Flexible Ramping Uncertainty Calculation in the Western Energy Imbalance Market (WEIM)

March 25, 2022

Prepared by

Hong Zhou

Contents

Table of Figures

Table of Tables

Acronyms

APS	Arizona Public Service
BANC	Balancing Authority of Northern California
BCHA	Powerex
BTM	Behind the Meter
FRD	Flexible Ramping Down
FRP	Flexible Ramping Product Requirements
FRU	Flexible Ramping Up
H	Histogram
IPCO	Idaho Power Company
ISO	Independent System Operator
L	Load Uncertainty
LADWP	Los Angeles Department of Water and Power
M	Mosaic Quantile
MW	Megawatt
NEVP	NV Energy
NL	Net Load Uncertainty
NWMT	Northwestern Energy
PACE	PacifiCorp East
PACW	PacifiCorp West
PGE	Portland General Electric
PNM	Public Service Company of New Mexico
PSEI	Puget Sound Energy
Q	Quantile
RTD	Real Time Dispatch
RTPD	Real Time Pre-Dispatch
S	Solar Uncertainty
SRP	Salt River Project
W	Wind Uncertainty
WEIM	Western Energy Imbalance Market

1. Introduction

Since November 1, 2016, the California Independent System Operator (ISO) has two Flexible Ramping Product Requirements (FRP) in place for the 15-minute and 5-minute markets. These products provide additional upward and downward flexible ramping capability to account for uncertainty due to gross load, wind and solar forecasting errors. The forecast uncertainty is measured by Net Load (NL), Net Load = Load (L) – Wind (W) – Solar (S). In each market, FRP needs to estimate both Flexible Ramping Up (FRU) and Flexible Ramping Down (FRD) requirements.

The current implemented approach is called Histogram, which uses the upper 97.5 and lower 2.5 percentiles of observed net load uncertainty from the previous rolling 40 with matching week days and rolling 20 days for matching weekend days to set the FRP. Within this approach there remains two main limitations observed; 1) no incorporation of future impact of weather conditions on the net load uncertainty and 2) the historical sample set utilized.

Within the FRP refinements stakeholder initiative, started on November 21st of 2019, the ISO proposed enhancements to the FRP formulation[1]. Uncertainty requirements, such as the FRP are important to further evaluate and enhance over time to ensure the market properly captures the uncertainty of net load. This has been an area of great interest in different research efforts in recent years. This technical paper describes the ISO's proposal to use Quantile Regression to incorporate weather information in estimating FRP, including the construction of the net load formulation and Mosaic Quantile Regression, the comparison of the current Histogram approach to the newly formed Mosaic Quantile Regression, the analysis of the overall benefit in the Mosaic Quantile Regression, and lastly a sensitivity analysis of some additional considerations the ISO is monitoring.

Section-4 and 5 show the empirical uncertainty patterns found in each component of net load. These sections help illustrate potential advantages that Mosaic Quantile Regression approach may bring over the current Histogram approach. Section 6 concentrates the effort to the development of Mosaic Quantile Regression. Section 7 also illustrates the performance comparisons between Histogram and Mosaic Quantile Regression approaches. The remaining sections discuss additional items for consideration such as, the historical data sampling scheme utilized, selection of mosaic model, and thresholds needed due to their importance and impact.

[1] Appendix C of the FRP enhancements provide a description of the quantile methodology. Available at http://www.caiso.com/InitiativeDocuments/AppendixC-QuantileRegressionApproach-FlexibleRampingProductRequirements.pdf

2. Notation

Term	Definition
FRD	Flexible Ramping Down
FRP	Flexible Ramping Product Requirements
FRU	Flexible Ramping Up
(H)[2]	Histogram - The Histogram approach to estimate the requirements for L, W, S, and NL
L	Load Uncertainty = RTD Bidding Load - RTD Advisory Load
M	Mosaic Quantile - The mosaic quantile approach to estimate the requirement for NL
NL	Net Load Uncertainty = L - W - S
Q	Quantile - The quadratic quantile approach to estimate the requirements for L, W, and S
RTD	Real Time Dispatch
RTPD	Real Time Pre-Dispatch
S	Solar Uncertainty = RTD Biding Solar - RTD Advisory Solar
W	Wind Uncertainty = RTD Biding Wind - RTD Advisory Wind

2 https://bpmcm.caiso.com/BPM%20Document%20Library/Market%20Operations/BPM_for_Market%20Operations_V77_redline.pdf

3. Quantile Regression

Regression models are widely used statistical methods to exploit the relationships between a dependent variable (*e.g.*, uncertainty) and independent variables (*e.g*. load, wind, and solar). There are many different kinds of regression models for different purposes. The most popular model is ordinary least square model, which focuses on the averages of the dependent variable. However, since the FRP consists estimating extreme percentiles, it needs to deploy a specific family of models, namely Quantile Regression, to achieve such an endeavor.

Figure 1 displays the ordinary least-square model and Quantile Regression model for two hypothetic variables: Y the dependent variable and X the independent variable; r_y, regression of y, is the estimated average line by least square regression, the red line in left graph; and q_y, Quantile Regression of y, is the estimated 95% line by Quantile Regression, the red line in Figure 2.

The model expressions for these two kinds of models are the same as $Y = a + b * X$. The ordinary least square model finds the best linear relationship between the average of Y and X, while the quantile model finds the best linear relationship between the 95 percentiles of Y and X.

Quadratic relationships can be modeled by $Y = a + b * X + c * X^2$

Figure 1: Ordinary Regressions

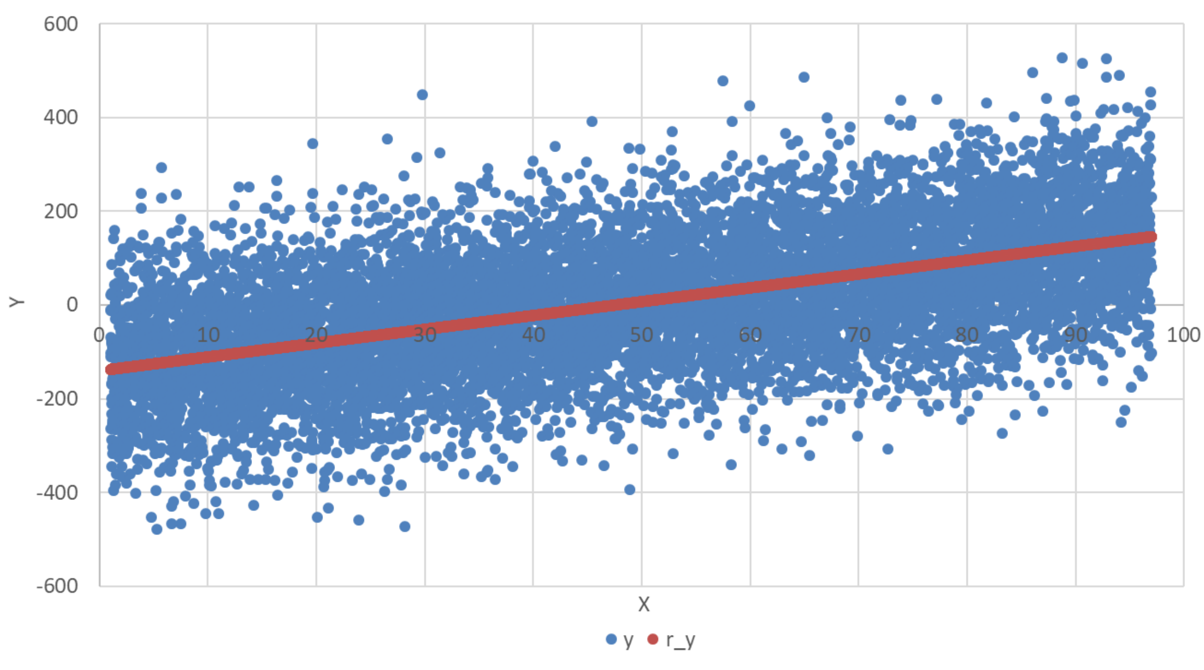
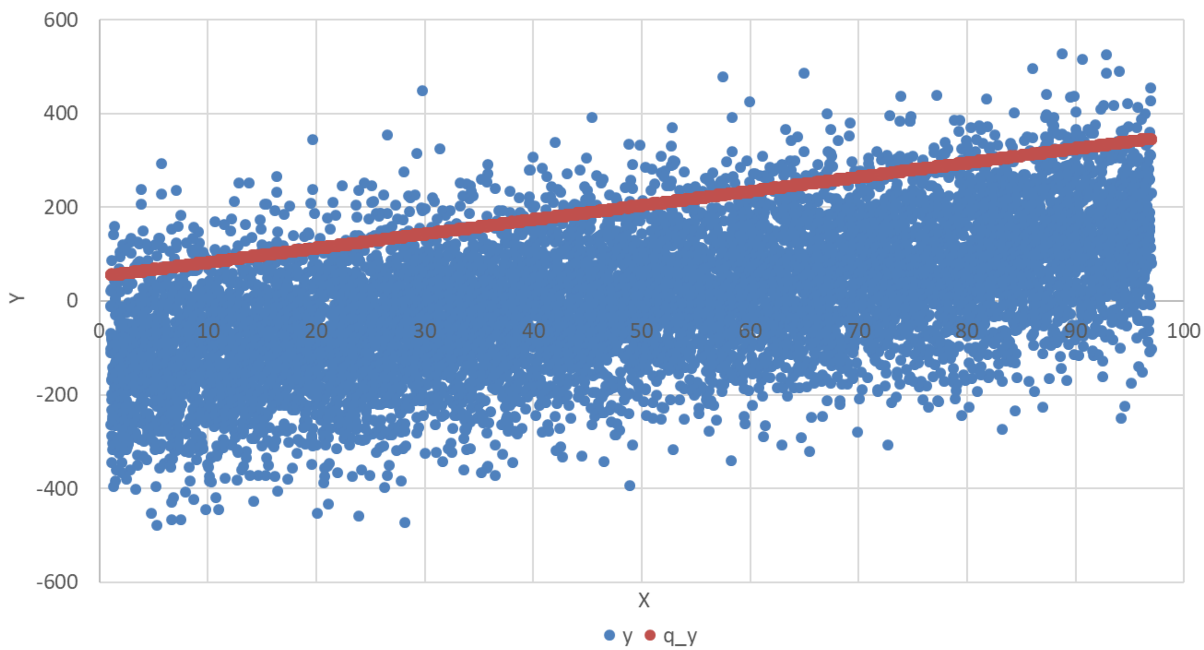


Figure 2 Quantile Regressions



4. Weather Information and Uncertainties

With the continuing growth of a resource mix that is impacted by weather conditions, it is important to use weather as an indicator of the uncertainty the grid may experience. Currently, at each stage of the ISO's markets, the load, wind, and solar forecasts utilize the latest Numerical Weather Prediction guidance in formulating the megawatt (MW) forecast. Within this proposal, the MW forecasts are treated as the condensed and the latest key weather information, these MW forecasts are readily available in the ISO when assessing the uncertainty of interest.

The existence of potential relationships between forecast uncertainty and the MW forecasts is crucial to the success of incorporating weather information into the estimates of FRP. The examination of each component, i.e., load, wind, and solar, reveals that forecast uncertainty seems to exhibit a quadratic pattern, i.e., the uncertainties are generally lower at two ends of spectrum of MW forecasts and reach high at the center of MW forecasts. The quadratic pattern is more pronounced in solar, followed by wind, and then load.

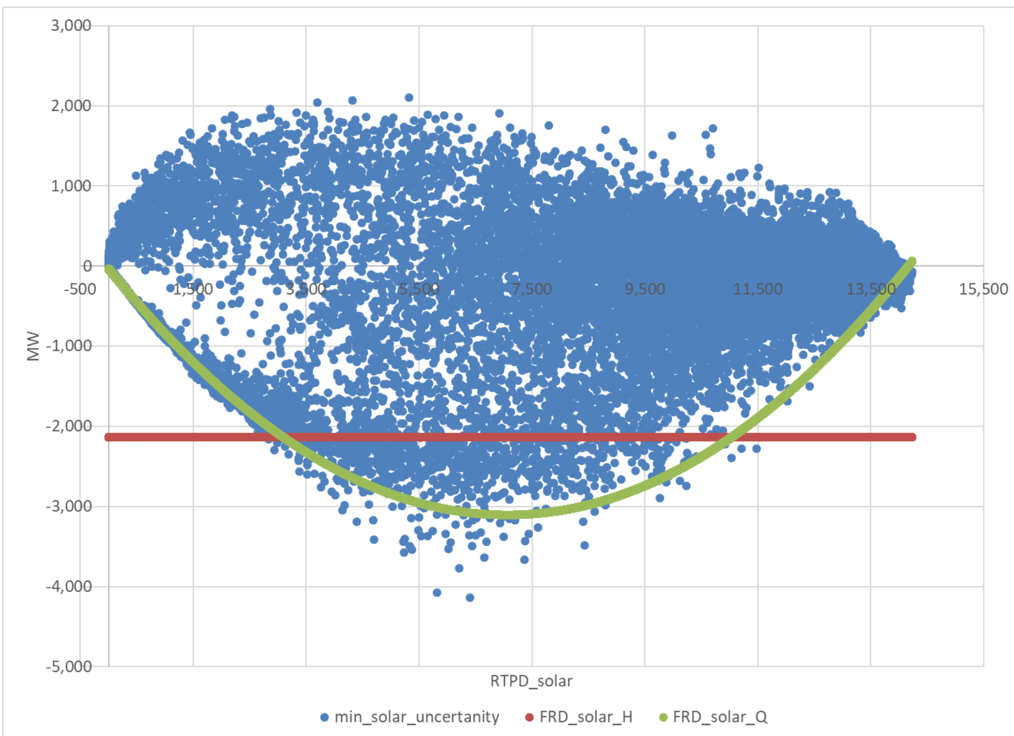
Solar uncertainties

Figure 3 examines the relationship between Solar Uncertainty (S) and Real Time Pre-Dispatch (RTPD) solar forecast in the ISO area in whole year of 2021. The patterns in other WEIMs have generally the same shape with varying degrees of curvatures as Figure 3.

The X axis represents RTPD solar forecasts, the Y axis represents the Real Time Dispatch (RTD) solar uncertainty defined as the difference between RTD binding solar forecast and RTPD advisory solar forecast, and where the minimum of the differences between the three RTD binding intervals and the RTPD advisory is counted for FRD, while the maximum is counted for FRU. The uncertainty notation of the minimum and the maximum will be dropped in the paper for the simplicity reason only.

The blue dots represent the solar uncertainty observed in a year. The red line represents 2.5% estimates by Histogram approach, whereas the green line represents the estimated 2.5% percentiles by quadratic Quantile Regression. The curvature exhibit by the green line shows the effect of Quadratic Regression Model to incorporate solar weather information. The observed uncertainty reaches a maximum level in the mid-range of the RTPD MW forecasts and gradually reduces to zero when RTPD forecast levels are either small or large.

Figure 3: Quadratic and histogram fit for solar uncertainties



With the model notations described in section 2, these observed curvatures in Figure 3 is modeled by the following quadratic Quantile Regression model,

$$solar\ uncertainty = a_s + b_s * solar_{RTPD} + c_s * \ solar_{RTPD}^2$$

Where $solar\ uncertainty = solar_{RTD} - \ solar_{RTPD}$, $solar_{RTD}$ and $solar_{RTPD}$ are the RTD solar and RTPD solar forecasts, respectively. The term $solar_{RTPD}^2$ is the square of RTPD solar forecast.

With the output denoted as S_q, the model is abbreviated as

$$S_q = a_s + b_s * solar_{RTPD} + c_s * solar_{RTPD}^2$$

The Histogram (H) estimate is attained when the linear and quadratic terms are disregarded

$$S_h = a_h$$

In other words, H estimate is a special Quantile Regression by dropping $Solar_{RTPD}$ and $Solar_{RTPD}^2$ in the model. The intercept (constant term) only regression model provides the H estimate, it will be denoted by S_h.

The solar Histogram estimates and the solar quadratic quantile estimates will also be used in the ISO's mosaic model.

Wind uncertainties

The impact of RTPD wind forecast on the Wind Uncertainty (W) is similar to the one observed for solar, especially on the low end of RTPD forecast Figure 4 shows the similar patterns as the solar counterpart.

In Figure 4, the X axis represents RTPD wind forecasts, the Y axis represents the RTD wind uncertainty defined as the difference between RTD binding wind forecast and RTPD advisory wind forecast, and where the minimum of three differences is counted for FRD, the maximum is used for FRU. The uncertainty notation of the minimum and the maximum is dropped to ease the description of formulation.

The blue dots represent the wind uncertainties observed in the year. The red line represents 2.5% estimates by Histogram approach, whereas the green line represents the estimated 2.5% percentiles by quadratic Quantile Regression (Q).

The wind quadratic quantile regression model is as follows

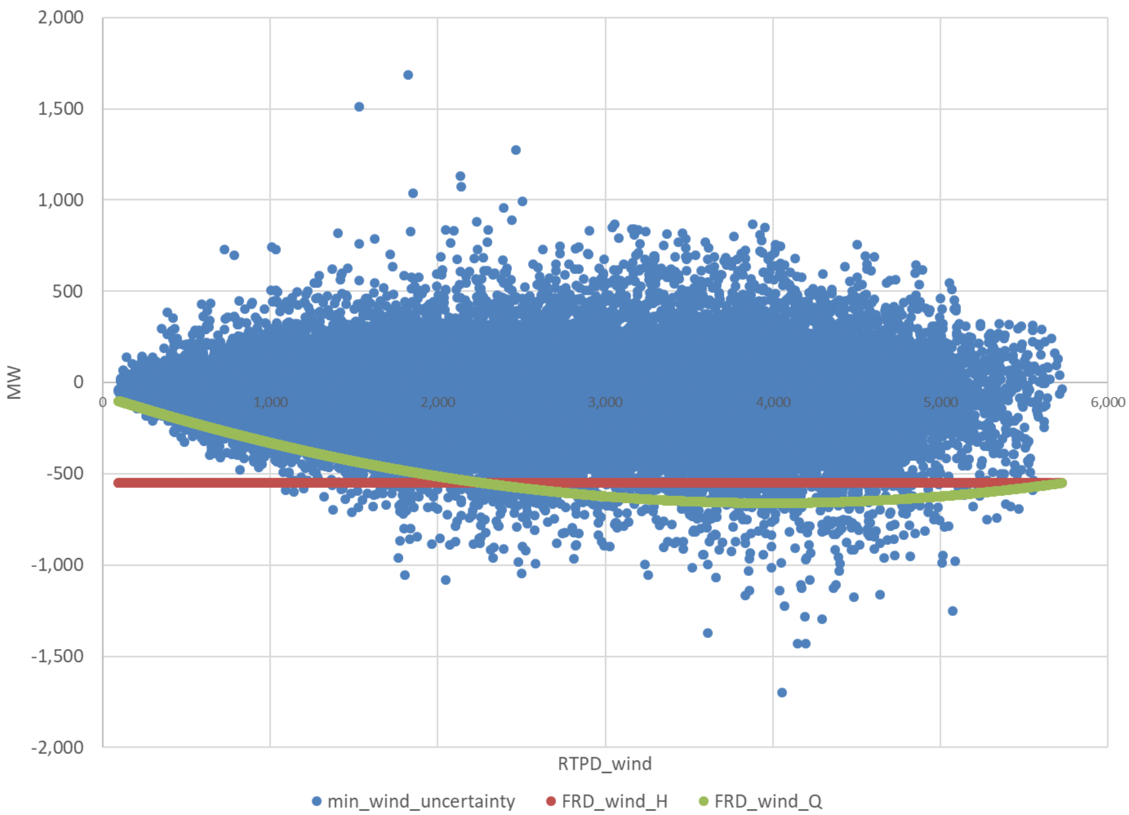
$$wind\ uncertainty = a_w + b_w * wind_{RTPD} + c_w * wind_{RTPD}^2$$

Where $wind\ uncertainty = wind_{RTD} - wind_{RTPD}$, $wind_{RTD}$ and $wind_{RTPD}$ are RTD wind and RTPD wind forecast, respectively. The term $wind_{RTPD}^2$ is the square of RTPD wind forecast.

With the output denoted as W_q , the wind model is abbreviated as

$$W_q = a_w + b_w * wind_{RTPD} + c_w * wind_{RTPD}^2$$

Figure 4: Wind uncertainties vs. RTPD wind forecasts



The wind Histogram estimates and the wind quadratic quantile estimates will also be used in the ISO's mosaic model.

5. Load Uncertainties

Similarly to wind and solar, RTPD load forecasts utilize weather information such as temperature, Behind-the-Meter (BTM) solar forecasts, and other key variables to produce a forecast. However, unlike the wind and solar, the energy consumption, i.e. load, is not only related to weather, but also heavily depends on human activity. Load has other factors driving error, such as COVID lockdowns and wild-fire smoke impacting the load levels.

In Figure 5 the blue dots represent the Load Uncertainty (L) observed in a year. The straight red line represents 2.5% estimates by Histogram approach, whereas the green curved line represents the estimated 2.5% percentiles by quadratic Quantile Regression (Q).

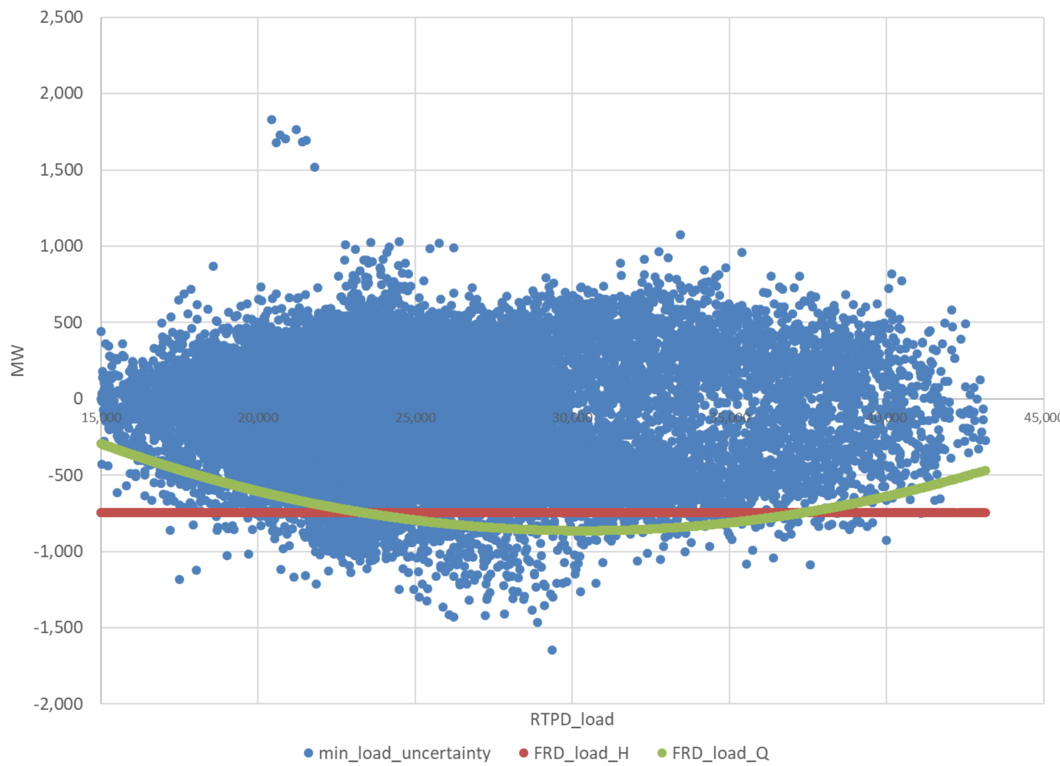
The load quadratic quantile regression model is as follows

$$load\ uncertainty = a_l + b_l * load_{RTPD} + c_l * load_{RTPD}^2$$

Where load $uncertainty = load_{RTD} - load_{RTPD}$, $load\,d_{RTD}$ and $load$ are RTD load and RTPD load forecast, respectively. The term $load_{RTPD}^2$ is the square of RTPD load forecast. With the output denoted as L_q y, the wind model is abbreviated as

$$L_q = a_w + b_w * load_{RTPD} + c_w * load_{RTPD}^2$$

Figure 5 Load uncertainties vs. RTPD load forecasts



The load Histogram estimates and the load quadratic quantile estimates will also be used in the ISO's mosaic model.

6. Mosaic Model for Net Load Uncertainty

In previous sections 4 and 5, each component (load, wind, and solar) has used a Quadratic form of the MW forecasts to estimate its forecast uncertainty. However, the success at the component level seems not be equally applicable to net load.

The ISO has explored different models before creating mosaic model as the final choice. The below are two examples:

1) $NL_q = \mathrm{a_{nl}} + b_{nl} * net_{load_{RTPD}} + \mathrm{c_{nl}} * \ net\_load^2_{RTPD}$,
2) $NL_q = \mathrm{a} + b_l * load_{RTPD} + c_l * \ \ load^2_{RTPD} +$

$$b_w * \ wind_{RTPD} + \mathrm{c_w} * \ wind^2_{RTPD} +$$

$$b\_s * solar_{RTPD} + \ c_s * solar^2_{RTPD}$$

The final choice of ISO's Quantile Regression model is called Mosaic Quantile Regression model. The results of the Mosaic Quantile Regression model exhibit better performance over other models like the above ones based on an array of performance measurements listed in the next section. This may be because the Mosaic Quantile Regression model directly utilizes the relationships found in each and every component shown in previous sections. Moreover, Mosaic Quantile Regression model has an element in modeling the complex interactions among load, wind, and solar.

The name of Mosaic Quantile Regression comes from the fact that it blend ingredients from its load, wind, and solar components.

There are two steps to construct Mosaic Quantile Regression model to estimate the uncertainty requirements

1) Define of the quadratic Quantile Regression model for its components,

$$L_q = a_l + b_l * load_{RTPD} + c_l * load^2_{RTPD},$$

$$W_q = a_w + b_w * wind_{RTPD} \ + c_w * wind^2_{RTPD},$$

$$S_q = a_s + b_s * solar_{RTPD} + c_s * \ solar^2_{RTPD}$$

Relative to the counterparts from Histogram estimates, L_h , W_h , and S_h , these quadratic estimates, L_q, W_q, and S_q can include weather information, respectively.

2) Construct a regression input variable based on quadratic Quantile Regression estimates from load, wind, and solar to incorporate weather information to modify net load uncertainty NL_h as follows

$$mosaic = \ NL_h + \left(L_q - L_h\right) - \left(W_q - W_h\right) - \left(S_q - S_h\right)$$

Then the final Mosaic Quantile Regression is as follows

$$NL_M \ = a_m + b_m * mosaic + c_m * mosaic^2$$

This model blends a few different pieces of estimates together and is the reason to be named Mosaic Quantile Regression model.

The following list may offer some perspectives to understand the Mosaic Quantile Regression model,

a) Mosaic Quantile Regression approach utilizes single quadratic inputs in stages as an alternative approach other than multivariate regression approach.
b) If no weather information has impact, i.e., $L_q = L_h$, $W_q = W_h$, and $S_q = S_h$, the proposed estimate NL_q will result the same estimates as obtained by the Histogram approach, i.e., $NL_q = NL_h$.
c) The item $NL_h - (L_h - W_h - S_h)$ from the Histogram estimates can be viewed as the effect of non-additive interactions among load, wind, and solar.
d) The marginal component curvatures L_q, W_q, and S_q contain weather information. $L_q - L_h$, $W_q - W_h$, and $S_q - S_h$ are major parts in the mosaic formation to provide modest modification of the Histogram estimate of net load uncertainty NL_h
e) The ISO's simulation study below shows some positive results for Mosaic Quantile Regression model, but it does not preclude the existence of other competitive or even better Quantile Regression models not using curvatures L_q, W_q, and S_q.
f) The Mosaic Quantile Regression square term may not offer too much additional benefit, see the discussion in a section 14.
g) The Mosaic Quantile Regression is to estimate extreme percentiles of uncertainty. Generally, any regression based estimates are not as stable as the one point estimate by Histogram approach. To ensure the feasibility of the regression based estimate, the ISO proposes to put bounds to cap the outputs as follows: NL_M is bounded by 1 percent and 99 percent estimates by Histogram approach.

Mosaic model example

This section provides an example by using the 2021 data in the ISO to illustrate the relationship between the estimates and load, wind, and solar components. Sections 4 and 5 discuss the Quantile Regression models for each component (solar, wind, and load), respectively, here the dependent variable is the net load uncertainty, and the independent variable is the mosaic constructed in this section. This section is a preamble for outcomes and visualization of the simulation study in section 7

In the following set of graphs, the blue dots are the observed net load uncertainty, red line is the Histogram estimate for FRU for this data set, and green dots are the Mosaic Quantile Regression estimate for this data set. X-axis is changing from net load, load, wind, to solar.

When compares to Histogram estimates, it can be seen that Mosaic Quantile Regression estimates are attempting to adapt various weather conditions. In the following figures, please note that a 99% cap, described above in g), is applied.

Figure 6 Estimates of Histogram vs. Net Load

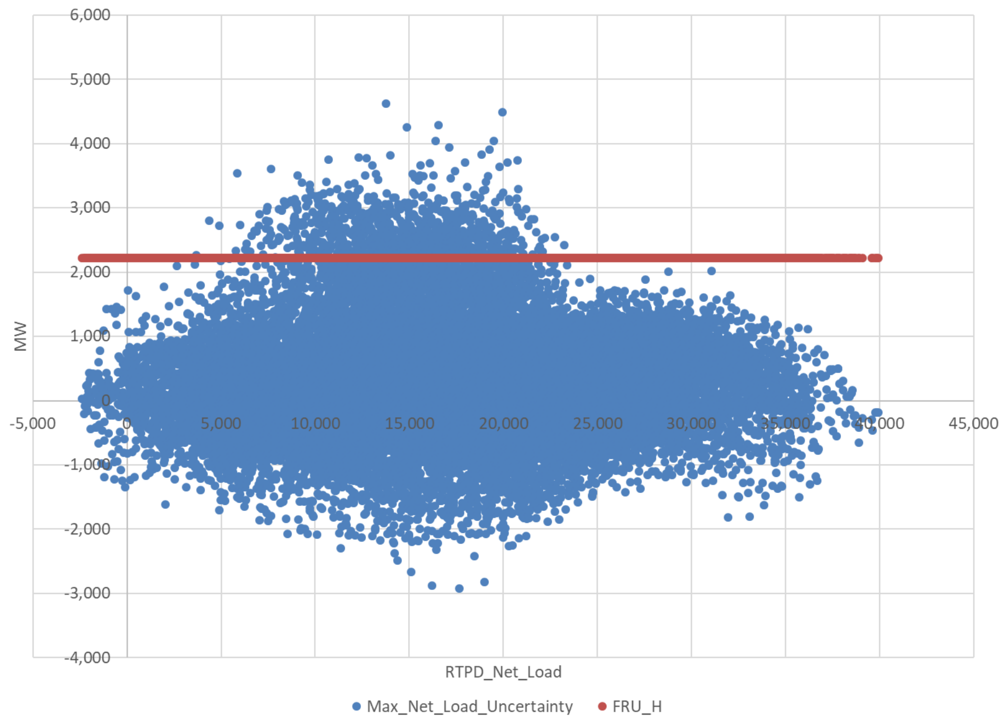


Figure 7 Estimates of Mosaic vs. Net Load

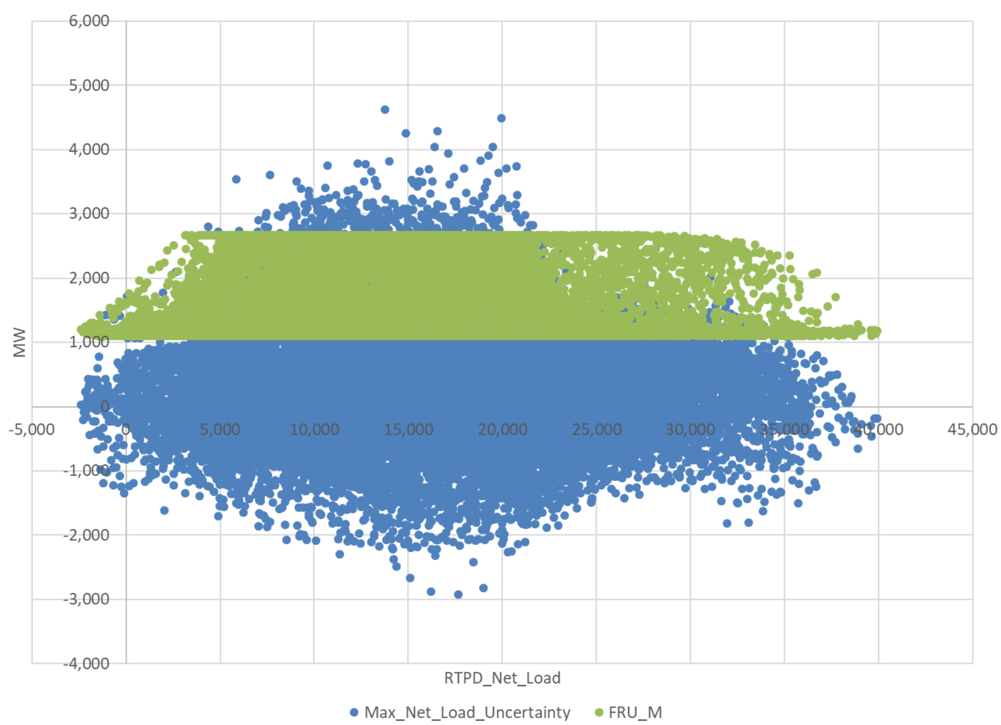


Figure 8 Estimates Histogram vs. Load

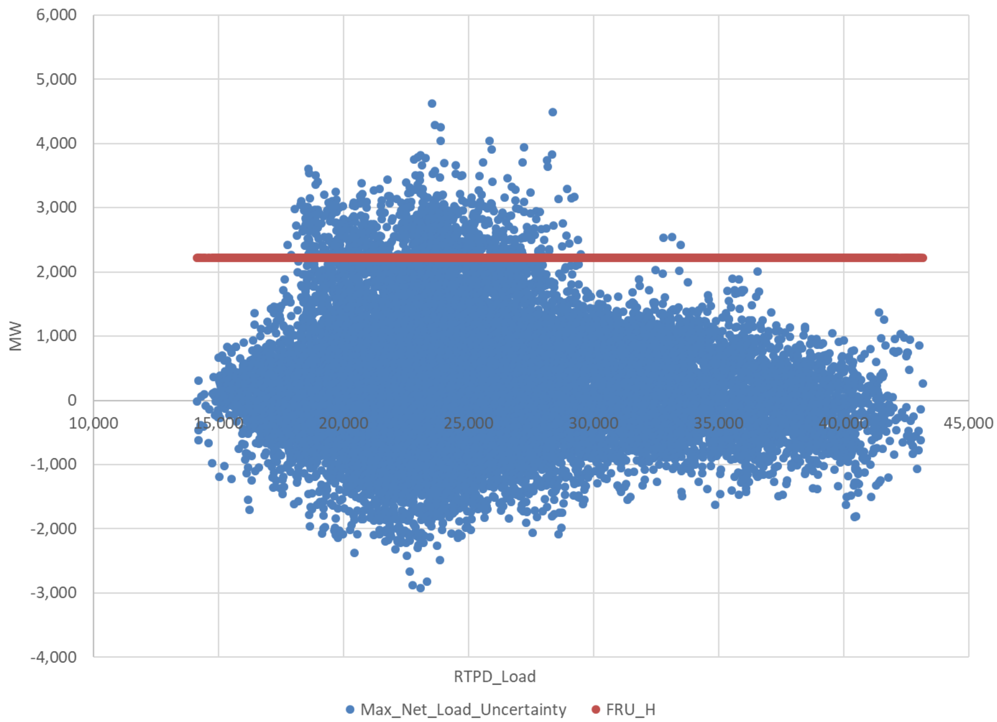


Figure 9 Estimates Mosaic vs. Load

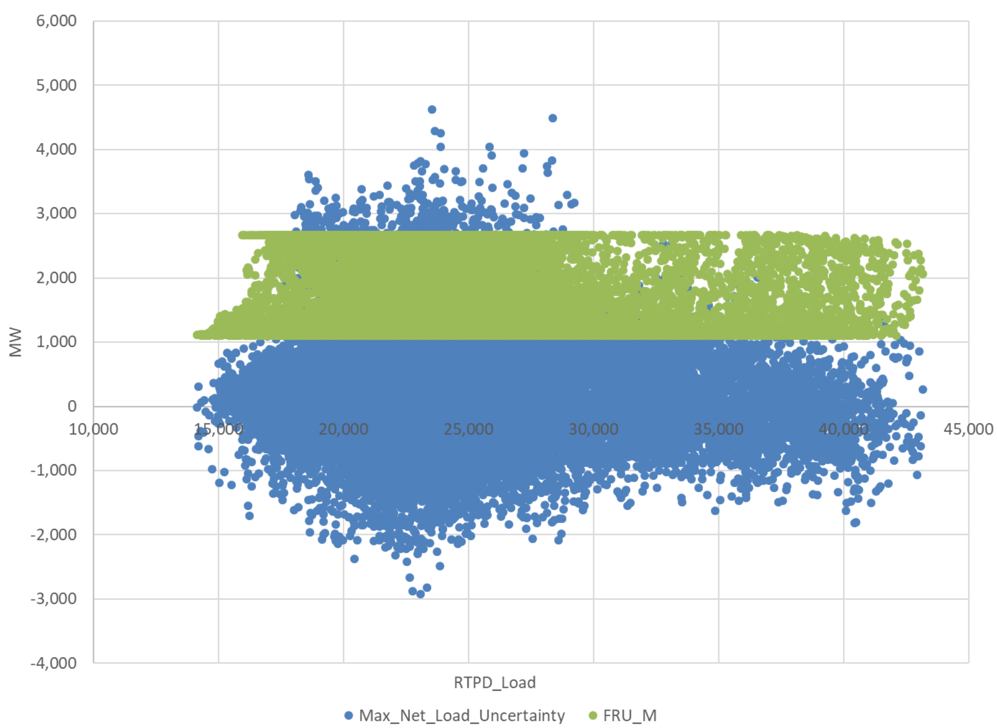


Figure 10 Estimates of Histogram vs. Wind

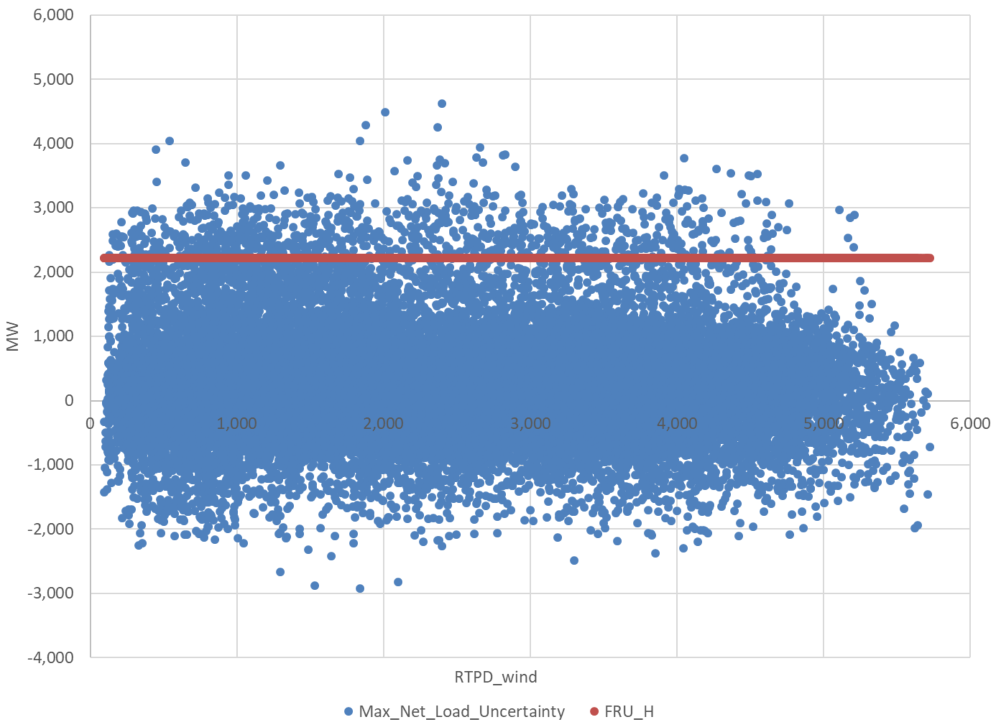


Figure 11 Estimates of Mosaic vs. Wind

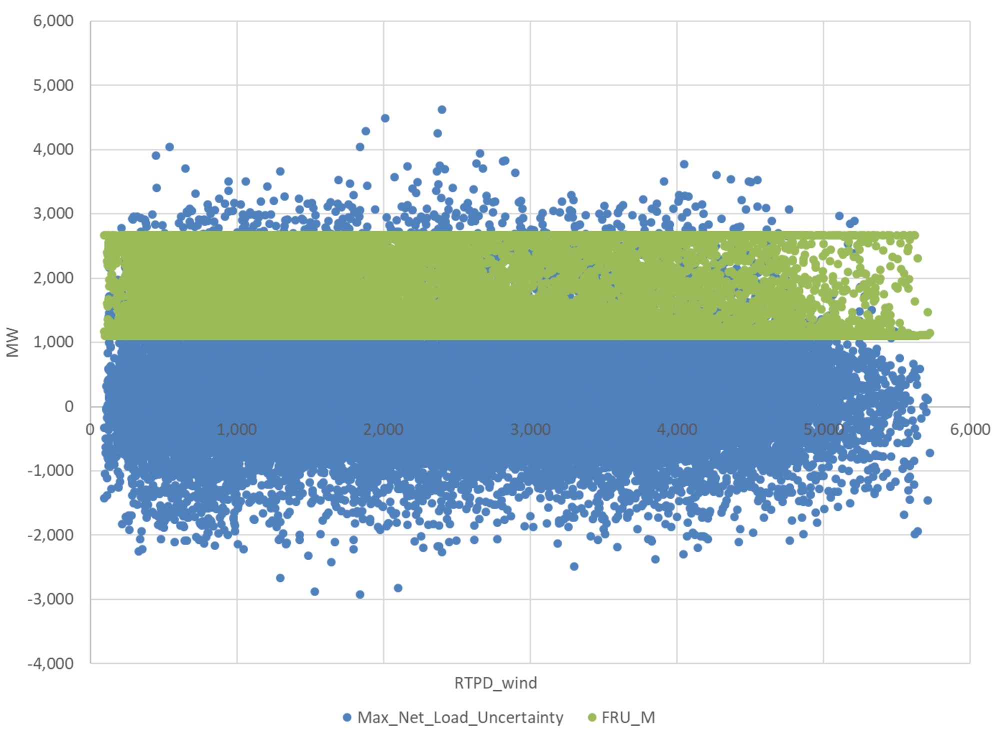


Figure 12 Estimates of Histogram vs. Solar

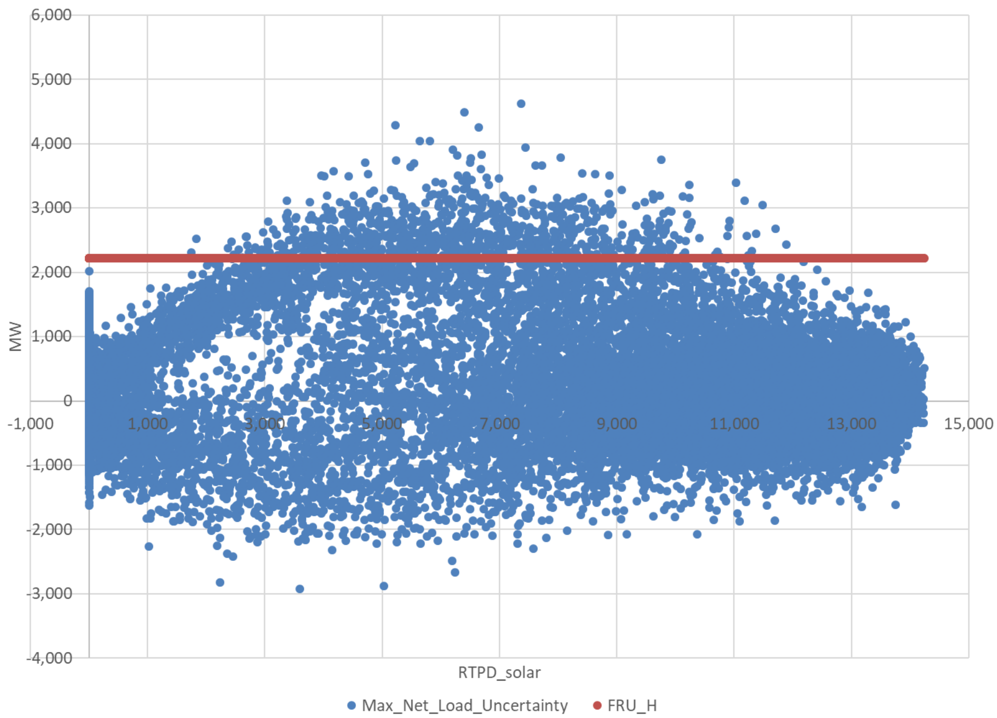
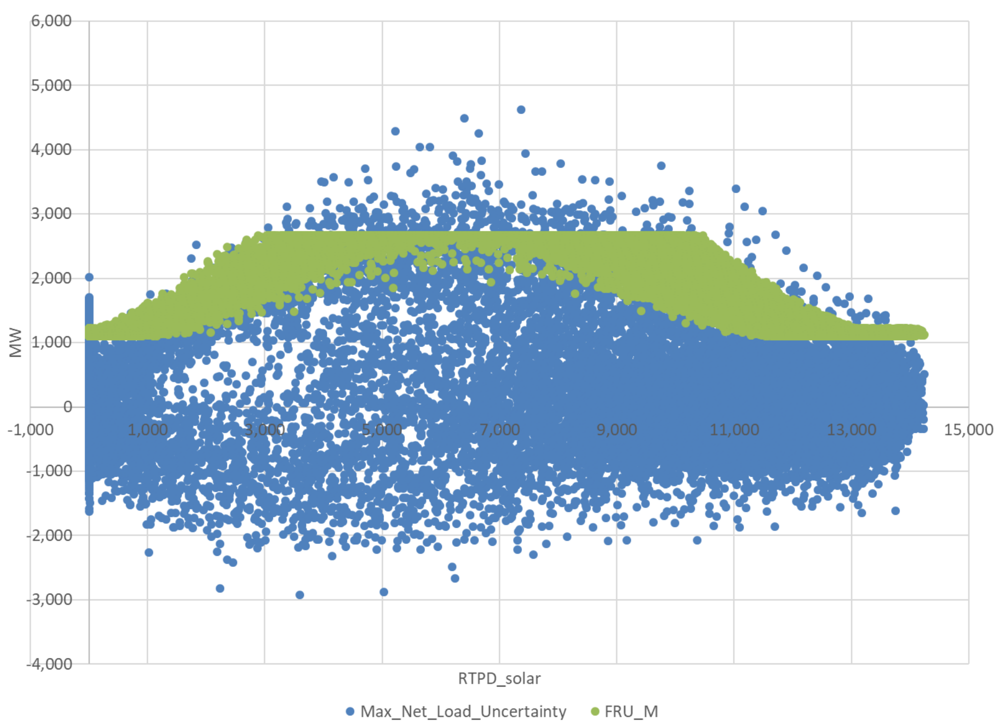


Figure 13 Estimates of Mosaic vs. Solar



7. Simulation Study

This section covers details of the simulation study by comparing requirements calculated with the current Histogram approach and the newly formed Mosaic Quantile Regression. Within the simulation study the ISO has ran results for the WEIM balancing areas, including ISO area. The study covers the period from January 1, 2021 to December 31, 2021, and relies on the historic data collected for the period from January 1, 2020 to December 31, 2021. The data set for each day in the study period consists of historical data with matching weekday and weekend with rolling 180 previous days. The uncertainty definition for each RTPD 15-minute interval is the maximum of the RTD binding forecasts in the three associated RTD 5-minute intervals minus RTPD advisory forecast for FRU, and the minimum of the RTD binding forecasts in the three associated RTD 5-minute intervals minus RTPD advisory forecast for FRD. The Table 1 below lists the load, wind, and solar profiles of the EIMs in this study.

Table 1 Profiles of EIMs Load, Wind, and Solar

BAA name	first wind date	first solar date	wind pmax	solar pmax	load max	wind pcent	solar pcent	load pcent
APS	10/1/2016	10/1/2016	399.20	560.00	7627.43	4.65%	6.52%	88.83%
BANC		4/3/2019		189.28	4387.95	0.00%	4.14%	95.86%
BCHA	4/4/2018		755.17		11887.00	5.97%	0.00%	94.03%
CISO	5/1/2014	5/1/2014	4386.19	5153.10	47484.16	7.69%	9.04%	83.27%
IPCO	4/4/2018	4/4/2018	718.57	289.50	4100.46	14.07%	5.67%	80.27%
LADWP	4/1/2021	4/1/2021	426.63	1154.00	4893.04	6.59%	17.83%	75.58%
NEVP	12/1/2015	12/1/2015	149.10	341.50	9444.36	1.50%	3.44%	95.06%
NWMT	6/16/2021	6/16/2021	462.54	18.70	1912.25	19.32%	0.78%	79.89%
PACE	10/15/2014	12/17/2015	1520.75	130.00	9413.13	13.75%	1.17%	85.08%
PACW	10/15/2014	6/1/2016	617.85	5.00	4227.01	12.74%	0.10%	87.16%
PGE	10/1/2017		716.70		4531.45	13.66%	0.00%	86.34%
PNM	4/1/2021	4/1/2021	957.00	392.20	2533.19	24.65%	10.10%	65.25%
PSEI	10/1/2016		375.10		5032.94	6.94%	0.00%	93.06%
SRP	4/1/2020	4/1/2020	128.00	136.00	7746.96	1.60%	1.70%	96.70%

The recently on-boarded WEIM entities, including Balancing Authority of Northern California (BANC), Los Angeles Department of Water and Power (LADWP), Northwestern Energy (NWMT), and Public Service of New Mexico (PNM) (highlighted in red and orange above) have only results for November and December 2021 reported in this simulation study, but they are excluded in later comparisons for the different sampling schemes. Salt River Project (SRP) (highlighted in red above) has joined the ISO as a WEIM starting on April 1, 2020, three months short of the data range (January 2020-December 2021) collected for other WEIMs, SRP is in the study with the full range of 2021, but not in comparison of different sampling scheme.

8. Performances and Benefits of Using the Mosaic Model

The following four measurements were developed to assess and compare the performance of different models.

i. Coverage: This is used to check the validity of a model, and is the coverage of observed uncertainty against the estimated requirement. The uncertainty requirement is targeted for 95%, which is achieved with 97.5% for upward and 2.5% for downward requirement.
ii. Requirement: This is the average of estimated requirement over a period of time.
iii. Closeness: This is defined as the average distance between the observed uncertainty and the estimated requirement.
iv. Exceeding: This is the average MW difference when the observed uncertainty is exceeding the estimated requirement. The exceeding reflects the reliability cost.

Since all these performance measurements are based on averages, they cannot reveal the effectiveness of incorporating weather information as shown the example in the section 6.

Table 2 below compares coverages of flexible ramping estimates between Histogram (H) and mosaic approaches. As shown above, coverage is used to check the validity of the model, if the target is to ensure the methodology remains within the 95^{th} percentile (97.5% for up and 2.5% for down) a 97% in coverage is achieved when the requirement is within the targeted range 97% of the time.

Table 2 Coverage: Histogram vs. Mosaic

BAA	FRU_H	FRU_M	FRD_H	FRD_M
APS	95.78%	95.27%	96.27%	95.27%
BANC	97.23%	96.70%	95.47%	96.32%
BCHA	96.84%	95.39%	96.86%	96.00%
CISO	96.41%	95.44%	96.47%	95.59%
IPCO	96.72%	95.70%	97.16%	96.23%
LADWP	98.07%	98.03%	97.69%	96.95%
NEVP	94.74%	94.22%	96.47%	95.60%
NWMT	96.56%	94.39%	97.86%	96.46%
PACE	96.31%	95.14%	96.07%	94.93%
PACW	97.09%	96.28%	96.58%	95.85%
PGE	97.14%	96.33%	97.17%	96.41%
PNM	98.92%	98.72%	99.69%	99.01%
PSEI	96.74%	95.76%	96.95%	96.11%
SRP	95.34%	94.95%	95.80%	94.70%

To further assess the average coverage shown above, Figure 14 and 15 below show the difference of the two methodologies by month. The monthly look of coverages highlights a smoother profile with the mosaic approach over the year, the number value 2 indicates the sample scheme in this simulation is the second sampling scheme to be discussed in section 10.

Figure 14 Examine seasonality of Histogram

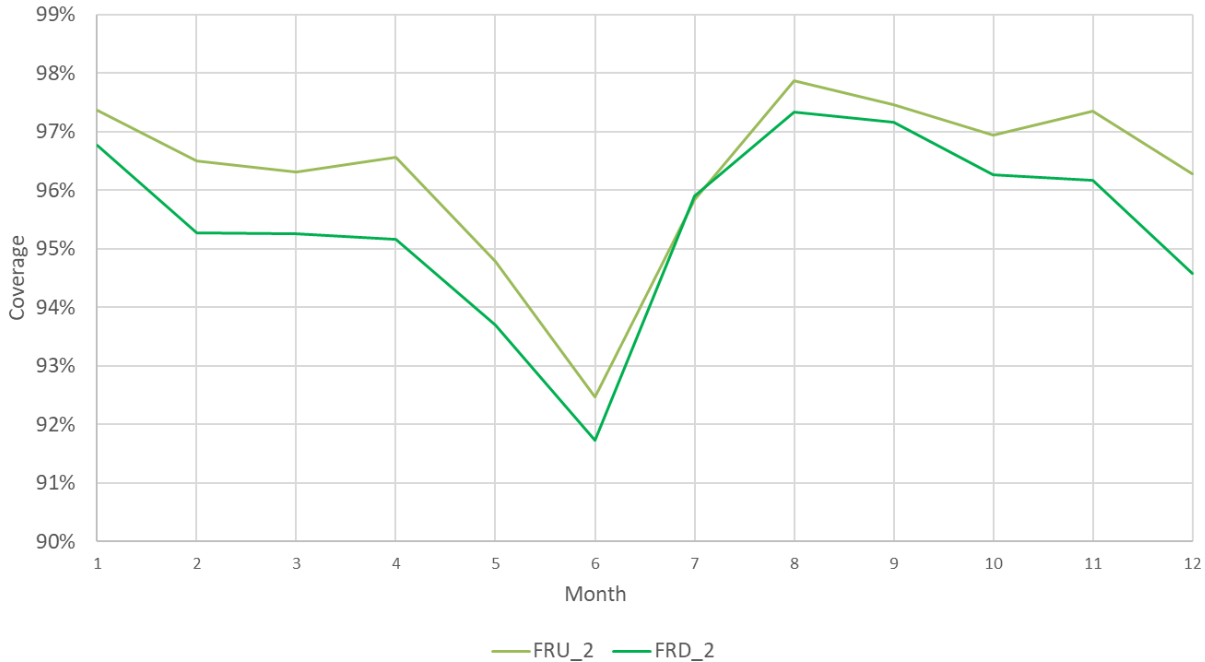


Figure 15 Examine seasonality of Mosaic

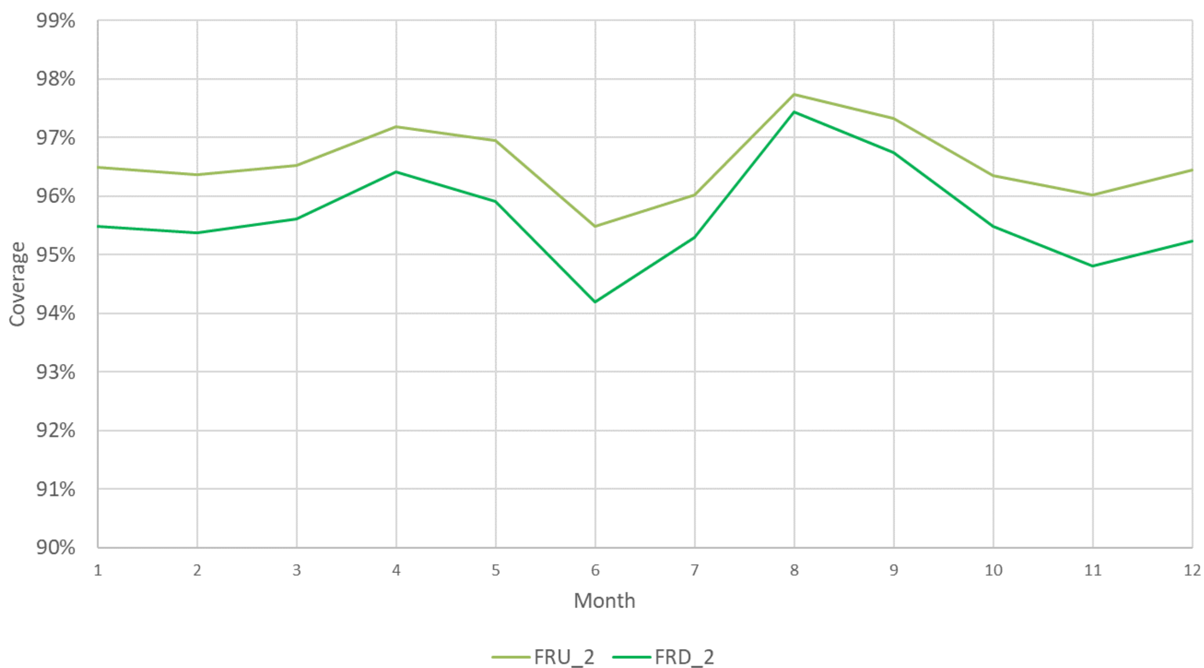


Table 3 below compares the requirements of flexible ramp estimates between the Histogram approach and the Mosaic approach. It shows that the Mosaic approach has a lower average requirement for all EIM areas over the one year study period. The reductions vary among WEIM areas because the inherent differences within the WEIM areas' profiles of load, wind, and solar as shown in Table 1. Even with these lower average requirement, Mosaic can still manage to achieve a comparable coverage as shown in Table 3.

Table 3 Requirement: Histogram vs. Mosaic

BAA	FRU_H	FRU_M	FRD_H	FRD_M
APS	150.68	135.89	-127.24	-117.99
BANC	60.52	41.45	-49.14	-43.81
BCHA	157.49	151.57	-169.00	-161.67
CISO	1142.37	1042.13	-943.51	-850.52
IPCO	105.89	101.74	-132.72	-124.42
LADWP	152.43	147.32	-148.52	-135.85
NEVP	165.02	141.58	-139.53	-129.69
NWMT	81.15	77.15	-98.52	-91.95
PACE	250.80	241.12	-286.39	-273.01
PACW	112.55	106.14	-98.53	-92.13
PGE	130.70	121.66	-118.67	-112.25
PNM	136.49	137.04	-166.43	-161.23
PSEI	94.00	90.04	-101.46	-98.19
SRP	113.68	102.66	-109.17	-97.01

Table 4 below compares the closeness of flexible ramp estimates between the Histogram and mosaic approaches. As described above, closeness is defined as the average distance between the observed uncertainty and the estimated requirement. This comparison shows that estimated requirements with the mosaic approach generally is closer to the observed uncertainty.

Table 4 Closeness: Histogram vs. Mosaic

BAA	FRU_H	FRU_M	FRD_H	FRD_M
APS	133.49	119.11	120.32	111.92
BANC	53.59	34.69	44.91	39.56
BCHA	138.48	134.05	148.10	141.64
CISO	891.44	798.23	931.16	843.43
IPCO	105.86	102.24	118.68	110.85
LADWP	138.40	133.07	136.17	123.58
NEVP	146.10	122.81	126.87	117.91
NWMT	80.56	77.46	89.32	83.16
PACE	239.31	231.18	256.79	245.03
PACW	99.34	93.23	93.11	86.90
PGE	113.65	105.10	118.71	112.60
PNM	128.85	129.68	155.25	150.30
PSEI	86.46	82.97	93.10	90.20
SRP	99.86	89.10	99.15	87.65

Lastly, Table 5 below compares the exceeding of flexible ramp estimate between the Histogram and the Mosaic approaches.

Table 5 Exceeding: Histogram vs. Mosaic

BAA	FRU_H	FRU_M	FRD_H	FRD_M
APS	39.97	39.99	38.30	39.22
BANC	14.47	14.73	6.46	7.73
BCHA	37.96	42.12	46.00	47.07
CISO	235.30	262.11	242.17	253.37
IPCO	37.52	34.85	34.26	32.12
LADWP	52.74	45.95	36.60	28.99
NEVP	45.73	42.87	39.94	42.00
NWMT	26.20	24.09	17.58	16.31
PACE	71.96	70.52	86.07	82.69
PACW	32.86	29.64	30.01	26.93
PGE	40.46	38.23	47.20	41.53
PNM	19.69	27.16	22.33	19.44
PSEI	28.59	27.58	29.28	27.76
SRP	26.51	27.11	32.15	31.78

The exceeding measures the MW difference when the actual uncertainty is exceeding the requirement. The lower exceeding means the requirement performs better at anticipating the realized uncertainty. The Mosaic Quantile Regression model with the square factor provides comparable but mixed results. There is another variation of the Mosaic Quantile Regression which produces smaller exceeding numbers see discussion in section 11.

Compare to the histogram estimates, the overall performance of mosaic regression estimate can be summarized as the follows: the coverage is comparable, the average requirement is slightly lower, while closeness slightly improves. Lastly, the exceedance slightly increase with the Mosaic relative to the current histogram. This can be addressed by using a different Mosaic model as described in section 11 of the paper.

To summarize the results of this simulation study, Mosaic methodology is effectively incorporating the weather conditions into the estimation of uncertainty requirements. It provides similar coverages to the Histogram estimates and reduces seasonality in the coverages, as shown in Figure 14 and Figure 15. All these benefits are achieved with lower requirement on average.

9. Daily Trends

This section provides a review of daily trends to compare the observed uncertainties to the estimated requirements by both Histogram and Mosaic approaches. In order to visualize the difference between H and M, the graph of load, wind, solar are provided. These graphs offer a glimpse of how Mosaic is reflecting changes of the load, wind, and solar components. These graphs in turn can also assist the ISO to fine tune the mosaic regression approach to be applicable in production, e.g., there is a need to add effective threshold to curb extreme estimates impacted by data error or outliers.

Figure 16 APS 06/05/2021

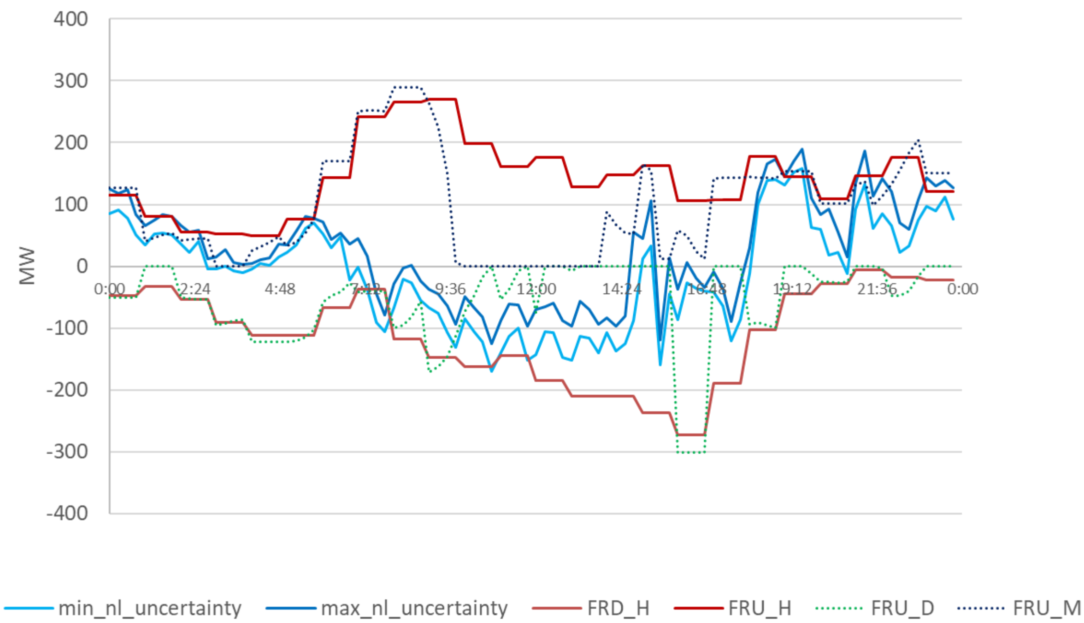


Figure 17 APS 06/05/2021

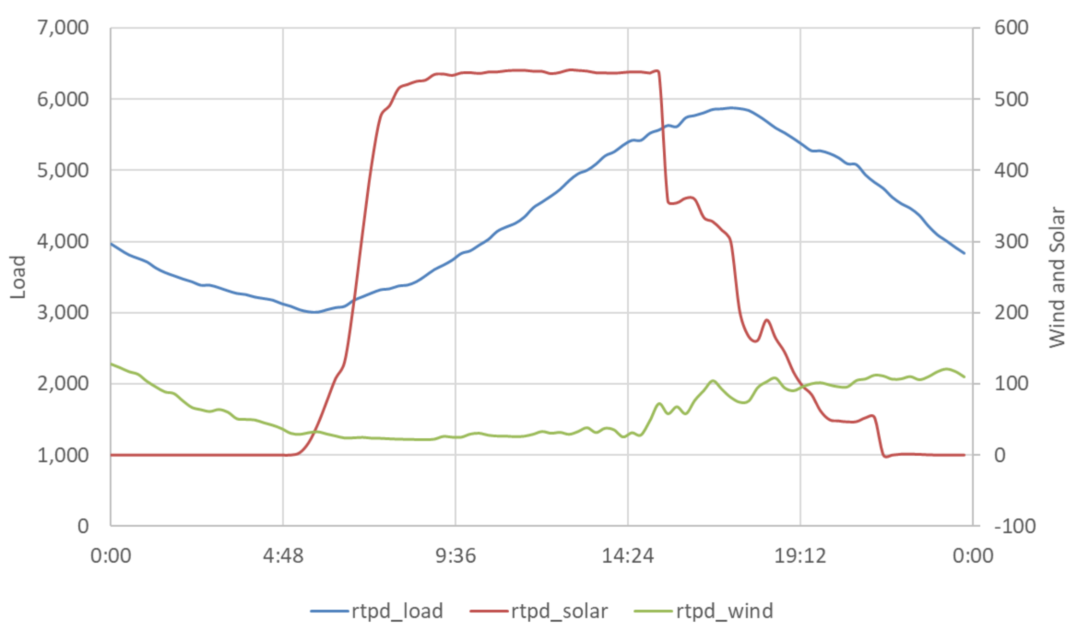


Figure 18 CISO 06/01/2021

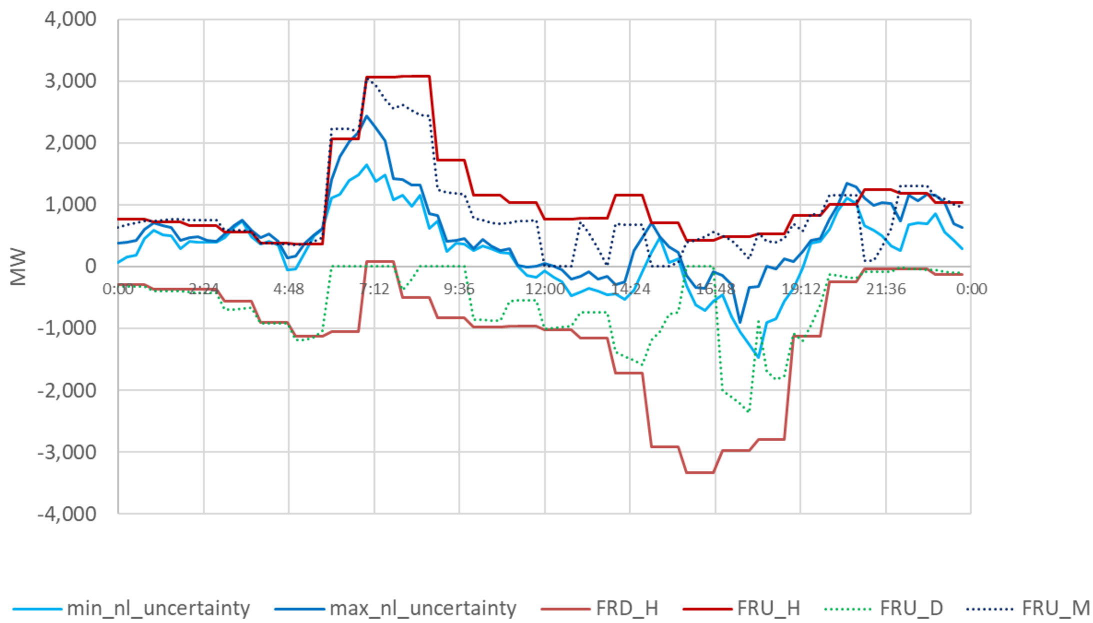


Figure 19 CISO 06/01/2021

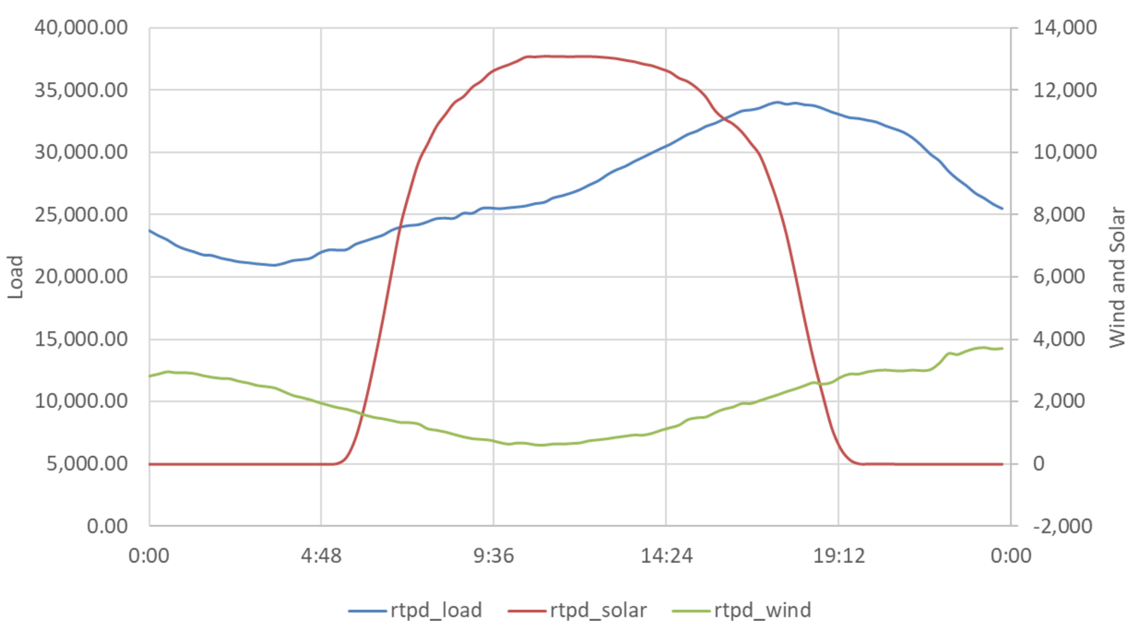


Figure 20 IPCO 08/01/2021

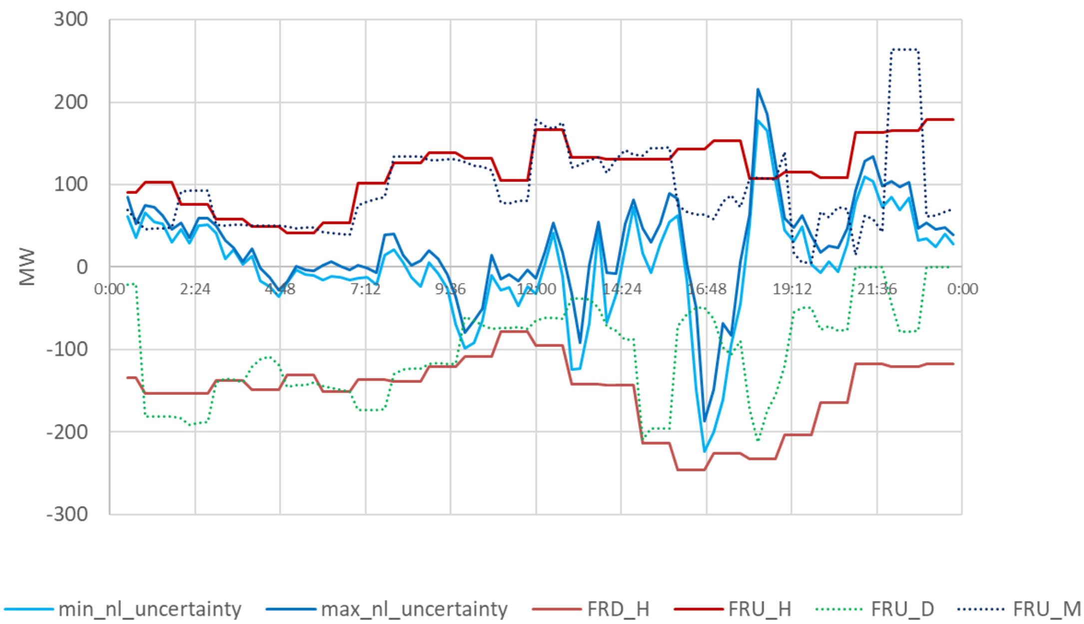


Figure 21 IPCO 08/01/2021

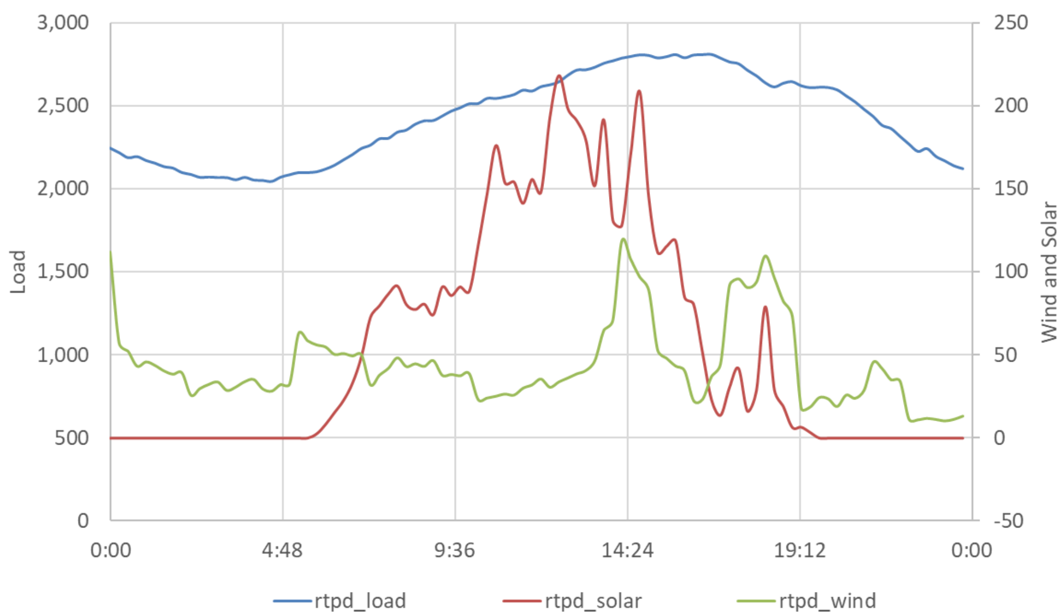


Figure 22 NEVP 05/01/2021

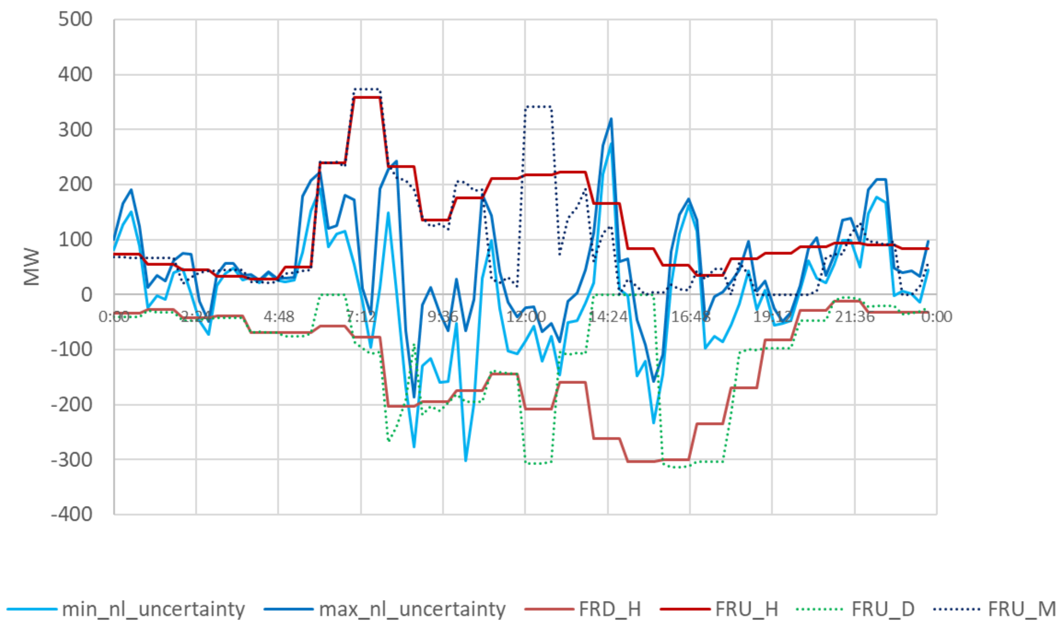


Figure 23 NEVP 05/01/2021

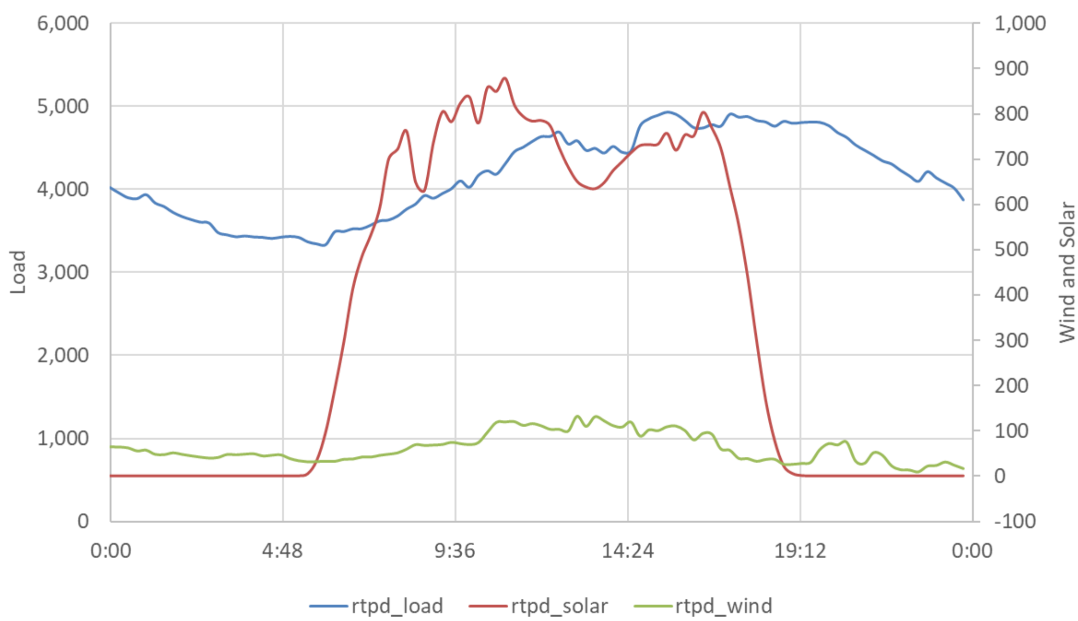


Figure 24 PACE 09/27/2021

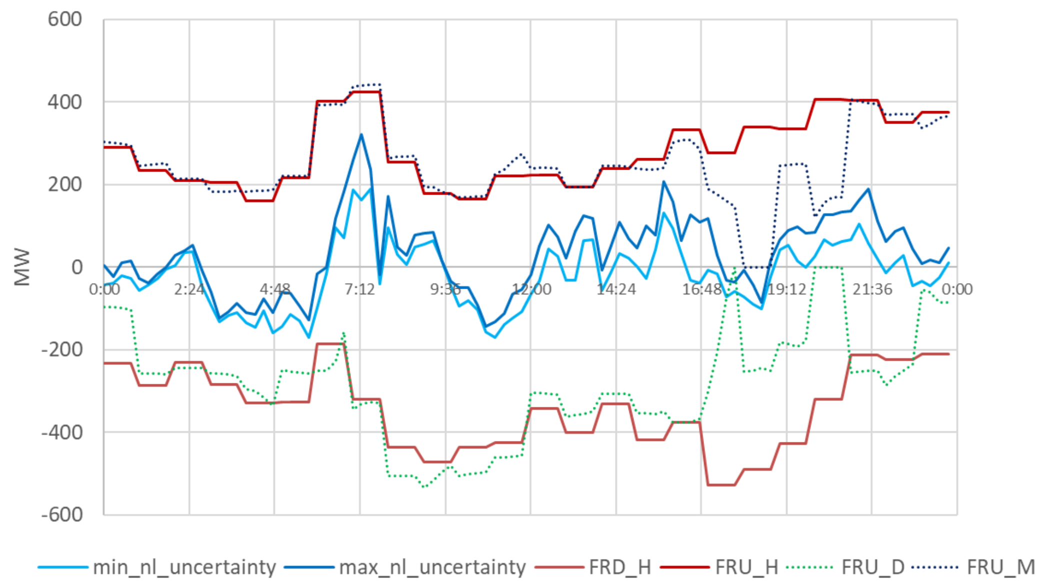


Figure 25 PACE 09/27/2021

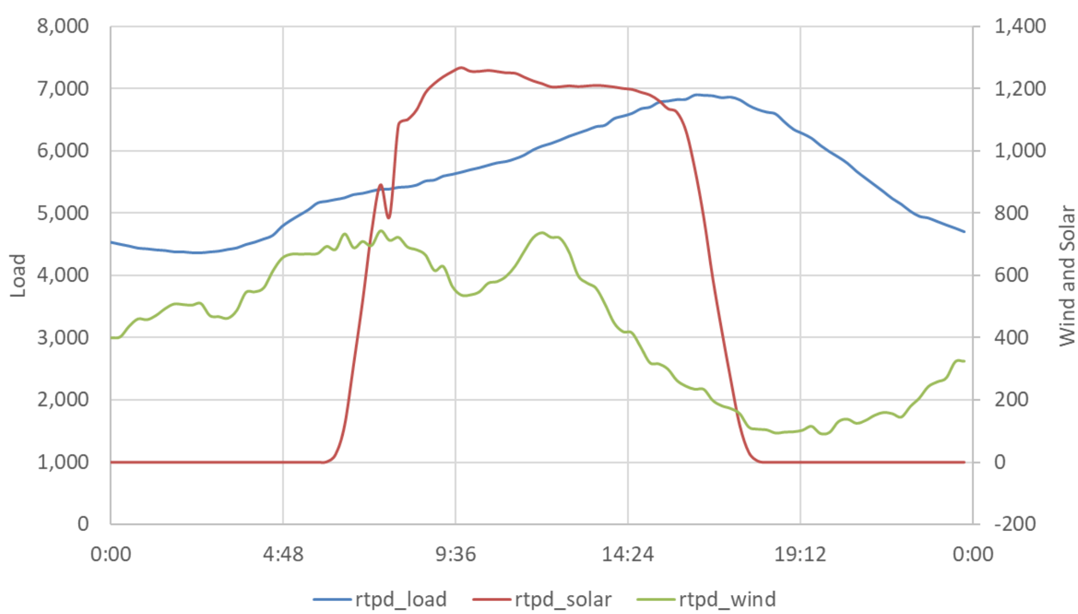


Figure 26 PACW 01/15/2021

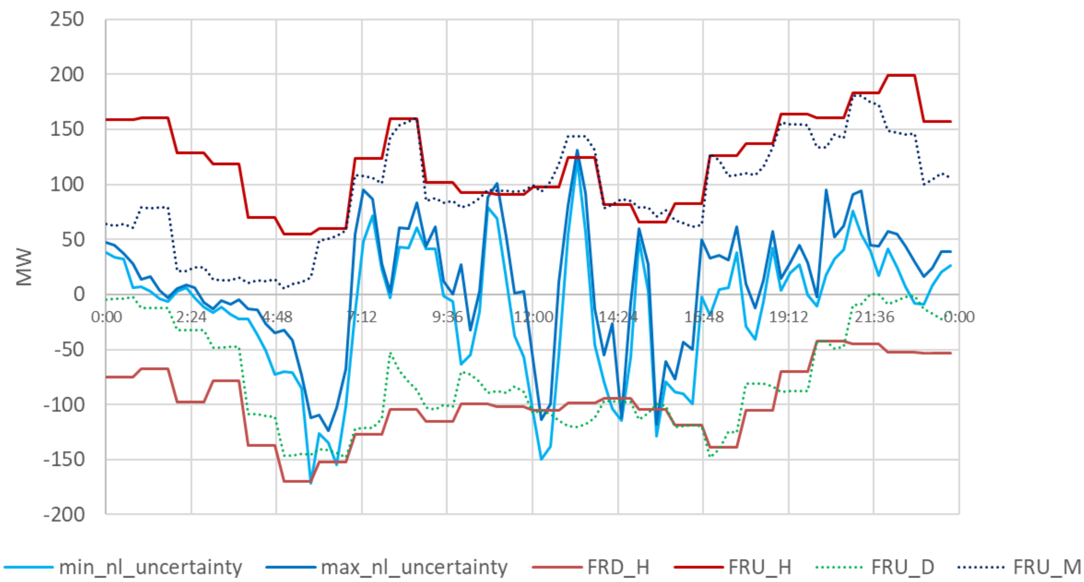


Figure 27 PACW 01/15/2021

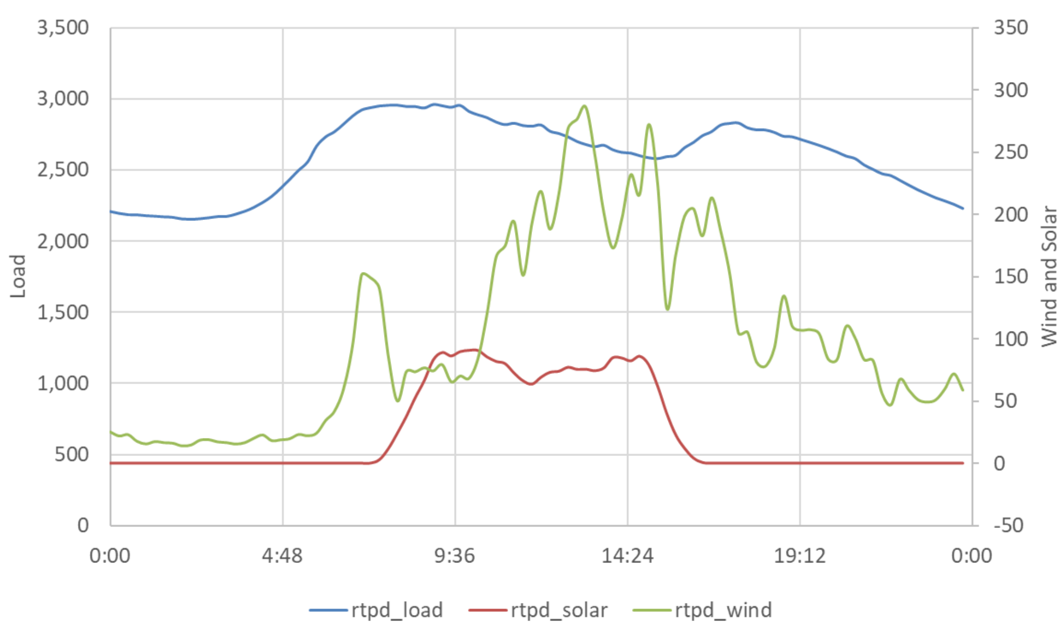


Figure 28 PGE 06/26/2021

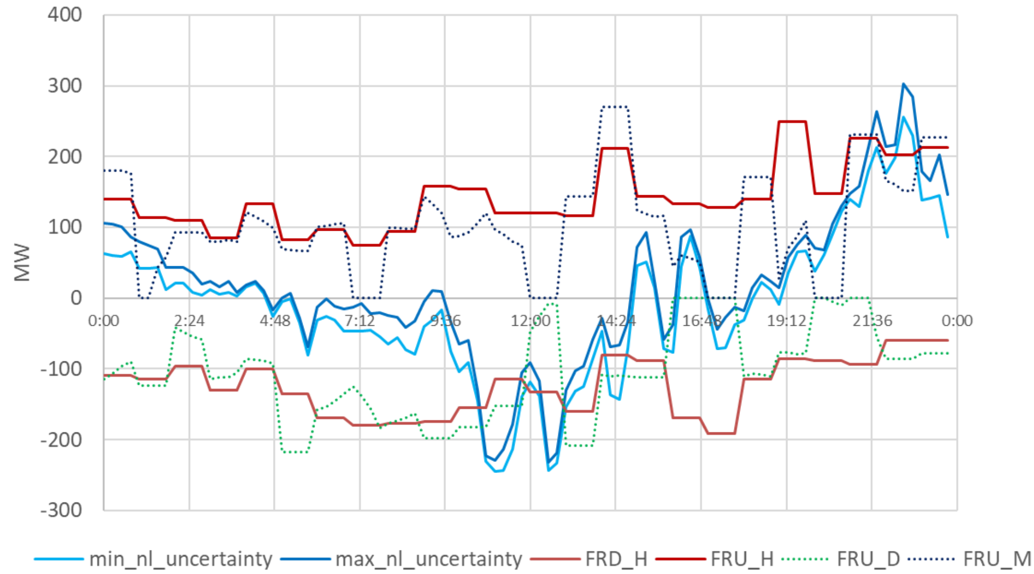


Figure 29 PGE 06/26/2021

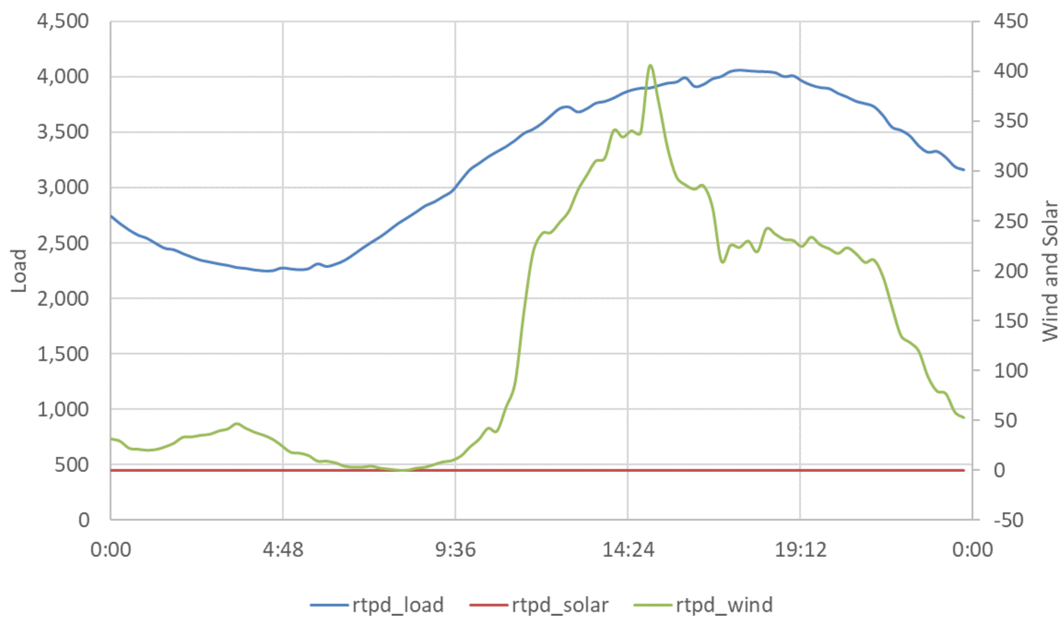


Figure 30 SRP 06/13/2021

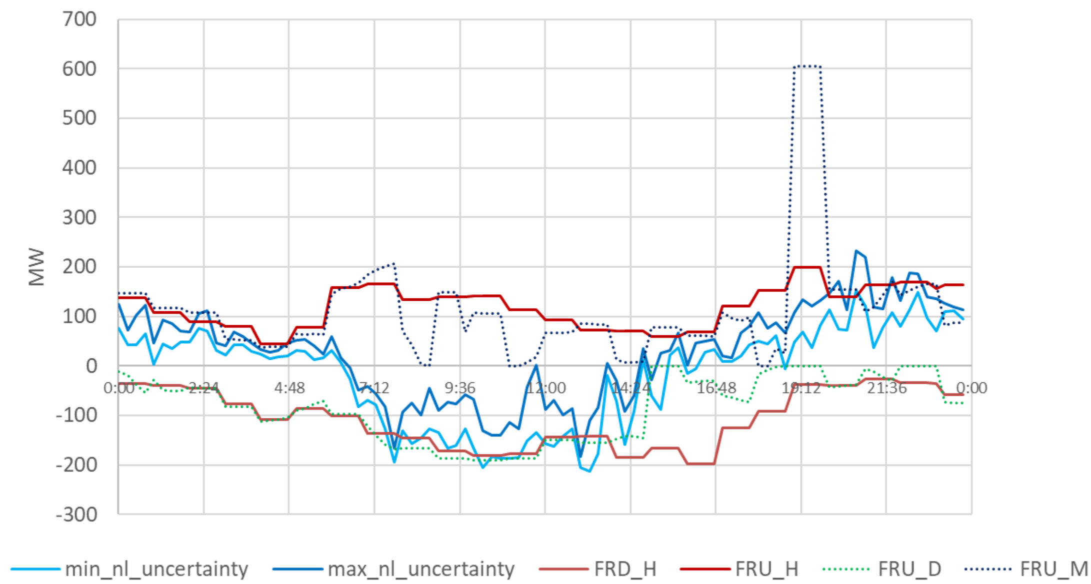
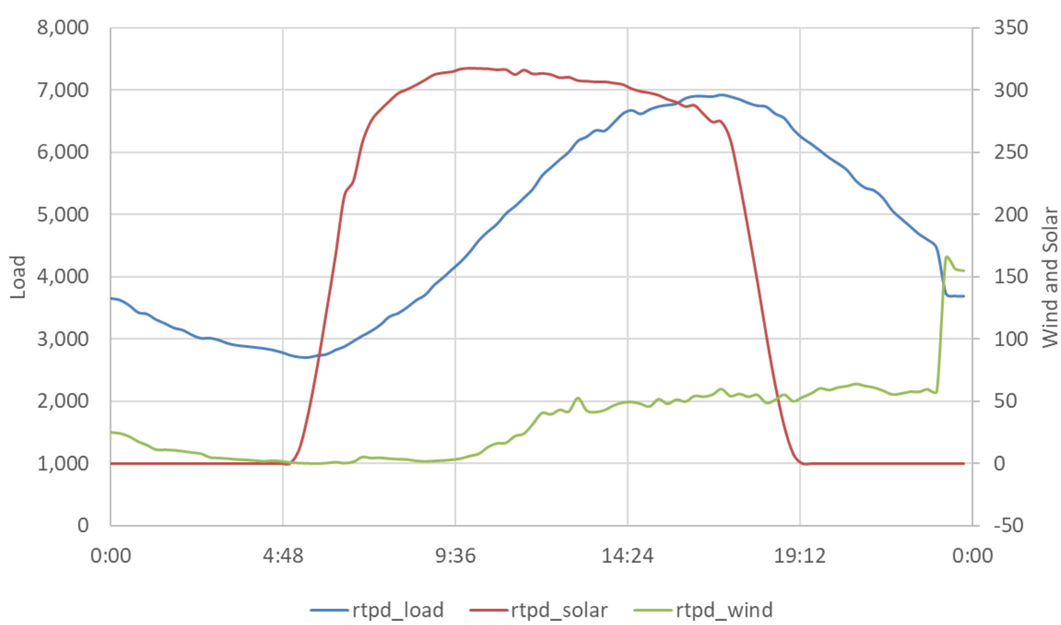


Figure 31 SRP 06/13/2021



10. Discuss Sampling Schemes

The choice of sampling scheme needs to balance out its feasibility and benefit. A sample consisting of fewer days in the data set may not adequately reflect the forecasted weather conditions and bring unwanted variability of the estimated requirement as well as the target coverage. A data sample with too many days may incorporate seasonal and weather conditions not reflective of the time being assessed and bring heterogeneity in the training data, it will lead larger production burden.

Figures 32 and 33, show that a small sample may bring better fitness to Quantile Regression model. The left graph uses the quadratic terms of RTPD as inputs in a year, the right graph adds the month as the stratification. A larger data set may offer some stability in regression, while a small sample may bring better fitness since it aligns up better to the current weather condition.

Figure 32 Month quantile regression

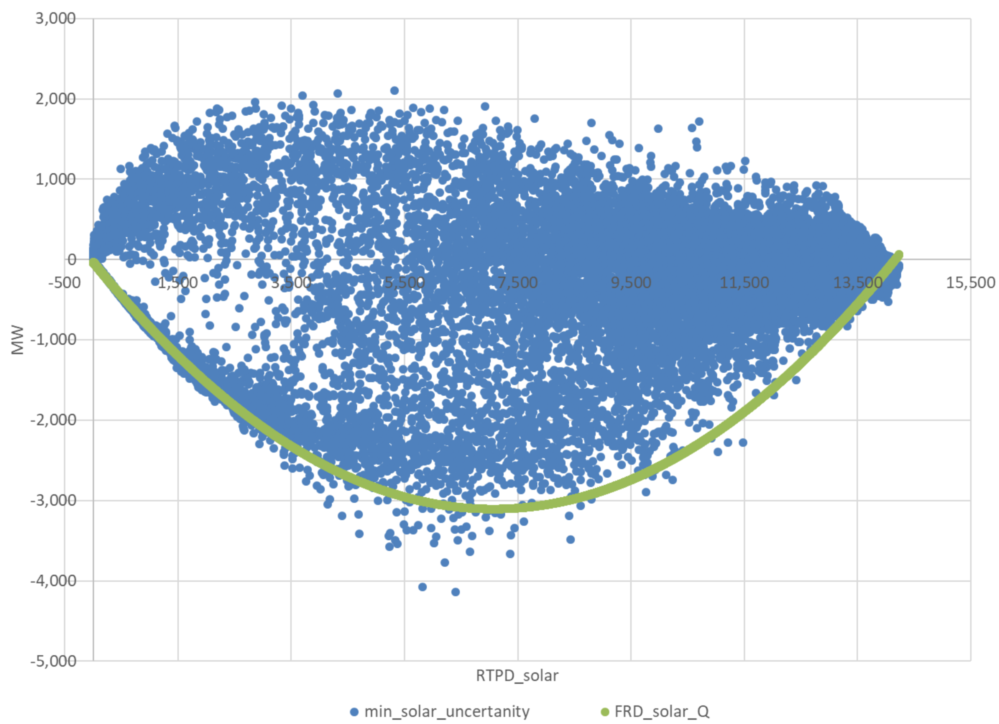
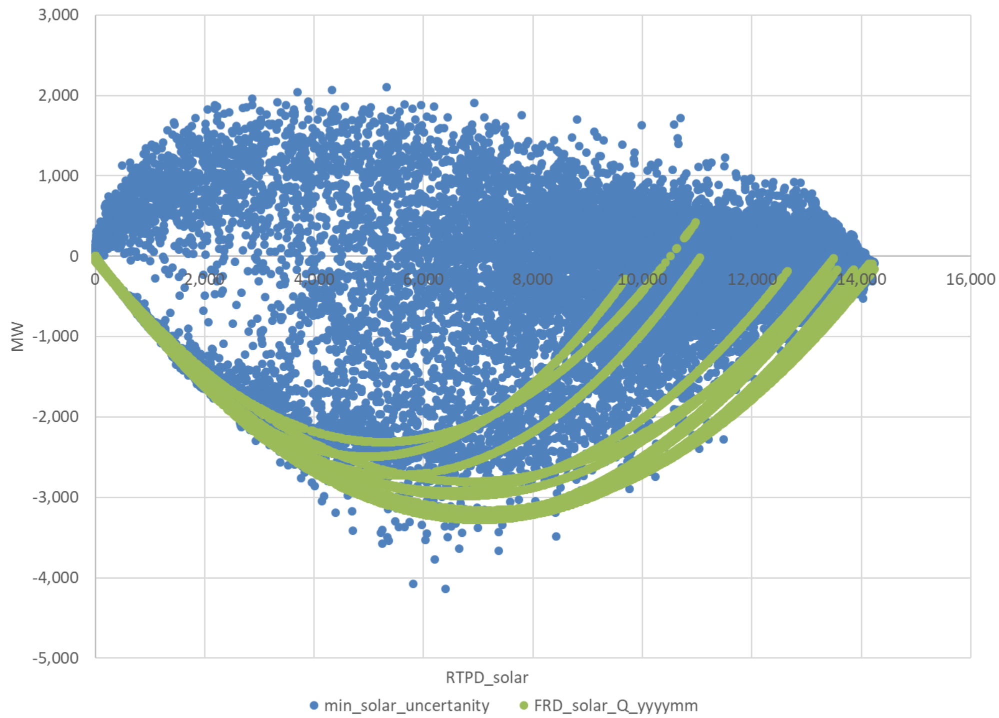


Figure 33 Yearly quantile regression



Three different sampling schemes were explored in this study.

Sampling scheme 1: Rolling previous 40 matching weekdays and 20 matching weekends. This is the sampling used in the ISO's current Histogram approach.

Sampling scheme 2: A fixed 180 rolling days with varying number of weekdays and weekend (holidays included). The increased sample size will bolster the robustness of regression computation.

There are many different choices of sampling schemes available to experiment with. Since these sample schemes may suffer impact from seasonality, ISO also explored a third approach.

Sampling scheme 3: In addition to the sampling scheme 1, the forward historical data in last year anchored from a date similar to the current day with matching weekday/weekend will be put into the training sample. The sampling scheme 3 balances out backwards and forward data for any given day. The third sampling schemes may require more storages but an implementation of smaller sample sizes can ease the computational burden.

The following graphs will help to better understand the complexity of the selection of sample schemes. Three schemes are considered, namely, 1) 40/20 weekday/weekend, 2) 180 d, and 3) balanced 40/40 weekday/weekend. Finally, the performance measurements for the EIMs are compared again to show average of coverage, requirement, closeness, and exceeding. These graphs are plotted throughout the following sections to show the additional considerations alongside each other.

Figures 34 and 35 display the coverage of various sample schemes over the months in the study period for both Histogram and Mosaic approaches. The sampling scheme 3 with red lines has provided the highest coverages in the simulation. Moreover, Mosaic approach generally produces the flatter with more stable coverages over the months, assisting with the seasonality changes.

Figure 34 Coverage comparison among the sampling schemes – Histogram approach

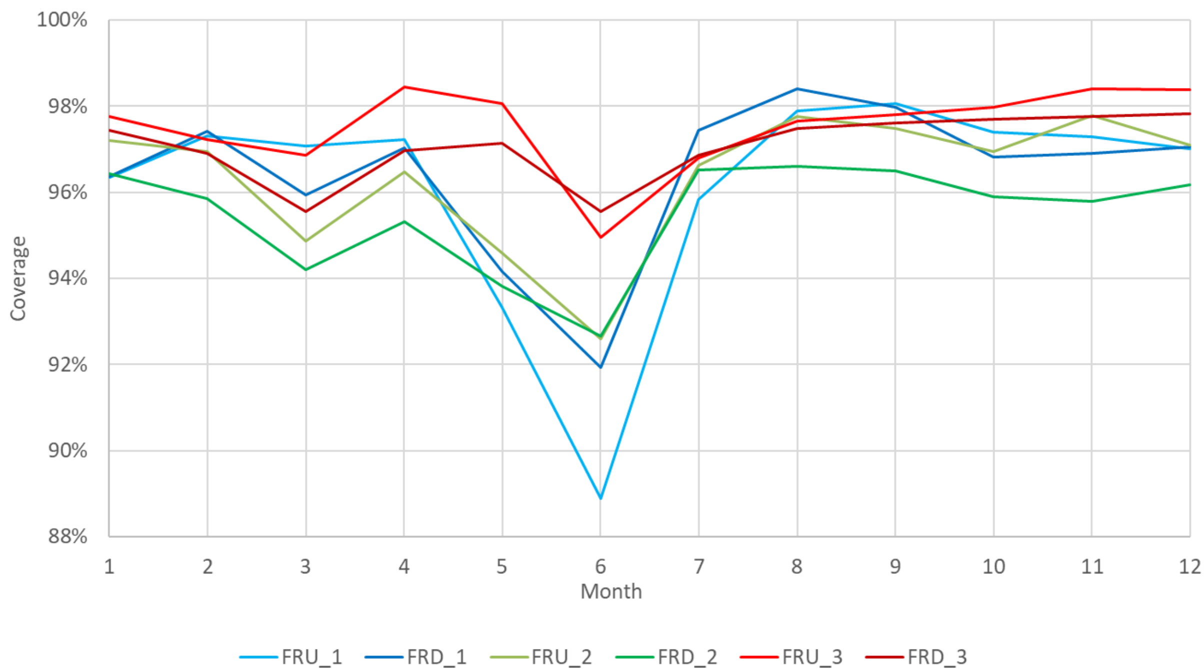
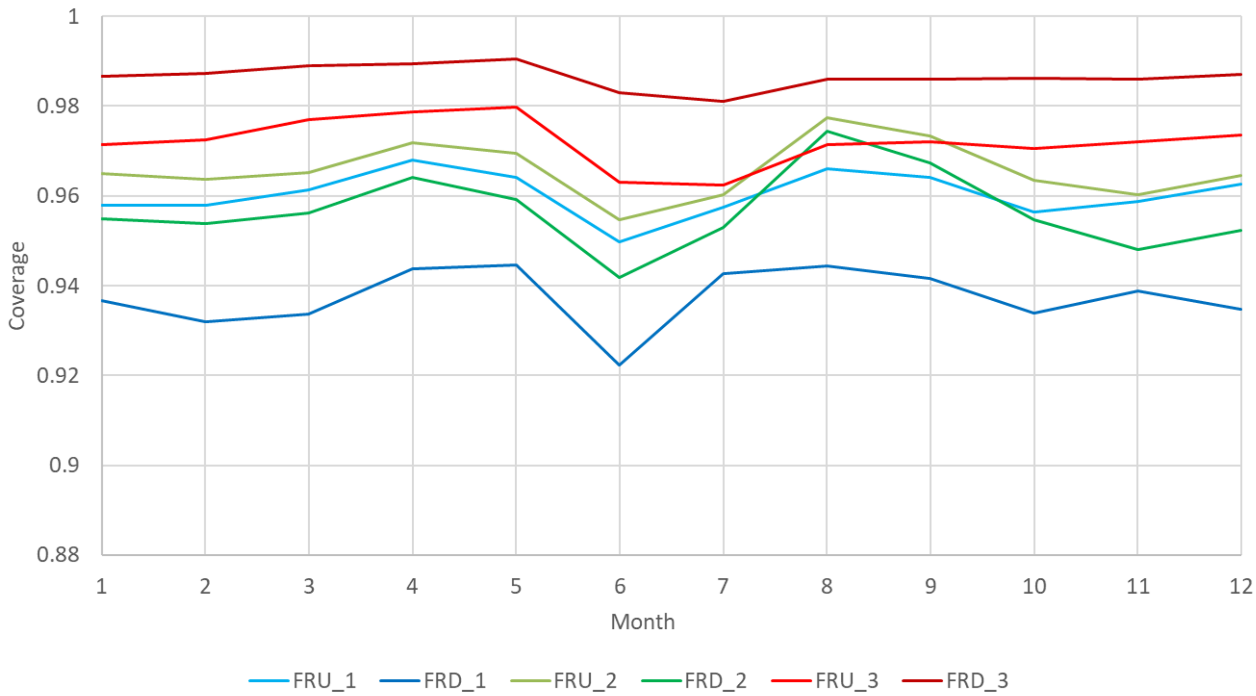


Figure 35 Coverage comparison among the sampling schemes - Mosaic approach



11. Linear Model as an Alternative

As described in earlier sections, there are two choices of input in the final mosaic model. The Mosaic model with square items is described in section 8 in the forms below

$$NL_M = a_m + b_m * mosaic + c_m * mosaic^2$$

This squared mosaic model produces a high number of exceedance for some WEIMs. The ISO has noticed that the linear Mosaic model can reduce such variability in exceedance estimates from the square Mosaic model.

$$NL_M = a_m + b_m * mosaic$$

To further combine the additional considerations together the performance matrix below highlights the linear and square Mosaic model alongside the sampling schemes mentioned in section 13. Finally, the performance measurements of coverage, requirement, closeness, and exceeding points out the linear model is another viable and simple model looking at the average of all the WEIMs are compared to show average of coverage, requirement, closeness, and exceeding between the two choices of the inputs into the final mosaic model.

The comparison of these three sampling scheme are in the following tables, 6 to 9, for the average of all EIMs with different sampling schemes. For example, the linear 1 indicates the combination linear mosaic model with sampling scheme 1.

Table 6 coverages for linear and square mosaic models

sampling_scheme	FRU_H	FRU_M	FRD_H	FRD_M
linear 1	95.93%	94.78%	96.04%	95.09%
linear 2	96.31%	95.99%	96.58%	96.14%
linear 3	97.07%	96.60%	97.20%	96.86%
sqaure 1	95.93%	93.33%	96.04%	93.75%
sqaure 2	96.31%	95.45%	96.58%	95.67%
sqaure 3	97.07%	98.56%	97.20%	98.65%

The linear and square model produce similar coverages.

Table 7 Requirements for linear and square mosaic models

sampling_scheme	FRU_H	FRU_M	FRD_H	FRD_M
linear 1	225.55	214.16	-203.16	-195.96
linear 2	242.32	228.79	-222.62	-211.25
linear 3	237.32	225.97	-217.34	-208.94
square 1	225.55	205.67	-203.16	-187.69
square 2	242.32	223.45	-222.62	-205.69
square 3	237.32	290.72	-217.34	-267.00

When looking at the average requirement, it is observed that the mosaic linear model has smaller average requirements as compared to the ones needed in the Flexible R Histogram approach. In addition, the square model has smaller average requirements, with the exception of sampling scheme 3 where it is observed they are slightly higher.

Table 8 Closeness for linear and square mosaic models

sampling_scheme	FRU_H	FRU_M	FRD_H	FRD_M
linear 1	189.11	179.04	191.41	185.25
linear 2	205.40	192.26	210.60	199.71
linear 3	199.53	188.53	204.13	195.97
square 1	189.11	173.15	191.41	179.19
square 2	205.40	187.80	210.60	194.81
square 3	199.53	251.14	204.13	252.17

The Mosaic linear model has smaller closeness number as compared to its counterpart in the Histogram approach. The square model has smaller closeness number except for sampling scheme *3.*

Table 9 Exceeding for linear and square mosaic models

sampling_scheme	FRU_H	FRU_M	FRD_H	FRD_M
linear 1	59.34	58.86	58.47	59.94
linear 2	59.70	58.37	62.52	61.82
linear 3	60.35	56.56	57.80	55.42
square 1	59.34	64.98	58.47	65.29
square 2	59.70	61.53	62.52	62.45
square 3	60.35	60.22	57.80	58.69

Finally, the Mosaic linear model has less exceeding as compared to the ones needed in the Histogram approach as well as the ones produced by the square model.

Overall Assessment

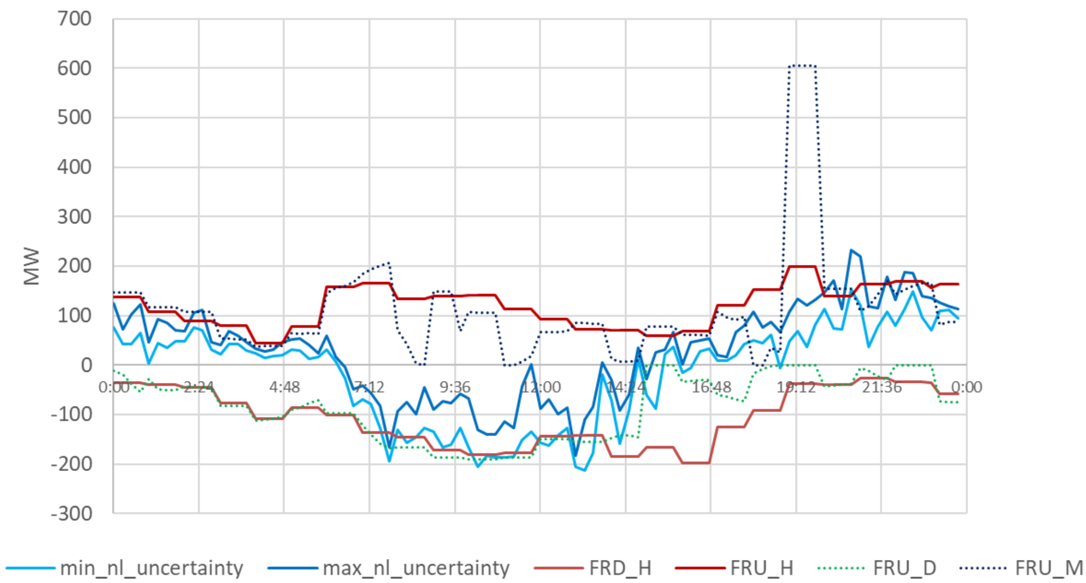
The performance tables above indicate the further simulation studies support the Mosaic linear model with sampling scheme 3 to be best overall choice as there are improvements within in all the performance measures monitored.

12. More Thresholds

Currently, within the Histogram approach the ISO calculates a threshold value for each WEIM Entity and the Total WEIM Balancing Area Authorities looking at the past 90 days at the 98th percentile. The thresholds were designed to handle errors within the data quality feeding the model, outlier errors within forecast, and also assist in providing guidelines for WEIM entities on the maximum up and down requirements they can experience. With Mosaic approach, a bound was developed within the automation, but further consideration after the simulation shows benefits in adding an additional threshold.

In this simulation study, the requirement sporadically observed occasional spikes even though the regression based estimate was bounded by 99% of Histogram estimates. See Figure 36 below,

Figure 36 More overall thresholds are needed



Both regression based estimates FRU and FRD each has two bounds, 99 and 1 percent estimates by Histogram estimate for FRU and FRD. With the inclusion of negative up uncertainty and positive down uncertainty in the data pool for calculating estimate, the zero value was used as a lower bound for FRU and upper bound for FRD, respectively. These zero bounds are noticeable in the above Figure as well as many Figures in the daily graph section. The zero bounds may pose some operational challenges since it may require sudden changes to the procurement of flexible ramping. It seems reasonable to add two additional thresholds, one positive lower threshold for FRU and one negative upper thresholds to replace the zero bounds.

13. Summary and Future Development

Overall this technical paper describes the ISO's proposal to use Mosaic Quantile Regression approach to incorporate weather information into the calculation of uncertainty requirements, including the construction of the net load formulation, and Mosaic Quantile Regression, the comparison of the current Histogram to the newly formed Mosaic Quantile Regression, metrics analyzing the overall impact in the Mosaic Quantile Regression, and lastly some sensitivity analysis on additional considerations the ISO is monitoring. The simulation results show the Mosaic Quantile Regression model generally produces smaller average requirements while keeping the comparable coverage.

Mosaic Quantile Regression approach dampens the impact of seasonality and shows less exceeding assisting with operational costs.

In addition, the ISO's proposed Mosaic methodology can be adapted for the future for ensemble or probabilistic estimates for any component in load, wind, and solar. The ISO will continue to develop and trial to see if the ensemble estimates further improve on the Mosaic Quantile Regression.

$$Ensemble = NL_m + (L_e - L_q) - (W_e - W_q) - (S_e - S_q)$$

$$NL_e = \mathrm{a_e} + \mathrm{b_e} * Ensemble$$

Where the subscript "e" stands for ensemble

14. Detailed Description of the Requirement Calculation

In order to increase transparency on the proposed Quantile methodology and enable interested parties to replicate the calculation, CAISO posted the step-by-step description of the methodology to calculate the FRP requirements. These details are captured in the Business Requirement Specification (BRS) posted on March 11, 2022. The document is available at

http://www.caiso.com/Documents/BusinessRequirementsSpecifications10-FlexibleRampProduct-RequirementsEnhancements.pdf