Flexible Ramping Product Refinements Initiative
Appendix C - Quantile Regression Approach

In this technical appendix, further details on quantile regression will be discussed, outlining the proposed methodology, as well as results observed when simulating the new methodology in comparison with the current histogram approach.

Table 1 below provides definitions that will be helpful in demonstrating the CAISO's proposal. All of the quantities are forecasted MW.

Table 1: Imbalance reserve requirement definitions

Term	Definition
Load Imbalance (L)	RTD Load – RTPD Load
Wind Imbalance (W)	RTD Wind – RTPD Wind
Solar Imbalance (S)	RTD Solar – RTPD Solar
Net Load Imbalance (NL)	L – W – S
Histogram (H)	The Histogram approach to estimate the requirement
Quantile (Q)	The Quantile approach to estimate the requirement
Sqr	The quadratic input variable, e.g. L_sqr = Load * Load

The definitions in Table 1 will be used throughout this appendix. For example, NL_H is the current CAISO requirement determined by the histogram approach, S_Q is the requirement determined by the quantile regression for the solar component, and other definitions such as S_H, L_Q, L_H, W_Q, and W_H will be used in the proposal of the regression formula below.

The advantage of applying the quantile regression approach will be illustrated graphically by using the solar component as an example. The CAISO has to run a simulation study using 2019 load, wind, and solar forecast data in RTPD and RTD. The solar component data is displayed in Figure 1 . The blue dots represent the RTD solar imbalance, the X-axis is RTPD solar forecast, the red line (S_H) is the requirement estimated by the histogram approach, and the green line (S_Q) maps the requirement based on the quantile regression approach. S_H is a straight line because the histogram approach does not utilize future forecast information. On the other hand, the curvature in S_Q demonstrates its ability to shape the requirement more effectively when input variables, such as RTPD solar forecasts here, have certain association with the RTD imbalances. Furthermore, it can be seen in this example that the requirement for solar is higher during the middle spectrum of the RTPD solar forecast, while less requirement is needed at the two ends of the RTPD solar forecasts. This follows, that the CAISO has observed higher uncertainty and forecast changes when solar is experiencing partly cloudy conditions. When solar is forecasted to have little generation or full generation, the uncertainty is less, as a result driving the requirement down during those forecasted conditions.

Similarly, the simple quantile regression with the RTPD forecast plus its quadratic term can enhance the estimation of the requirement for load and wind components, respectively.

The proposed quantile regression will use historical data from some previous configurable time period. The actual number of days used needs to provide balance between the sample size (i.e., number of data entries) and the staleness of the data used to set the requirement. For instance, the sun angles for solar resources change by day, wind and load also have patterns varying over a year.

When both histogram and quantile regression are applied based on monthly data, it will further improve the estimate of the requirement Figure 2 clearly shows quantile regression benefits more on this monthly stratification, since the monthly S_H are clustered together and S_Q has more curvatures varying from month to month.

Figure 1: Flexible Ramping Requirements by Histogram (H) and Quantile Regression (Q)

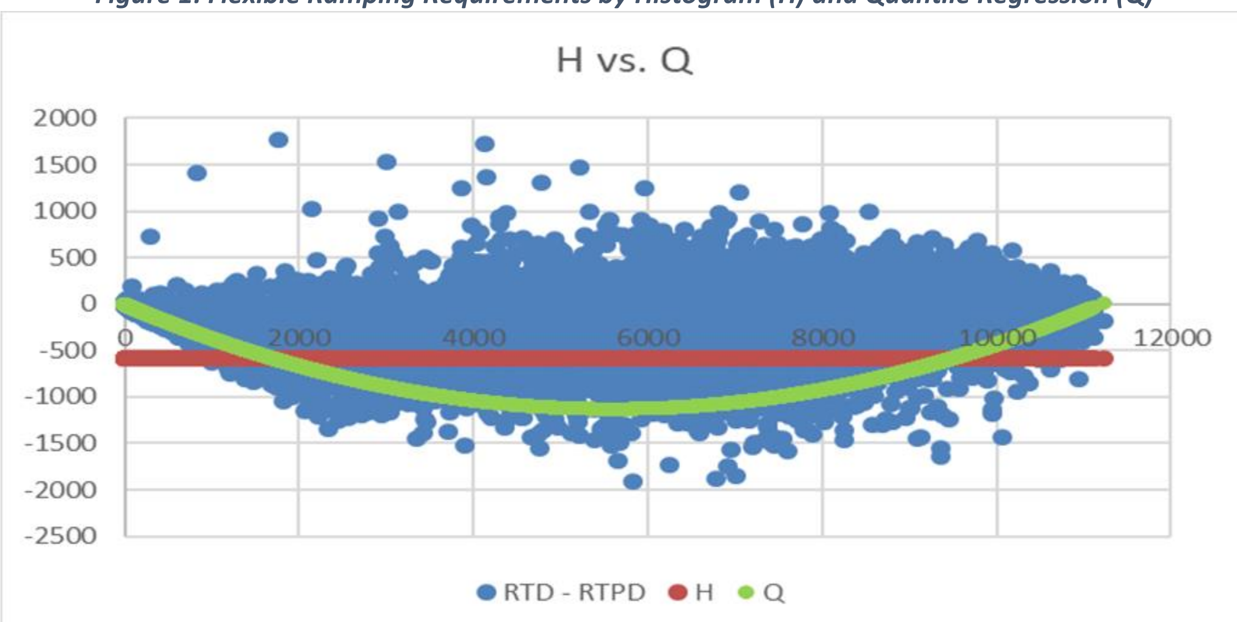
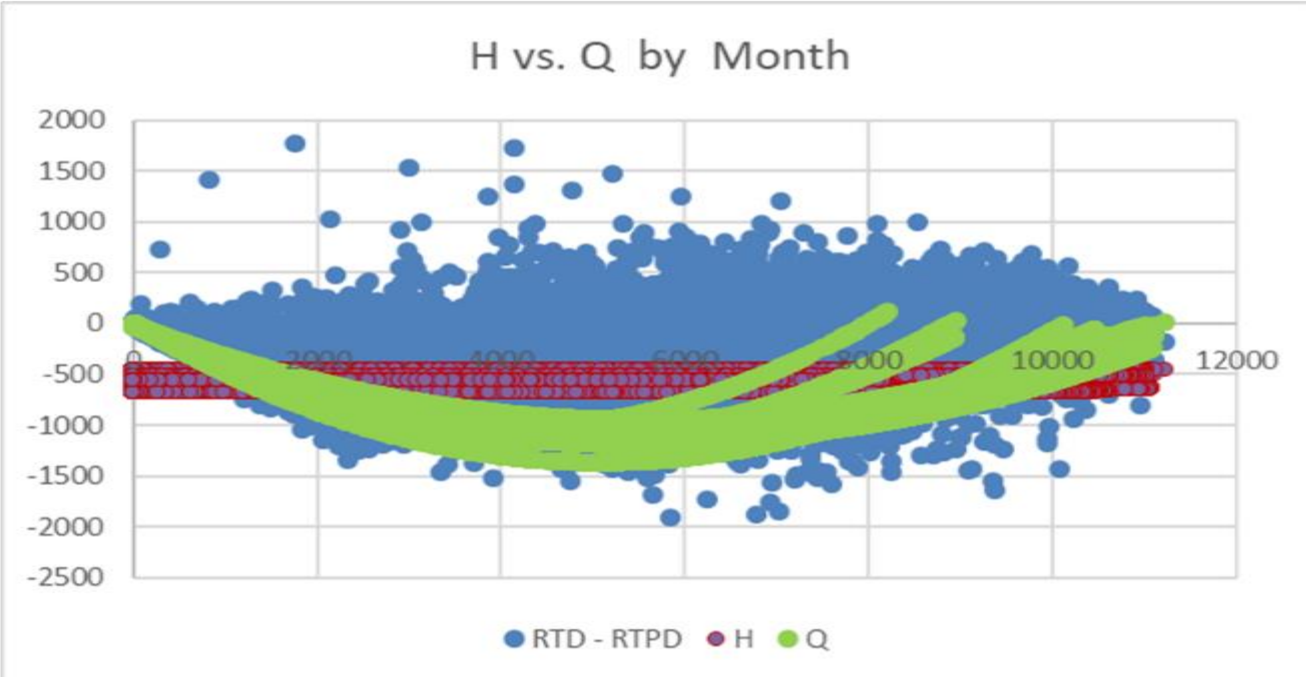


Figure 2: Monthly Flexible Ramping Requirements by Histogram (H) and Quantile Regression (Q)



The models of the component-wise quantile regression used in the CAISO proposal are listed as follows:

S_Q = RTPD_Solar + RTPD_Solar_sqr

W_Q = RTPD_Wind + RTPD_Wind_sqr

L_Q = RTPD_Load + RTPD_Load_sqr

The CAISO has noticed that S_Q, W_Q, and L_Q are better estimates of the requirements for each component itself than their counterparts S_H, W_H, and L_H, respectively. The CAISO proposes to use a blend of the estimators above to create a new input variable, MOSAIC, as follows,

MOSAIC = $NL_H - (L_H - W_H - S_H) + \left(L_Q - W_Q - S_Q\right)$

MOSAIC will bridge the load, wind, and solar requirements of each component to the net load requirement, and $NL_H - (L_H - W_H - S_H)$ will help to adjust the naïve estimate $\left(L_Q - W_Q - S_Q\right)$. The output of the quantile regression (NL_Q) on MOSAIC will be bounded by two configurable parameters; γ_1 and γ_2. The bounded $NL_Q$is the new enhancement of the flexible ramping requirement. The bounds ensure the stability of market operation and avoid over procurement.

Figures 3-6 below show the enhancement gained by the MOSAIC quantile regression model in the CAISO's simulation study. MOSAIC quantile regression has a strong feature over the histogram approach, as well as other model options, i.e., MOSAIC quantile regression offers a good fit not only on net load forecast, but also on the forecast of its component load, wind, and solar.

Figure 3 and Figure 4 show the output NL_H and NL_Q on the net load forecasts for the month July in 2019. The difference between the flat red dots in Figure 3 and the curved green dots in Figure 4 shows that Q has provided a better fit for the net load forecast.

The solar component is selected to show the fit by NL_H and NL_Q. It can be seen in Figure 5 and Figure 6 that the quantile regression still preserves the curvature exhibited in Figure 2. The similar observation also holds true for the load and the wind components, respectively.

Figure 3: Net Load Requirement by H along the Net Load RTPD Forecast

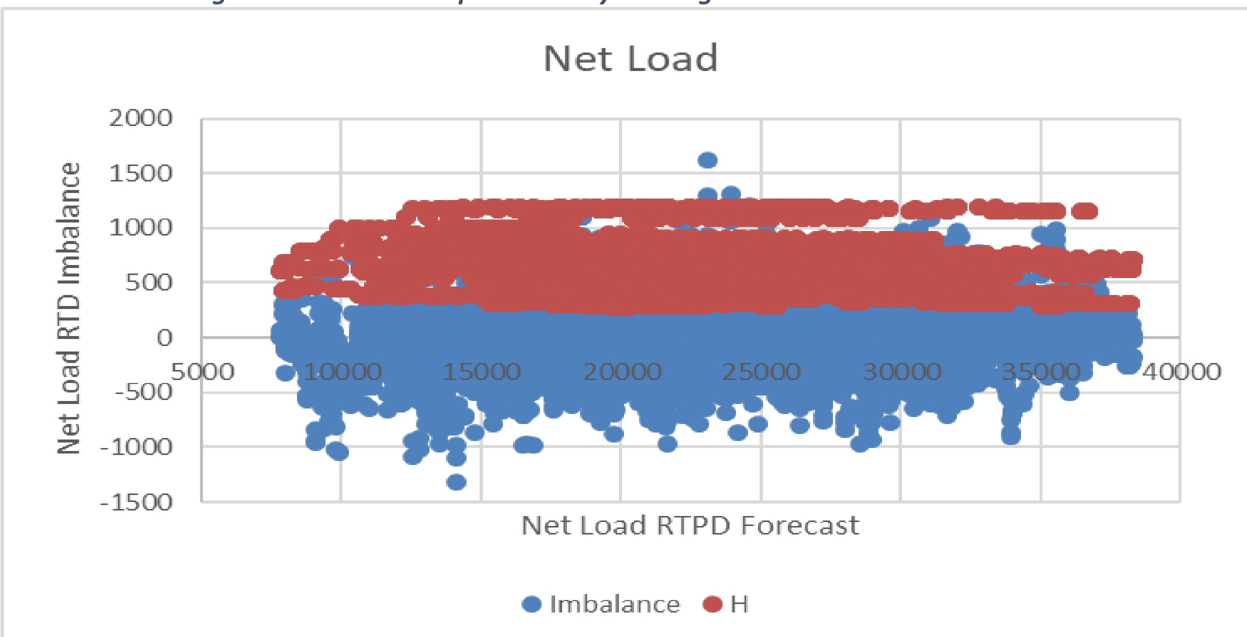


Figure 4: Net Load Requirement by Q along the Net Load RTPD Forecast

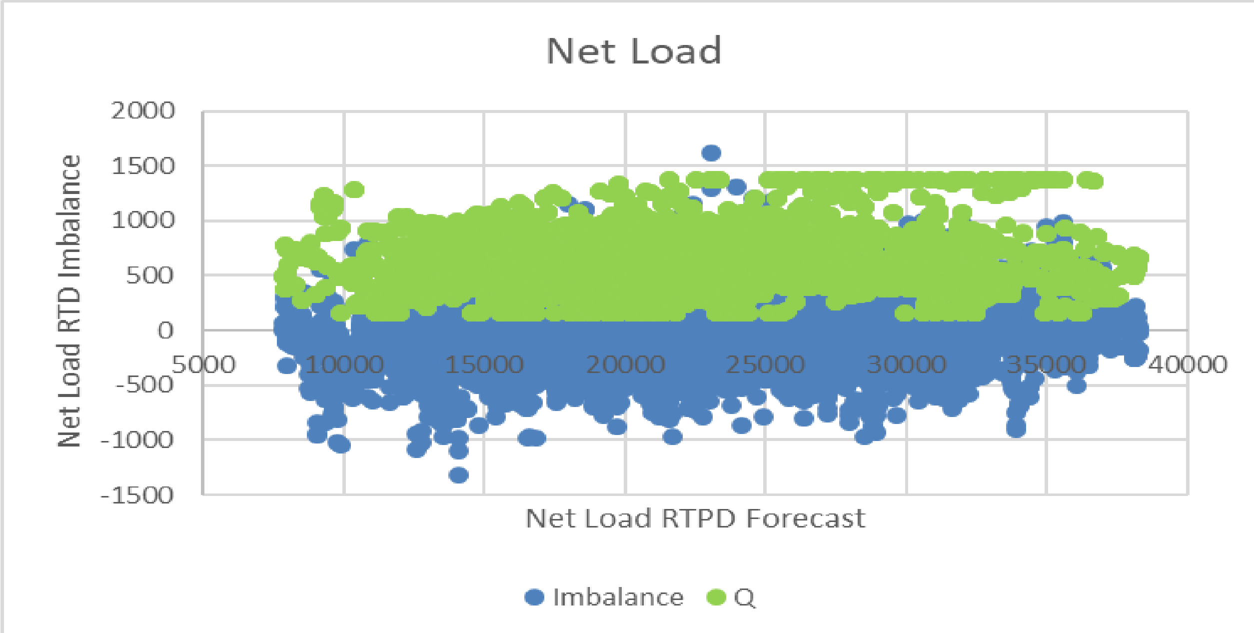


Figure 5: Net Load Requirement by H along the RTPD Solar Forecast

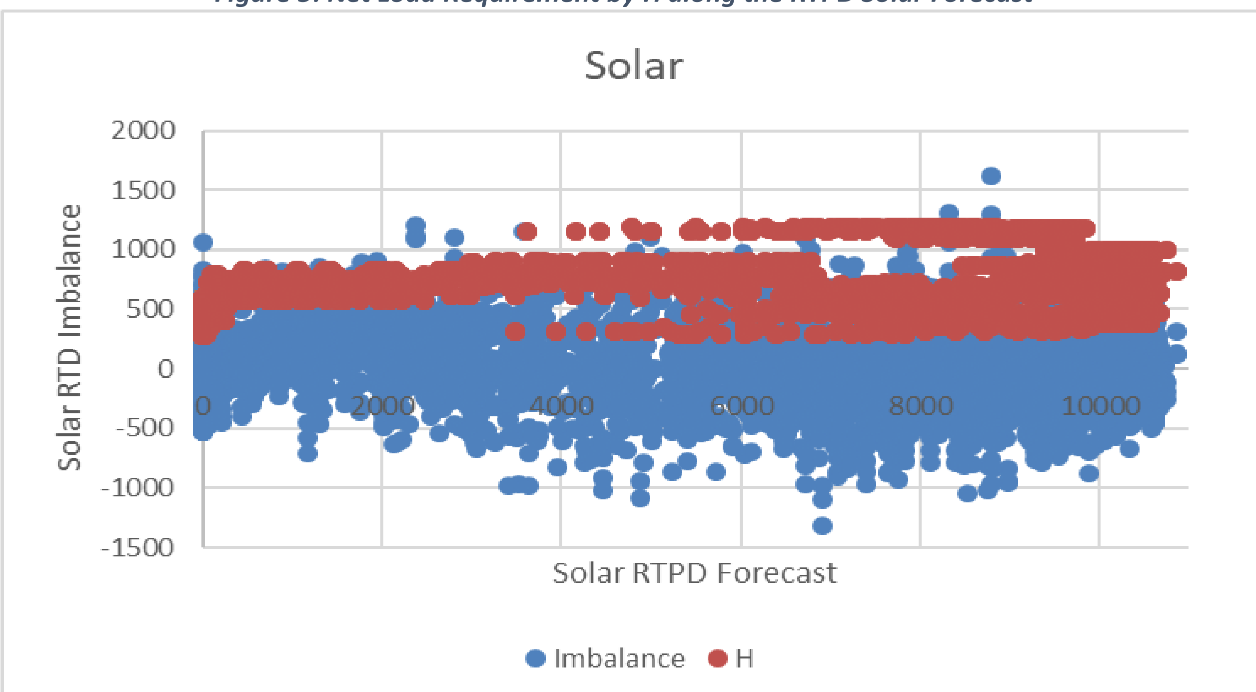
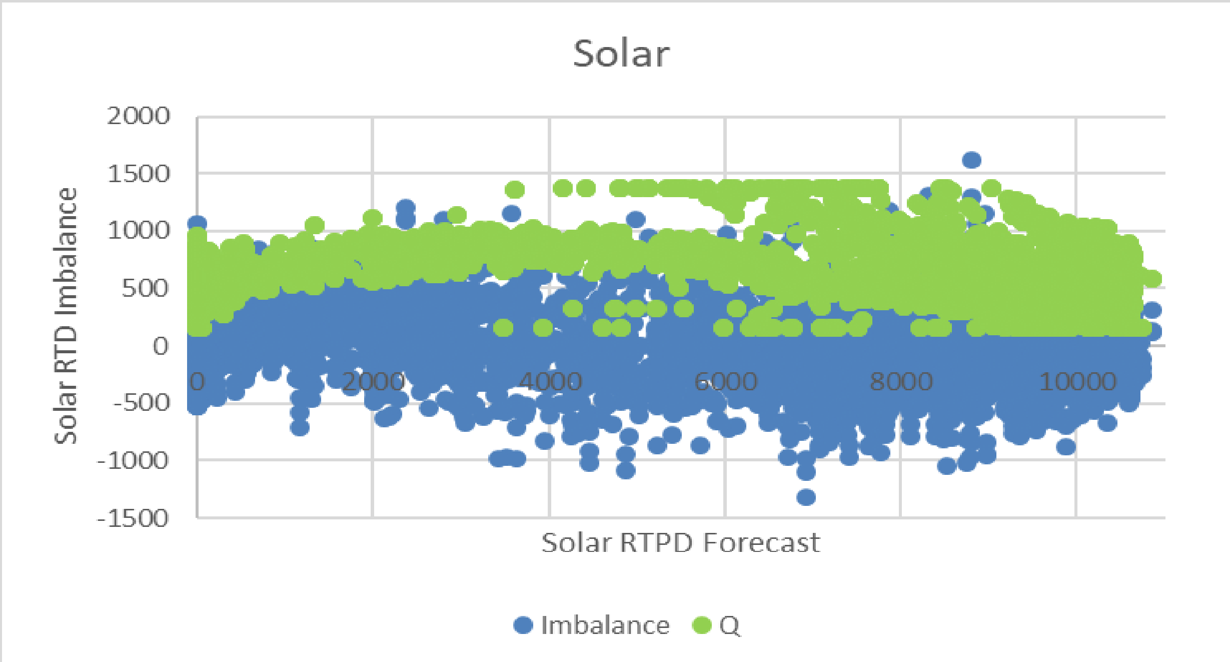


Figure 6: Net Load Requirement by Q along the RTPD Solar Forecast



Performance Measurements and Simulation Results

Besides the above graphical representations that compare the histogram and quantile regression approaches, the CAISO has also designed a matrix with four measurements to evaluate the performance of these two approaches to calculate the flexible ramping requirements. The four measurements are as follows:

1. Coverage: The percentage of the observed imbalance exceeding the requirement. This measurement is used to see how much deviation there is from the nominal level, say, 97.5.
2. Requirement: The average amount of the calculated requirement.
3. Closeness: The average of the distance between the observed and the requirement.
4. Exceeding: The average of the amount when the observed imbalance is exceeding the requirement.

In the CAISO's simulation study for the period of January 1, 2019 through December 31, 2019, six Energy Imbalance Market balancing authority areas (EIM BAAs) were included. Each hour in a day has 40 previous days of data to run the analysis. Table 2 summarizes the performance measurements for the histogram (H) as well as the quantile (Q) approaches.

Table 2: Comparing Performances of Histogram (H) and Quantile Regression (Q) approaches

	Coverage		Requirement		Closeness		Exceeding	
BAA	H	Q	H	Q	H	Q	H	Q
AZPS	96.87%	96.17%	122.72	117.17	144.24	139.08	49.56	45.65
CISO	96.71%	96.10%	602.85	547.13	595.46	540.99	175.07	163.74
IPCO	97.16%	96.80%	66.02	61.58	67.61	63.08	24.84	20.75
NEVP	97.00%	96.08%	70.63	62.02	78.05	69.79	29.10	26.77
PACE	96.99%	96.57%	108.79	107.11	110.65	109.08	36.86	33.97
PACW	97.19%	96.86%	59.33	53.81	58.40	52.70	23.51	18.35

The CAISO has tested other quantile models, including the following:

1. Forecasted Net Load plus its quadratic term,
2. Lump together all the components, Load, Wind, and Solar, plus the counterparts of their quadratic terms.

The MOSAIC blending model is selected based on its performance. Moreover, the CAISO also believes that MOSAIC blending is a tool that can be used in the future for further improvement when Numerical Weather Prediction Ensembles are available for one or all the components of load, wind, and solar.

The CAISO finds that the MOSAIC quantile regression can enhance the overall performance of the flexible ramping product requirements. Table 2 shows the evidences that all EIM BAAs have different degrees of enhancement by the MOSAIC quantile regression approach.

In summary, the MOSAIC quantile regression can achieve three goals: it will reduce average requirement and at the same, more importantly, it provides the more requirement when it is needed. Lastly, the less measurement of the exceeding means the MOSAIC quantile regression offers more accurate estimate of the flexible ramping requirement.