Western Energy Imbalance Market Resource Sufficiency Evaluation Metrics Report covering July 2023

August 31, 2023

Prepared by: Department of Market Monitoring

California Independent System Operator

1 Report overview

As part of the Western Energy Imbalance Market (WEIM) resource sufficiency evaluation enhancements stakeholder initiative, DMM is providing additional information and analysis about resource sufficiency evaluation performance, accuracy, and impacts in regular monthly reports.[1] This report provides metrics and analysis covering July 2023 and is organized as follows:

- Section 2 provides an overview of the flexible ramp sufficiency and bid-range capacity tests.
- Section 3 provides an overview of the changes implemented as part of phase 2 of resource sufficiency evaluation enhancements. This includes new analysis on the following topics:
    - Adjustment for lower priority exports in CAISO's resource sufficiency evaluation.
    - Implementation of Assistance Energy Transfers.
- Section 4 summarizes the frequency and size of resource sufficiency evaluation failures.
- Section 5 provides an overview and analysis on the quantile regression method for calculating uncertainty in the resource sufficiency evaluation. This method was implemented on February 1.
- Section 6 summarizes demand-response based load adjustments used in the resource sufficiency evaluation.
- Section 7 summarizes WEIM import limits and transfers following a resource sufficiency evaluation failure.

DMM continues to welcome feedback on existing or additional metrics and analysis that WEIM entities and other stakeholders would find most helpful. Comments and questions may be submitted to DMM via email at DMM@caiso.com.

[1] California ISO, *EIM Resource Sufficiency Evaluation Enhancements Straw Proposal*, August 16, 2021: http://www.caiso.com/InitiativeDocuments/StrawProposal-ResourceSufficiencyEvaluationEnhancements.pdf

2 Overview of the flexible ramp sufficiency and capacity tests

As part of the Western Energy Imbalance Market (WEIM) design, each balancing area (including the California ISO) is subject to a resource sufficiency evaluation. The evaluation is performed prior to each hour to ensure that generation in each area is sufficient without relying on transfers from other balancing areas. The evaluation is made up of four tests: the power flow feasibility test, the balancing test, the bid range capacity test, and the flexible ramp sufficiency test.

The market software automatically limits transfers into a balancing area from other WEIM areas if a balancing area fails either of the following two tests:

- **The bid range capacity test (capacity test)** requires that each area provides incremental bid-in capacity to meet the imbalance between load, intertie, and generation base schedules.
- **The flexible ramp sufficiency test (flexibility test)** requires that each balancing area has enough ramping flexibility over an hour to meet the forecasted change in demand as well as uncertainty.

If an area fails either the flexible ramp sufficiency test or bid range capacity test in the *upward* direction, WEIM transfers into that area cannot be *increased*.[2] Similarly, if an area fails either test in the *downward* direction, transfers out of that area cannot be *increased*.

Bid range capacity test

The *bid range capacity test* requires that each area provide incremental (or decremental) bid-in capacity to meet the imbalance between load, intertie, and generation base schedules. Equation 2 shows the different components and mathematical formulation of the bid range capacity test. As shown in Equation 2, the requirement for the bid range capacity test is calculated as the *load forecast* plus *export base schedules* minus *import and generation base schedules*. Intertie uncertainty was removed on June 1, 2022.

Equation 2. Bid Range Capacity Test Formulation

$$Requirement = \underbrace{Load}_{\text{Load forecast}} + \underbrace{Export_{base} - Import_{base} - Generation_{base}}_{\text{Intertie and generation base schedules}}$$

If the requirement is positive, then the area must show sufficient incremental bid range capacity to meet the requirement, and if the requirement is negative, then sufficient decremental bid range capacity must be shown.

The bid range capacity used to the meet the requirement is calculated relative to the base schedules. For the California ISO (CAISO), the "base" schedules used in the requirement are the advisory schedules from the last binding 15-minute market run. For all other WEIM areas, the export, import, and generation schedules used in the requirement are the base schedules submitted as part of the hourly resource plan.

[2] If an area fails either test in the upward direction, net WEIM imports during the interval cannot exceed the greater of either the base transfer or optimal transfer from the last 15-minute market interval.

Since the bid range capacity is calculated relative to the base schedules, the upward capacity test can generally be expressed as follows:[3]

$$\underbrace{Generation_{maximum} + Net\ Import_{maximum}}_{\text{Upward capacity}} \geq \underbrace{Load}_{\text{Load forecast (requirement)}}$$

Incremental bid-in generation capacity is calculated as the range between the generation base schedule and the economic maximum, accounting for upward ancillary services and any de-rates (outages). Other resource constraints including start-times and ramp rates are not considered in the capacity test; 15-minute dispatchable imports and exports are included as bid range capacity.

Flexible ramp sufficiency test

The *flexible ramp sufficiency test* requires that each balancing area has enough ramping resources to meet expected upward and downward ramping needs in the real-time market without relying on transfers from other balancing areas. Each area must show sufficient ramping capability from the start of the hour to each of the four 15-minute intervals within the hour.

Equation 1 shows the different components and mathematical formulation of the flexible ramp sufficiency test. As shown in Equation 1, the requirement for the flexible ramp sufficiency test is calculated as the *forecasted change in load* plus the *uncertainty component* minus two components: (1) the *diversity benefit* and (2*) flexible ramping credits.* Any undersupply infeasibility in the last 15-minute market interval is also accounted for in the flexibility test requirement as of June 1, 2022.

Equation 1. Flexible Ramp Sufficiency Test Formulation

$$Up\ Requirement = \Delta Load + Up\ uncertainty - \min\begin{bmatrix} Net\ import\ capability, \\ Diversity\ benefit + Up\ credit \end{bmatrix} + Undersupply\ infeasibility$$

$$Down\ Requirement = \underbrace{-\Delta Load}_{\text{Change in load forecast}} + \underbrace{Down\ uncertainty}_{\text{Net load uncertainty}} - \underbrace{\min\begin{bmatrix} Net\ export\ capability, \\ Diversity\ benefit + Down\ credit \end{bmatrix}}_{\text{Discounts: diversity benefit and credit reduction capped by transfer capability}} - \underbrace{Undersupply\ infeasibility}_{\text{Undersupply infeasibility in last 15-minute market interval, excluding imbalance conformance}}$$

The diversity benefit reflects that system-level flexible ramping needs are typically smaller than the sum of the needs of individual balancing areas because of reduced uncertainty across a larger footprint. As a result, balancing areas receive a prorated diversity benefit discount based on this proportion.

The flexible ramping credits reflect the ability to reduce exports from a balancing area to increase upward ramping capability or to reduce imports to increase downward ramping capability.

[3] DMM has identified cases when the existing incremental approach for the capacity test relative to base schedules does not equal maximum capacity expected under a total approach. The incremental bid-range capacity can be positive only. If maximum capacity at the time of the test run is below base schedules, this difference will not be accounted for in the test. For more information, see DMM's *Comments on EIM Resource Sufficiency Evaluation Enhancements Issue Paper,* September 8, 2021: https://stakeholdercenter.caiso.com/Common/DownloadFile/25df1561-236b-4a47-9b1c-717b4a9cf9f0

As shown in Equation 1, the reduction in the flexibility test requirement because of any diversity benefit or flexible ramping credit is capped by the area's net import capability for the upward direction, or net export capability for the downward direction.

Last, as part of phase 1 of *resource sufficiency evaluation enhancements*, the flexibility test requirement now includes any undersupply infeasibility (power balance constraint relaxation) from the 15-minute market solution immediately prior to the resource sufficiency evaluation hour. This amount excludes any operator imbalance conformance.

As of February 1, 2023, the uncertainty component used in the flexible ramp sufficiency test is calculated using a regression method which considers forecasted net load currently on the system.[4] The measured uncertainty reflects extreme historical net load errors (95 percent confidence interval) adjusted to reflect forecasted conditions. The net load error observations used to calculate uncertainty in the resource sufficiency evaluation are measured from the difference between (1) binding 5-minute market net load forecasts and (2) the corresponding advisory 15-minute market net load forecast.

[4] California ISO, *Flexible Ramping Product Refinements Final Proposal*, August 31, 2020: http://www.caiso.com/InitiativeDocuments/FinalProposal-FlexibleRampingProductRefinements.pdf

3 Resource sufficiency evaluation enhancements phase 2

Phase 2 of the resource sufficiency evaluation enhancements was implemented on July 1, 2023. This included the following enhancements:

- **Adjustment for real-time low-priority and economic exports in CAISO's resource sufficiency evaluation.** These exports are no longer strictly counted as part of CAISO's demand obligation.
- **Implementation of Assistance Energy Transfers (AET).** This new option gives balancing areas access to excess WEIM supply that may not have been available otherwise following a resource sufficiency evaluation failure. Balancing areas can opt into AET to prevent their WEIM transfers from being limited during a test failure but will be subject to an ex-post surcharge.

More detailed information on each of these enhancements is discussed in the following sections.

Adjustment for lower priority exports in CAISO's resource sufficiency evaluation

Export schedules in the market can be based on economic bids or be self-scheduled (price-taking). The market defines different levels of prioritization for self-scheduled exports. The highest priority is given to Existing Transmission Contract (ETC) and Transmission Ownership Right (TOR) export schedules. Next, exports that are *not* supported by resource adequacy capacity are given *high-priority*. *Low-priority* exports are those supported by resource adequacy capacity. Within this category, export schedules that clear the residual unit commitment process can be self-scheduled in the real-time market with day-ahead priority (DA-LPT). Real-time low-priority price-taking (RT-LPT) exports are instead self-scheduled directly in real-time.

RT-LPT and economic exports that clear the hour-ahead scheduling process (HASP) are effectively no longer counted against CAISO's obligation in the resource sufficiency evaluation. During phase 1 of the initiative, analysis by the ISO showed the potential for advisory WEIM imports to support additional exports in HASP.[5] These hourly exports would then be counted against CAISO in the resource sufficiency evaluation but may not have existed without WEIM imports to balance these. Further, it was identified that these real-time low-priority and economic exports could be curtailed by CAISO operators during tight system conditions subject to operator judgement and consistent with good utility practices.[6] As a result, these export schedules were adjusted in CAISO's capacity and flexibility tests on July 1, 2023. In effect, only higher-priority exports as well as exports that were scheduled through the ISO's residual unit commitment process are counted in CAISO's demand obligation.[7]

Adjustment for lower priority exports in the capacity test

The upward capacity test requires that each balancing area show incremental available capacity to meet the imbalance between load, intertie, and generation "base" schedules. For CAISO, the "base" schedules used in the requirement are the advisory schedules from the last binding 15-minute market run, which includes all exports that cleared the hour-ahead scheduling process. Prior to the enhancements in July, all export schedules were included in CAISO's obligation (or test requirement) as extra demand while

[5] California ISO, *Interaction of Hourly Intertie Schedules and WEIM Transfers*, April 26, 2022: http://www.caiso.com/InitiativeDocuments/AnalysisReport-InteractionofHourlyIntertieSchedulesandTransfers-WEIMResourceSufficiencyEvaluationEnhancements.pdf

[6] California ISO, *WEIM RSE Enhancements Phase 2 Second Revised Final Proposal*, December 6, 2022: http://www.caiso.com/InitiativeDocuments/SecondRevisedFinalProposal-WEIMResourceSufficiencyEvaluationEnhancementsPhase2.pdf

[7] Including Existing Transmission Contract (ETC) and Transmission Ownership Right (TOR) export schedules.

only *15-minute dispatchable* export schedules were counted as incremental available capacity since these could be dispatched down.

Following the enhancements, all export schedules are still included in CAISO's demand obligation, but both real-time low-priority and economic export schedules are now counted as incremental available capacity on the other side of the equation. This effectively removes these exports from the capacity test as the additional exports in the test requirement are cancelled out by additional available capacity.[8] For example, assume that there is a real-time low priority export schedule of 75 MW in HASP. The 75 MW would still be counted as part of the obligation in the capacity test requirement, but this would be offset by 75 MW of upward available capacity counted towards meeting the requirement.

Figure 3.1 summarizes export schedules included in CAISO's capacity test requirement during peak hours (17 through 21) between July 17 and July 27. During this period, forecasted load reached over 40,000 MW. The blue and green bars show real-time low priority and economic exports that are now effectively removed from the capacity test obligation, as they are also counted as upward available supply reflecting potential decremental dispatch or curtailment of these exports. These amounts are showed separately in Figure 3.2. The remaining categories show Existing Transmission Contract (ETC), Transmission Ownership Right (TOR), high priority, and day-ahead low priority exports that are still counted in CAISO's test requirement. During the peak hours of this period, an average of just over 1,500 MW in real-time low priority and economic exports were effectively removed from CAISO's capacity test obligation.

During July, this enhancement had no impact on CAISO passing the *capacity* test. CAISO would have still passed the upward capacity test during all intervals in July had real-time low priority and economic exports been strictly counted against CAISO's demand obligation.

[8] DMM has identified cases in which the upward available capacity counted for a low-priority or economic export exceeds the contribution of the export to the capacity test requirement. This can occur in the first 15-minute interval of the evaluation hour when the export schedule is increasing based on the difference between the more granular ramp-constrained dispatch and the hourly block schedule. A fix is underway.

Figure 3.1 Export schedules in CAISO's capacity test requirement (peak hours, July 17 - July 27)

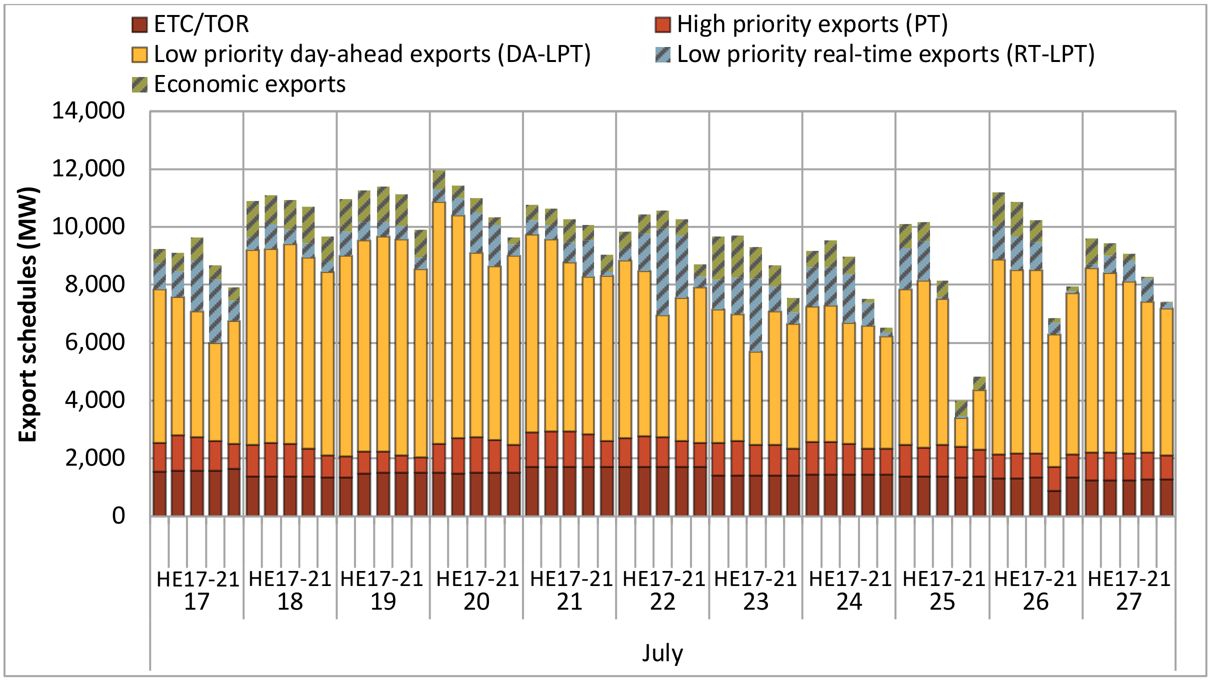
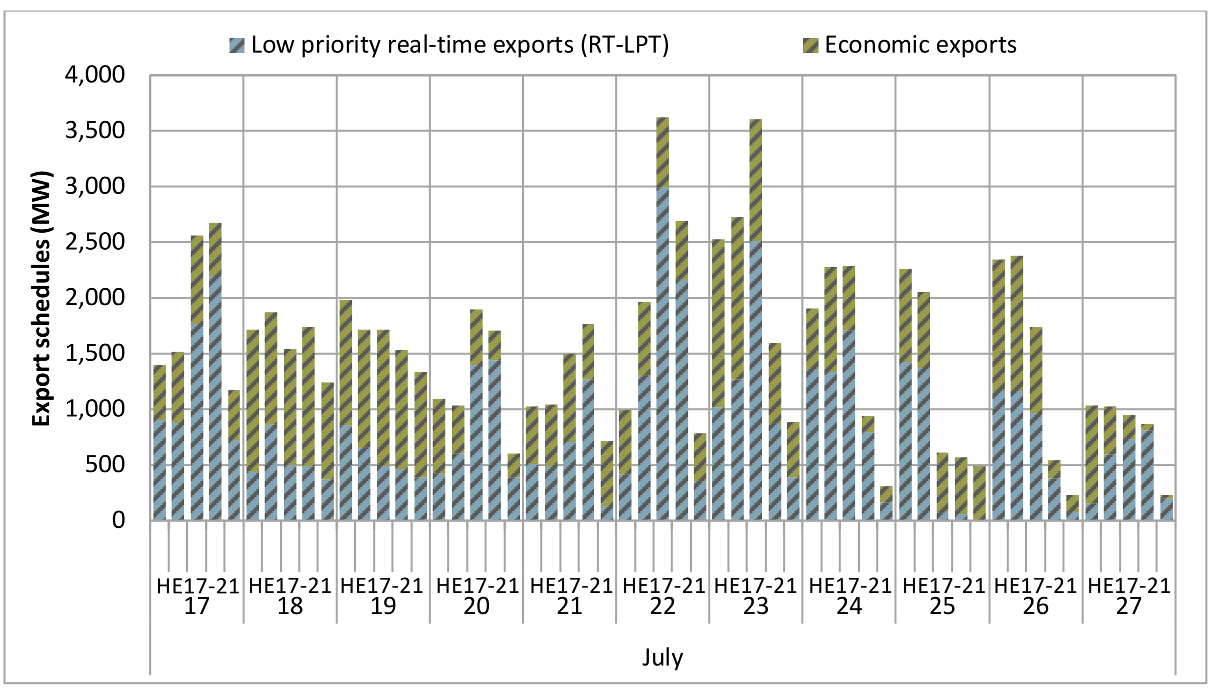


Figure 3.2 Real-time low-priority and economic exports in CAISO's capacity test requirement (peak hours, July 17 - July 27)



Adjustment for lower priority exports in the *flexibility test*

The flexibility test requires that each balancing area show sufficient flexibility over the evaluation hour to meet the forecasted change in demand as well as net load uncertainty. Flexible ramping capacity counted towards the requirement includes both economic energy bids which can be dispatched, as well as changes in fixed energy schedules from the binding interval immediately prior to the evaluation hour to each of the four intervals within the evaluation hour.

Prior to the enhancements, any change to an *hourly* intertie schedule from the previous hour to the evaluation hour was accounted for in the total flexible ramping capacity. So, an increase to an hourly export schedule (or decrease to an hourly import schedule) from the previous hour to the next would be counted as negative upward ramping capacity.[9] *15-minute-dispatchable* exports were instead counted as upward ramping capacity based on the schedule in the market interval immediately prior to the evaluation hour. This reflects that these exports schedules can be decrementally dispatched down to zero.

Following the enhancements, all real-time low-priority (RT-LPT) and economic export schedules are now counted as upward flexible capacity based on its schedule in the market interval immediately prior to the evaluation hour. For example, assume that in HASP there is RT-LPT export schedule of 50 MW in the previous hour and 75 MW during the evaluation hour. Prior to the enhancements, this change would have been counted as 25 MW of negative upward ramping capacity because of the fixed increase in demand. Following the enhancements, this would be counted as 50 MW of positive upward ramping capacity. All export schedules of higher priority than RT-LPT and economic are still counted in the flexibility test based on the change in the hourly schedule.

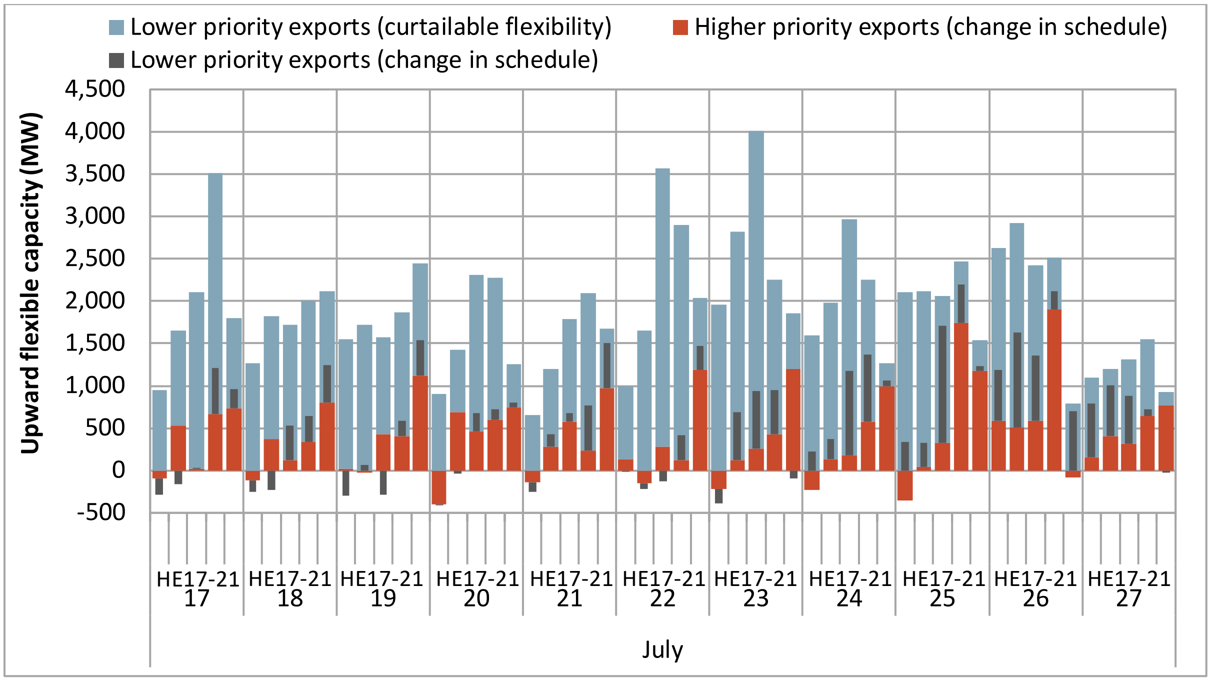
Figure 3.3 summarizes upward flexible ramp capacity counted from exports during peak hours (17 through 21) between July 17 and July 27. The blue bars show flexible capacity counted from *both* RT-LPT and economic exports while the red bars show flexible capacity counted from all other exports of higher priority. For all of the exports of higher priority (red bars), flexible capacity reflects the *change* from the export schedule in the binding interval immediately prior to the evaluation hour to the expected schedule during the evaluation hour. A decrease in the export schedule is shown as positive upward flexible capacity. The flexible capacity from the exports of higher priority would not have been impacted by the enhancements.

The blue bars show upward flexible capacity from RT-LPT and economic exports following the enhancement. These amounts reflect export schedules from the binding interval immediately prior to the evaluation hour that can be curtailed (increased upward flexibility) during the evaluation hour. The gray bars instead show exports of lower priority that would have been counted as upward flexible capacity prior to the enhancements based on the *change* in the export schedule. The difference between the gray and blue bars therefore show increased upward flexible capacity because of the enhancement. During the peak hours of this period, the enhancement increased CAISO upward flexible capacity by around 1,180 MW on average.

During July, this enhancement had no impact on CAISO passing the *flexibility* test. CAISO would have still passed the upward flexibility test during all intervals in July had real-time low priority and economic exports been counted based on their change in schedule.

[9] Alternatively, a decrease to an hourly export schedule (or increase to an import schedule) from the previous hour to the next would be counted as positive upward ramping capacity.

Figure 3.3 Upward flexible ramp capacity from exports in the flexibility test (peak hours, July 17 - July 27)



Assistance Energy Transfers

Assistance Energy Transfers (AET) gives balancing areas access to excess WEIM supply that may not have been available otherwise following an upward resource sufficiency evaluation failure. Without AET, a balancing area failing either the upward flexibility or upward capacity test would have net WEIM imports limited to the greater of either the base transfer or the optimal transfer from the last 15-minute market interval. Balancing areas can voluntarily opt in to the AET program to prevent their WEIM transfers from being limited during an upward resource sufficiency evaluation failure, but will be subject to an ex-post surcharge. Balancing areas must opt in or opt out of the program in advance of the trade date.[10]

The Assistance Energy Transfer surcharge is applied during any interval in which an opt-in balancing area fails the upward flexibility or capacity test. The surcharge is calculated as the *applicable real-time assistance energy transfer* times the real-time bid cap.[11] The applicable AET quantity is based on the lesser of either (1) the tagged dynamic WEIM transfers or (2) the amount by which the balancing area failed the resource sufficiency evaluation. If the tagged dynamic WEIM transfers are less than the amount by which the balancing area failed the resource sufficiency evaluation, then the applicable AET quantity is also reduced by a credit. The credit is either upward available balancing capacity for WEIM entities or cleared regulation up for the CAISO balancing area.

---

[10] Assistance Energy Transfer designation requests are submitted to Master File as *opt-in* or *opt-out* and include both a start and end date. The standard timeline to implement an opt-in or opt-out request is at least five business days in advance of the start date. An *emergency* opt-in request is also available should reliability necessitate this for two business days in advance of the start date. For more information, see: https://bpmcm.caiso.com/Pages/ViewPRR.aspx?PRRID=1525&IsDlg=0

[11] The soft bid cap is $1,000/MWh and can increase to the hard bid cap of $2,000/MWh under certain conditions.

Opting in to the Assistance Energy Transfer program does not guarantee that the balancing area will achieve additional WEIM supply following a resource sufficiency evaluation failure (compared to opting out of the program). It only removes the import limit that would have been in place following a test failure, allowing the market to freely and optimally schedule WEIM transfers based on supply and demand conditions in the system. If the import limit following a test failure was set high such that it is not restricting the optimal solution, then opting in or opting out of the program will have no effect on WEIM import supply in that interval. In this scenario, the balancing area will still be subject to an ex-post surcharge. This is shown in Figure 3.4.

The example in Figure 3.4 highlights the impact of assistance energy transfer opt-in, when the import limit following a resource sufficiency evaluation failure is high (and not restrictive) or low (and restrictive). In this example, the optimal WEIM import in the 5-minute market without any limit is 100 MW and the balancing area failed the resource sufficiency evaluation by 150 MW. The example also assumes that there is no available balancing capacity or regulation credit such that the applicable assistance energy transfer quantity equals the 5-minute market transfer.

The example highlights two conditions — in which the WEIM import limit following the resource sufficiency evaluation failure is high at 200 MW or low at 25 MW. In the cases when the WEIM import limit is *high* (and not restrictive on the solution), opting in to the program has no effect on achieving additional energy, but would result in an ex-post surcharge. Alternatively, in the cases when the WEIM import limit is *low* (and restricting transfers), opting in to the program will achieve additional energy (75 MW in this example). In both cases, opting in to receiving assistance energy transfers results in the same surcharge.

Figure 3.4 Example — impact of assistance energy transfer opt-in when WEIM import limits are high or low

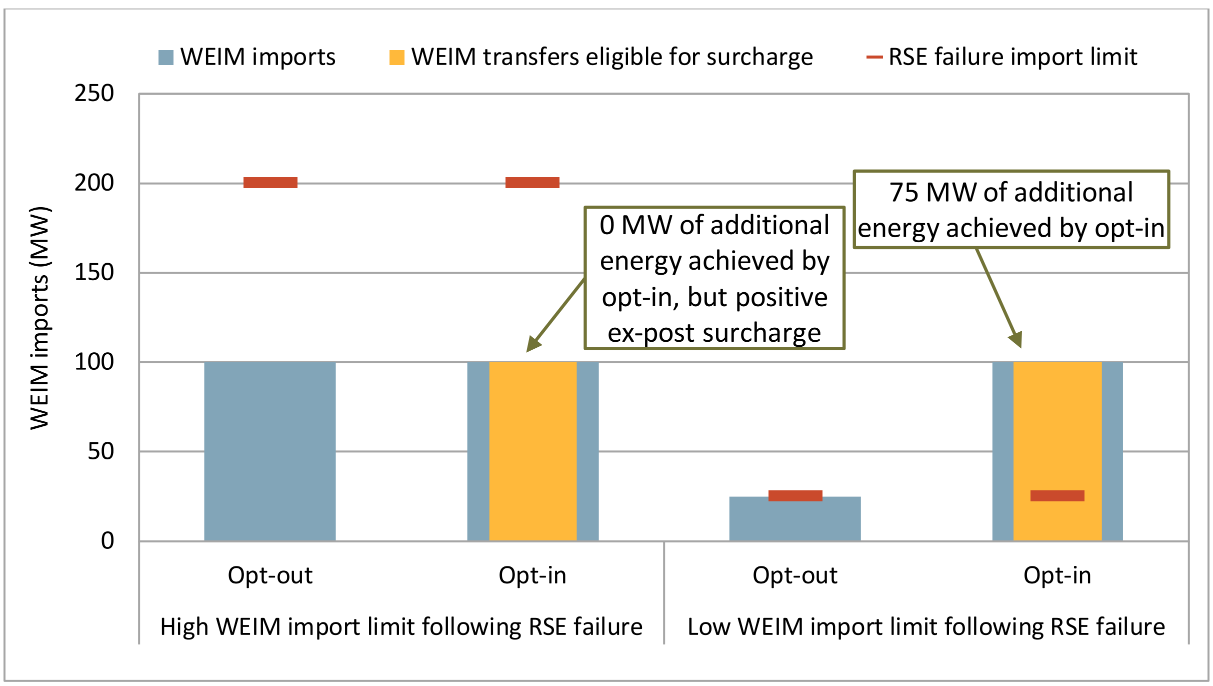
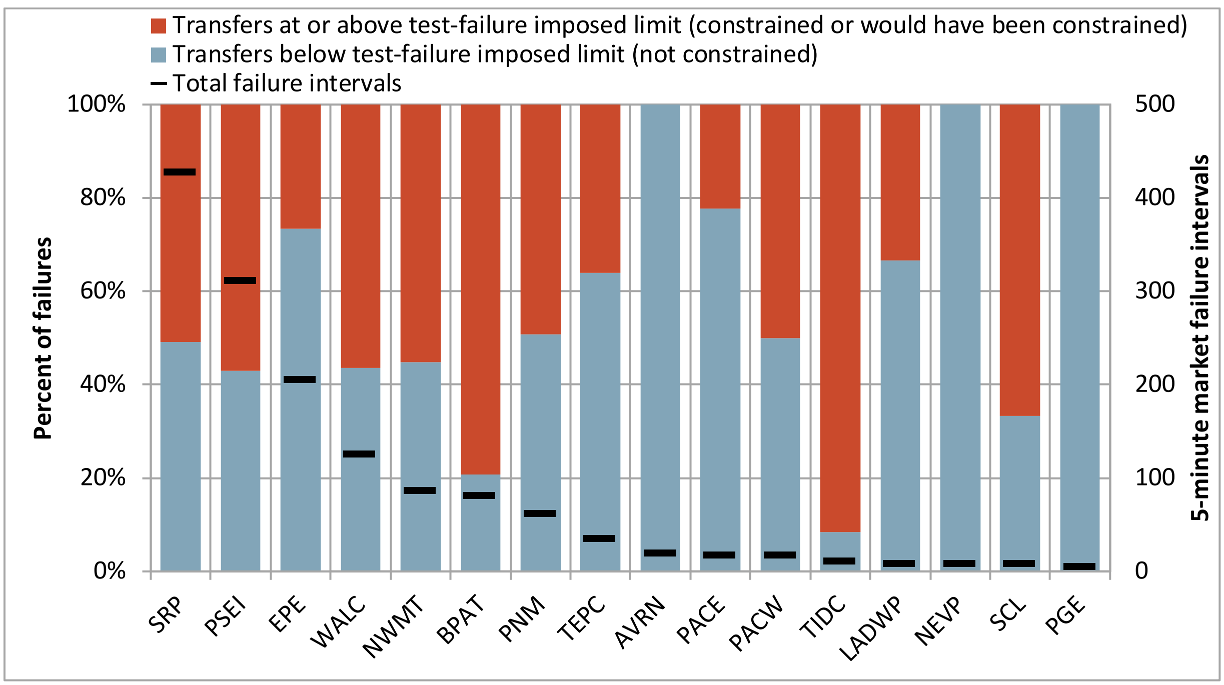


Figure 3.5 summarizes import limits following a resource sufficiency evaluation failure for both balancing areas that opted out or opted in to the assistance energy transfer program.

- The black horizontal line (right axis) shows the number of 5-minute market intervals with either a capacity or flexibility test failure.
- The blue bars (left axis) show the percent of 5-minute market failure intervals in which the resulting transfers after failing the resource sufficiency evaluation were *below* the import limit that was imposed (or would have been imposed for opt-in balancing areas). During these intervals, the optimal WEIM transfer was below the import limit such that an opt-in designation would have no effect on achieving additional energy.
- In all other failure intervals (red bars), the resulting transfers were either constrained to the limit imposed after failing the test or would have been constrained by the limit without an opt-in designation. During these intervals, additional WEIM transfers into the balancing area could have occurred through participation in the assistance energy transfer program.

Figure 3.5 Percent of upward 5-minute failure intervals in which WEIM imports were constrained or would have been constrained by test failure limits (July 2023)



4 Frequency of resource sufficiency evaluation failures

This section summarizes the frequency and shortfall amount for bid-range capacity test and flexible ramping sufficiency test failures.[12] If a balancing area fails either (or both) of these tests, then transfers between that and the rest of the WEIM areas are limited.

Figure 4.1 through Figure 4.4 show the percent of 15-minute intervals in which each WEIM area failed the upward capacity or the flexibility tests as well as the average shortfall of those test failures.[13] Figure 4.5 through Figure 4.8 provide the same information for the downward direction. The dash indicates that the area did not fail the test during the month.

In July:

- Salt River Project failed the upward flexibility test during around 3.7 percent of intervals and the upward capacity test during around 2.8 percent of intervals.
- Puget Sound Energy failed the upward flexibility test during around 2.6 percent of intervals and the upward capacity test during around 1.5 percent of intervals.
- El Paso Electric failed the upward flexibility test during around 2.1 percent of intervals.

Net load uncertainty — which is added to the flexibility test requirement — was adjusted on February 1, 2023 as part of flexible ramping enhancements. The uncertainty was adjusted to incorporate current load, solar, and wind forecast information using a technique called *mosaic quantile regression*. This regression combines both histogram and quantile regression models to estimate the lower and upper extremes of uncertainty that might materialize. For more information on this regression, see Section 3. The capacity test currently does not include any net load uncertainty adder in the requirement.

Figure 4.9 summarizes the overlap between failure of the upward capacity and the flexibility tests during the month. The black horizontal line (right axis) shows the number of 15-minute intervals with either a capacity or a flexibility test failure for each WEIM area. The areas are shown in descending number of failure intervals. The bars (left axis) show the percent of the failure intervals that meet the condition.

Figure 4.10 shows the same information for the downward direction. Areas that did not fail either the capacity or the flexibility tests during this period were omitted from the figure. Across both directions, the flexibility test was more often the source of the resource sufficiency evaluation failure.

[12] Results in this section exclude known invalid test failures. These can occur because of a market disruption, software defect, or other errors.

[13] Results in these figures reflect the final resource sufficiency evaluation (40 minutes prior to the evaluation hour).

Figure 4.1 Frequency of upward capacity test failures (percent of 15-minute intervals)

	2022								2023						
	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Arizona Publ. Serv.	—	—	—	—	—	—	—	0.1	0.4	0.5	0.7	0.2	0.0	0.1	—
Avangrid												0.0	—	—	—
Avista	—	0.2	0.2	0.0	—	—	—	0.1	—	—	—	0.1	0.0	—	—
BANC	—	—	—	0.0	0.3	—	—	—	—	—	—	—	—	—	—
BPA	—	0.1	—	0.0	0.5	—	—	0.4	—	—	—	0.2	—	0.3	0.4
California ISO	—	—	—	—	0.1	—	—	—	—	—	—	—	—	—	—
El Paso Electric												0.0	0.1	0.3	0.8
Idaho Power	—	—	—	0.2	0.2	—	—	—	—	—	—	0.0	0.1	—	—
LADWP	—	—	0.0	—	—	—	—	—	0.1	—	—	—	—	—	0.1
NorthWestern En.	—	—	—	0.1	0.1	—	0.2	0.1	0.3	0.1	—	—	—	—	0.3
NV Energy	0.1	0.0	0.1	—	—	—	—	—	—	—	—	—	0.0	—	0.0
PacifiCorp East	—	—	—	—	0.1	—	—	0.3	—	—	—	—	—	—	0.0
PacifiCorp West	0.2	0.0	1.0	0.2	0.0	—	0.0	0.0	0.1	0.1	—	—	—	—	—
Portland Gen. Elec.	—	—	—	0.1	—	—	0.3	—	—	0.0	0.0	0.1	0.4	0.1	0.0
Powerex	—	—	—	0.2	—	—	0.0	—	—	—	—	—	0.1	—	—
PSC of New Mexico	—	—	—	—	—	—	—	—	—	—	0.7	0.3	0.2	0.0	—
Puget Sound En.	0.0	0.2	—	—	0.2	0.1	0.0	—	—	0.0	0.2	—	0.1	0.5	1.5
Salt River Proj.	1.0	0.2	0.2	0.4	0.4	0.2	0.0	0.0	1.0	0.4	1.1	0.9	0.2	0.0	2.8
Seattle City Light	—	—	0.2	0.1	0.2	0.0	0.0	0.2	0.0	0.1	—	—	—	—	0.1
Tacoma Power	0.1	0.0	0.0	0.2	0.0	—	—	—	0.0	0.1	0.1	—	0.1	—	—
Tucson Elec. Pow.	—	—	—	0.1	—	—	—	—	0.1	0.0	—	—	—	—	0.3
Turlock Irrig. Dist.	—	0.1	—	—	—	—	—	0.2	—	—	—	0.0	—	—	0.1
WAPA DSW												2.3	0.8	0.7	1.1

Figure 4.2 Average shortfall of upward capacity test failures (MW)

	2022								2023						
	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Arizona Publ. Serv.	—	—	—	—	—	—	—	32	316	41	817	637	35	192	—
Avangrid												1	—	—	—
Avista	—	6	27	5	—	—	—	9	—	—	—	20	1	—	—
BANC	—	—	—	37	264	—	—	—	—	—	—	—	—	—	—
BPA	—	81	—	8	336	—	—	68	—	—	—	55	—	238	118
California ISO	—	—	—	—	141	—	—	—	—	—	—	—	—	—	—
El Paso Electric												6	8	88	20
Idaho Power	—	—	—	60	37	—	—	—	—	—	—	23	12	—	—
LADWP	—	—	0	—	—	—	—	—	49	—	—	—	—	—	10
NorthWestern En.	—	—	—	86	64	—	91	40	56	68	—	—	—	—	70
NV Energy	67	2	36	—	—	—	—	—	—	—	—	—	53	—	3
PacifiCorp East	—	—	—	—	124	—	—	293	—	—	—	—	—	—	116
PacifiCorp West	11	50	24	36	4	—	17	16	9	84	—	—	—	—	—
Portland Gen. Elec.	—	—	—	1	—	—	25	—	—	38	13	1	19	12	24
Powerex	—	—	—	142	—	—	50	—	—	—	—	—	131	—	—
PSC of New Mexico	—	—	—	—	—	—	—	—	—	—	52	24	106	5	—
Puget Sound En.	1	27	—	—	13	24	5	—	—	15	22	—	26	45	29
Salt River Proj.	44	51	41	214	132	30	17	26	44	44	54	30	38	1	65
Seattle City Light	—	—	15	9	7	5	2	16	2	16	—	—	—	—	2
Tacoma Power	2	1	3	6	0	—	—	—	0	5	0	—	2	—	—
Tucson Elec. Pow.	—	—	—	20	—	—	—	—	65	1	—	—	—	—	54
Turlock Irrig. Dist.	—	104	—	—	—	—	—	1	—	—	—	2	—	—	8
WAPA DSW												133	74	5	18

Figure 4.3 Frequency of upward flexibility test failures (percent of 15-minute intervals)

	2022								2023						
	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Arizona Publ. Serv.	—	—	0.0	0.1	—	—	0.1	0.4	0.9	1.8	2.5	1.1	0.2	0.1	—
Avangrid												1.0	0.7	0.1	0.2
Avista	0.5	1.0	0.5	0.1	—	0.1	—	0.1	—	0.0	0.0	0.2	0.2	0.0	—
BANC	—	—	—	—	0.3	—	—	—	—	—	—	—	0.1	—	—
BPA	0.9	3.1	3.3	1.0	1.1	0.2	0.1	0.4	—	0.1	0.6	0.2	1.2	0.3	1.3
California ISO	—	—	—	0.1	0.5	0.0	—	—	—	—	—	—	—	—	—
El Paso Electric												0.8	0.7	0.3	2.1
Idaho Power	—	—	0.2	0.2	0.5	—	0.1	—	0.0	0.1	0.3	0.3	0.5	0.1	—
LADWP	—	—	—	—	0.1	0.1	—	—	—	0.3	—	0.1	0.0	0.1	0.0
NorthWestern En.	—	0.1	0.3	1.0	0.2	—	0.5	0.8	0.3	0.1	0.2	0.8	0.3	0.2	1.0
NV Energy	0.8	0.2	—	0.1	0.1	0.1	0.2	0.0	0.1	0.3	0.0	0.1	0.1	0.0	0.1
PacifiCorp East	0.1	0.1	0.2	0.1	—	0.1	—	0.0	0.1	—	0.0	0.1	—	0.0	0.2
PacifiCorp West	0.1	0.0	—	0.1	0.1	—	0.1	—	0.1	0.1	—	0.1	0.6	0.0	0.2
Portland Gen. Elec.	—	0.0	0.4	0.1	0.1	0.2	1.0	0.1	0.0	0.1	0.0	0.1	1.5	0.7	0.1
Powerex	—	—	—	0.3	0.1	—	—	—	—	0.2	—	—	—	—	—
PSC of New Mexico	0.1	—	0.4	—	0.0	0.2	0.1	0.8	0.2	—	1.2	5.1	0.9	0.6	0.7
Puget Sound En.	—	0.1	0.4	0.2	0.3	—	0.0	—	—	0.1	0.8	0.2	1.0	0.6	2.6
Salt River Proj.	0.2	0.5	0.6	1.1	0.6	0.6	0.5	0.8	3.5	1.2	1.7	2.0	0.6	0.2	3.7
Seattle City Light	—	—	0.2	0.0	0.2	—	0.1	0.0	—	0.1	—	—	—	—	—
Tacoma Power	0.1	0.1	0.0	0.1	0.1	—	0.2	—	0.2	0.1	0.2	—	0.1	—	—
Tucson Elec. Pow.	0.1	—	—	—	0.4	0.0	—	0.2	0.3	0.3	0.3	0.1	0.1	—	0.2
Turlock Irrig. Dist.	—	—	—	—	0.1	—	—	1.2	—	—	—	0.0	—	—	0.1
WAPA DSW												2.7	0.7	0.8	0.3

Figure 4.4 Average shortfall of upward flexibility test failures (MW)

	2022								2023						
	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Arizona Publ. Serv.	—	—	28	28	—	—	15	65	154	77	288	119	36	76	—
Avangrid												79	13	9	20
Avista	29	26	19	30	—	5	—	11	—	13	12	35	14	4	—
BANC	—	—	—	—	237	—	—	—	—	—	—	—	64	—	—
BPA	68	71	50	56	232	43	42	114	—	36	62	99	82	164	114
California ISO	—	—	—	684	671	53	—	—	—	—	—	—	—	—	—
El Paso Electric												24	15	123	19
Idaho Power	—	—	13	34	45	—	14	—	5	51	28	42	24	46	—
LADWP	—	—	—	—	36	9	—	—	—	45	—	21	30	56	51
NorthWestern En.	—	10	15	22	83	—	45	30	44	16	40	14	33	11	32
NV Energy	66	89	—	80	88	41	91	69	60	69	29	164	59	24	52
PacifiCorp East	77	9	34	43	—	16	—	13	53	—	101	47	—	18	36
PacifiCorp West	24	5	—	31	28	—	62	—	14	79	—	30	146	2	35
Portland Gen. Elec.	—	8	72	25	16	19	46	12	39	16	9	61	49	37	27
Powerex	—	—	—	318	101	—	—	—	—	86	—	—	—	—	—
PSC of New Mexico	33	—	70	—	22	38	39	36	19	—	45	47	26	21	35
Puget Sound En.	—	49	46	17	21	—	29	—	—	45	46	29	59	48	55
Salt River Proj.	43	89	45	156	72	61	38	67	47	39	48	54	72	53	77
Seattle City Light	—	—	17	2	8	—	4	6	—	23	—	—	—	—	—
Tacoma Power	206	6	3	5	3	—	16	—	6	3	6	—	21	—	—
Tucson Elec. Pow.	22	—	—	—	44	5	—	97	67	28	31	36	30	—	35
Turlock Irrig. Dist.	—	—	—	—	3	—	—	6	—	—	—	1	—	—	12
WAPA DSW												71	122	21	9

Figure 4.5 Frequency of downward capacity test failures (percent of 15-minute intervals)

	2022								2023						
	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Arizona Publ. Serv.	0.0	0.0	—	—	—	—	—	0.1	—	—	0.6	—	—	—	—
Avangrid												—	—	—	—
Avista	—	—	0.2	—	—	0.0	—	—	—	—	—	0.0	—	—	—
BANC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	—	—	—	—	0.1	—	—	—	—	0.1	—	0.2	0.1	—	—
California ISO	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
El Paso Electric												0.2	0.1	0.3	0.2
Idaho Power	0.6	—	—	—	—	—	—	—	—	—	—	—	—	0.0	—
LADWP	0.2	—	—	—	—	—	—	—	0.1	—	—	—	—	0.0	—
NorthWestern En.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
NV Energy	0.1	0.5	—	—	—	—	—	—	—	—	—	0.1	0.1	0.6	0.1
PacifiCorp East	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PacifiCorp West	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Portland Gen. Elec.	—	0.0	—	—	—	—	—	—	—	—	—	—	—	—	—
Powerex	0.1	—	—	0.0	—	—	0.0	—	—	—	—	—	—	—	0.0
PSC of New Mexico	0.1	—	—	—	—	—	—	—	—	—	0.1	0.3	—	—	—
Puget Sound En.	0.0	0.7	0.1	—	—	—	—	—	—	—	—	—	0.1	—	—
Salt River Proj.	0.4	0.5	0.1	0.2	1.1	0.2	0.3	—	0.4	1.5	0.2	0.3	0.6	0.4	0.7
Seattle City Light	—	0.0	0.1	—	0.2	—	—	—	—	0.1	—	—	—	—	—
Tacoma Power	0.1	—	0.6	0.3	—	0.1	—	0.2	—	0.2	0.1	—	—	—	0.0
Tucson Elec. Pow.	—	0.0	—	—	—	—	—	—	—	—	—	—	—	—	—
Turlock Irrig. Dist.	—	—	—	—	—	—	—	—	—	0.1	—	—	—	—	—
WAPA DSW												0.2	—	0.8	0.1

Figure 4.6 Average shortfall of downward capacity test failures (MW)

	2022								2023						
	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Arizona Publ. Serv.	33	19	—	—	—	—	—	146	—	—	210	—	—	—	—
Avangrid												—	—	—	—
Avista	—	—	52	—	—	14	—	—	—	—	—	2	—	—	—
BANC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	—	—	—	—	31	—	—	—	—	435	—	12	99	—	—
California ISO	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
El Paso Electric												91	8	11	15
Idaho Power	7	—	—	—	—	—	—	—	—	—	—	—	—	4	—
LADWP	34	—	—	—	—	—	—	—	16	—	—	—	—	19	—
NorthWestern En.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
NV Energy	53	41	—	—	—	—	—	—	—	—	—	14	42	124	51
PacifiCorp East	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PacifiCorp West	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Portland Gen. Elec.	—	23	—	—	—	—	—	—	—	—	—	—	—	—	—
Powerex	175	—	—	13	—	—	12	—	—	—	—	—	—	—	15
PSC of New Mexico	6	—	—	—	—	—	—	—	—	—	9	233	—	—	—
Puget Sound En.	61	31	19	—	—	—	—	—	—	—	—	—	26	—	—
Salt River Proj.	41	46	8	72	27	11	14	—	15	15	6	79	27	35	39
Seattle City Light	—	2	7	—	6	—	—	—	—	7	—	—	—	—	—
Tacoma Power	3	—	5	8	—	33	—	2	—	4	7	—	—	—	1
Tucson Elec. Pow.	—	6	—	—	—	—	—	—	—	—	—	—	—	—	—
Turlock Irrig. Dist.	—	—	—	—	—	—	—	—	—	0	—	—	—	—	—
WAPA DSW												9	—	12	13

Figure 4.7 Frequency of downward flexibility test failures (percent of 15-minute intervals)

	2022								2023						
	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Arizona Publ. Serv.	0.5	0.2	—	—	0.1	0.2	0.2	0.1	0.9	0.5	2.1	0.7	1.2	0.1	—
Avangrid												0.1	—	—	—
Avista	—	0.1	—	—	0.1	0.2	—	0.0	—	—	0.1	0.1	0.1	—	—
BANC	0.1	0.1	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	0.1	0.2	—	0.0	0.3	—	0.2	0.2	—	0.0	0.1	0.6	5.5	0.0	0.4
California ISO	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
El Paso Electric												0.2	0.9	1.9	0.5
Idaho Power	0.4	—	—	0.0	—	—	—	—	—	—	0.9	0.2	—	—	—
LADWP	—	—	—	—	—	—	—	—	0.1	—	—	—	—	—	—
NorthWestern En.	0.5	1.9	0.2	—	—	—	0.0	0.1	—	0.0	—	—	0.2	0.2	—
NV Energy	1.3	2.0	0.6	0.2	0.5	0.5	0.6	0.1	0.1	0.1	0.1	0.0	0.1	0.4	0.1
PacifiCorp East	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PacifiCorp West	0.1	0.4	0.5	—	—	0.1	—	0.0	—	—	—	0.0	0.2	0.0	—
Portland Gen. Elec.	—	0.2	—	—	—	—	—	—	—	—	—	—	—	—	—
Powerex	0.3	0.2	—	0.1	0.1	0.1	—	—	0.1	0.1	—	0.2	—	—	0.0
PSC of New Mexico	1.8	0.7	0.0	0.0	0.2	0.2	0.1	—	0.0	—	0.4	1.6	2.1	—	0.1
Puget Sound En.	0.2	2.3	0.1	—	—	0.1	—	—	—	—	—	—	0.8	—	—
Salt River Proj.	0.4	0.5	0.2	0.2	1.0	0.2	0.9	0.3	1.4	3.3	1.0	0.3	0.1	0.1	0.1
Seattle City Light	0.1	0.3	0.1	0.8	0.3	—	0.2	0.6	0.1	0.2	0.0	0.3	0.0	0.3	0.4
Tacoma Power	0.3	—	0.5	0.2	—	—	—	0.1	—	0.2	0.1	—	—	—	0.0
Tucson Elec. Pow.	—	—	—	—	—	—	0.0	—	—	—	—	—	—	—	—
Turlock Irrig. Dist.	0.1	0.5	0.1	0.1	—	—	0.1	—	0.1	0.1	0.1	0.1	0.4	—	—
WAPA DSW												2.7	0.5	0.7	0.1

Figure 4.8 Average shortfall of downward flexibility test failures (MW)

	2022								2023						
	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul
Arizona Publ. Serv.	58	33	—	—	81	20	28	31	46	45	49	33	64	14	—
Avangrid												13	—	—	—
Avista	—	20	—	—	11	20	—	26	—	—	16	12	29	—	—
BANC	5	15	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	212	55	—	4	149	—	77	191	—	72	78	102	741	27	62
California ISO	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
El Paso Electric												8	15	30	36
Idaho Power	55	—	—	13	—	—	—	—	—	—	31	11	—	—	—
LADWP	—	—	—	—	—	—	—	—	34	—	—	—	—	—	—
NorthWestern En.	12	27	14	—	—	—	2	16	—	17	—	—	39	16	—
NV Energy	49	98	151	59	58	43	28	62	83	104	90	22	13	96	120
PacifiCorp East	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PacifiCorp West	55	28	11	—	—	12	—	22	—	—	—	6	44	7	—
Portland Gen. Elec.	—	18	—	—	—	—	—	—	—	—	—	—	—	—	—
Powerex	257	244	—	87	62	86	—	—	23	30	—	48	—	—	85
PSC of New Mexico	144	34	3	9	40	16	15	—	16	—	115	112	75	—	15
Puget Sound En.	54	33	47	—	—	11	—	—	—	—	—	—	38	—	—
Salt River Proj.	62	34	54	155	42	42	113	38	52	54	84	45	49	23	172
Seattle City Light	7	11	10	21	10	—	24	39	10	28	6	6	30	15	7
Tacoma Power	14	—	5	4	—	—	—	8	—	3	4	—	—	—	2
Tucson Elec. Pow.	—	—	—	—	—	—	14	—	—	—	—	—	—	—	—
Turlock Irrig. Dist.	5	6	3	2	—	—	5	—	6	6	14	8	4	—	—
WAPA DSW												55	8	16	12

Figure 4.9 Upward capacity/flexibility test failure intervals by concurrence (July 2023)

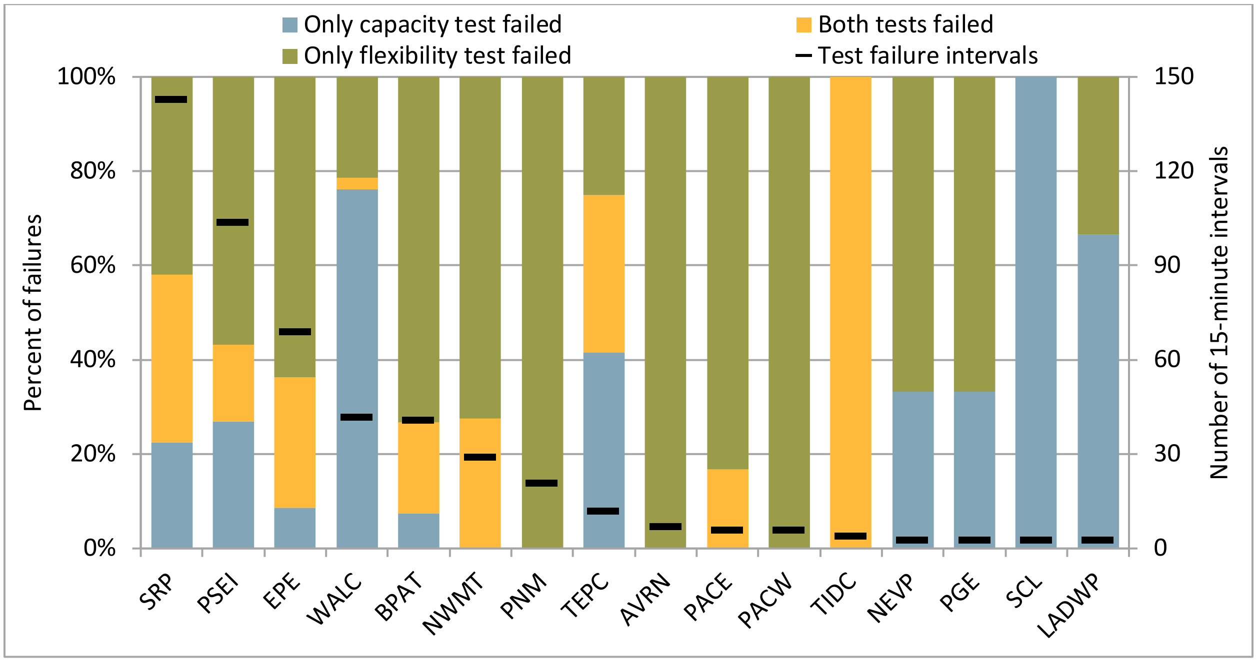
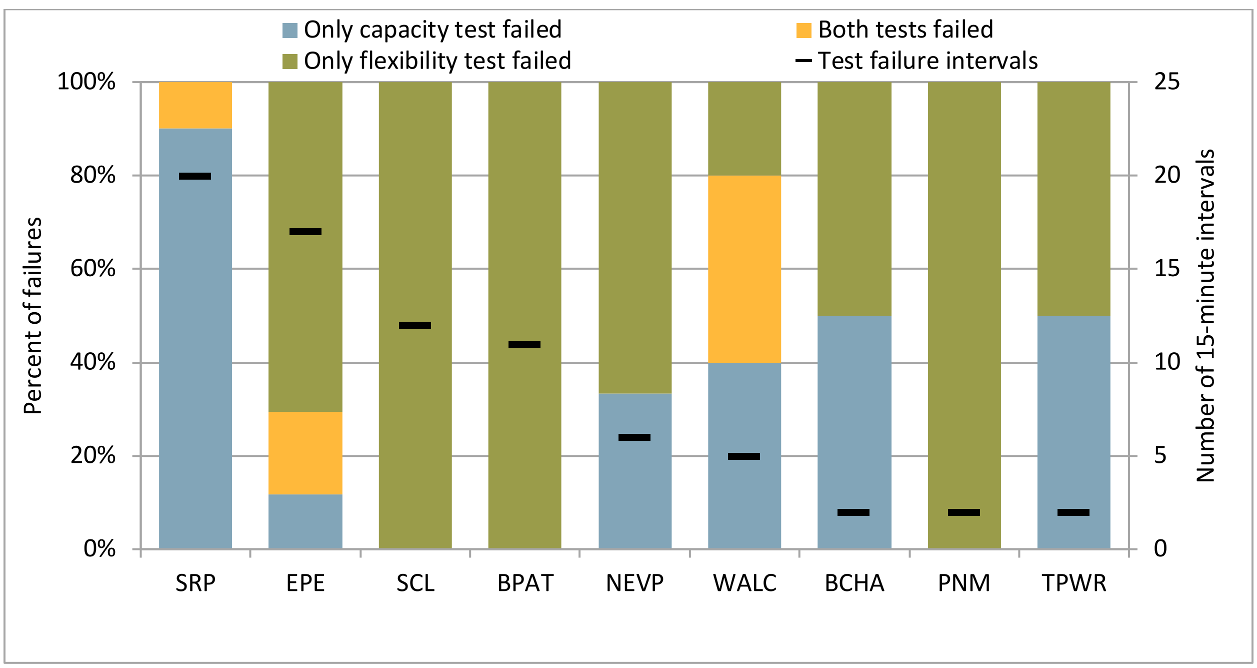


Figure 4.10 Downward capacity/flexibility test failure intervals by concurrence (July 2023)



Impact of earlier runs of the resource sufficiency evaluation on market results

There are three runs of the resource sufficiency evaluation, at 75 minutes (first run), 55 minutes (second run), and 40 minutes (final run) prior to each evaluation hour. The first and second runs are sometimes considered the *advisory runs* with the results of the final evaluation at 40 minutes prior considered the *binding run*. The previous section summarized the frequency of resource sufficiency evaluation failures in the final run. However, the results in the earlier runs of the resource sufficiency evaluation can also impact binding market results in several key ways. These are discussed below.

Nodal flexible ramping capacity procurement in the first 15-minute interval of each hour

Flexible ramping product nodal procurement in the first 15-minute market interval of each hour is dependent on the second run of the resource sufficiency evaluation at 55 minutes prior to the evaluation hour.

The results of the resource sufficiency evaluation are used as an input for the flexible ramping product. As part of the enhancements implemented on February 1, the real-time market will enforce an area-specific uncertainty target for balancing areas that fail the resource sufficiency evaluation. This target can only be met by flexible capacity within that area. In contrast, flexible capacity for the group of balancing areas that pass the resource sufficiency evaluation are pooled together to meet the uncertainty target for the rest of the system.

Deliverable flexible capacity awards are produced through two deployment scenarios that adjust the expected net load forecast in the *following* interval by the lower and upper ends of uncertainty that might materialize. This ensures that upward and downward flexible capacity awards do not violate transmission or transfer constraints. A consequence of this is that binding flex ramp awards in the first 15-minute market interval of each hour are now dependent on the second run of the resource sufficiency evaluation at 55 minutes prior to the evaluation hour — based on the latest information available at the time of this market run.

Figure 4.11 and Figure 4.12 summarize the first interval of each evaluation hour during the month with a failure in the second (T-55) or final (T-40) resource sufficiency evaluation.[14] This reflects failure of *either* the flexibility or capacity test in the second or final run. The red and yellow bars show intervals with a failure in the second evaluation (T-55) and whether the balancing area ultimately failed or passed in that interval based on the final evaluation results at 40 minutes prior to the hour. The dashed blue region instead shows cases in the first interval of the hour when the balancing area passed the second evaluation (T-55) but failed the final evaluation (T-40). In these intervals, the balancing area would have been included in the pass-group for the purpose of procuring flexible ramping capacity. The pass-group uncertainty requirement includes any diversity benefit of reduced uncertainty over a larger footprint.

[14] Areas that did not fail in the first interval of a resource sufficiency evaluation at T-55 or T-40 during this period were omitted from these figures.

Figure 4.11 Upward resource sufficiency evaluation failures in first 15-minute interval of hour (July 2023)

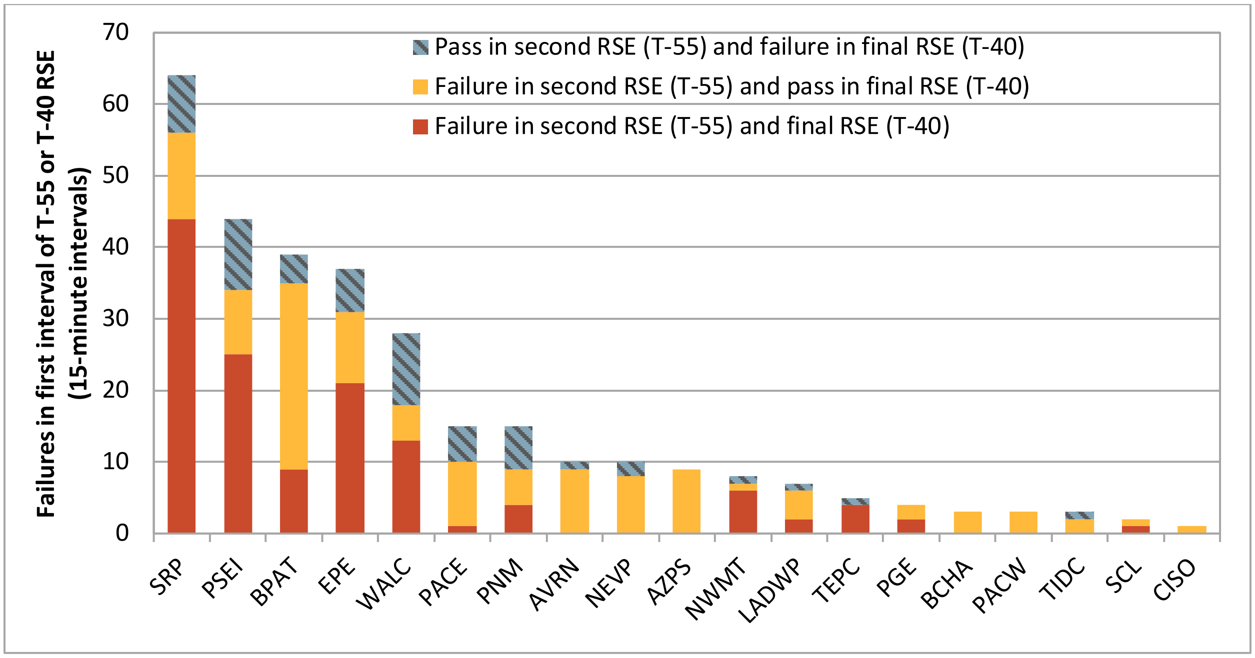
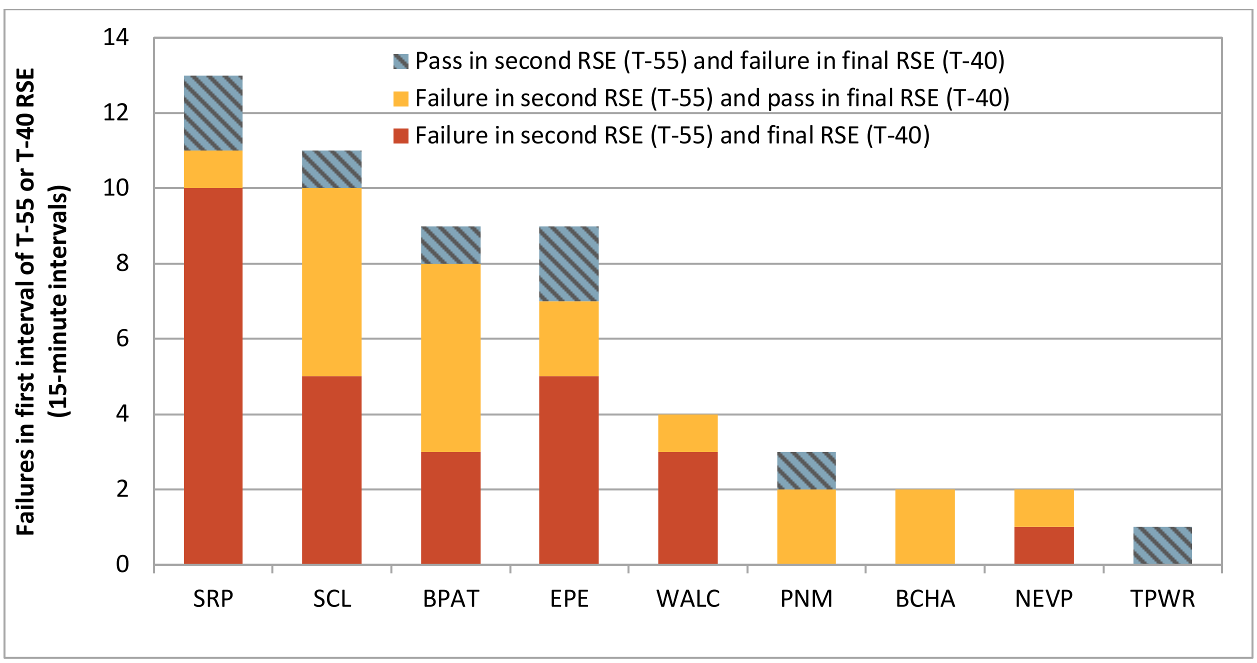


Figure 4.12 Downward resource sufficiency evaluation failures in first 15-minute interval of hour (July 2023)



Calculating uncertainty for balancing areas passing the resource sufficiency evaluation

Uncertainty estimates created for the group of balancing areas that pass the resource sufficiency evaluation in the *first and second* interval of each hour are based on earlier test results.

As part of the enhancements implemented on February 1, uncertainty is now calculated based on regression results that use historical data to predict uncertainty relative to load, solar, and wind forecasts.[15] Once all of the regressions are complete, the regression outputs can be combined with current forecast information to calculate uncertainty for each interval.

For a single balancing area that failed the resource sufficiency evaluation, these regressions can be performed in advance and local uncertainty targets can be readily determined based on current forecast information. However, for instead the group of balancing areas that pass the resource sufficiency evaluation (known as the pass-group), the regression procedure needs to first determine which balancing areas make up this group so that it can perform the regression using historical data accordingly for that group.

To perform the regressions to estimate the pass-group uncertainty, the composition of balancing areas in this group is based on earlier test results for the first and second 15-minute market interval of each hour. In the first interval, the results from the earliest resource sufficiency evaluation (T-75) is used to define the pass-group. In the second interval, the results from the second resource sufficiency evaluation (T-55) is used to define the pass-group. This is based on the latest information available at the time of this process.

However, the current weather information that is ultimately combined with the regression results to calculate uncertainty are instead consistent with the group of balancing areas in the pass-group for flexible ramping capacity procurement. This is based on the second run of the resource sufficiency evaluation (T-55) for interval 1 and the final resource sufficiency evaluation (T-40) for intervals 2 to through 4. Table 4.1 summarizes this inconsistency by showing which resource sufficiency evaluation run is used for each interval and process.

Table 4.1 Source of pass-group for calculating uncertainty and procuring flexible ramping capacity

15-minute market interval	Current weather information for calculating uncertainty and flex ramp procurement	Regression inputs and outputs
1	Second run (T-55)	First run (T-75)
2	Final run (T-40)	Second run (T-55)
3	Final run (T-40)	Final run (T-40)
4	Final run (T-40)	Final run (T-40)

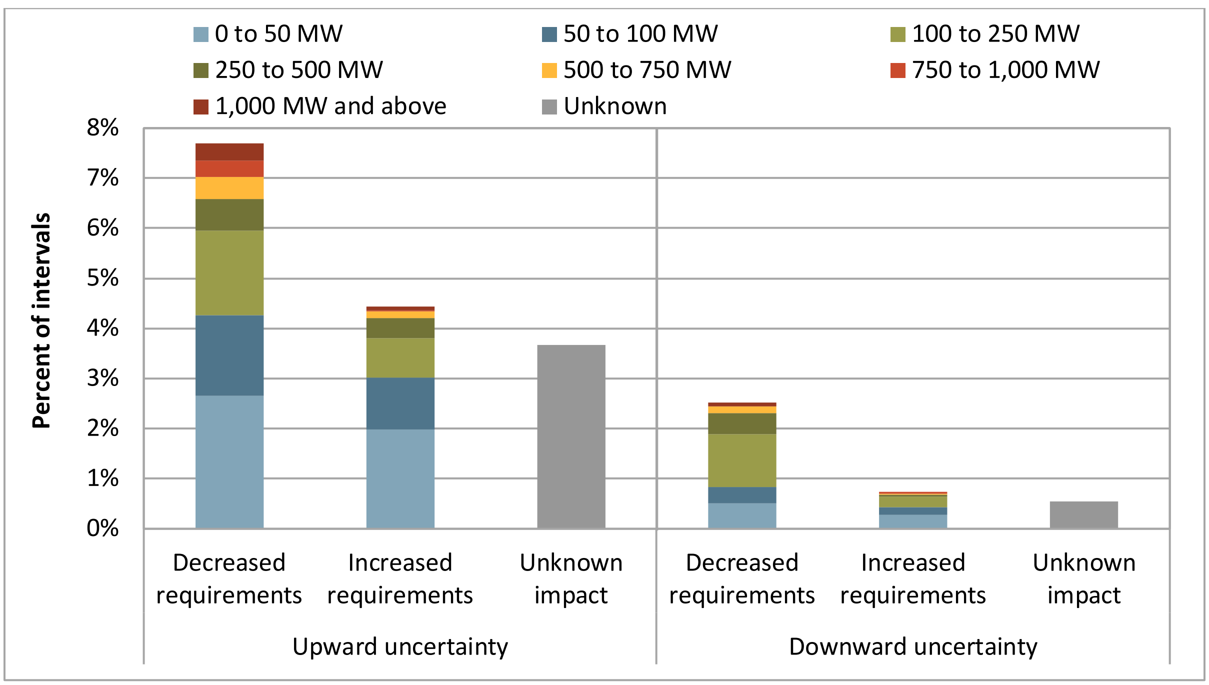
Using an inconsistent composition of balancing areas in the pass-group between the forecast and regression information can create significant swings in the calculated uncertainty for this group. For example, if you have a model to predict uncertainty based on forecast information of all but one balancing area passing the test (based on earlier test results), but then combine this with current forecast information of all balancing areas (based on later test results), then the calculated uncertainty can be disconnected from forecasted conditions in the system. DMM has requested that the ISO consider options to resolve inconsistencies in the composition of balancing areas in the pass-group.

[15] The calculation of uncertainty is described in more depth in the following section.

During about 17 percent of intervals during the month, the composition of balancing areas in the pass-group between the current forecast information and regression information were inconsistent for either upward or downward uncertainty. Figure 4.13 summarizes the impact of this inconsistency on pass-group uncertainty requirements in cases when the composition of balancing areas differed between the two sets of data. Figure 4.13 shows the percent of intervals in which the market uncertainty requirements (with inconsistent balancing areas in the pass-group) were higher or lower than counterfactual uncertainty requirements with a consistent composition of balancing areas in the pass-group.[16] These results are shown separately for the following categories to highlight the impact of this inconsistency on uncertainty requirements.

- **Decreased requirements** indicate that market uncertainty requirements for the pass-group were lower as a result of inconsistent balancing areas in the pass-group.
- **Increased requirements** indicate that market uncertainty requirements for the pass-group were higher as a result of inconsistent balancing areas in the pass-group.
- **Unknown impact** indicates that there was an inconsistent composition of balancing areas in the pass-group but data was not available to calculate the impact.

Figure 4.13 Impact of pass-group inconsistency on uncertainty requirements (July 2023)



[16] This analysis accounts for any thresholds that capped or would have capped calculated uncertainty requirements.

Additional impacts of earlier resource sufficiency evaluation failures on market results

Each real-time market run will use the latest resource sufficiency evaluation results available to optimize resources and energy transfers in the WEIM accordingly. This includes future advisory intervals that can be impacted by earlier runs of the resource sufficiency evaluation. In particular, the hour-ahead market includes resources and transfers in the WEIM footprint with transfer limits potentially impacted from test failures from the first run of the resource sufficiency evaluation at 75 minutes prior to the evaluation hour.

5 Net load uncertainty in the resource sufficiency evaluation

Net load uncertainty is included in the requirement of the flexible ramp sufficiency test (flexibility test) to capture additional flexibility needs that may be required in the evaluation hour due to variation in either load, solar, or wind forecasts. This calculation was adjusted on February 1 using a method called *mosaic quantile regression*. This section summarizes how uncertainty is currently calculated, the results of the uncertainty calculation, and how it compares with actual error between forecasts used in the tests and in the real-time market.

Calculating net load uncertainty in the resource sufficiency evaluation

Histogram method

Uncertainty used in the resource sufficiency evaluation was previously calculated by selecting the 2.5th and 97.5th percentile of observations from a distribution of historical net load forecast errors. This is known as the *histogram method*. The historical error observations in the distribution were the difference between binding 5-minute market net load forecasts and corresponding advisory 15-minute market net load forecasts.[17] Prior to February 1, 2023, the weekday distributions used data for the same hour from the previous 40 weekdays while weekend distributions instead used same-hour observations from the previous 20 weekend days. The histogram approach did not factor in any current load, solar, or wind forecast information. Under this approach, uncertainty could have been set by historical outlier observations uncorrelated with current market conditions such as an extreme historical observation in which wind forecasts were significant while wind forecasts in the evaluation hour were minimal.

Mosaic quantile regression method

The calculation for net load uncertainty was adjusted on February 1, 2023 as part of flexible ramping enhancements. The uncertainty was adjusted to incorporate current load, solar, and wind forecast information using a method called *mosaic quantile regression*.

Regression is a statistical method used to study the relationship between two or more variables, such as the relationship between the load or renewable forecasts (independent variables) and uncertainty (dependent variable). Ordinary Least Squares is widely used to estimate the *mean* relationship between these variables (i.e. the average value of the dependent variable as a function of the independent variable). In contrast, quantile regression is a variation of regression that is useful when interested in the relationship between the independent variable(s) and different *percentiles* of the dependent variable. For example, the relationship between the load or renewable forecasts and the 97.5th percentile of uncertainty.

[17] In comparing the 15-minute observation to the three corresponding 5-minute observations, the minimum and maximum net load errors were used as a separate observation in the distribution.

The chosen regression method is a two-step procedure to forecast the lower and upper extremes of net load uncertainty that might materialize. The initial quantile regressions determine the relationship between the forecasts (load, solar, and wind) and the extremes of uncertainty (load, solar, and wind). In a simple linear regression, the relationship between the dependent variable Y and the independent variable X takes the basic form of $Y = bX$ where the outcome of the regression, b, explains how much Y changes for every one unit increase in X (e.g. If b is two, then y is predicted to be twice X). For calculating uncertainty as a function of the forecast, the quantile regressions are instead defined in the quadratic form ($Y = aX^2 + bX + c$). The initial regressions are shown below for upward net load uncertainty.[18]

Equation 1. Initial quantile regressions for upward net load uncertainty

$$
\begin{aligned}
Load\ uncertainty^{max} &= a_l^{97.5}(load)^2 + b_l^{97.5}(load) + c_l^{97.5} + \varepsilon \qquad (\tau = 0.975)\\
Solar\ uncertainty^{min} &= a_s^{2.5}(solar)^2 + b_s^{2.5}(solar) + c_s^{2.5} + \varepsilon \qquad (\tau = 0.025)\\
Wind\ uncertainty^{min} &= a_w^{2.5}(wind)^2 + b_w^{2.5}(wind) + c_w^{2.5} + \varepsilon \qquad (\tau = 0.025)
\end{aligned}
$$

Dependent variable: load, solar, and wind uncertainty — minimum or maximum difference between binding 5-minute market forecasts and advisory 15-minute market forecasts in each 15-minute market interval

Independent variable: advisory 15-minute market forecasts for load, solar, and wind in each interval

Error term (ε): variation in dependent variable that is not explained by independent variable

Quantile parameter (τ): determines the level of the quantile regression being estimated (high: 97.5 percentile, low: 2.5 percentile)

The uncertainty regressions use a distribution of historical forecast observations from the previous 180 days — separate for each balancing area, hour, and day-type (weekday or weekend/holiday). For the resource sufficiency evaluation, uncertainty in the distributions is the difference between binding 5-minute market forecasts and corresponding advisory 15-minute market forecasts.[19] The outcome of these regressions are the coefficients *a*, *b*, and *c*, that define the relationships between the forecasts and the extreme end of uncertainty that might materialize.[20] These coefficients can then be combined with the historical 15-minute forecast data to create a distribution of predicted values for load, solar, and wind uncertainty which is needed for the second step of the calculation. This is shown below for upward net load uncertainty.

[18] Equations 1 to 5 are for calculating *upward* net load uncertainty. *Downward* net load uncertainty is instead based on the lower end of load uncertainty, and upper end of solar and wind uncertainty that might materialize.

[19] In comparing the 15-minute observation to the three corresponding 5-minute observations, the maximum load errors and minimum wind and solar errors are used to calculate upward net load uncertainty. Or, minimum load errors and maximum wind and solar errors for downward net load uncertainty.

[20] The coefficient *c* is also known as the intercept. It shows the value of the dependent variable when all independent variables are equal to zero.

Equation 2. Predicted values for upward net load uncertainty

$$\hat{L}_Q^{97.5} = a_l^{97.5}(load)^2 + b_l^{97.5}(load) + c_l^{97.5}$$
$$\hat{S}_Q^{2.5} = a_s^{2.5}(solar)^2 + b_s^{2.5}(solar) + c_s^{2.5}$$
$$\widehat{W}_Q^{2.5} = a_W^{2.5}(wind)^2 + b_W^{2.5}(wind) + c_W^{2.5}$$

Predicted values: predicted 97.5th percentile of load uncertainty and 2.5th percentile of solar and wind uncertainty based on regression coefficients and historical distribution

***Regression coefficients*:** parameters "a", "b" and "c" that define the relationship between the forecasts and the extreme end of uncertainty that might materialize

The *mosaic* element of the regression combines the predicted forecasts above with the histogram method. For the histogram estimates, the 180-day distributions are again used to calculate the lower and upper ends of uncertainty, based on the 2.5th and 97.5th percentiles in the distribution. The combination of the predicted values and the histograms extremes in the mosaic variable are intended to capture the incremental weather effect of using predicted information relative to the histogram approach. Here, the calculation modifies the histogram net load by adding the predicted values and subtracting the histogram outcomes for each uncertainty type individually.[21] This is shown below for upward net load uncertainty:

Equation 3. Mosaic variable for upward net load uncertainty

$$mosaic^{97.5} = NL_H^{97.5} + \left(\left(\hat{L}_Q^{97.5} - L_H^{97.5}\right) - \left(\hat{S}_Q^{2.5} - S_H^{2.5}\right) - \left(\widehat{W}_Q^{2.5} - W_H^{2.5}\right)\right)$$

***Upward mosaic variable*:** intermediate variable for final regression

97.5th percentile of net load uncertainty from histogram

Predicted values: predicted load, solar, and wind uncertainty from initial quantile regressions (using historical distribution)

Load, solar, and wind uncertainty from histograms

Once the mosaic variable is calculated for each interval in the distribution, the software runs a final regression to predict net load uncertainty. Again, the quantile regression method looks for the extreme values of the data (at the 2.5th and 97.5th percentiles) such that the output reflects the upper and lower boundaries of the future uncertainty. Therefore, the predicted values obtained from the quantile regression models are expected to estimate the range in which net load uncertainty is likely to materialize. The final regression is shown below:

21 The mosaic variable can be thought of as the modified net load.

Equation 4. Mosaic regression for upward net load uncertainty

$$Net\ load\ uncertainty^{max} = a_m^{97.5}(mosaic^{97.5})^2 + b_m^{97.5}(mosaic^{97.5}) + c_m^{97.5} + \varepsilon \qquad (\tau = 0.975)$$

Dependent variable: net load uncertainty — maximum difference between binding 5-minute market forecasts and advisory 15-minute market forecasts in each 15-minute market interval

Independent variable: mosaic variable in each 15-minute market interval (from previous step)

Error term (ε): variation in dependent variable that is not explained by independent variable

Quantile parameter (τ): determines the level of the quantile regression being estimated (high: 97.5 percentile)

Once all of the regressions are complete, the regression output coefficients can be combined with current forecast information to calculate uncertainty for each interval. For the flexibility test, this forecast information is the same load, solar, and wind forecasts which are considered in the resource sufficiency evaluation for calculating ramping capacity and test requirements. The latest forecasts at the time of the second pass of the resource sufficiency evaluation at 55 minutes prior to the evaluation hour are held constant for the final test at 40 minutes prior to the hour. The final equations for combining the current forecast information with the regression coefficients and histogram extremes to calculate upward uncertainty for each interval are shown below.

Equation 5. Calculation of upward uncertainty from current forecast information

$$\hat{L}_{current}^{97.5} = a_l^{97.5}(load_{current})^2 + b_l^{97.5}(load_{current}) + c_l^{97.5}$$
$$\hat{S}_{current}^{2.5} = a_s^{2.5}(solar_{current})^2 + b_s^{2.5}(solar_{current}) + c_s^{2.5}$$
$$\hat{W}_{current}^{2.5} = a_w^{2.5}(wind_{current})^2 + b_w^{2.5}(wind_{current}) + c_w^{2.5}$$

$$mosaic_{current}^{97.5} = NL_H^{97.5} + \left(\left(\hat{L}_{current}^{97.5} - L_H^{97.5}\right) - \left(\hat{S}_{current}^{2.5} - S_H^{2.5}\right) - \left(\hat{W}_{current}^{2.5} - W_H^{2.5}\right)\right)$$

$$Net\ load\ uncertainty_{current}^{97.5} = a_m^{97.5}\left(mosaic_{current}^{97.5}\right)^2 + b_m^{97.5}\left(mosaic_{current}^{97.5}\right) + c_m^{97.5}$$

The performance of the mosaic quantile regression method depends on whether there is a meaningful relationship between net load uncertainty and the mosaic variables created from historical and predicted values. DMM is currently in the process of evaluating whether there is a strong relationship between these variables.

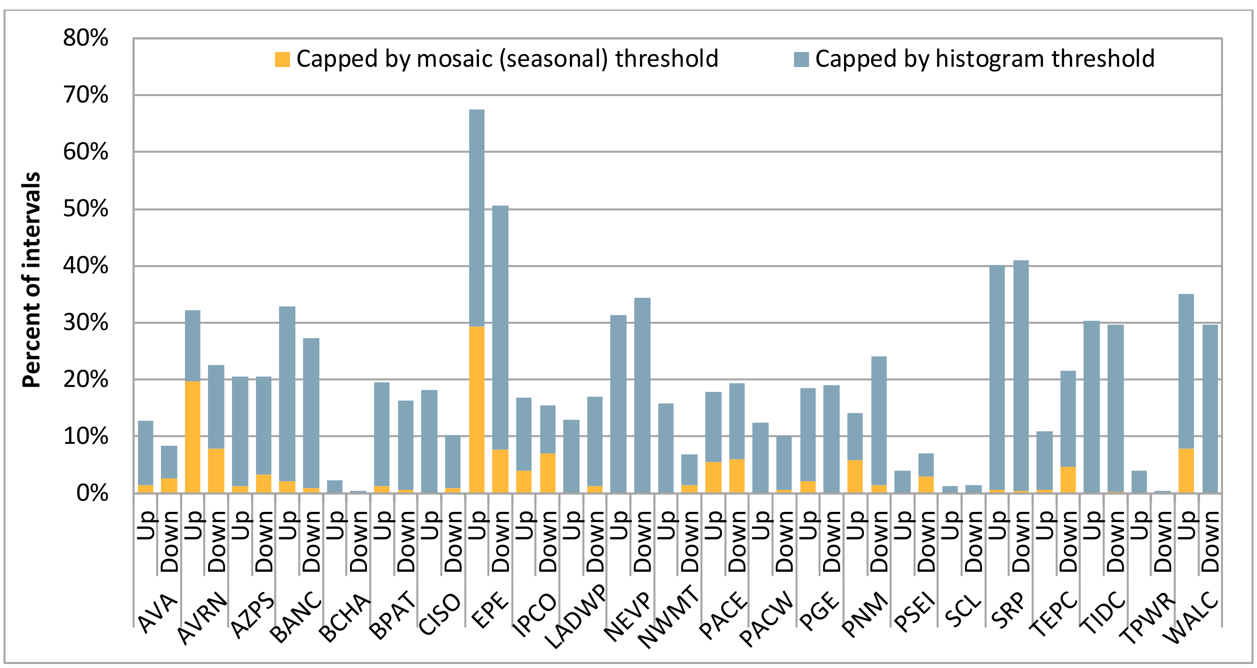
Thresholds for capping uncertainty

Uncertainty calculated from the quantile regressions are capped by the lesser of two thresholds. The thresholds are designed to help prevent extreme outlier results from impacting the final uncertainty. The *histogram* threshold is pulled for each hour from the 1st and 99th percentile of net load error observations from the previous 180 days.[22] The *mosaic* (or seasonal) threshold is updated each quarter and is calculated based on the 1st and 99th percentile using the quantile regression method and observations over the previous 90 days. Here, each hour is calculated separately and the greatest upward and downward uncertainty across all hours sets the mosaic threshold for each hour of the same direction.

Figure 5.1 shows the percent of test intervals in which the upward or downward uncertainty calculated by the quantile regression was capped by either the mosaic or histogram threshold during the month. During July, the calculated uncertainty for El Paso Electric uncertainty was capped by either of the thresholds in around 67 percent of intervals for upward uncertainty and 51 percent of intervals for downward uncertainty. For most of the balancing areas, the histogram threshold capped the calculated uncertainty more frequently compared to the mosaic threshold.

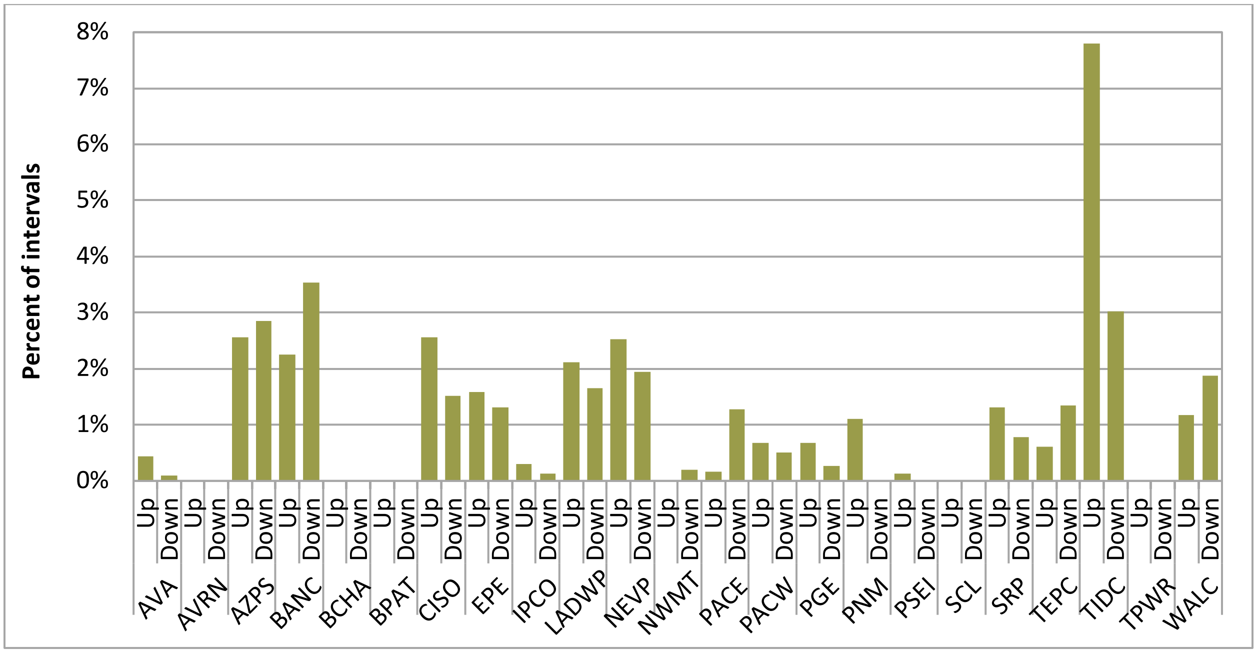
A threshold is also in place that sets the *floor* for uncertainty at 0.1 MW in both directions. The upward and downward uncertainty is therefore set near zero when the uncertainty calculated from the quantile regression would be negative. Figure 5.2 shows the percent of test intervals in which the quantile regression uncertainty was set near zero by this threshold during the month. In particular, the floor set the calculated upward uncertainty for Turlock Irrigation District near zero during almost 8 percent of intervals.

Figure 5.1 Quantile regression uncertainty capped by mosaic or histogram thresholds (July 2023)



[22] The histogram threshold is updated every day. The distributions are separate for each hour and day type (weekday or weekend/holiday).

Figure 5.2 Quantile regression uncertainty set near zero by mosaic threshold (July 2023)



Using uncertainty from the flexible ramping product in the resource sufficiency evaluation

The calculation of uncertainty in the flexibility test continues to be measured similarly to the 15-minute market flexible ramping product — based on the difference between binding 5-minute market forecasts and corresponding advisory 15-minute market forecasts. The quantile regression uses the historical sample of 5-minute and 15-minute market observations to create hourly coefficients that define the relationship between the forecasts and uncertainty. The resource sufficiency evaluation and flexible ramping product uncertainty calculations for a single balancing area use the same hourly coefficients, but are combined with the current forecast information for each time horizon.[23]

The calculated uncertainty is based on the 2.5th and 97.5th percentile for downward and upward uncertainty, respectively. The 95 percent confidence interval for the uncertainty requirement in the flexible ramping product was designed to capture the upper end of uncertainty needs, such that it could be optimally relaxed based on the trade-off between the cost of procuring additional flexible ramping capacity and the expected cost of a power balance constraint relaxation. In the resource sufficiency evaluation, this trade-off is not considered, and the upper end of uncertainty is instead required in full to pass both tests. DMM has asked the CAISO and stakeholders to consider whether the 95 percent confidence interval, or another, is most appropriate for the tests.[24]

---

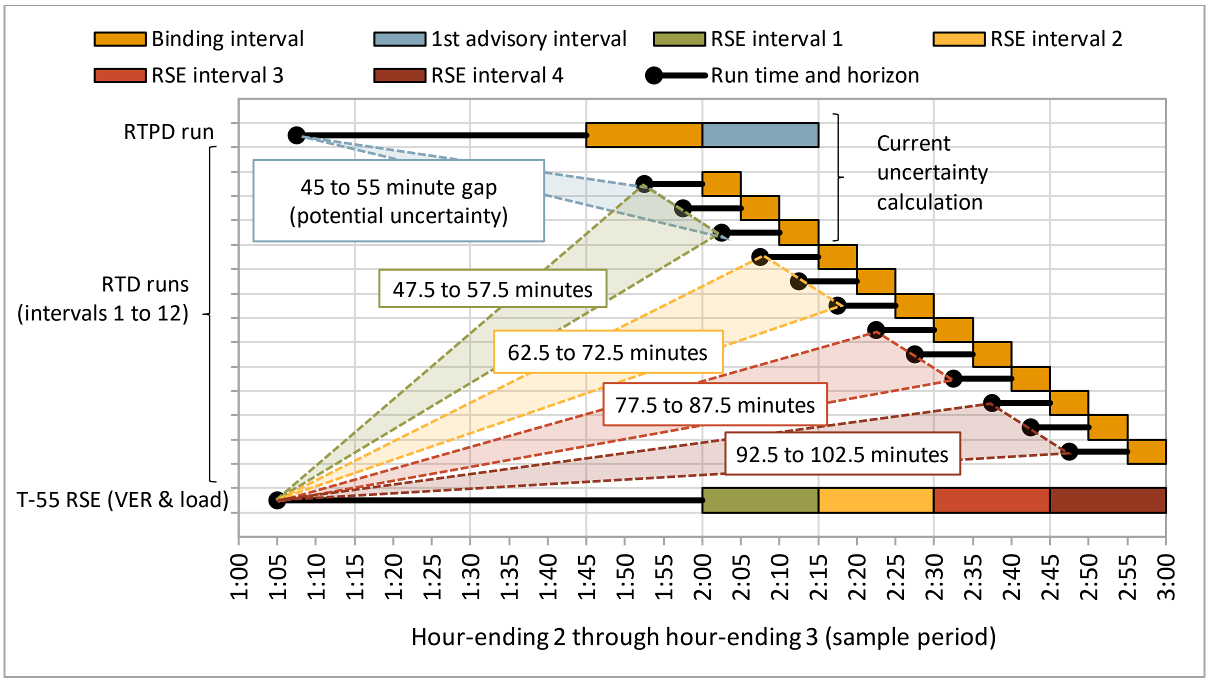
[23] A balancing-area-specific flexible ramping product uncertainty requirement will be enforced for any balancing area that failed the resource sufficiency evaluation.

[24] Department of Market Monitoring, *Comments on EIM Resource Sufficiency Evaluation Enhancements Issue Paper*, September 8, 2021: http://www.caiso.com/Documents/DMM-Comments-on-EIM-Resource-Sufficiency-Evaluation-Enhancements-Issue-Paper-Sep-8-2021.pdf

Further, the resource sufficiency evaluation occurs in a different timeframe than the 15-minute market. Figure 5.3 illustrates the current uncertainty calculation — based on net load error between an advisory 15-minute market interval and corresponding binding 5-minute market intervals — as well as how it compares with the timeframe of the resource sufficiency evaluation. The current uncertainty calculation captures 45 to 55 minutes of potential uncertainty from the 15-minute market run to three corresponding 5-minute market runs. In contrast, when comparing the VER and load forecast values used in each interval of the resource sufficiency evaluation to corresponding 5-minute intervals, there exists a larger gap for uncertainty to materialize.[25]

In comparing the first 15-minute test interval to corresponding 5-minute market intervals, the timeframe and potential for net load uncertainty is similar to the timeframe of the 15-minute market flexible ramping product uncertainty calculation. In the later test intervals, the gap between the predicted forecasts at the time of the resource sufficiency evaluation and the real-time forecasts widens, reaching above 100 minutes.

Figure 5.3 Comparison of current uncertainty calculation to the timeframe of the RSE



Results of quantile regression uncertainty in the resource sufficiency evaluation

Figure 5.4 summarizes the histogram uncertainty (pulled from the 2.5th and 97.5th percentile of observations in the hour from the previous 180 days) and the final uncertainty from the mosaic quantile regression during the month for CAISO. The green and blue lines show the *average* upward and downward uncertainty from each method while the areas around the lines show the minimum and

[25] The figure shows the resource sufficiency evaluation run time at 55 minutes prior to the hour. While the financially binding test is run at 40 minutes prior to the hour, the VER and load forecasts used in the final test are pulled from the advisory test performed at T-55.

maximum amount over the month. The dashed red and yellow lines in Figure 5.4 show the average histogram and mosaic thresholds, respectively, during the month.

Figure 5.5 summarizes actual error between net load forecasts used in the resource sufficiency evaluation and those used in the 5-minute market for CAISO during the month. The distributions in each interval were created from the difference between the 5-minute market net load and *net load in the corresponding test interval*. Here, a higher net load error reflects higher load (or lower renewables) in real-time, relative to the tests.

For comparison, the blue lines in Figure 5.5 show the average upward and downward uncertainty used in the tests during the same period (per the quantile regression output). Again, the blue areas around the lines show the minimum and maximum amounts for each hour. This metric therefore highlights net load error from the time horizon of the resource sufficiency evaluation and how well it fits within the current construct of uncertainty.

Figures covering the same information for all WEIM entities are provided further below. Overall, uncertainty calculated from the quantile regression approach were often comparable to those calculated with the histogram approach, though with the quantile regression approach tending to be lower across most hours and balancing areas.

Figure 5.4 CAISO resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

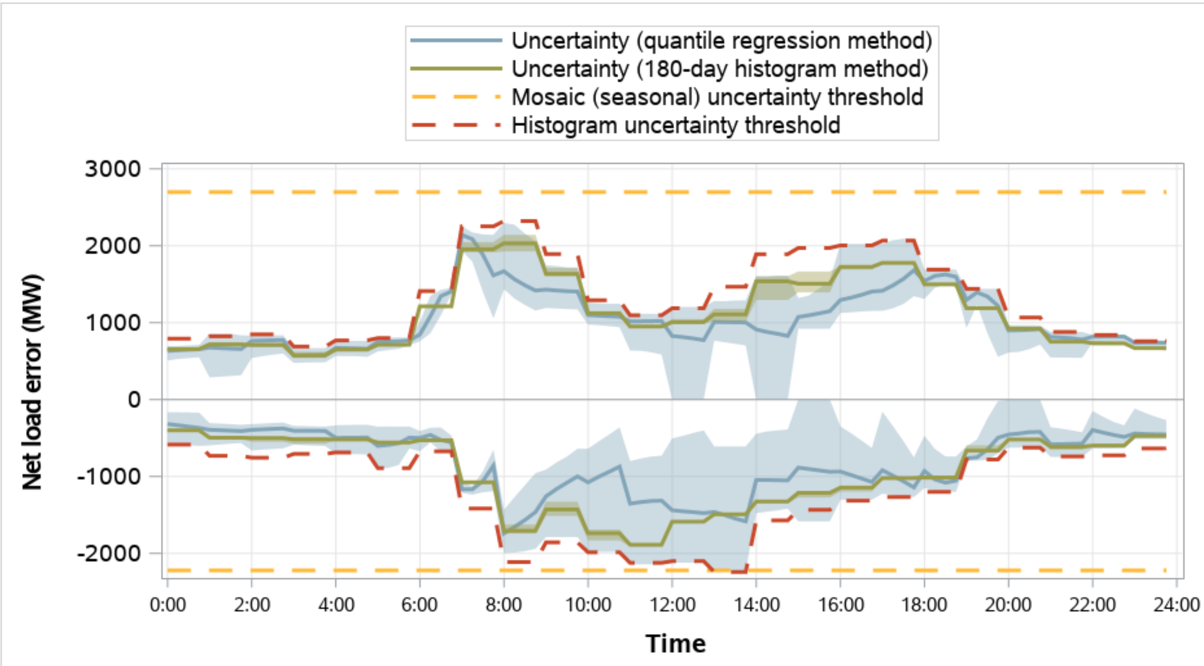


Figure 5.5 CAISO distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

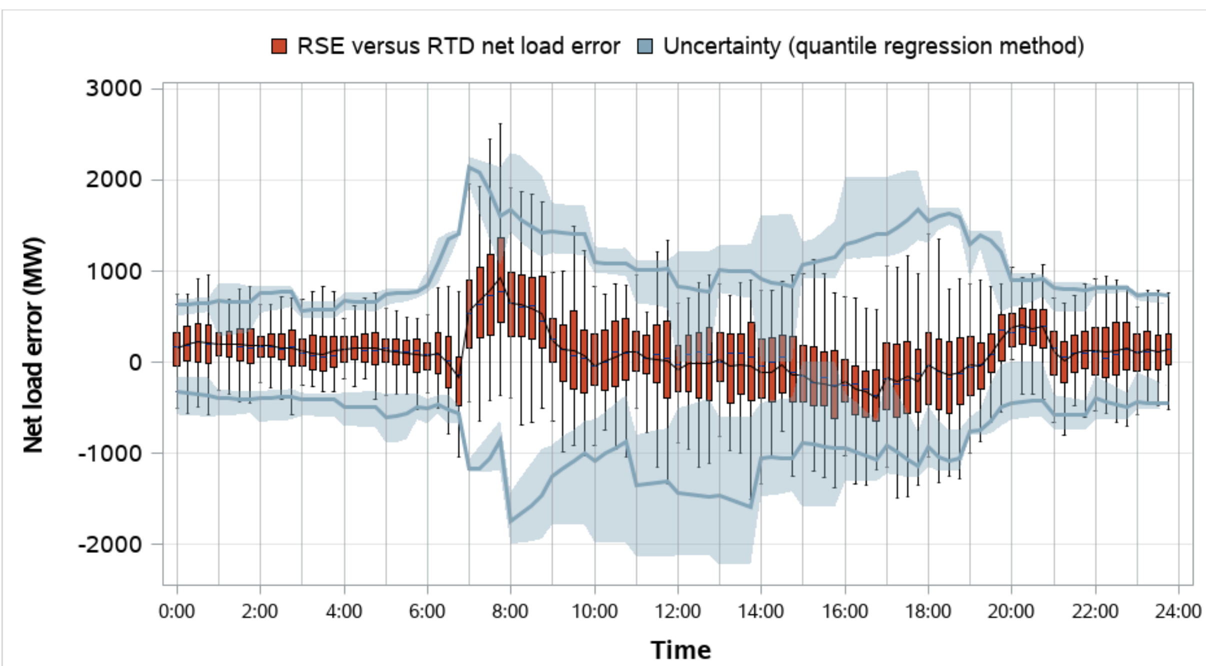


Figure 5.6 Arizona Public Service resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

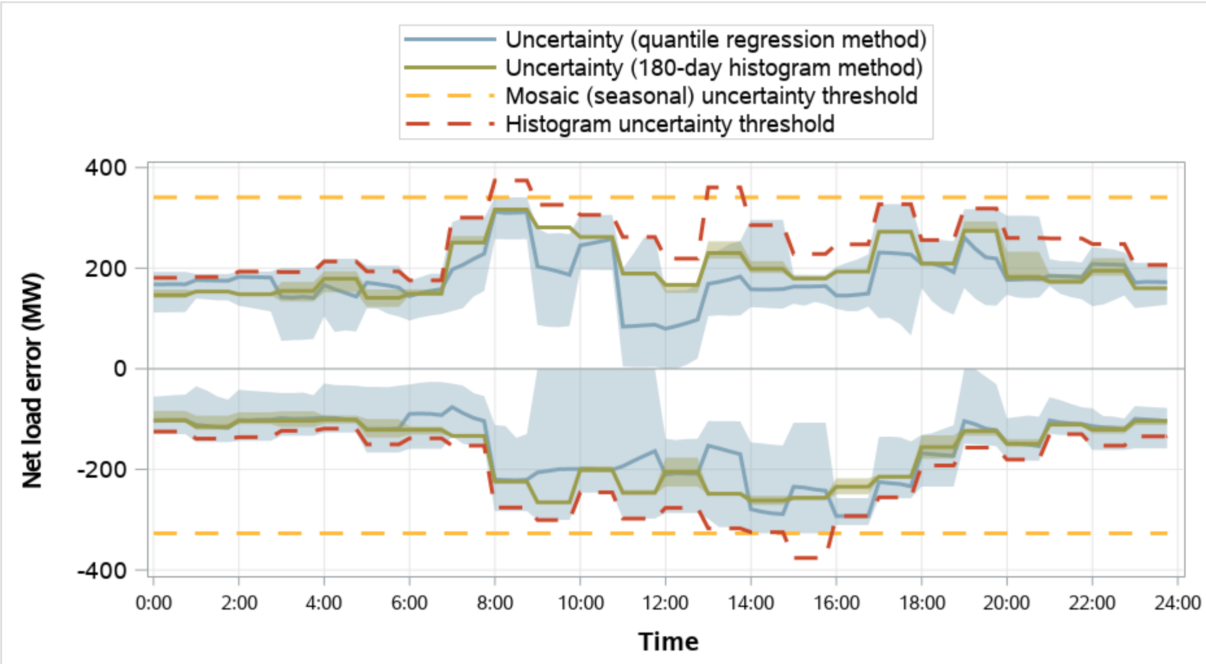


Figure 5.7 Arizona Public Service distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

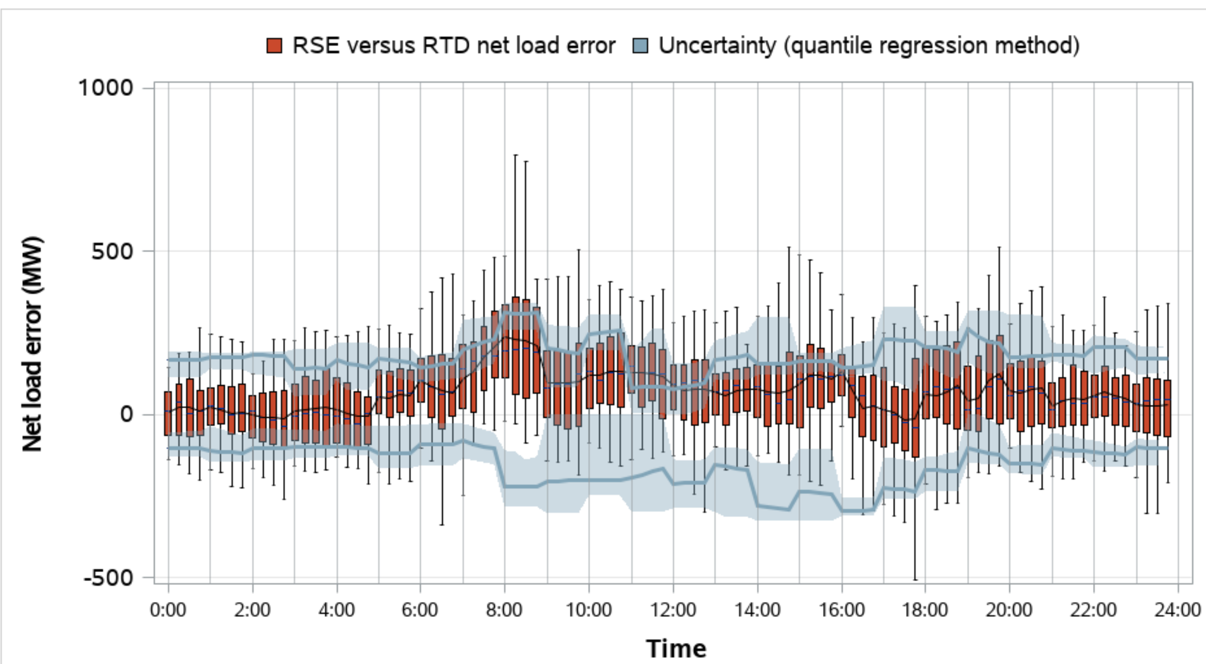


Figure 5.8 Avangrid resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

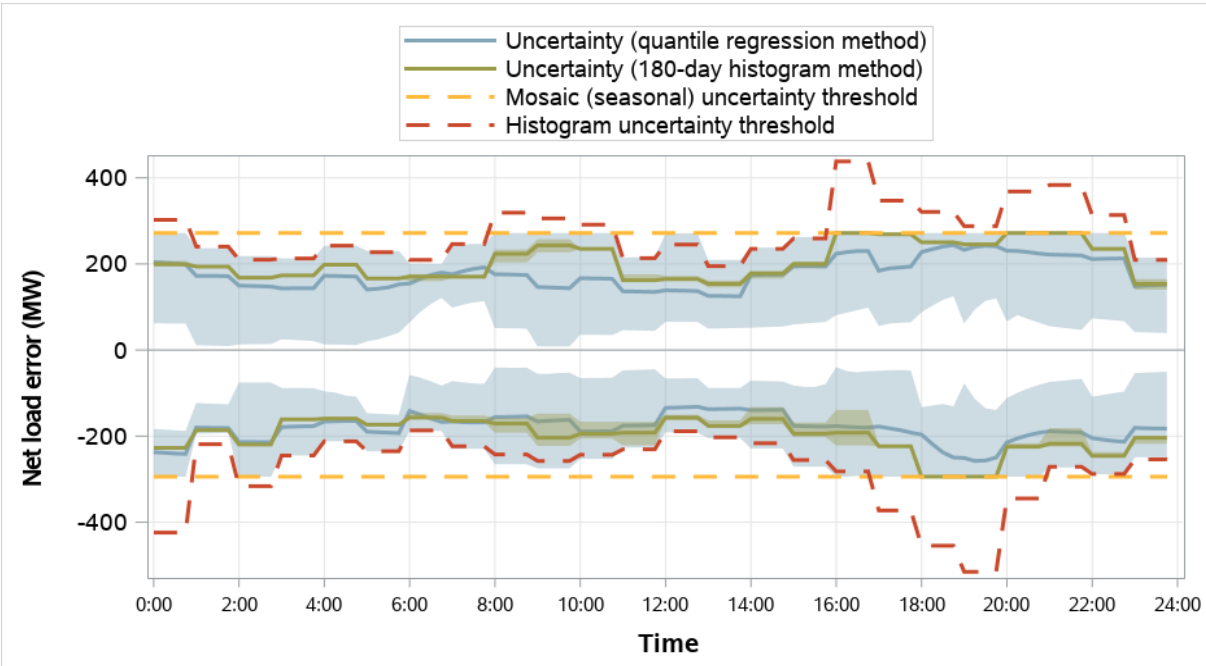


Figure 5.9 Avangrid distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

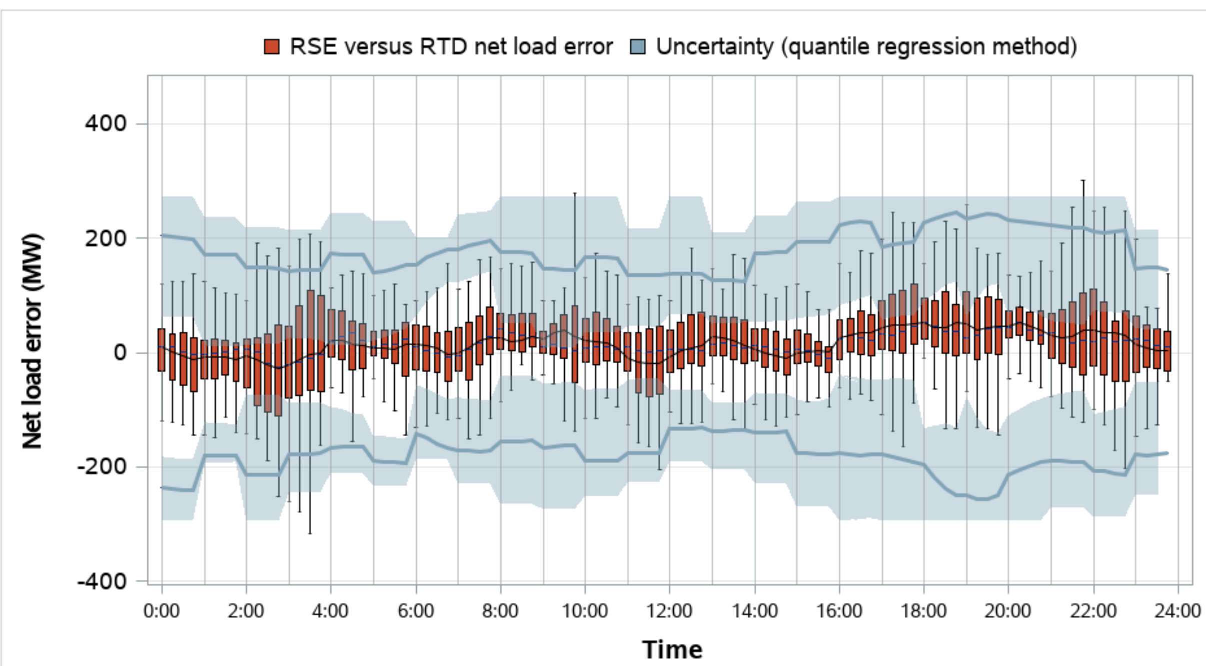


Figure 5.10 Avista resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

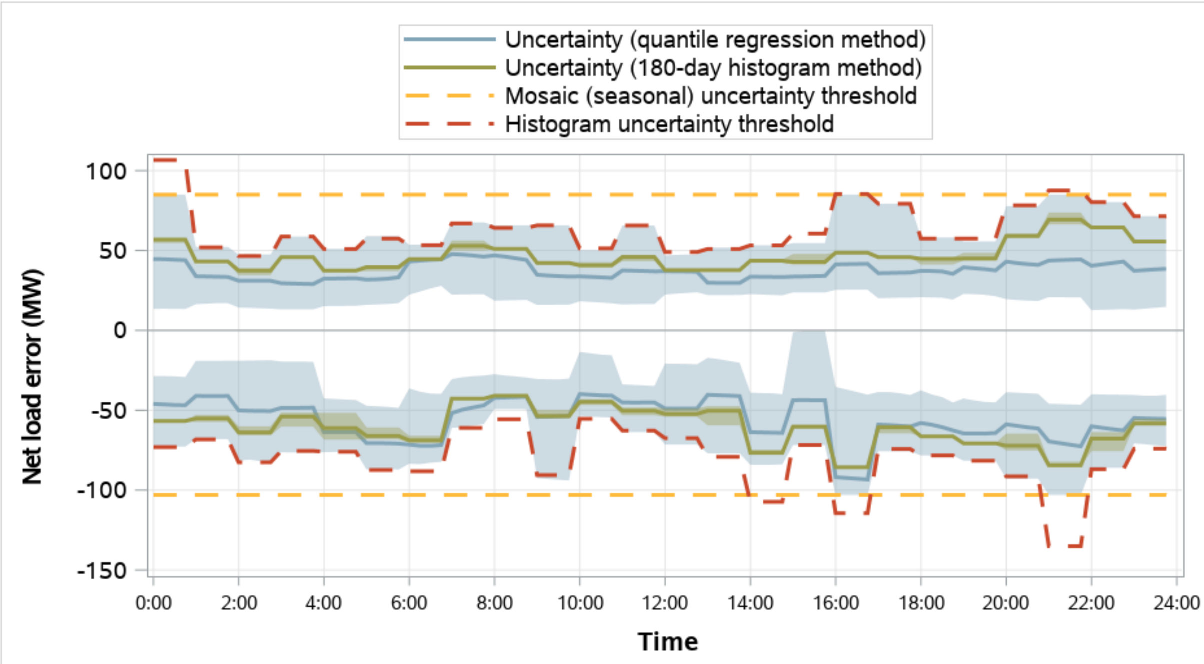


Figure 5.11 Avista distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

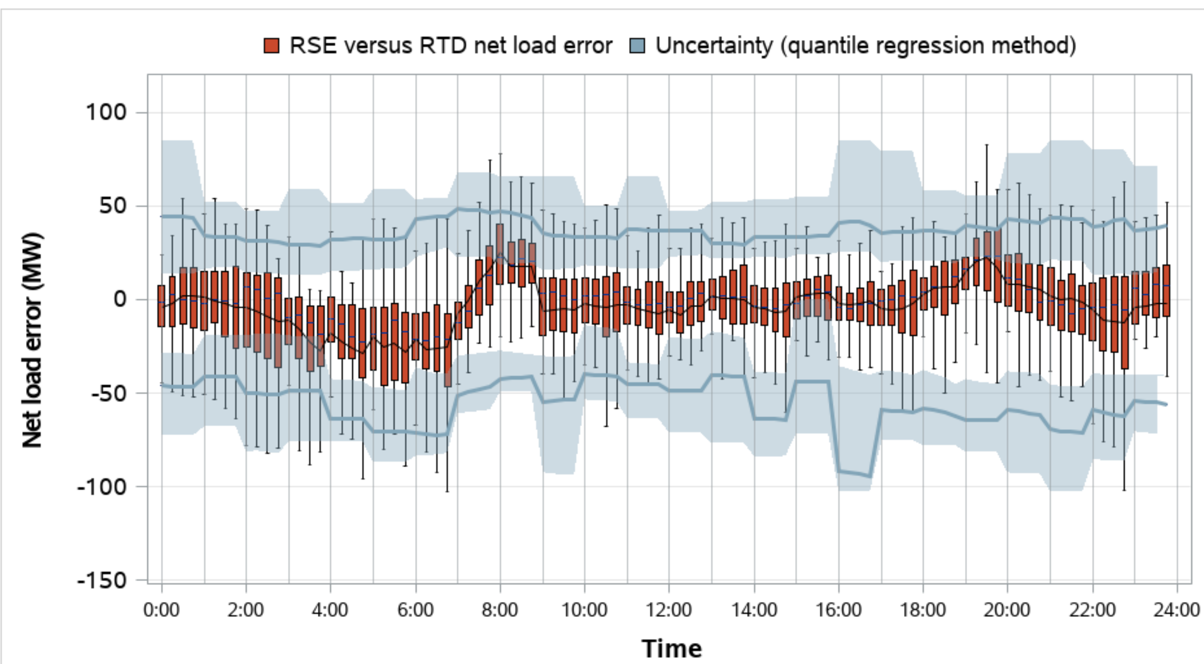


Figure 5.12 BANC resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

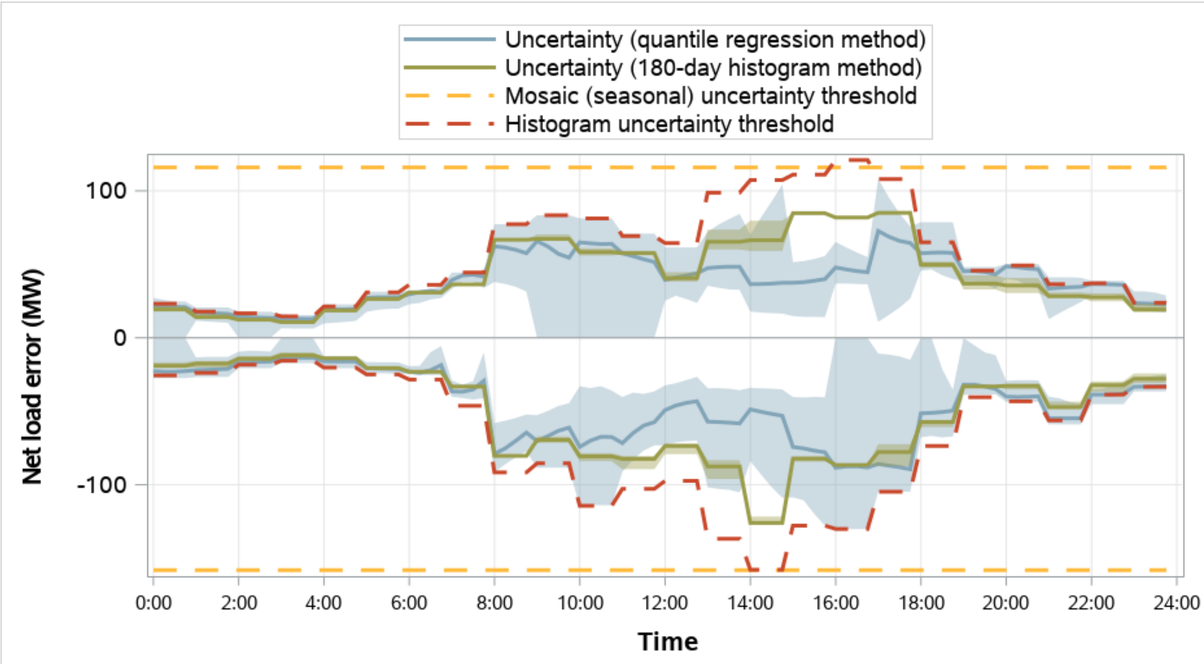


Figure 5.13 BANC distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

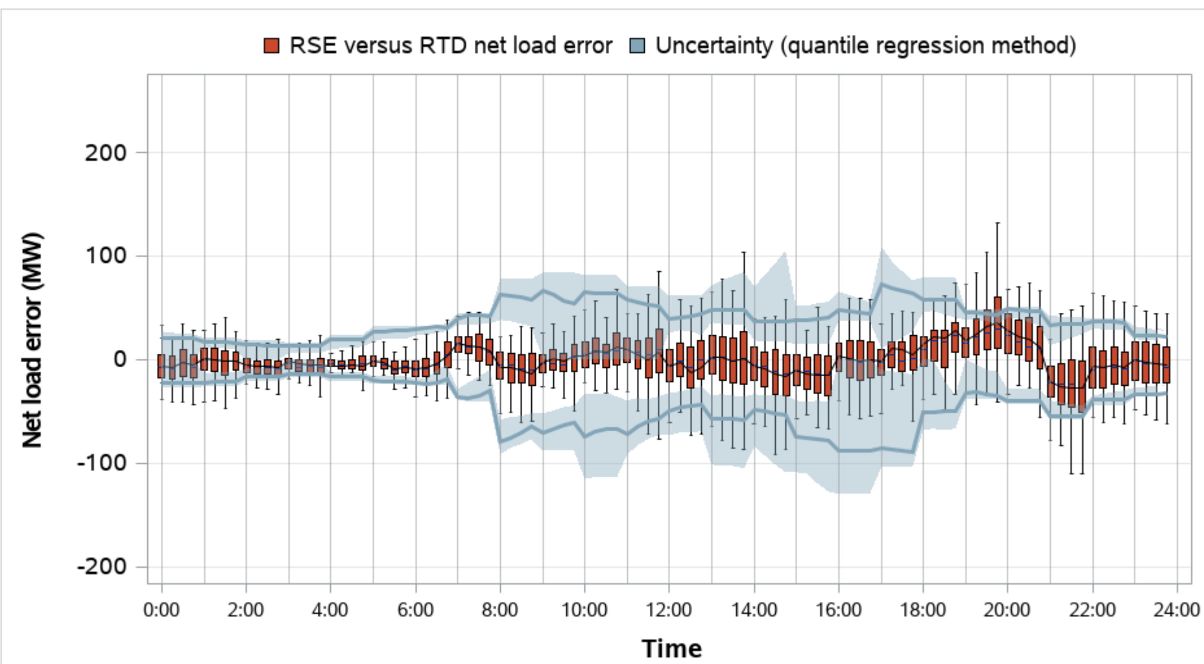


Figure 5.14 BPA resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

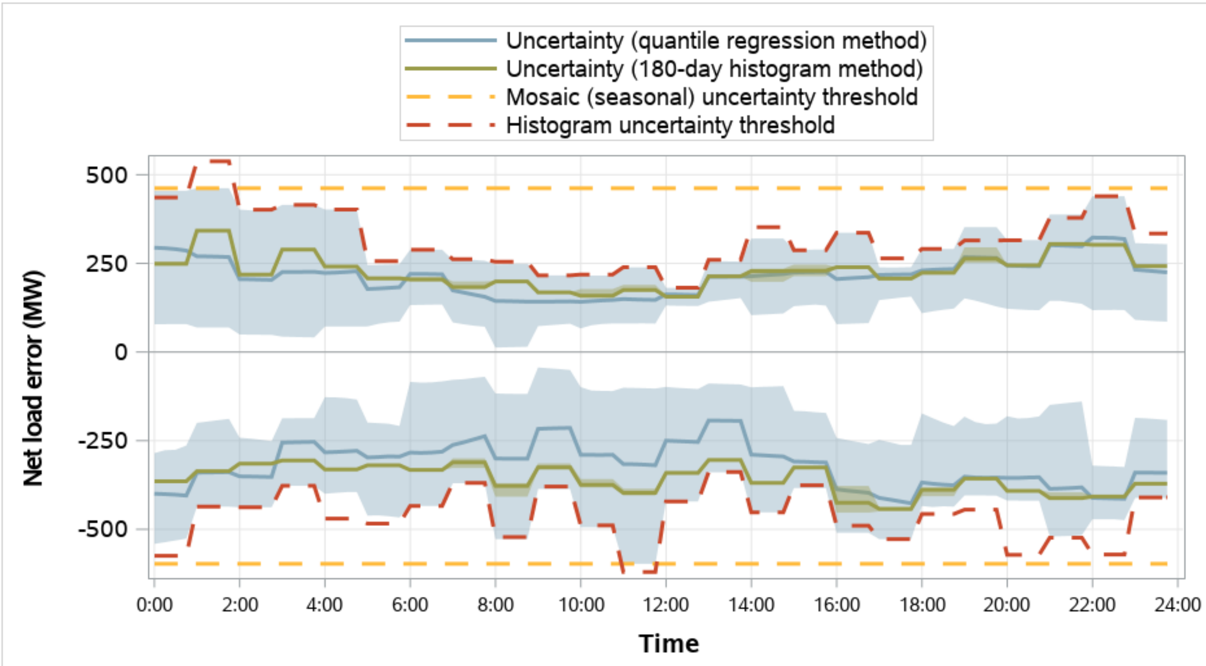


Figure 5.15 BPA distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

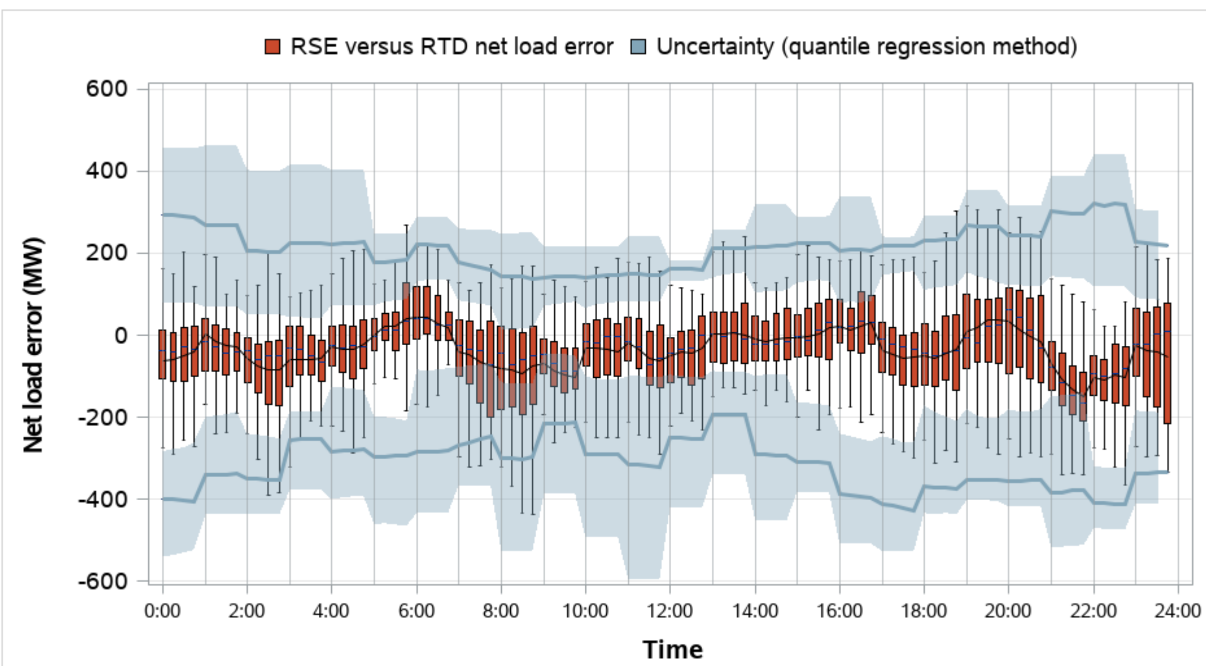


Figure 5.16 El Paso Electric distribution resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

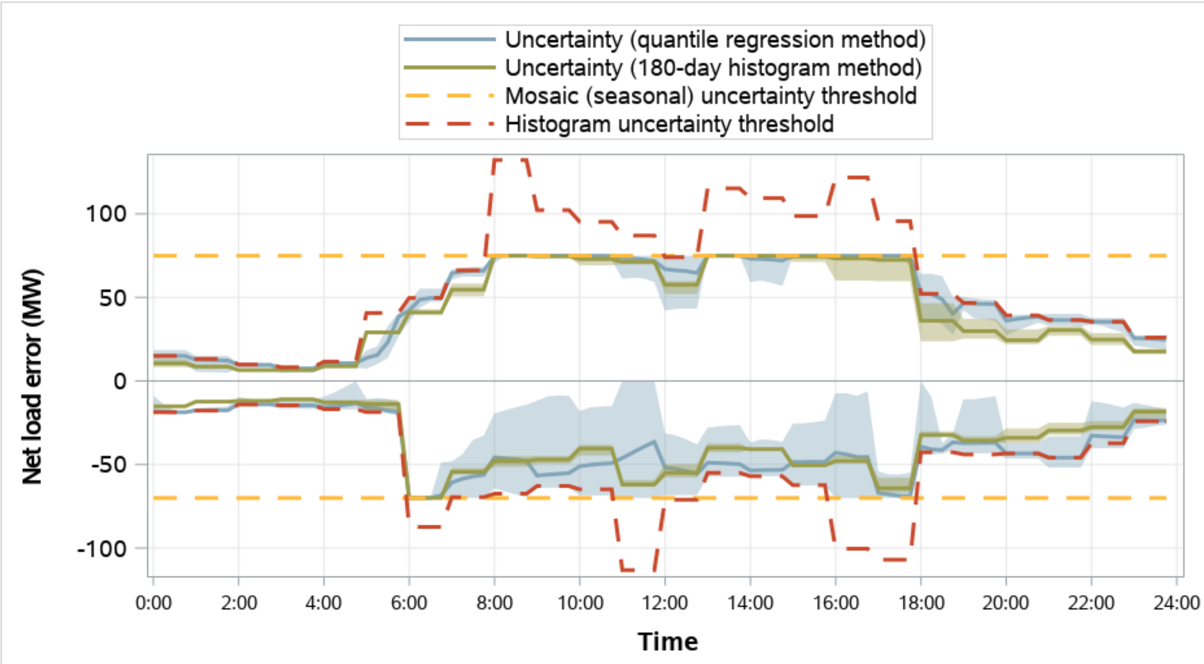


Figure 5.17 El Paso Electric distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

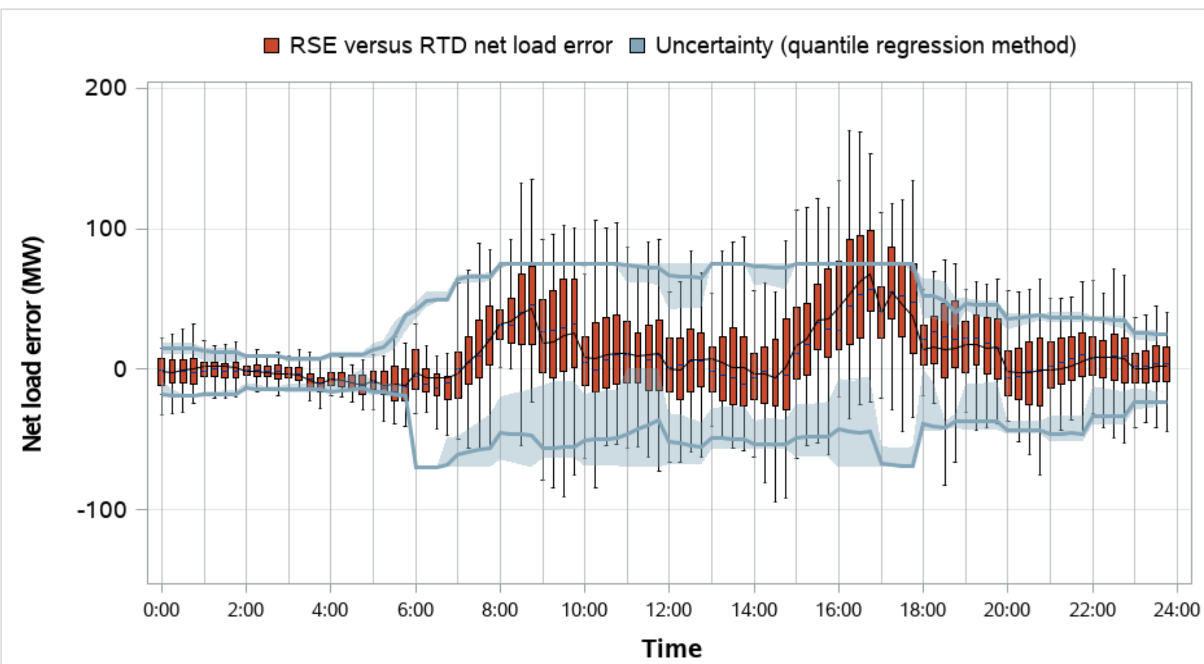


Figure 5.18 Idaho Power distribution resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

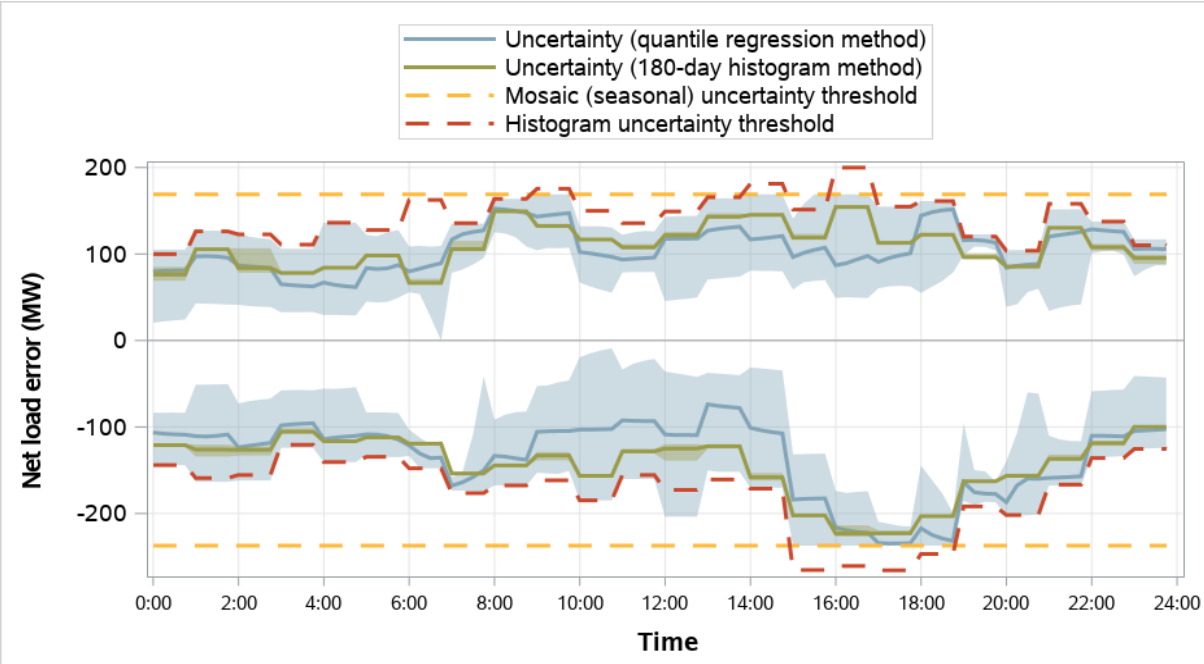


Figure 5.19 Idaho Power distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

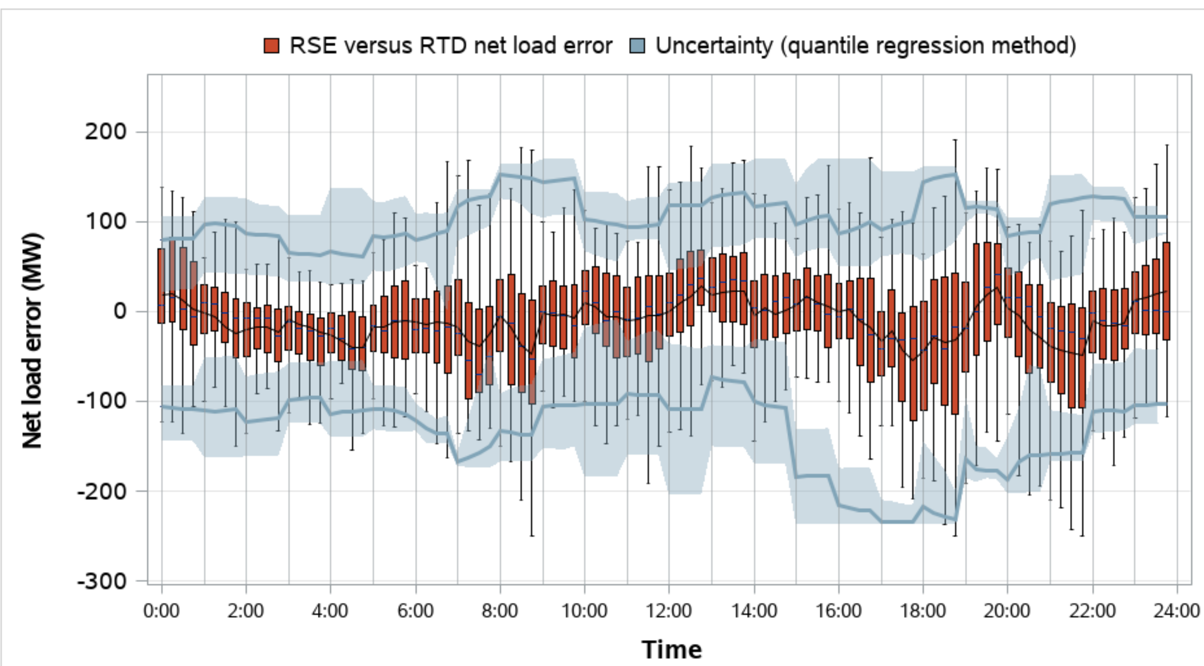


Figure 5.20 LADWP resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

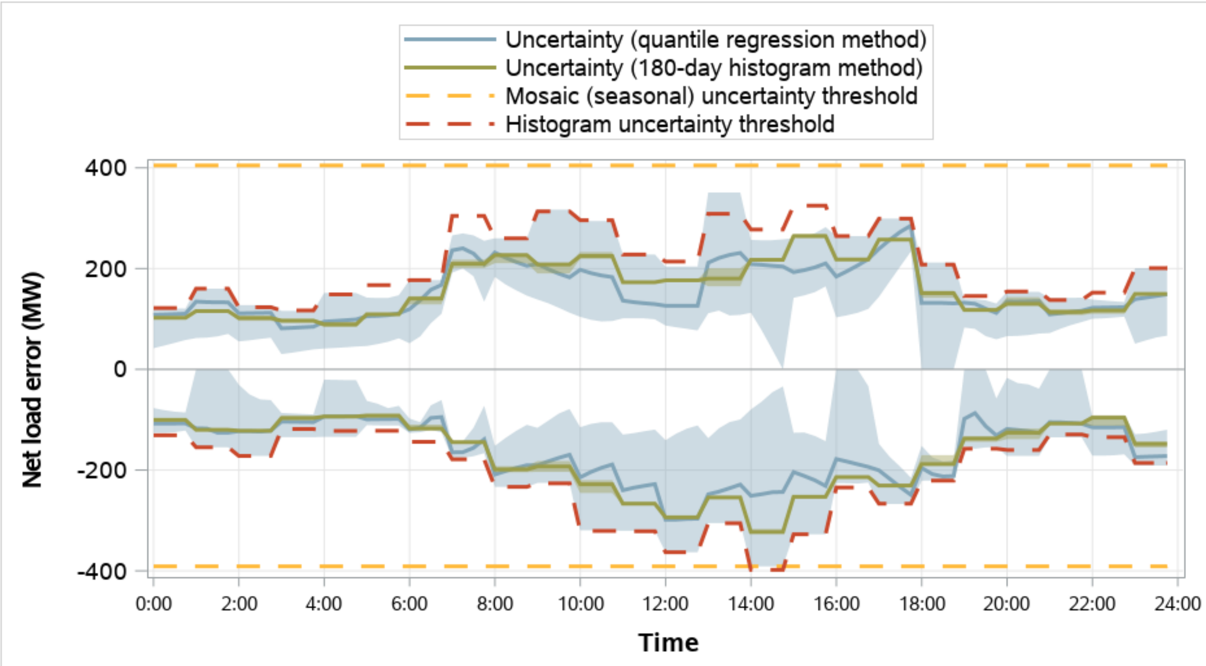


Figure 5.21 LADWP distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

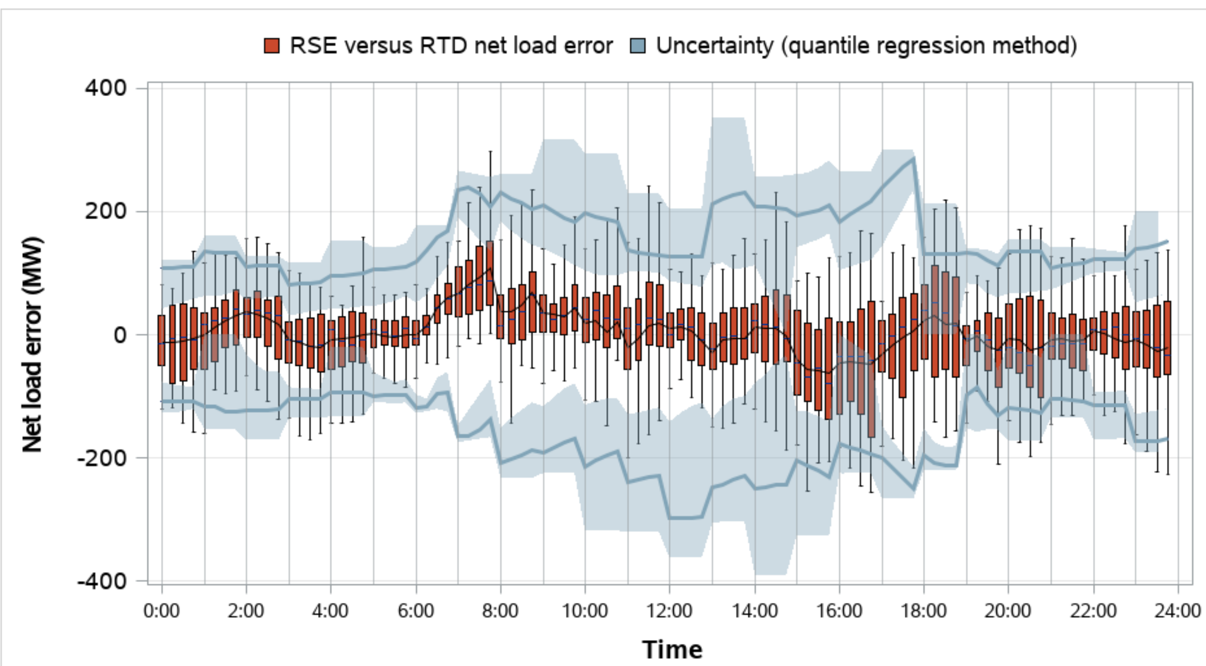


Figure 5.22 NorthWestern Energy average uncertainty by component (weekdays, July 2023)

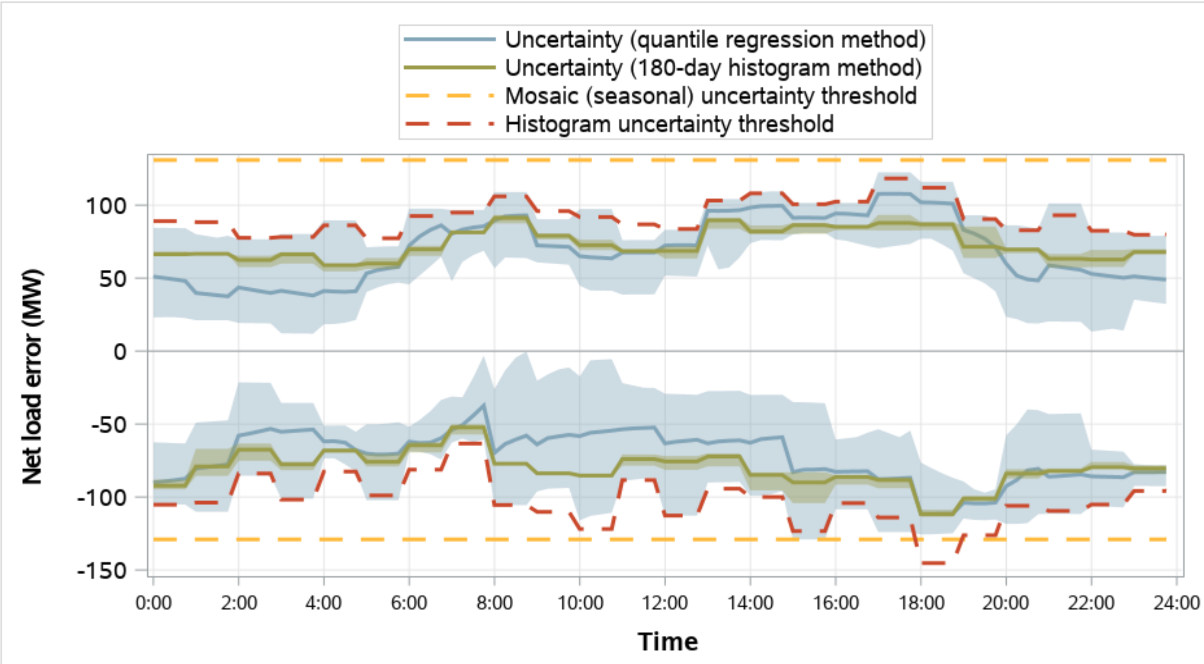


Figure 5.23 NorthWestern Energy distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

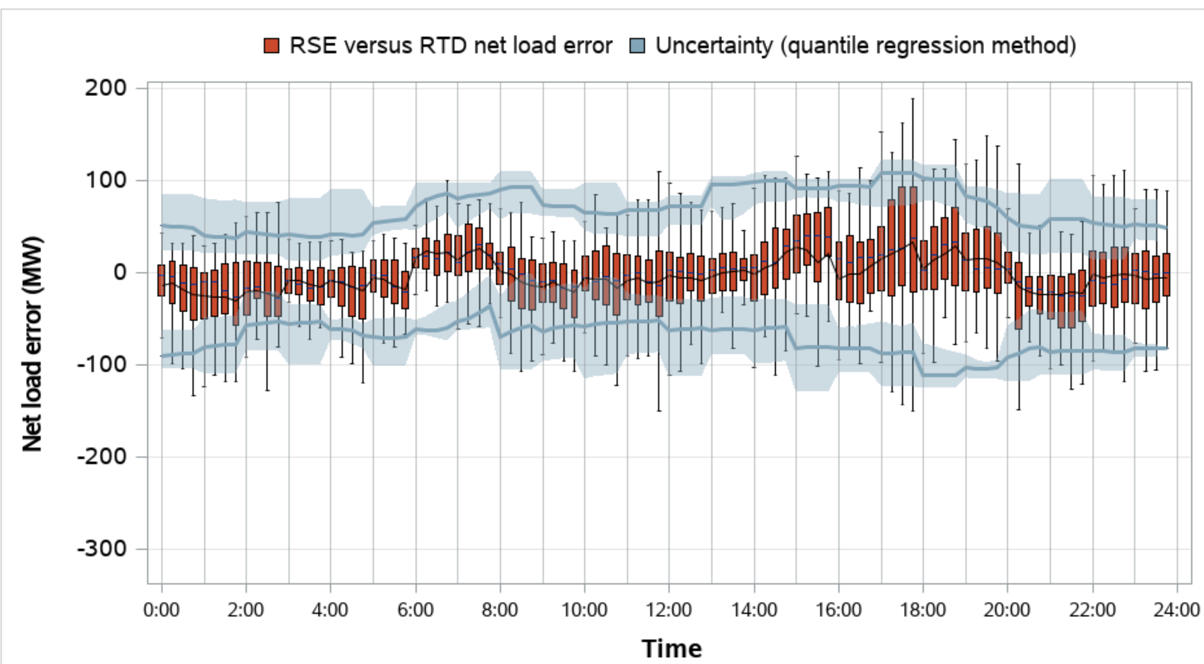


Figure 5.24 NV Energy resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

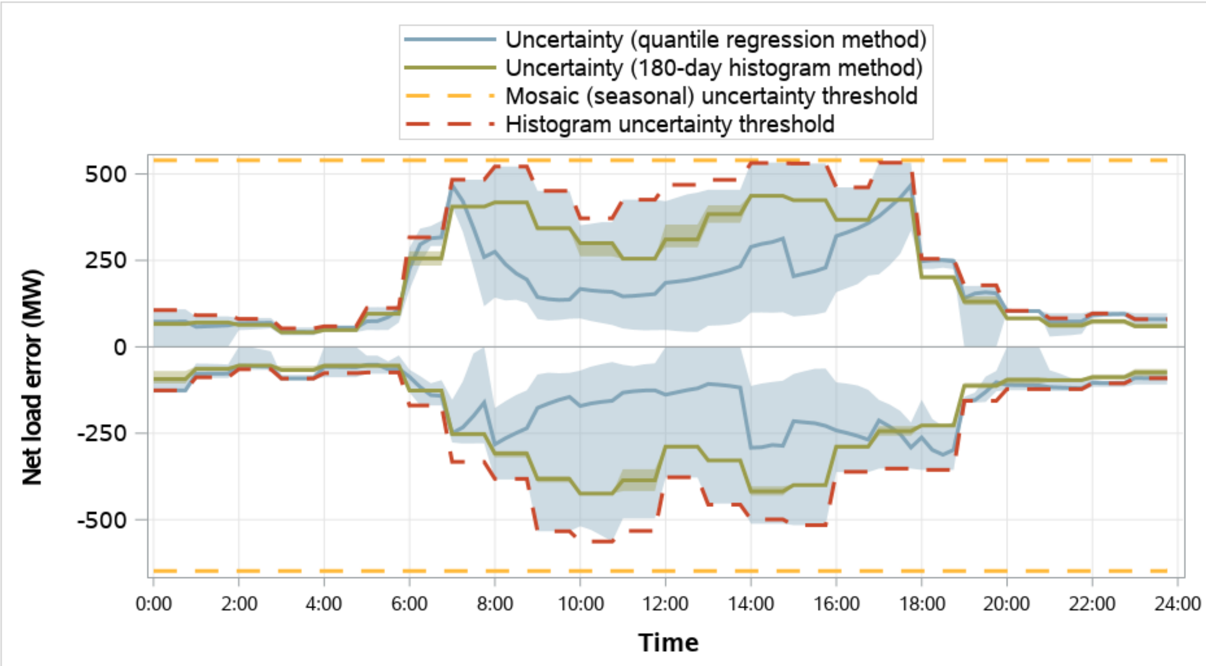


Figure 5.25 NV Energy distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

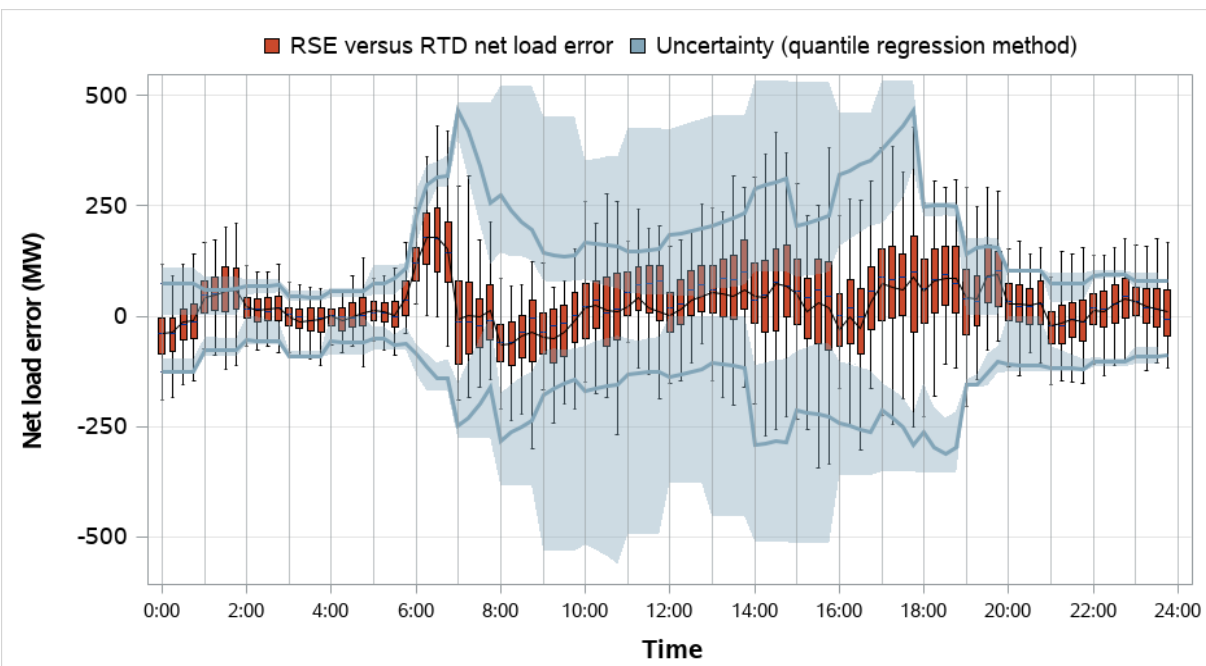


Figure 5.26 PacifiCorp East resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

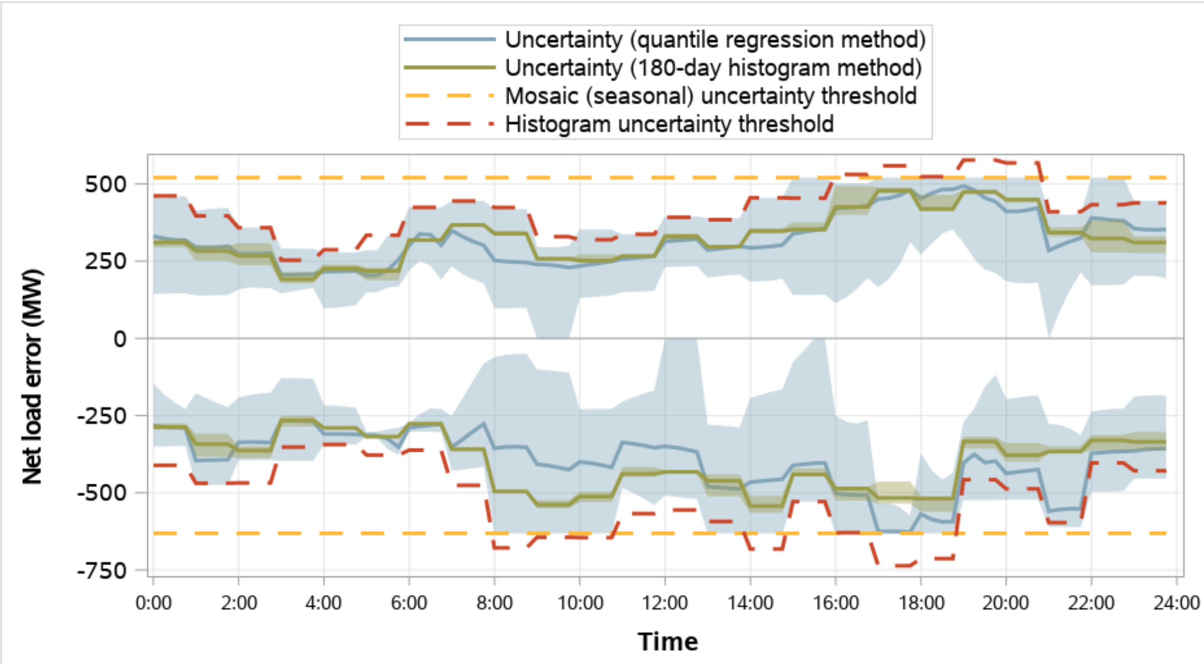


Figure 5.27 PacifiCorp East distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

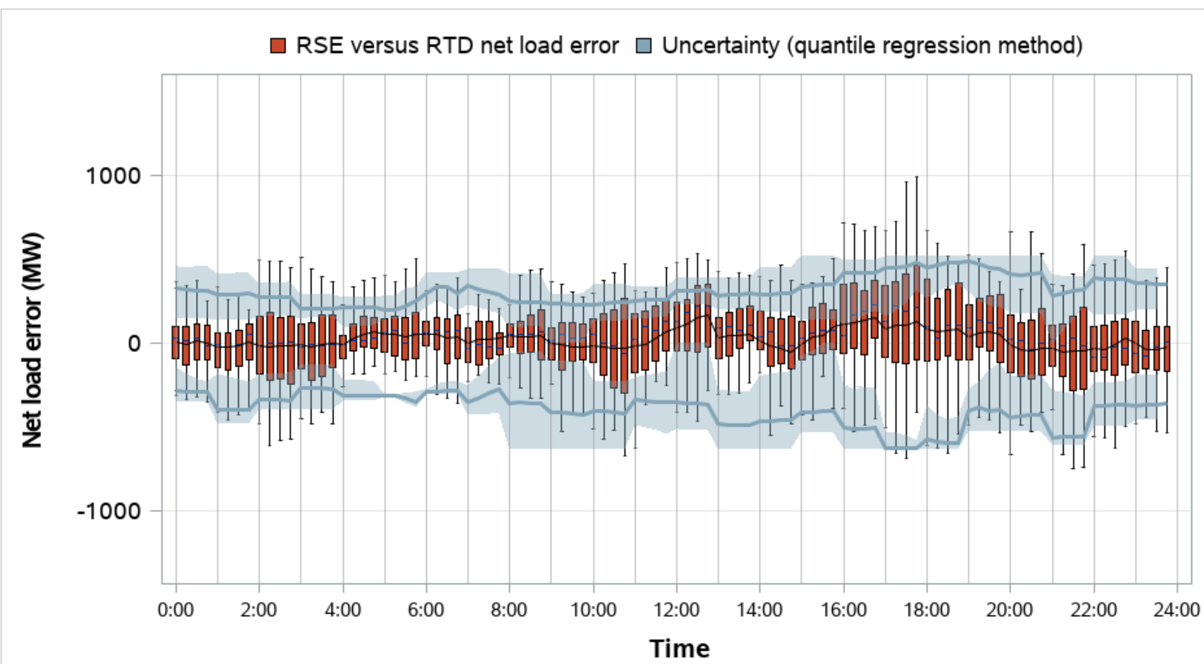


Figure 5.28 PacifiCorp West resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

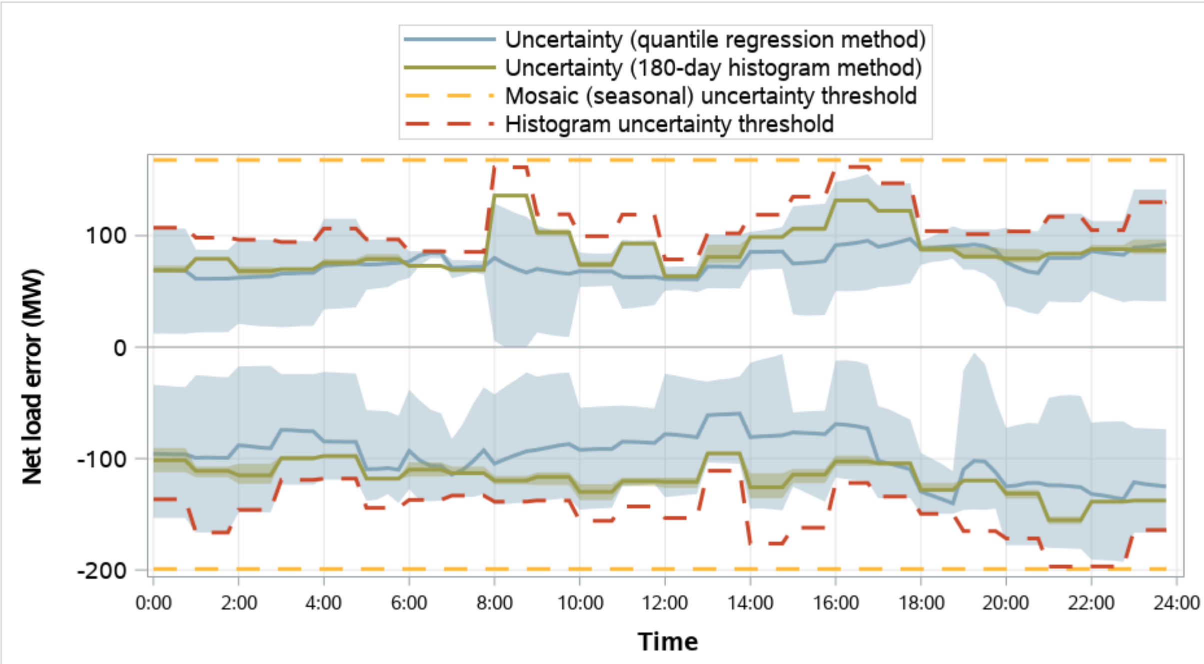


Figure 5.29 PacifiCorp West distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

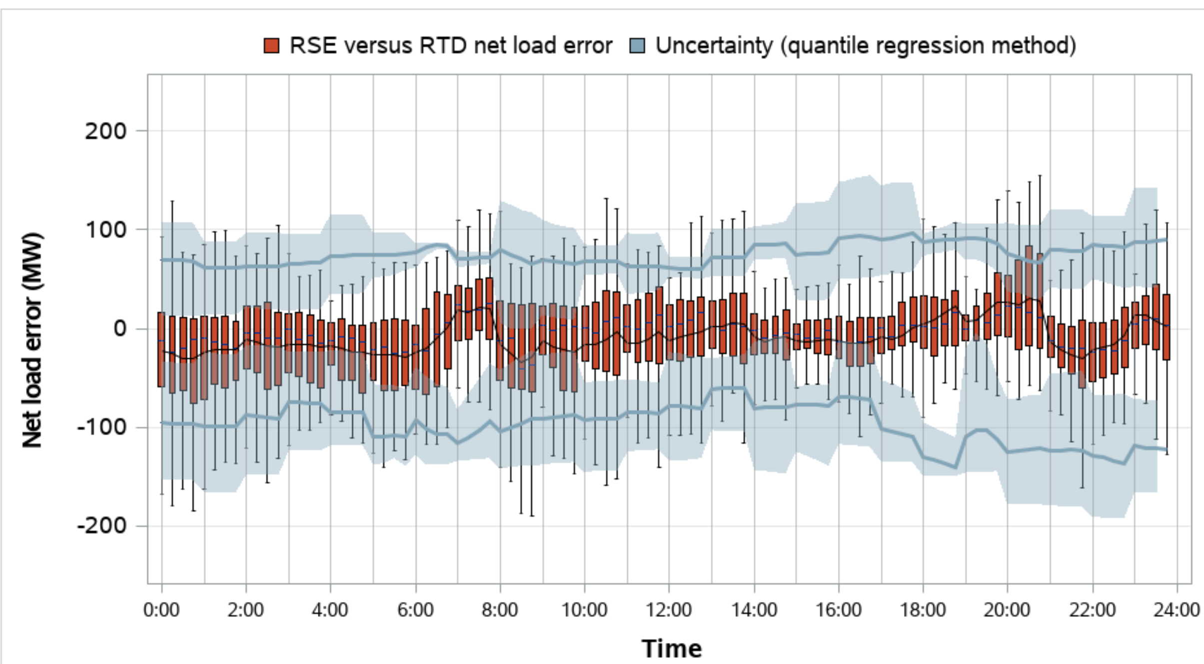


Figure 5.30 Portland General Electric resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

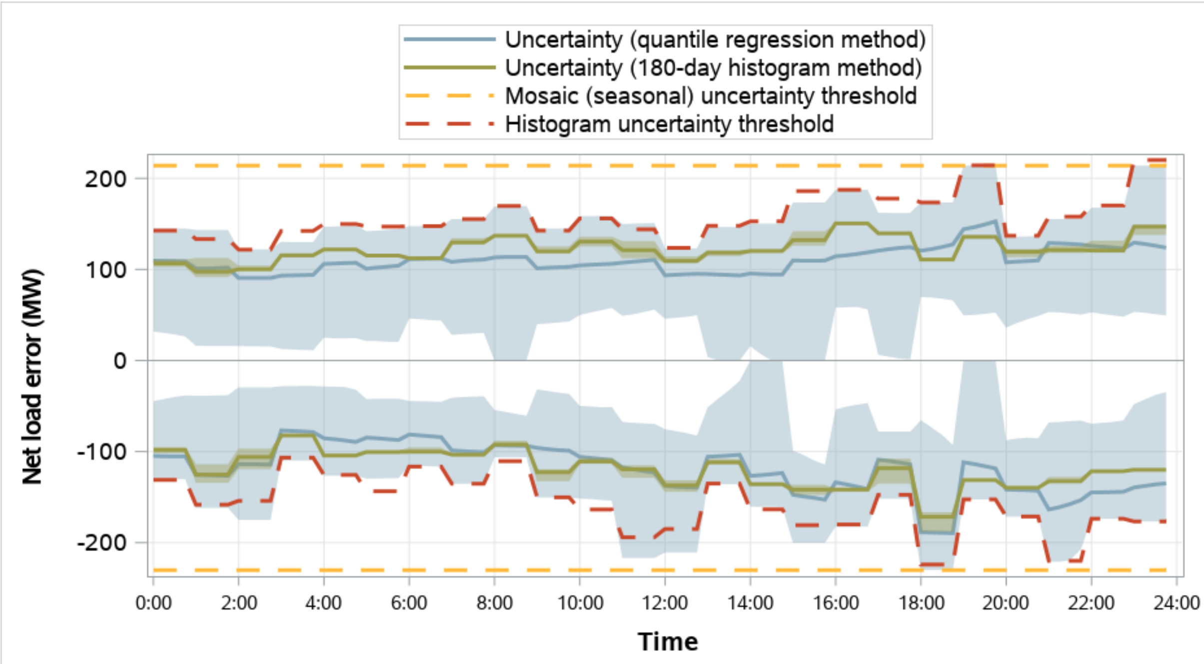


Figure 5.31 Portland General Electric distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

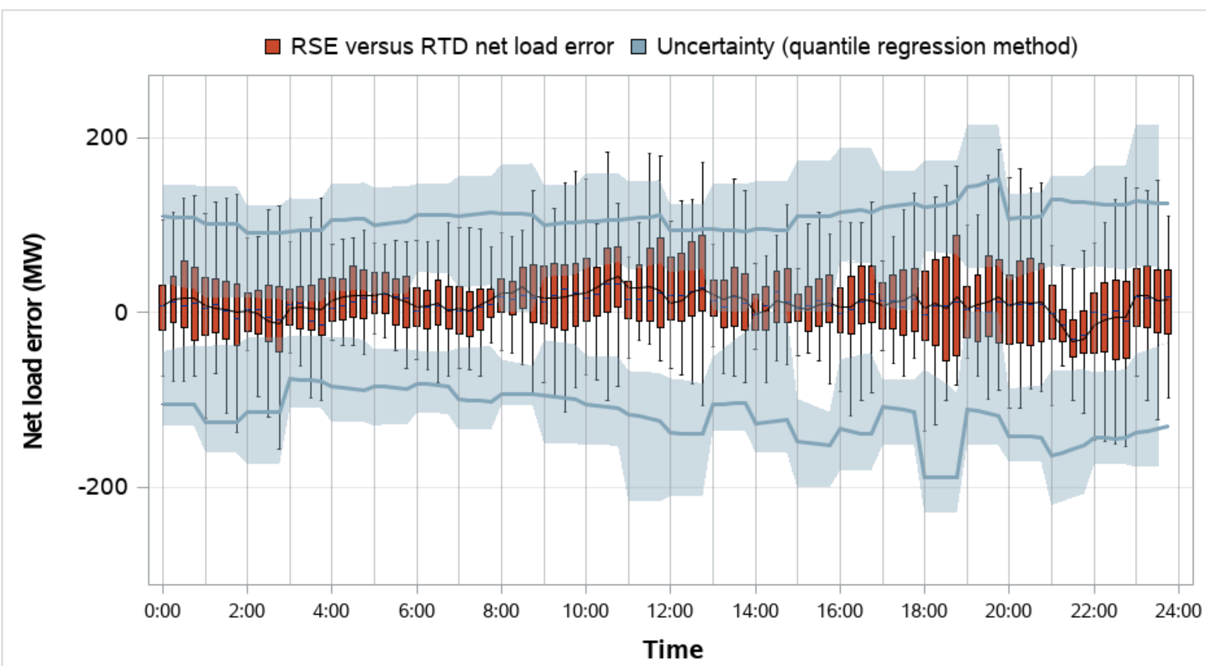


Figure 5.32 Powerex resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

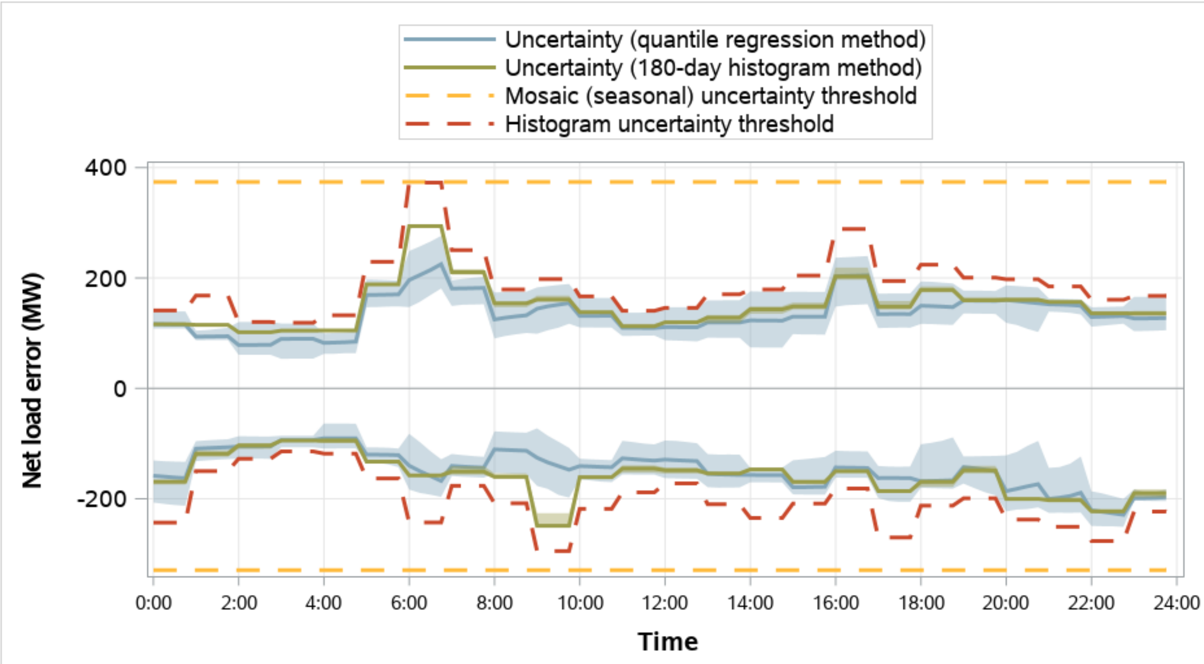


Figure 5.33 Powerex distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

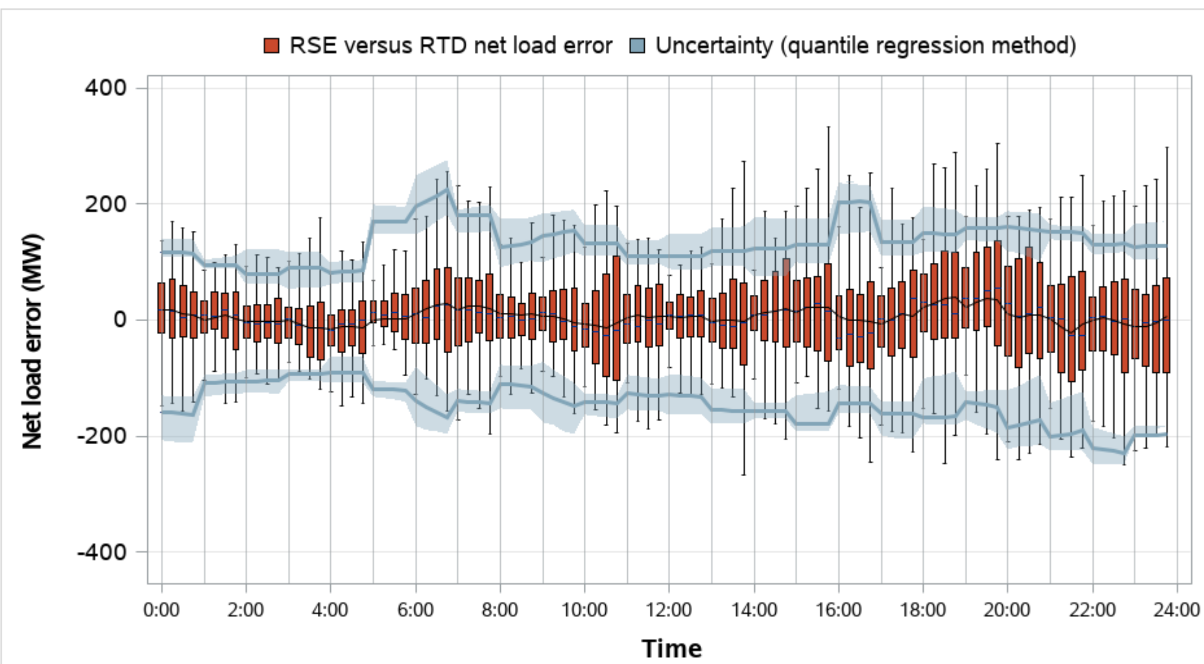


Figure 5.34 PNM resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

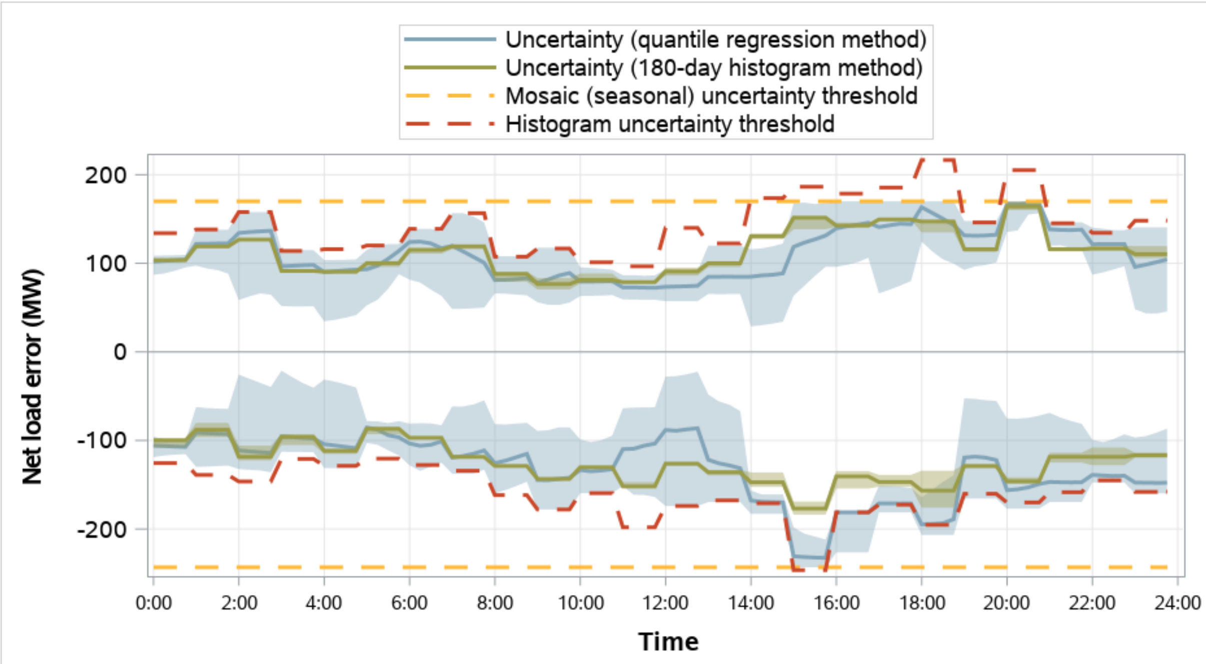


Figure 5.35 PNM distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

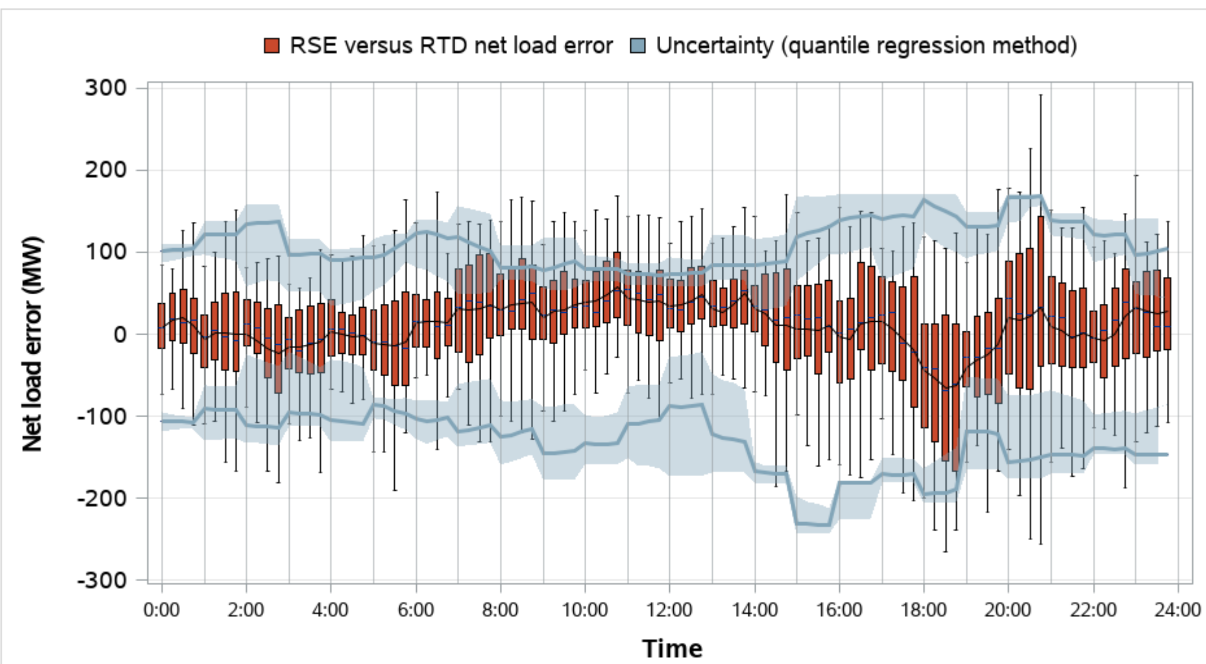


Figure 5.36 Puget Sound Energy resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

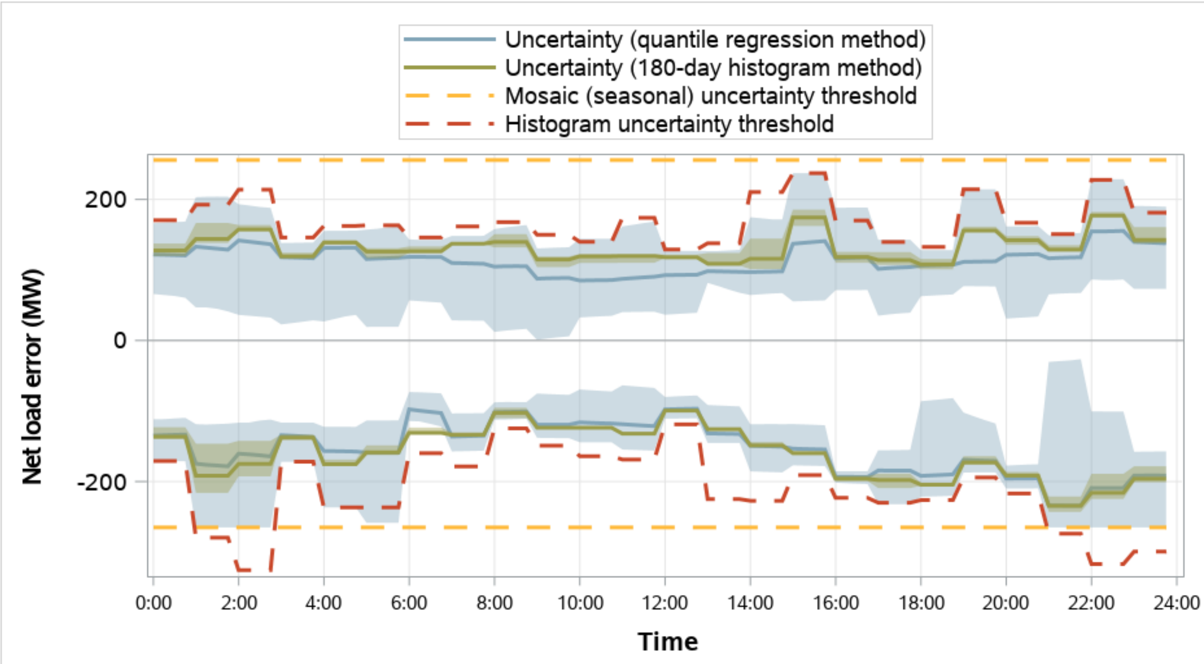


Figure 5.37 Puget Sound Energy distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

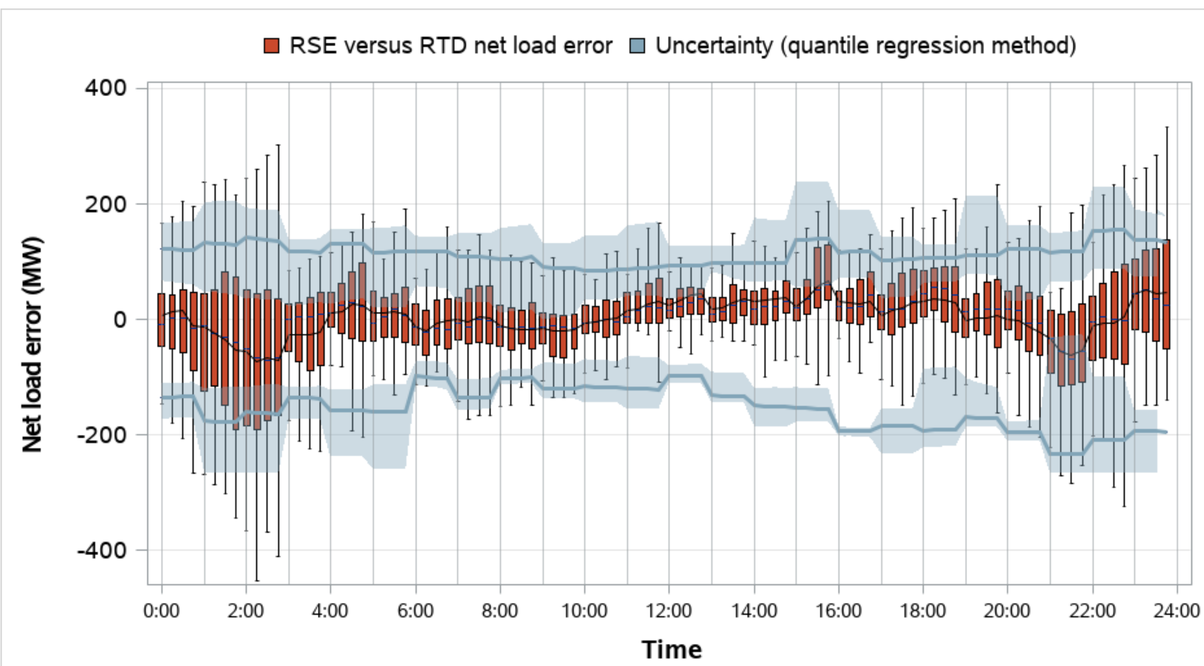


Figure 5.38 Salt River Project resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

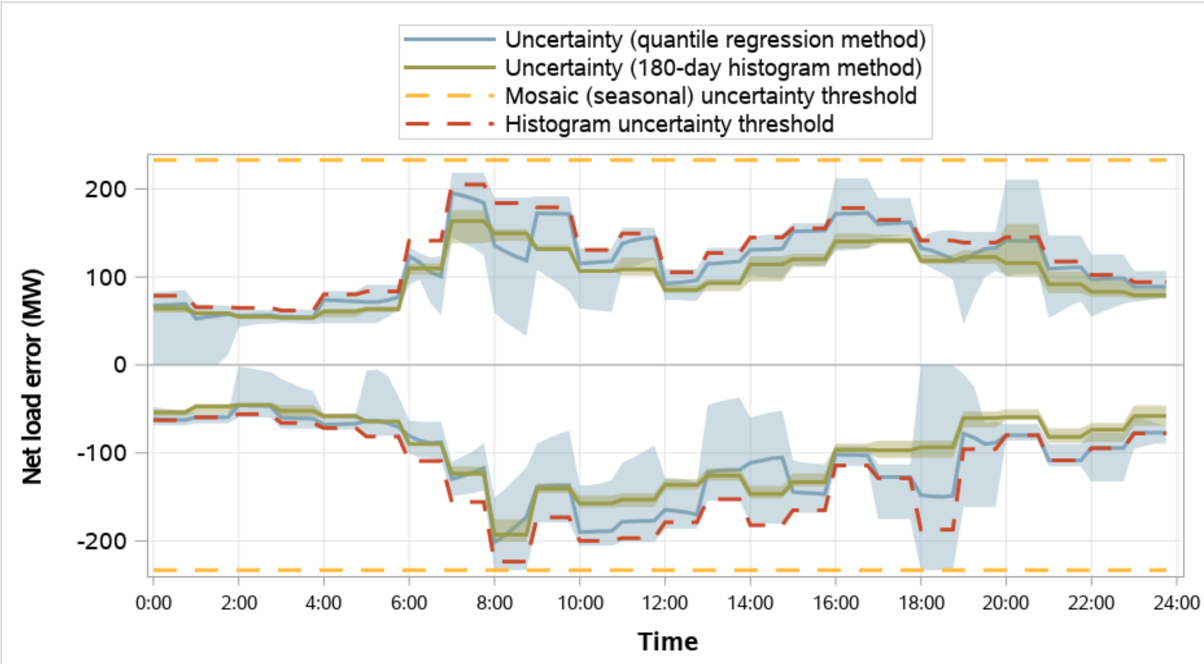


Figure 5.39 Salt River Project distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

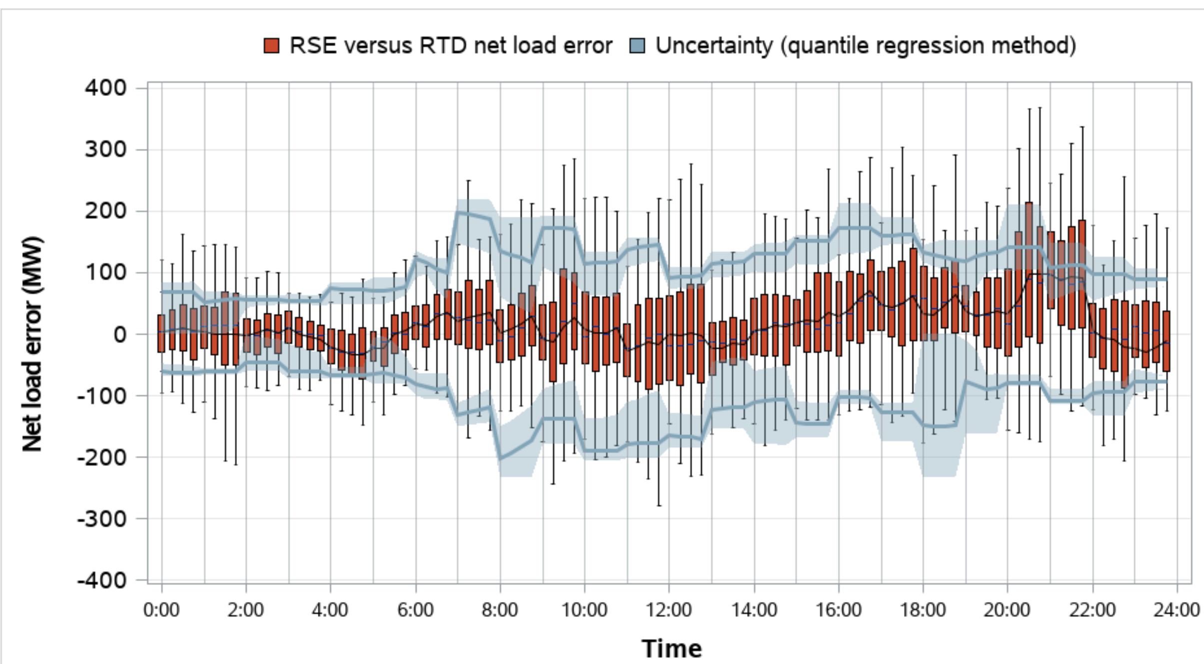


Figure 5.40 Seattle City Light resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

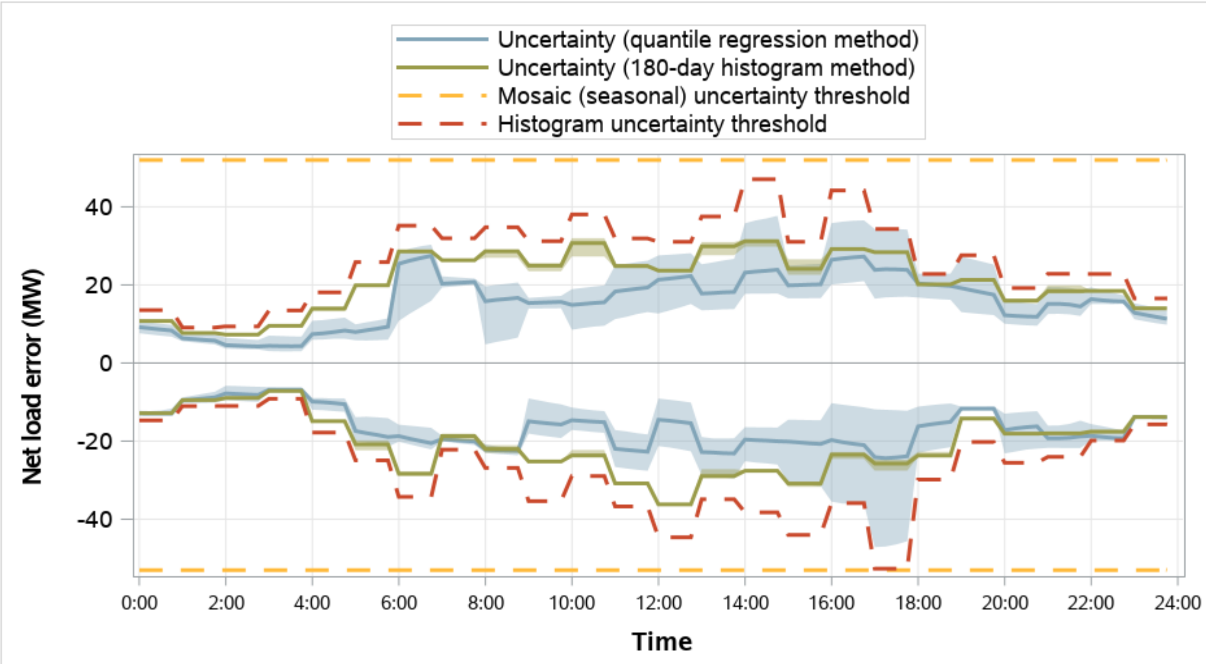


Figure 5.41 Seattle City Light distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

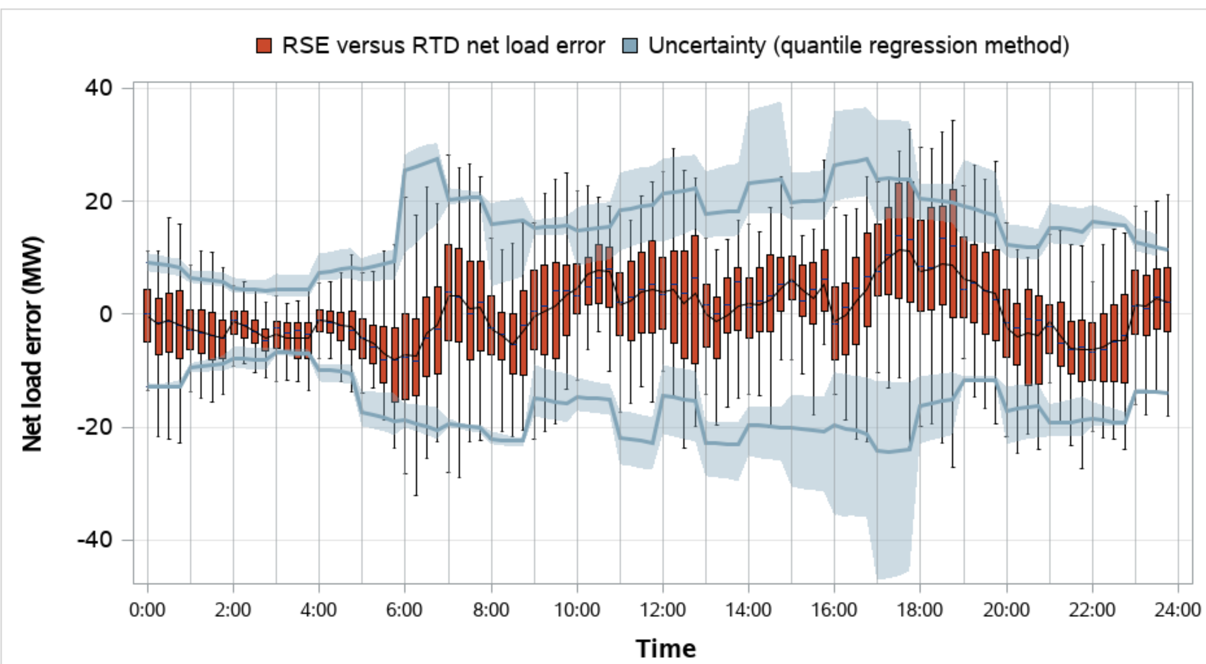


Figure 5.42 Tacoma Power resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

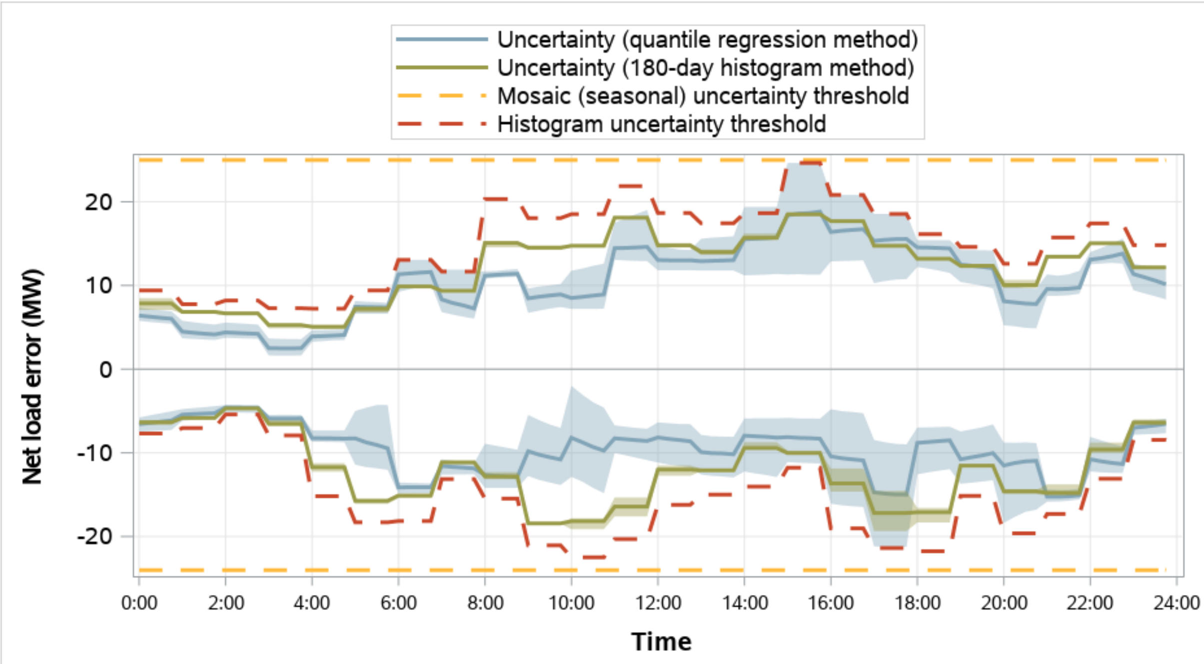


Figure 5.43 Tacoma Power distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

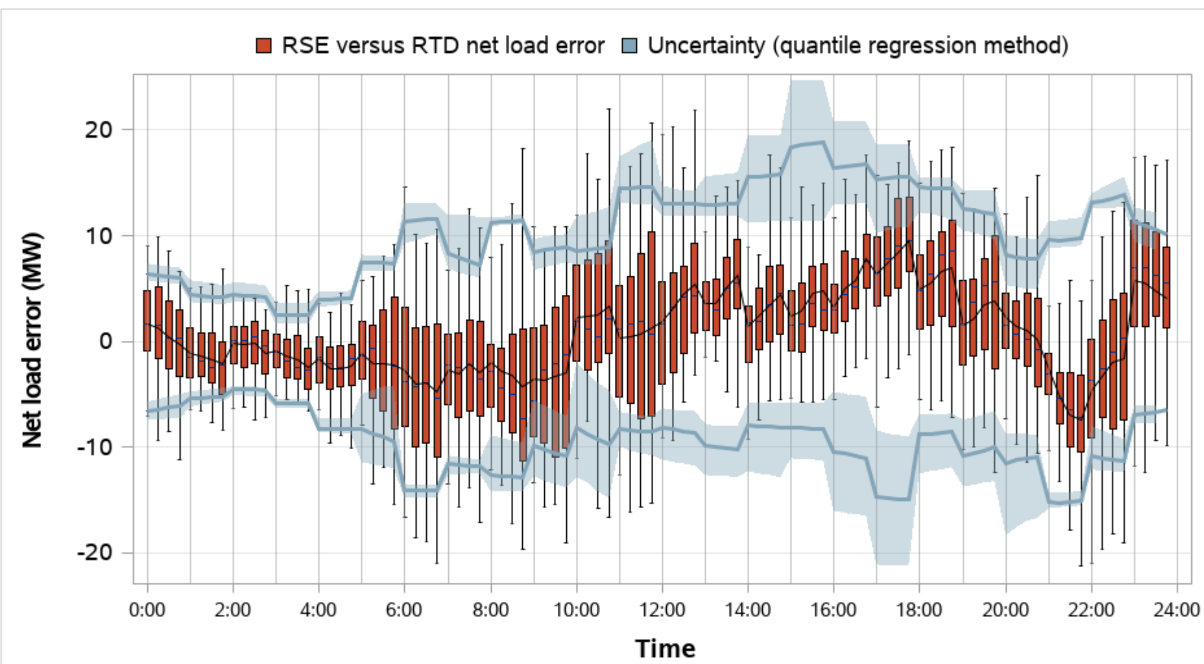


Figure 5.44 Tucson Electric Power resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

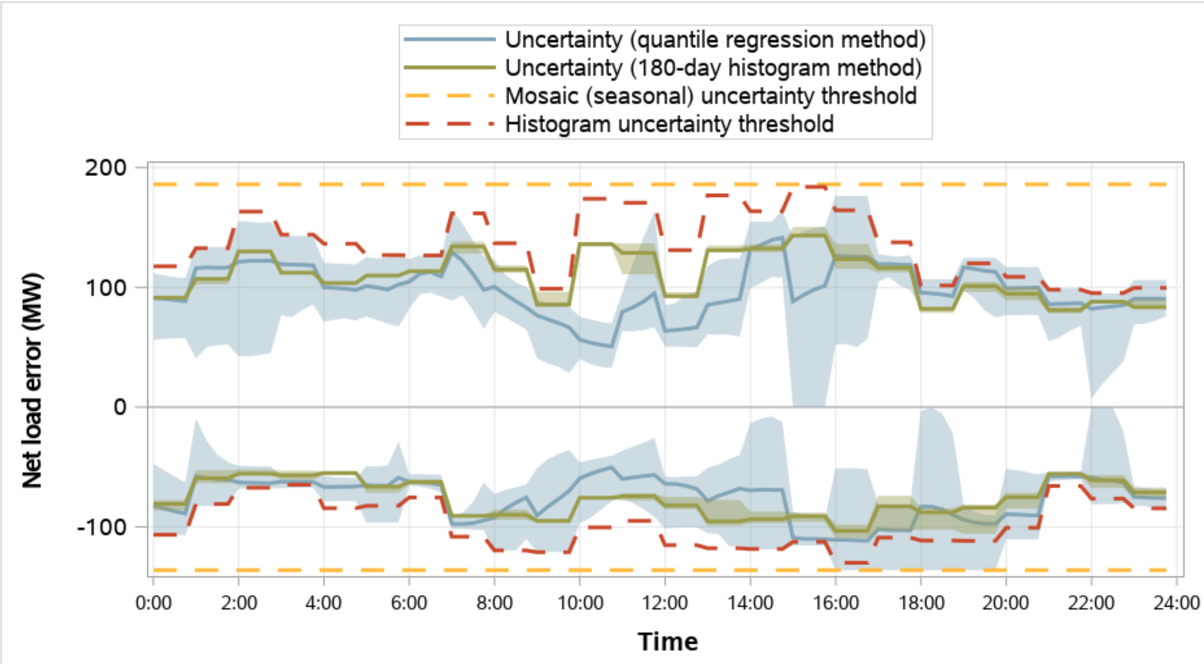


Figure 5.45 Tucson Electric Power distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

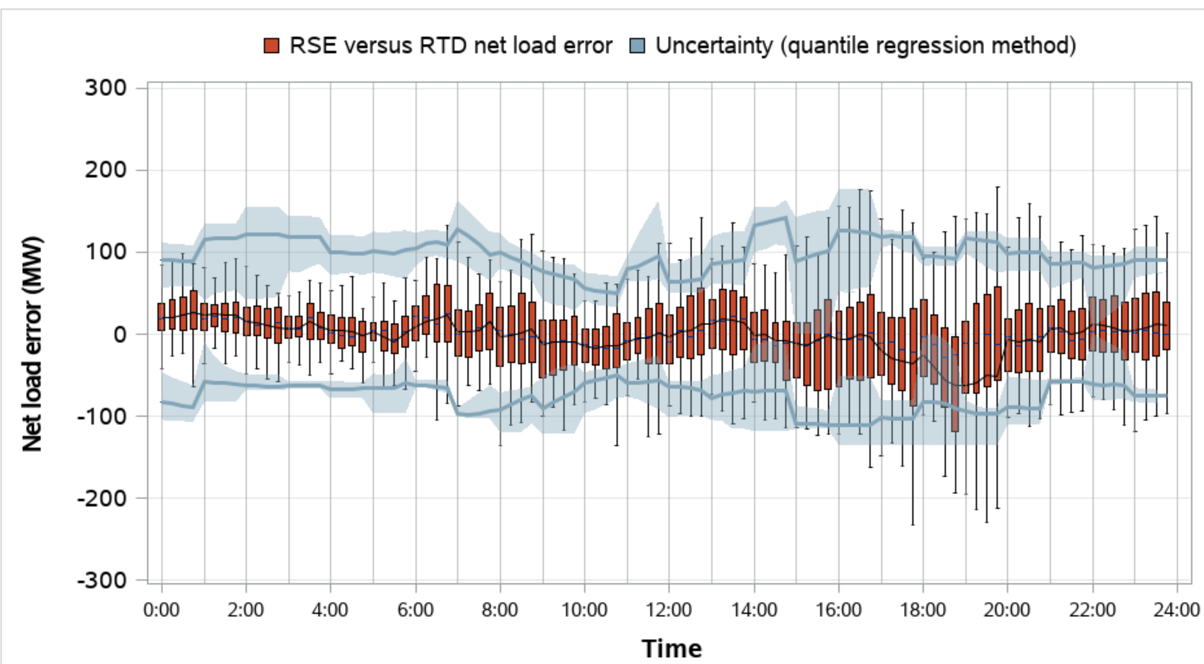


Figure 5.46 Turlock Irrigation District resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

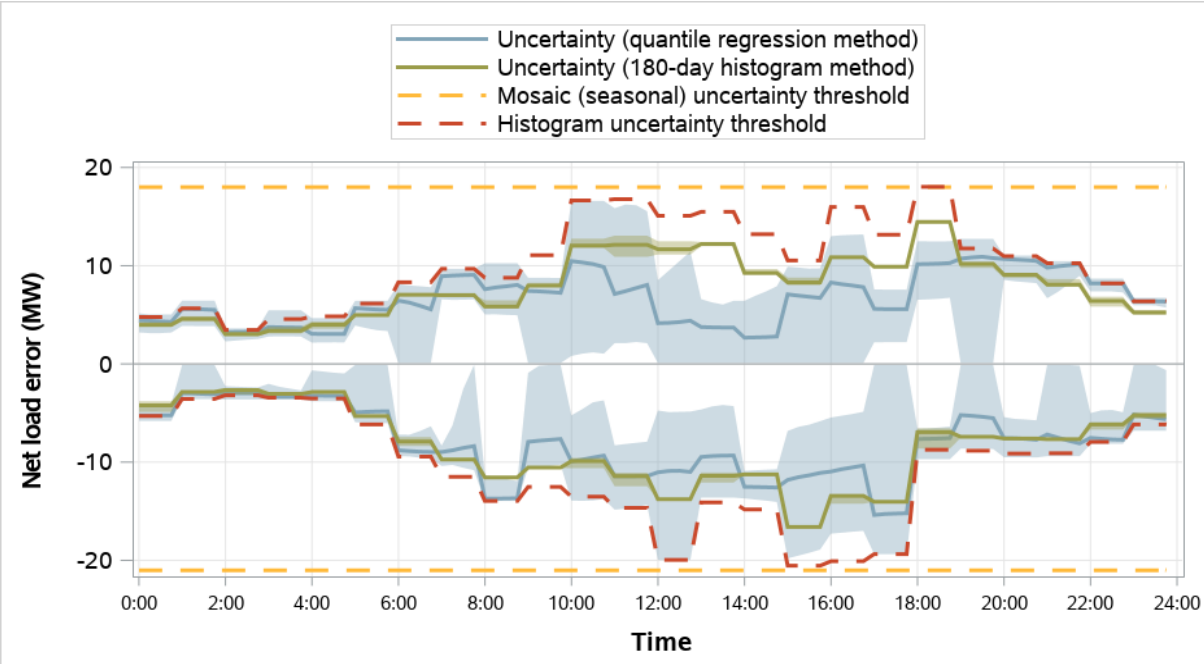


Figure 5.47 Turlock Irrigation District distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)

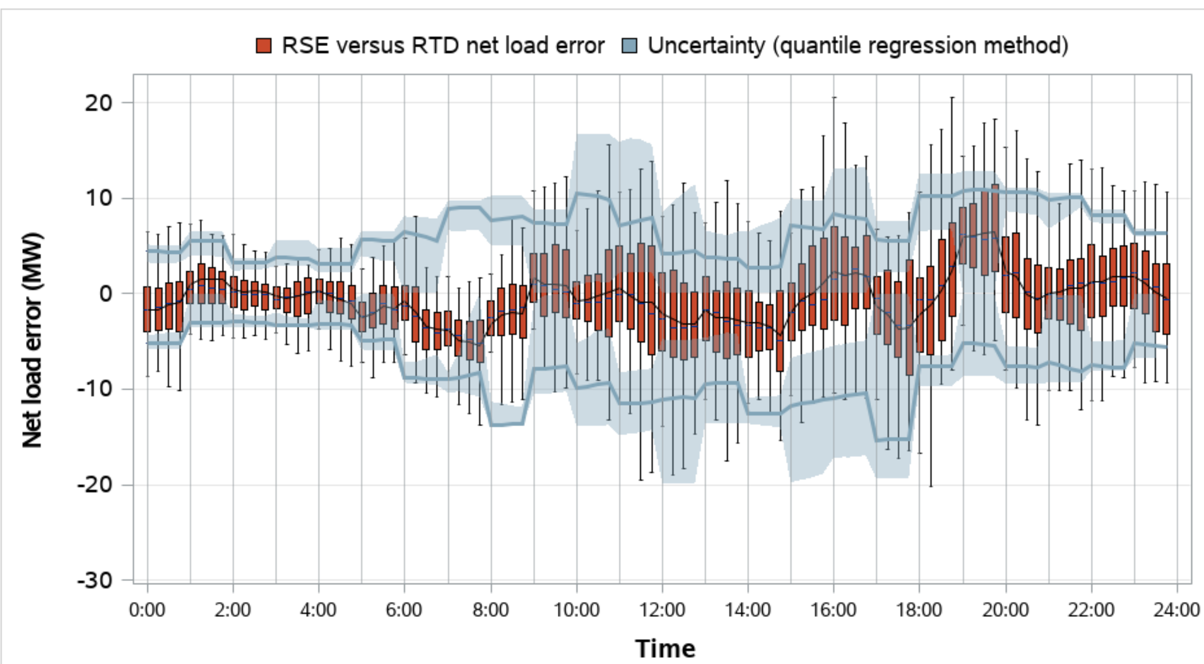


Figure 5.48 WAPA Desert Southwest resource sufficiency evaluation uncertainty requirements (weekdays, July 2023)

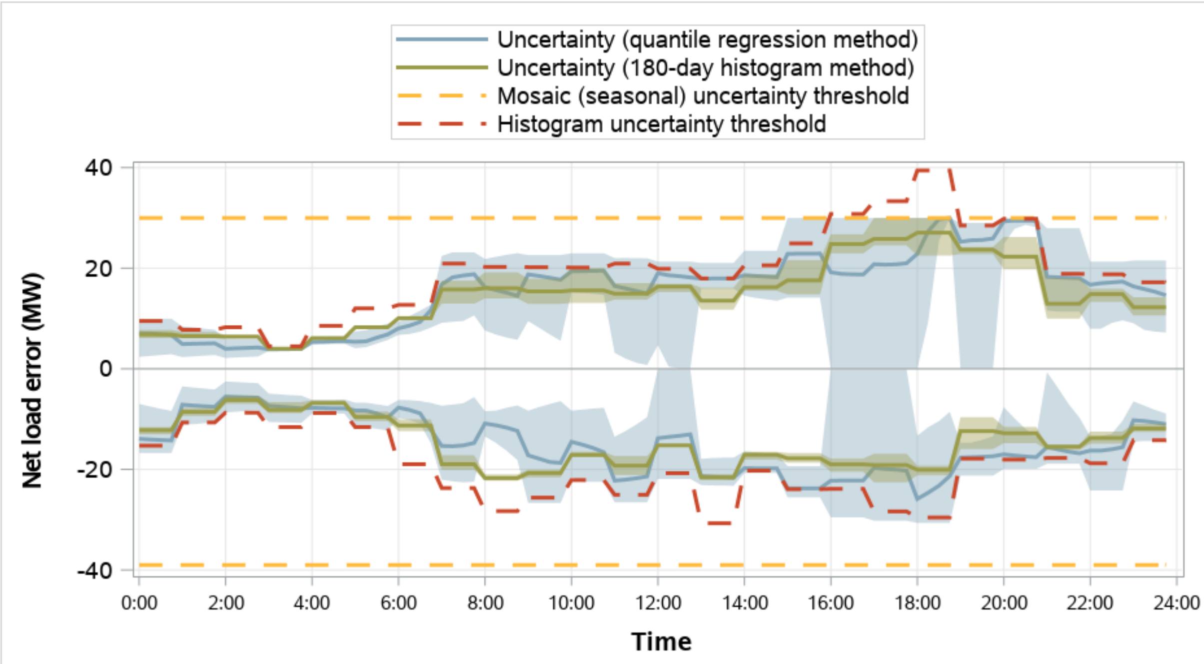
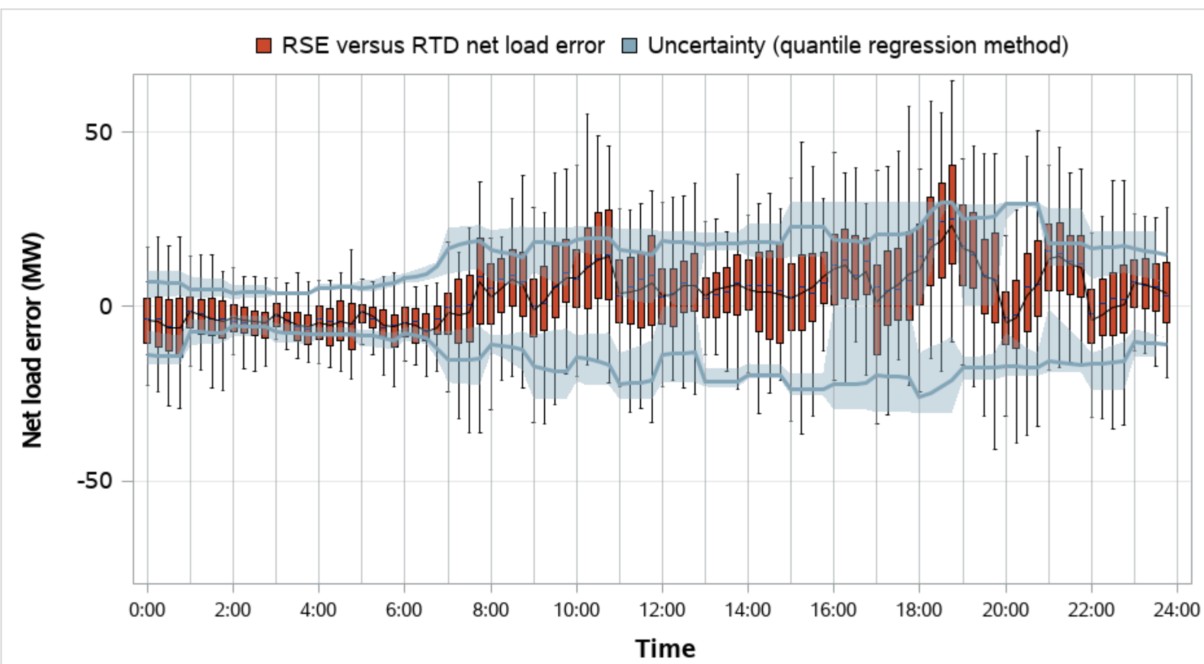


Figure 5.49 WAPA Desert Southwest distribution of RSE and RTD net load error and comparison to RSE uncertainty (weekdays, July 2023)



Performance measurements of quantile regression uncertainty

Table 5.1 summarizes the average requirements calculated using both the histogram and mosaic quantile regression methods. On average across all hours, the uncertainty calculated from the regression method was less than the histogram method for most of the balancing areas. The exceptions were El Paso Electric, Portland General Electric, Salt River Project, and WAPA Desert Southwest where uncertainty from the regression method was slightly higher than the histogram method on average for upward or downward uncertainty.

Table 5.2 summarizes the *actual net load error* — as measured by the difference between binding 5-minute market net load forecasts and *net load forecasts in the resource evaluation* — and how that compares to the mosaic regression uncertainty requirements for the same interval.[26] The left side of the table summarizes the closeness of the actual net load error to the uncertainty requirements when the actual net load error was within (or covered) by the upward and downward requirements.[27] The calculated uncertainty from the mosaic regression covered only 68 percent of actual net load error for Arizona Public Service, compared to 77 percent from the histogram method. For WAPA Desert Southwest, the calculated uncertainty from the mosaic regression covered 66 percent of actual net load error, compared to 68 percent from the histogram method. This result is in part driven by the thresholds in place which cap calculated uncertainty. For all other balancing areas, the mosaic regression requirements covered between 77 and 94 percent of actual net load errors. The right side of the table summarizes when the actual net load error instead exceeded upward or downward uncertainty requirements.

Table 5.3 shows the same information except with requirements calculated from the histogram method. Coverage from the histogram method was typically more than the mosaic method, with the exception of the coverage for El Paso Electric and Salt River Project.

[26] In comparing the 15-minute resource sufficiency evaluation forecasts to the three corresponding 5-minute forecasts, all three observations of error were used as a separate observation for calculating coverage, closeness, and exceedance.

[27] To the extent that the actual net load error averages around zero MW, this measurement largely matches the upward and downward uncertainty requirements.

Table 5.1 Average uncertainty requirements in the resource sufficiency evaluation (July 2023)

	Upward uncertainty			Downward uncertainty		
Balancing area	**Histogram**	**Mosaic**	***Difference***	**Histogram**	**Mosaic**	***Difference***
Arizona Public Service	194.3	170.8	-23.5	-163.3	-150.9	12.4
Avangrid	206.8	183.2	-23.6	-195.6	-185.8	9.8
Avista	47.5	36.2	-11.3	-61.6	-55.5	6.1
BANC	43.5	40.0	-3.5	-51.2	-47.0	4.3
Bonneville Power Admin.	225.5	218.9	-6.5	-360.7	-322.6	38.1
California ISO	1,148.1	1,023.0	-125.1	-940.5	-818.5	122.0
El Paso Electric	43.2	46.1	2.9	-35.4	-37.7	-2.3
Idaho Power	110.1	104.8	-5.3	-146.2	-133.7	12.6
LADWP	161.1	149.4	-11.7	-173.6	-165.3	8.3
NorthWestern Energy	73.7	66.4	-7.3	-78.5	-67.6	10.8
NV Energy	209.6	174.5	-35.0	-201.3	-154.1	47.3
PacifiCorp East	314.8	301.2	-13.6	-392.4	-366.9	25.5
PacifiCorp West	88.2	78.4	-9.8	-121.2	-100.8	20.4
Portland General Electric	119.4	111.9	-7.5	-127.1	-128.0	-0.9
Powerex	151.2	133.3	-17.9	-157.6	-145.0	12.6
PNM	110.6	103.9	-6.6	-127.2	-125.5	1.7
Puget Sound Energy	132.2	111.9	-20.2	-165.0	-156.6	8.5
Salt River Project	99.2	109.8	10.7	-94.8	-102.6	-7.7
Seattle City Light	20.8	15.6	-5.2	-20.5	-15.8	4.7
Tacoma Power	12.0	10.1	-1.9	-11.8	-9.2	2.6
Tucson Electric Power	107.7	98.7	-9.1	-80.6	-79.2	1.4
Turlock Irrigation District	7.8	6.2	-1.6	-8.2	-7.9	0.3
WAPA Desert Southwest	14.4	15.0	0.6	-14.6	-14.4	0.2

Table 5.2 Actual net load error compared to mosaic regression uncertainty requirements (July 2023)

	Actual net load error falls within calculated uncertainty requirements			Actual net load error exceeds ...			
				upward requirement		downward requirement	
Balancing area	**Percent of intervals**	**Distance to up requirement (MW)**	**Distance to down requirement (MW)**	**Percent of intervals**	**Amount (MW)**	**Percent of intervals**	**Amount (MW)**
Arizona Public Service	68%	144.9	186.8	23%	90.1	9%	59.4
Avangrid	94%	165.7	205.4	5%	46.1	2%	48.7
Avista	88%	38.2	54.4	7%	10.9	5%	21.5
BANC	83%	43.4	48.6	7%	20.6	10%	19.6
Bonneville Power Admin.	94%	254.1	293.5	3%	46.0	3%	42.3
California ISO	90%	947.7	929.2	6%	278.0	4%	216.5
El Paso Electric	80%	40.1	46.2	12%	18.9	9%	18.9
Idaho Power	90%	112.7	128.4	6%	37.7	4%	38.1
LADWP	90%	149.8	172.9	5%	48.6	5%	47.3
NorthWestern Energy	88%	64.6	70.0	4%	23.1	7%	27.5
NV Energy	82%	167.9	182.1	14%	60.3	5%	57.0
PacifiCorp East	80%	284.2	397.3	13%	107.7	7%	170.1
PacifiCorp West	89%	88.2	93.0	6%	24.0	5%	25.8
Portland General Electric	93%	104.5	136.8	5%	23.7	2%	32.8
Powerex	89%	133.3	146.2	7%	45.6	4%	51.4
PNM	84%	99.0	135.9	11%	28.9	6%	52.4
Puget Sound Energy	86%	111.2	157.9	9%	39.3	5%	61.6
Salt River Project	77%	104.7	116.5	14%	57.5	8%	52.2
Seattle City Light	88%	15.0	17.0	6%	4.1	6%	3.2
Tacoma Power	87%	9.4	10.2	7%	2.4	6%	2.4
Tucson Electric Power	87%	95.6	85.8	6%	31.6	7%	61.8
Turlock Irrigation District	80%	7.3	7.7	11%	3.8	9%	4.4
WAPA Desert Southwest	66%	14.1	17.0	19%	9.0	15%	8.6

Table 5.3 Actual net load error compared to histogram uncertainty requirements (July 2023)

Balancing area	Actual net load error falls within calculated uncertainty requirements			Actual net load error exceeds …			
				upward requirement		downward requirement	
	Percent of intervals	Distance to up requirement (MW)	Distance to down requirement (MW)	Percent of intervals	Amount (MW)	Percent of intervals	Amount (MW)
Arizona Public Service	77%	156.7	204.7	17%	78.9	6%	63.1
Avangrid	95%	189.0	215.3	3%	62.6	2%	60.8
Avista	92%	48.5	61.0	3%	16.4	4%	28.1
BANC	84%	48.3	52.3	8%	20.6	9%	15.4
Bonneville Power Admin.	97%	259.0	328.0	2%	63.5	2%	52.1
California ISO	95%	1,050.5	1,050.0	3%	205.9	2%	266.9
El Paso Electric	74%	38.3	43.8	15%	17.2	11%	16.6
Idaho Power	92%	117.7	140.5	5%	39.5	3%	41.9
LADWP	93%	159.6	179.6	4%	35.6	3%	35.7
NorthWestern Energy	90%	73.4	78.8	5%	28.1	5%	33.8
NV Energy	85%	204.3	238.8	11%	53.7	4%	58.2
PacifiCorp East	83%	287.5	419.5	11%	123.4	6%	178.3
PacifiCorp West	92%	96.9	113.0	4%	29.6	4%	31.6
Portland General Electric	94%	111.0	136.2	4%	30.5	2%	41.1
Powerex	92%	149.7	159.5	5%	46.0	3%	52.0
PNM	87%	99.8	139.4	7%	29.7	6%	50.4
Puget Sound Energy	90%	130.4	166.3	6%	52.0	4%	61.2
Salt River Project	75%	92.7	107.8	15%	61.0	10%	51.1
Seattle City Light	94%	20.2	21.6	3%	3.6	3%	2.7
Tacoma Power	93%	11.2	12.7	4%	2.2	3%	2.6
Tucson Electric Power	91%	102.6	86.3	4%	29.3	5%	72.9
Turlock Irrigation District	86%	8.5	8.1	7%	2.8	7%	3.4
WAPA Desert Southwest	68%	13.1	16.1	19%	9.2	13%	8.1

Variability of quantile regression uncertainty

Prior to February 2023, uncertainty used in the resource sufficiency evaluation was known in advance of the trade date based on the lower and upper percentiles of observations over the historical period for the same hour (*histogram approach*). Under this approach, the uncertainty was also the same in each interval for the evaluation hour. The *mosaic quantile regression* approach combines regression results with current load, solar, and wind forecast information to calculate uncertainty in each 15-minute interval of the evaluation hour. With this approach, the regression coefficients for individual balancing areas are known in advance, but the exact uncertainty is dependent on current forecast information. A natural consequence of this is that calculated uncertainty has greater variability and is more difficult to predict in advance. This section looks at the variability of net load uncertainty using metrics reflecting several different perspectives.

Changes in uncertainty between resource sufficiency evaluation runs

Figure 5.50 shows the difference in the calculated upward uncertainty from the first run of the resource sufficiency evaluation at 75 minutes prior to the evaluation hour, to the second run of the resource sufficiency evaluation at 55 minutes prior to the evaluation hour. Figure 5.51 shows the same information for downward uncertainty. Load and renewable forecasts are held fixed between the second (T-55) and final (T-40) resource sufficiency evaluations such that uncertainty is also unchanged between these runs. Therefore, these figures summarize how effective the T-75 uncertainty is in predicting the final uncertainty used in the resource sufficiency evaluation. The dashed gray region shows effectively no difference from the first resource sufficiency evaluation (less than one MW change). The regions above or below this show increased or decreased uncertainty relative to the T-75 results. The uncertainty difference from the first run of the resource sufficiency evaluation was typically less than 10 MW. More significant increases in the uncertainty requirement also occurred in rare instances and may lead to unexpected resource sufficiency evaluation failures.

Figure 5.50 Megawatt change in upward quantile regression uncertainty between T-75 and T-55 resource sufficiency evaluation runs (July 2023)

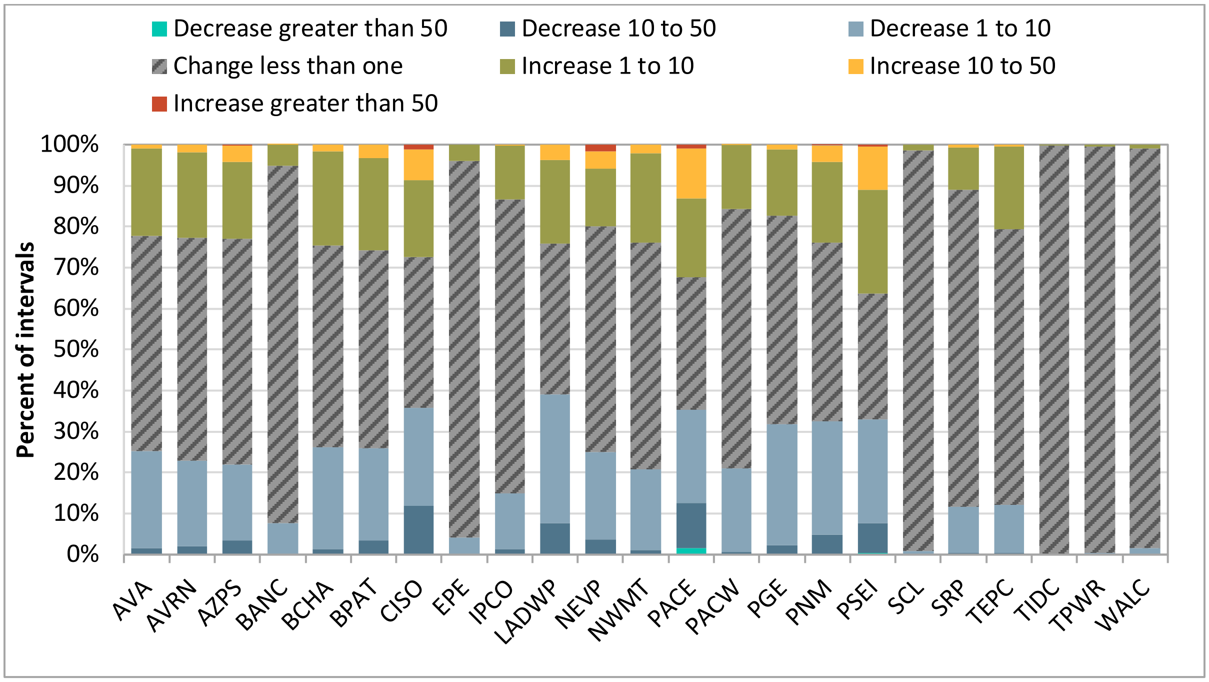
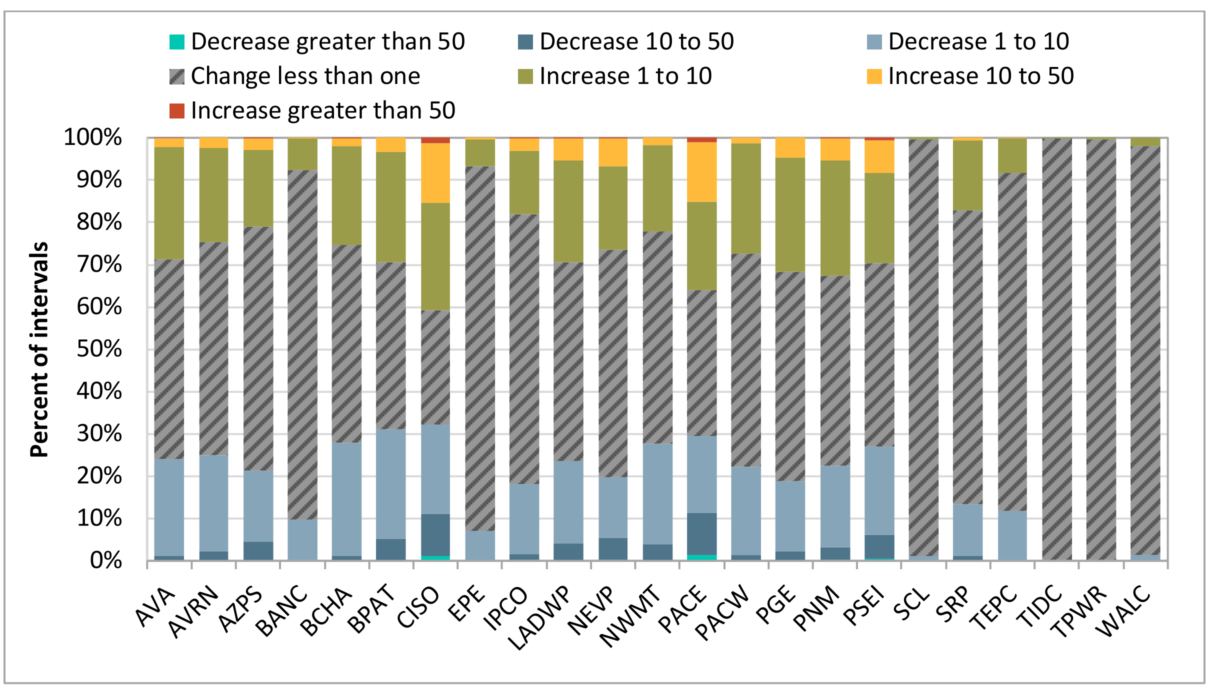


Figure 5.51 Megawatt change in downward quantile regression uncertainty between T-75 and T-55 resource sufficiency evaluation runs (July 2023)



Movement of net load uncertainty between intervals

This section summarizes how much movement exists between adjacent intervals for the net load uncertainty used in the final resource sufficiency evaluation. Prior to February, the uncertainty used in the tests was the same throughout each hour. Following the enhancements, uncertainty used in each 15-minute interval of the evaluation hour are more variable as an expected consequence of using forecast conditions in each interval to effect uncertainty. Each interval of the evaluation hour still uses the same underlying historical distribution and regression outputs, but can vary based on the forecast information. Figure 5.52 summarizes the absolute change of upward uncertainty between adjacent intervals *within the hour* (intra-hour movement). Intra-hour uncertainty movement for most balancing areas was typically small, at less than 10 MW in most intervals.

Figure 5.53 instead shows the absolute change of upward uncertainty *between* hours (inter-hour movement). This is expected to show greater variability compared to intra-hour movement because the underlying historical distribution and regression outputs also differ between hours in addition to the forecast information. Figure 5.54 shows the same information for inter-hour uncertainty movement based on the histogram approach. The inter-hour movement from histogram uncertainty was similar, but generally showed lower levels of change across all balancing areas.

Figure 5.52 Intra-hour movement of upward quantile regression uncertainty — absolute MW change (July 2023)

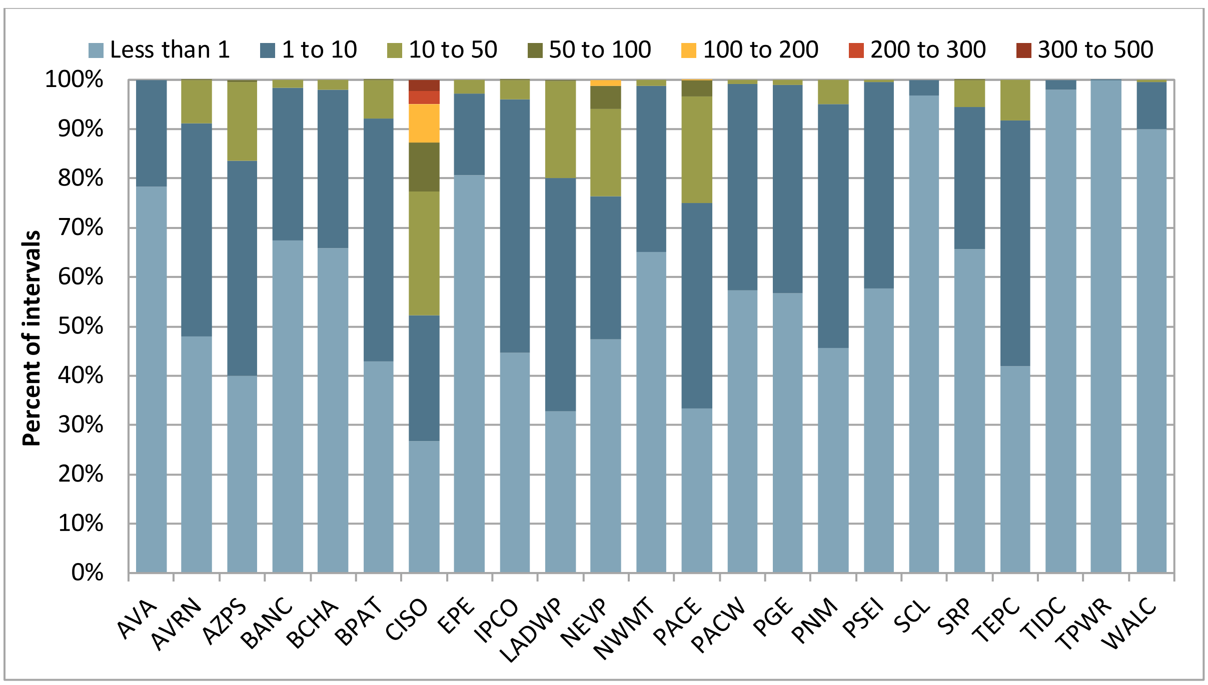


Figure 5.53 Inter-hour movement of upward quantile regression uncertainty — absolute MW change (July 2023)

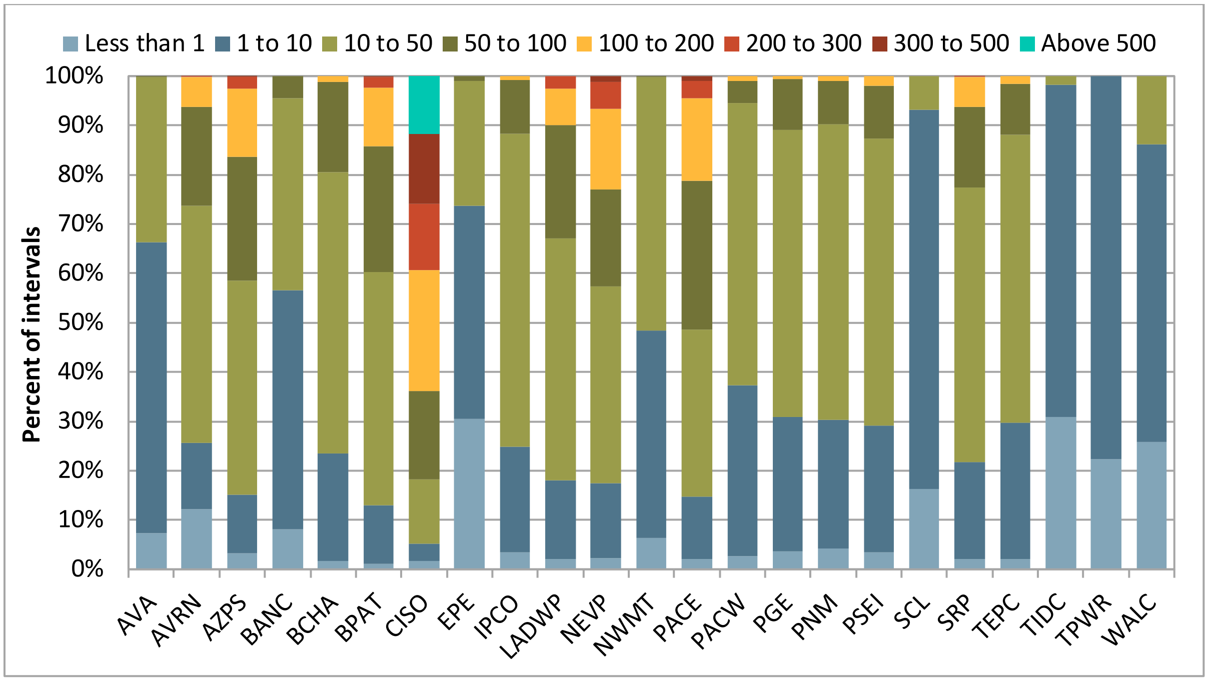
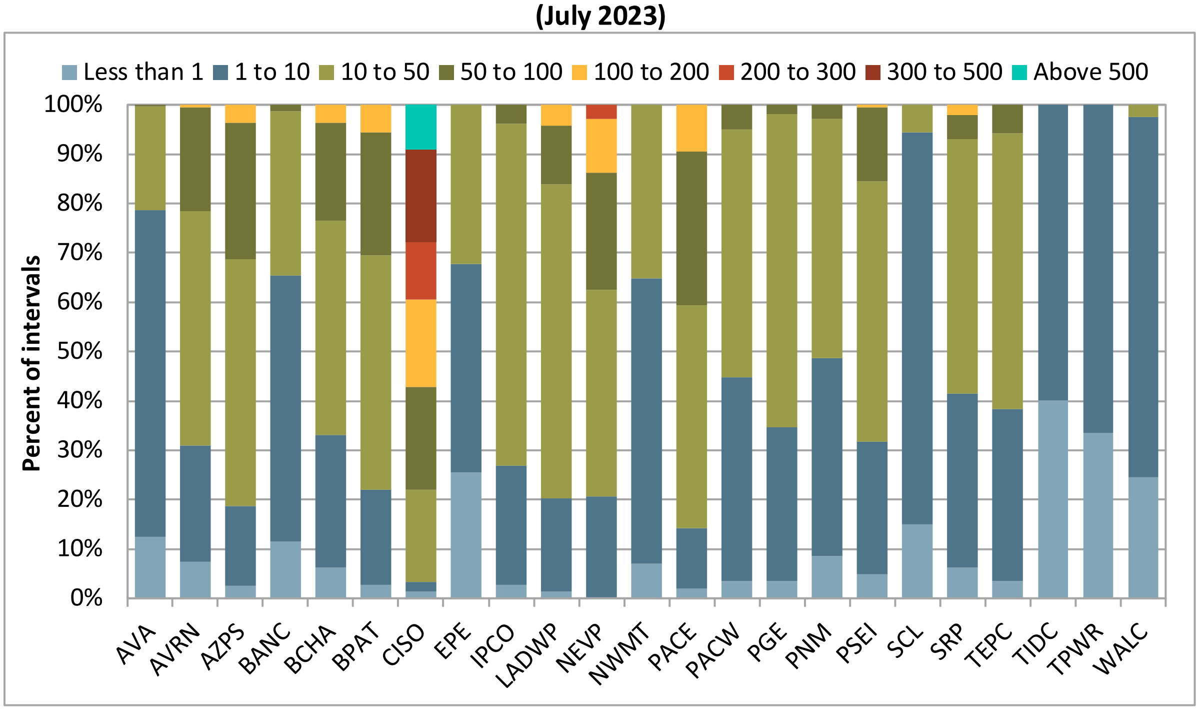


Figure 5.54 Inter-hour movement of upward histogram uncertainty — absolute MW change (July 2023)



6 Demand-response-based load adjustments in the resource sufficiency evaluation

As part of phase 1 of the resource sufficiency evaluation enhancements initiative, the California ISO implemented a new feature to allow WEIM entities to submit load forecast adjustments to reflect demand response programs which could not be accounted for otherwise in the real-time market. This adjustment is included in both the capacity and flexibility tests, and impacts the load used in the requirements of both tests.

The use of demand-response-based load adjustments in the resource sufficiency evaluation increased during July. Table 6.1 summarizes the frequency and size of these adjustments. A negative adjustment reflects a lower load forecast as a result of a demand response program. This will decrease the requirement for the upward capacity and flexibility tests, but will increase the requirement for the downward tests. The adjustments can also be entered as a positive load adjustment. This can reflect additional demand because of expected pre-cooling or post-demand-response-event increases (sometimes referred to as snapback).

The feature to adjust the load forecast in the tests based on a demand-response program was used by five balancing areas in July: Arizona Public Service, Idaho Power, NV Energy, PacifiCorp East, and Portland General Electric. NV Energy was the most frequent user of this feature, with negative adjustments in 39 hours during July at around -53 MW on average (or -97 MW at its lowest). NV Energy has also submitted positive adjustments following the demand-response events. This occurred during 41 hours at 28 MW on average.

Figure 6.1 and Figure 6.2 show all hourly demand-response-based load adjustments between hours 17 and 21 during the first and second half of July, respectively. Most of the demand-response-based adjustments occurred during these peak hours.

During this period, these adjustments allowed Idaho Power to pass the flexibility test in five intervals that would have been failures had the demand response program not been accounted for in the test. This also occurred for Arizona Public Service in one interval. These adjustments did not have any impact on any area passing the capacity test that would have been a failure otherwise.

Table 6.1 Frequency and average size of demand-response-based load adjustments in the resource sufficiency evaluation (July 1 - July 31, 2023)

	Negative **demand-response-based load adjustment**				***Positive*** **demand-response-based load adjustment**			
Balancing area	**Total hours**	**Percent of hours**	**Average adjustment**	**Lowest adjustment**	**Total hours**	**Percent of hours**	**Average adjustment**	**Highest adjustment**
Arizona Public Service	19	2.6%	-51.0	-95	19	2.6%	13.9	43
Idaho Power	29	3.9%	-38.0	-90	—	—	—	—
NV Energy	39	5.2%	-53.3	-97	41	5.5%	28.0	62
PacifiCorp East	4	0.5%	-60.0	-65	—	—	—	—
Portland General Electric	6	0.8%	-40.7	-77	—	—	—	—

Figure 6.1 Demand-response-based load adjustments included in the resource sufficiency evaluation (peak hours, July 1 - July 15, 2023)

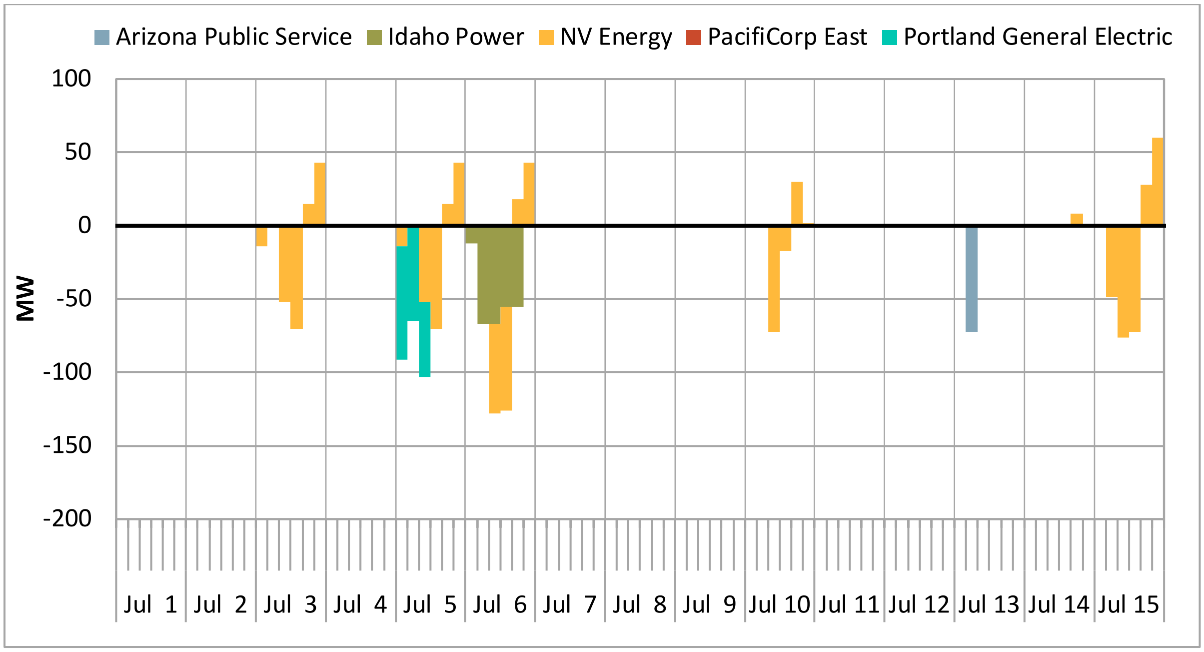
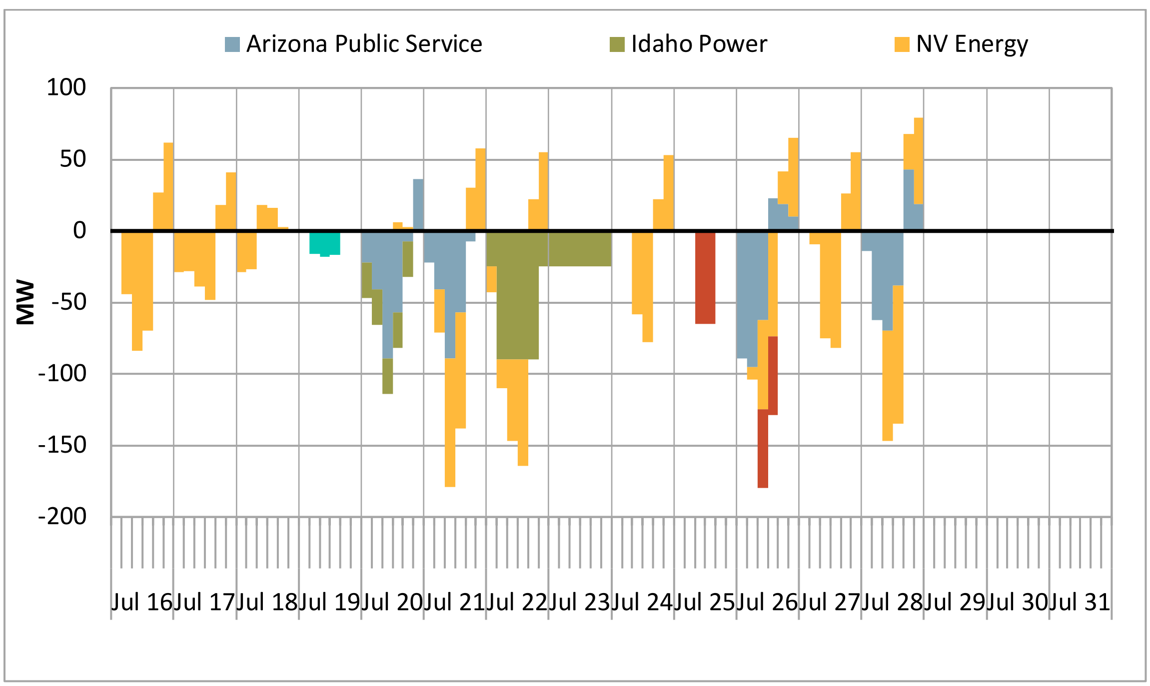


Figure 6.2 Demand-response-based load adjustments included in the resource sufficiency evaluation (peak hours, July 16 - July 31, 2023)



7 WEIM import limits following test failure

This section summarizes the import limits that are imposed when a WEIM entity fails either the bid-range capacity or the flexible ramping sufficiency test in the upward direction.

Balancing areas can voluntarily opt in to receiving assistance energy transfers. When a balancing area opts in to the program, their WEIM transfers will not be affected by any limits that would exist following an upward resource sufficiency evaluation failure — allowing the market to freely and optimally schedule WEIM transfers based on supply and demand conditions in the system. The import limits summarized in this section cover both balancing areas that opted out or opted in to the assistance energy transfer program. For balancing areas that opted in to the program, these limits reflect what would have been in place had the balancing area not opted in.

When either test fails in the upward direction, imports will be capped at the greater of (1) the base transfer or (2) the transfer from the last 15-minute market interval. Figure 7.1 summarizes the import limits after failing either test by the source of the limit. The black horizontal line (right axis) shows the number of 15-minute intervals with either a capacity or a flexibility test failure while the bars (left axis) show the percent of failure intervals in which the WEIM import limit was capped by either the base transfer or the last 15-minute market transfer. In some cases, the import limit after failing the test (i.e. the greater of the base transfer or last 15-minute interval transfer) is at or above the unconstrained total import capacity. In these cases, the import limit imposed after failing the test has no impact.

Figure 7.1 Upward capacity/flexibility test failure intervals by source of import limit (July 2023)

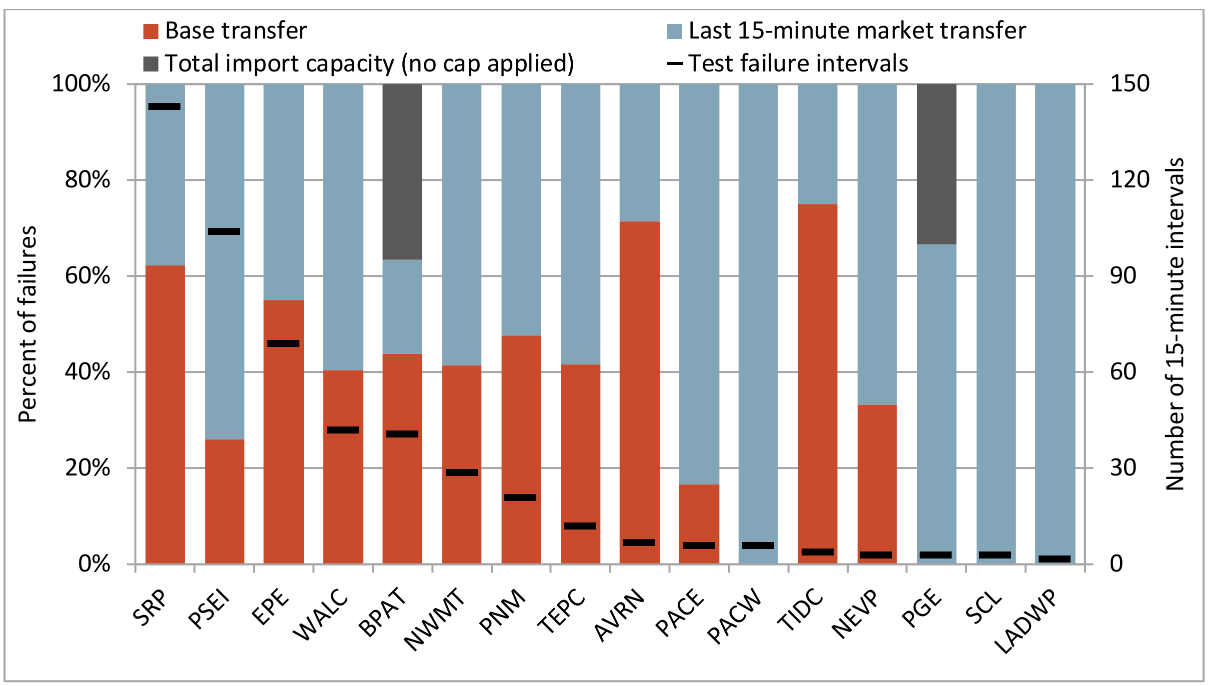


Figure 7.2 summarizes dynamic WEIM import limits above base transfers (fixed bilateral transactions between WEIM entities) after failing either test in the upward direction.[28] From this perspective, the incremental WEIM import limit after a test failure is set by the greater of (1) zero or (2) the transfer from the last 15-minute market interval minus the current base transfer. Therefore, the dynamic import limits show the incremental flexibility available through the WEIM after a resource sufficiency evaluation failure. The black horizontal line (right axis) shows the number of 15-minute intervals with an import limit imposed after a test failure. Areas without any upward test failures during the month were excluded.

Figure 7.2 Upward capacity/flexibility test failure intervals by dynamic import limit (July 2023)

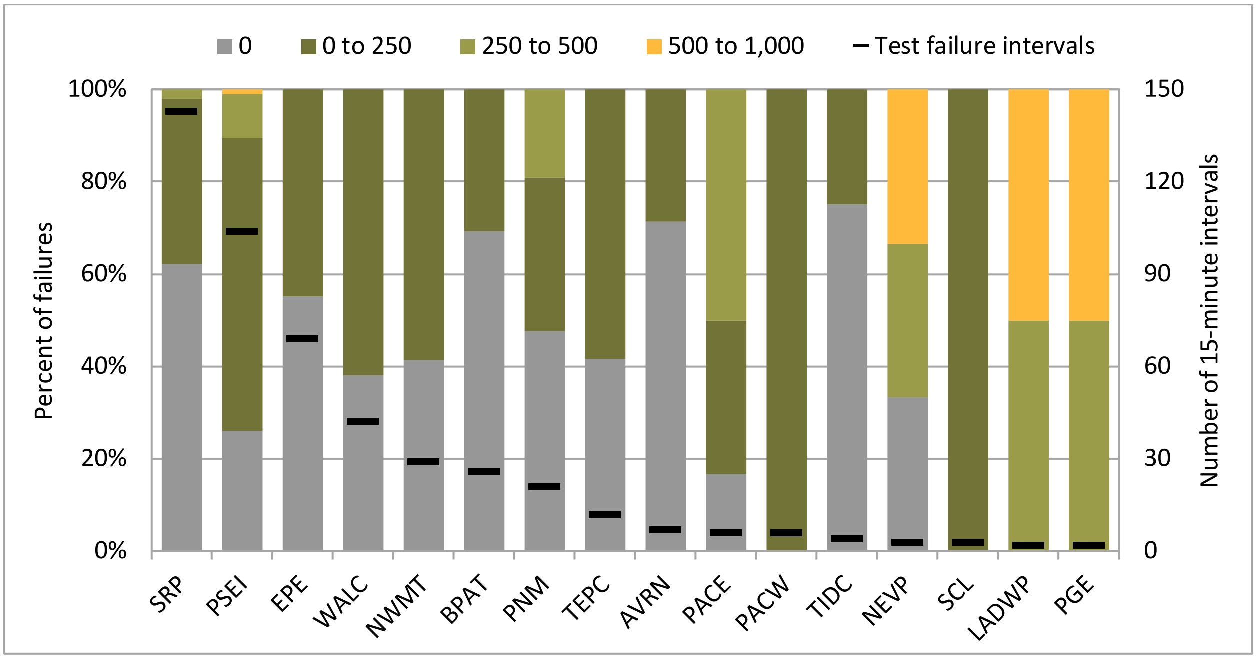


Figure 7.3 summarizes whether the import limit that was imposed after failing either test in the upward direction impacted market transfers (or would have impacted market transfers had the balancing area not opted in to the assistance energy transfer program).[29] The black horizontal line (right axis) shows the number of 15-minute market intervals with either a capacity or flexibility test failure. The blue bars (left axis) show the percent of failure intervals in which the resulting transfers after failing the resource sufficiency evaluation were *below* the import limit that was imposed (or would have been imposed for opt-in balancing areas). In all other failure intervals (yellow bars), the resulting transfers were either constrained to the limit imposed after failing the test or would have been constrained by the limit without an opt-in designation. These results are shown separately for the 15-minute (FMM) and 5-minute (RTD) markets.

[28] Test failure intervals in which an import limit was not imposed because it was at or above the unconstrained total import capacity were excluded from this summary.

[29] Test failure intervals in which an import limit was not imposed because it was at or above the unconstrained total import capacity were excluded from this summary.

Figure 7.3 Percent of upward failure intervals in which WEIM imports were constrained or would have been constrained by test failure limits (July 2023)