i,t}^{g,s,h} + \alpha_i + \varepsilon_{i,t}^{g,s,h}$$

Where:

Term	Description
$Imapct_{i,t}^{g,s,h}$	The impact for customer i, at time t, within the industry group g, sub-LAP s, and hour h
$\beta_0^{g,s,h}$	Intercept, specific to the unique combinations of industry, sub-LAP, and hour
$\widehat{ReferencekWh}_{i,t}^{g,s,h}$	The predicted reference load using the reference load model
α_i	Individual (site-level) fixed effects, capturing the time-invariant differences in baseline consumption for each customer i
$\varepsilon_{i,t}^{g,s,h}$	Error term

Table 33: Ex-Ante Models Tested for Accuracy

Explanatory Variables	Models Tested for Accuracy	
	1	2
Estimated Reference Load	X	X
Cooling Degree Days 65 °F Base		X

To evaluate the model performance when estimating impacts, DSA separated DR events into three groups and conducted a cross-validation process. Each model was trained three times, with one group left out in each iteration. The left-out group served as an out-of-sample testing set, allowing DSA to generate predictions using each model specification. The accuracy and precision of each model were assessed by comparing the predicted impacts to the observed impacts. Table 34 shows the aggregated average hourly impacts and predictions paired with their error metrics. Best Model in this context reflects the predictions made when each sub-LAP and industry combination uses the most accurate model.

Table 34: Ex-Ante Model Accuracy

Model Number	Average Aggregated Impact (MW)	Average Aggregated Prediction (MW)	Average Error	% Bias	Absolute % bias	RMSE	RRMSE	R squared
Best Model	7340.89	7353.00	12.10	0.16%	0.16%	2355.41	32%	92%
1	7340.89	7324.43	-16.46	-0.22%	0.22%	2429.07	33%	91%
2	7340.89	6861.97	-478.93	-6.52%	6.52%	2790.29	38%	89%

Figure 25 highlights the accuracy of using the best model for each sub-LAP and industry. Overall, the predictions aligned with the observed impacts across the full spectrum of temperature conditions.

Figure 25: Aggregated Hourly Impacts Over Temperature – Best Model Predictions vs Actual