Western Energy Imbalance Market Resource Sufficiency Evaluation Metrics Report covering Q4 2024

January 24, 2025

Prepared by: Department of Market Monitoring

California Independent System Operator

1 Report summary

As part of the Western Energy Imbalance Market (WEIM) design, each balancing area is subject to a *resource sufficiency evaluation* (RSE). The evaluation is performed prior to each hour to ensure that generation capacity and flexibility in each area is sufficient without relying on transfers from other WEIM balancing areas. In this report, DMM provides additional information and analysis about resource sufficiency evaluation performance, accuracy, impacts, and enhancements during the fourth quarter of 2024.

Report highlights

Resource sufficiency evaluation failures

- **The Public Service Company of New Mexico (PNM) failed the upward flexibility test relatively frequently in November, during around 7 percent of intervals.** The frequency of capacity and flexibility test failures remained low across all other balancing areas for the quarter.

Enhancements in determining uncertainty for pass-group

- **On June 25, 2024, the ISO made an improvement for determining the group of balancing areas passing the resource sufficiency evaluation in advance of the regressions for calculating uncertainty for the pass-group.**
- In some intervals, the regressions for calculating the uncertainty requirement for the pass-group must be performed before the final set of balancing areas in this group are known. An improvement in the process increased the consistency between (1) the group of balancing areas used to determine the regression coefficients for the pass-group and (2) the group of balancing areas whose forecast information gets combined with those coefficients to determine the uncertainty requirement.
- Following this enhancement, the set of balancing areas in the pass-group between the regression and current-forecast-information differed in around 7 percent of intervals, compared to around 18 percent of intervals prior to the enhancement.
- Additional improvements should still be considered to address any remaining inconsistency between these two sets of information.

Assistance energy transfers

- **Eight balancing areas were opted in to assistance energy transfers (AET) during the fourth quarter.** Five of these areas (Avangrid, Idaho Power, NorthWestern Energy, PacifiCorp East, and PacifiCorp West) failed the resource sufficiency evaluation during at least one interval while opted in to the program, gaining access to additional WEIM supply that would not have been available otherwise. The other three areas (California ISO, NV Energy, and WAPA Desert Southwest) did not fail the resource sufficiency evaluation while participating in the program.

Quantile regression method for calculating uncertainty

- **The regression model's predicted uncertainty for the resource sufficiency evaluation covered the realized uncertainty much less for intervals at the end of the evaluation hour than for intervals at the beginning of the hour.** This is because the model is designed to predict uncertainty in forecasts that are produced only 45 to 55 minutes before real-time. However, the time horizon of the resource sufficiency evaluation includes four intervals, produced between 47.5 and 102.5 minutes before real-time.

CAISO non-participating pump load

This report also highlights non-participating pump loads in the ISO balancing area that are not included in the ISO area resource sufficiency evaluation.

- **Non-participating pump load is included in the ISO area real-time market requirement, but is not included in the resource sufficiency evaluation.** This can contribute to conditions in which the ISO passes the resource sufficiency evaluation while an Energy Emergency Alert is issued (such as during July 2023). DMM continues to recommend that the ISO and stakeholders consider whether non-participating pump load should be included in the resource sufficiency evaluation. This would better align the conditions in the real-time market with the conditions considered in the resource sufficiency evaluation.

Organization of the report

- Section 2 **summarizes the frequency and size of resource sufficiency evaluation failures.**
- Section 3 **summarizes the impact of advisory resource sufficiency evaluation runs**. This section describes improvements made for determining the group of balancing areas that pass the resource sufficiency evaluation in advance of the regressions for calculating uncertainty.
- Section 4 **summarizes the use of assistance energy transfers (AET).** This section includes new analysis on AET costs as well as a review of widespread reliance on AET during tight west-wide conditions.
- Section 5 **summarizes uncertainty used in the flexible ramp sufficiency test.**
- Section 6 **provides an overview of demand differences that can exist between the real-time market and resource sufficiency evaluation.** CAISO non-participating pump load is included in the real-time market but not in the resource sufficiency evaluation.
- Section 7 **summarizes WEIM import limits and transfers following a resource sufficiency evaluation failure.**
- Appendix A **provides a technical overview of the flexible ramp sufficiency and bid range capacity tests.**
- Appendix B **provides an overview of the mosaic quantile regression method for calculating uncertainty.**

DMM welcomes 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.

2 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.[1] 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 2.1 through Figure 2.4 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.[2] Figure 2.5 through Figure 2.8 provide the same information for the downward direction. The dash indicates that the area did not fail the test during the month.

In the fourth quarter:

- Public Service Company of New Mexico (PNM) failed the upward flexibility test relatively frequently, in around 2.5 percent of intervals. This was most common during November (around 7 percent of intervals). PNM also failed the upward capacity test in around 1.2 percent of intervals.
- All other balancing areas failed each test type in less than one percent of intervals.

Figure 2.9 shows the change in the percent of intervals with an upward test failure from the fourth quarter of 2023 to the fourth quarter of 2024. Figure 2.10 shows the same information for downward test failures.

Figure 2.11 summarizes the overlap between failure of the upward capacity and the flexibility tests during the quarter. 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 2.12 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.

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

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

Figure 2.1 Frequency of upward capacity test failures (percent of 15-minute intervals)

	2023			2024											
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Arizona Publ. Serv.	0.0	—	0.1	—	—	—	—	—	—	—	—	—	—	—	0.1
Avangrid	—	—	—	—	—	—	—	—	—	0.1	—	—	—	—	—
Avista	0.0	0.1	—	0.3	0.1	—	—	—	—	0.1	—	—	0.1	—	—
BANC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	—	—	—	0.3	—	—	—	—	—	—	—	—	—	—	—
California ISO	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
El Paso Electric	0.1	—	—	—	—	0.1	0.2	0.6	0.1	0.3	0.1	0.0	—	0.1	0.0
Idaho Power	—	0.1	—	0.0	—	—	—	—	—	—	—	—	—	—	—
LADWP	—	—	0.0	0.1	0.0	—	0.0	0.0	—	0.1	0.3	—	—	—	—
NorthWestern En.	—	—	—	—	0.1	—	—	—	—	—	—	0.3	—	—	0.3
NV Energy	0.0	—	—	—	—	—	—	0.1	0.0	0.1	0.0	—	—	—	—
PacifiCorp East	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PacifiCorp West	—	—	—	0.8	0.0	—	0.1	0.0	—	—	—	—	0.1	0.3	0.0
Portland Gen. Elec.	0.0	0.6	—	—	—	—	—	0.0	0.1	0.0	—	—	—	—	—
Powerex	0.0	0.0	—	—	—	—	—	—	—	—	—	—	—	—	—
PSC of New Mexico	0.1	—	0.1	—	—	—	0.1	0.1	0.1	0.3	0.1	—	0.4	3.1	—
Puget Sound En.	0.7	1.0	0.2	0.8	0.1	0.2	0.3	0.2	—	0.2	0.1	—	0.1	—	0.1
Salt River Proj.	0.8	0.2	0.1	0.1	0.1	0.2	0.1	—	0.2	0.1	0.1	0.2	0.1	—	—
Seattle City Light	0.1	0.6	—	0.5	—	—	0.4	—	0.0	0.4	0.1	0.1	0.3	0.0	0.1
Tacoma Power	0.1	0.0	—	—	—	0.3	—	0.0	—	—	—	—	0.0	0.1	—
Tucson Elec. Pow.	0.2	—	—	—	—	—	—	—	—	0.0	—	0.0	—	—	—
Turlock Irrig. Dist.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
WAPA DSW	0.3	0.4	0.1	—	—	0.1	—	0.5	0.3	0.2	0.2	—	0.1	—	—

Figure 2.2 Average shortfall of upward capacity test failures (MW)

	2023			2024											
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Arizona Publ. Serv.	58	—	160	—	—	—	—	—	—	—	—	—	—	—	62
Avangrid	—	—	—	—	—	—	—	—	—	396	—	—	—	—	—
Avista	2	13	—	31	9	—	—	—	—	11	—	—	8	—	—
BANC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	—	—	—	176	—	—	—	—	—	—	—	—	—	—	—
California ISO	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
El Paso Electric	18	—	—	—	—	4	3	15	14	27	17	12	—	9	4
Idaho Power	—	58	—	4	—	—	—	—	—	—	—	—	—	—	—
LADWP	—	—	10	71	1	—	19	6	—	54	61	—	—	—	—
NorthWestern En.	—	—	—	—	12	—	—	—	—	—	—	40	—	—	16
NV Energy	12	—	—	—	—	—	—	39	38	55	57	—	—	—	—
PacifiCorp East	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PacifiCorp West	—	—	—	51	8	—	25	8	—	—	—	—	89	66	15
Portland Gen. Elec.	17	228	—	—	—	—	—	9	12	10	—	—	—	—	—
Powerex	2	6	—	—	—	—	—	—	—	—	—	—	—	—	—
PSC of New Mexico	48	—	49	—	—	—	46	8	28	48	22	—	135	38	—
Puget Sound En.	48	89	41	78	15	52	53	18	—	71	27	—	14	—	6
Salt River Proj.	56	23	10	17	38	22	29	—	233	76	24	68	29	—	—
Seattle City Light	5	563	—	18	—	—	9	—	0	7	4	7	10	12	19
Tacoma Power	5	0	—	—	—	119	—	2	—	—	—	—	2	3	—
Tucson Elec. Pow.	12	—	—	—	—	—	—	—	—	16	—	8	—	—	—
Turlock Irrig. Dist.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
WAPA DSW	7	282	78	—	—	1	—	5	6	7	13	—	4	—	—

Figure 2.3 Frequency of upward flexibility test failures (percent of 15-minute intervals)

	2023			2024											
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Arizona Publ. Serv.	—	0.2	0.1	0.2	0.1	0.5	0.1	0.3	—	0.0	0.0	—	—	—	0.1
Avangrid	0.1	0.1	0.2	0.2	0.1	0.1	0.0	0.2	0.5	0.2	—	0.1	0.5	0.1	0.4
Avista	0.1	0.1	—	0.1	—	0.1	—	—	—	0.1	—	—	0.0	—	—
BANC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	0.1	—	—	0.4	0.0	—	0.1	0.1	0.1	0.3	0.3	0.0	—	0.1	0.1
California ISO	—	—	—	—	—	—	—	—	—	0.0	—	—	—	—	—
El Paso Electric	0.4	0.2	0.1	0.3	0.0	1.0	0.9	1.0	0.9	0.6	0.8	0.3	0.2	0.3	0.4
Idaho Power	0.1	—	—	1.1	—	0.1	0.6	0.6	0.1	0.1	—	—	0.0	—	—
LADWP	—	—	0.1	0.1	—	0.1	0.4	0.1	0.0	0.3	0.3	0.0	—	0.1	—
NorthWestern En.	0.2	0.0	0.1	0.5	0.1	0.0	0.0	0.1	0.3	0.2	—	0.4	0.1	0.2	0.2
NV Energy	—	0.1	0.0	—	0.1	0.0	—	0.1	—	—	—	—	—	—	—
PacifiCorp East	—	—	—	—	—	—	0.0	0.0	—	—	0.1	—	0.1	0.0	—
PacifiCorp West	0.0	0.0	0.1	1.0	—	0.1	—	—	0.1	—	—	—	—	0.3	0.1
Portland Gen. Elec.	0.6	0.0	—	—	—	0.0	—	0.2	0.2	—	—	0.0	0.1	—	—
Powerex	—	—	—	0.2	—	—	—	—	—	0.6	—	—	—	—	—
PSC of New Mexico	1.9	1.9	0.3	2.0	2.3	0.4	1.8	1.1	1.2	1.0	1.0	0.9	0.3	7.1	0.2
Puget Sound En.	1.3	1.9	0.5	0.8	0.1	0.2	0.4	0.5	0.5	0.7	0.3	—	0.4	—	0.5
Salt River Proj.	0.6	0.4	0.2	0.2	0.1	0.7	0.4	0.1	0.3	0.3	0.4	0.5	0.2	—	—
Seattle City Light	0.0	—	—	0.3	—	0.1	0.1	0.1	—	—	—	0.0	0.1	0.1	—
Tacoma Power	0.2	0.0	—	0.1	0.0	0.4	0.0	0.0	—	—	—	—	—	0.1	0.0
Tucson Elec. Pow.	0.1	0.2	0.1	0.0	0.2	—	0.1	0.1	—	0.1	0.3	0.7	0.2	0.1	0.1
Turlock Irrig. Dist.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
WAPA DSW	0.3	0.5	0.1	1.1	2.5	3.5	0.3	0.8	0.2	—	—	—	0.2	—	0.1

Figure 2.4 Average shortfall of upward flexibility test failures (MW)

	2023			2024											
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Arizona Publ. Serv.	—	102	23	27	55	65	35	83	—	208	33	—	—	—	139
Avangrid	60	8	7	12	20	30	19	15	20	273	—	17	108	27	127
Avista	21	15	—	66	—	14	—	—	—	8	—	—	10	—	—
BANC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	5	—	—	153	73	—	5	9	47	49	43	32	—	47	28
California ISO	—	—	—	—	—	—	—	—	—	388	—	—	—	—	—
El Paso Electric	17	10	13	9	24	23	17	17	14	10	15	21	13	7	6
Idaho Power	10	—	—	32	—	17	27	28	33	17	—	—	7	—	—
LADWP	—	—	69	26	—	52	200	50	48	32	67	74	—	75	—
NorthWestern En.	11	4	5	16	24	3	11	12	26	72	—	40	18	18	22
NV Energy	—	22	19	—	137	136	—	80	—	—	—	—	—	—	—
PacifiCorp East	—	—	—	—	—	—	27	44	—	—	27	—	32	59	—
PacifiCorp West	22	25	22	104	—	9	—	—	6	—	—	—	—	177	13
Portland Gen. Elec.	25	2	—	—	—	23	—	44	22	—	—	12	7	—	—
Powerex	—	—	—	106	—	—	—	—	—	1345	—	—	—	—	—
PSC of New Mexico	56	38	39	52	50	37	48	43	83	55	55	67	107	88	48
Puget Sound En.	86	42	27	29	16	71	12	17	22	53	45	—	32	—	18
Salt River Proj.	43	63	151	66	50	44	76	92	209	36	103	52	37	—	—
Seattle City Light	6	—	—	22	—	16	19	6	—	—	—	5	16	21	—
Tacoma Power	9	2	—	6	7	73	2	13	—	—	—	—	—	5	8
Tucson Elec. Pow.	13	13	55	6	13	—	10	25	—	12	15	28	14	7	15
Turlock Irrig. Dist.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
WAPA DSW	16	143	12	12	27	33	18	12	18	—	—	—	53	—	53

Figure 2.5 Frequency of downward capacity test failures (percent of 15-minute intervals)

	2023			2024											
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Arizona Publ. Serv.	—	—	0.8	0.1	0.0	0.1	0.2	—	—	—	—	0.4	—	—	—
Avangrid	—	0.3	—	—	—	—	—	—	—	—	—	—	—	—	—
Avista	—	—	—	—	—	0.1	—	—	—	—	—	—	0.0	—	—
BANC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0.1
California ISO	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
El Paso Electric	—	—	—	0.2	—	0.4	0.2	0.4	0.3	0.1	0.1	0.0	0.1	0.0	—
Idaho Power	—	—	—	—	—	—	0.5	—	—	—	—	—	—	—	—
LADWP	—	—	—	—	—	—	—	—	—	—	—	0.0	—	—	—
NorthWestern En.	—	—	—	—	—	—	—	—	—	—	—	0.1	—	0.0	—
NV Energy	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PacifiCorp East	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PacifiCorp West	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Portland Gen. Elec.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Powerex	—	—	—	—	—	—	0.0	—	—	—	—	—	—	0.1	—
PSC of New Mexico	—	—	—	—	—	—	—	—	—	0.1	—	—	—	0.2	—
Puget Sound En.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Salt River Proj.	0.1	—	—	—	0.1	0.1	0.4	0.7	—	—	0.2	—	—	—	0.3
Seattle City Light	—	0.1	0.2	0.0	—	—	—	—	—	0.3	—	—	—	0.1	0.0
Tacoma Power	—	—	—	—	—	—	—	—	0.0	0.1	0.0	—	—	—	—
Tucson Elec. Pow.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Turlock Irrig. Dist.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
WAPA DSW	0.2	0.2	—	—	—	—	—	—	—	0.1	—	—	—	—	—

Figure 2.6 Average shortfall of downward capacity test failures (MW)

	2023			2024											
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Arizona Publ. Serv.	—	—	176	18	16	9	100	—	—	—	—	69	—	—	—
Avangrid	—	93	—	—	—	—	—	—	—	—	—	—	—	—	—
Avista	—	—	—	—	—	264	—	—	—	—	—	—	1	—	—
BANC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	—	—	—	—	—	—	—	—	—	—	—	—	—	—	105
California ISO	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
El Paso Electric	—	—	—	4	—	7	6	10	20	5	15	6	7	1	—
Idaho Power	—	—	—	—	—	—	12	—	—	—	—	—	—	—	—
LADWP	—	—	—	—	—	—	—	—	—	—	—	7	—	—	—
NorthWestern En.	—	—	—	—	—	—	—	—	—	—	—	4	—	1	—
NV Energy	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PacifiCorp East	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
PacifiCorp West	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Portland Gen. Elec.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Powerex	—	—	—	—	—	—	25	—	—	—	—	—	—	86	—
PSC of New Mexico	—	—	—	—	—	—	—	—	—	5	—	—	—	186	—
Puget Sound En.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Salt River Proj.	46	—	—	—	20	2	21	33	—	—	14	—	—	—	24
Seattle City Light	—	15	3	1	—	—	—	—	—	8	—	—	—	4	1
Tacoma Power	—	—	—	—	—	—	—	—	2	2	0	—	—	—	—
Tucson Elec. Pow.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Turlock Irrig. Dist.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
WAPA DSW	6	2	—	—	—	—	—	—	—	7	—	—	—	—	—

Figure 2.7 Frequency of downward flexibility test failures (percent of 15-minute intervals)

	2023			2024											
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Arizona Publ. Serv.	—	—	0.3	0.1	0.1	0.2	0.1	—	—	—	—	—	—	0.3	0.2
Avangrid	—	—	—	0.1	—	—	—	—	—	—	—	—	—	—	—
Avista	—	0.1	—	—	0.0	—	—	—	—	0.1	—	0.0	0.6	—	—
BANC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	0.2	—	—	0.4	0.1	—	0.0	0.1	0.1	—	—	—	—	—	0.8
California ISO	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
El Paso Electric	—	0.2	0.3	0.3	0.2	0.4	0.8	0.7	0.1	—	0.1	—	0.0	—	0.2
Idaho Power	—	0.1	—	—	—	0.0	1.0	—	—	—	—	—	—	—	—
LADWP	—	—	—	—	—	—	—	—	—	—	—	0.1	—	—	—
NorthWestern En.	—	—	—	0.2	—	0.1	—	0.3	0.2	0.2	0.1	0.0	2.2	0.2	0.1
NV Energy	0.1	0.1	—	—	—	0.1	0.0	—	0.1	—	—	—	—	—	—
PacifiCorp East	0.1	—	—	—	0.2	0.0	0.5	0.2	0.0	0.0	—	0.1	—	—	—
PacifiCorp West	—	0.1	—	—	—	0.2	—	—	—	—	—	—	0.0	0.0	—
Portland Gen. Elec.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Powerex	0.1	—	0.1	—	0.1	0.4	0.0	—	—	1.1	0.2	—	—	0.1	—
PSC of New Mexico	0.4	0.2	0.2	0.9	0.9	0.4	0.0	0.6	0.1	0.1	0.0	0.9	0.3	2.0	0.1
Puget Sound En.	—	—	—	—	—	—	—	—	—	0.1	—	—	—	—	—
Salt River Proj.	—	0.1	0.0	0.1	0.1	0.7	0.7	0.7	0.0	—	—	—	0.1	—	0.5
Seattle City Light	—	0.8	0.2	0.2	0.1	0.1	0.2	—	0.1	0.5	0.1	—	0.0	0.2	0.1
Tacoma Power	—	0.0	—	—	0.0	—	—	—	—	—	—	—	—	—	—
Tucson Elec. Pow.	—	—	—	—	0.1	—	—	—	—	—	—	—	—	—	—
Turlock Irrig. Dist.	—	0.1	—	—	0.0	—	—	0.2	0.0	—	0.0	—	0.1	0.1	—
WAPA DSW	0.8	0.2	0.1	0.3	0.1	0.0	0.0	—	—	0.1	0.0	0.1	—	0.0	—

Figure 2.8 Average shortfall of downward flexibility test failures (MW)

	2023			2024											
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Arizona Publ. Serv.	—	—	84	72	53	116	94	—	—	—	—	—	—	88	102
Avangrid	—	—	—	18	—	—	—	—	—	—	—	—	—	—	—
Avista	—	27	—	—	1	—	—	—	—	25	—	21	27	—	—
BANC	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
BPA	192	—	—	243	104	—	36	190	125	—	—	—	—	—	164
California ISO	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
El Paso Electric	—	7	8	7	7	10	31	11	9	—	9	—	1	—	34
Idaho Power	—	4	—	—	—	3	89	—	—	—	—	—	—	—	—
LADWP	—	—	—	—	—	—	—	—	—	—	—	34	—	—	—
NorthWestern En.	—	—	—	22	—	31	—	21	8	84	8	4	20	29	7
NV Energy	59	156	—	—	—	94	23	—	279	—	—	—	—	—	—
PacifiCorp East	8	—	—	—	36	35	68	76	125	118	—	98	—	—	—
PacifiCorp West	—	6	—	—	—	30	—	—	—	—	—	—	18	5	—
Portland Gen. Elec.	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
Powerex	421	—	160	—	84	2528	5	—	—	1225	620	—	—	192	—
PSC of New Mexico	36	20	44	55	37	21	42	59	9	109	14	82	74	77	75
Puget Sound En.	—	—	—	—	—	—	—	—	—	16	—	—	—	—	—
Salt River Proj.	—	41	1	44	27	62	36	59	23	—	—	—	30	—	80
Seattle City Light	—	45	8	64	4	9	37	—	5	9	11	—	13	30	4
Tacoma Power	—	2	—	—	2	—	—	—	—	—	—	—	—	—	—
Tucson Elec. Pow.	—	—	—	—	94	—	—	—	—	—	—	—	—	—	—
Turlock Irrig. Dist.	—	39	—	—	1	—	—	5	2	—	3	—	6	3	—
WAPA DSW	14	8	8	66	22	16	3	—	—	9	4	9	—	8	—

Figure 2.9 Change in percent of intervals with an upward resource sufficiency evaluation failure (Q4 2023 to Q4 2024)

	Flexibility test			Capacity test		
WEIM entity	**Q4 2023**	**Q4 2024**	**Difference**	**Q4 2023**	**Q4 2024**	**Difference**
Arizona Publ. Serv.	0.1%	0.0%	-0.1%	0.0%	0.0%	0.0%
Avangrid	0.1%	0.3%	0.2%	0%	0%	0%
Avista	0.1%	0.0%	0.0%	0.1%	0.0%	0.0%
BANC	0%	0%	0%	0%	0%	0%
BPA	0.0%	0.1%	0.0%	0%	0%	0%
California ISO	0%	0%	0%	0%	0%	0%
El Paso Electric	0.2%	0.3%	0.1%	0.0%	0.0%	0.0%
Idaho Power	0.0%	0.0%	0.0%	0.0%	0%	0.0%
LADWP	0.0%	0.0%	0.0%	0.0%	0%	0.0%
NorthWestern En.	0.1%	0.2%	0.1%	0%	0.1%	0.1%
NV Energy	0.0%	0%	0.0%	0.0%	0%	0.0%
PacifiCorp East	0%	0.0%	0.0%	0%	0%	0%
PacifiCorp West	0.0%	0.1%	0.1%	0%	0.2%	0.2%
Portland Gen. Elec.	0.2%	0.0%	-0.2%	0.2%	0%	-0.2%
Powerex	0%	0%	0%	0.0%	0%	0.0%
PSC of New Mexico	1.4%	2.5%	1.2%	0.1%	1.2%	1.1%
Puget Sound En.	1.2%	0.3%	-1.0%	0.6%	0.1%	-0.6%
Salt River Proj.	0.4%	0.1%	-0.3%	0.4%	0.0%	-0.3%
Seattle City Light	0.0%	0.1%	0.0%	0.2%	0.2%	-0.1%
Tacoma Power	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%
Tucson Elec. Pow.	0.1%	0.1%	0.0%	0.1%	0%	-0.1%
Turlock Irrig. Dist.	0%	0%	0%	0%	0%	0%
WAPA DSW	0.3%	0.1%	-0.2%	0.3%	0.0%	-0.3%

Figure 2.10 Change in percent of intervals with a downward resource sufficiency evaluation failure (Q4 2023 to Q4 2024)

	Flexibility test			Capacity test		
WEIM entity	Q4 2023	Q4 2024	Difference	Q4 2023	Q4 2024	Difference
Arizona Publ. Serv.	0.1%	0.2%	0.1%	0.3%	0%	-0.3%
Avangrid	0%	0%	0%	0.1%	0%	-0.1%
Avista	0.0%	0.2%	0.2%	0%	0.0%	0.0%
BANC	0%	0%	0%	0%	0%	0%
BPA	0.1%	0.3%	0.2%	0%	0.0%	0.0%
California ISO	0%	0%	0%	0%	0%	0%
El Paso Electric	0.1%	0.1%	-0.1%	0%	0.0%	0.0%
Idaho Power	0.0%	0%	0.0%	0%	0%	0%
LADWP	0%	0%	0%	0%	0%	0%
NorthWestern En.	0%	0.8%	0.8%	0%	0.0%	0.0%
NV Energy	0.1%	0%	-0.1%	0%	0%	0%
PacifiCorp East	0.0%	0%	0.0%	0%	0%	0%
PacifiCorp West	0.0%	0.0%	0.0%	0%	0%	0%
Portland Gen. Elec.	0%	0%	0%	0%	0%	0%
Powerex	0.1%	0.0%	0.0%	0%	0.0%	0.0%
PSC of New Mexico	0.3%	0.8%	0.5%	0%	0.1%	0.1%
Puget Sound En.	0%	0%	0%	0%	0%	0%
Salt River Proj.	0.0%	0.2%	0.2%	0.0%	0.1%	0.1%
Seattle City Light	0.3%	0.1%	-0.2%	0.1%	0.0%	-0.1%
Tacoma Power	0.0%	0%	0.0%	0%	0%	0%
Tucson Elec. Pow.	0%	0%	0%	0%	0%	0%
Turlock Irrig. Dist.	0.0%	0.1%	0.0%	0%	0%	0%
WAPA DSW	0.4%	0.0%	-0.3%	0.1%	0%	-0.1%

Figure 2.11 Upward capacity/flexibility test failure intervals by concurrence (October–December 2024)

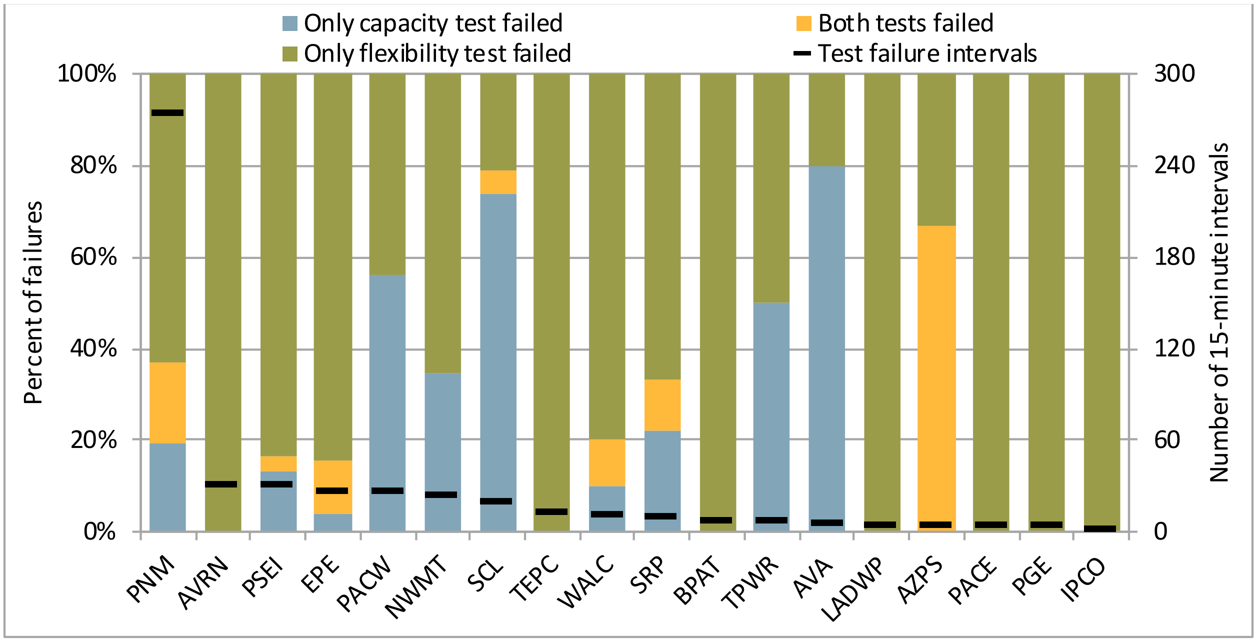
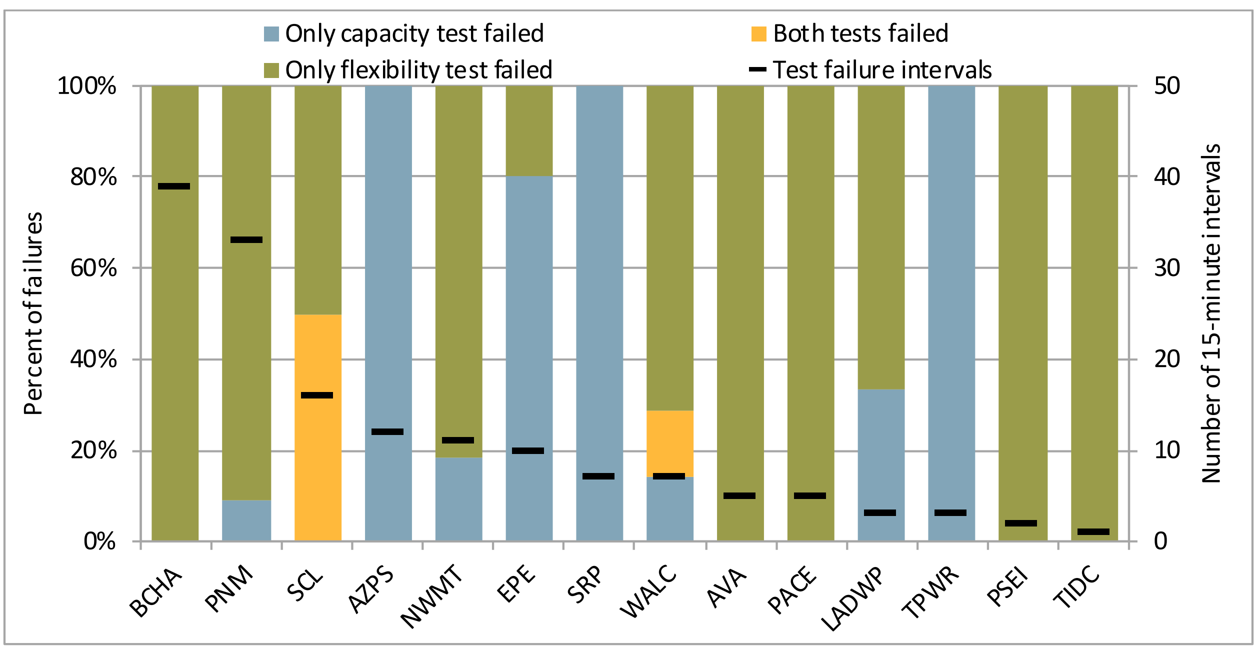


Figure 2.12 Downward capacity/flexibility test failure intervals by concurrence (October–December 2024)



3 Impact of advisory resource sufficiency evaluation runs

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. The real-time market enforces 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 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 3.1 and Figure 3.2 summarize the *first* interval of each evaluation hour during the quarter and the frequency of a failure in the second (T-55) or final (T-40) resource sufficiency evaluation.[3] This reflects failure of *either* the flexibility or capacity test in the second or final run. The red and yellow bars show instances 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.

[3] 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 3.1 Upward resource sufficiency evaluation failures in first 15-minute interval of hour (October–December 2024)

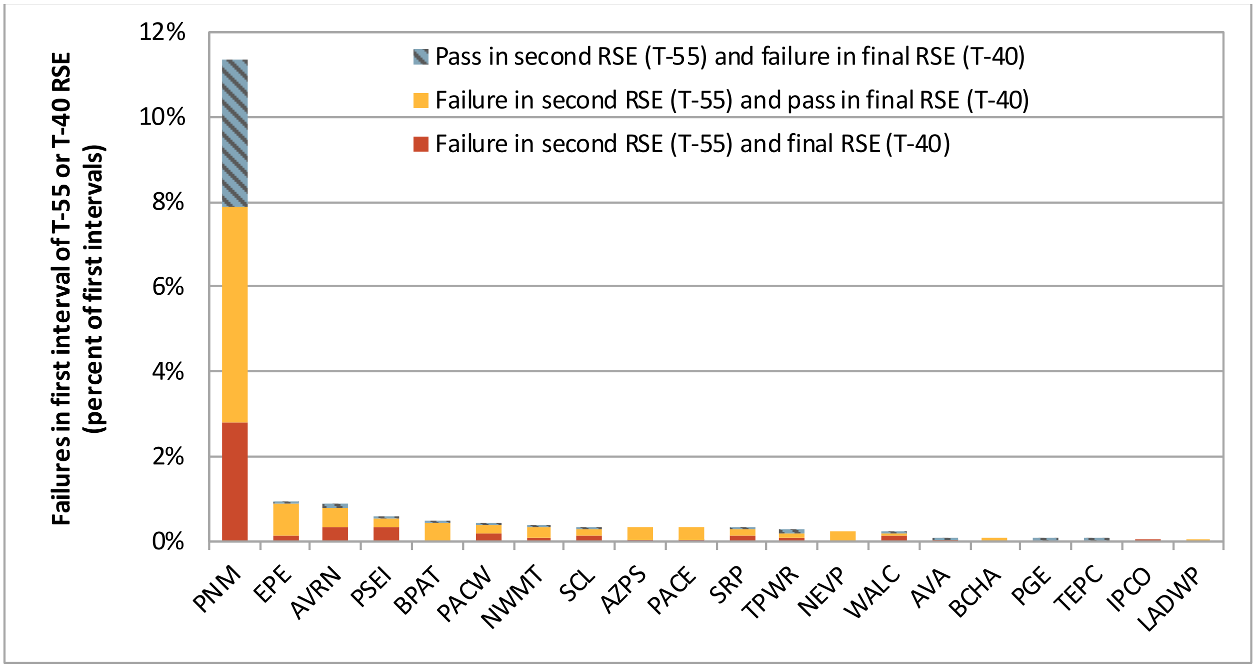
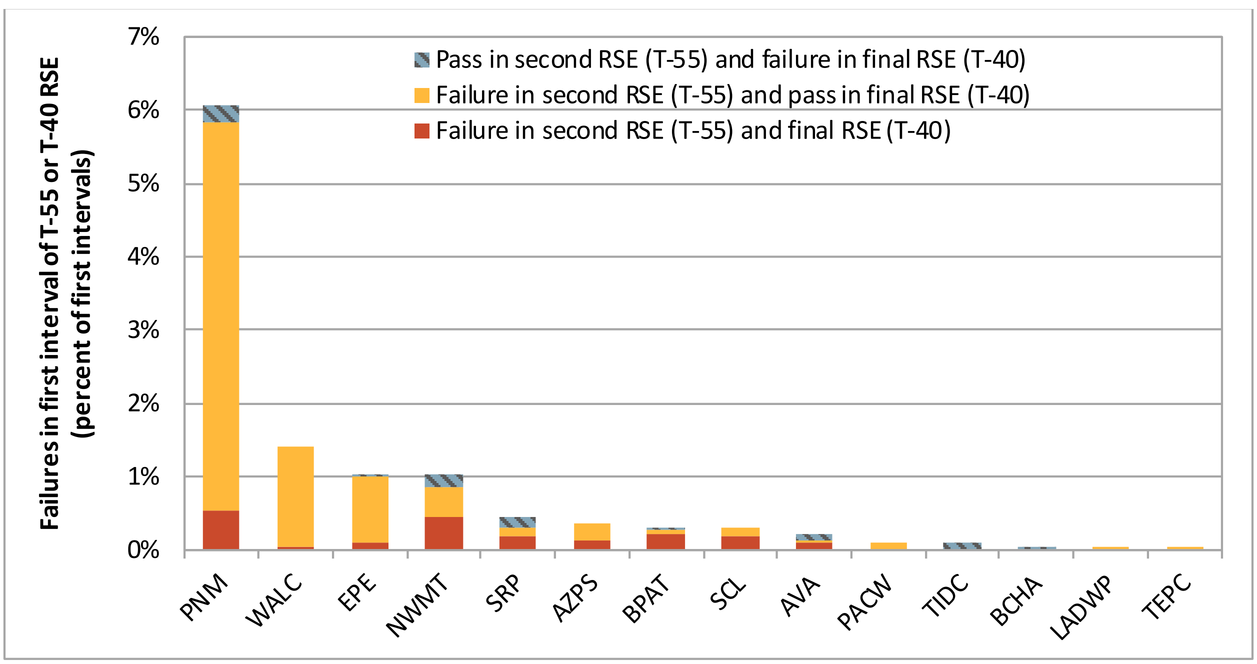


Figure 3.2 Downward resource sufficiency evaluation failures in first 15-minute interval of hour (October–December 2024)



Improvement for calculating uncertainty within the group of balancing areas that pass the tests

The ISO made improvements to the set of balancing areas considered in the pass-group for performing the regressions that estimate the relationship between forecast information and uncertainty.

As part of the enhancements implemented on February 1, 2023, uncertainty is calculated based on regression results that use historical data to predict uncertainty relative to load, solar, and wind forecasts. 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 in each interval so that it can perform the regression using historical data accordingly for that group.

To perform the regressions for the pass-group, the set of balancing areas in this group is sometimes estimated from preliminary test results based on information available at the time of this process. Then in the present, when the current forecast information is combined with the regression information to calculate uncertainty, a different set of balancing areas in the pass-group may be used based on changes in the results of the later resource sufficiency evaluation runs.

On June 25, 2024 the ISO made an improvement to the timing in which the resource sufficiency evaluation results are pushed in advance of the regressions that are performed to calculate pass-group uncertainty. In some intervals, the regressions for calculating the uncertainty requirement for the pass-group must be performed before the final set of balancing +areas in this group are known. The enhancement improved the consistency between (1) the group of balancing areas used to determine the regression coefficients for the pass-group and (2) the group of balancing areas whose forecast information gets combined with those coefficients to determine the uncertainty requirement.

Table 3.1 summarizes this inconsistency and the improvement made on June 25. The set of balancing areas in the pass-group for the current weather information that is ultimately combined with the regression results to calculate uncertainty and procure flexible capacity, 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 through 4. However, prior to June 25, *the regressions* were based on the results from the earliest resource sufficiency evaluation (T-75) to define the pass-group for the first interval of each hour, while the results from the second resource sufficiency evaluation (T-55) were used to define the pass-group for the second interval of each hour.

Starting on June 25, 2024 the set of balancing areas in the pass-group between the regression information and the current forecast information became more consistent. For the second interval of each hour, the regressions now use the results from the final resource sufficiency evaluation (consistent with forecast information). For the first interval of each hour, the regressions now use the results from the first or second resource sufficiency evaluation depending on the timing of various market processes (sometimes consistent with forecast information). DMM recommends that additional improvements be made to resolve inconsistencies in the set of balancing areas in the pass-group for the first interval of each hour.

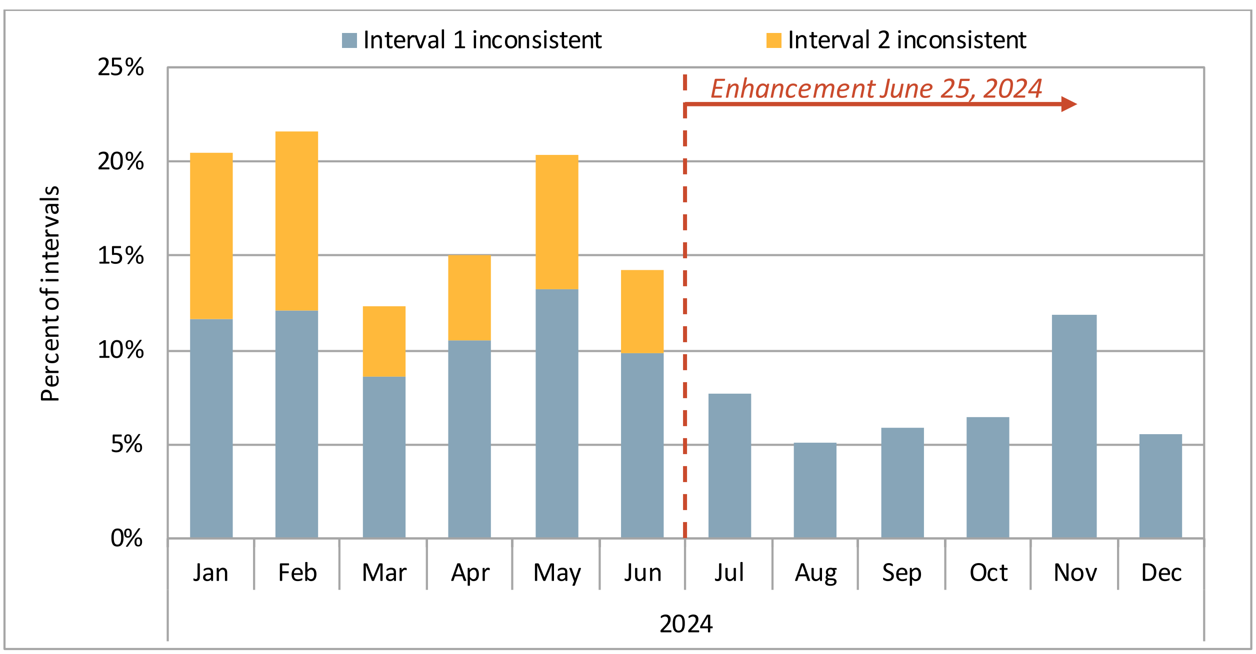
Table 3.1 Source of pass-group for calculating uncertainty and procuring flexible ramping capacity (prior to and after June 25, 2024)

15-minute market interval	Regression inputs and outputs		Current weather information for calculating uncertainty and flex ramp procurement
	(*prior to* June 25, 2024)	(*after* June 25, 2024)	
1	First run (T-75)	First run (T-75) or second run (T-55)	Second run (T-55)
2	Second run (T-55)	Final run (T-40)	Final run (T-40)
3	Final run (T-40)	Final run (T-40)	Final run (T-40)
4	Final run (T-40)	Final run (T-40)	Final run (T-40)

Using an inconsistent set 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 any of the historical data.

Figure 3.3 shows the percent of intervals by month in which the set of balancing areas in the pass-group differed between the regression information and current forecast information. The figure also shows whether it was the first or second interval of the hour that had the inconsistency. The enhancement removed the potential for inconsistency in interval 2 and improved the consistency in interval 1. Following the enhancements, the set of balancing areas in the pass-group differed in around 7 percent of intervals, compared to around 18 percent of intervals prior to the enhancements in 2024.

Figure 3.3 Percent of intervals in which the set of balancing areas in the pass-group differed between the current forecast information and regression information



4 Assistance energy transfers

Assistance energy transfers (AET) give 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.[4]

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.[5] The applicable AET quantity is based on the lesser of either (1) the dynamic WEIM transfers or (2) the amount by which the balancing area failed the resource sufficiency evaluation. If the 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 ISO balancing area.

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.

Table 4.1 shows the days in which a balancing area was opted in to receiving assistance energy transfers during the quarter. Eight balancing areas were opted in to the program on at least one day during this period: Avangrid, CAISO, Idaho Power, NorthWestern Energy, NV Energy, PacifiCorp East, PacifiCorp West, and WAPA Desert Southwest.[6] Avangrid, NorthWestern Energy, and NV Energy were opted in to AET during all days during the quarter (92 days).

[4] 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

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

[6] The CAISO balancing area can opt in to assistance energy transfers based on upcoming system conditions and operator experience. For more information, see the Business Practice Manual for the Western Energy Imbalance Market, section 11.3.2: https://bpmcm.caiso.com/Pages/BPMDetails.aspx?BPM=Energy%20Imbalance%20Market

Table 4.1 Assistance energy transfer opt-in designations by balancing area (October–December 2024)

Balancing area	Period opted in to receiving assistance energy transfers	Days opted in to AET
Avangrid	Oct. 1 - Dec. 31	92
California ISO	Oct. 1 - Oct. 8, Nov. 7	9
Idaho Power	Oct. 1 - Oct. 31, Nov. 6 - Dec. 31	87
NorthWestern Energy	Oct. 1 - Dec. 31	92
NV Energy	Oct. 1 - Dec. 31	92
PacifiCorp East	Oct. 24 - Dec. 31	69
PacifiCorp West	Oct. 24 - Dec. 31	69
WAPA Desert Southwest	Oct. 1 - Oct. 15	15

Table 4.2 summarizes all balancing areas that were opted in to assistance energy transfers on at least one day during the quarter and the subsequent impact following a resource sufficiency evaluation failure. First, the table shows the number of 15-minute intervals in which a balancing area failed the resource sufficiency evaluation after opting in to AET. These are the intervals in which the WEIM import limit following the test failure was removed—giving the WEIM entity access to WEIM supply that may not have been available otherwise. During the quarter, six balancing areas (Avangrid, Idaho Power, NorthWestern Energy, PacifiCorp East, and PacifiCorp West) failed the resource sufficiency evaluation during at least one interval while opted in to the program.

Table 4.2 also shows the percent of failure intervals in the 5-minute market in which the balancing area achieved additional WEIM imports due to opting in to AET. The table also shows the average, maximum, and total WEIM imports added in the 5-minute market because of AET. During the quarter, Avangrid failed the resource sufficiency evaluation during 30 intervals while opted in to receiving assistance energy transfers. Avangrid achieved an additional 15 MW on average during these intervals (and maximum of 221 MW).

Table 4.3 summarizes the total cost from assistance energy transfers.[7] AET is settled during any interval in which the balancing area both opted in to receiving assistance energy transfers and failed the resource sufficiency evaluation. The applicable quantity that is settled for AET is based on the lower of the resource sufficiency evaluation insufficiency or the WEIM imports.[8] The price is the real-time bid cap, typically $1,000/MWh.

[7] Information is based on settlement values available at the time of drafting. Updates can occur regularly within the settlements timeline, starting with T+9B (trade date plus nine business days) and T+70B, as well as others up to 36 months after the trade date.

[8] If the 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 ISO balancing area.

Table 4.2 Resource sufficiency evaluation failures during assistance energy transfer opt-in (October–December 2024)

Balancing area	Days opted in to AET	RSE failures under AET (15-min. intervals)	Percent of failure intervals with additional WEIM imports due to AET	Average WEIM imports added (MW)	Max WEIM imports added (MW)	Total WEIM imports added (MWh)
Avangrid	92	30	17%	15	221	111
California ISO	9	0	N/A	N/A	N/A	N/A
Idaho Power	87	1	0%	0	0	0
NorthWestern Energy	92	23	15%	3	43	20
NV Energy	92	0	N/A	N/A	N/A	N/A
PacifiCorp East	69	1	0%	0	0	0
PacifiCorp West	69	21	40%	29	235	151
WAPA Desert Southwest	15	0	N/A	N/A	N/A	N/A

Table 4.3 Cost of assistance energy transfers (October–December 2024)

Balancing area	RSE failures under AET (15-min. intervals)	Total WEIM imports added (MWh)	Total cost of assistance energy transfers	Total cost per added WEIM imports
Avangrid	30	111	$20,412	$183
California ISO	0	N/A	N/A	N/A
Idaho Power	1	0	$601	*
NorthWestern Energy	23	20	$49,958	$2,517
NV Energy	0	N/A	N/A	N/A
PacifiCorp East	1	0	$14,655	*
PacifiCorp West	21	151	$157,240	$1,040
WAPA Desert Southwest	0	N/A	N/A	N/A

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, 2023 using a method called *mosaic quantile regression*. Details on the calculation are included in Appendix B. This section summarizes 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.

Thresholds for capping uncertainty

Uncertainty calculated from the quantile regressions is 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 updated each day and pulled for each hour from the 1st and 99th percentile of net load error observations from the 180-day period.[9] The seasonal threshold is updated each quarter and is calculated based on the 1st and 99th percentile using observations over the previous 90 days. Here, each hour is calculated separately, and the greatest upward and downward uncertainty across all hours sets the seasonal 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 seasonal or histogram threshold during the quarter. Averaging across all balancing areas, the thresholds capped the calculated upward and downward uncertainty in around 13 percent of intervals. In the large majority of cases with capped uncertainty, the *histogram* threshold capped the uncertainty.

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 quarter.

[9] The histogram threshold also uses symmetric sampling, from historical observations from the previous 90 days as well as the next 90 days minus one year.

Figure 5.1 Quantile regression uncertainty capped by mosaic or histogram thresholds (October–December 2024)

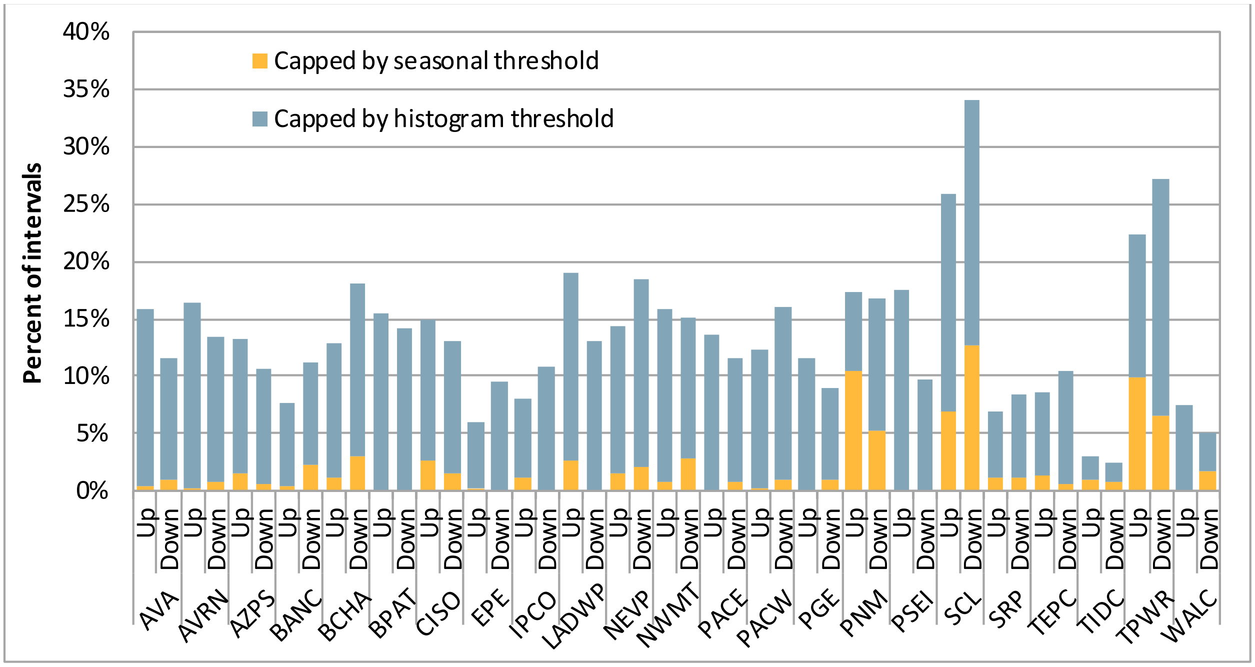
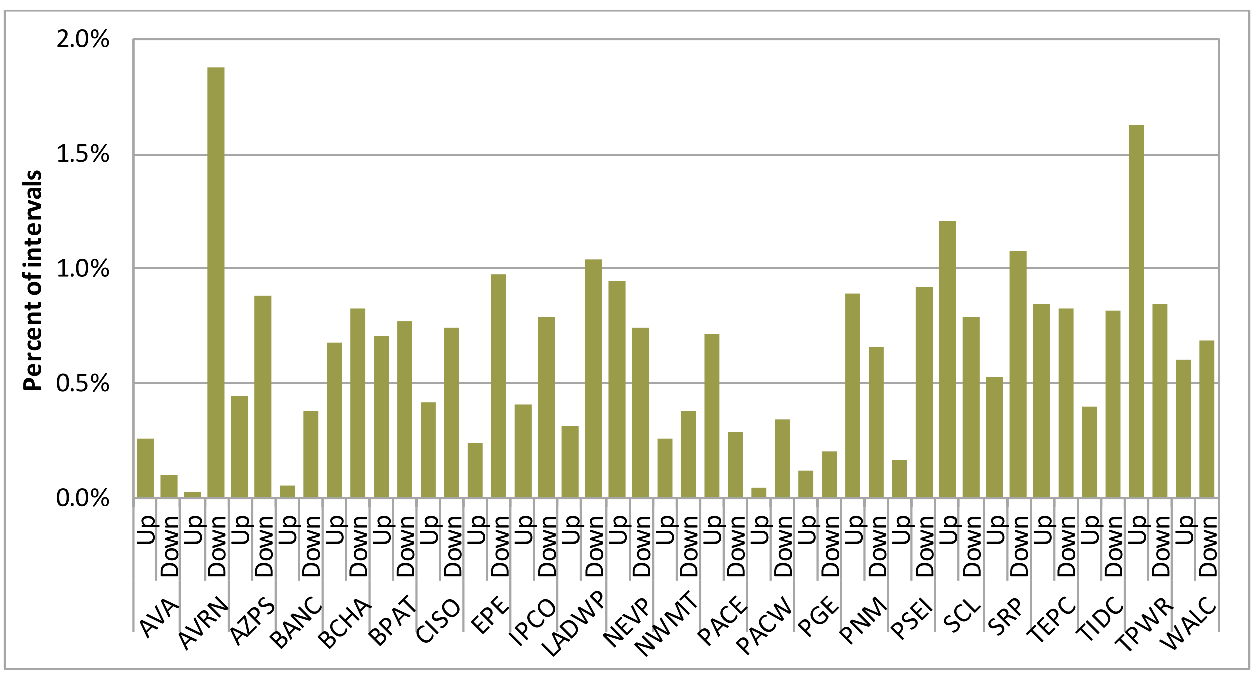


Figure 5.2 Quantile regression uncertainty set near zero by mosaic threshold (October–December 2024)



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

The regression model used for the resource sufficiency evaluation is currently designed to predict uncertainty in forecasts produced only 45 to 55 minutes before real-time. However, the time horizon of the resource sufficiency evaluation includes four intervals, typically produced between 47.5 and 102.5 minutes before real-time.

The resource sufficiency evaluation uses exactly the same underlying historical data to perform the regressions and calculate uncertainty as the flexible ramping product in the 15-minute market.[10] This data is based on the difference from advisory forecasts in the 15-minute market to the corresponding binding forecasts in the 5-minute market. The regressions use this data to produce hourly coefficients that define the relationship between the forecasts and uncertainty. This calculation reflects 45 to 55 minutes in which uncertainty may materialize between the applicable 15-minute and 5-minute market runs.

However, the resource sufficiency evaluation occurs over a different timeframe than what is considered for procuring 15-minute market flexible capacity. Figure 5.3 illustrates the timeframe of uncertainty considered for the flexible ramping product in the 15-minute market, and how it compares with the timeframe of the resource sufficiency evaluation.[11] For the flexible ramping product, the calculation is designed to capture uncertainty that may materialize around a single upcoming (advisory) interval. However, the resource sufficiency evaluation considers forecast information from four 15-minute intervals within an hour. When comparing the forecast values used in each interval of the resource sufficiency evaluation to corresponding 5-minute market intervals, there exists a larger gap of time for uncertainty to materialize.

In comparing the first 15-minute test interval of the RSE to corresponding 5-minute market intervals, the timeframe and potential for net load uncertainty to materialize is similar to the timeframe of the 15-minute market flexible ramping product uncertainty calculation. However, 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. The current determination of the regression coefficients for predicting net load uncertainty for the resource sufficiency evaluation (based on short-term historical data) does not capture the increased net load uncertainty associated with the longer-term horizon of this market process.[12]

This inconsistency results in lower performance in the rate of coverage provided by the uncertainty component in the resource sufficiency evaluation. Figure 5.4 shows the average coverage rate across all balancing areas by interval. Here, coverage is measured as the percent of intervals when realized uncertainty from the forecasts considered in the resource sufficiency evaluation to the 5-minute market forecasts fell within the calculated uncertainty requirement for the same interval. The calculated uncertainty covered the realized uncertainty much less for intervals at the end of the hour compared to

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

[11] The figure shows the time horizon for the resource sufficiency evaluation ran 55 minutes prior to the hour (T-55 RSE). While the final test is run at 40 minutes prior to the hour, the load and renewable forecasts used in the final test are held fixed from the forecasts in the T-55 RSE. This is intended to reduce unexpected failures that would be caused by forecast variation between the T-55 and T-40 resource sufficiency evaluations.

[12] The resource sufficiency evaluation and flexible ramping product uncertainty calculations for a single balancing area use the same hourly regression coefficients (produced from the same short-term historical data) but are combined with the current forecast information at the time of each market process to determine the final uncertainty. Here, longer-term forecast information at the time of the resource sufficiency evaluation is combined with the short-term regression coefficients.

the beginning of the hour because the current calculation is not designed to capture uncertainty that can realize over a longer-term horizon.

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

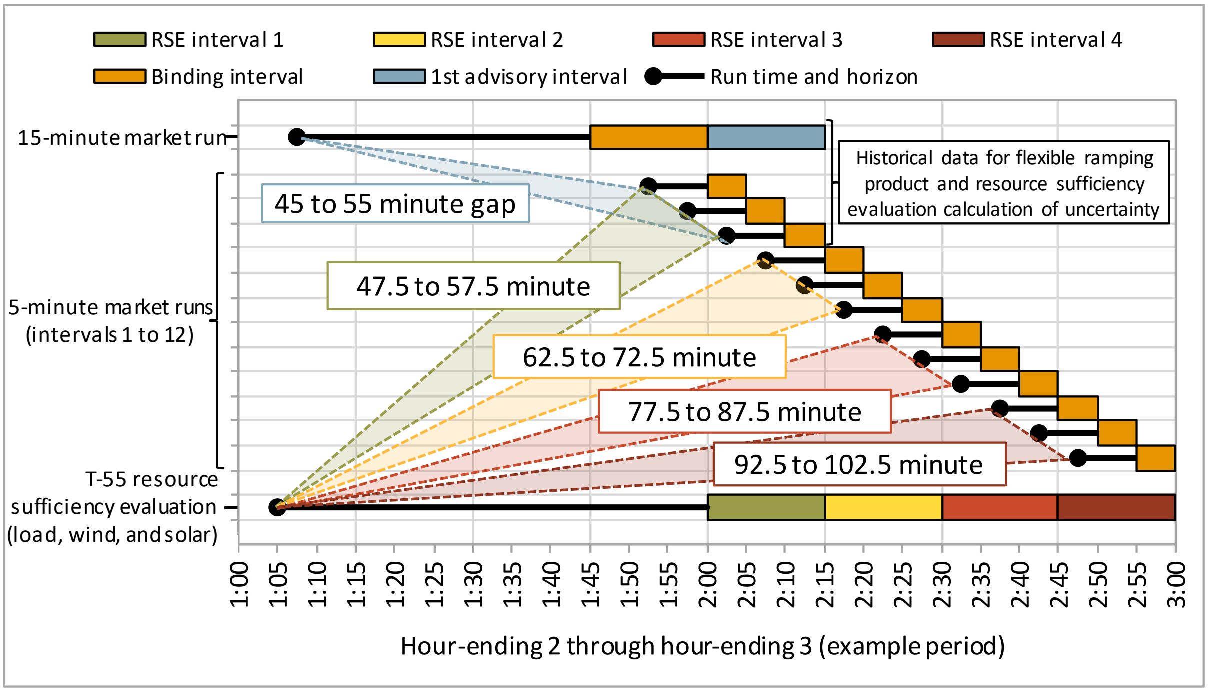
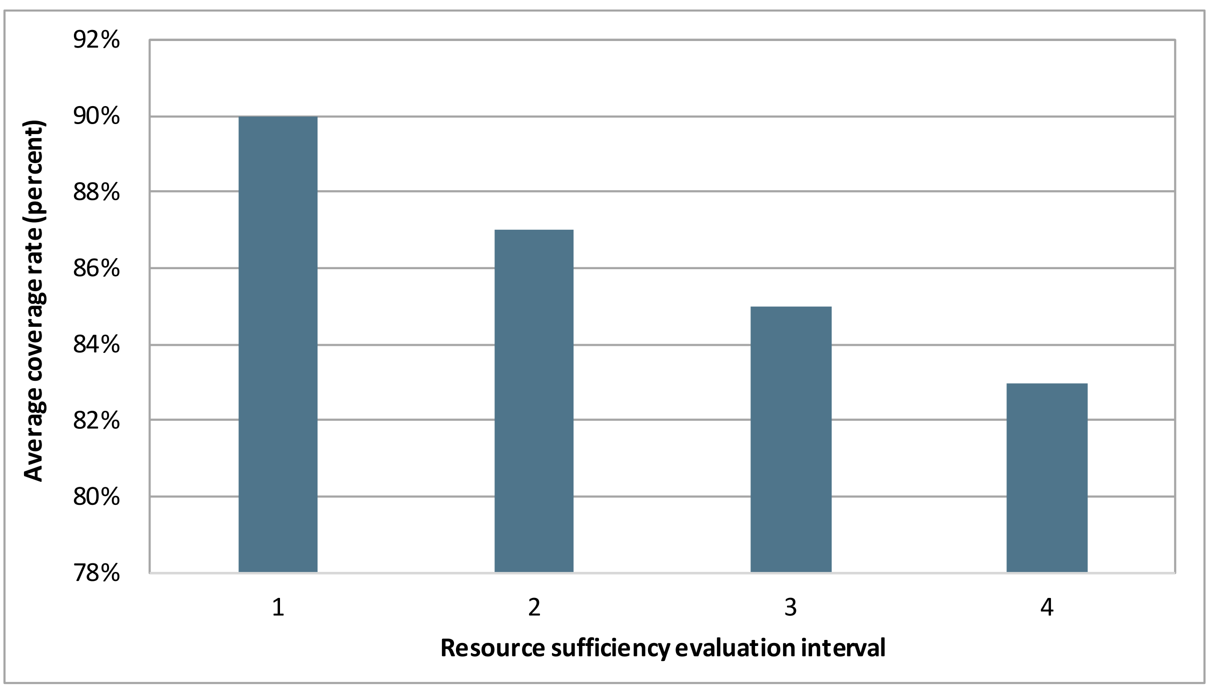


Figure 5.4 Average coverage rate by resource sufficiency evaluation interval (October–December 2024)



Results of quantile regression uncertainty in the resource sufficiency evaluation

Figure 5.5 through Figure 5.27 show 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 for all balancing areas during the fourth quarter. 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 maximum amount over the quarter (range of uncertainty in each interval). The dashed red and yellow lines show the average histogram and seasonal thresholds, respectively, during the quarter.

Overall, the uncertainty outcomes from the mosaic quantile regression approach were often comparable to those calculated with the prior histogram approach. The mosaic quantile regression approach tends to be somewhat lower on average across most hours and balancing areas. However, results of the mosaic quantile regression approach vary more widely, including periods with zero uncertainty.

Figure 5.5 Arizona Public Service resource sufficiency evaluation uncertainty requirements (October–December 2024)

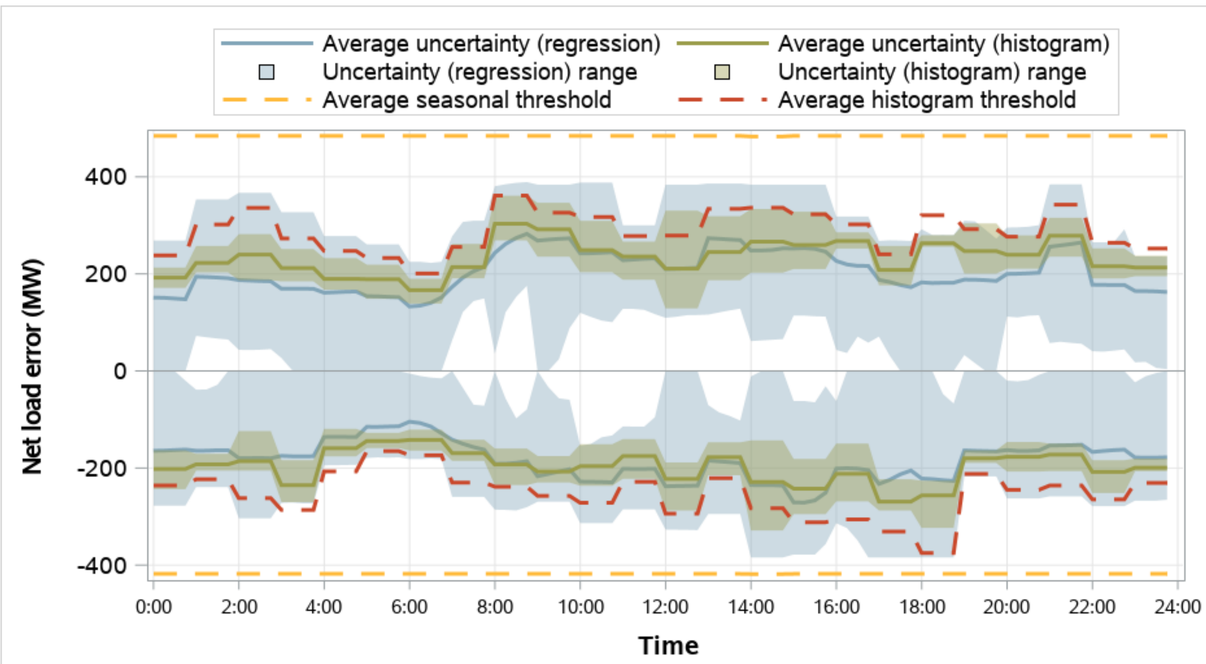


Figure 5.6 Avangrid resource sufficiency evaluation uncertainty requirements (October–December 2024)

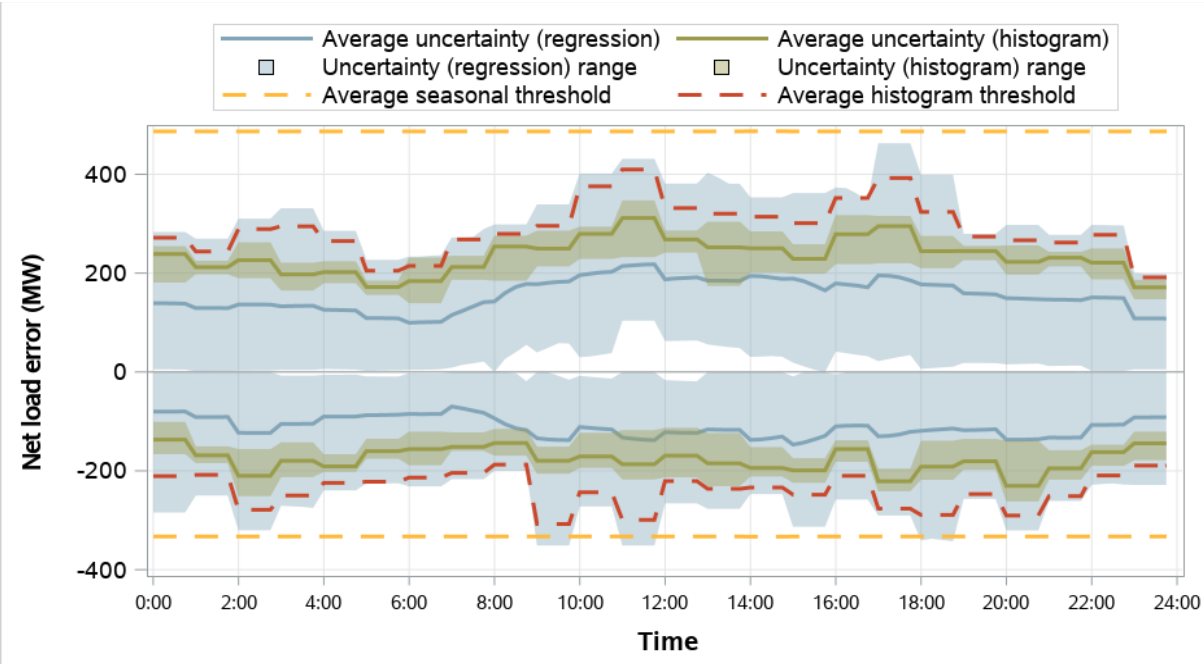


Figure 5.7 Avista resource sufficiency evaluation uncertainty requirements (October–December 2024)

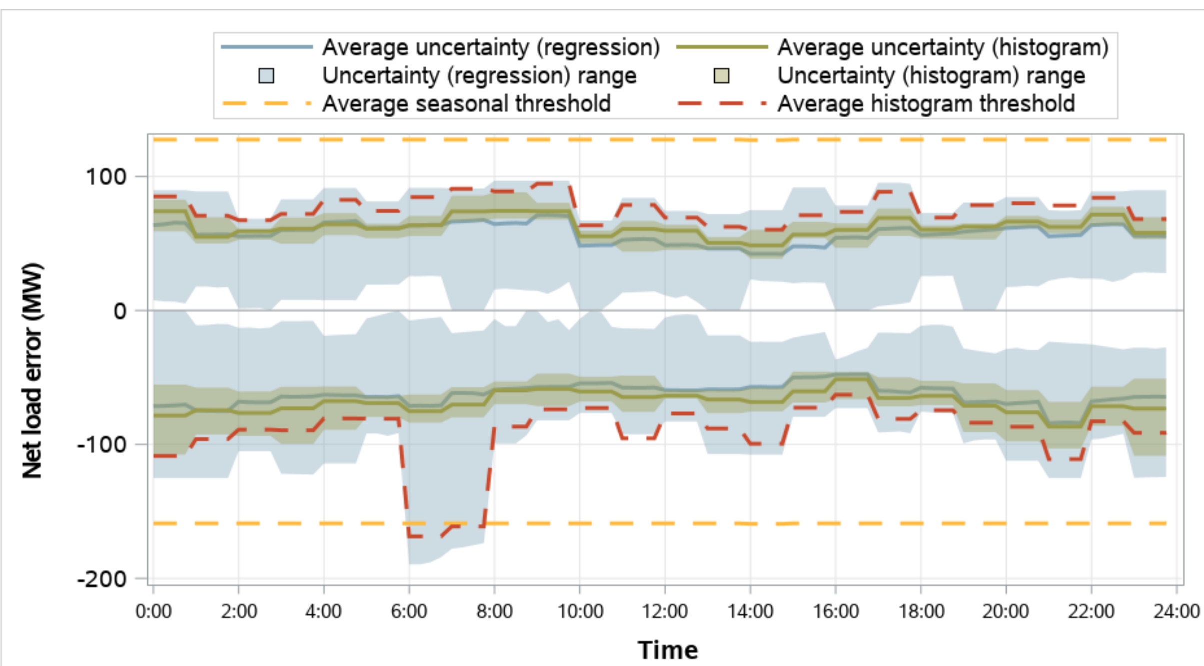


Figure 5.8 BANC resource sufficiency evaluation uncertainty requirements (October–December 2024)

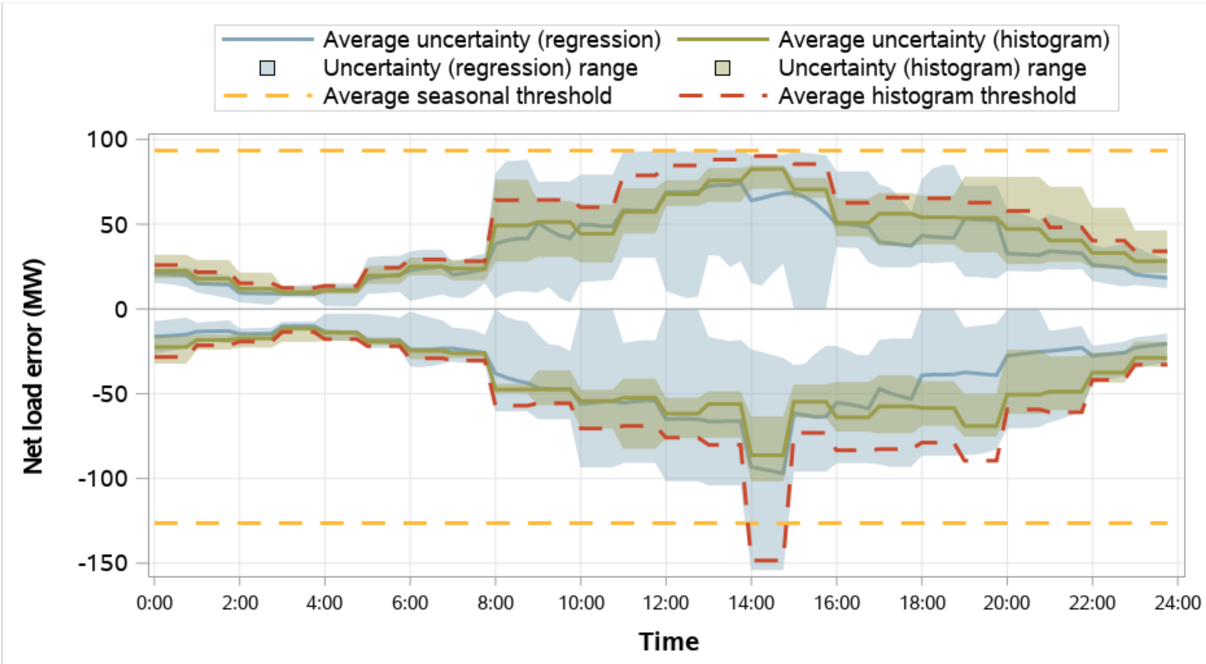


Figure 5.9 BPA resource sufficiency evaluation uncertainty requirements (October–December 2024)

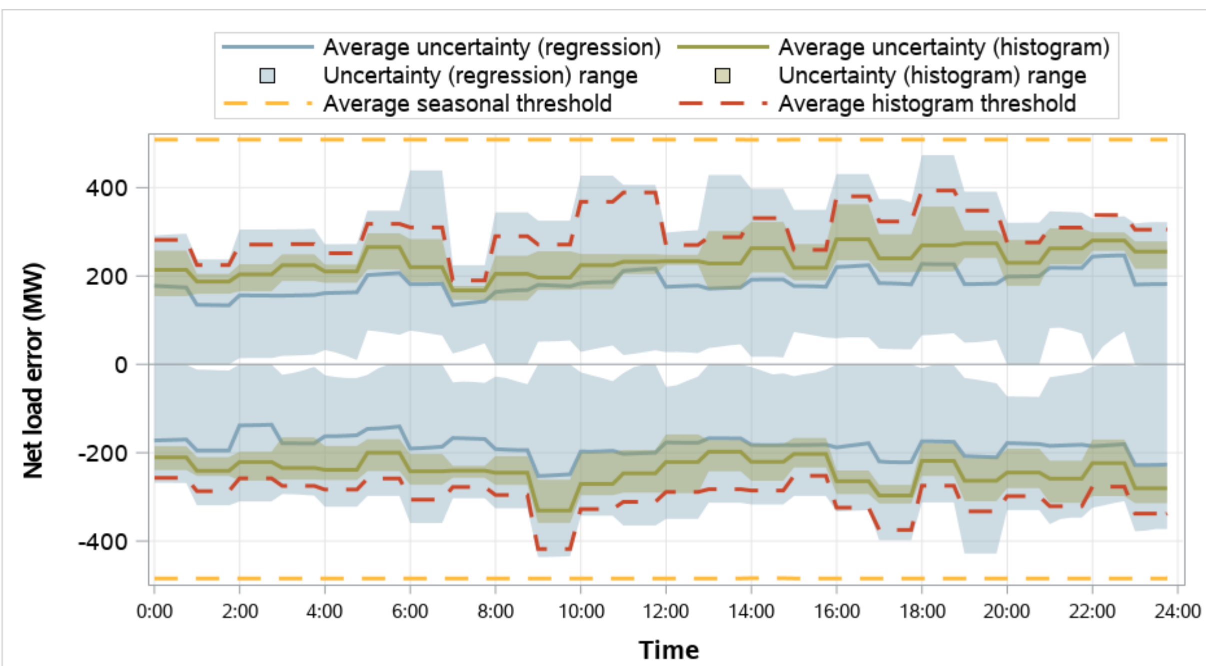


Figure 5.10 California ISO resource sufficiency evaluation uncertainty requirements (October–December 2024)

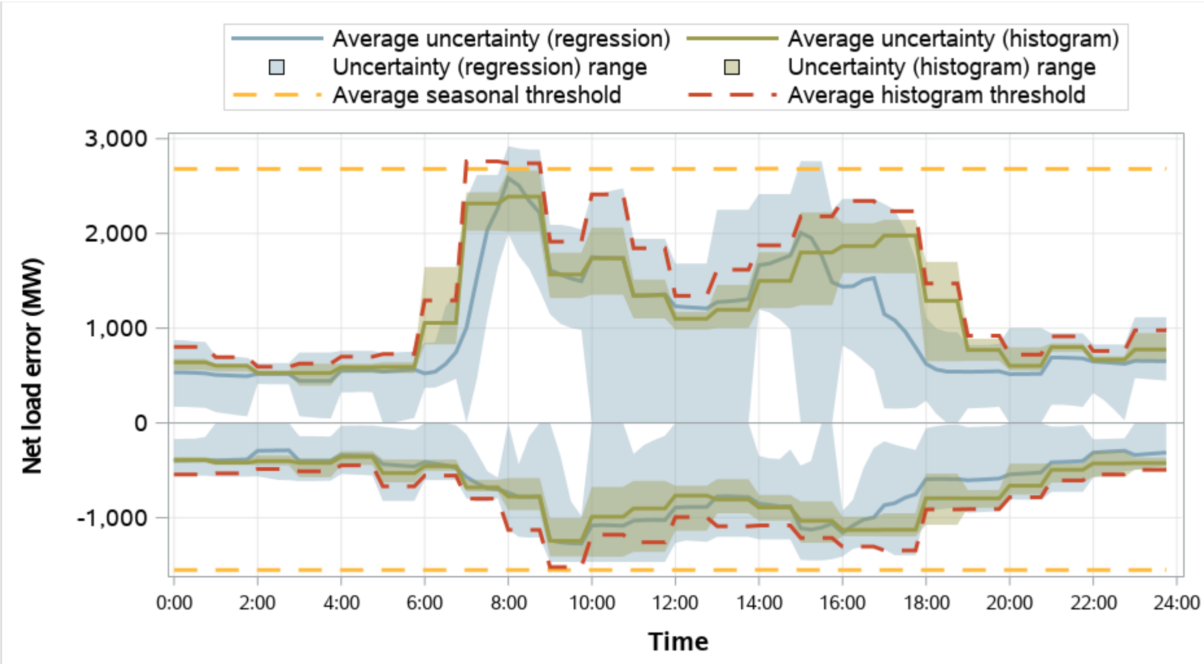


Figure 5.11 El Paso Electric resource sufficiency evaluation uncertainty requirements (October–December 2024)

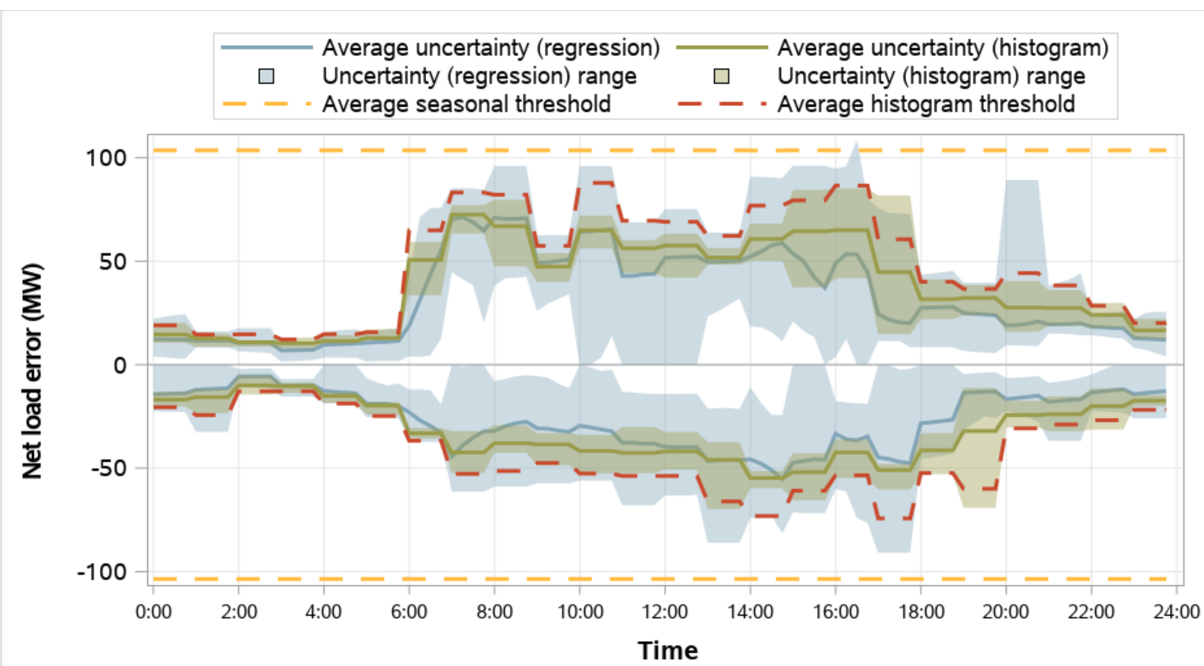


Figure 5.12 Idaho Power resource sufficiency evaluation uncertainty requirements (October–December 2024)

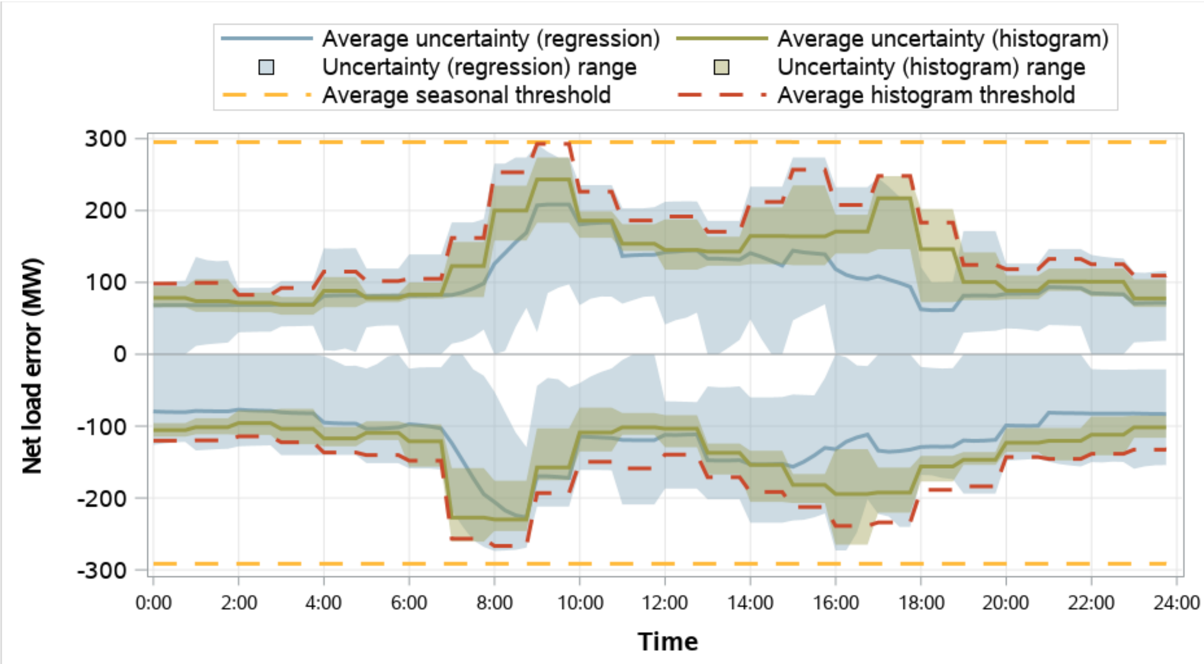


Figure 5.13 LADWP resource sufficiency evaluation uncertainty requirements (October–December 2024)

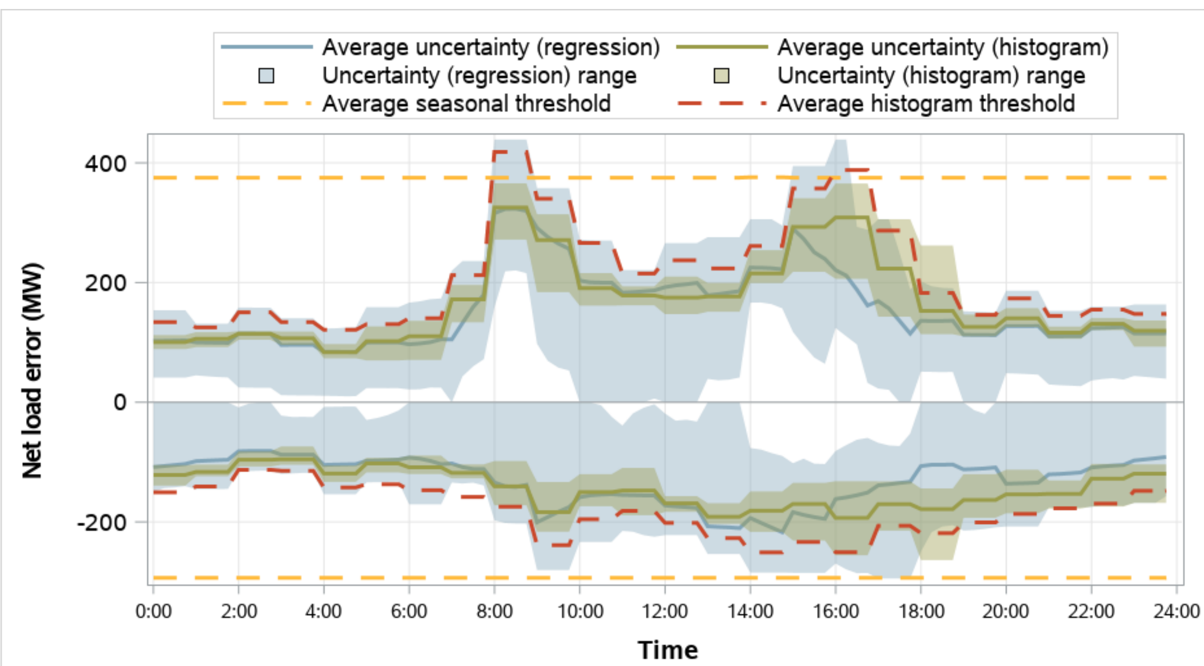


Figure 5.14 NorthWestern Energy resource sufficiency evaluation uncertainty requirements (October–December 2024)

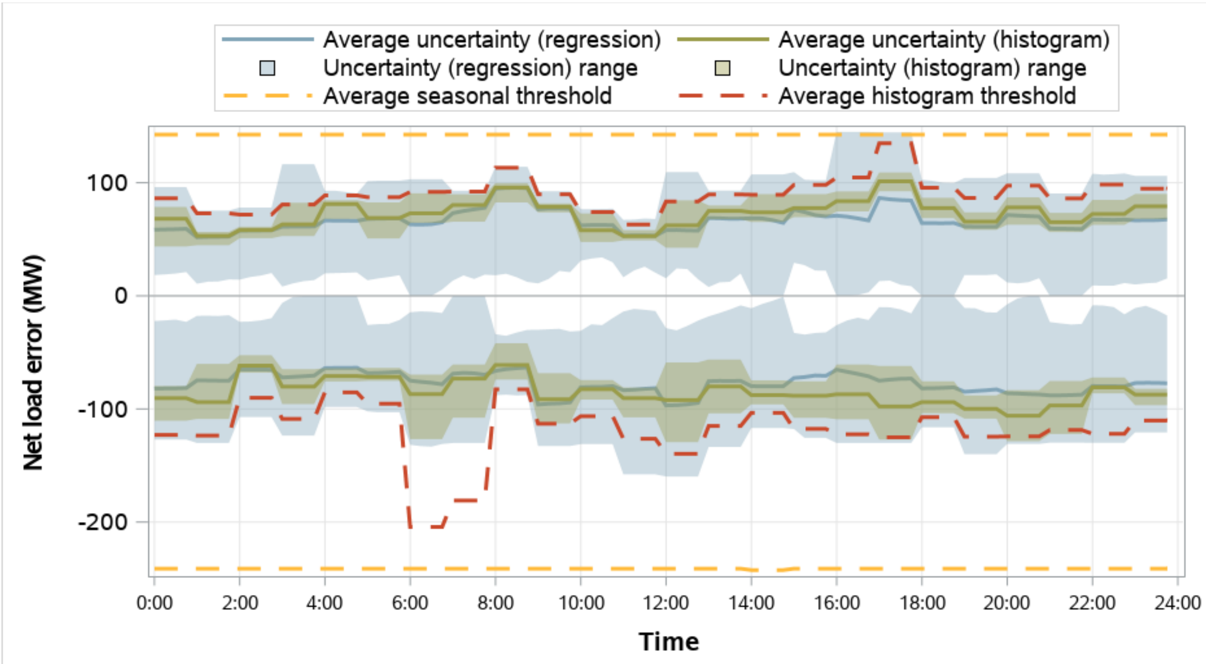


Figure 5.15 NV Energy resource sufficiency evaluation uncertainty requirements (October–December 2024)

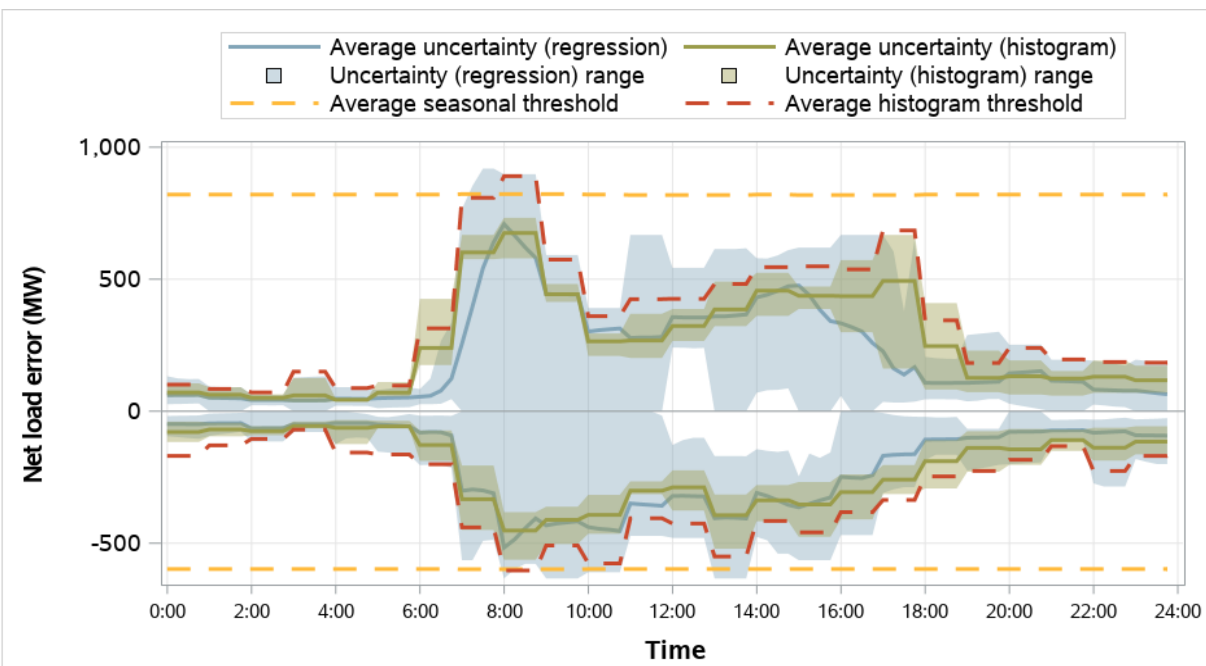


Figure 5.16 PacifiCorp East resource sufficiency evaluation uncertainty requirements (October–December 2024)

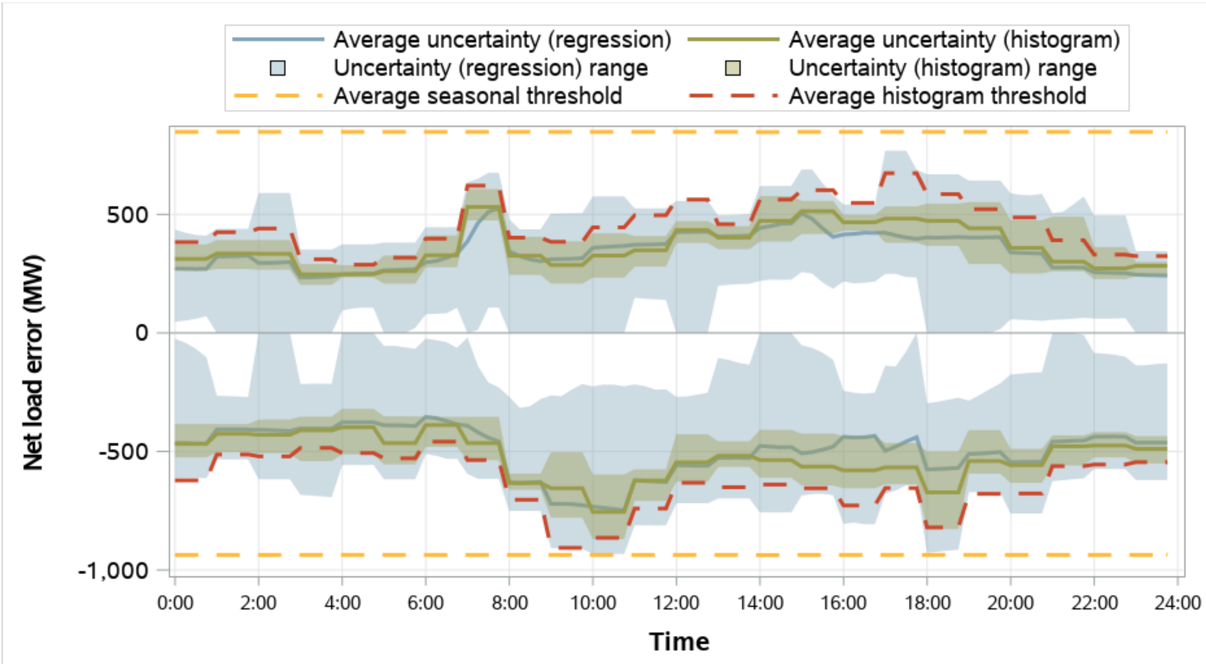


Figure 5.17 PacifiCorp West resource sufficiency evaluation uncertainty requirements (October–December 2024)

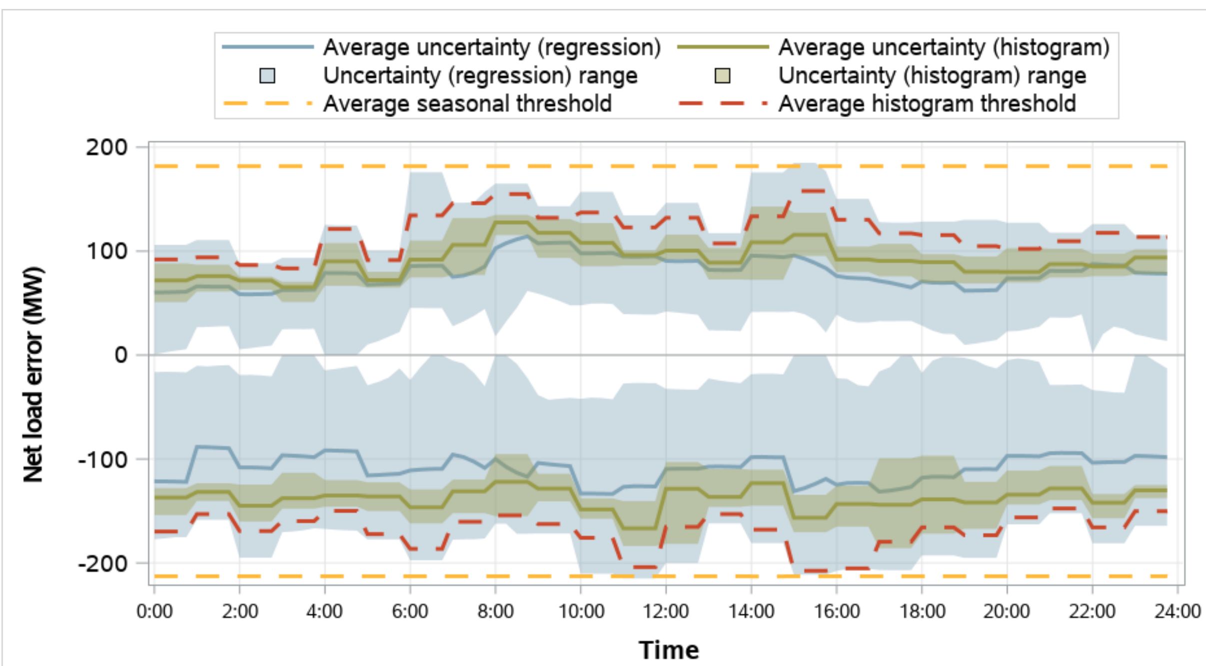


Figure 5.18 Portland General Electric resource sufficiency evaluation uncertainty requirements (October–December 2024)

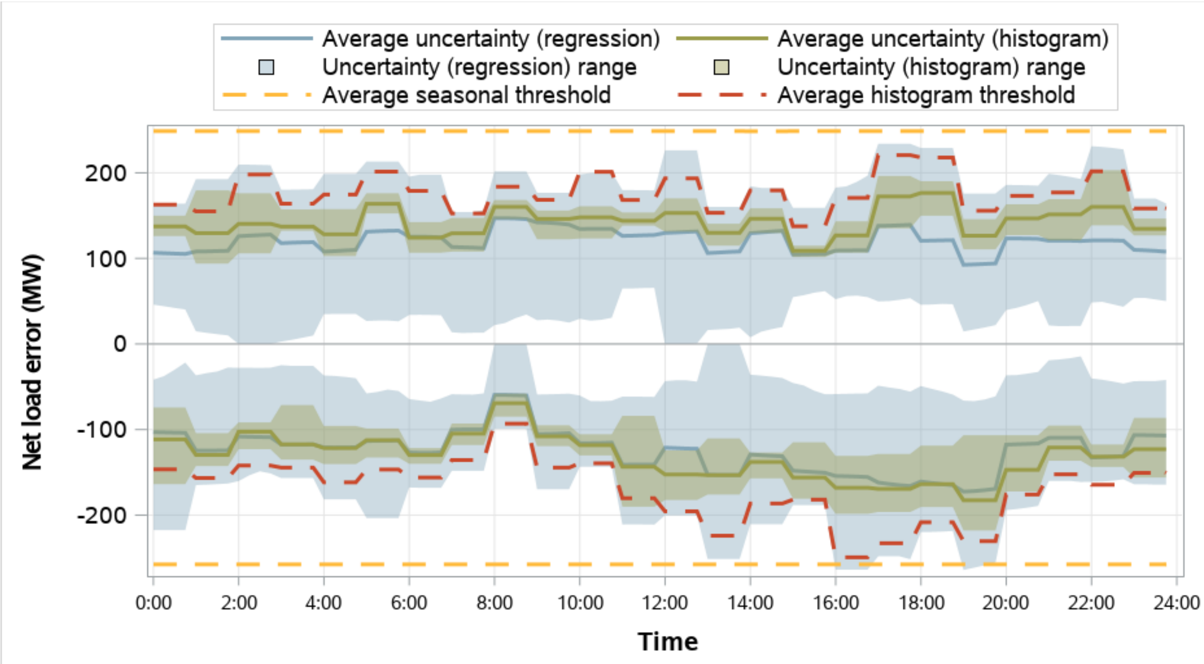


Figure 5.19 Powerex resource sufficiency evaluation uncertainty requirements (October–December 2024)

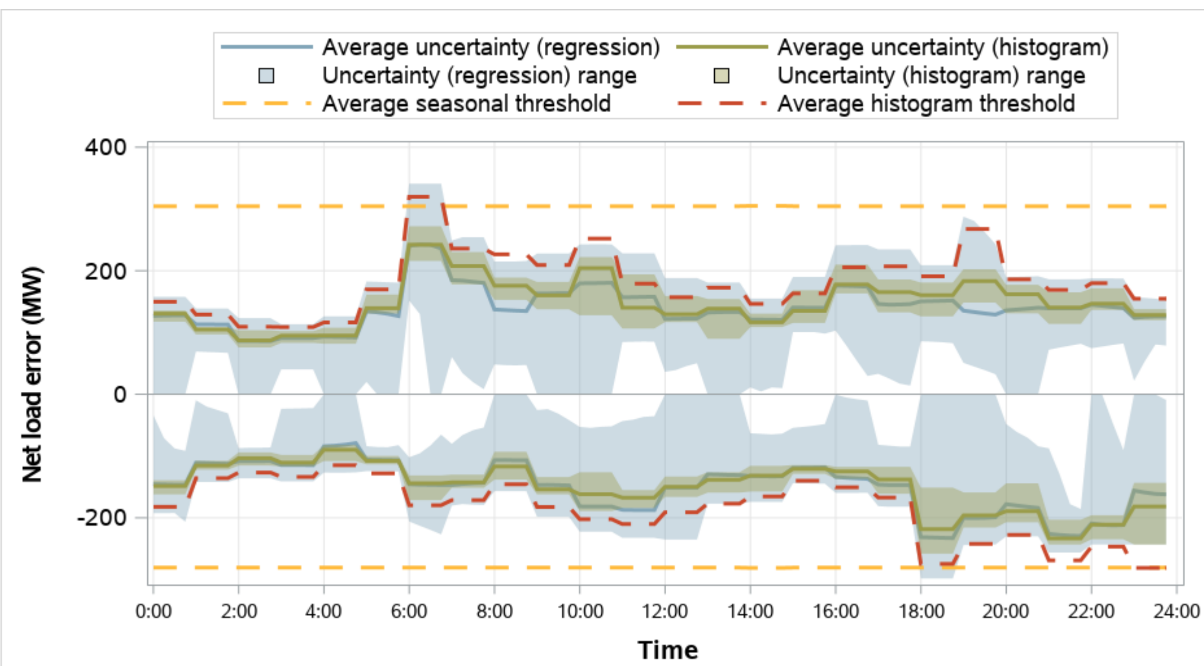


Figure 5.20 PNM resource sufficiency evaluation uncertainty requirements (October–December 2024)

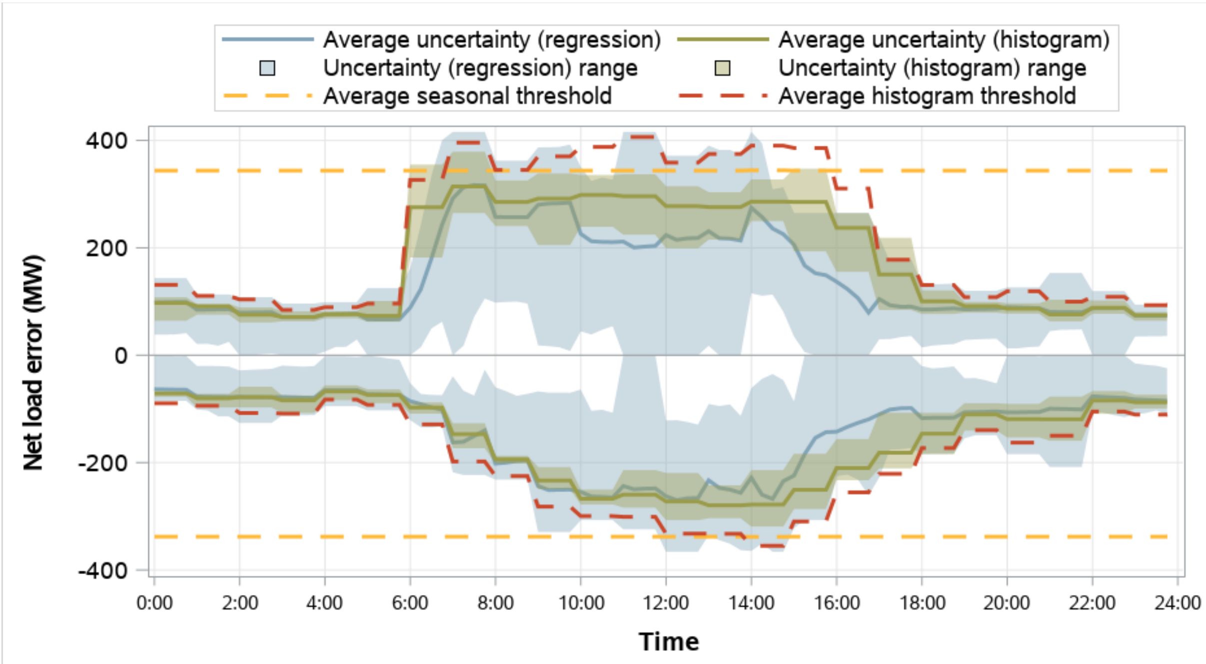


Figure 5.21 Puget Sound Energy resource sufficiency evaluation uncertainty requirements (October–December 2024)

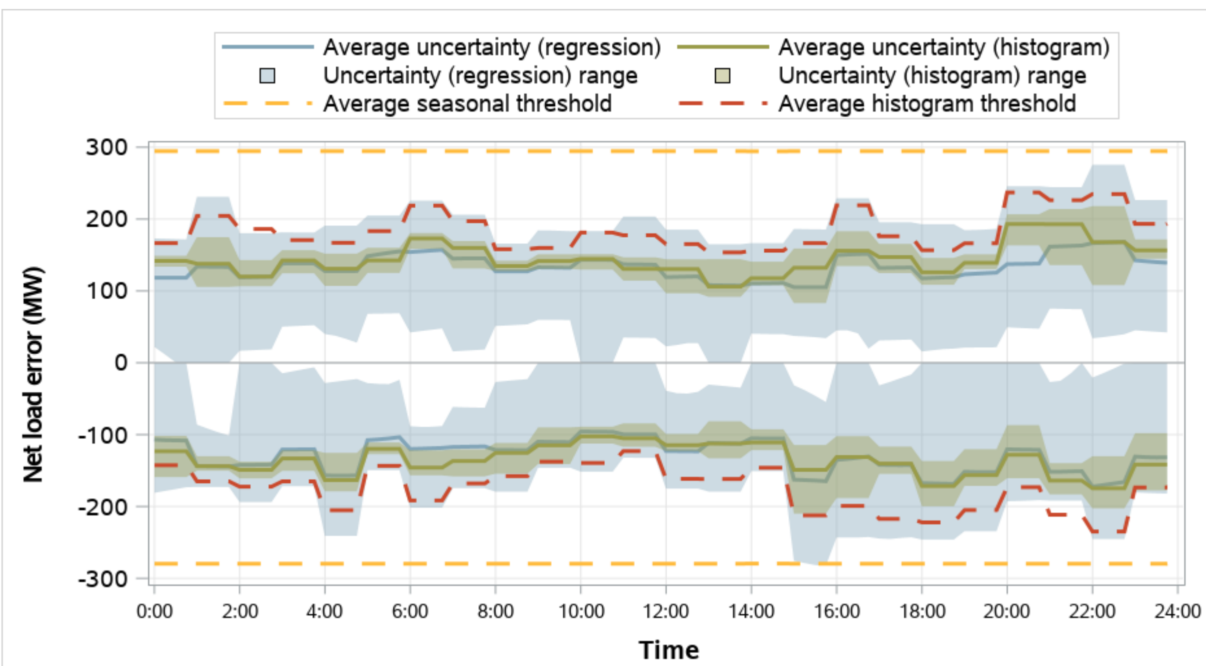


Figure 5.22 Salt River Project resource sufficiency evaluation uncertainty requirements (October–December 2024)

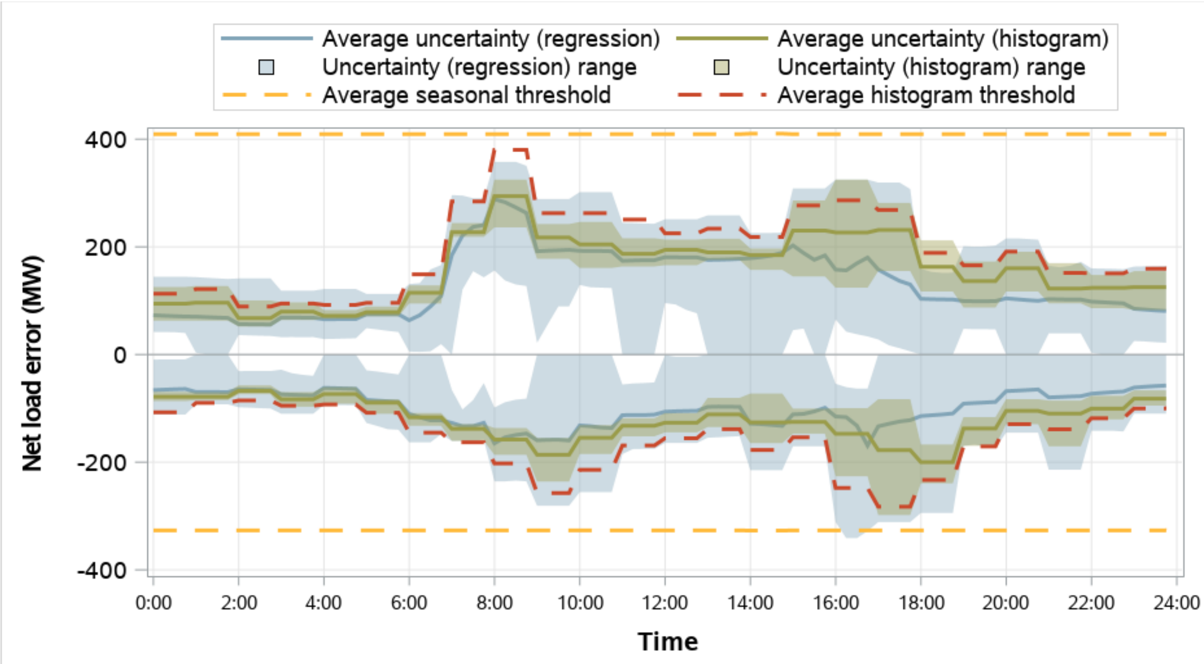


Figure 5.23 Seattle City Light resource sufficiency evaluation uncertainty requirements (October–December 2024)

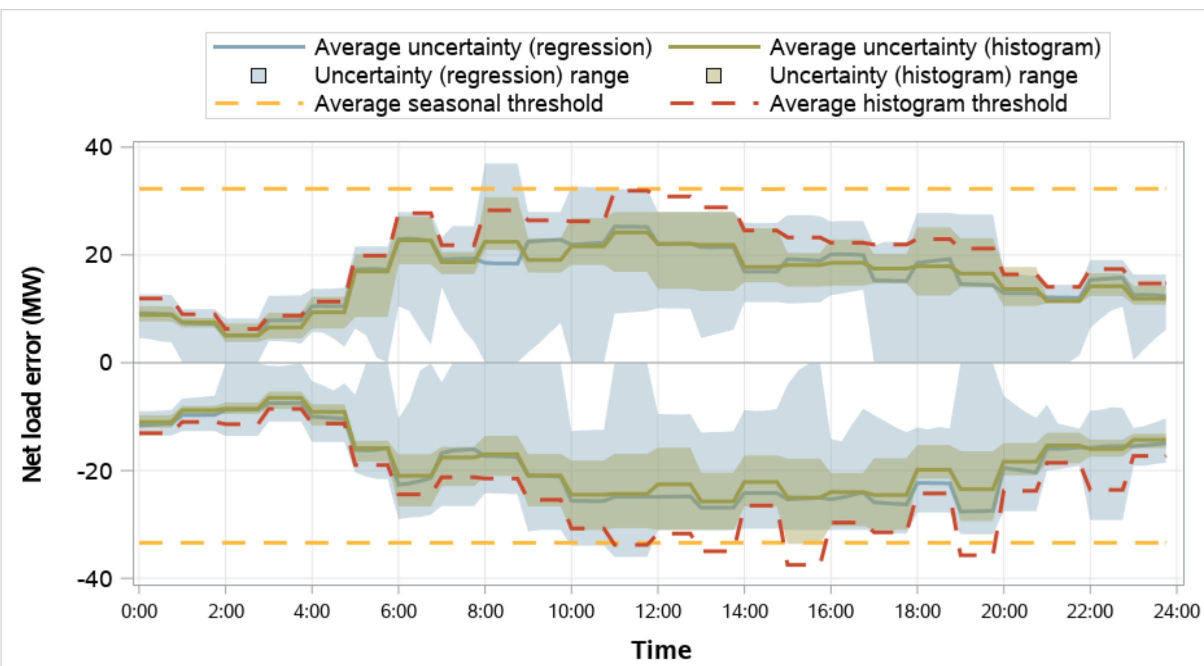


Figure 5.24 Tacoma Power resource sufficiency evaluation uncertainty requirements (October–December 2024)

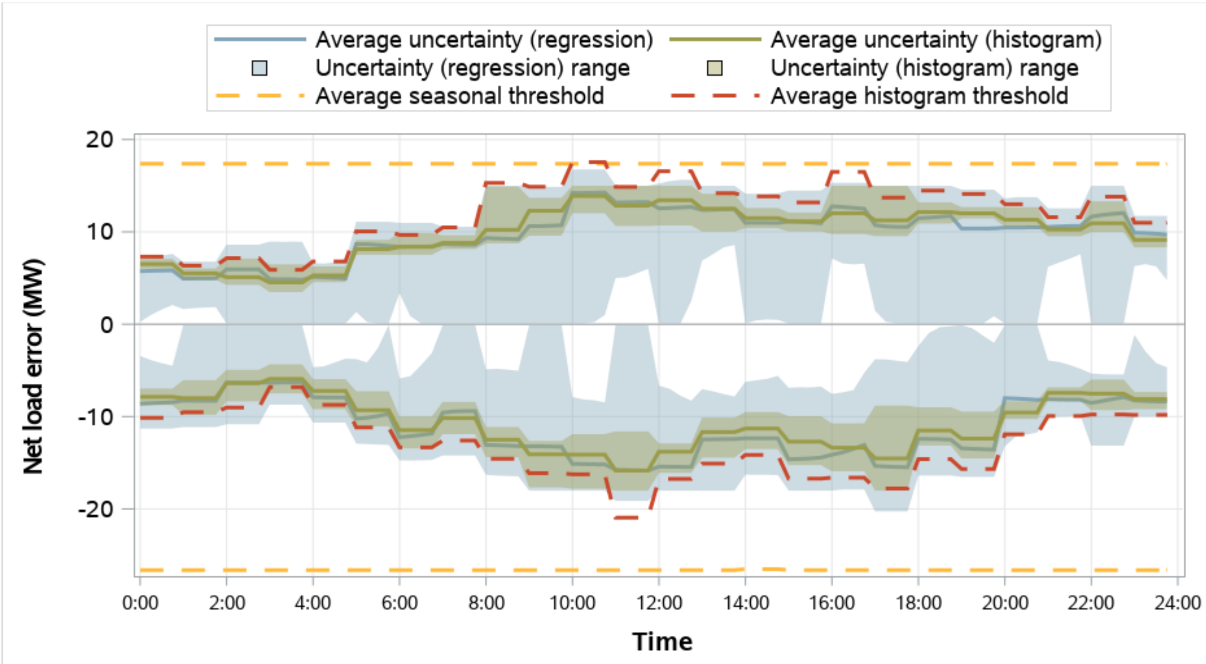


Figure 5.25 Tucson Electric Power resource sufficiency evaluation uncertainty requirements (October–December 2024)

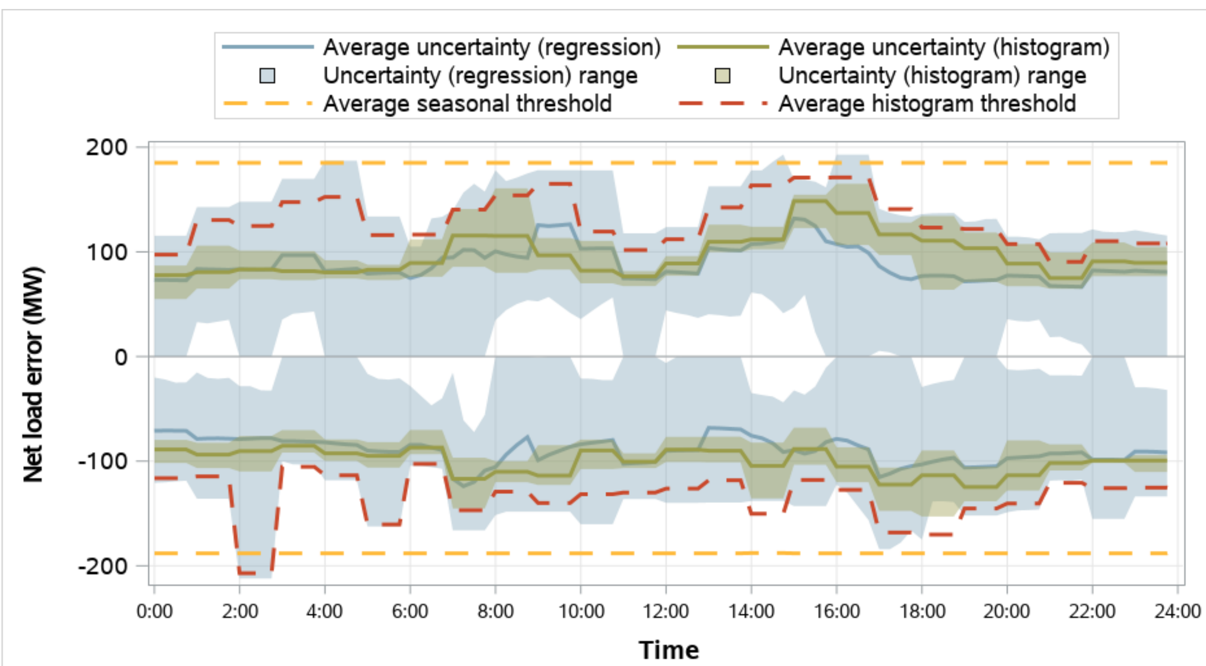


Figure 5.26 Turlock Irrigation District resource sufficiency evaluation uncertainty requirements (October–December 2024)

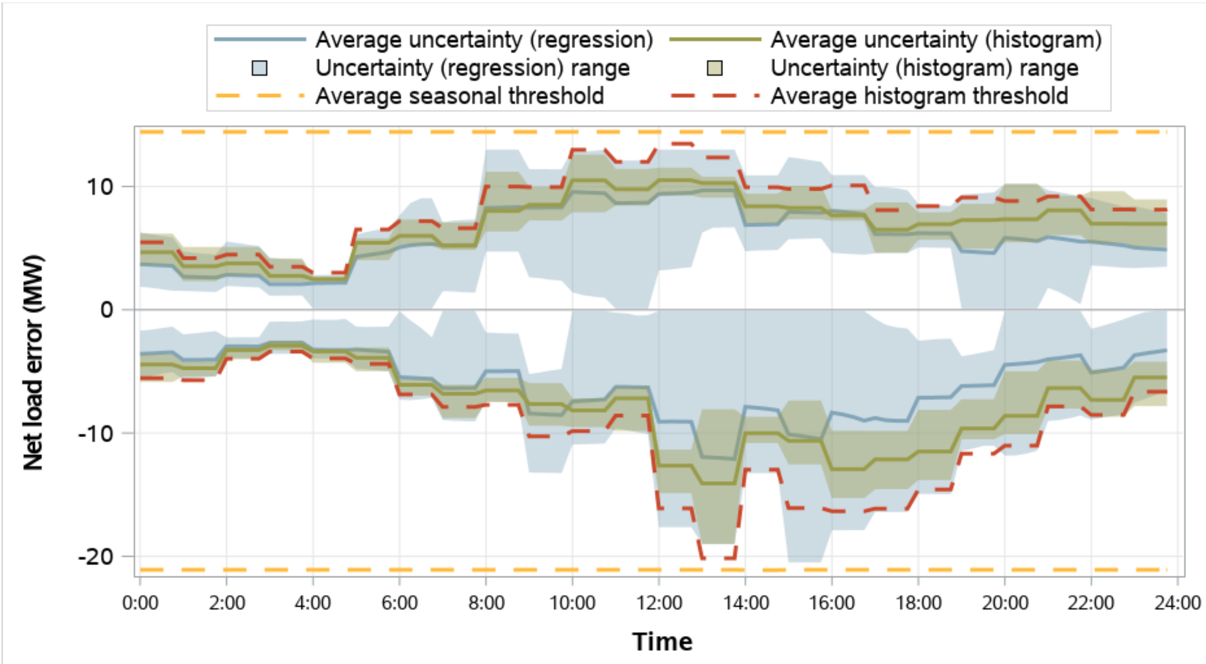
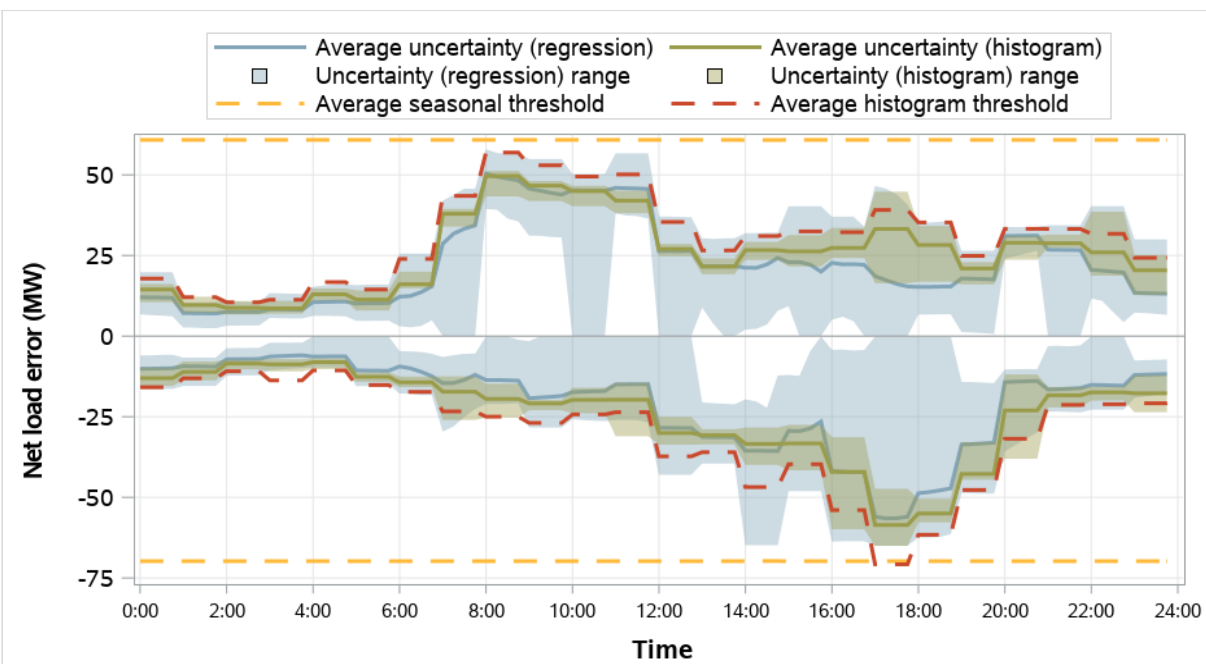


Figure 5.27 WAPA Desert Southwest resource sufficiency evaluation uncertainty requirements (October–December 2024)



Performance measurements of quantile regression uncertainty

Table 5.1 summarizes the average requirements and coverage for uncertainty in the resource sufficiency evaluation using both the histogram and mosaic quantile regression methods. In this table, *uncertainty* shows the average uncertainty component considered in the upward and downward flexibility test requirements. *Coverage* measures how frequent realized uncertainty—as measured by the difference from net load forecasts in the resource sufficiency evaluation to those in the binding 5-minute market—fell within the calculated uncertainty requirements for the same interval. On average across all hours, the uncertainty calculated from the regression method was less than the histogram method for almost all of the WEIM balancing areas.

The regression-method covered only 67 percent of uncertainty that ultimately realized in the real-time market for Arizona Public Service. Across all other balancing areas, the uncertainty calculated for use in the resource sufficiency evaluation from the regression method covered between 82 and 91 percent of realized uncertainty. The calculated uncertainty is designed to predict uncertainty in forecasts only 45 to 55 minutes before real-time while the resource sufficiency evaluation covers four 15-minute intervals, typically produced between 47.5 and 102.5 minutes before real-time. The shorter-term horizon of the design can contribute to lower coverage of realized uncertainty in the resource sufficiency evaluation.[13]

Table 5.1 Average resource sufficiency evaluation uncertainty requirements and coverage (October–December 2024)

	Upward uncertainty			Downward uncertainty			Coverage		
Balancing area	**Histogram**	**Mosaic**	**Difference**	**Histogram**	**Mosaic**	**Difference**	**Histogram**	**Mosaic**	**Difference**
Arizona Public Service	234	205	-29	198	184	-14	71%	67%	-4%
Avangrid	235	157	-79	177	112	-66	93%	87%	-6%
Avista	63	58	-5	69	63	-6	93%	89%	-3%
BANC	42	37	-5	43	38	-5	89%	86%	-3%
Bonneville Power Admin.	233	185	-48	242	186	-57	91%	85%	-6%
California ISO	1,178	1,025	-153	704	656	-48	91%	88%	-3%
El Paso Electric	39	33	-6	32	26	-6	93%	86%	-7%
Idaho Power	128	105	-23	138	117	-21	91%	86%	-5%
LADWP	168	156	-12	144	131	-13	92%	89%	-3%
NorthWestern Energy	73	67	-6	85	77	-8	92%	89%	-2%
NV Energy	261	216	-45	216	196	-19	94%	89%	-5%
PacifiCorp East	366	346	-21	526	496	-30	93%	91%	-2%
PacifiCorp West	92	80	-12	138	109	-29	91%	86%	-4%
Portland General Electric	143	121	-22	132	125	-7	93%	90%	-3%
Powerex	149	140	-8	150	150	1	89%	87%	-2%
PNM	178	145	-34	158	141	-16	93%	89%	-4%
Puget Sound Energy	144	134	-11	136	130	-6	92%	90%	-2%
Salt River Project	159	132	-27	121	100	-21	93%	87%	-5%
Seattle City Light	16	16	0	18	19	1	83%	83%	0%
Tacoma Power	10	10	0	11	11	0	83%	82%	-1%
Tucson Electric Power	97	89	-8	101	90	-11	93%	88%	-4%
Turlock Irrigation District	7	6	-1	8	6	-2	91%	83%	-7%
WAPA Desert Southwest	26	23	-3	24	21	-3	89%	85%	-5%

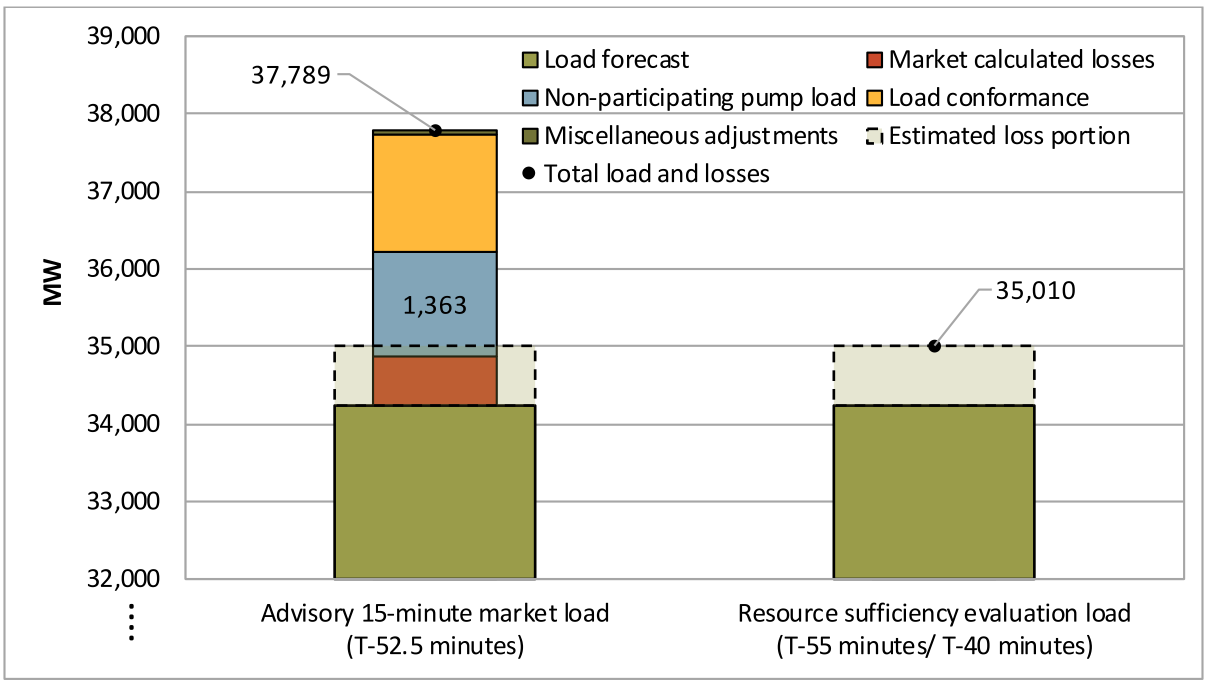
[13] For more information, see the section, Using uncertainty from the flexible ramping product for the resource sufficiency evaluation.

6 Additional demand in the real-time market compared to the resource sufficiency evaluation

The real-time market and resource sufficiency evaluation use different measurements for the total load. The resource sufficiency evaluation uses the raw (or initial) real-time load forecast directly in the requirement for both the capacity and the flexibility test. However, in the real-time market, the software adds operator load conformance, adds non-participating pump load, removes the portion that is estimated to be from losses, and finally recalculates the losses in the market.[14]

This is illustrated below in Figure 6.1 for the ISO area during an example interval. The example compares the total load and losses between the resource sufficiency evaluation with a corresponding advisory interval from the latest 15-minute market run.[15] In this example, the raw load forecast used in both cases (35,010 MW) was identical based on the timing of when the two market processes were run.

Figure 6.1 Example — difference between load used in the real-time market and in the resource sufficiency evaluation (CAISO, July 20, 2023. Hour-ending 23. Interval 1.)



The potential inclusion of load conformance was discussed as part of a resource sufficiency evaluation enhancements stakeholder process. In this process, the ISO confirmed no changes in the tests to account for load conformance, following findings that the use of load conformance does not regularly benefit any balancing area from passing the resource sufficiency evaluation.[16]

---

[14] The total load also adjusts for a few other miscellaneous components that cannot be accounted for elsewhere. The amounts here are typically small.

[15] Load and renewable forecasts are held fixed between the second run of the resource sufficiency evaluation (T-55) and final run (T-40).

[16] *EIM Resource Sufficiency Evaluation Enhancements Phase 2 Straw Proposal*, California ISO, July 1, 2022: http://www.caiso.com/InitiativeDocuments/StrawProposal-WEIMResourceSufficiencyEvaluationEnhancementsPhase2.pdf

Non-participating pump load within the ISO balancing area is not counted in the resource sufficiency evaluation. This is pumping load that is bid and scheduled as non-participating load in the day-ahead market, and included as a component of the total load in the real-time market optimization. This pumping load can be significant (above 1,000 MW).

Non-participating pump load is included in the real-time market but not in the resource sufficiency evaluation. This can create differences in the conditions observed between both processes. This can also be a factor in hours during which the ISO passes the resource sufficiency evaluation while an Energy Emergency Alert (EEA) is issued.

Other factors can also contribute to this outcome. First, rapidly evolving and declining conditions might prompt an EEA, but may not be observed by the resource sufficiency evaluation based on the latest information in advance of the evaluation hour. Also, real-time low priority and economic exports that clear the hour-ahead scheduling process would be included in the real-time market as additional demand, but are no longer counted as such in the resource sufficiency evaluation because of enhancements implemented on July 1, 2023.

DMM recommends that the ISO and stakeholders consider whether non-participating pump load should be included in the resource sufficiency evaluation. This would better align the conditions in the real-time market with the conditions considered in the resource sufficiency evaluation.

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 (October–December 2024)

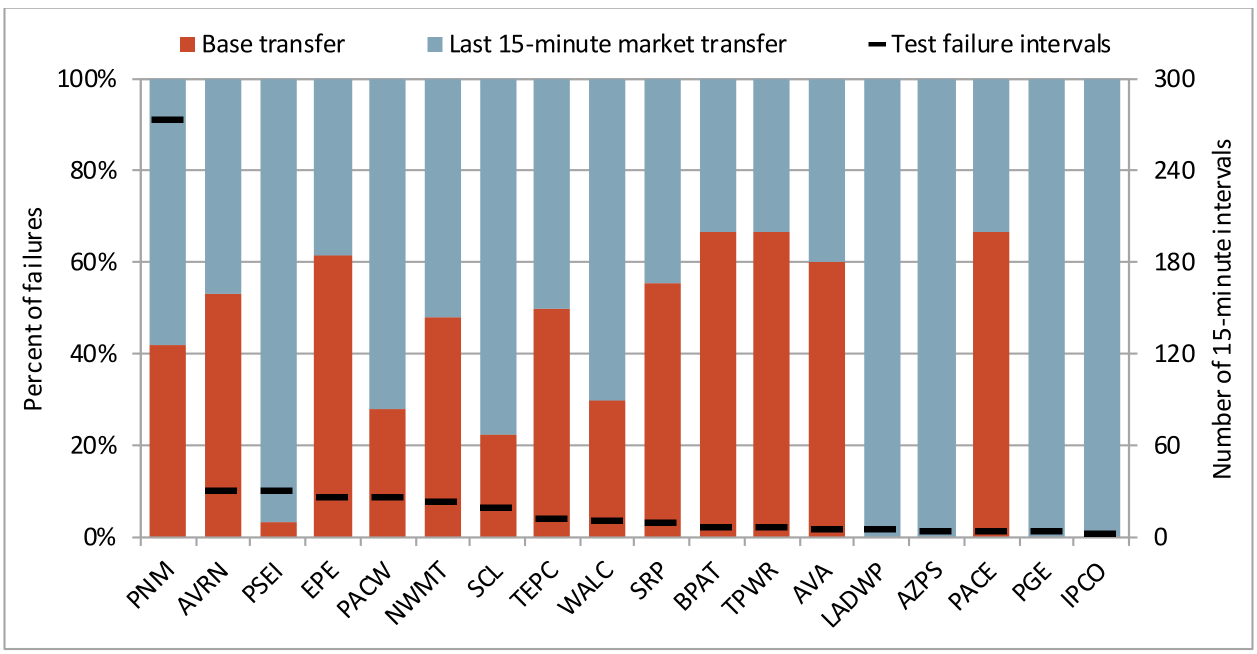


Figure 7.2 summarizes dynamic WEIM import limits above base transfers after failing either test in the upward direction.[17] From this perspective, the incremental WEIM import limit after a test failure is set by the greater of (1) zero or (2) the transfer from the last 15-minute market interval minus 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 quarter were excluded.

Figure 7.2 Upward capacity/flexibility test failure intervals by dynamic import limit (October–December 2024)

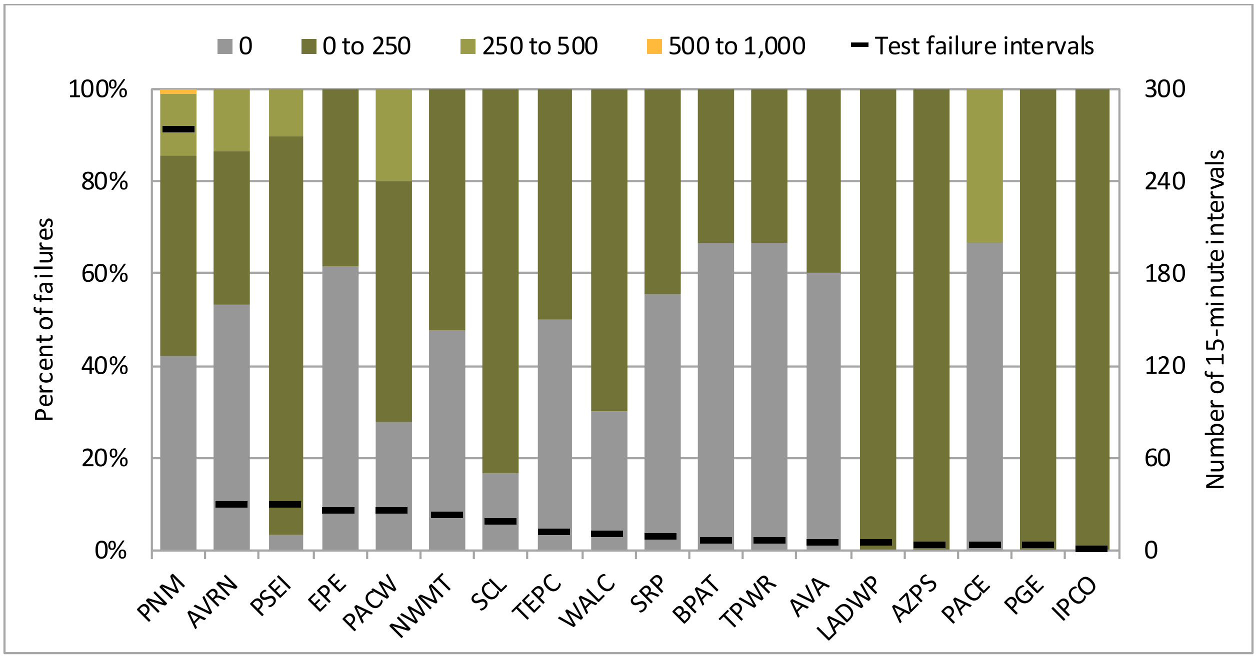
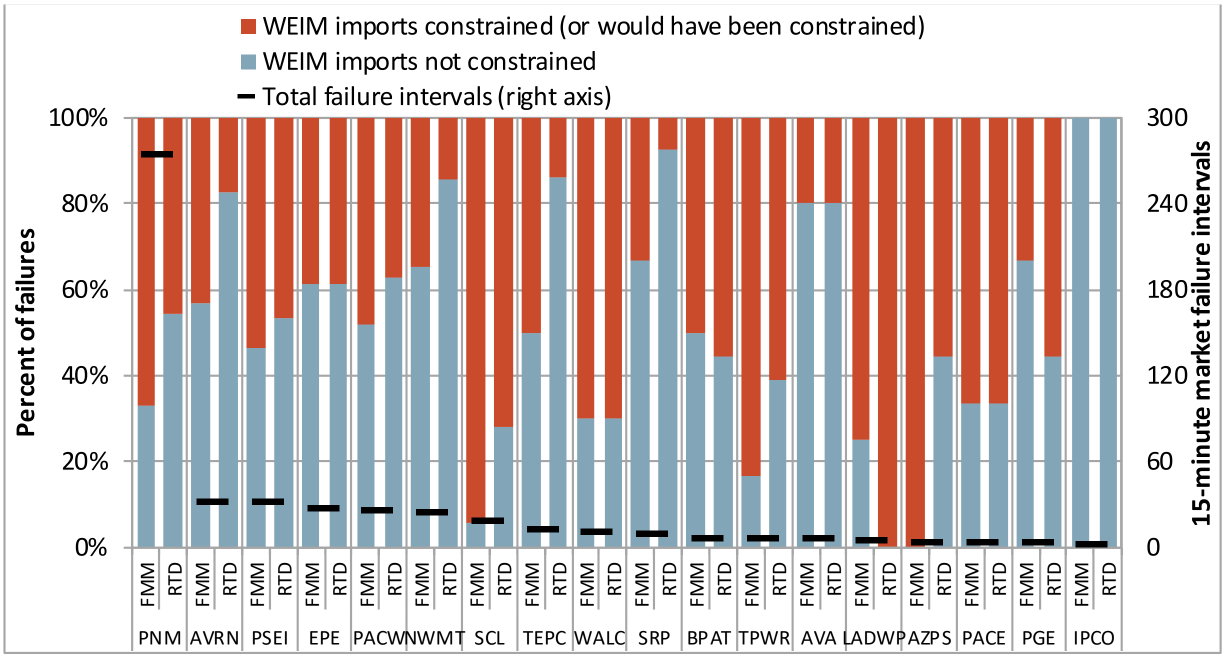


Figure 7.3 summarizes whether the import limit that was imposed after failing either test in the upward direction impacted market transfers (or would have impacted market transfers had the balancing area not opted in to the assistance energy transfer program).[18] The black horizontal line (right axis) 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 (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. These results are shown separately for the 15-minute (FMM) and 5-minute (RTD) markets.

[17] 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.

[18] 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 (October–December 2024)



Appendix A — 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 provide 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*.[19] 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 A.1 shows the different components and mathematical formulation of the bid range capacity test. As shown in Equation A.1, 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 A.1 Bid range capacity test requirement

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

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 balancing area, 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

[19] 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.

resource plan. Since the bid range capacity is calculated relative to the base schedules, the upward capacity test can generally be expressed as shown in Equation A.2.[20]

Equation A.2 Bid range capacity test reformulation

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

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 A.3 shows the different components and formulation of the flexible ramp sufficiency test requirement. 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 since June 1, 2022.

Equation A.3 Flexible ramp sufficiency test requirement

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

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

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.

[20] 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

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.

As shown in Equation A.3 above, 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.

Since 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.[21] 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.

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

Appendix B — Calculating net load uncertainty in the tests

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.[22] 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.

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 each type 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 in Equation B.1 for upward net load uncertainty.[23]

[22] 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.

[23] 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.

Equation B.1 Initial quantile regressions for upward net load uncertainty

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

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: 97.5th percentile, low: 2.5th percentile)

The uncertainty regressions use a distribution of historical forecast observations from 180 days, separate for each balancing area and hour. As of August 14, 2024, the historical observations are from two combined periods: (1) the previous 90 days, and (2) the next 90 days minus one year.[24] 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.[25] 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.[26] 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 in Equation B.2 for upward net load uncertainty.

Equation B.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

[24] Changes to Net-Demand Uncertainty Requirement Calculation Methodology in Flexible Ramping Product effective trade date 8/14/24: https://www.caiso.com/notices/changes-to-net-demand-uncertainty-requirement-calculation-methodology-in-flexible-ramping-product-effective-trade-date-8-14-24

[25] 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.

[26] 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.

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 histogram 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.[27] This is shown below in Equation B.3 for upward net load uncertainty.

Equation B.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 in Equation B.4 below.

Equation B.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: 97.5th 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

[27] The mosaic variable can be thought of as the modified net load.

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 in Equation B.5 below.

Equation B.5 Calculation of upward uncertainty from current forecast information

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

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

$$Net\ load\ uncertainty^{97.5}_{current} = a_m^{97.5}\left(mosaic^{97.5}_{current}\right)^2 + b_m^{97.5}\left(mosaic^{97.5}_{current}\right) + c_m^{97.5}$$

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 has published a more detailed review of the mosaic quantile regression approach.[28] DMM finds that the regression model has limited predictive capability for forecasting net load uncertainty.

[28] *Review of mosaic quantile regression for estimating net load uncertainty*, Department of Market Monitoring, November 20, 2023: http://www.caiso.com/Documents/Review-of-the-Mosaic-Quantile-Regression-Nov-20-2023.pdf