

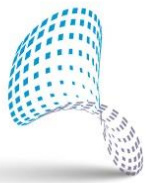
**DOCKETED**

<b>Docket Number:</b>	21-DR-01
<b>Project Title:</b>	Supply Side Demand Response
<b>TN #:</b>	248194
<b>Document Title:</b>	Josh Bode Comments - DSA Comments on CEC Draft Report Qualifying Capacity of Supply-Side DR
<b>Description:</b>	N/A
<b>Filer:</b>	System
<b>Organization:</b>	Josh Bode
<b>Submitter Role:</b>	Public
<b>Submission Date:</b>	12/20/2022 6:10:05 PM
<b>Docketed Date:</b>	12/21/2022

*Comment Received From: Josh Bode*  
*Submitted On: 12/20/2022*  
*Docket Number: 21-DR-01*

**DSA Comments on CEC Draft Report Qualifying Capacity of  
Supply-Side DR**

*Additional submitted attachment is included below.*



# Demand Side Analytics

DATA DRIVEN RESEARCH AND INSIGHTS

## CEC Qualifying Capacity of Supply-Side Demand Response Working Group Final Report Comments



Prepared By: Demand Side Analytics

December 20, 2022

## TABLE OF CONTENTS

1	Introduction .....	2
2	Observations on CEC Report and Working Group Process .....	3
3	DSA Positions on CEC Recommendations .....	7
4	DSA Concerns with CEC Recommended Approach.....	11
5	Request to Revise DSA Proposal Summary.....	15
5.1	CURRENT SUMMARY OF DSA PROPOSAL .....	15
5.2	DSA PREFERRED PROPOSAL SUMMARY .....	16

# 1 INTRODUCTION

Decision 22-06-050 OP11 states “*The California Energy Commission (CEC) Working Group is requested to continue to develop long-term recommendations for a new demand response (DR) qualifying capacity (QC) methodology for the 2025 Resource Adequacy (RA) year, consistent with the Reform Track framework adopted in this decision*”. The California Energy Commission (CEC) developed a report detailing the CEC’s final findings and recommendations to the CPUC based on proposals submitted to the CEC’s working group on supply-side demand response. This document provides Demand Side Analytics’ (DSA) comments on the working group process and our position on the recommendations provided by the CEC.

For clarity, we break our comments into the following sections:

- Overall observations of the CEC report and working group process
- DSA positions on CEC recommendations
- DSA concerns with the CEC report and proposal
- Request for revisions to the characterization of the DSA proposal by the CEC

## 2 OBSERVATIONS ON CEC REPORT AND WORKING GROUP PROCESS

DSA has several concerns about the working group process and the recommendations provided by the CEC. We do not feel that the recommendations provided by the CEC accurately reflect the proposals and discussions that took place during the working group meetings.

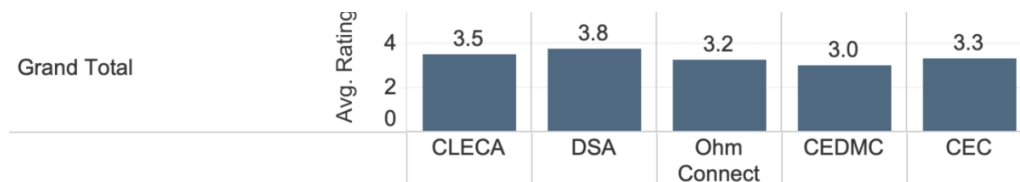
- **The CEC was not an unbiased working group moderator.** They were an active stakeholder, not a moderator. Specifically, a single CEC commissioner advisor submitted his own proposal (which is extremely unusual), moderated the discussion, and made the final decisions. The role of the CEC was not about gathering information from various parties. A working group with various stakeholders should be facilitated by an unbiased party without a stake in the outcome, and that is not what took place here. As soon as the CEC introduced its own proposal, it could not be a neutral fact-finding moderator. Because the CEC advisor was the moderator, a proponent, and the decision maker, stakeholders were not fully at ease discussing misgivings and limitations of the CEC proposal.
- **The CEC report imposed a litmus test that was not communicated or transparent and was not included in the working group principles or the CPUC guidance.** Specifically, the CEC rejected from consideration any proposal that did not discuss penalties and relied on historical performance as the basis for determining the qualifying capacity of supply-side demand response. The litmus test was not communicated to the working group; it was not in the principles that the parties had agreed to for scoring proposals, and; it was not in the guidance provided by the CPUC. As a result, the CEC only seriously considered two proposals that collectively scored the lowest on the working group principles as ranked by stakeholders. Those proposals were the CEC commissioner advisor's own proposal and CEDMEC's proposal. Notably, the DSA proposal ranked highest and was specifically formulated so it could be converted from one that relied on historical performance to a proposal that relied on DR provider upfront capability forecasts and backend rigorous analysis and penalties. However, without clear communication of the litmus test to the working group, such adjustments could not be implemented. Had the CEC communicated the automatic rejection of proposals without penalties or that relied on historical performance, the working group stakeholders would have submitted different proposals
- **The CEC report ignored how proposals aligned with the working group principles as scored by the workshop stakeholders.** The feedback provided during the workshop by the stakeholders did not favor the CEC proposal. All the stakeholders were invited to rank how each of the proposals submitted aligned with the agreed-on principles by the stakeholders. The majority of stakeholder entities replied to the survey. The eleven (11) respondents are shown in the table below (Table 1) and include a wide range of perspectives and interests. The stakeholders, as a group, definitively scored the two proposals that built on existing load impact protocols the

highest based on agreed upon principles. When the stakeholder survey results were revealed, the moderator – who submitted their own proposal and scored the proposals – repeatedly insinuated that stakeholders must have misunderstood the principles and attempted to re-score the results, which ranked the CEC and CEDMEC proposals the lowest. This is remarkably unusual. In their report, the CEC summarily dismisses the stakeholder feedback and preferences. The stakeholders were not confused by the principles, they simply interpreted the principles and applied them to the proposals. What the CEC did not mention in their report is that there were numerous meetings debating the principles and their wording. Everyone scoring the proposals had the context, knowledge, and expertise to identify how well the proposals aligned with the principles. In their report, the CEC does not discuss the alignment of their recommendation with the working group principles. It insinuates that it incorporates other proposals, but at its core, it is the CEC commissioner advisor’s proposal with minimal substantive modifications. It does not reflect a consensus or the positions from the working group.

Table 1: Overview of stakeholders who provided workshop feedback

Entity Description	Survey Participants
Investor-owned utilities	<ul style="list-style-type: none"> <li>PG&amp;E</li> <li>SCE</li> <li>SDG&amp;E</li> </ul>
Regulators	<ul style="list-style-type: none"> <li>California Energy Commission</li> </ul>
Market operators	<ul style="list-style-type: none"> <li>CAISO</li> </ul>
Third-Party DR vendors and implementers	<ul style="list-style-type: none"> <li>SunRun</li> <li>LEAP</li> <li>Olivine</li> </ul>
Stakeholder organizations	<ul style="list-style-type: none"> <li>California Efficiency + Demand Management Council</li> <li>California Large Energy Consumers Association (CLECA)</li> </ul>
Evaluators	<ul style="list-style-type: none"> <li>Demand Side Analytics (DSA)</li> </ul>

Figure 2: Overall, Alignment of Proposals with Working Group Principles as Scored by Stakeholders



- The core of the CEC recommendations were not tested in practice and were not fully vetted by the members of the working group.** The best practice is to test methods before adopting them and then phasing them in. Most members of the working group noted they found the CEC proposal confusing and unclear and had to re-read it multiple times. The actual CEC proposal is unclear because modifications are made in the report itself. There is no one location that

contains the end-to-end proposal and the terminology in the report is inconsistent with the terminology in the CECs proposal. The code used to develop the CEC methodology was not shared with other working group members and has not been replicated or implemented by any other working group members. This goes against one of the key workshop principles of developing a transparent and understandable methodology. As far as we know, only one person, the CEC advisor, tested the methodology using data from a program that is not representative of most DR programs. The CEC's proposed methodology is a radical departure from tested approaches and throws out more than 10-years of past DR performance and evaluations. The proposed approach should be tested on real-world examples and on a variety of DR program types to determine whether any specific program types or entities are unfairly penalized by the new methodology. This would allow the stakeholders and regulatory bodies to be better informed when deciding whether to adopt the CEC's proposed methodology.

- **The DR QC methodology should only be adopted once it has been tested. If adopted, the approach should be phased in over multiple years with a process for improvements and corrections.** The CEC proposal is simply not ready for implementation. Based on DSAs' discussions, evaluators, investor-owned utilities, large customer associations, and CAISO do not believe it has been adequately tested and do not believe it is ready for implementation under the timeline proposed. The CEC proposes some adjustments to their initial approach in their report, but it will not be possible to resolve these technical concerns without a clear process that can be implemented by other parties for testing purposes. The CEC has not shared their code, testing data, or provided enough detail for other parties to implement their approach. We recommend that the CEC share their process, data, and testing code so that the process is more clear and can be tested by other members of the working group and applied to a wider range of programs.
- **The CEC recommended approach has multiple technical flaws and cannot be adopted in its current form.** These issues are discussed in more detail in Section 4. At broad level, upfront predictions with penalties can work, but they require upfront checks, rigorous measurement, and enforceable penalties. The CEC proposal provides unbounded flexibility upfront, relies on baseline settlement heuristics that are known to be imprecise and flawed, and does not propose a viable enforcement mechanism. It also adds needless complexity when it is not needed.
- **The CEC report recommends removing the CPUC's resource adequacy oversight role and reassigning it to CAISO.** This is a radical departure from the CPUC's role in the DR QC process. Specifically, it recommends that CAISO be responsible for determining the DR QC and enforcing penalties. Very importantly, it is not clear whether CAISO is allowed to enforce penalties or whether they are interested in taking on an enforcement role. CAISO oversees the energy markets and settlement. It currently does not have oversight over capacity and capacity funds are in bi-lateral contracts between load-serving entities and DR providers. Thus, it is unclear how the CAISO could enforce penalties when the capacity money is not in the CAISO market.



Based on the inherent flaws in the working group process and the CEC proposal, we will recommend to the CPUC that it reject the CEC's report and recommendations. If the CPUC wants to shift the DR paradigm from one based on historical performance to one based upfront forecasts with back end rigor and penalties, we recommend that the CPUC indicate so directly, and allow for better, more feasible proposals to be submitted. Those proposals should transparently undergo proper testing before they are adopted and phased in.

### 3 DSA POSITIONS ON CEC RECOMMENDATIONS

The table below summarizes our stances on each of the recommendations made in the CEC proposal.

Table 3: Summary of DSA's Positions on CEC Recommendations

#	Recommendation Component	DSA's Position
1	Adopt an incentive-based approach (move away from LIP ex-ante forecast)	If adopted, we recommend that stakeholders be allowed to submit proposals that conform with this principle since it was not a part of the working group's principles and was not in the CPUC guidance.
2	Adopt the capacity shortfall penalty incentive mechanism with forced outage adder	If adopted, we recommend that stakeholders be allowed to submit proposals that conform with this principle since it was not a part of the working group's principles and was not in the CPUC guidance.
3a	Adopt the ex-ante capability profile	Agree that upfront forecasts of capability are needed but recommend making them a time-temperature matrix. A time temperature matrix is a forecast of resource capability as a function of temperature, event start, event duration, and hours into the event. Unlike the ex-ante capability proposed, a time-temperature matrix allows for direct comparisons of forecasted capability and actual event performance. The ex-ante capability proposal is inadequate for weather sensitivity resources where the number of hours into the event can be a bigger predictive factor than temperature.
3b	Adopt the ex post regression approach proposed by CEC staff	Disagree. First, the name "ex-post regression" is a misnomer and confusing. What the CEC proposes is a prescriptive regression on ex post results to determine the DR QC delivered. The analysis proposed by the CEC is needlessly prescriptive, complicated, and unneeded. It is simply false precision. If parties are required to submit an upfront forecast of reduction capability by temperature, hour, and other factors, one should be able to directly compare the forecast to the demand reductions delivered under the same set of conditions. DSA proposed a metric that does this, the performance alignment metric (PAM) and a bid alignment metric (BAM), which are simpler and easier to calculate. If the CPUC adopts an upfront forecast with backend evaluation rigor and penalties, we recommend that Delivered Qualifying Capacity simply be the DR capability forecast at the worst-day of month temperatures times the performance alignment metric times the bid alignment metric. As noted in the DSA proposal, additional testing is needed before adoption.
4	Require resources to show takeback (no distinction	Agree. DR capacity should include all spillover effects, whether positive or negative.

	between event hours and spillover)	
5a	Require DR providers to submit capability profiles	Agree that upfront forecasts of capability are needed but recommend making them a time-temperature matrix. A time temperature matrix is a forecast of resource capability as a function of temperature, event start, event duration, and hours into the event. Unlike the ex-ante capability proposed, a time-temperature matrix allows for direct comparisons of forecasted capability and actual event performance. The ex-ante capability proposal provided is inadequate for weather sensitivity resources because the number of hours into the event is often a bigger predictive factor than temperature itself.
5b	Require DR providers to submit a “slice-of-day” table to summarize QC values	Agree
6	Eliminate unnecessary reporting requirements for QC determination (Reuse of evaluation plans and eliminate public review process)	Agree that it should be simplified. However, we would recommend still requiring drop-down menu based annual evaluation plans and removing the public review.
7	Plan to produce final QC numbers by June 1, preceding the RA compliance year	Agree
8	Adopt streamlined QC approval criteria.	Agree
9a	Implement the proposed penalty mechanism and exempt DR from the RAAIM	Agree. If a penalty mechanism is adopted, DR should be exempt from RAAIM.
9b	Should CAISO be responsible for implementing the penalty mechanism?	No. The CPUC should be responsible for the penalty mechanism. CAISO has limited leverage over penalties since it only settles the energy component, which is minimal. The capacity funding, which makes up the bulk of the money, is in bi-lateral contracts between the load serving entities and DR providers.
9c	Should CAISO be responsible for determining QC?	No. The CPUC should be responsible for determining QC. Otherwise, it is relegating authority over resource adequacy and capacity to CAISO.
10	Consider phase-in of incentive-based approach over time (1 year)	Agree, but the adopted DR QC method should be tested in a transparent manner before it is phased in. The current proposal has not undergone adequate testing and was not transparent. We also recommend a 3-year phase-in with 1 <sup>st</sup> year free, 2 <sup>nd</sup> year at 50% penalties, and 3 <sup>rd</sup> year at 75% penalties.

11a	Are settlement baselines accurate enough to use as the default for assessing DR QC performance?	No. The evidence shows that settlement baselines are imprecise heuristics for load impacts. The 2017 CAISO baseline working group report concluded that even small control groups are more than twice as precise as the best day-matching or weather-matching baselines. CAISO more recent study on “Demand Response Advanced Measurement Methodology” concluded what DR evaluators have known since early 2010, namely that comparison groups are more precise and accurate than day-matching settlement baselines. The proposal to use the CAISO settlement baselines moves us backwards, not forwards, and makes imprecise measurement the default. We are not in disagreement with the need to further standardize options or the need to narrow the gap between settlement baselines and evaluations. However, the right direction is to move CAISO baselines to be closer to evaluation methods rather than make CAISO day-matching baselines the default, and to work to allow different types of comparison group data to be made publicly available.
11b	Should settlement baselines be used for assessing DR QC performance?	No. Settlement baselines are simplistic heuristics and highly imprecise. In addition, they are often incorrectly applied. The CAISO rules require an aggregate-first baseline, meaning that the resource loads are first aggregated first, and the baseline days are identified, the baseline is constructed, and same day adjustments are applied. Many parties implement individual first baselines, meaning they pick different baseline days for each site, calculate a baseline for each site, and apply the adjustment at the individual site level, and then aggregate the results. This latter approach does not follow the CAISO baseline rules and leads to upward bias.
11c	Should there be alternatives? (If so, when are alternatives allowed)	Yes. A difference-in-difference calculation with a comparison group should be the default. The current CAISO day matching baselines should be an option of last resort. Difference-in-differences with comparison groups are the current default for the evaluation of most mass market programs, but third party aggregators sometimes deviate from the norm. In general, you can standardize the measurement to four or five options that cover nearly all types of DR and rank them in order of priority.
12a	Adopt bid normalization for load impact in assessing DR QC performance (BNLI)?	No, it’s only necessary if the unneeded DR performance regression is adopted. If the CEC approach is adopted, we propose to modify the equation so it’s symmetric.
12b	Adopt CEC metric for bid normalization (BNLI)?	No, it’s only necessary if the unneeded DR performance regression is adopted. As noted earlier, the analysis proposed by the CEC is needlessly prescriptive, complicated, and unneeded. If parties are required to submit an upfront forecast of reduction capability by temperature, hour, and other factors, one should be able to directly compare the forecast to the demand reductions delivered under the same

		set of conditions. DSA proposed a metric that does this, the performance alignment metric (PAM) and a bid alignment metric (BAM), which are simpler and easier to calculate.
13	Reduce the threshold required for midyear QC update.	Agree. DR providers should be able to provide updates more frequently.
14	Eliminate the components of the PRM adder associated with operating reserves and load forecast	No position. We note that folks fundamentally misunderstand the PRM and how it is calculated. It's inaccurate to parse it into components and the component are incorrectly described as error. In practice, the PRM is based on simulating a loss of load probability model and assessing the percentage buffer from 1-in-2 demand needed to meet a 1-in-10 loss of load expectation. The loss of load expectation tends to occur when conditions are load hotter than 1-in-2 (weather and load variability) and when resources shortages are bigger than typical (forced outage variability, import variability, and hydro/solar/wind variability).
15	Convert the forced outage adder to a multiplier applied in the effective capacity formula	No position.
16	Maintain the distribution loss factor adder in QC values	Agree. Since DR is at the meter it avoid the need to transmit power over distribution lines and thus avoided line losses.
17	Update transmission loss factors and include the adder as a credit	Agree. Since DR is at the meter it avoid the need to transmit power over transmission lines and thus avoided line losses.

## 4 DSA CONCERNS WITH CEC RECOMMENDED APPROACH

The CEC's recommendations and approach have multiple inherent technical flaws or other implementation issues. We describe each of the proposal concepts in greater detail in the table below, along with our concerns around the concept.

Table 4: Summary of Proposal Issues

CEC Approach Concept	High Level Description of Concept	Issue
Customer Baselines	The CEC recommends requiring settlement baselines to calculate ex post impacts unless a DRP can prove that there is a more accurate method for calculating impacts.	Multiple studies have found customer baselines to be less accurate compared to other evaluation methodologies. This is not a preferred method for many programs and should not be required. It is an option now for ex-post evaluation and should continue to be an option. CAISO conducted a study with Recurve in 2021 that found that control groups are better than baselines when estimating demand response. In 2017, the CAISO Baseline Working Group Accuracy study found that control groups are twice as precise as the best baseline methods, even with smaller sample sizes. Changes to settlement going forward should be focused on improving the ability of all parties to be able to access to control group or comparison data rather than requiring all parties to use a method that has been found to be inferior to control groups in multiple studies. In energy efficiency, the use of granular aggregated profiles (time series) for non-participants is becoming the norm and is as accurate and often better than individually matched controls.
Bid-Normalized Load Impacts (BNLI)	Under a full dispatch, the BNLI is equal to the delivered load impacts. Under a partial dispatch, the bid amount is adjusted by the ratio of delivered load impacts to the bid amount, but this ratio is always capped at 1. The only time BNLI can exceed the bid is	This metric is only needed if performing a DR performance regression, which we feel is unnecessary. Additionally, the Commission Report recommends the use of an alternative BNLI metric that was not included in the CEC's original proposal and was shown to the working group prior to the CEC making its recommendations.

	when load impacts exceed the bid, regardless of the dispatch amount.	
Ex-post Regression (or availability regressions)	Develop a linear regression model of adjusted load impacts on temperature over each month (or other grouping) by hour of day.	<p>First, the name “ex-post regression” is a misnomer and confusing. It is also unclear what the CEC proposes. The initial proposal included an “availability regression” which seemed to be used to determine the capability profile. The report refers to an ex-post regression, which appears to be a prescriptive regression on ex-post results to determine the DR QC delivered.</p> <p>In general, the CEC proposal is difficult to understand and appears to be flawed. The initial version seems to weight bids the same as actual performance during events, which seems incorrect on first principles. In the proposal, it is unclear whether a resource that CAISO only dispatches from 6-8 PM can claim availability from 4-9 PM. Also, in the proposal, it appears a DR provider can game the system by never dispatching for a continuous 4, 5, or 6 hour event. Thus, a provider can dispatch for a single hour many times and claim the higher reduction that occur in the first hour, but are not sustainable over a longer event duration. Further testing of this methodology is required to make it clear whether the methodology can work.</p> <p>We also believe the analysis proposed by the CEC is needlessly prescriptive, complicated, and unneeded. If parties are required to submit an upfront forecast of reduction capability by temperature, hour, and other factors, one should be able to directly compare the forecast to the demand reductions dispatched and delivered under the same set of conditions without modeling.</p>
DR Capability Profiles	Develop a continuous relationship between a resource’s load impacts and temperature conditions for each hour based on the availability regression described above.	The capability profiles can be difficult to understand and implement in practice. A time-temperature matrix would achieve the same goal while also being simpler, easier to understand, and better reflecting the capability of weather-sensitive resources.

Flexibility on Capability Profiles	Rather than the current LIP submission, DR providers can submit a “capability profile” which shows the minimum load impacts a DR provider expects of a resource under varying temperature conditions.	The requirements of the capability profiles are vague. We concerned that this could allow DR providers to overestimate their capability for an entire season prior to any evaluation at the end of the summer.
Granularity of DR performance assessment	The report does not make a recommendation on this critical issue.	DR delivers capacity, it does not produce energy but reduces or shifts it. As such, performance should be assessed at the same level that capacity is delivered, at a portfolio level. In most cases, the focus is on system capacity, but DR can also be used to meet local capacity area (LCA) requirement. The only reason DR resources are split into distinct granular resource IDs is in order to fit into the energy market. However, the more DR is disaggregated, the noisier measurements becomes. The more you aggregate, the more precise the measurement becomes. Noisy, granular measurement increases the risk of penalties due to imprecise measurement rather than true underperformance. It would be incorrect to pretend DR is an energy product and assess its performance at highly disaggregated level.
The DR evaluation are expensive	The CEC report repeatedly asserts that DR evaluation are expensive without evidence	<p>The CEC report simply asserts the third-party claims that difficult, expensive, and opaque without any evidence. To put this in perspective, Ohm comment claimed 185.5 MW of DR resources for August 1-in-2 peaks in 2023, in its latest filing. At \$100/kW-year that is the equivalent of over \$18.5 Million. The cost of an evaluation is \$50k-\$100k per program (&lt;1% of revenue). Leap Frog Power claimed 179 MW of ex-ante capability. Most third parties are claiming well over 50 MWs of demand response capability, which means their evaluation costs are at most 1-2% of the capacity revenue.</p> <p>Most of the complaints in the working were less about the actual evaluations themselves but about the regulatory oversight, the inability to claim MWs they had not yet built, and about the process.</p>



		Nowhere did the CEC demonstrate that its approach is less expensive than the current approach. The approach simply has less oversight and allows third party aggregators to claim large amounts of DR they have not yet built while relying on untested penalty enforcement mechanisms
The CEC does not require any proof that DR providers can deliver the DR capacity claimed. It relies entirely on a penalty mechanism.	<b>Other ISO's require demonstration that a DR provider can deliver the claimed DR capabilities</b>	In other jurisdictions, ISO's do not simply accept claimed DR capability, but have a qualification process where the DR providers have to show ex-ante how they plan to deliver those reductions claimed and that they have built the resources. The CEC proposal has unbounded flexibility upfront without adequate rigor or oversight.

## 5 REQUEST TO REVISE DSA PROPOSAL SUMMARY

DSA does not feel that our proposal is accurately summarized in the CEC report. In this section we include the current summary of the DSA proposal in the report and a revised summary, which we feel more accurately characterizes our proposal.

### 5.1 CURRENT SUMMARY OF DSA PROPOSAL

The Demand Side Analytics (DSA) proposal is fundamentally an application of the LIPs to the new slice-of-day framework. The first element of the proposal states: “The [LIPs] should be retained but modified to address the 24-hour slice-of-day framework.” Such modifications include updating planning temperatures to the “worst day” as defined in the RA program, allowing DR resources the flexibility to provide capacity value based on need (in contrast to the static availability assessment hours), and accounting for spillover in nonevent hours (including negative load impacts or takeback).

Much of the DSA proposal focuses on standardization of reporting requirements and outputs. For all resources, a 24-slice-of-day table would show hourly load impacts for the worst day in each month. Each load impact estimate in the table would be the hourly capacity value eligible for RA in that hour and month. For weather-sensitive resources, DSA proposes production of a time-temperature matrix of load impacts as an upon-request output. The time-temperature matrix could also disaggregate load impacts by event start time or hours into event, if needed.

The proposal includes supplemental components apparently for informational purposes. DSA proposes that a central planning authority produce a “reliability risk heatmap” for each compliance year that will help DR providers align resources and programs with system need but does not directly affect either ex ante or ex post capacity valuation.

The proposal also includes two ex post performance metrics:

- The **bid alignment metric** measures the extent to which resources bid as expected based on the associated 24-slice-of-day table or time-temperature matrix or both.
- The **performance alignment metric** measures the extent to which resources perform as expected when dispatched.

Like the risk heatmap, these metrics do not appear to directly impact ex ante or ex post capacity valuation. DSA writes of both metrics: “we recognize that stakeholders may want additional discussion and the opportunity to test it in practice before it is adopted,” suggesting it may be integrated into the QC methodology, but a description of how it would do so is not provided.

DSA also included suggestions for aligning evaluation of load impacts in the planning space with evaluation used for settlement. These elements of the proposal are included in the section “Alignment of Operational and Planning Spaces” of this report.

For more information, see the DSA proposal posted to CEC Docket 21-DR-01.

## 5.2 DSA PREFERRED PROPOSAL SUMMARY

The Demand Side Analytics (DSA) proposal is fundamentally an application of the LIPs to the new slice-of-day framework. The first element of the proposal states: “The [LIPs] should be retained but modified to address the 24-hour slice-of-day framework.” Such modifications include updating planning temperatures to the “worst day” as defined in the RA program, allowing DR resources the flexibility to provide capacity value based on need (in contrast to the static availability assessment hours), and accounting for spillover in nonevent hours (including negative load impacts or takeback). **DSA also notes that evaluation of ex-post load impacts (rather than settlement heuristics) should be used as the basis for assessing performance since they are more accurate and have a long history using a standard output template. CAISO should allow evaluation results to be used for settlement as long the evaluation is produced in advance, and the results are produced within the settlement period.**

Much of the DSA proposal focuses on standardization of reporting requirements and outputs. For all resources, a 24-slice-of-day table would show hourly load impacts for the worst day in each month. Each load impact estimate in the table would be the hourly capacity value eligible for RA in that hour and month. For weather-sensitive resources, DSA proposes production of a time-temperature matrix of load impacts **as an upon-request output. 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 can be used to compare ex-ante predictions to actual reduction delivered during events under the same set of conditions.**

The proposal includes supplemental components **that can be used to better align operations and planning values/ apparently for informational purposes.** DSA proposes that a central planning authority produce a “reliability risk heatmap” for each compliance year that will help DR providers align resources and programs with system need but does not directly affect either ex-ante or ex-post capacity valuation.

The proposal also includes two ~~ex-post~~ performance metrics:

- The **bid alignment metric** measures the extent to which resources bid **into the market align with the forecasted resource capability,** as documented in the ~~expected based on the associated 24-slice-of-day table or time-temperature matrix or both.~~
- The **performance alignment metric** measures the extent to which resources perform as expected when dispatched. **It compares the predicted resource capability, as documented in a time temperature matrix, to the actual reductions delivered under the same set of conditions.**

Like the risk heatmap, these metrics aim **to better align operations and planning values** ~~do not appear to directly impact ex-ante or ex-post capacity valuation.~~ DSA writes of both metrics: “we recognize that stakeholders may want additional discussion and the opportunity to test it in practice before it is adopted.” ~~suggesting it may be integrated into the QC methodology, but a description of how it would~~

~~do so is not provided.~~ DSA did not wish to recommend the application of metrics before they had been fully tested and vetted.

The DSA proposal was intentionally designed so it could be converted into an up-front forecast (using the time-temperature matrix), with backend evaluation rigor, and penalties, if needed. The metrics proposed were developed to enable a calculation of delivered DR capability by multiplying the DR capability under planning conditions by the performance alignment metric and the bid alignment metric. It was designed with the working group principles and CPUC guidance which, at the time, did not include requirements for ex-ante forecasts with backend penalties. Overall, the DSA proposal was scored the highest by the stakeholder survey in its alignment with the agreed upon principles.

DSA also included [additional](#) suggestions for aligning evaluation of load impacts in the planning space with evaluation used for settlement. These elements of the proposal are included in the section “Alignment of Operational and Planning Spaces” of this report.

For more information, see the DSA proposal posted to CEC Docket 21-DR-01.