DOCKETED	
Docket Number:	21-DR-01
Project Title:	Supply Side Demand Response
TN #:	246240
Document Title:	Demand Side Analytics - Final Draft Proposal
Description:	N/A
Filer:	Courtney Wagner
Organization:	California Energy Commission
Submitter Role:	Commission Staff
Submission Date:	9/28/2022 1:19:24 PM
Docketed Date:	9/28/2022

FINAL DRAFT PROPOSAL

Demand Response Qualifying Capacity Working Group Proposal

Prepared by:
Demand Side Analytics
San Diego Gas & Electric

TABLE OF CONTENTS

1 INTRODUCTION

This report is prepared for consideration by members of the California Energy Commission (CEC) working group on the Qualifying Capacity (QC) of demand response (DR) resources. The proposal supersedes the draft proposal submitted on April 28, 2022. Pacific Gas and Electric (PG&E), Southern California Edison (SCE), San Diego Gas & Electric (SDG&E), and Demand Side Analytics (DSA) met for over 12 work sessions. The group met to formulate workable solutions, develop models, run applied tests, and gather feedback from specific entities. In this process, SCE asked questions and provided feedback. PG&E cannot endorse the proposal at this time because they need more time to vet it with multiple organizational units. After additional discussion, PG&E will determine whether to support the proposal in whole or in part. In addition, the team gathered feedback from CAISO, CPUC, and CLECA to ensure we understood their concerns and to identify areas that needed further refinement.

As the resource mix in California is changing to meet de-carbonization goals, the need for flexible resources has increased. As such, it is now more important than ever that we accurately estimate the capability of DR resources for planning and operations. However, the process for developing the qualifying capacity value of DR resources is complex. DR resources include a wide range of technologies and customer segments. They can vary in shape, weather sensitivity, and operating limitations such as the maximum event duration, annual hours of dispatch, and the number of consecutive dispatch days. While there are many aspects of developing the value of DR resources that will continue to be discussed, we seek to answer three main questions in this proposal:

1) **How do we determine the ex-ante DR capability under different conditions?** Specifically, how can we develop ex-ante values that can be used for planning and also reflect how DR is expected to perform under a range of operating conditions?
2) **How are the characteristics of DR accounted for in determining the slice of day values by month and day?** Specifically, how does the approach account for the coincidence of DR with resource needs and for its limitations on availability, event duration, and frequency of dispatch?
3) **How do we measure DR performance?** How can we measure whether the ex-ante values used for planning align with bids and actual event performance?

In addition to answering these questions, our proposal has a few overarching goals:

- Provide greater transparency
- Produce a framework that accounts for the characteristics of each resource, including coincidence with reliability risk, weather sensitivity, resource availability, maximum event duration, and limitation on the annual hours, monthly hours, and consecutive days of dispatch
- Generate greater alignment between DR planning and operations;
- Produce estimates of DR capability that align with the slice of day framework;

- Ensure accurate measurement of the demand reductions delivered
- Ensure accurate estimation of resource capabilities under planning conditions
- Develop standardized metrics for measuring if bids and actual event performance align with the ex-ante values

The proposal focuses on technical aspects of how to align the DR outputs to fit the 24-Slice of day resource adequacy framework. It does not discuss how and where to simplify the Load Impact Protocols, or how and where to simplify the process and shorten the timelines. However, we are open to both simplifying the Load Impact Protocols and process improvements to reduce the burden on DR providers, the CPUC, and other stakeholders.

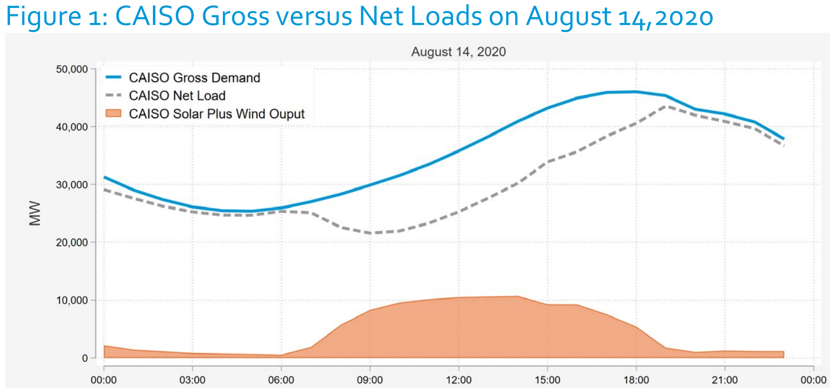
The remainder of the proposal is divided into two main sections and six technical appendices. Section 2 presents the California context and motivation. Section 3 contains the proposal, which we have intentionally kept concise. The appendices include technical detail. Appendices A and B describe how to produce the slice day table and the workbook used to test the process. Appendices C describes how to produce a time-temperature matrix, and Appendix D contains an applied example. Appendices E and F provide examples of a performance alignment metric and a bid alignment metric for illustrative purposes.

2 CALIFORNIA CONTEXT AND MOTIVATION

The fundamental nature of how electricity is generated, transmitted, distributed, and used in California changed substantially in the past ten years and will continue to evolve in the next decade. The single largest change affecting California's electric grid is the de-carbonization goals. The penetration of intermittent utility-scale renewable generation, mostly in the form of large solar power facilities and wind farms, has grown substantially in the past decade. In 2021, solar resources delivered up to 13,000 MW and wind resources exceeded 6,000 MW.[1] In addition, residential households and businesses are installing behind-the-meter solar, installing battery storage, and increasingly adopting electric vehicles.

Historically, the electric grid infrastructure has been sized to meet the aggregate peak demand of end users with a reserve margin for extreme weather or unforeseen outages. The electric system is unique in that it is necessary to balance supply and demand at all times. An imbalance can lead to cascading outages and compromise the reliability of the entire grid. Because electricity storage was prohibitively expensive in the past, enough supply capacity and flexibility had to be built to accommodate peak demands, and enough reserves had to be maintained to withstand unforecasted changes in the supply-demand balance (e.g., generator and transmission outages). However, the technology for energy storage has evolved, and the costs are declining. California's generation interconnection queue includes a large amount of battery storage.

Figure 1: CAISO Gross versus Net Loads on August 14,2020



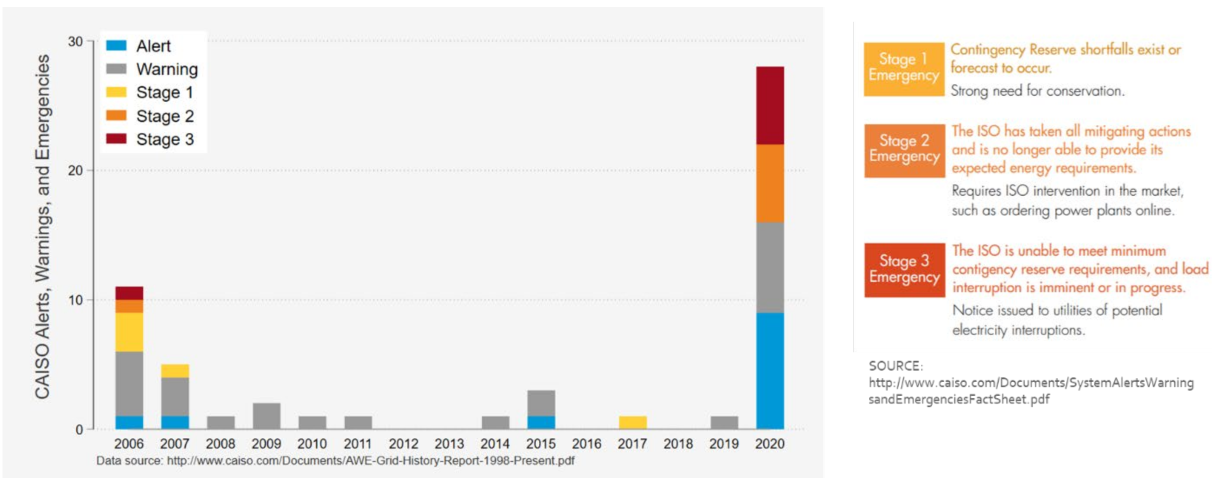
The introduction of large-scale solar and wind has led to fundamental changes in planning the electric grid. The focus has shifted from planning for gross peak demand to net peak demand – electricity demand minus large-scale solar and wind. The grid must now focus on having sufficient dispatchable resources to meet the demand that cannot be met using solar and wind resources. Figure 1 illustrates the concept of net loads versus gross demand. It shows the electric demand and the wind and solar production on August 14, 2020, a day when California had experienced a shortage in resources. While gross demand peaks in the late afternoon, net loads peak a couple of hours later, when solar production declines as the sun sets. The ongoing changes lead to a cleaner supply mix, but also affect the magnitude and type of resources and grid services required to maintain

[1] CAISO press release. http://www.caiso.com/Documents/California-ISO-Hits-All-Time-Peak-of-More-Than-97-Percent-Renewables.pdf

reliability. They place a premium on flexible resources: enough flexibility is needed to adjust supply to meet fluctuations in demand and fill gaps when solar and wind power are unavailable.

In 2020, California experienced a confluence of extreme weather and widespread fires, leading to a historic number of CAISO emergency events, including rolling blackouts. The emergencies occurred due to a mix of high demand, unusual weather conditions, lower than forecasted solar output, operator forecasting error, and planning paradigms focused on gross demand rather than net loads. Demand response played a critical role in helping reduce demand when resources were needed. In 2020, the resources shortages did not occur when gross peak demand was at its highest but later in the evening when net loads (demand minus solar and wind) peaked.

Figure 2: Historical CAISO Alerts, Warnings, and Emergencies



2.1 DEMAND RESPONSE RESOURCES AND CAISO PEAKING PATTERNS

Historically, demand response programs have been designed to reduce peak demand and offset the need for additional peaking capacity. When, where, how often, and for how long DR resources are needed are evolving due to the introduction of large amounts of intermittent renewable resources.

A fundamental characteristic of power system planning is that a small number of hours drive a significant share of costs. Electric prices climb sharply when the grid is strained due to high demand, generator outages, transmission outages, fluctuations in power output, or forecast error. Resource shortages typically occur due to high net load demand levels and a combination of generator outages, transmission outages, low imports, or unforecasted fluctuations in solar or wind output.

Figure 3 shows the concentration of CAISO high net load hours and days. The panel to the left is a load duration curve, which ranks the top 5% of hours based on net loads from highest to lowest. The panel to the right shows the hourly patterns on the ten days with the highest CAISO net loads. Net loads are the primary driver of resource capacity needs and are highly concentrated. The net loads in roughly 1% of the hours in the year drive the need for 18% of the capacity resources (over 9,000 MW with the reserve margin). Moreover, the timing of the high net loads is concentrated in the summer months and on specific hours. Figure 4 shows a heat map of CAISO net loads in 2020. Even in unusual years, such as 2020, the risk of resource shortages is concentrated in a limited number of hours in the summer months and driven by heatwaves.

Figure 3: CAISO Concentration of High Net Load Hours and Days

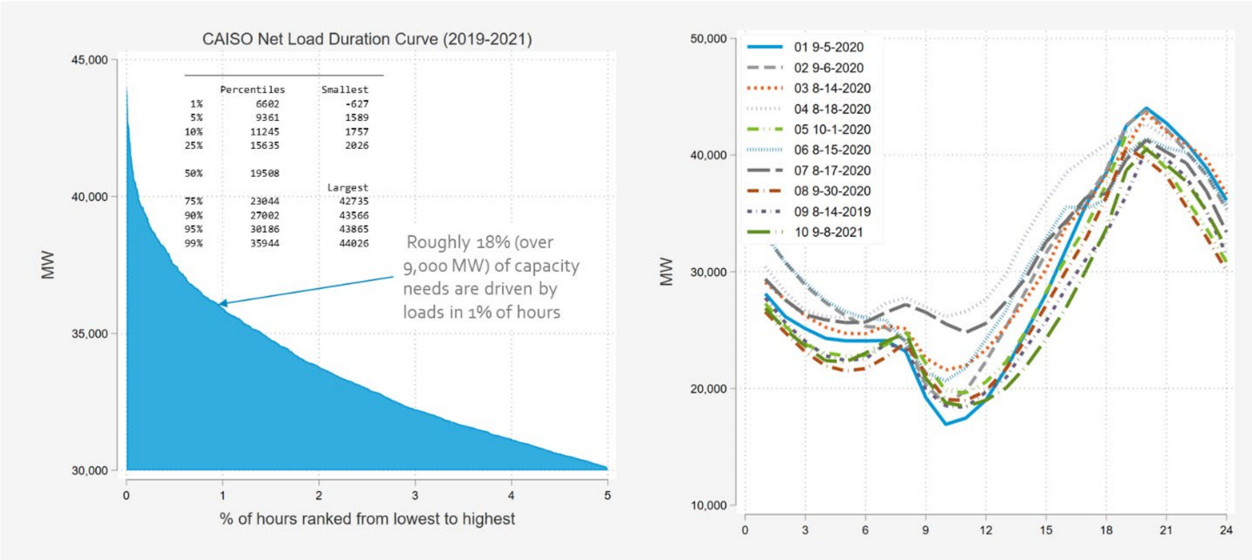
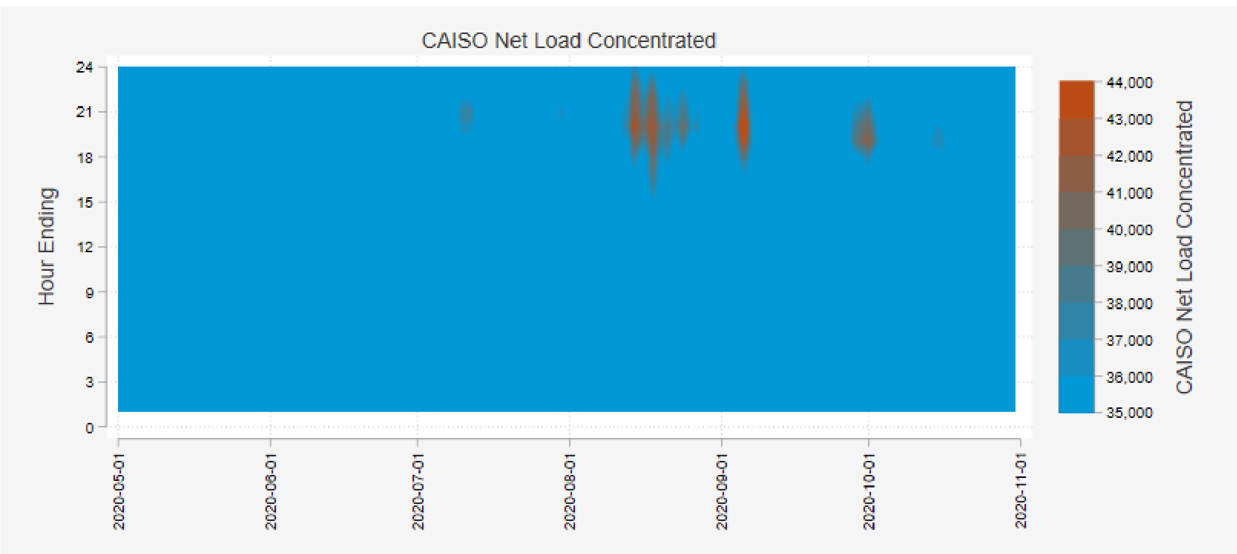
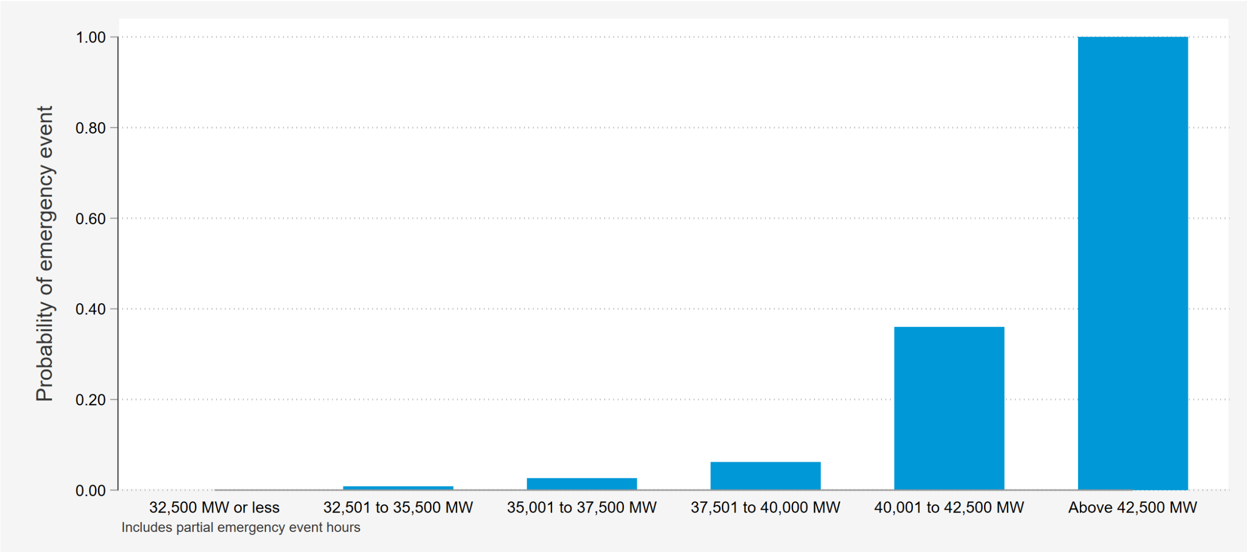


Figure 4: Heat Map of CAISO Net Load in Summer 2020



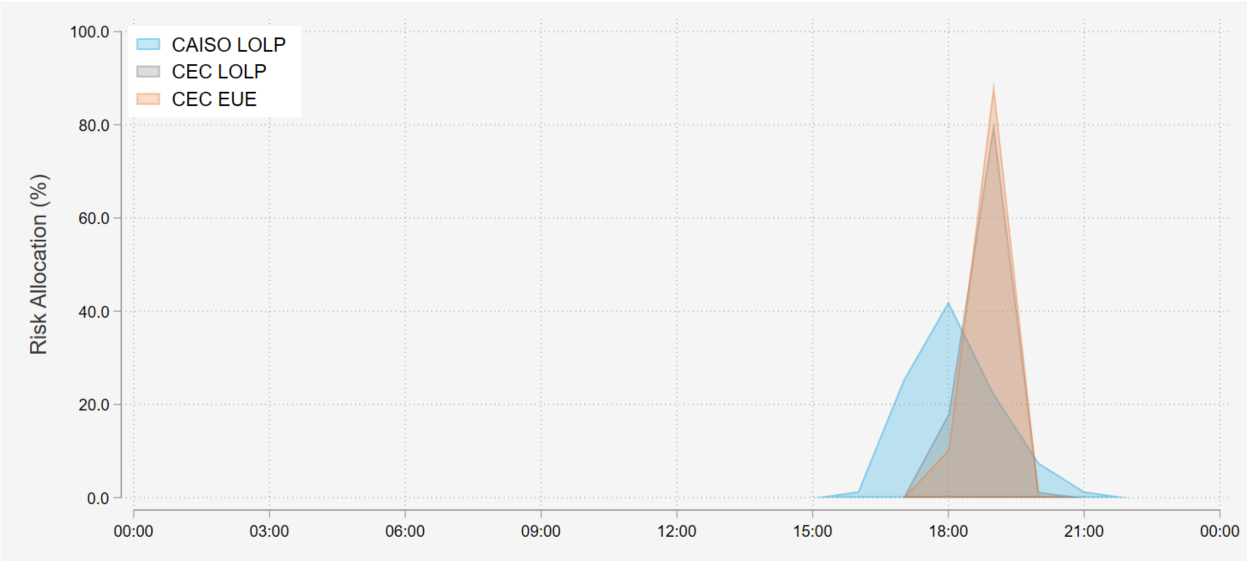
High net loads are closely related to resource shortages, as measured by CAISO emergency notices, which are directly linked to the available reserve margin. Figure 5 shows the relationship. The probability of resource shortages in 2019-2021 was directly linked to net loads. The risk of resource shortages was highest when loads exceeded 40,000 MW.

Figure 5: 2019-2021 Relationship Between Net Loads and CAISO Emergency Events



In 2021 and 2022, CAISO and the CEC, respectively, conducted reliability planning studies and quantified the risk of resource shortfalls using loss-of-load probabilities (LOLP) or expected unserved energy (EUE). The results from the studies also indicate the risk of resource adequacy shortages is highly concentrated in a limited number of hours.

Figure 6: Risk of Capacity Shortfalls is Highly Concentrated in Limited Hours



To help meet resource adequacy requirements, DR resources need to be dispatched in the right months and right hours when net loads are high. Because net loads drive planning needs, the framework of DR qualifying capacity must account for the level of solar and wind penetration. DR includes a wide range of resources ranging from residential thermostats and behind-the-meter batteries to large industrial customers, each with differing capabilities on when, how often, how long, and how much demand reduction they can deliver. It is our position that any resource adequacy and qualifying capacity framework must properly incorporate and model the use limitations of DR resources and their coincidence with resource needs. DR resources also interact with battery storage. Both resources effectively aim to shave the net load duration curve, targeting the hours when resources are needed most. Higher amounts of peak shaving resources effectively mean that the resources must be dispatched more often to shave the load duration curve.

The main takeaways are simple:

- **Planning has shifted from gross loads to net loads.** Wind and solar are effectively the base supply resource but are inherently intermittent.
- **Electricity infrastructure costs are currently driven by net loads which are highly concentrated, peaking on a limited number of hours and days.** Over 9,000 MW of capacity resources (18%) are needed due to high net loads in less than 1% of hours.
- **Empirically, high net loads are closely linked to resource shortages.** The likelihood of shortages increases as net loads grow.
- **To deliver resource adequacy, DR resources need to be dispatched in the months and hours when net loads are high**. Because net loads drive planning needs, the DR QC framework must account for the level of solar and wind penetration. DR resources are not needed for all the roughly 720 hours each month to ensure resource adequacy.
- **DR also interacts with battery storage** since both resources have use limitations; target the hours when resources are needed most; and aim to shave the net load duration curve.

2.2 SHORTCOMINGS OF THE CURRENT FRAMEWORK

The current framework is often referred to as the Load Impact Protocol (LIP) framework. The Load Impact Protocols were designed to produce standardized outputs to use to track historical performance and to inform planning and resource adequacy. The protocols themselves did not specify how load impacts should be used for resource adequacy, did not limit the ability to run updates, and did not set the timelines for approval. Subsequent decisions by the CPUC led to the use of the average hourly impact over the availability window (4-9) to produce a monthly qualifying capacity value and set forth the process and timelines for approval of qualifying capacity. At the time the LIP protocols were approved in 2008, resource adequacy was driven by system gross peak loads, while today they are driven by net loads.

Table 1: Summary of Current Process

Component	Detail
What were the actual demand reductions delivered under the conditions called (ex-post impacts)?	For simplicity, these are called ex-post impacts. The goal is to provide the most accurate estimate of the delivered demand reductions. Most evaluations conduct accuracy tournaments testing different models, and many rely on matched control groups with difference-in-differences using smart meter data. The protocols require producing hourly results for each event in a standardized format, including information about the number of participants called, event start and end times, weather conditions, and confidence intervals. It also requires validation of the accuracy of the method used to produce the load impacts. Notably, the CAISO settlement does not match the evaluation results. CAISO settlement usually relies on heuristic methods – e.g., same hour average for the past ten (10) days – which can be implemented quickly and is easier for customers to understand.
What is the magnitude of program resources available under standard planning conditions (ex-ante impacts)?	For simplicity, these are called ex-ante impacts. They rely on developing a predictive model using hourly reductions from historical events, typically the most recent three years. The objective is to model how reductions vary as a function of weather, hour-of-day, hours into the event, and other factors (e.g., cycling strategy, location, etc.). This model is then used to predict demand reduction capability for each hour under 1-in-2 and 1-in-10 weather conditions and standardized dispatch hours that align with resource adequacy planning (currently 4-9 PM). The results are hourly tables with the load reduction capability for each month for 1-in-2 and 1-in-10 weather years
What value is used to determine the qualifying capacity?	Even though the outputs are hourly, the CPUC currently uses the average for the 4–9 PM time period under 1-in-2 utility peak conditions to determine the qualifying capacity for each month. The CPUC also specifies minimums a DR resource must meet to qualify for capacity. Currently, DR resources must be available Monday through Saturday for four (4) consecutive hours between 4 PM and 9 PM, and at least 24 hours per month from May to September. The DR qualifying capacity per customer relies on the load impact evaluation from two years ago (e.g., the 2023 qualifying capacity is based on the 2021 evaluation). Demand response providers can update enrollments, but only under limited circumstances.

In practice, utilities, CAISO, planners, operators, and program managers need to understand the magnitude of resources available for different hours under various temperature conditions, for different start times, and for different event durations. Actual events reflect on-the-ground decisions and do not always align with planning values. Specifically, actual weather conditions do not frequently match the 1-in-2 and 1-in-10 weather year planning conditions, and the event start times and durations often differ from the 4-9 PM resource adequacy window. Moreover, DR events are called for multiple reasons –testing or evaluation, economic dispatch, and reliability-related alerts, warnings, and emergencies. The current process has several limitations, summarized in Table 2.

Table 2: Limitations of Current DR Qualifying Capacity Framework

Limitation	Explanation
It does not incorporate the hourly capability of the resources	The current approach uses the average hourly load impacts from 4-9 PM under 1-in-2 peaking conditions for each month. It does not reflect the hourly load reduction capability, even though ex-ante values are produced on an hourly basis.
It does not fully factor in the coincidence of the resource shape with the risk of capacity shortages	The risk of capacity shortages is highly concentrated on specific hours when net loads are high, as shown by the recent CAISO and CEC reliability studies. Many DR resources are also tied to an underlying load shape – e.g., air conditioners or C&I load – and some of those resources deliver larger demand reductions when weather and demand are more extreme. Simply put, not all hours between 4–9 PM are equal. Thus, the coincidence of the DR resources with the hours when the risk is highest should be a critical component of determining the DR qualifying capacity value.
Is difficult to assess if performance during operations and bids into CAISO and align with the planning values	Actual events reflect on-the-ground decisions and do not always align with planning conditions. The actual weather conditions often do not frequently match the 1-in-2 or 1-in-10 weather conditions, and the event start times and durations often differ from the 4-9 PM resource adequacy window. Because of the format of the outputs, it can be difficult to compare the resource capability under planning conditions to bids or to compare them to the performance during actual events. This is particularly true for weather-sensitive programs that deliver lower reductions on milder days and larger reductions on hotter days when resources are needed most. In addition, the comparisons are sometimes inconsistent about whether the behind-the-meter demand reduction are scaled up to account for transmission and distribution line losses or the planning reserve margin. Last but not least, evaluation results are often used to assess performance, which does not always match the CAISO settlement. CAISO settlement is typically conducted using heuristics – day matching baselines – which are easy to understand and easy to compute. By contrast, evaluations often use accuracy tournaments, control groups, and techniques such as difference-in-difference regression models.
It lacks the flexibility needed for the 24-hour slice of day resource adequacy framework	The existing framework aligns well with the new 24-hour slice of day resource adequacy framework in several aspects. The demand response capability is produced by month and hour for standardized 1-in-2 and 1-in-10 system peak conditions and reflects spillover effects (e.g., pre-cooling and snapback). However, DR providers will need the flexibility to target the hours that maximize value and coincide with need. A standard 4–9 PM dispatch window may not be adequate. Several DR resources can deliver reductions for more than five hours and are also available outside of 4–9 PM. Resources that experience performance decay, such as thermostat control programs, can maximize value by avoiding early dispatch and targeting the most critical hours.

3 PROPOSAL

The current method requires producing hourly results for each event in a standardized format, including information about the number of participants called, event start and end times, weather conditions and confidence intervals. It also requires DR providers to produce estimates of DR capability by month and hour for 1-in-2 and 1-in-10 planning conditions that are ground in actual event performance when possible. The core elements of the existing framework align well with the slice-of-day resource adequacy framework, which also requires estimates of resource capability by month and hour-of-day. The proposal has nine main elements.

1

The Load Impact Protocols (LIP) should be retained but modified to address the 24-hour slice-of-day framework. Specifically, the protocols should continue to:

- Require that ex-ante load impact be grounded on actual event demand reductions when possible
- Require reporting of hourly load impacts for each event in a standard hourly format
- Require reporting of resource capability under planning conditions (ex-ante impacts) on an hourly basis for each month
- Provide flexibility in methods and models for ex-post evaluation and ex-ante impacts. Based on over a decade of applied experience, it is clear that no single ex-ante model fits all programs. Rather than focus on the models, we believe the focus should be on standardized outputs and transparency.

The team is open to modifications to simplify, add transparency, and further standardize outputs. The team is also open to streamlining the process to make it more concise and timely. However, modifications to the load impact protocols require technical expertise and testing – they should be done with caution.

2

Modifications to the Load Impact Protocols should include:

- Aligning weather conditions with the worst day of the month as defined in resource adequacy.
- Allowing DR providers flexibility to target the hours that maximize value and coincide with need (i.e., don't force everyone into 4–9 PM) while taking into account:
  - ✓ The coincidence of the resource with the risk of capacity shortages
  - ✓ The availability of the resource as defined by the program rules (e.g., 12–9 PM) by month, hour, and weekday/weekend conditions
  - ✓ Max event duration
  - ✓ Spillover effects such as snapback, pre-cooling, or persistence of load reductions beyond the event window (for non-residential).
  - ✓ Minimum requirements for annual maximum dispatch hours, monthly maximum dispatch hours, and maximum consecutive days

- Ensuring the load impacts for the worst day of the month is an output of the ex-ante impacts
- Produce a summary 24-slice of day table that shows impact of the resource for all 24 hours for each of the 12 months on the worst day. The table must meet the resource adequacy requirements, match the load impact protocol tables, and include all spillover effects. The below table serves as an example.

Hour	January	February	March	April	May	June	July	August	Septembe	October	Novemb	Decemb
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	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
17	0.00	0.00	0.00	0.00	46.44	86.60	89.22	91.81	89.57	82.31	65.14	0.00
18	0.00	0.00	16.84	52.53	36.20	74.64	80.02	84.14	79.52	74.41	47.92	0.00
19	0.00	0.00	15.13	39.60	24.94	66.29	71.19	77.54	69.96	65.96	22.69	0.00
20	0.00	0.00	9.50	18.11	11.44	44.24	53.00	59.54	49.12	38.86	5.29	0.00
21	0.00	0.00	0.44	5.29	2.30	23.76	34.99	37.54	31.93	18.11	1.29	0.00
22	0.00	0.00	0.00	4.49	0.00	24.70	41.44	40.84	39.15	19.94	8.29	0.00
23	0.00	0.00	0.00	-0.06	0.00	-7.40	-14.44	-10.95	-13.50	-5.95	-4.68	0.00
24	0.00	0.00	0.00	0.00	0.00	-1.66	-4.31	-4.56	-6.64	-2.80	-1.75	0.00

- Production of a Time-Temperature Matrix for weather-sensitive resources using a standard output format upon request. A time-temperature matrix quantifies the relationship between demand reductions, temperature conditions, the hour of the day, event start times, and hours into an event. It is based on the same model used to produce ex-ante impacts under planning conditions. Including a time-temperature matrix would better reflect the range of the resource capabilities that are not captured by a single planning value for each month (or a 24-hour profile for each month) and help bridge the gap between operations and planning

We note that the request for flexibility in choosing the event window is not without boundaries. The resource needs to fulfill the minimum resource adequacy requirements, as they are defined.Once the minimum requirements are met, the DR provider can choose additional hours to show DR impacts. However, once a DR provider has elected the hours to show reductions, it cannot modify them since it fundamentally alters the 24-hour slice of day stack. To illustrate, consider a resource that can be dispatched for six event hours and assume resource adequacy requires resource between 4 PM – 9 PM, with a four-hour duration minimum. The resource could elect to show a 3–9 PM or a 4–10 PM reduction window. In both cases, it would need to include all spillover effects, whether positive or negative, for all 24 hours (if any).

The goal of the modifications is to show the full effects of the DR, good and bad, across all 24 hours, consistent with the 24-slice-day framework.

3

The long-term DR qualifying capacity methodology should be applicable to both supply-side and load-modifying DR resources

4

A single entity (CPUC, CEC, CAISO) should produce the reliability risk heatmap in advance (e.g., 18 months before the RA compliance year). This enables DR providers to adjust programs and slice-of-day estimates to coincide with the hours when resource needs are greatest.

5

The ex-post load impact from evaluations should be used as the basis for performance:

- The impacts are more use the best available method and typically rely on an accuracy tournament or matched control groups with difference-in-differences
- There is a long history of load reductions in a standard template (since 2008)

6

CAISO should allow evaluation results to be used for settlement if:

- The evaluation plan is produced in advance of the season
- The results are produced within the settlement period
- The statistical analysis code to produce the results is made available to CAISO for replication

7

Develop a standardized performance alignment metric. The main objective of this metric is to assess if the actual performance during operations aligns with the historical forecasted capability at the meter, given the conditions actually experienced during operations and the resources dispatched. By design, the metric is centered on 1.00, with values above 1.00 indicating overperformance and values below 1.00 indicating underperformance. We introduce an applied example of calculating the metric. Still, we recognize that stakeholders may want additional discussion and the opportunity to test it in practice before it is adopted. The metric and workbook with underlying calculations would be available to the CPUC, CEC, and CAISO upon request.

8

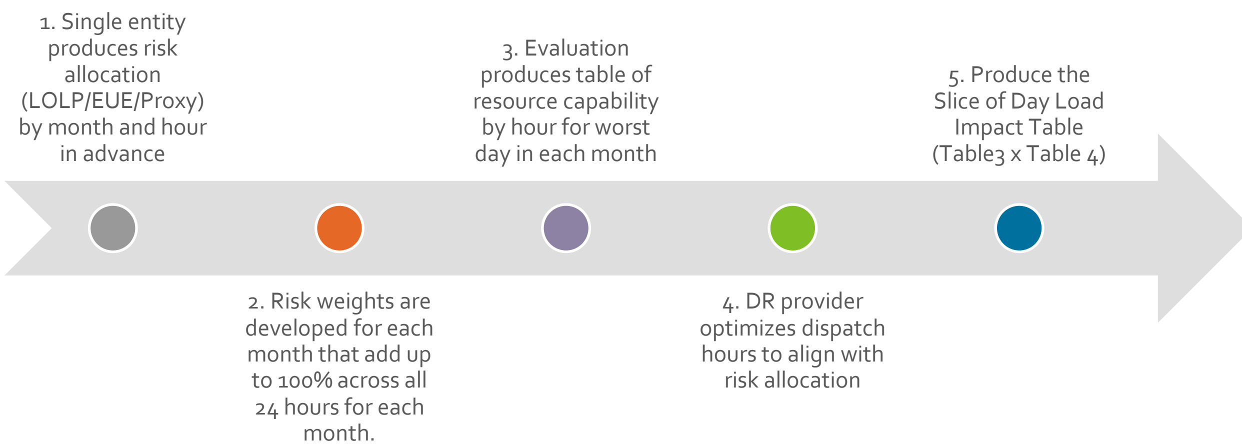
Develop a standardized bid alignment metric. The main objective of this metric is to assess if the bids align with the historical forecasted capability, given the conditions actually experienced. By design, the metric is centered on 1.00, with values above 1.00 indicating overperformance and values below 1.0 indicating underperformance. We introduce an applied example of calculating the metric. However, we recognize that stakeholders may want additional discussion and the opportunity to test it in practice before it is adopted. The metric and workbook with underlying calculations would be available to the CPUC, CEC, and CAISO upon request.

9

Work out the methodology for the monthly qualifying capacity value in the Resource Adequacy Working group, starting with the one on September 21, 2022.

APPENDIX A: PRODUCING THE SLICE OF DAY TABLE

The figure below outlines the key steps for producing a slice of day table. Each step is outlined in greater detail with an example in the table on the following page. The process can also be used to produce monthly qualifying capacity values consistent with the slice of day framework.



Description	Example

1 Single entity produces risk allocation (LOLP/EUE/Proxy) by month and hour in advance

- Must be provided in advance to allow DR providers to adjust programs/rules
- Entity could be CPUC, CEC, or CAISO
- Output can be:
  - ✓ LOLP
  - ✓ EUE
  - ✓ LOLP proxy
  - ✓ EUE proxy
- Team is providing an open data, open code option as a backup.

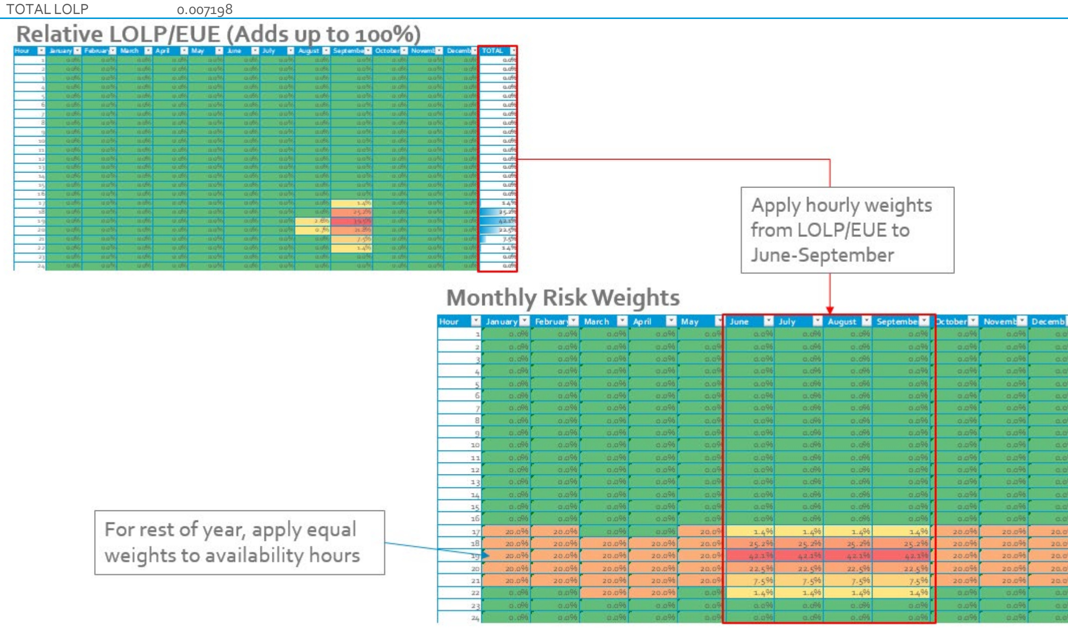
LOLP 2023 (Produced in 2021-22)

Hour	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	-	-	-	-	-	-	-	-	-	-	-	-
2	-	-	-	-	-	-	-	-	-	-	-	-
3	-	-	-	-	-	-	-	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-	-	-	-	-	-
8	-	-	-	-	-	-	-	-	-	-	-	-
9	-	-	-	-	-	-	-	-	-	-	-	-
10	-	-	-	-	-	-	-	-	-	-	-	-
11	-	-	-	-	-	-	-	-	-	-	-	-
12	-	-	-	-	-	-	-	-	-	-	-	-
13	-	-	-	-	-	-	-	-	-	-	-	-
14	-	-	-	-	-	-	-	-	-	-	-	-
15	-	-	-	-	-	-	-	-	-	-	-	-
16	-	-	-	-	-	-	-	-	-	-	-	-
17	-	-	-	-	-	-	-	-	0.000098	-	-	-
18	-	-	-	-	-	-	-	-	0.001814	-	-	-
19	-	-	-	-	-	-	-	0.000190	0.002843	-	-	-
20	-	-	-	-	-	-	-	0.000047	0.001569	-	-	-
21	-	-	-	-	-	-	-	-	0.000539	-	-	-
22	-	-	-	-	-	-	-	-	0.000098	-	-	-
23	-	-	-	-	-	-	-	-	-	-	-	-
24	-	-	-	-	-	-	-	-	-	-	-	-

TOTAL LOLP 0.007198

2 Risk weights are developed for each month that add up to 100%.

- Produce separate values for summer and winter
  - ✓ Summer value is based on LOLP
  - ✓ Winter value applies equal weights to each availability hour
- 2 values suggested because:
  - ✓ The EUE/LOLP results do not produce resource shortages each month
  - ✓ Values tend to be concentrated in 1-2 months in the outputs, which does not account for monthly variation across years



	Description	Example
3	Evaluation produces a table of resource capability by hour for the worst day in each month for all hours when the resource is available and snapback/spillover and event decay multipliers (Table 3). The table reflect the value or the first hour of dispatch.	(table below)

Hour	January	February	March	April	May	June	July	August	September	October	Novemb	Decemb
1	0.00	0.00	0.00	0.00	0.00	9.06	6.30	7.97	15.24	11.90	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00	0.00	2.67	6.59	14.40	12.21	0.00	0.00
3	0.00	0.00	0.00	0.00	0.00	0.00	0.93	5.89	11.66	8.99	0.00	0.00
4	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.45	5.44	9.81	0.00	0.00
5	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.41	4.59	7.97	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	7.79	6.94	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.31	6.52	3.53	0.00	0.00
8	0.00	0.00	0.00	0.00	0.00	0.00	1.56	1.27	7.75	2.03	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	7.71	0.12	5.89	14.63	6.42	0.00	0.00
10	0.00	0.00	0.00	0.00	0.00	31.10	10.43	18.70	30.66	15.21	1.62	0.00
11	0.00	0.00	0.00	6.56	0.00	50.04	25.66	35.94	43.31	37.46	18.17	0.00
12	0.00	0.00	2.14	15.50	6.84	60.77	47.63	53.23	56.70	56.44	38.72	0.00
13	0.00	0.00	9.09	34.84	18.63	71.59	64.05	68.25	66.71	68.84	56.75	0.00
14	0.00	0.00	15.07	47.29	30.61	80.23	75.98	79.07	76.73	76.47	63.22	0.00
15	0.00	0.00	18.78	57.80	38.04	87.80	83.39	85.95	82.84	79.38	71.28	0.00
16	0.00	0.00	21.83	63.98	44.32	90.16	86.94	90.38	87.40	81.95	68.92	0.00
17	0.00	0.00	23.23	63.85	46.44	86.60	89.22	91.81	89.57	82.31	65.14	0.00
18	0.00	0.00	16.84	52.53	36.20	74.64	80.02	84.14	79.52	74.41	47.92	0.00
19	0.00	0.00	15.13	39.60	24.94	66.29	71.19	77.54	69.96	65.96	22.69	0.00
20	0.00	0.00	9.50	18.11	11.44	44.24	53.00	59.54	49.12	38.86	5.29	0.00
21	0.00	0.00	0.44	5.29	2.30	23.76	34.99	37.54	31.93	18.11	1.29	0.00
22	0.00	0.00	0.00	4.49	0.00	24.70	41.44	40.84	39.15	19.94	8.29	0.00
23	0.00	0.00	0.00	0.12	0.00	14.79	28.87	21.89	27.01	11.90	9.36	0.00
24	0.00	0.00	0.00	0.00	0.00	5.53	14.36	15.21	22.14	9.32	5.85	0.00

4 DR provider optimizes dispatch hours to align with risk allocation (Table 4)

- Dispatch should include multipliers for event decay, pre-cooling, snap back, and spillover as relevant
- Can dispatch during availability hours or at different times depending on what is optimal for the resource

Hour	January	February	March	April	May	June	July	August	Septembe	October	Novembe	December
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
13	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
15	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
16	-0.5	-0.5	0.0	0.0	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5
17	1.0	1.0	-0.5	-0.5	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
18	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
19	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
20	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
21	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
22	-1.0	-1.0	1.0	1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0	-1.0
23	0.0	0.0	-1.0	-1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
24	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Description	Example
5 Produce the Slice of Day Load Impact Table (Table 3 x Table 4)	

Hour	January	Februar	March	April	May	June	July	August	Septembe	October	Novemb	Decemb
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	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0.00	0.00	0.00	0.00	-22.16	-45.08	-43.47	-45.19	-43.70	-40.97	-34.46	0.00
17	0.00	0.00	-11.62	-31.93	46.44	86.60	89.22	91.81	89.57	82.31	65.14	0.00
18	0.00	0.00	16.84	52.53	36.20	74.64	80.02	84.14	79.52	74.41	47.92	0.00
19	0.00	0.00	15.13	39.60	24.94	66.29	71.19	77.54	69.96	65.96	22.69	0.00
20	0.00	0.00	9.50	18.11	11.44	44.24	53.00	59.54	49.12	38.86	5.29	0.00
21	0.00	0.00	0.44	5.29	2.30	23.76	34.99	37.54	31.93	18.11	1.29	0.00
22	0.00	0.00	0.00	4.49	0.00	-24.70	-41.44	-40.84	-39.15	-19.94	-8.29	0.00
23	0.00	0.00	0.00	-0.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

APPENDIX B: SLICE OF DAY APPLIED EXAMPLE

Slice of Day and
Monthly QC example

APPENDIX C: PRODUCING A TIME TEMPERATURE MATRIX

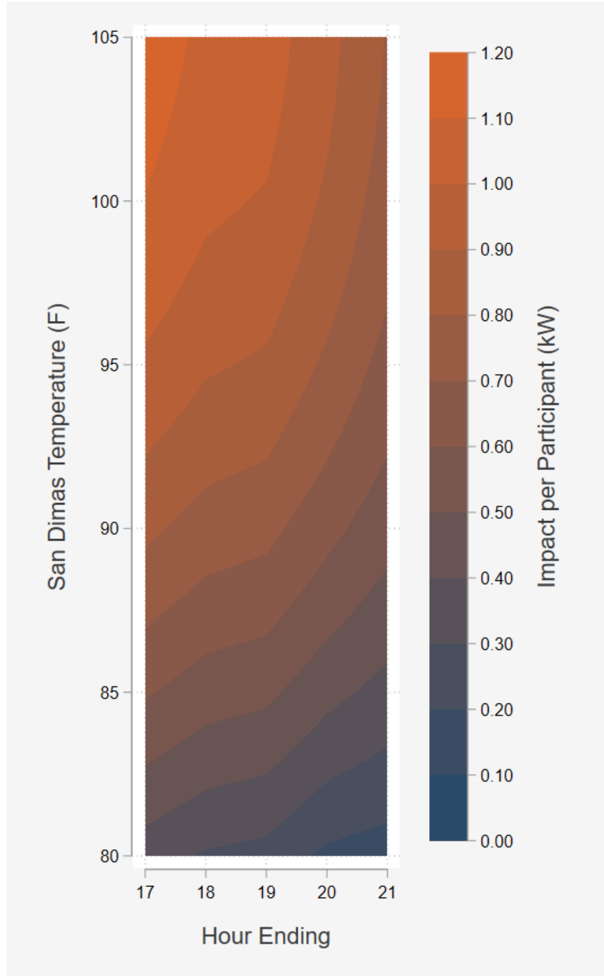
A Time-temperature quantifies the relationship between demand reductions, temperature conditions, hour of the day, event start times, and hours into an event. Importantly, a TTM is developed using the same predictive model used to produce the ex-ante planning impacts under standard conditions. Including a time-temperature matrix would better reflect the range of the resource capabilities for these different conditions that are not captured by a single planning value for each month (or a 24-hour profile for each month). A TTM has multiple uses:

- It can be used for operations and bidding.
- It can be used to compare the historical ex-ante forecasts to the bids submitted,
- It can be used to compare actual event performance to historical event forecasts, and
- It can be used to simulate the resource availability for different weather years, a common application in planning

Figure 7 shows example outputs of a simple TTM developed for SCE's Summer Discount Plan Residential (SDP-R) Program. For this program, the only independent variables used to develop the TTM were temperature (indexed to the San Dimas weather station) and hour of day. Impacts shown in the matrix are static and represent the expected participant-level impact for a territory-wide event for the given hour and temperature.

Figure 7: SDP-R Time-Temperature Matrix

Temp	Hour Ending 17	18	19	20	21
105	1.16	1.08	1.05	0.93	0.79
104	1.15	1.07	1.04	0.93	0.79
103	1.14	1.06	1.03	0.92	0.78
102	1.13	1.05	1.02	0.91	0.77
101	1.11	1.04	1.01	0.90	0.76
100	1.09	1.02	0.99	0.88	0.75
99	1.08	1.00	0.97	0.87	0.74
98	1.06	0.98	0.95	0.85	0.72
97	1.03	0.96	0.93	0.83	0.70
96	1.01	0.94	0.91	0.81	0.69
95	0.98	0.91	0.89	0.78	0.66
94	0.96	0.89	0.86	0.76	0.64
93	0.93	0.86	0.83	0.73	0.62
92	0.89	0.82	0.80	0.70	0.59
91	0.86	0.79	0.76	0.67	0.57
90	0.82	0.76	0.73	0.63	0.54
89	0.78	0.72	0.69	0.60	0.51
88	0.74	0.68	0.65	0.56	0.47
87	0.70	0.64	0.61	0.52	0.44
86	0.66	0.59	0.57	0.48	0.40
85	0.61	0.55	0.52	0.43	0.37
84	0.56	0.50	0.48	0.38	0.33
83	0.51	0.45	0.43	0.34	0.29
82	0.46	0.40	0.38	0.29	0.24
81	0.41	0.35	0.32	0.23	0.20
80	0.35	0.29	0.27	0.18	0.15



The method for calculating a time-temperature matrix is relatively straightforward. The first step for calculating a time-temperature matrix is to develop a model that predicts impacts for the average customer as a function of temperature. This will be the same model that is used to develop weather-normalized ex-ante impacts as a part of the annual reporting process for demand response. Below is a sample equation for modeling impacts as a function of temperature. This is the equation that was used to predict impacts for the TTM in Figure 7.

$$Impact_i = \beta_0 + \beta_1 * Temp + \beta_2 * Temp^2 + \beta_3 * hour * Temp + \varepsilon_i$$

Model Term	Description
$Impact_i$	Average impact in kW during interval i
β_0	The model intercept
Temp	Temperature at San Dimas Weather Station
$Temp^2$	Square of Temperature at San Dimas Weather Station
Hour * Temp	Interaction term between hour and temperature
β_1-β_3	Regression coefficients
ε_i	Error term

Once the model has been developed, the matrix is created by predicting impacts for the expected temperature range you would expect the program to operate in (in the above example the temperature ranges from 80⁰F - 105⁰F) and for the expected operating hours of the program (in the above example the operating hours range from 4-9 PM). For programs where there is event decay the matrix can also include variation in impacts based on the event hour.

Due to the varied nature of DR resources, it is important to require standard formatting so that different resources can be compared to one-another. The load impact protocols currently require standardized reporting of performance during actual events (ex-post impacts) and require the standardized reporting of hourly demand reduction capability for standardized monthly system peak days conditions (ex-ante impacts). We recommend that any additional data provided also require standardized reporting.

The actual model underlying the TTM and ex-ante impacts can vary due to the diversity of programs, but the outputs need to be standardized to include the same columns and use pre-specified weather stations by Sub-LAP. Below is the recommended data structure for the model outputs. The key outputs include the resource, the location, the event start time and duration, the hour of the event, and the average daily temperature. In this output we include the per-unit impact so that the impacts can be scaled if enrollment changes.

Table 3: Time Temperature Matrix Standard Output Format

Resource Name	Location (Sub-LAP)	Hour of Day	Event Hour	Start Time	Avg. Temperature	Event Duration	Forecasted per Unit Impact (kW)
Resource A	SCEC	20	1	7 PM	90	5	1.19
Resource A	SCEC	21	2	7 PM	90	5	1.10
Resource A	SCEC	22	3	7 PM	90	5	1.06
Resource A	SCEC	23	4	7 PM	90	5	0.96
Resource A	SCEC	24	5	7 PM	90	5	0.80
Resource A	SCEC	20	1	7 PM	89	4	1.16
Resource A	SCEC	21	2	7 PM	89	4	1.07
Resource A	SCEC	22	3	7 PM	89	4	0.99
Resource A	SCEC	23	4	7 PM	89	4	0.97
Resource A	SCEC	19	1	6 PM	89	4	1.18

Resource A	SCEC	20	2	6 PM	89	4	1.09
Resource A	SCEC	21	3	6 PM	89	4	1.00
Resource A	SCEC	22	4	6 PM	89	4	0.89
Resource A	SCEC	19	1	6 PM	88	4	1.10
Resource A	SCEC	20	2	6 PM	88	4	1.03
Resource A	SCEC	21	3	6 PM	88	4	1.00
Resource A	SCEC	22	4	6 PM	88	4	0.88
Resource A	SCEC	18	1	5 PM	88	4	1.15
Resource A	SCEC	19	2	5 PM	88	4	1.03
Resource A	SCEC	20	3	5 PM	88	4	1.02
Resource A	SCEC	21	4	5 PM	88	4	0.91

APPENDIX D: TIME TEMPERATURE MATRIX EXAMPLE

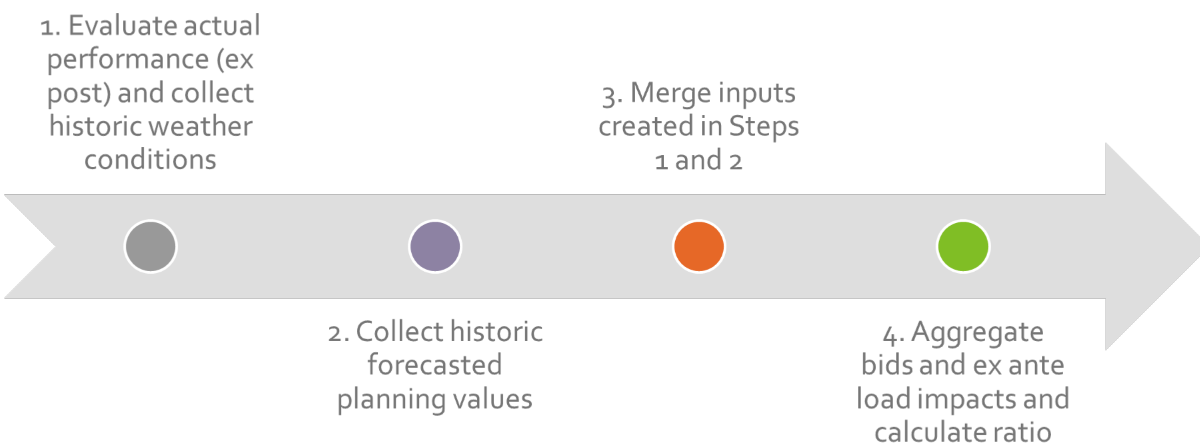
TTM Example.xlsm

APPENDIX E: PERFORMANCE ALIGNMENT METRIC

The performance alignment metric aims to determine whether actual performance during operations aligns with the forecasted capability used for planning (ex-ante impacts). The example metric is a ratio between the historic performance (ex post impacts) and the planning values developed from the historic ex ante model for the same weather and dispatch conditions. This comparison would be done for all events awarded for a given evaluation year. A ratio of 1.0 would indicate perfect alignment between performance and planning, a value greater than 1.0 would indicate that the actual performance during operations is greater than the values indicated by the planning model, and a value less than 1.0 would indicate that the actual performance for operating conditions is lower than the values indicated by the planning model.

The main concept is creating a standardized metric that is easy for all parties to understand and has a transparent calculation method. This metric can let implementers, planners, and CAISO know if there needs to be an adjustment to the planning model in the long term so that there is greater alignment between actual performance and the forecasted performance.

The figure below illustrates the key steps for developing the comparison between ex ante values and bid values. We discuss each step in greater detail in the table below.



	Step	Example
1	**Evaluate actual program performance (ex post impacts) and collect historic weather conditions.** Evaluate historic program performance for all events, as is typically done for a DR evaluation. Historic weather conditions are typically collected as a part of this process.	Below is an example of the outputs from an ex post evaluation. The results need to include the event hour, per unit impact (in kW), and weather conditions for each event.

Southern California Edison

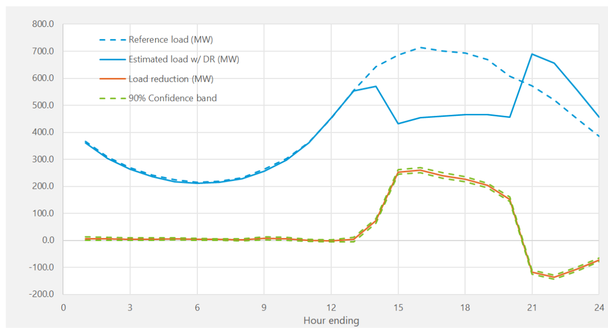
2020 Ex Post Load Impacts - SDPR

Table 1: Menu options

Program	SDP-R
Type of result	Aggregate
Category	ALL
Subcategory	All
Event date and hours	2020-08-18 (13:40-19:48 PM)

Table 2: Event day information

Event start	1:40 PM
Event end	7:48 PM
Total sites	199,557
Total devices	232,734
Total AC tonnage	846,367
Event window temperature (F)	100.6
Full event hours load reduction (MW)	236.82
Full event hours % load reduction	34.2%
All event hours load reduction (MW)	201.29
All event hours % load reduction	29.9%



Hour ending	Reference load (MW)	Estimated load w/ DR (MW)	Load reduction (MW)	% Load reduction	Avg temp (F, site weighted)	Uncertainty adjusted impact - Percentiles 5th	95th	Std. error	T-statistic
1	367.43	360.50	6.93	1.9%	82.03	0.87	12.99	3.68	1.88
2	309.42	304.00	5.42	1.8%	80.85	-0.09	10.92	3.35	1.62
3	268.75	263.89	4.85	1.8%	79.91	-0.06	9.77	2.99	1.62
4	241.34	236.43	4.90	2.0%	78.86	0.51	9.30	2.67	1.84
5	224.46	217.72	6.74	3.0%	77.74	3.08	10.41	2.23	3.03
6	214.65	210.65	4.00	1.9%	76.93	0.55	7.45	2.10	1.91
7	218.21	214.71	3.50	1.6%	76.54	0.09	6.91	2.07	1.69
8	231.63	228.79	2.83	1.2%	76.36	-1.11	6.78	2.40	1.18
9	264.30	256.54	7.77	2.9%	76.27	2.84	12.69	2.99	2.59
10	302.48	296.85	5.63	1.9%	78.31	0.23	11.02	3.28	1.72
11	362.18	361.54	0.64	0.2%	83.13	-3.27	4.55	2.38	0.27
12	451.59	452.23	-0.64	-0.1%	88.85	-4.55	3.27	2.38	-0.27
13	556.08	552.48	3.60	0.6%	93.73	-4.31	11.51	4.81	0.75
14	643.39	570.51	72.89	11.3%	97.47	64.47	81.30	5.11	14.25
15	685.85	432.43	253.43	37.0%	101.13	244.52	262.34	5.42	46.78
16	713.48	453.87	259.61	36.4%	101.79	250.37	268.85	5.62	46.21
17	700.20	459.69	240.50	34.3%	102.56	230.83	250.18	5.88	40.89
18	692.66	465.69	226.97	32.8%	100.00	217.89	236.06	5.52	41.09
19	669.46	465.87	203.59	30.4%	97.36	195.43	211.75	4.96	41.06
20	608.10	456.08	152.02	25.0%	94.97	143.84	160.19	4.97	30.59
21	572.36	689.30	-116.94	-20.4%	91.93	-124.62	-109.26	4.67	-25.04
22	519.66	655.21	-135.55	-26.1%	87.98	-143.28	-127.82	4.70	-28.83
23	451.22	557.60	-106.37	-23.6%	85.33	-113.81	-98.93	4.52	-23.53
24	385.11	457.05	-71.94	-18.7%	83.61	-78.32	-65.56	3.88	-18.55
Daily	Reference load (MWh)	Estimated load w/ DR (MWh)	Energy savings (MWh)	% Change	Avg. Daily Weighted temp (F)	Uncertainty adjusted impact - Percentiles 5th	95th	Std. error	T-statistic
Daily kWh	10654.01	9619.63	1034.38	10.8%	87.24	1001.68	1067.08	19.88	52.03

	Step	Example
2	**Collect historic forecasted planning values**	At a high level, the inputs are summarized below. The data inputs are intentionally structured so they can merge with the inputs developed in step 1.

Component	Weather Sensitive Resources	Non-Weather Sensitive Resources
Forecasted per unit load reduction capability (kW)	Table by hour of day and average daily temperature bins (TTM)	Table by hour of day and month (ex ante load impact tables)

	Step	Example
3	**Merge dataset from Steps 1 and 2.**	Below is an example for 10 hours of the merged inputs for a weather-sensitive resource.

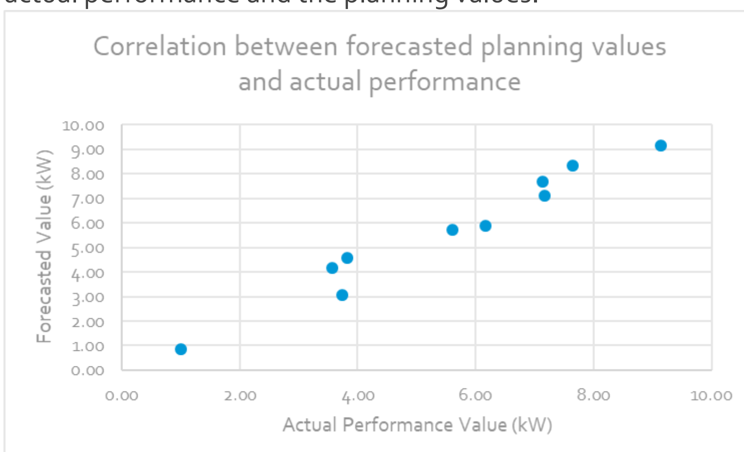
Non-weather-sensitive inputs are merged with bids for top 100 hours based on day type, month, and event hour. Weather-sensitive inputs are merged based on temperature and event hour.

Resource Name	Date	Hour	Start Time	Temp	Event Duration	Actual per Unit Performance (kW)	Forecasted Planning per Unit Value (kW)
Resource A	6/15/2021	21	6 PM	90	4	13.06	12.41
Resource A	6/16/2021	20	6 PM	81	4	88.16	74.65
Resource A	6/16/2021	21	6 PM	81	4	67.91	76.65
Resource A	6/17/2021	20	6 PM	81	4	26.19	28.83
Resource A	6/17/2021	21	6 PM	81	4	45.51	46.91
Resource A	6/17/2021	22	6 PM	81	4	37.78	37.05
Resource A	6/18/2021	20	6 PM	75	4	83.90	72.99
Resource A	6/18/2021	21	6 PM	75	4	15.65	13.78
Resource A	7/8/2021	20	6 PM	80	4	28.10	29.85
Resource A	7/8/2021	21	6 PM	80	4	91.51	93.64

5 **Aggregate bids and ex ante load impacts and calculate ratio.**

The average actual performance kW is divided by the total ex ante kW predicted to be available for all events. We use average impacts instead of aggregate MW as often the entire DR resource is not dispatched during operations. The result produces a ratio that assesses how well the actual performance aligned with the values produced by the ex-ante model.

Below is a comparison of the sample actual performance and forecasted values across all event hours for the DR season. As expected, the two are highly correlated, which indicates good alignment between the actual performance and the planning values.



Below is an example of the summed inputs and ratio calculation. A value of 1.0 indicates perfect alignment. A value greater than 1.0 indicates that the actual performance is greater than the values in the

$$Metric = \frac{\sum Actual\ Performance\ kW}{\sum Forecasted\ Ex\ Ante\ kW}$$	planning model. A value less than 1.0 indicates that the actual performance is lower than the values in the planning model.

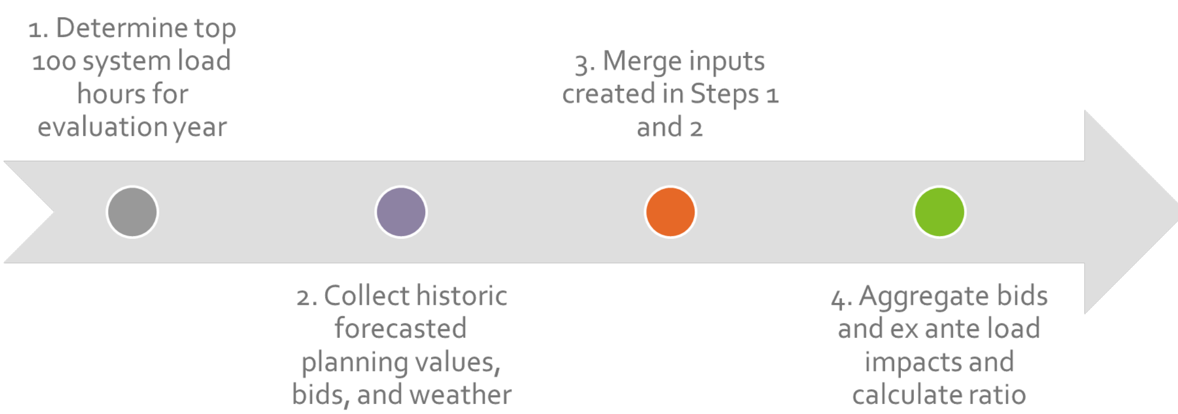
	Average Actual Performance (kW)	Average Forecasted Planning TTM Value (kW)
TOTAL	4.68	4.49
RATIO	1.04	

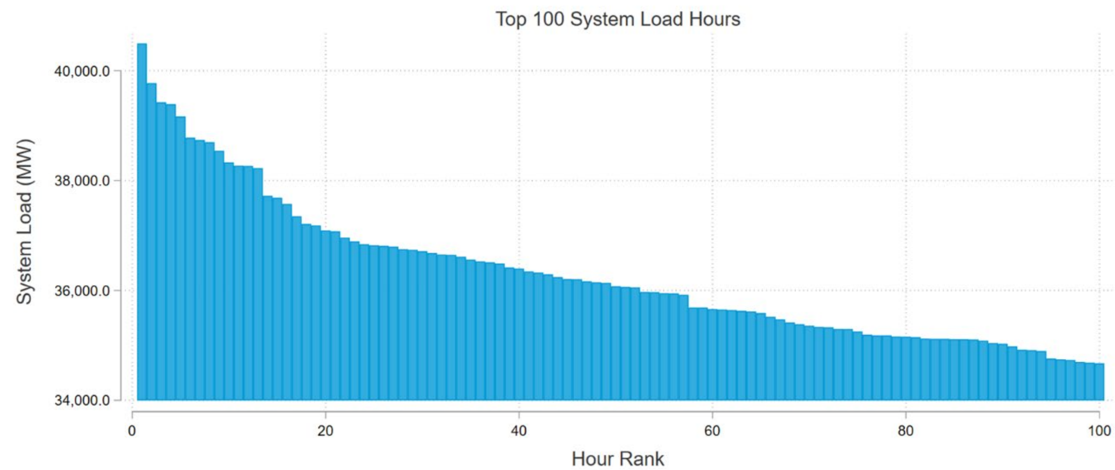
APPENDIX F: BID ALIGNMENT METRIC

The bid alignment metric aims to determine whether historic bids align with the forecasted capability used for planning (ex-ante impacts). The example metric is a ratio between the historic bidding values and the capability forecasted by the historic ex-ante model. We recommend narrowing the comparison to the top 100 net load hours for each year for simplicity and because these hours are when DR resources are most needed. A ratio of 1.0 indicates full alignment between operations and planning, a value greater than 1.0 means that the bid values were greater than the capability forecasted by the ex-ante model, and a value less than 1.0 would indicate that the bid values are lower than the values indicated by the planning model.

The goal of the bid alignment metric is to use a standardized metric that is easy for all parties to understand and has a transparent calculation method. This metric can let implementers, planners, and CAISO know if there needs to be an adjustment to the planning model or the bidding process to improve alignment.

The figure below illustrates the key steps for developing the comparison between ex-ante values and bid values. The table below details the steps to produce the bid alignment metric.



	Step	Example
1	**Determine the Top 100 hours for the evaluation year**. Use the CAISO net system load to determine the top 100 hours for the year.	Below is a graph of the top 100 system load hours for 2021. 
2	**Collect historic forecasted planning values, bids, enrollments, and weather**	At a high level, the inputs are summarized below. The data inputs are intentionally structured so they can merge with the inputs developed in step 1.
3	**Merge dataset from Steps 1 and 2.**	Below is an example of 10 hours of the merged inputs for a weather-sensitive resource.

Component	Weather Sensitive Resources	Non-Weather Sensitive Resources
Historic forecasted planning values	Table by hour-of-day, and average daily temperature bins, event start, and hours into the event (TTM)	Table by hour-of-day and month (ex-ante load impact tables)
Historic bids for evaluation	Table that includes date, hour, number of customers, and bid MW for CAISO. The bids should be aggregated across sub-LAPS since the goal is to assess if the capacity is being bid into the market.	
Historic Weather Conditions	Table that includes the date, hour, and temperature.	

Non-weather-sensitive inputs are merged with bids for the top 100 hours based on day type, month, and event hour. Weather-sensitive inputs are merged based on temperature and event hour.

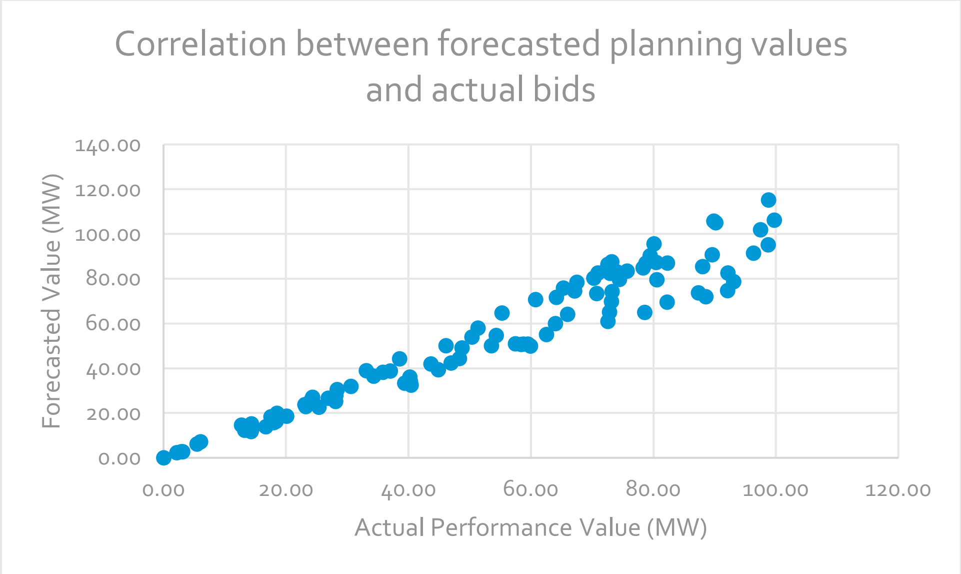
Resource Name	Date	Hour	Start Time	Temp	Event Duration	Bid Value (MW)	Forecasted Planning Value MW (Ex-Ante /TTM)
Resource A	6/15/2021	21	6 PM	90	4	15.03	12.63
Resource A	6/16/2021	20	6 PM	81	4	78.76	88.96
Resource A	6/16/2021	21	6 PM	81	4	26.96	26.56
Resource A	6/17/2021	20	6 PM	81	4	19.74	19.80
Resource A	6/17/2021	21	6 PM	81	4	39.89	42.84
Resource A	6/17/2021	22	6 PM	81	4	7.84	8.52
Resource A	6/18/2021	20	6 PM	75	4	99.25	113.17
Resource A	6/18/2021	21	6 PM	75	4	87.18	79.30
Resource A	7/8/2021	20	6 PM	80	4	54.47	62.25
Resource A	7/8/2021	21	6 PM	80	4	36.61	37.77

5 **Aggregate bids and ex-ante load impacts and calculate the ratio.**

The total bid MW available for the top 100 hours is divided by the total ex-ante MW predicted to be available for the top 100 hours. The result produces a ratio that assesses how well the bid values aligned with the values produced by the ex-ante model.

$$Metric = \frac{\sum Historic\ Bid\ MW}{\sum Forecasted\ Planning\ MW}$$

Below is a comparison of the sample bid values and forecasted values across the top 100 hours. The two are highly correlated in the example, indicating good alignment between the actual bids and the forecasted planning values.



Below is an example of the summed inputs and ratio calculation. A value of 1.0 indicates perfect alignment. A value greater than 1.0 indicates that the bid values were greater than the forecasted planning values. A value less than 1.0 indicates that the bid values were lower than the forecasted planning values.

	Bid Value (MW)	Ex Ante TTM Value (MW)
TOTAL	4,882	4,629
RATIO	1.05	