California Independent System Operator Corporation

California ISO

Review of mosaic quantile regression for estimating net load uncertainty

November 20, 2023

Department of Market Monitoring

TABLE OF CONTENTS

1 Executive Summary

In February, the California ISO implemented a new calculation of real-time net load uncertainty used in the flexible ramping product demand and in the Western Energy Imbalance Market resource sufficiency evaluation. This new calculation, the *mosaic quantile regression*, uses historic data to estimate the relationship between extreme outcomes of net load uncertainty, and load, solar, and wind forecasts.

This new method begins with the understanding that while future uncertainty is unknown, there is available data on load, solar, and wind forecasts that can be used to forecast this uncertainty. Quantile regression analysis is used to quantify the historical relationship between these elements and the extreme outcomes (or 95th percentile) of uncertainty. This quantified relationship is then combined with current load, solar, and wind forecasts to estimate the upper and lower bounds of future uncertainty, with a 95 percent chance of encompassing the actual outcomes.

The quality of this regression method hinges on two key points. First, are load, solar, and wind forecasts significantly related to extreme outcomes of net load uncertainty? Second, if so, is quantile regression adept at quantifying this relationship? DMM replicated the ISO methodology and conducted statistical analysis to evaluate these two key points. Key findings of this analysis include the following:

- **The mosaic quantile regression model has limited predictive capability** for forecasting net load uncertainty. DMM conducted statistical tests to determine whether coefficients of the series of quantile regressions were significantly different from zero. The coefficients represent the measured relationship between variables at extreme percentiles. If these coefficients are statistically indistinguishable from zero, it suggests a weak or inconsistent relationship. Tests conducted based on DMM's replication of the model found only 35 percent of the coefficients were statistically different from zero for the pass-group 15-minute product between February and September.
- **The results from the mosaic quantile regression closely resemble the histogram model,** highlighting the weak relationship between net load uncertainty and the mosaic variable, the selected predictor for net load uncertainty. When this variables have no meaningful relationship, regression forecasts tend to converge toward basic historical percentiles of uncertainty (histogram method). DMM's test, involving a weakly related variable, validate this convergence.
- **The quantile regression relies on a limited sample size, which may compromise the validity of the regression model.** The ISO sampling method for 15-minute uncertainty forecasts results in about 205 observations for weekend forecasts and 514 for weekday forecasts. Quantile regression assigns varying weight to a subset of these observations based on the target percentiles in the model. The current mosaic model targets the 2.5th percentile, giving substantial weight to a limited number of extreme observations in estimation. This results in an effective sample size in about 5 observations for weekends and 13 observations for weekdays.
- **The performance of the quantile regression and histogram methods is quite similar, with neither method showing clear dominance in accuracy or efficiency.** The histogram method is slightly more reliable, meaning its estimated boundaries are more likely to capture the actual uncertainty. In contrast, the mosaic quantile regression is slightly more efficient, resulting in lower estimated requirements than those of the histogram. However, the disparity in reliability and efficiency between the two models is minimal.
- **The mosaic quantile regression model has a specification error involving the use of unintended information in forecasting, along with unnecessary steps that do not influence the outcome.** Further details can be found in the Section 6.

Based on this analysis and prior comments, DMM provides the following recommendations for consideration by the ISO and stakeholders:

- **DMM recommends that the ISO develop separate models of net load uncertainty for each of the five current and potential future market features in which the ISO incorporates uncertainty.** These include: 1) the flexible ramping product, 2) the resource sufficiency evaluation, 3) the extended flexible ramping product, 4) imbalance reserve up/down, and 5) the uncertainty captured in operator adjustment to the residual unit commitment demand. Each of these market features involve different applications, forecasting horizons and timelines of net load uncertainty.

- **The infrastructure developed to support the mosaic quantile regression can be adapted to support other model formulations.** A thorough review of other options would include an evaluation of different sampling methods, independent variable selection and functional form. DMM has begun evaluating potential refinements including: 1) pooling estimation across hours, 2) using a more conventional forecasting approach than the quantile regression, and 3) replacing the multi-stage mosaic variable regression with a regression on net load uncertainty directly. These results will be published in a future report.

- **DMM recommends that the ISO consider reverting to the histogram approach for the current flexible ramping product and the resource sufficiency evaluation** while an improved model or approach is developed. This would provide requirements that capture a high level of uncertainty equivalent to that of the quantile regression model, but in a much more predictable way for current market participants. This could also reduce the additional processing time that the quantile regression approach adds to the real-time market.

- **If the ISO retains the existing mosaic quantile regression, the limited sample size and technical errors in the formulation should be resolved.** DMM suggests that the histogram component of net load, load, solar, and wind could be removed from the construction of the mosaic variable. Model results remained the same without these components. To address the misspecification issue in the initial quantile regression, the dependent variables could remain load, solar, and wind uncertainty, but the independent variable should be adjusted by using one-interval or one-hour lagged 15-minute forecasts. This modification prevents the overlap of the same 15-minute forecast in both the right- and left-hand sides of the regression equation. DMM recommends extending the sample period beyond 180 days, but with appropriate seasonal controls for load forecasts. Unlike renewable forecasts, which are currently adjusted based on installed capacity, load forecasts lack an equivalent method. DMM observed that during the summer months, the current load forecast stands out as an outlier compared to the load forecasts in the sample, resulting in extremely large forecasts.

- **There is value in having a good approach for estimating the different probabilities of different net load realizations for market products that are procured based on a demand curve.** These three products include: 1) the current flexible ramping product, 2) the proposed day-ahead imbalance reserve up, and 3) the extended flexible ramping products. The demand curve for each of these products is explicitly based on the probabilities of different net load realizations, combined with the estimated costs of different levels of procurement.

- **DMM recommends that a much more simplified approach be considered for incorporating uncertainty into resource sufficiency evaluations**. This includes the current WEIM resource sufficiency test, as well as the resource sufficiency tests for the Extended Day-Ahead Market (EDAM).

These requirements used in these tests are not based on a demand curve or any other specific reliability standard. Instead, the uncertainty component of these tests is more akin to a reserve capacity margin that is agreed upon by all balancing areas. Thus, DMM has recommended that a different – and much more simplified – approach be considered for setting the uncertainty component of these tests.

- **For resource sufficiency evaluations, DMM suggests that requirements could be based on a very simplified approach based on the amount of load, wind, and solar each hour within each balancing area.** This type of simplified approach would allow balancing areas to know what their resource sufficiency evaluation requirements would be, and plan accordingly to meet these requirements. This would avoid the problems created by the very high variability and uncertainty about resource sufficiency evaluation requirements that participants face under the current quantile regression approach.

2 Background

Net load uncertainty refers to variation between an actual and expected forecast of load net variable solar and wind generation. Estimates of net load uncertainty are added to requirements (or a demand curve) for market products designed to reserve capacity to meet net load uncertainty. Reasonable forecasts provide the market with sufficient reserves and provide market participants with confidence that the capacity being reserved is appropriate.

The current flexible ramping products procure ramping capacity based on very short-term net load uncertainty (i.e. uncertainty that might materialize one 15-minute interval in the future for the 15-minute product or that might materialize one 5-minute interval in the future for the 5-minute product.). The 15-minute estimate of net load uncertainty is also used in the Western Energy Imbalance Market resource sufficiency evaluation.

New uncertainty products are being developed to capture uncertainty between the day-ahead and the 15-minute market.[1] An extended flexible ramping product is also being developed to capture real-time uncertainty two to four hours in advance, as part of the price formation enhancement initiative. DMM supports the development of these new extended uncertainty products. However, these new products will make forecasting uncertainty significantly more important than it is in the current market design and may require different estimation techniques.

The quantile regression approach was developed as an enhancement, intended to improve estimates of net load uncertainty used in these market products. The performance of the mosaic quantile regression model being used was tested against other models before deployment. However, the data used to develop the quantile regression model was not representative of the uncertainty in the market at the time or since deployment.[2] Regression diagnostics are not provided to the ISO by the vendor that developed the regression model used in the market software.

DMM replicated the model as closely as possible based on written technical requirements. As explained in detail below, coefficients are not statistically significant. Without significant coefficients, the choice of optimization may create very different results. On the basis of these replicated results and observed market outcomes, DMM recommends that the ISO reevaluate the use of the mosaic quantile regression for both existing and future market products.

[1] The imbalance reserve product is being developed as part of the Day-Ahead Market Enhancements.

[2] The data used in the model selection process was constructed incorrectly, defining the uncertainty by subtracting the wrong interval's advisory value. Any conclusions from this review should be revisited with data from the market, which does not have this issue. California ISO, *Flexible Ramping Uncertainty Calculation in the Western Energy Imbalance Market (WEIM)*, March 25, 2022: http://www.caiso.com/InitiativeDocuments/Analysis-FlexibleRampingUncertaintyCalculationintheWesternEnergyImbalanceMarket.pdf

3 Forecasting net load uncertainty: histogram and quantile regression

This section presents the ISO methodologies for forecasting net load uncertainty, highlighting two primary techniques: the histogram and the mosaic quantile regression. The empirical assessment of the performance of both techniques begins in Section 4.

Histogram and regression are statistical techniques that determine the mathematical strength and nature of the relationship between one variable, and one or more variables. Forecasting is the process of predicting future values based on past and present data, often using the relationships identified through regression.

Regardless of the complexity of a regression model, its purpose is to estimate the conditional average (or quantile value) of the net load uncertainty. This conditional average refers to the expected value of net load uncertainty, given certain values or conditions of the independent variables. On the other hand, the histogram method calculates the unconditional or regular average (quantile value) of net load uncertainty irrespective of any independent variables.

The primary difference between an unconditional and a conditional average (or quantile) is that the former does not take into consideration any additional information, while the latter does.[3] If an unconditional and a conditional average (quantile) are similar, this means that the conditions being considered do not have an impact on the overall average (quantile). In other words, the condition being applied is not associated with net load uncertainty.

The ISO utilizes a mosaic variable to predict net load uncertainty. Essentially, the mosaic quantile regression estimates the conditional quantile of net load uncertainty based on the value of the mosaic variable, as reflected in historical data.[4]

The ISO leverages this historical relationship between uncertainty and the mosaic variable to forecast unknown net load uncertainty. For example, with a relationship of 0.5, the ISO can compute future net load uncertainty at the upper or lower quantile. This is achievable because the mosaic variable is available for future periods. Therefore, if the relationship is 0.5, the conditional quantile of uncertainty is calculated as the future mosaic variable multiplied by 0.5. The underlying assumption is that the historical relationship between the mosaic variable and net load uncertainty will persist in the future period for which the forecast is being made.

[3] The histogram and regression methods answer different questions when processing data. The histogram method produces the response to: “What is the average net load uncertainty over the last 180 days?” The regression method answers: “What is the average net load uncertainty when renewable forecasts are high?” This principle holds true for quantile regression. In this case, the quantile regression identifies the conditional quantile values (0.975 and 0.025) of net load uncertainty, rather than the average.

For example, consider a scenario in which there are only two renewable levels: high and low. The histogram method calculates an overall average of net load uncertainty, without distinguishing between high and low levels of renewable energy. The regression method identifies the conditional averages for both high and low renewable scenarios. The regression coefficient indicates the difference between two conditional averages. After identifying these conditional averages, the regression then calculates a weighted average of the two. The weights assigned to each conditional average are determined by the distribution of high and low renewable scenarios. If the data splits evenly between high and low scenarios, each conditional average would be assigned a weight of 0.5.

[4] For instance, if the historical data reveals a mosaic quantile regression coefficient of 0.5, then the historical relationship is 0.5.

3.1 Net load uncertainty

Net load uncertainty refers to the unpredictable nature of net load (load minus utility scale, wind, and solar forecasts) in the real-time market. The real-time market uses two net load uncertainty definitions. For the 5-minute flexible ramping product, uncertainty is the difference between net load in the binding 5-minute interval and in the corresponding advisory 5-minute interval.

For both the 15-minute flexible ramping product and the WEIM resource sufficiency evaluation, uncertainty is the difference between binding 5-minute market forecasts and corresponding advisory 15-minute market forecasts.

3.2 Histogram method

Before February 1, 2023, the ISO calculated uncertainty by selecting the 2.5th and 97.5th percentile of observations from a distribution of historical net load errors. This is known as the *histogram method*. For the 15-minute market product and the resource sufficiency evaluation, the historical net load error observations in the distribution are defined as the difference between the binding 5-minute market net load forecasts and corresponding advisory 15-minute market net load forecasts.[5]

Prior to February 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.

3.3 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.

This new regression method is a two-step procedure to forecast the lower and upper extremes of net load uncertainty that might materialize. The first step in this procedure is to create the mosaic variable, the primary predictor for net load uncertainty. This process itself requires multiple regressions, and the results of these regressions serve as inputs for the mosaic variable. These 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).

---

[5] In comparing the 15-minute observation to the three corresponding 5-minute observations for the 15-minute market product, the minimum and maximum net load errors were used as a separate observation in the distribution. The 5-minute market product instead used the difference between a binding 5-minute market net load forecast and advisory 5-minute market net load forecast.

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 in Equation 1 below for 15-minute market upward net load uncertainty.[6]

Equation 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: 97.5 percentile, low: 2.5 percentile)

The uncertainty regressions use a distribution of historical forecast observations from the previous 180 days — separate for each balancing area, hour, and day-type (weekday or weekend/holiday). For the 15-minute product and resource sufficiency evaluation, uncertainty in the distributions is the difference between binding 5-minute market forecasts and corresponding advisory 15-minute market forecasts.[7]

The outcome of these regressions are the coefficients *a*, *b*, and *c*, that define the relationships between these forecasts and the extreme percentile of uncertainty that might materialize.[8] These coefficients can then be combined with the historical 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 in Equation 2 below for upward net load uncertainty.

[6] Equations 1 to 6 are for calculating *15-minute market 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. *5-minute market* net load uncertainty is instead based on the difference between a binding and advisory 5-minute market net load forecast.

[7] 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.

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

Equation 2 Predicted values for upward net load uncertainty

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

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

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

After the initial quantile regression generates predicted values, the subsequent step involves constructing the *mosaic variable*. 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. Here, the calculation modifies the histogram net load by adding the predicted values and subtracting the histogram outcomes for each uncertainty type individually.[9] This is shown in Equation 3 below for upward net load uncertainty:

Equation 3 Mosaic variable for upward net load uncertainty

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

Upward mosaic variable: intermediate variable for final regression

97.5th percentile of net load uncertainty from histogram

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

Load, solar, and wind uncertainty from histograms

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

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

Equation 4 Mosaic regression for upward net load uncertainty

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

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

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

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

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

After completing the mosaic quantile regression in Equation 4, the output coefficients represent the numerical relationship between the extreme end of net load uncertainty and mosaic variables over the last 180 days.

The final question is: what will the net load uncertainty be in the future? Equation 4 displays the relationship between net load uncertainty and every possible value of the mosaic variable. When the mosaic variable is set to 0, inserting this value into the equation yields the upper boundary of net load uncertainty. When the current forecast (the target forecasting interval) presents a mosaic variable value of 100, entering 100 into the mosaic variable portion of the equation determines the upper boundary.[10]

The final equation for combining the current forecast information with the regression coefficients and histogram extremes to calculate upward uncertainty for each interval are shown in Equation 5.

Equation 5 Calculation of upward uncertainty from current forecast information

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

Equation 5 lays out the relationship between known variables and those to be determined by forecasting. The right-hand side highlights the current upper boundary of uncertainty, which is not known but can be estimated using the equation on the left-hand side. On the left, the coefficients, which are already known, are derived from the mosaic quantile regression. The current mosaic variable value is also known. It is constructed based on the current load, solar, and wind forecasts, combined with coefficients from the initial quantile regression.

[10] The 15-minute market product and resource sufficiency evaluation use the same regression coefficients but are combined with slightly different forecast information from each time horizon. The flexibility test uses the same load, solar, and wind forecasts, which are considered in the resource sufficiency evaluation for calculating ramping capacity and test requirements. The latest forecasts at the time of the second pass of the resource sufficiency evaluation at 55 minutes prior to the evaluation hour, are held constant for the final test at 40 minutes prior to the hour. The 15-minute and 5-minute market uncertainty calculations for the flexible ramping product instead use the advisory forecasts in the next interval — the same interval in which the deployment scenarios are run to determine feasible flexible capacity awards.

Equation 6 Calculation of the current mosaic variable

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

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

Equation 6 sets forth the method to determine the current mosaic value. Begin by collecting the latest forecasts for load, solar, and wind. Subsequently, using the coefficients from the initial quantile regression generates the predicted values. By integrating these predictions with the histogram values, the current mosaic values are obtained.

It is important to note that this process does not necessitate any additional regressions. It merely involves merging the previous regression coefficients, which are individual figures, with the current forecast values.

In summary, regression captures the past 180-day relationship between uncertainty and the mosaic variable. When forecasting, this quantified relationship is used to predict future upper uncertainties. So, if today's mosaic value is 100, forecasting would identify the 97.5th percentile from previous days that also had a mosaic value of 100.

Equation 7 below represents the simplified steps of mosaic quantile regression. First, three initial quantile regressions are performed to create the mosaic variable. Second, the mosaic quantile regression determines the historical correlation between the extreme levels of uncertainty and the mosaic variable. Finally, uncertainty is calculated by blending this historical relationship with the current mosaic variable.

Equation 7 Simplified steps of mosaic quantile regression

1. Initial quantile regression to estimate γ (assume estimated $\gamma = 0.1$)

   $X_i = Z_i\gamma + e_i$

   Find predicted values $\hat{X}_i = Z_i \cdot 0.1$

2. Mosaic quantile regression to estimate β

   $Y_i = \hat{X}_i\beta + \epsilon_i$

3. Forecast $Y_{current}$ with current $\hat{X}_{current}$ (assume estimated $\beta = 0.5$)

   $Y_{current} = \hat{X}_{current} \cdot 0.5$

4 Comparing performance: mosaic quantile regression vs. histogram

This section evaluates the performance of two forecasting models – the histogram and the mosaic quantile regression. The analysis uses data from February and September 2023, focusing on the flexible ramping product uncertainty for the group of balancing areas that passed the resource sufficiency evaluation. The analysis reveals a significant similarity in performance between the two models.

In this and upcoming sections, DMM evaluates both the histogram and quantile regression based on the same sample size from the past 180 days. The objective is to evaluate and compare the accuracy and efficiency of the two distinct models using an identical sample size.

4.1 Out-of-sample performance of mosaic quantile regression

Overall, the out-of-sample performance of the quantile regression and histogram methods is quite similar, with neither method showing clear dominance in accuracy or efficiency. A subtle distinction lies in their strengths: the histogram method is more reliable, meaning its estimated boundaries are more likely to capture the actual uncertainty. In contrast, the mosaic quantile regression is more efficient, resulting in lower estimated requirements than those of the histogram.

The coefficients derived from the mosaic quantile regression are calculated based on the previous 180 days of data. Assessing how well these coefficients align with the same dataset is termed in-sample performance.

Out-of-sample forecasting evaluates a model's predictive accuracy using data that was not incorporated during the regression. Specifically, it assesses how accurately the forecasted flexible ramping requirement aligns with the actual net load uncertainty over a specified forecasting interval. Evaluating out-of-sample performance is the most definitive method to assess the effectiveness of a forecasting model. The mosaic quantile regression and histogram are specifically designed to target the upper and lower boundaries of actual (future) net load uncertainty, capturing the actual values within these limits 95 percent of the time.

This paper mainly evaluates the performance of regression with specified thresholds. Uncertainty calculated from the regression method are capped by three different thresholds. The thresholds are designed to help prevent outliers from influencing the final requirement and ensure that the flexible ramp up product remains positive and the flexible ramp down product stays negative.[11] These thresholds primarily influence the results of the mosaic quantile regression.

Table 4.1 presents the four main out-of-sample performance metrics for both histogram and quantile regression models. The table summarizes the flexible ramping product uncertainty for the group of balancing areas that passed the resource sufficiency evaluation (known as the pass-group) during February and September in 2023.

[11] The California ISO uses three kinds of threshold. The first is a seasonal threshold, representing the top 1 percent of uncertainty over the past 90 days, measured across all hours. The second, a 180-day threshold, is similar but spans the last 180 days, calculated hourly. The final threshold is a floor, ensuring the flexible ramp up (FRU) requirement stays above zero and flexible ramp down (FRD) below zero. If quantile regression results in a negative FRU or positive FRD value, it is adjusted to zero.

Table 4.2 provides the same performance metrics, focusing solely on the evening peak hours performance between hour-ending 17 and 21.

The hour out-of-sample performance metrics are:

1. *Coverage*: This metric evaluates whether the model's predicted upper and lower bounds include the actual uncertainty. The goal is to capture 95 percent of the actual uncertainty within these predicted boundaries.
2. *Requirement*: Represent the average estimated requirement (boundaries).
3. *Closeness*: Measures the average distance between observed uncertainty and the estimated requirement (boundaries).
4. *Exceedance*: Indicates the average closeness when the actual uncertainty falls outside the predicted boundaries.

Table 4.1 Average pass-group uncertainty coverage and requirements (all hours, including thresholds, February-September 2023)

Market	Type	Coverage		Requirement		Closeness		Exceedance	
		Mosaic	Histogram	Mosaic	Histogram	Mosaic	Histogram	Mosaic	Histogram
15-minute	Upward	96%	97%	1,367	1,510	1,337	1,464	320	310
	Downward	96%	97%	1,243	1,363	1,414	1,518	465	456
5-minute	Upward	97%	97%	258	267	279	288	77	80
	Downward	97%	97%	275	286	278	289	87	88

Table 4.2 Average pass-group uncertainty coverage and requirements (hour-ending 17-21, including thresholds, February-September 2023)

Market	Type	Coverage		Requirement		Closeness		Exceedance	
		Mosaic	Histogram	Mosaic	Histogram	Mosaic	Histogram	Mosaic	Histogram
15-minute	Upward	95%	96%	1,765	1,801	1,604	1,632	424	375
	Downward	97%	98%	1,308	1,352	1,630	1,655	382	319
5-minute	Upward	97%	97%	309	290	329	312	102	105
	Downward	96%	96%	290	278	297	285	102	101

The result shows that the histogram and mosaic quantile regression displayed comparable coverage levels, at approximately 96 percent. In both markets, the histogram slightly outperformed the quantile regression, showing a 1 percent better coverage in the upward and downward direction.

During evening peak hours 17 to 21 (Table 4.2), coverage is similar between the models in both markets, around 96 percent, with the histogram model showing a slightly better performance in the 15-minute market.

The quantile regression model shows a lower requirement in both the 5-minute and 15-minute markets. On average, it requires about 100-150 MW less in the 15-minute market and 10 MW less in the 5-minute market.

During evening peak hours, the quantile regression model continues to have a lower requirement, though the difference is smaller. In the 15-minute market, the gap in requirement between models is less than 50 MW, and in the 5-minute market, it is around 10 MW.

The quantile regression model exhibits a lower closeness, approximately 100 MW lower in the 15-minute market and about 15 MW lower in the 5-minute market. However, this gap in closeness diminishes during the evening peak hours, with around a 25 MW difference in the 15-minute market and a 15 MW difference in the 5-minute market.

The exceedance levels are similar between the two models, but the histogram method shows a slightly lower exceedance in the 15-minute market, about 10 MW less, while the quantile regression has a lower exceedance in the 5-minute market, around 2 MW less.

During evening peak hours, this disparity become more pronounced in the 15-minute market: the histogram method exhibits around 60 MW less exceedance. In the 5-minute market, the quantile regression has lower exceedance by about 2 MW.

4.2 Graphical analysis of mosaic quantile regression vs. histogram performance

Uncertainty outcomes produced by the mosaic quantile regression closely resemble the conventional histogram method. This section presents a visual representation of the coverage rates from both forecasting models for February and September 2023. DMM presents the metrics for the quantile regression both with and without the application of the thresholds.

The quantile models are not designed to forecast the actual uncertainty, but rather the upper and lower boundaries of potential uncertainty. Consequently, establishing a direct correlation between the actual and forecasted outcomes may not be straightforward. While the ideal forecast model should show a correlation where high actual uncertainty corresponds with high upper bounds, and vice versa, this may not always be the case. An upper bound forecast can display a weaker correlation with actual uncertainty because it takes into account the volatility of the actuals, making the relationship less linear.

Therefore, in this section, DMM's evaluation focused on assessing the degree of similarity or disparity between these two forecasting models. The shape of the presented scatter plots serve as an indicator of each model's characteristics, including whether they provided relatively constant forecasts or showed more variation when actual uncertainty was low or high.

Figure 4.1 and Figure 4.2 display scatter plots for the actual net load uncertainty and forecasted results generated by the histogram method, and the mosaic quantile regression, respectively, for the group of balancing areas that passed the resource sufficiency evaluation during the 8-month period. Each marker is a distinct 15-minute interval's actual uncertainty and its corresponding forecasted upper boundary of uncertainty.

The red 45-degree line demonstrates the point of parity where the forecasted upper boundary of uncertainty aligns with the actual uncertainty. Any markers situated above this line, on the top-left side, indicate instances where the forecasted upper bounds are higher than the actuals. Conversely, markers positioned below the line, on the bottom-right side, represent situations where forecasted upper bounds fall short of the actuals.

Figure 4.1 Flexible ramping up coverage of histogram (February-September 2023)

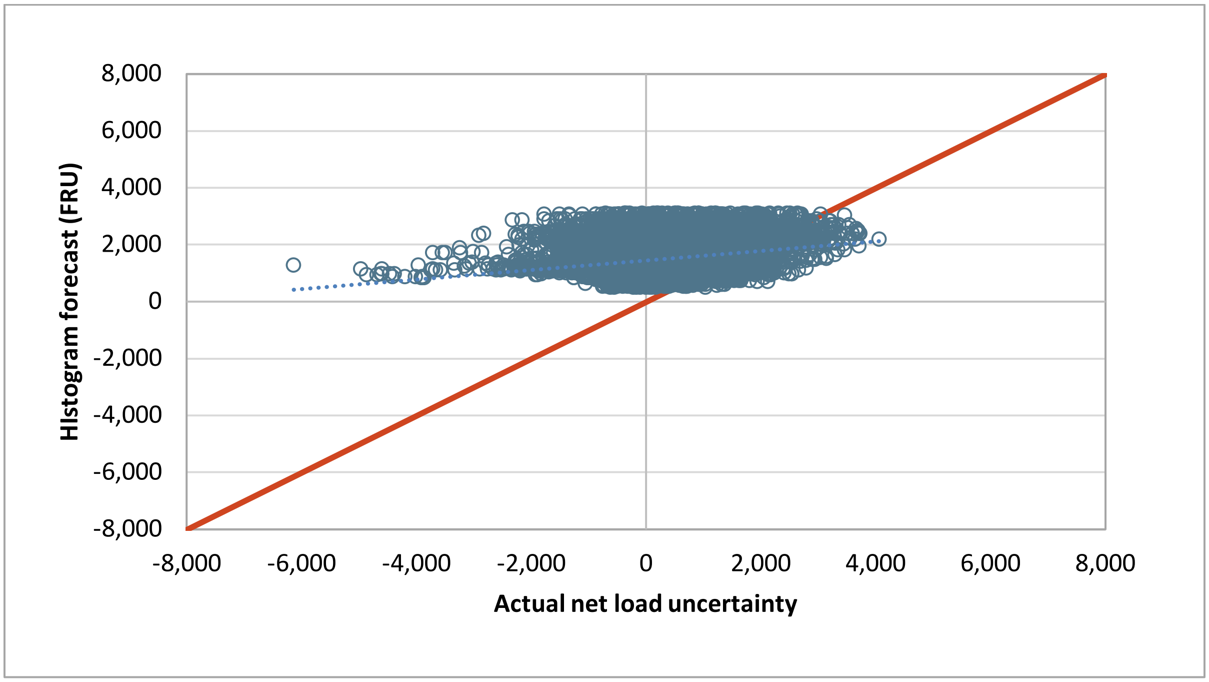


Figure 4.2 Flexible ramping up coverage of mosaic quantile regression forecast (with thresholds, February-September 2023)

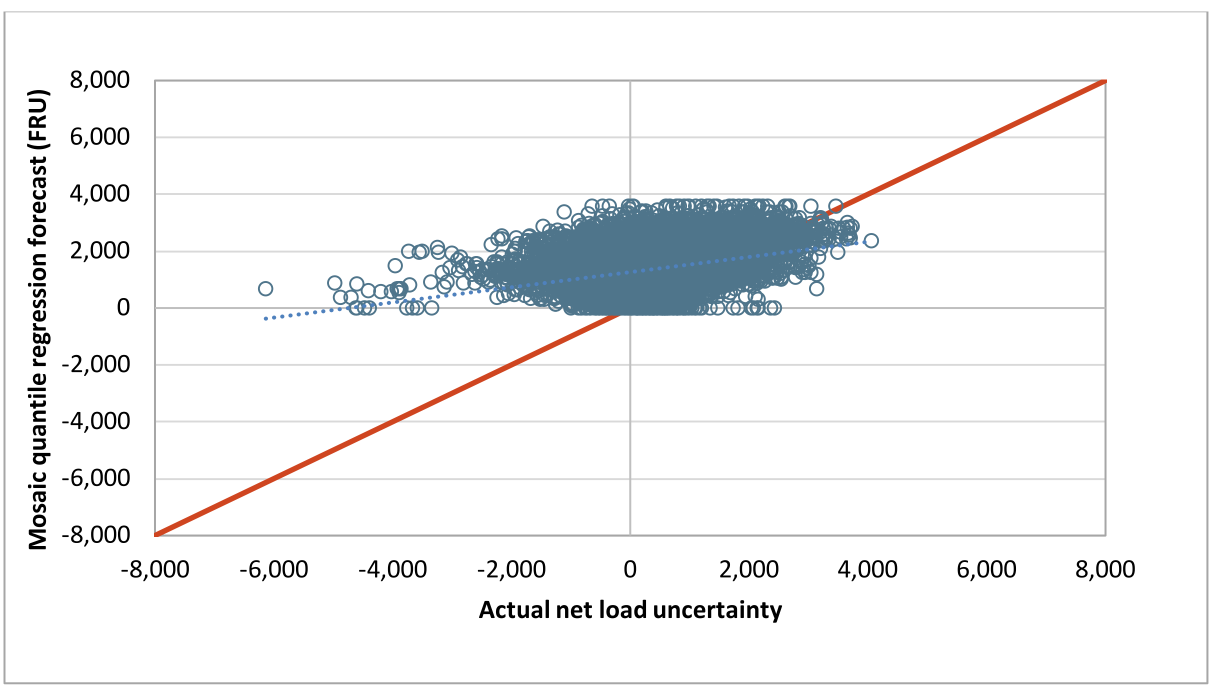


Figure 4.3 Flexible ramping up coverage of histogram and mosaic quantile regression forecast (with thresholds, February-September 2023)

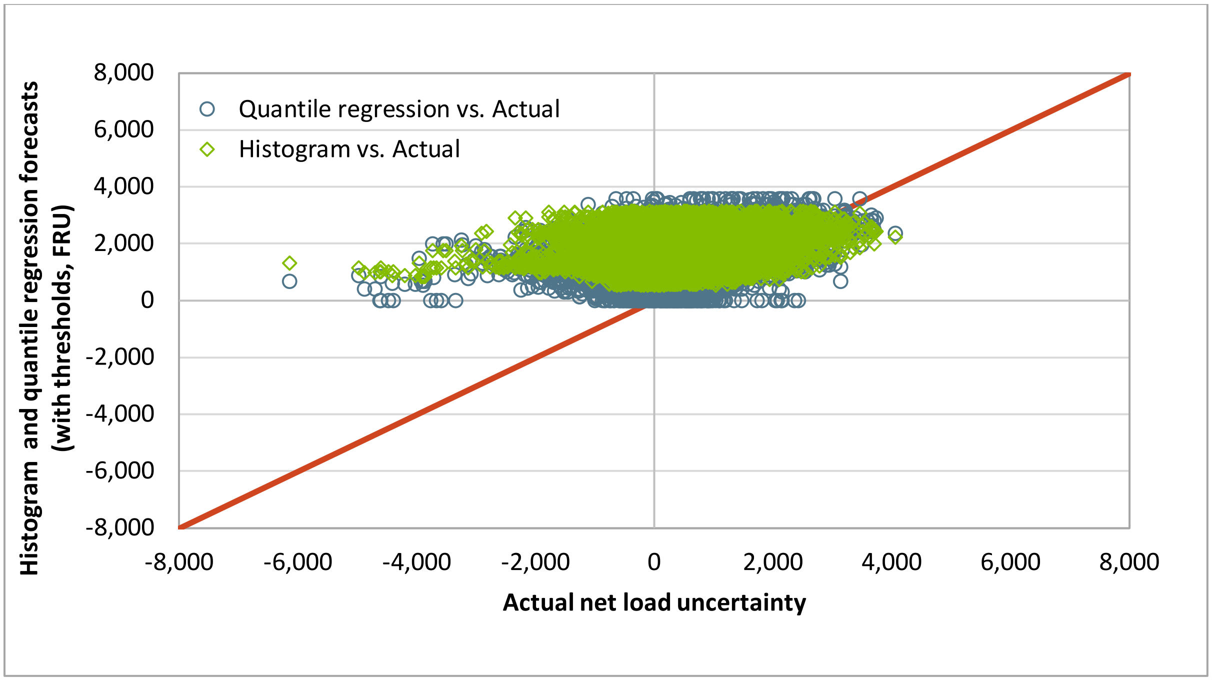


Table 4.1 in the previous section demonstrates that both the histogram and quantile regression method maintain a 96 percent coverage rate in their 15-minute upward flexible ramping (FRU) forecast. The 96 percent coverage rate is consistent with the graphical data in the figures, as most of the markers are above the 45-degree red line.

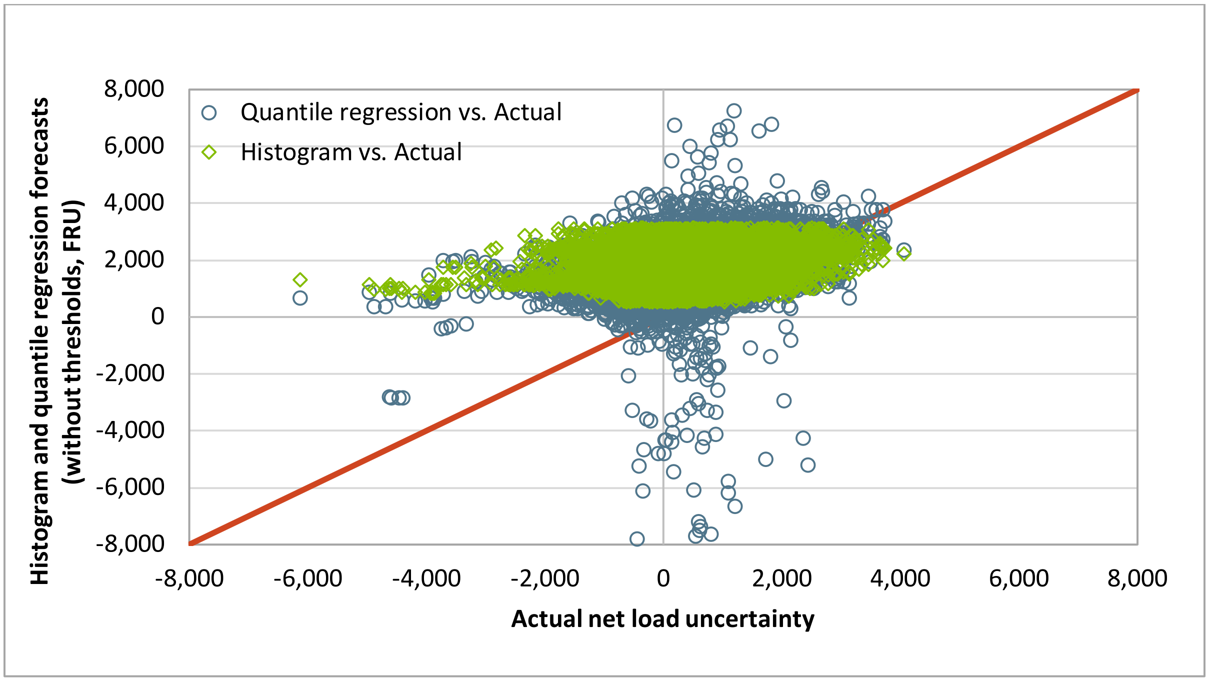
These scatter plots show the characteristics of two forecasting models. Figure 4.1 shows a scatter plot for the actual uncertainty and histogram forecast outcome. This forecast model examines historical net load uncertainty and selects the highest 97.5th percentile value. Since this highest percentile value tends not to change significantly over time, the forecasted outcome on the vertical axis remains consistent between 1,000 and 3,000. A possible interpretation of the plot shape is to view it as a horizontal line. This implies that the forecasted value, around 2,000 for instance, remains constant irrespective of the varying actual uncertainty across different intervals.

Figure 4.2 presents a scatter plot of the actual uncertainty alongside the output from the mosaic quantile regression. Upon visual examination, the two scatter plots display a high degree of similarity, exhibiting analogous characteristics. In both cases, the upper bounds remained consistent irrespective of whether the actual uncertainty was higher or lower.

Figure 4.3 displays overlapping scatter plots of the histogram and quantile regression. In this chart, blue markers represent the quantile regression forecasts and its corresponding actual uncertainty, while green markers indicate the histogram.

The resemblance between the histogram and quantile regression models might stem from the thresholds applied to the quantile regression forecasts. Figure 4.4 illustrates overlapping scatter plots of these two methods without the thresholds. This chart reveals a similar overall pattern but with a larger disparity, particularly when the actual uncertainty is around zero. Without thresholds, the quantile regression produces extreme forecasts. The chart's vertical axis is truncated to between -8,000 MW and 8,000 MW to manage the extreme values from the quantile regression. The uncapped uncertainty from the regression method have ranged from -870,000 to 400,000 MW.

Figure 4.4 Flexible ramping up coverage of histogram and mosaic quantile regression forecast (no thresholds, February-September 2023)



Additional evidence supports the degree of similarity between these two methods. Figure 4.5 and Figure 4.6 demonstrate the correlation between the outcomes of the two forecast models in upward flexible ramping (FRU) and downward flexible ramping (FRD), respectively. They both exhibit a correlation of 0.75.

Figure 4.5. Flexible ramping up correlation between histogram forecast and mosaic regression forecast (with threshold, February-September 2023)

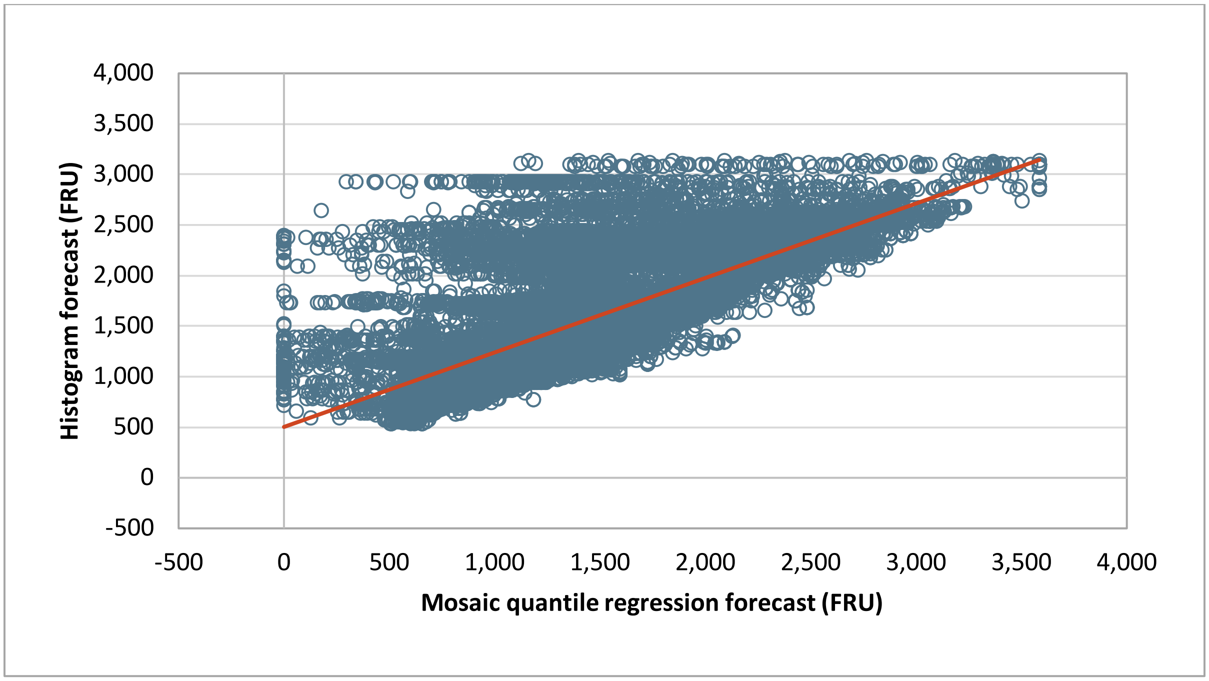
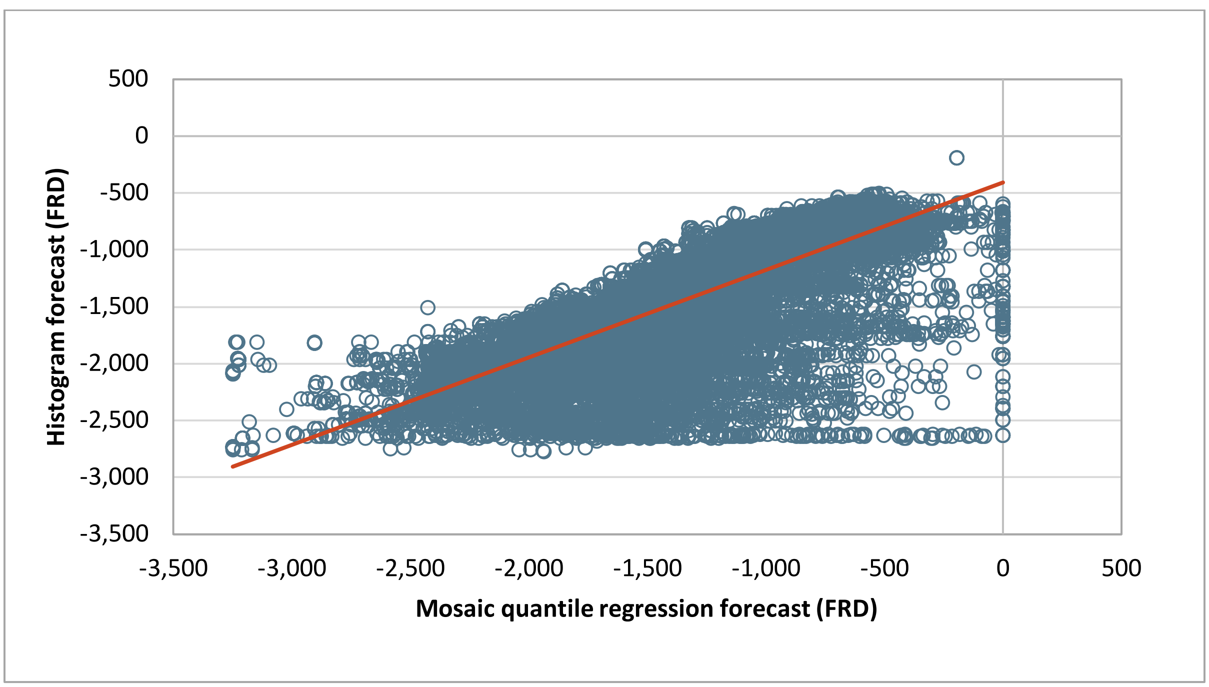


Figure 4.6. Flexible ramping down correlation between histogram forecast and mosaic regression forecast (with threshold, February-September 2023)



Similarity between the outcomes of the two models is not mere coincidence. Instead, it could signify limited forecasting power of the mosaic quantile regression. The next section explores this issue in more detail, offering additional empirical evidence.

5 Statistical examination of mosaic quantile regression

This section provides a comparative review of the new mosaic quantile regression methodology compared to the prior approach, which includes the following:

- Reviews the characteristics of net load uncertainty,
- Compares the performance of the new and prior methods in predicting the boundary of that uncertainty,
- Explains the similarities between the results of both methods,
- Evaluates the ability of the quantile regression to forecast uncertainty, and
- Summarizes the statistical properties of estimated coefficients.

5.1 Characteristics of net load uncertainty

This section focuses on the 15-minute net load uncertainty, rather than the 5-minute market uncertainty.

This section finds that net load uncertainty in any interval is highly correlated with net load uncertainty in the interval immediately preceding. However, the predictive power of this lagged uncertainty declines significantly when forecasting one or two hours into the future.

Figure 5.1 shows the relationship between net load uncertainty in the 15-minute market and its corresponding 15-minute lagged equivalent during February and September 2023. Blue dots indicate the values of two variables, illustrating their relationship or pattern. The red line shows a linear trend that summarizes the relationship of net load uncertainty to its own lagged value.

Net load uncertainty is highly correlated with its 15-minute lagged value (0.78). The high correlation indicates that when the net load uncertainty is high, the lagged uncertainty is similarly elevated, and the converse also holds true. This strong relationship positions the lagged uncertainty as an effective predictor for forecasting. However, this correlation decreases swiftly as the length of the lag increases. As shown in Figure 5.2, with a one-hour lag, correlation is significantly lower (0.4). Figure 5.3 indicates the correlation even decreases to 0.08 with a 3-hour lag.

Figure 5.4 illustrates the correlation between net load uncertainty and its respective lags over multiple time periods. The figure suggests that the net load uncertainty is more independent and less influenced by previous values. This fast decaying autocorrelation does not necessarily imply the absence of other patterns, but it may be necessary to explore multiple methods to identify any remaining systematic component of the net load uncertainty.

Figure 5.1 Correlation between net load uncertainty and its 15-minute lagged equivalent (February-September 2023)

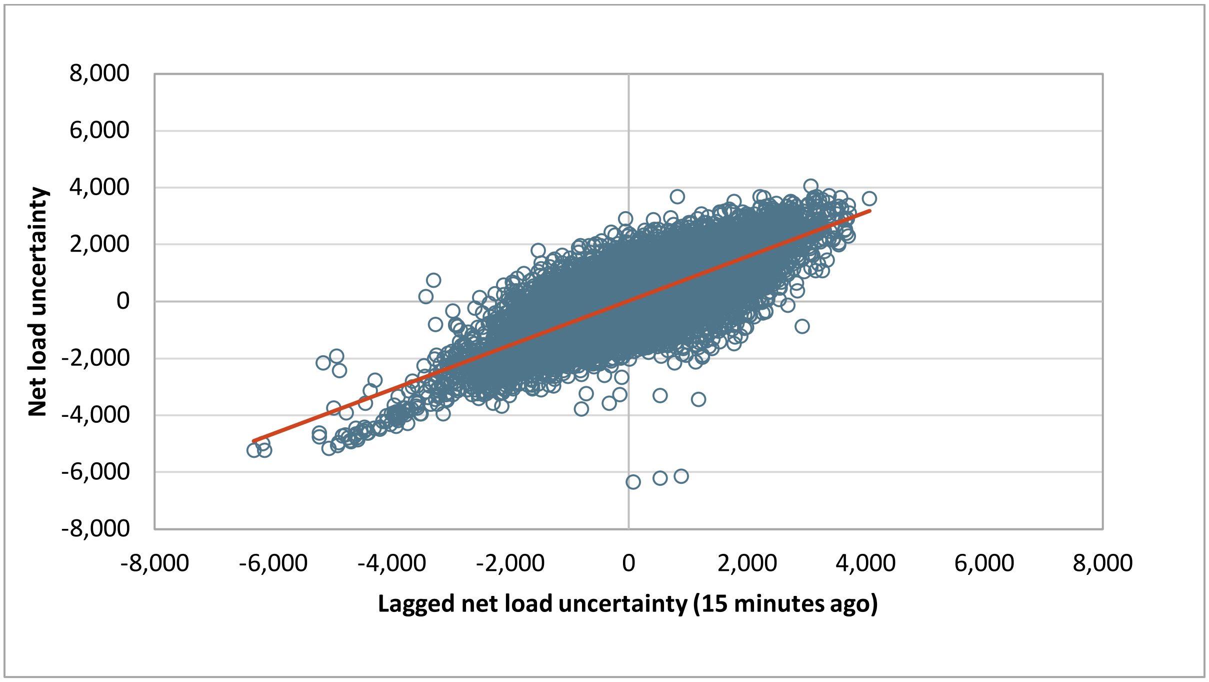


Figure 5.2 Correlation between net load uncertainty and its 1-hour lagged equivalent (February-September 2023)

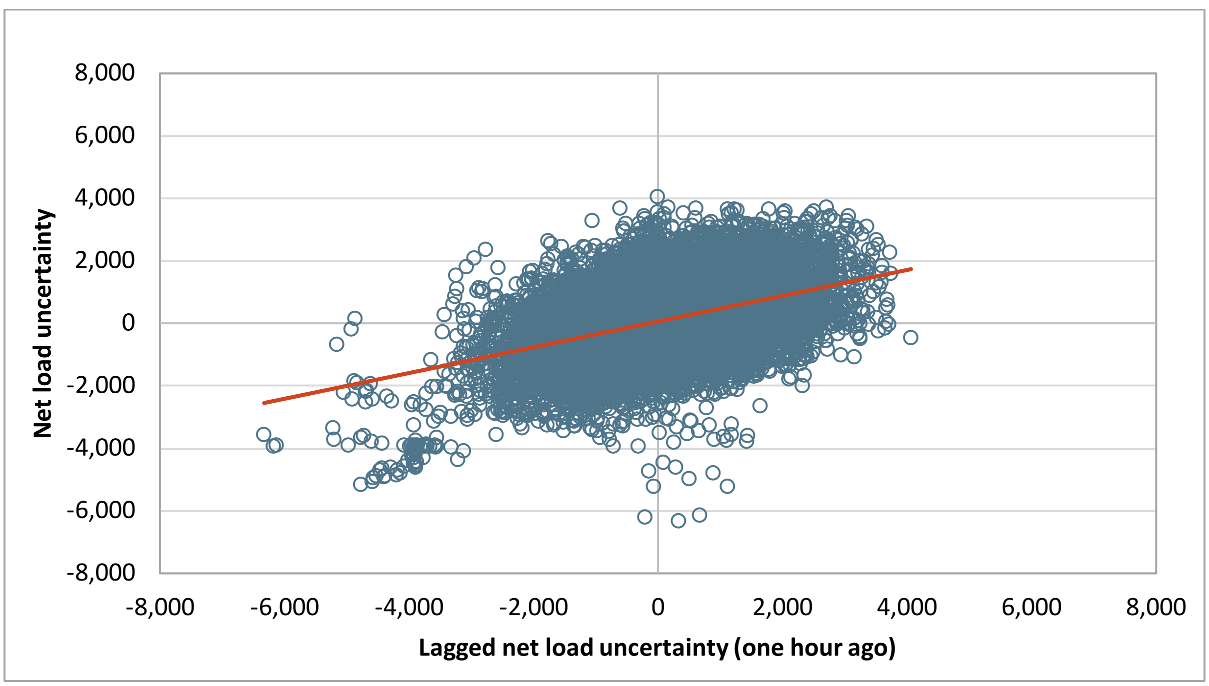


Figure 5.3 Correlation between net load uncertainty and its 3-hours lagged equivalent (February-September 2023)

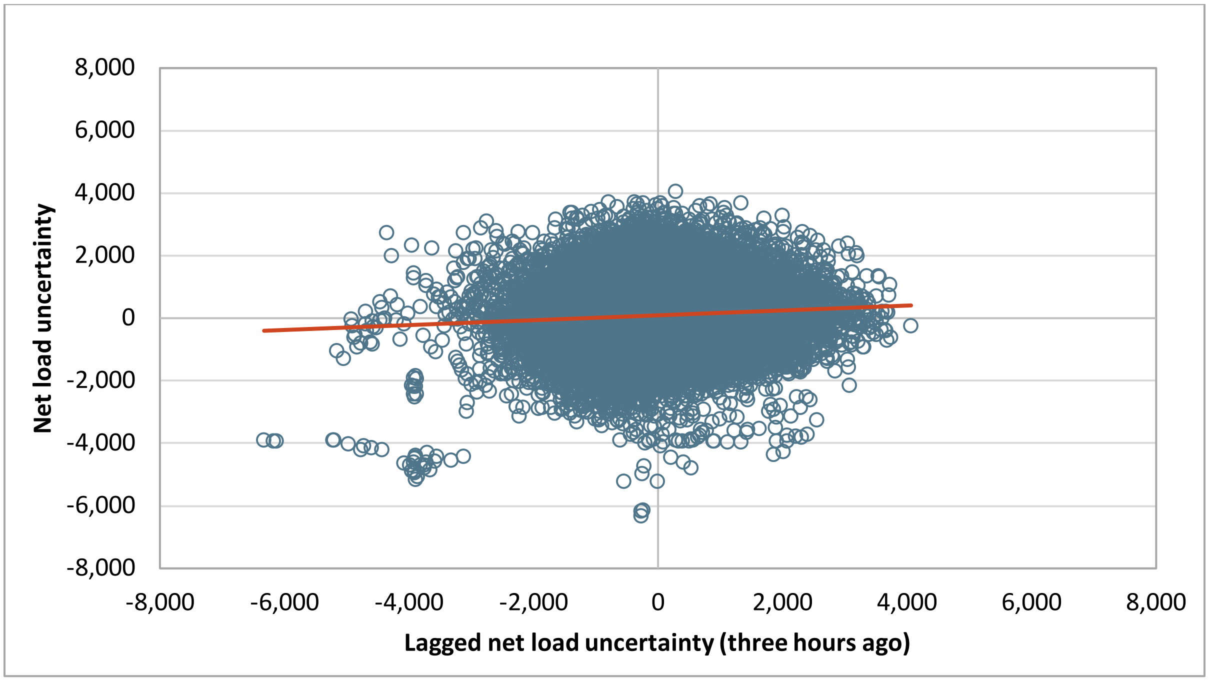
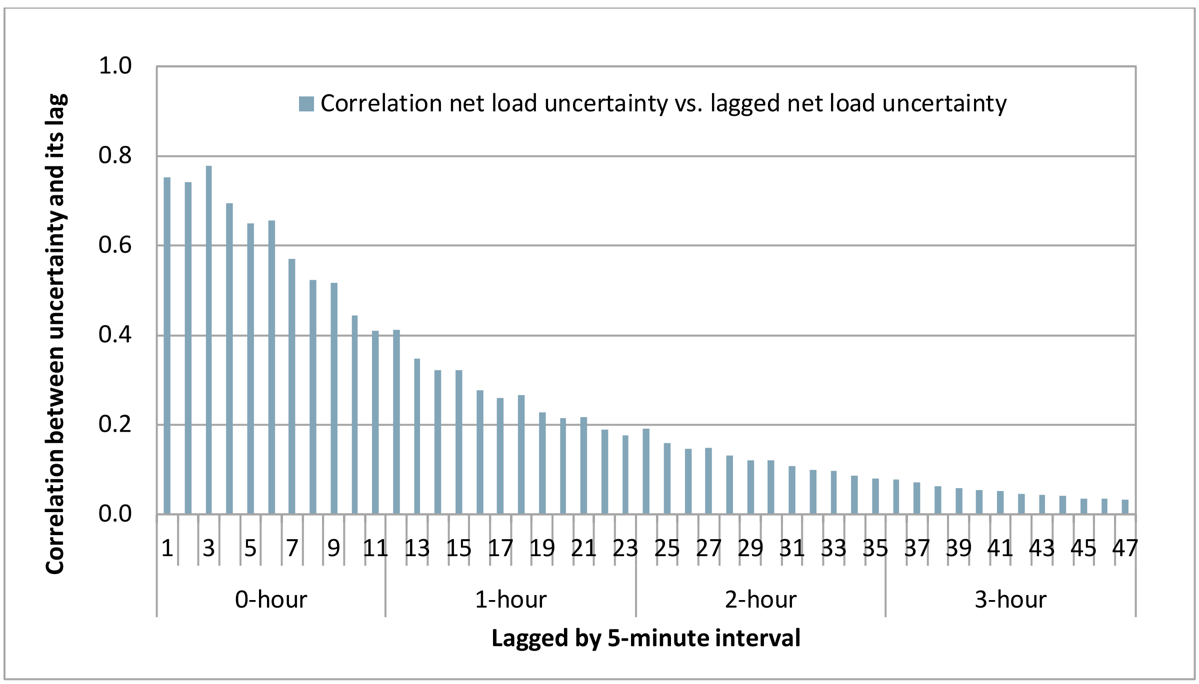


Figure 5.4 Autocorrelation of 15-minute net load uncertainty (February-September 2023)



Another noteworthy characteristic of net load uncertainty is its distribution shape. Table 5.1, which uses kurtosis to measure the distribution shape, helps summarize this varying shape of net load uncertainty of the RSE pass group in the upward direction. In a normal distribution, kurtosis is typically around 3. Numbers higher than 3 indicate a distribution with heavier tails and more pronounced peaks, meaning more extreme values are present than in a normal distribution. Conversely, kurtosis values that are less than 3 suggest a flatter distribution with fewer extremes.

For the 15-minute uncertainty, the distribution varies significantly by the hour. Uncertainty distribution after hour-ending 7, the majority of data points are gathered close to the mean, indicating a narrower range of variation. Extreme values are less common, suggesting perhaps more predictable pattern of uncertainty within a smaller range around the mean.

In the cases of uncertainty before hour-ending 7, the distribution shows a significant prevalence of extreme values or outliers. This means that the data points are more spread out with heavier tails, suggesting a higher level of variability and less predictability in the uncertainty during these hours.

In the 5-minute market, although kurtosis varies by hour, the overall distribution generally does not exhibit heavier tails compared to a normal distribution. However, there are exceptions at specific hours, notably hour-ending 6, 18, and 19, where kurtosis is notably higher. The hour-ending 19 is particularly extreme, with a kurtosis of 260. This means there is a significant occurrence of extreme values at this hour, suggesting extremely high variability and unpredictability.

Skewness describes the asymmetry of a distribution around its mean. In a normal distribution, the skewness is zero, indicating that the distribution is symmetrical.

Table 5.1 shows skewness is nearly zero on average. However, 15-minute uncertainty distribution shows a skewness of -2 from hour-ending 1 to 7. This means that the tail on the left side of distribution is wider or fatter than the tail on the right side. It suggests a tendency for more frequent occurrences of values that are lower than the mean, and that the distribution of these values is not symmetrical around the mean. This can imply certain patterns or behaviors specific to this time in the 15-minute market.

Table 5.1 Features of hourly net load uncertainty distribution: skewness and kurtosis (upward pass-group, February-September 2023)

Hour	Kurtosis (normal: 3)		Skewness (normal: 0)	
	15-minute	5-minute	15-minute	5-minute
1	21.4	0.8	-2.3	0.0
2	27.3	0.5	-3.0	0.0
3	19.7	5.6	-2.4	0.1
4	20.2	2.6	-2.4	0.2
5	16.1	2.0	-2.1	-0.2
6	17.6	15.2	-2.1	-1.2
7	9.3	0.6	-0.8	0.1
8	0.1	0.9	0.0	0.1
9	0.4	1.2	0.0	-0.3
10	0.9	1.5	0.0	-0.4
11	0.6	1.6	-0.6	-0.4
12	0.7	2.3	-0.5	-0.3
13	1.2	3.0	-0.6	-0.3
14	2.4	1.5	-0.5	0.0
15	0.5	1.7	0.0	0.1
16	0.3	1.6	0.3	0.3
17	0.4	1.8	0.4	0.2
18	-0.1	26.9	0.3	-1.4
19	0.1	259.7	0.1	8.8
20	0.3	2.6	0.1	-0.1
21	0.7	2.7	0.0	-0.3
22	0.1	1.8	-0.2	-0.1
23	0.0	8.5	-0.1	0.1
24	0.3	1.1	0.1	-0.1
Average	5.9	14.5	-0.7	0.2

Figure 5.5 shows the distribution of the 15-minute market net load uncertainty using a kernel density function. The blue bars show the actual distribution of net load uncertainty for all hours. The orange lines are the kernel distribution of the same data, which is a smoothed-out version of the blue bars. The black dashed lines represent a comparable normal distribution that uses the same mean and standard error as the net load uncertainty data.

The resulting shape of the actual net load uncertainty exhibits a higher peak than a normal distribution, suggesting a greater concentration of data points near the mean. Additionally, it also features a heavy tail that skews in the negative direction, indicating a longer and thicker tail on the lower end of distribution. The long negative tail in the distribution primarily arises from the net load uncertainty observed between hour-ending 1 and 7, the periods with high kurtosis of around 20.

Figure 5.6 presents the 15-minute uncertainty distribution at hour-ending 20, where the kurtosis is 0.3. The uncertainty distribution closely resembles a normal distribution, indicating a more even spread of values around the mean with fewer outliers or extreme variations.

Figure 5.5 15-minute net load uncertainty density function vs. normal distribution (upward pass-group, all hours, February-September 2023)

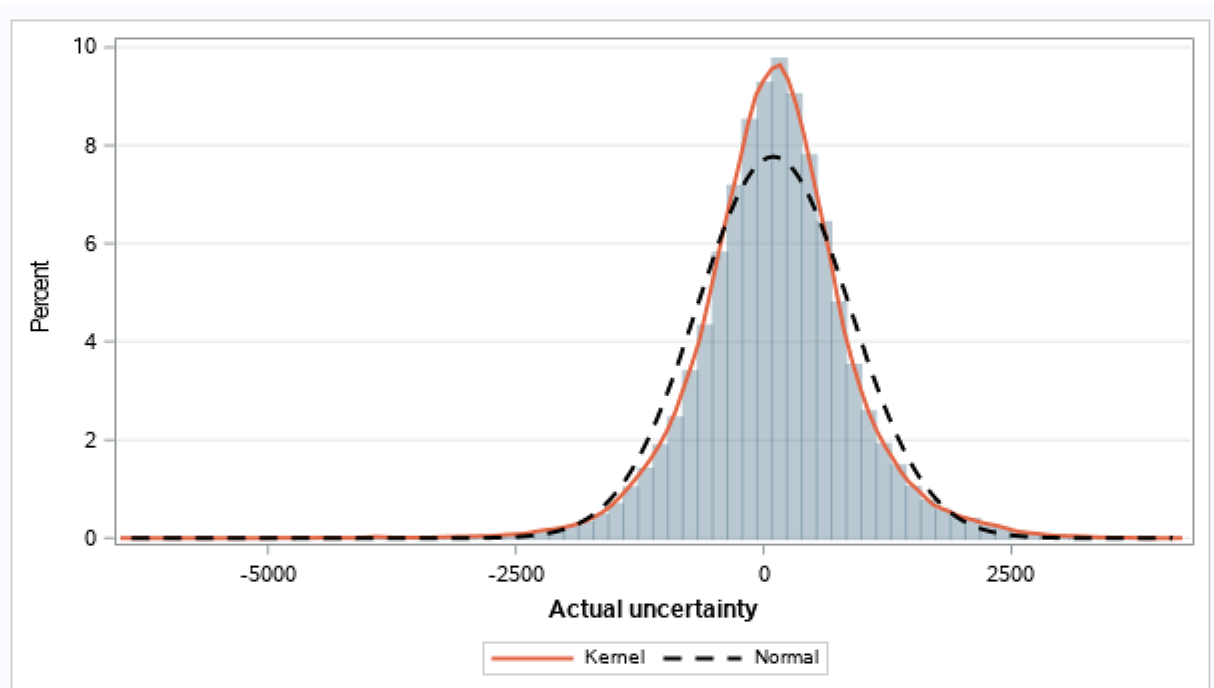


Figure 5.6 15-minute net load uncertainty density function vs. normal distribution (upward pass-group, hour-ending 20, February-September 2023)

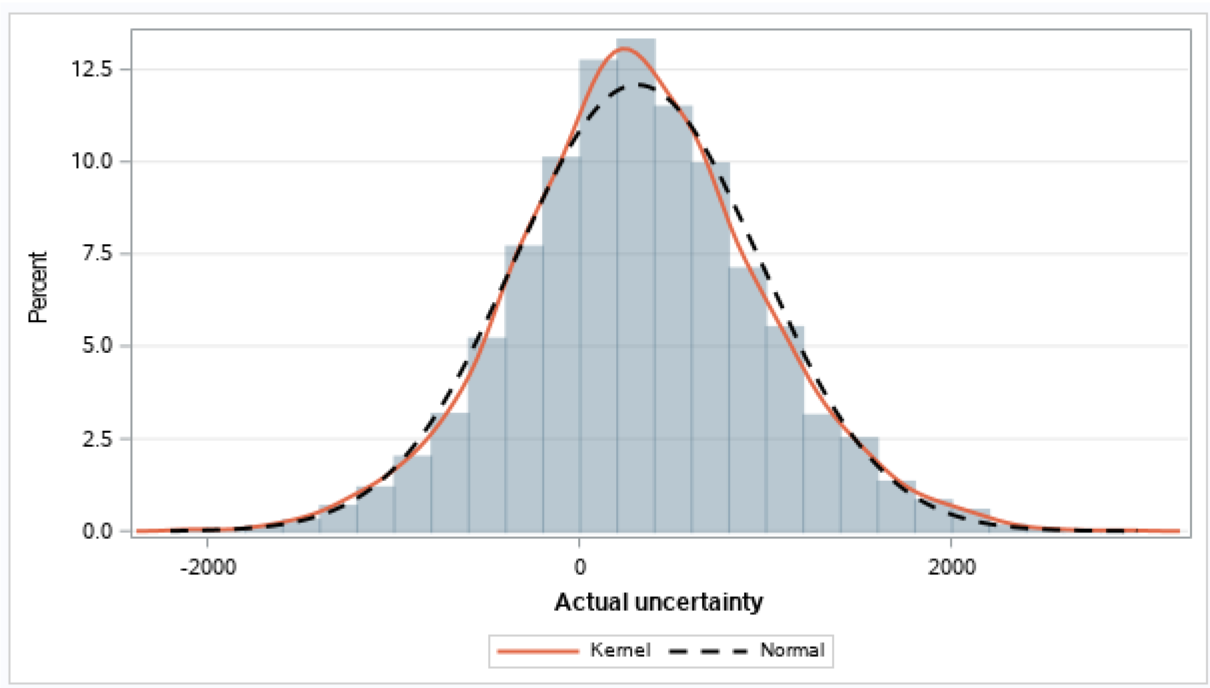


Figure 5.7 displays the 5-minute uncertainty distribution across all hours, highlighting some extreme ends of tails. These particularly long tails are primarily a result of the data for hour-ending 19, which has an exceptionally high kurtosis of 260. For most other hours, where the kurtosis is less than 3, the tails of the distribution would be much shorter.

The 5-minute uncertainty for hour-ending 20, as shown in Figure 5.8, is more clustered around the mean, with shorter tails than the overall hourly distribution. This hour's kurtosis at 0.3 suggests a normal-like distribution, but the long negative tail indicates a presence of outliers in the negative direction.

Figure 5.7 5-minute net load uncertainty density function vs. normal distribution (upward pass-group, all hours, February-September 2023)

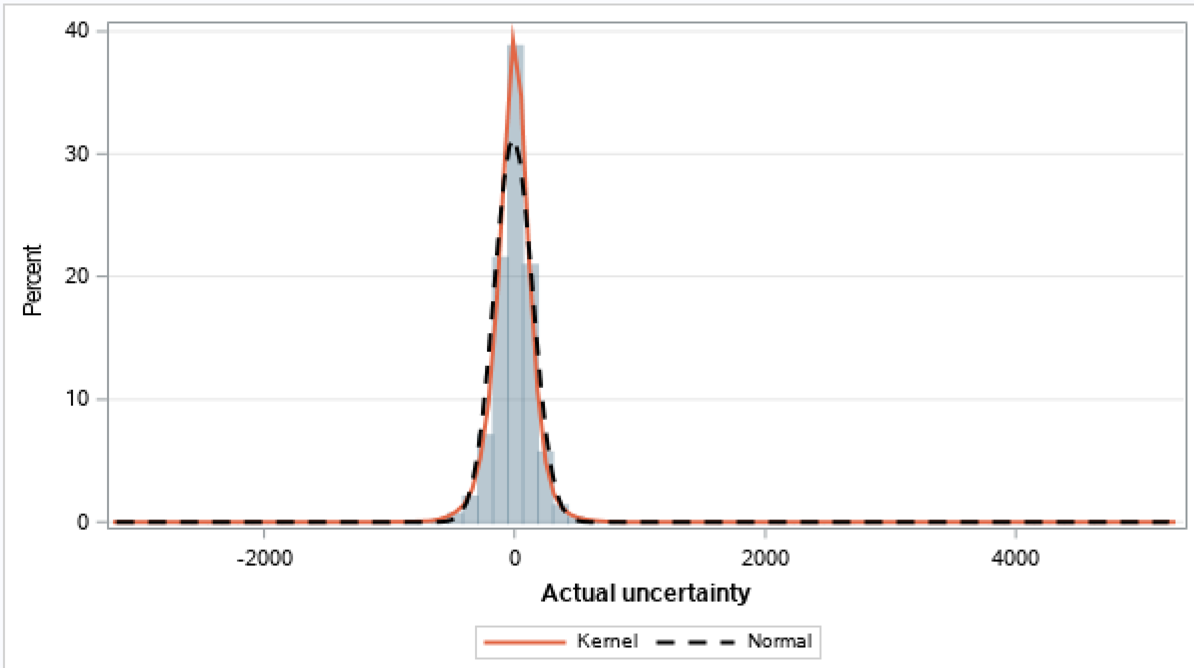
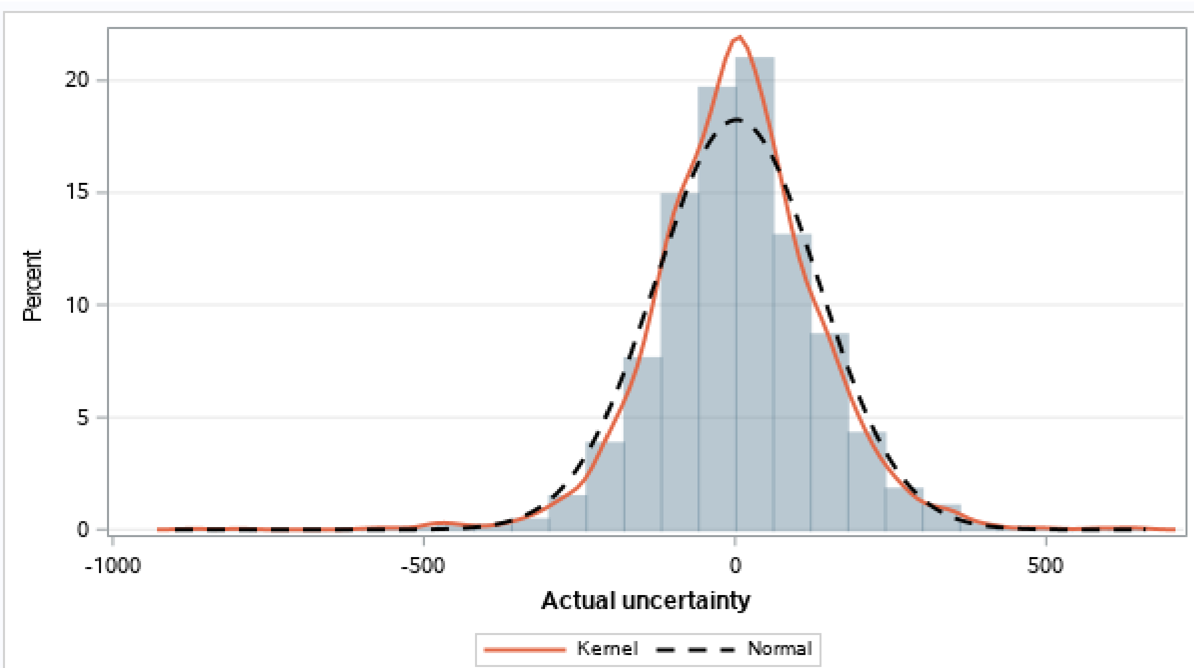


Figure 5.8 5-minute net load uncertainty density function vs. normal distribution (upward pass-group, hour-ending 20, February-September 2023)



5.2 Mosaic variable's forecasting power on net load uncertainty

This section visually demonstrates the forecasting power of the mosaic quantile regression. The mosaic variable serves as the primary predictor of net load uncertainty in the ISO calculation. This variable is an aggregate of the variations in load, solar, and wind uncertainty (forecasting error) explained by 15-minute load, solar, and wind forecast data.

According to the ISO, this method integrates weather information and 15-minute forecast data to provide a more precise depiction of fluctuating uncertainty conditions.[12] However, the ISO does not offer theoretical or empirical evidence linking higher or lower 15-minute forecast data with specific levels of net load uncertainty.[13] This leaves an unresolved gap in understanding how effectively these mosaic variables can predict net load uncertainty. To address this, DMM performed an empirical assessment of the forecasting power of the mosaic variable in predicting net load uncertainty.

DMM utilizes a scatter plot to illustrate the prediction power of mosaic quantile regression on net load uncertainty. A scatter plot is an effective tool to visualize the outcome of a regression. It depicts the relationship between two variables as a collection of points, with each point representing an observation of net load uncertainty and mosaic variable for each interval.

When the mosaic variable has no correlation with net load uncertainty, the points in the scatter plot appear randomly dispersed, forming a pattern akin to a cloud or a circle. Conversely, when there is a strong correlation, the points tend to line up in a distinct direction. A positive correlation results in an upward slope, while a negative correlation presents as a downward slope. Thus, the shape and direction of the scatter plot provides key insights into the relationship between the variables under consideration.

The potential correlation between net load uncertainty and the mosaic variable at extreme percentiles is critical. This relationship forms the primary basis for using quantile regression in forecasting. The regression captures the historical relationship between variables. Forecasting, then, involving leveraging this past relationship to predict future values.

A strong correlation implies that as one variable changes, the other does, too, in a predictable manner. This predictability is the core of the effectiveness of a forecast. If two variables are highly correlated, a change in one can be used to predict a corresponding change in the other. Thus, identifying and understanding correlations between variables is crucial for accurate and reliable forecasting.

Figure 5.9 shows a scatter plot with system-wide 15-minute upward net load uncertainty on the vertical axis and the mosaic variable on the horizontal axis for an example interval.[14] These plots summarize flexible ramping product uncertainty for the group of balancing areas that passed the resource

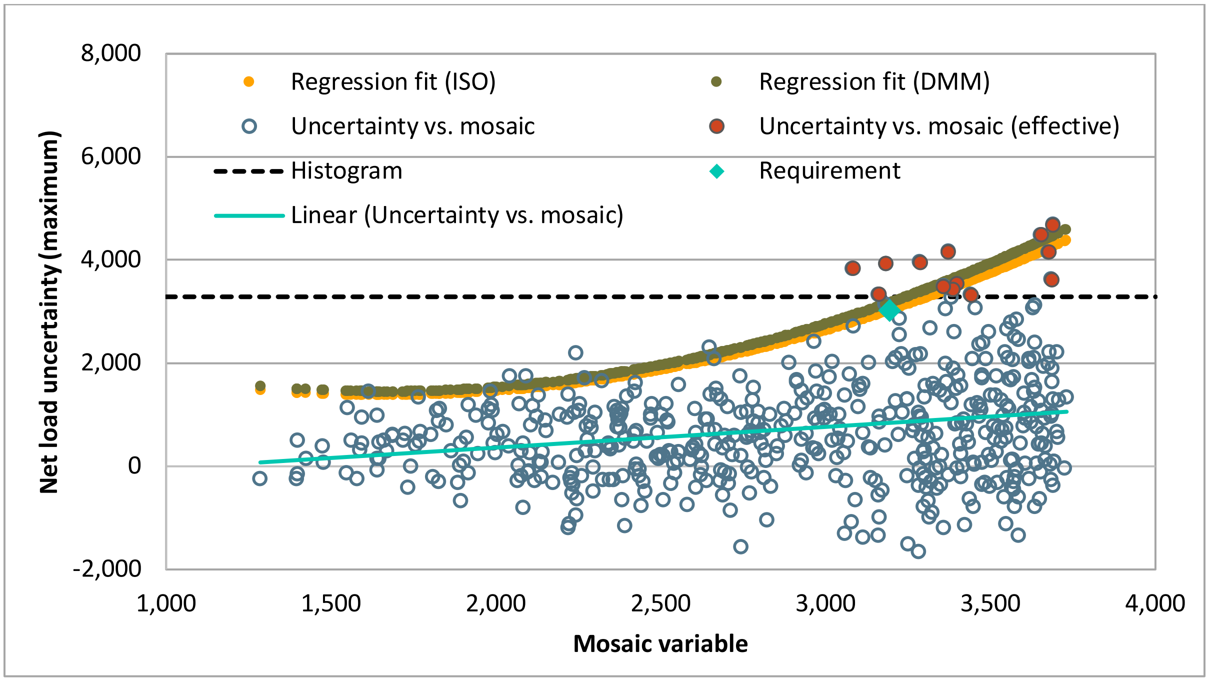
[12] The detailed documentation by the ISO on the implementation of mosaic quantile regression can be accessed from the following link: http://www.caiso.com/InitiativeDocuments/AnalysisFlexibleRampingUncertaintyCalculationintheWesternEnergyImbalanceMarket.pdf

[13] While the 15-minute forecast might be useful in scenarios where there is high renewable generation in the middle of the day and consequently high uncertainty, this explanation does not hold up when the ISO selects regression samples within the hour. Since the samples used are from the past 180 days of the same hour, the across-hour correlation between uncertainty and the 15-minute forecast data is not captured in the regression.

[14] The 15-minute net load uncertainty is calculated as the difference between the binding 5-minute market vs. the advisory 15-minute market net load forecast. In calculating upward uncertainty, the maximum of three corresponding 5-minute observations for the respective 15-minute market forecast intervals is employed.

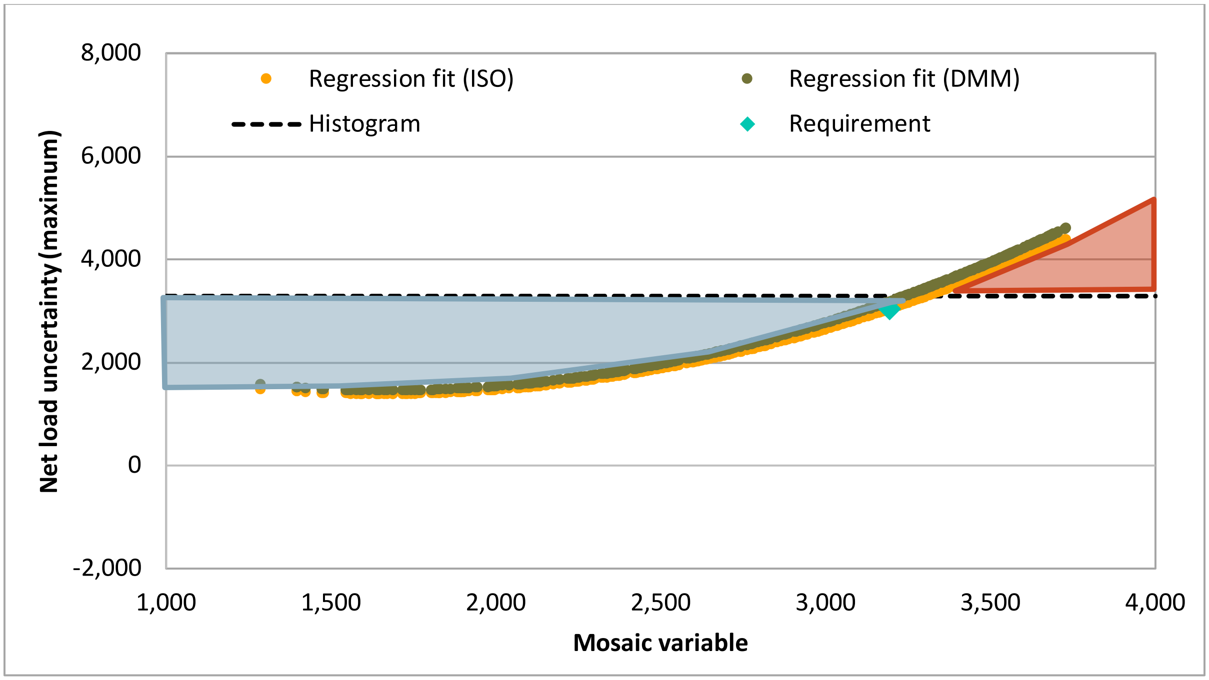
sufficiency evaluation (RSE) for the first interval in hour-ending 9 on February 15, 2023. Each blue marker on the plot corresponds to the net load uncertainty and the mosaic variable from the previous 180 days.[15] Overall, the light blue trend line represents the correlation between uncertainty and the mosaic variable.

Figure 5.9 Mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 9, February 15, 2023)



[15] To elaborate, the ISO initially selects data from the past 180 days for the mosaic quantile regression. Subsequently, if the forecasting day is a weekday, only weekday data from these 180 days are selected. Similarly, if the forecasting day is a weekend, only weekend data are chosen. The final observation depicted in the scatter plots and quantile regression are curated from these selected data points from the past 180 days, according to the day type.

Figure 5.10 Potential outcomes of mosaic quantile regression vs. histogram forecast



The green and yellow lines are the fitted values from the mosaic quantile regression.[16] The fitted value from the quantile regression signifies the estimated 97.5th percentile of the dependent variable's distribution for a given value of the mosaic variable. In other words, the fitted values show the correlation between uncertainty and the mosaic variable for those observations where uncertainty exceeds the 97.5th percentile.[17] These points are denoted by red markers in the scatter plot.

These fitted values represent the conditional quantiles of net load uncertainty. They illustrate the potential extreme end of uncertainty for a specific interval, given particular values of the mosaic variables. For example, according to Figure 5.9, if the mosaic variable for hour-ending 9 on February 15, 2023 was 2,500, then the corresponding conditional quantile values suggest that the extreme end of uncertainty could potentially reach up to 2,000.

[16] The regression fits presented in Figure 5.9 and Figure 5.10 are determined by the coefficients of the quantile regression. Coefficient B primarily captures mostly the linear trend of red markers, illustrating the linear relationship between the mosaic variable and the upper extremes of net load uncertainty. Coefficient A introduces a curvature to the regression fit. A positive value for coefficient A results in an increasingly exponential shape for the fit, signifying that the level of uncertainty grows more rapidly as the value of the mosaic variable increases. On the contrary, a negative value for coefficient A leads to a decreasing shape of the regression fit, indicating a slower rate of uncertainty growth as the value of the mosaic variable rises. Lastly, coefficient C, which serves as the intercept, pinpoints where the regression fit initiates when the value of the x-axis is zero. This represents the level of uncertainty in the absence of any influence from the mosaic variable.

[17] In the quantile regression model, when the percentile is 0.975, the shape of the fitted values primarily depends on the highest 2.5 percent of uncertainty distribution (above 97.5th percentile). The implication is that the shape of the fitted values is influenced by the top 2.5 percent of the uncertainty distribution, represented by the red markers. The remaining 2.5 percent of the shape is determined by the other data points, represented by the blue markers.

The blue diamond marker on the fitted values represents the actual mosaic value based on the current forecast information from the first interval in hour-ending 9 on February 15, along with the upper bound of the uncertainty. This upper bound serves as the upward flexible ramping requirement for the pass-group.

DMM is able to replicate the ISO quantile regression for this particular interval. This is shown by the proximity of the yellow and green fitted values in the figure. DMM employs the Simplex Algorithm to estimate the quantile regression.[18] Unfortunately, information regarding the specific algorithm employed by the ISO is not public.

The discrepancy between DMM's replication and the output by the ISO is likely attributed to the small effective sample size, specifically, those observations above the 97.5th percentile. As these effective observations primarily determine the forecast outcome, a small sample size could lead to differing outcomes, depending on the optimization method applied.

In the analysis depicted in Figure 5.9, only 13 observations were effective, constituting a very small sample size. The effective observation means that the quantile regression assigns a significant weight (97.5 percent) to these observations when estimating coefficients. The limited sample size is not exclusive to this specific interval. All quantile regressions for weekdays exhibit a similar sample size, ranging between 13 and 16 observations.[19] For weekends, the effective number of observations is typically even smaller, falling between 5 and 7. As a result, the forecasting outcomes may not be particularly meaningful or reliable.

The black horizontal lines represent the histogram output. When compared with the quantile regression fitted values, it shows that the quantile regression can potentially yield either a lower or higher requirement. If the mosaic variable for this particular interval and date falls below 3,300, then the histogram requirement is higher than that produced by the quantile regression. This is shown in the blue area in Figure 5.10. Conversely, if the mosaic variable for this particular interval exceeds 3,300, the histogram requirement is lower than the quantile regression requirement, as depicted by the red area in the figure.

Overall, Figure 5.9 shows a degree of correlation between the extreme end of uncertainty and the mosaic variable, despite the overall correlation – depicted by the blue linear trend – being relatively weak. The regression fit displays a distinct trend, with the slope exhibiting exponential growth as the mosaic variable increases. These fitted values capture extreme values of uncertainty corresponding to each value of the mosaic variable. However, the overall linear trend's slope is nearly flat, signifying that

[18] Quantile regression differs from Ordinary Least Squares (OLS) regression in its approach to estimating the relationship between variables. OLS focuses on minimizing the sum of the squared residuals, which leads to a mathematical equation that has a direct analytical solution. This is because the squared residuals from a quadratic function, which is differentiable and convex, make it easy to find a global minimum. On the other hand, quantile regression seeks to minimize the sum of absolute deviations, which does not result in a quadratic function. As a result, it does not have a simple, differentiable equation with an analytical solution. Instead, it results in an optimization problem that is solved numerically using optimization algorithms, such as the Simplex Algorithm or Interior Point method.

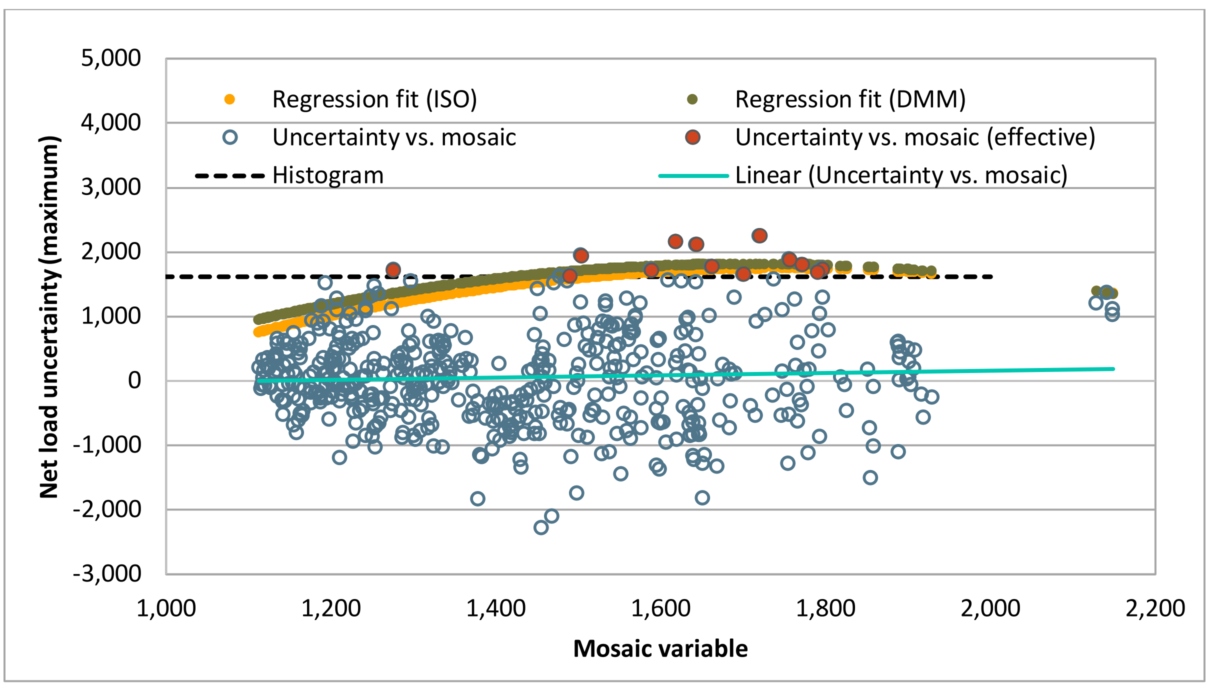
[19] For the 15-minute uncertainty quantile regression, the process begins by retaining a 180-day sample for each hour. This result in approximately 720 intervals, given that each hour contains four 15-minute intervals. Subsequently, the data is categorized based on the day type of forecast – segregating weekdays from weekends and holidays. For weekdays, the sample equates to roughly 180 x 4 x (5/7), reflecting the proportion of weekdays in a week. In contrast, for weekends, the sample size is round 180 x 4 x (2/7), corresponding to the weekends. Note that these figures may vary, depending on the number of weekends and holidays within the 180-day span.

the level of uncertainty remains largely unchanged despite increases or decreases in the mosaic variable. This observation underscores a relatively weak or perhaps even negligible relationship between the mosaic variable and net load uncertainty on average, while still allowing for meaningful variation at the extremes.

While this figure only depicts one interval on a single day in February, the majority of these observations (represented by markers) will likely remain in the sample for the next six months. This is due to the ISO practice of utilizing data from the past 180 days for its quantile regression, thereby ensuring the continued relevance of these observations in forecasting models over a substantial period.

Figure 5.11 highlights a weak predictive power of the quantile regression for hour-ending 13 on March 1, 2023.[20] A difference observed in these regression fits compared to the previous regression fits is the almost flat slope, which exhibits no indications of a linear or quadratic trend. This is visible in both the yellow line representing the ISO output, and the green line signifying DMM's replication. Such a flat trendline suggests minimal relationship between the mosaic variable and the extremes of net load uncertainty. Additionally, the overall trend between the two variables, represented by the blue linear line, is flat. This indicates that the correlation between net load uncertainty and the mosaic variable for this particular interval is essentially non-existent. In other words, changes in the mosaic variable do not correspond to changes in net load uncertainty, underscoring the limited predictive power of the mosaic variable in this context.

Figure 5.11 Mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 13, March 1, 2023)



When the regression fit closely aligns with the histogram outcomes (represented by a black dashed horizontal line), it indicates a lack of correlation between the mosaic variables and net load uncertainty. In such a scenario, the mosaic variables appear to operate as random numbers in the context of

[20] The sample groups in this case also include the BAAs that have passed the resource sufficiency test.

forecasting uncertainty. This can be understood considering that while the quantile regression provides a conditional quantile given the value of mosaic variable, the histogram depicts an unconditional or normal quantile value of uncertainty. Therefore, when these two forecasting models yield similar outputs, it suggests that the predictive power of the mosaic variable in relation to net load uncertainty is effectively negligible.

Figure 5.12 shows an example of how to interpret similar outcomes between the histogram and quantile regression. In a scenario where DMM artificially uses a random number from normal distribution instead of the mosaic variable, the quantile regression fit becomes almost identical to the histogram outcome. This provides evidence that when the mosaic variable loses its predictive power, or effectively becomes a random number, the resulting quantile regression becomes indistinguishable from the unconditional quantile given by the histogram.

Furthermore, the scatter plot depicting the relationship between uncertainty and the random number forms an elliptical shape. This shape is typically representative of an absence of correlation between two variables. This elliptical shape in the scatter plot is anticipated, as random numbers are not by definition correlated with uncertainty. The procedure employed for this random number regression mirrors that of the mosaic quantile regression in all respects except one: the independent variable is sourced from a normal distribution of random numbers with a mean of zero and a standard deviation of 1,000, rather than being a mosaic variable, even though the same equation and algorithm were applied.

Figure 5.13 shows the regression fit during the evening ramping hours on March 15, 2023. The shape of both the scatter plot and the regression fit closely mirror those in the random number example presented in Figure 5.12. The scatter plot takes on an elliptical form, and the regression fit displays only a minimal degree of quadratic shape, appearing largely flat. The fit aligns almost perfectly with the histogram outcome. This suggests a lack of substantial correlation between the mosaic variable and net load uncertainty for this particular interval.

Figure 5.12 Random number regression fit (hour-ending 13, March 1, 2023)

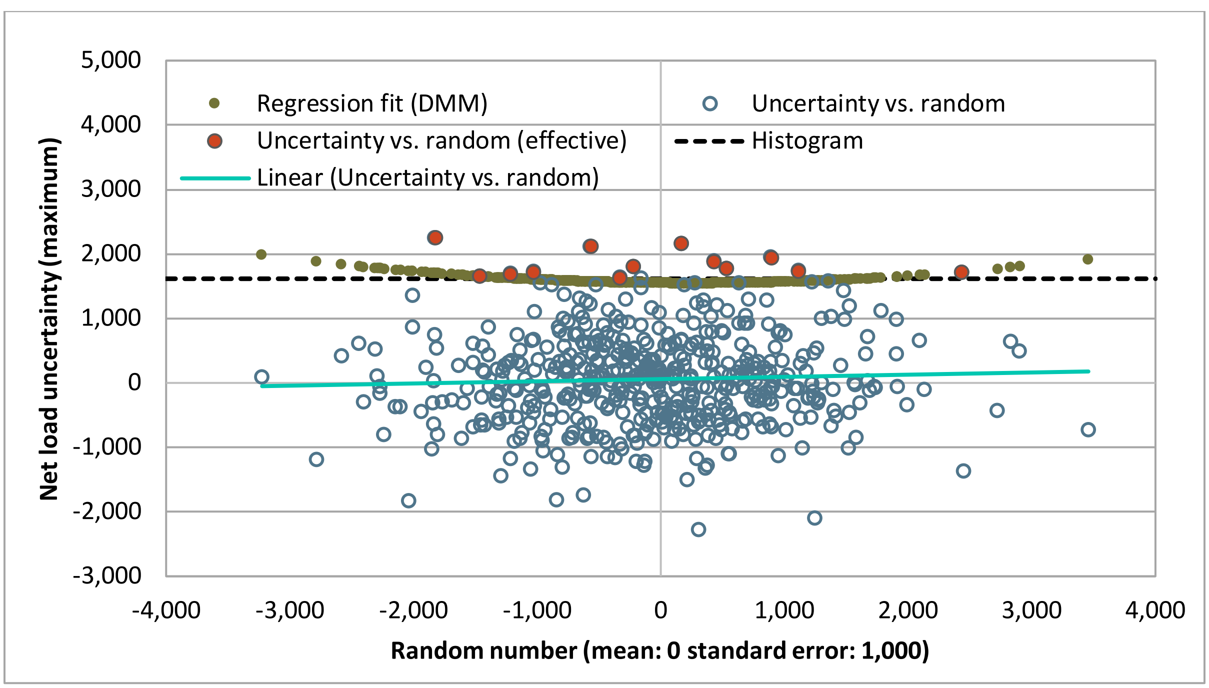
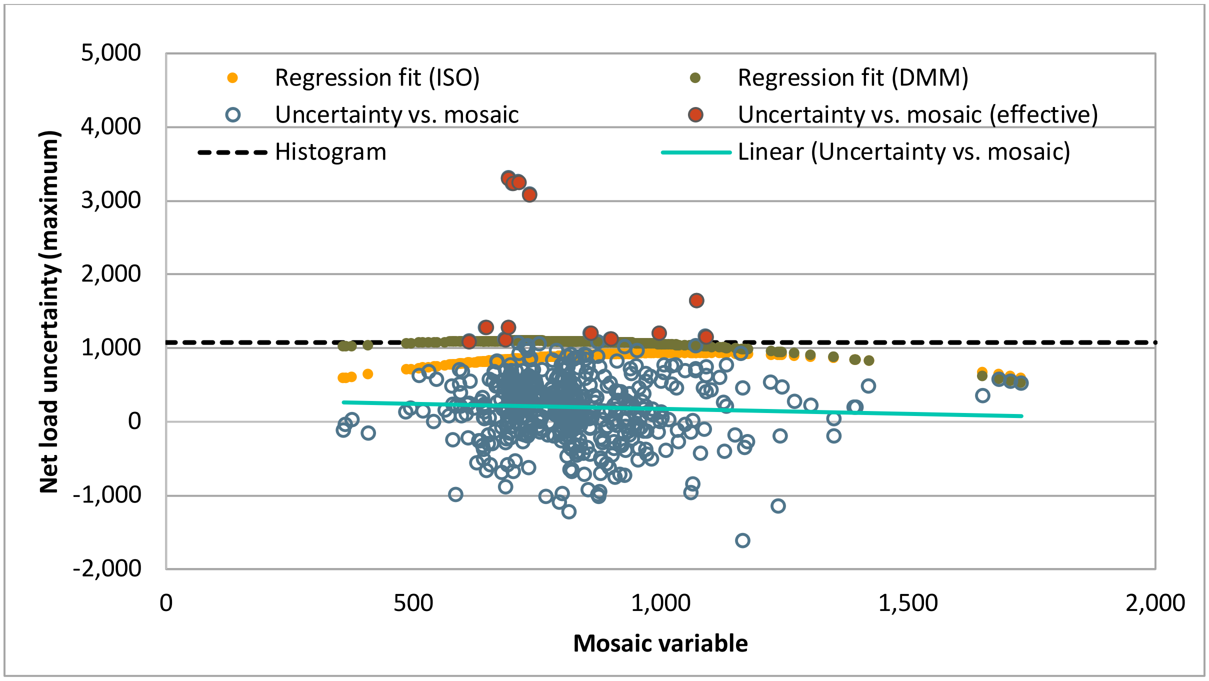


Figure 5.13 Mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 20, March 15, 2023)



While the scatter plots provided above only cover specific intervals during the first quarter of 2023, the similarity of outcomes from the quantile regression and histogram across all intervals in this quarter suggest a consistent trend. This underlines the limited predictive power of the mosaic variable on uncertainty. It is important to understand that this resemblance between the two forecasting models' outcomes is not coincidental, but rather a reflection of the low correlation between the mosaic variable and uncertainty. The example charts above serve as visual explanations of the underlying mechanism.

The following *six figures* illustrate mosaic regression fits for selected days of each month, encompassing a mix of weekdays and weekends. These figures also show that when the mosaic variable correlates with uncertainty, the forecasts deviate from the histogram outcomes. However, in cases where there is no meaningful correlation – either on average or only at the extreme ends of uncertainty – the forecasts resemble those of histogram.

Figure 5.14 Mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 19, April 15, 2023)

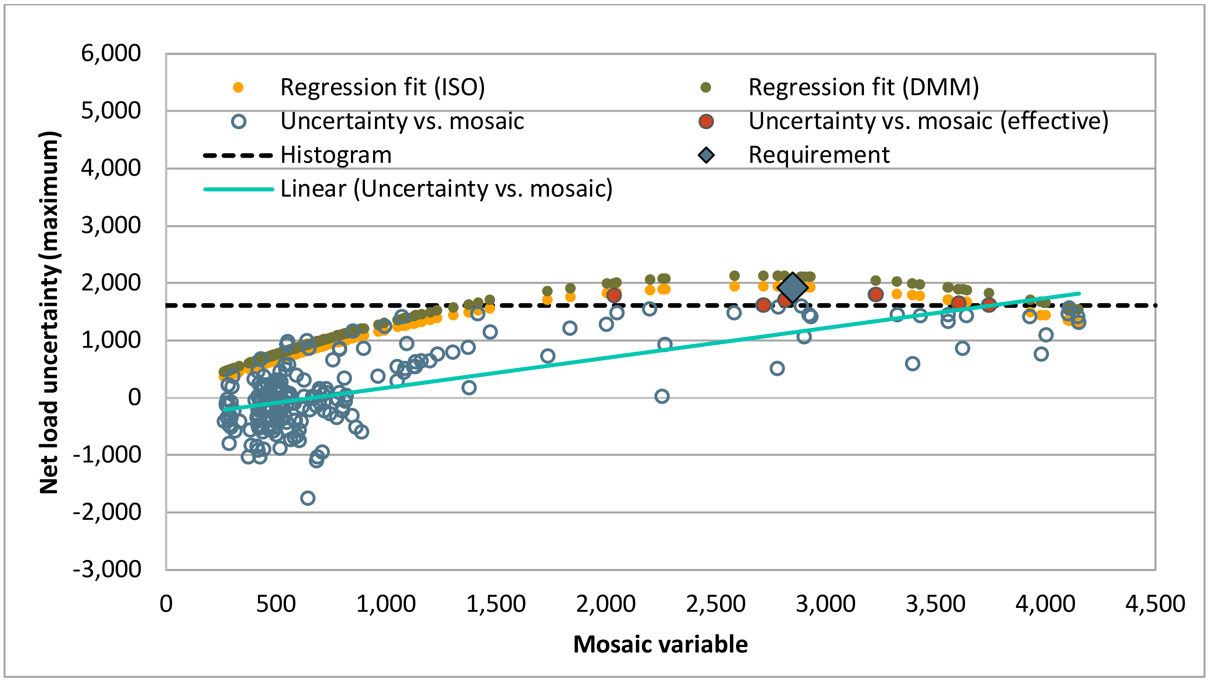


Figure 5.15 Mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 19, May 15, 2023)

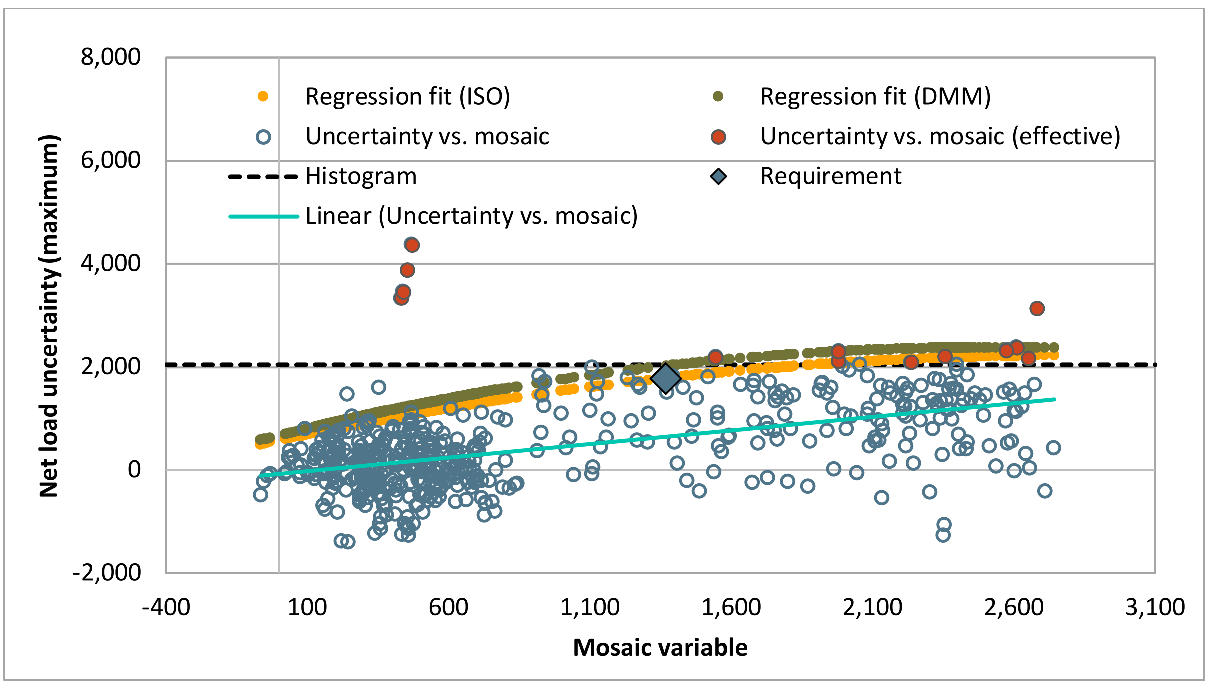


Figure 5.16 Mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 19, June 15, 2023)

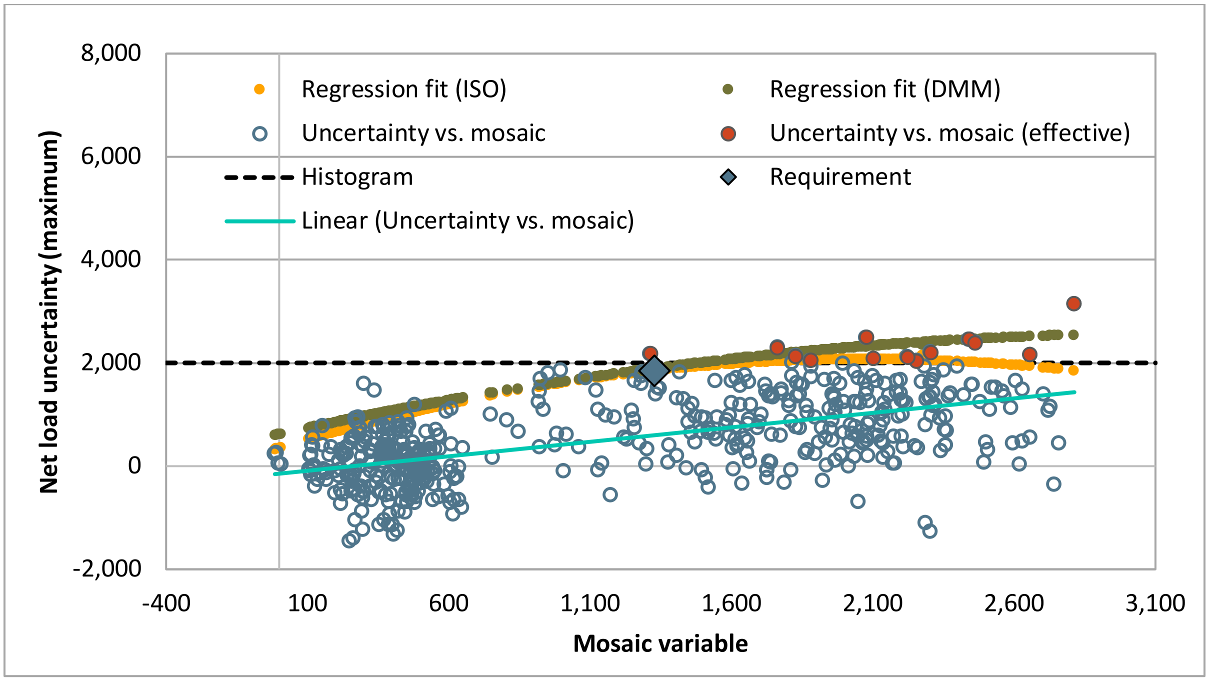


Figure 5.17 Mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 19, July 15, 2023)

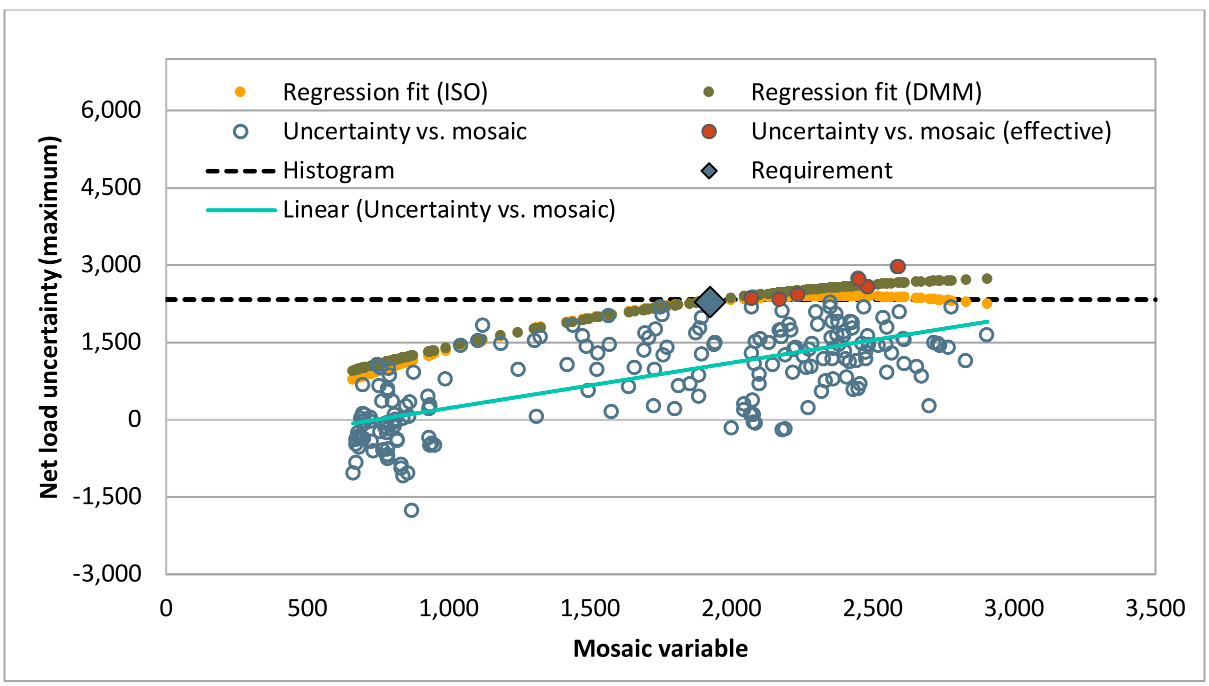


Figure 5.18 Mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 19, August 15, 2023)

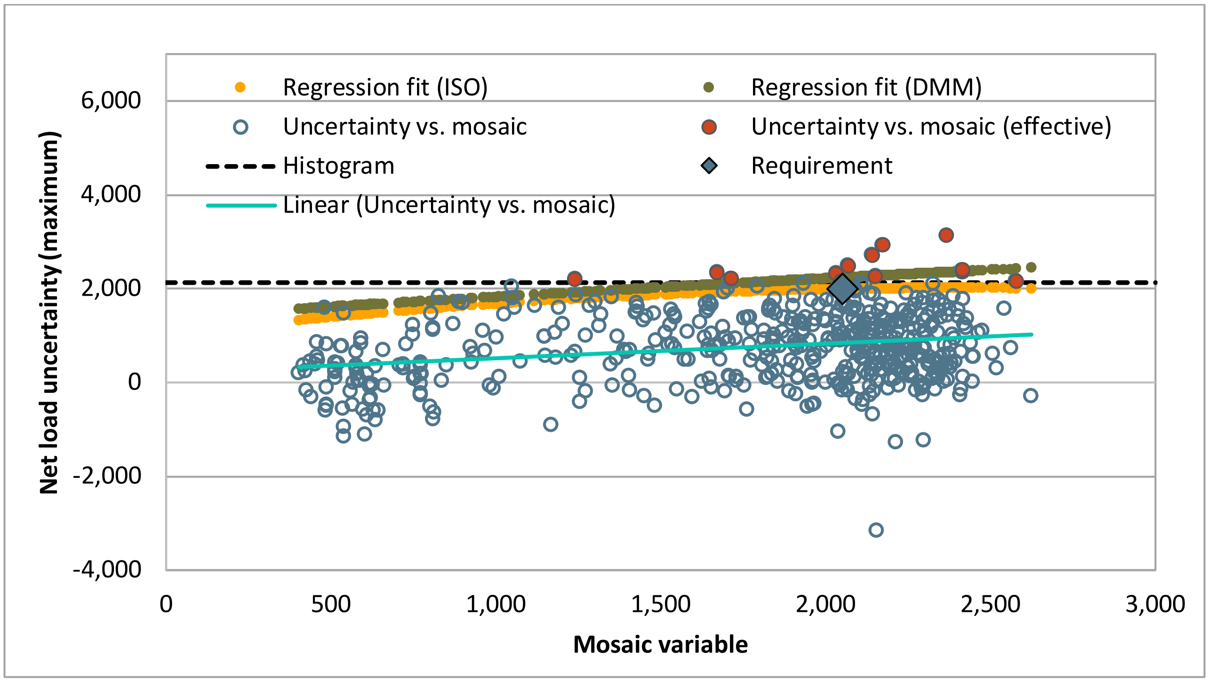
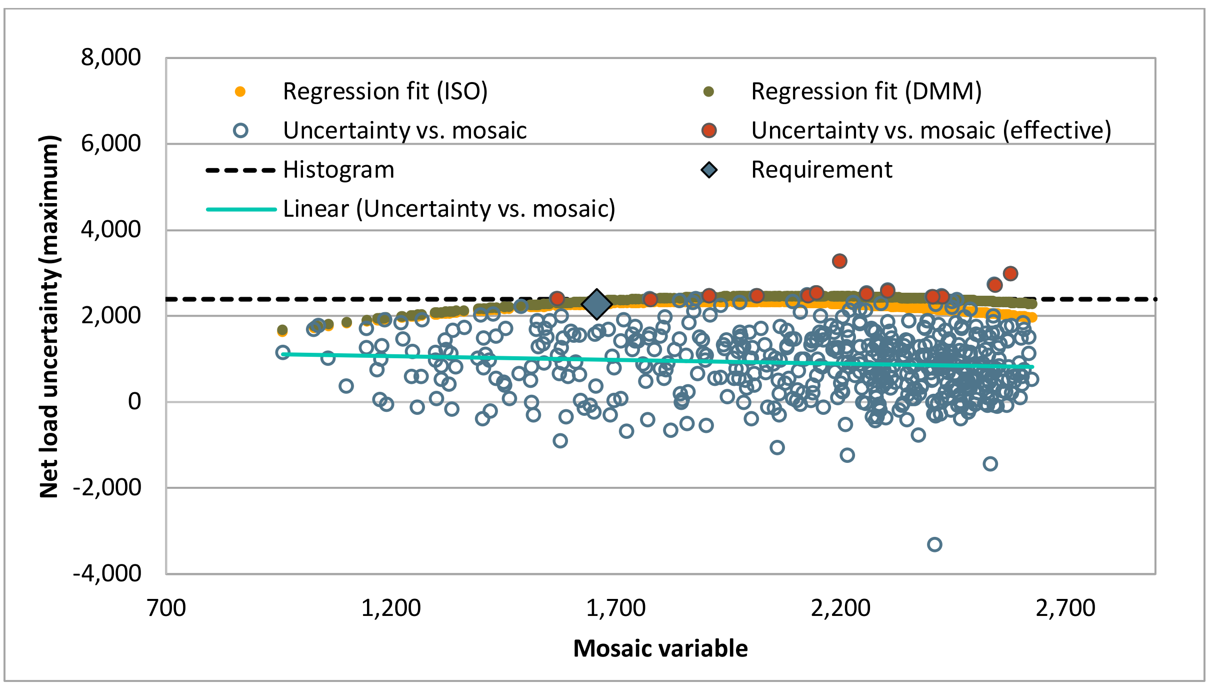


Figure 5.19 Mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 19, September 15, 2023)



5.3 Assessing the reliability of coefficients in quantile regression

This section employs statistical analysis to further examine this correlation, providing a more comprehensive understanding beyond visual scatter plot interpretations. DMM conducts statistical tests to determine whether the coefficients of the series of quantile regression in FRP are significantly different from zero.

This test outcome does not imply that the coefficients are zero. Rather, it highlights the regression's struggle to identify a definite connection between the extreme end of uncertainty and the mosaic variable. This ambiguity in findings could also stem from a small effective sample size because it gets harder to detect and confirm relationships in the data.

Regression analysis might seem to produce a single coefficient that summarizes the relationship between uncertainty and the mosaic variable, but it estimates a distribution of coefficients. The coefficient used in forecasting is typically the mean of this distribution. The true nature of the relationship is unknown. Instead, the regression identifies the likely range of coefficients, providing a range of possible outcomes rather than a single value. This method acknowledges the uncertainty inherent in data analysis.

If the distribution is wide, the mean is not accurately representing the actual relationship. On the other hand, if the coefficient distribution is tightly clustered around the mean, then it is more likely that the true relationship closely aligns with this mean.

If a test indicates that the coefficients are not statistically different from zero, it suggests that the estimated coefficient is in a range containing zero. That can happen if either the coefficient distribution is quite wide or if the mean of the coefficient distribution could be zero, rather than the specific figure initially given by the regression.

When a coefficient is statistically different from zero, it suggests that the observed association is likely to be a true effect rather than a result of chance or sampling variability. This reliability is essential in forecasting, as it provides confidence that the patterns identified in the input data will persist or recur in the future. Without statistical significance, the relationship might be too weak to forecast.

It is important to note that the test result indicates either a weak or a strong relationship, specifically at the extreme ends of uncertainty in relation to the mosaic variable. This is because the coefficients derived from mosaic quantile regression do not represent an average relationship but are specific to the upper and lower 2.5 percentiles of uncertainty and their corresponding mosaic variables.

Since quantile regression does not have a straightforward analytical solution for standard errors of coefficients, DMM employs the bootstrap method, a resampling-based approach used for estimating the sampling distribution of coefficients.[21] By repeatedly sampling from the last 180 days of dataset and re-estimating model, DMM can empirically derive the distribution of the coefficients and their standard

[21] Bootstrap resampling is a statistical technique that involves drawing repeated samples from the original data, and recalculating the statistic or model for each resampled data set. This bootstrap approach is frequently used in quantile regression because the standard errors estimated this way are robust to the specific assumptions about the underlying distribution of errors. For further information, refer to *Koenker, R., Chernozhukov, V., He, X., & Peng, L. (Eds.). (2017). Handbook of quantile regression.*

errors. Once such distribution is available, DMM conducts a Wald test to determine whether these coefficients are significantly different from zero.[22]

It is important to clarify that the tests conducted here are based on coefficients from DMM's replication of the model. Although DMM's replication aligns closely with the ISO quantile regression in terms of fit and requirement, the coefficient values do not always become identical or close to those of the ISO. Thus, the result might not exactly match the ISO calculations. This is due to the small effective sample size and the unknown nature of the algorithm that the ISO employs for quantile regression.

Table 5.2 and Table 5.3 show the percentage of coefficients that are statistically different from zero for each hour between February and September. The data here summarizes the coefficients used to calculate flexible ramping product uncertainty requirements for the group of balancing areas that passed the resource sufficiency evaluation.[23]

In these tables, DMM includes not only coefficients from the mosaic quantile regression, but also coefficients from the load, wind, and solar quantile regression, which were initially run to construct the mosaic variables. The coefficients A and B are particularly worth testing for statistical significance. Coefficient B captures the linear trend between the dependent and independent variables, indicating the direction and rate of change in the dependent variable for a one unit change in the independent variable. Coefficient A adds curvature to the trend, determining the extent to which the relationship between the variables deviates from linearity.[24]

Table 5.2 reveals that out of all the coefficients in DMM's replication for calculating upward uncertainty (97.5 quantile regression), only 35 percent were statistically significant and meaningfully different from zero. This indicates that these 35 percent of coefficients demonstrate a relationship between the uncertainty and 15-minute forecast data, providing meaningful information for the regression model. The remaining 65 percent of coefficients were not distinguishable from zero, suggesting that they lack a significant impact on the uncertainty and do not contribute substantially to the model's predictive power.

[22] The Wald test is a statistical test commonly used in conjunction with bootstrapping to evaluate the significance of coefficients in a regression model. For further information, refer to *Davidson, R., & MacKinnon, J. G. (1999). Bootstrap testing in nonlinear models. International Economic Review, 40(2), pp. 487-508.*

[23] The significance level in hypothesis testing represents the threshold at which one is willing to reject the possibility that the coefficients are zero, given the evidence from the sample data. A significance of level 0.1 is a less stringent level, meaning it requires less evidence to consider a coefficient significantly different from zero. A more stringent level, such as 0.01 would require stronger evidence against the null hypothesis – in this case, the assumption that the coefficients are zero.

[24] Coefficient C, the intercept, is a constant term in the regression equation. It represents the expected value of the dependent variable when all the independent variables are zero. While the intercept is a valuable component of a regression model, it often does not undergo significance testing because its value does not influence the relationships between the dependent and independent variables.

Table 5.2 Percentage of statistically non-zero coefficients for upward pass-group uncertainty (February-September 2023)[25]

Hour	Load		Solar		Wind		Mosaic		Average
	coef A	coef B	coef A	coef B	coef A	coef B	coef A	coef B	
1	13%	12%	40%	43%	9%	14%	16%	16%	20%
2	23%	19%	31%	55%	46%	47%	29%	32%	35%
3	21%	25%	19%	45%	44%	36%	14%	16%	27%
4	58%	61%	4%	39%	1%	7%	20%	21%	26%
5	67%	70%	5%	3%	17%	17%	19%	20%	27%
6	10%	16%	81%	99%	11%	11%	10%	13%	31%
7	19%	18%	90%	100%	37%	45%	8%	16%	42%
8	28%	30%	94%	99%	44%	67%	16%	4%	48%
9	14%	16%	66%	55%	29%	40%	20%	23%	33%
10	7%	8%	96%	88%	55%	68%	13%	17%	44%
11	37%	37%	72%	49%	65%	79%	10%	18%	46%
12	14%	13%	81%	66%	24%	47%	30%	46%	40%
13	20%	23%	88%	69%	23%	58%	9%	22%	39%
14	25%	29%	89%	81%	20%	32%	12%	9%	37%
15	36%	36%	89%	88%	11%	7%	10%	25%	38%
16	36%	36%	87%	66%	16%	18%	20%	26%	38%
17	48%	49%	92%	91%	47%	46%	8%	38%	52%
18	44%	47%	81%	85%	68%	73%	31%	51%	60%
19	26%	26%	89%	100%	20%	28%	34%	61%	48%
20	39%	42%	51%	79%	28%	42%	12%	30%	40%
21	23%	24%	34%	33%	23%	26%	16%	18%	24%
22	47%	48%	42%	55%	20%	36%	6%	5%	32%
23	21%	23%	51%	65%	8%	20%	13%	14%	27%
24	46%	46%	54%	57%	13%	6%	7%	5%	29%
Average	30%	31%	64%	67%	28%	36%	16%	23%	37%

The performance of the mosaic coefficients, which determine the flexible ramp requirement, exhibited the lowest significance. Only 20 percent of these coefficients were statistically non-zero, indicating a weak relationship between the mosaic variable and the net load uncertainty. For the majority of hours, less than 20 percent of the mosaic coefficients were statistically non-zero, while there was slight improvement during the evening ramping hours between 5 p.m. and 7 p.m.

The load coefficients showed fluctuating performance hour to hour, with hour-ending 10 having only 7 percent of coefficients that were statistically non-zero, while the following hour had over 37 percent of coefficients as statistically non-zero. This pattern was observed across all hours.

[25] In Table 5.2 and Table 5.3, analyses have been conducted excluding any missing values.

The wind coefficients exhibited a similar trend, with over 70 percent of coefficients being statistically significant around 5 p.m. However, the overall performance of the wind coefficients was 26 percent.

The solar coefficients demonstrated the highest rate of statistically non-zero coefficients, averaging at 65 percent. Particularly during the morning and evening ramping hours, over 90 percent of the solar coefficients were statistically non-zero. Since the regression only samples the same hour across the 180 days, it indicates the solar quantile regression incorporated the changing timing of sunrise and sunset across the past 180 days.

Table 5.3 summarizes the coefficients for calculating downward uncertainty, as replicated by DMM. Much like the results from the 97.5 quantile regression, DMM found that on average, only 32 percent of these coefficients were statistically different from zero. This implies that throughout most intervals, no significant variation was observed between uncertainty and the 15-minute forecast data.

Table 5.3 Percentage of statistically non-zero coefficients for downward pass-group uncertainty (February-September 2023)

Hour	Load		Solar		Wind		Mosaic		Average
	coef A	coef B	coef A	coef B	coef A	coef B	coef A	coef B	
1	17%	14%	0%	99%	37%	37%	9%	15%	28%
2	9%	8%	22%	95%	72%	76%	14%	14%	39%
3	7%	8%	0%	73%	33%	33%	38%	35%	29%
4	8%	12%	0%	80%	25%	42%	20%	26%	27%
5	2%	3%	0%	99%	36%	38%	7%	9%	24%
6	41%	38%	91%	98%	25%	22%	16%	15%	43%
7	16%	17%	99%	100%	39%	41%	46%	56%	52%
8	23%	27%	95%	95%	50%	58%	41%	34%	53%
9	17%	19%	24%	6%	20%	38%	18%	30%	22%
10	59%	61%	34%	42%	43%	49%	36%	56%	48%
11	33%	33%	22%	27%	16%	23%	16%	42%	27%
12	15%	15%	8%	9%	9%	19%	11%	50%	17%
13	9%	9%	28%	32%	9%	21%	23%	52%	23%
14	13%	9%	7%	5%	56%	65%	23%	37%	27%
15	15%	15%	24%	21%	28%	33%	13%	38%	23%
16	13%	14%	73%	63%	20%	37%	15%	20%	32%
17	11%	12%	90%	88%	51%	69%	23%	25%	46%
18	11%	13%	96%	91%	25%	28%	33%	42%	42%
19	25%	26%	93%	99%	13%	21%	24%	39%	43%
20	26%	27%	81%	99%	23%	28%	15%	27%	41%
21	12%	13%	25%	71%	6%	21%	15%	11%	22%
22	7%	6%	10%	58%	18%	28%	23%	18%	21%
23	11%	12%	0%	81%	13%	13%	21%	31%	23%
24	3%	4%	3%	96%	44%	41%	16%	21%	28%
Average	17%	17%	39%	68%	30%	37%	22%	31%	32%

The mosaic quantile regression revealed that only 27 percent of intervals exhibited statistically significant non-zero coefficients. In certain hours, less than 10 percent of the coefficients differed significantly from zero. Typically, during the morning ramping hours, between hour-ending 6 and 8, DMM observed 30-50 percent of the coefficients to be statistically significant. In the middle of the day, the mosaic quantile regression captured some linear trend, with coefficient B providing statistically significant results for about 50 percent of intervals. However, performance declined post-hour-ending 20, where less than 20 percent of intervals displayed statistically significant results.

DMM's load coefficients demonstrated that, on average, only 17 percent of intervals yielded statistically significant coefficients. The bulk of these significant coefficients was concentrated during specific hours, typically in the morning, ranging from 17 to 61 percent. However, the remaining hours exhibited less than 5 percent of intervals with statistically significant results.

In comparison, the solar and wind quantile regression performed relatively better, despite still representing a low percentage of statistically significant coefficients – 54 percent for solar and 34 percent for wind on average. The solar coefficients displayed the most dramatic hourly fluctuations in significant coefficients. During morning and evening ramp-up hours, all coefficients were statistically significant, but this fell to 20-30 percent from hour-ending 9 through hour-ending 14. Similar to the 97.5 solar quantile regression, this pattern might indicate variations caused by changing sunrise and sunset times over the past 180 days, which the regression seems to capture.

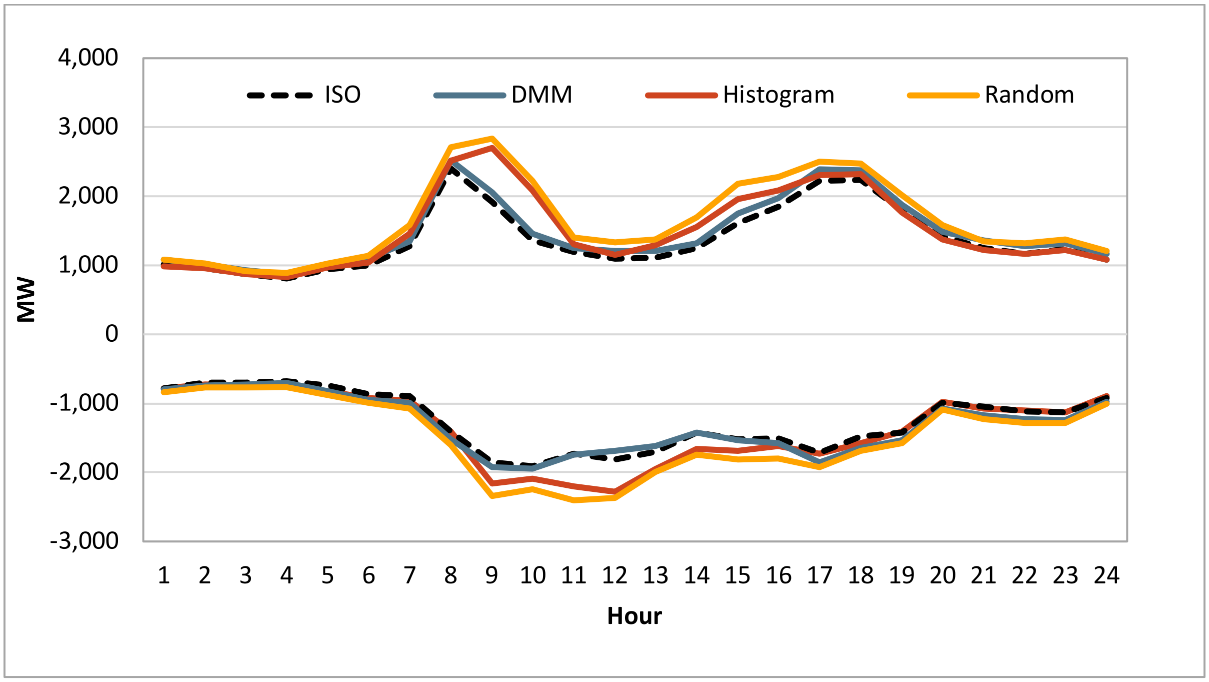
On the other hand, the wind quantile regression demonstrated relatively high performance from hour-ending 1 to hour-ending 10, and again from hour-ending 14 to hour-ending 18, with over 40 percent of intervals consistently demonstrating statistically significant coefficients.

Overall, DMM replication of both 97.5 and 2.5 quantile regression found that only 35 percent of coefficients were statistically different from zero. Although this 35 percent figure may not precisely mirror the ISO calculation, it does raise concern about the actual performance of the mosaic quantile regression. Relying on non-statistically significant coefficients for forecasting suggests an unsubstantiated level of predictability. These coefficients failed to demonstrate a clear, statistically significant relationship with the uncertainty, compromising their predictive reliability. Use of non-significant coefficients may result in inaccurate or misleading forecasts.

Figure 5.20 illustrates the implications of using a non-significant coefficient in mosaic quantile regression. It compares the hourly average flexible ramping product requirements (with thresholds) derived from four distinct quantile regression outcomes. The data comprises all intervals from February and September 2023 for balancing areas that passed the resource sufficiency evaluation. The black dashed line represents the forecasted upper and lower bounds of net load uncertainty as determined by the ISO mosaic quantile regression, calculated using their coefficients. The blue line depicts DMM's replication, which used DMM's replicated coefficients. The red line shows the histogram requirement based on the past 180 days of net load uncertainty.

Lastly, the yellow line shows forecasted outcomes based on the random number quantile regression of net load uncertainty. In this case, the procedure mirrors that of the ISO mosaic quantile regression. However, instead of using the mosaic variable as the independent variable, DMM used a random number with a mean of zero and a standard deviation of 1,000.

Figure 5.20 Hourly average flexible ramping requirement across different models (with threshold, February-September 2023)



The replication performed by DMM, depicted in the blue line, closely mirrored the ISO forecast, represented by the black dashed line. The random number quantile regression forecast also aligned closely with these two forecasts and the histogram outcomes. By design, the random number quantile regression produces statistically non-significant coefficients because the random numbers are uncorrelated with any variables, including net load uncertainty.

Figure 5.20 shows that the regression analysis based on a weak relationship tends to converge with the histogram method. This convergence is primarily because the quantile regression estimates conditional percentiles, whereas the histogram method estimates unconditional percentiles. When the conditional percentile is weak, such as the random number regression, the forecast leads to the unconditional outcomes.

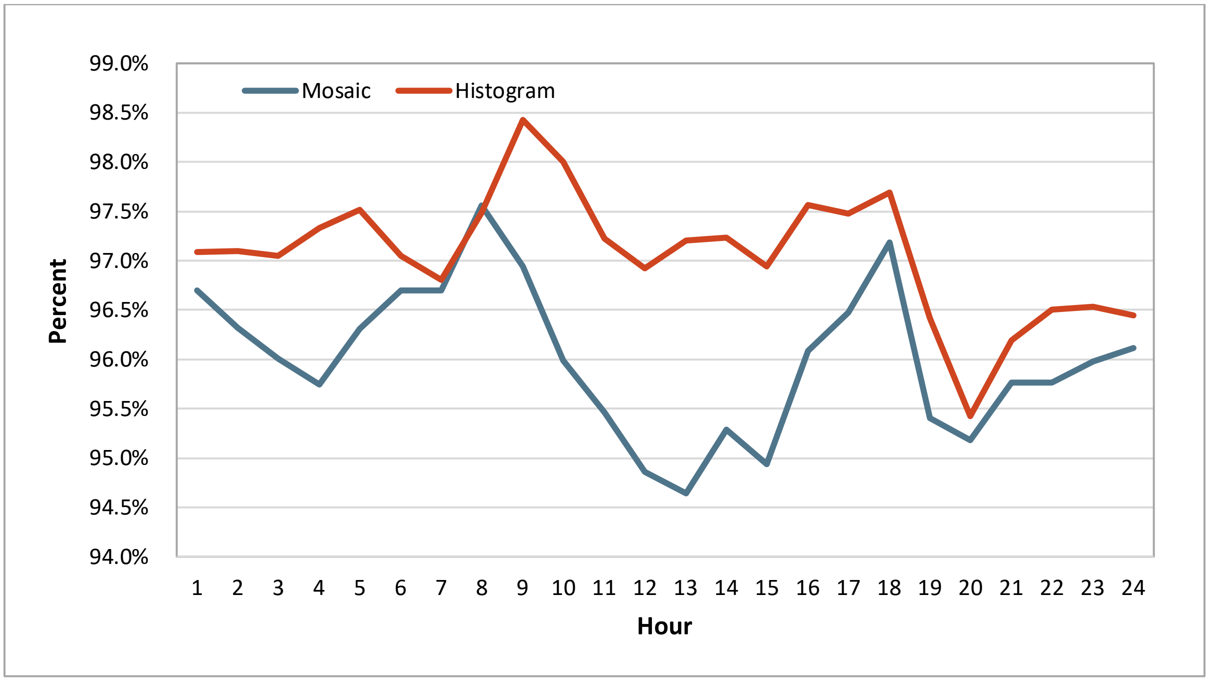
The similarities between the ISO outputs and the random number approach, despite the former using the mosaic variable, suggest that the mosaic variables have a weak relationship with net load uncertainty, particularly during hours when random number outcomes align with the mosaic regression results. This observation aligns with the previously noted 77 percent of non-statistically significant coefficients from DMM's mosaic quantile regression, further emphasizing the potential limitation of using non-significant coefficients in a forecast model.

In summary, a close resemblance between the mosaic quantile regression and histogram outputs, along with the random number regression, implies that the mosaic variable lacks a meaningful relationship with net load uncertainty.

It is noteworthy that the mosaic quantile regression diverges from the histogram and random number approaches between hour-ending 8 and 11. These deviations are meaningful, with approximately

500 MW differences. This indicates quantile regression is capturing a certain relationship with net load uncertainty at the cost of reduced coverage, as illustrated in Figure 5.21.

Figure 5.21 Hourly average coverage rates: histogram vs. mosaic quantile regression (with threshold, February-September 2023)

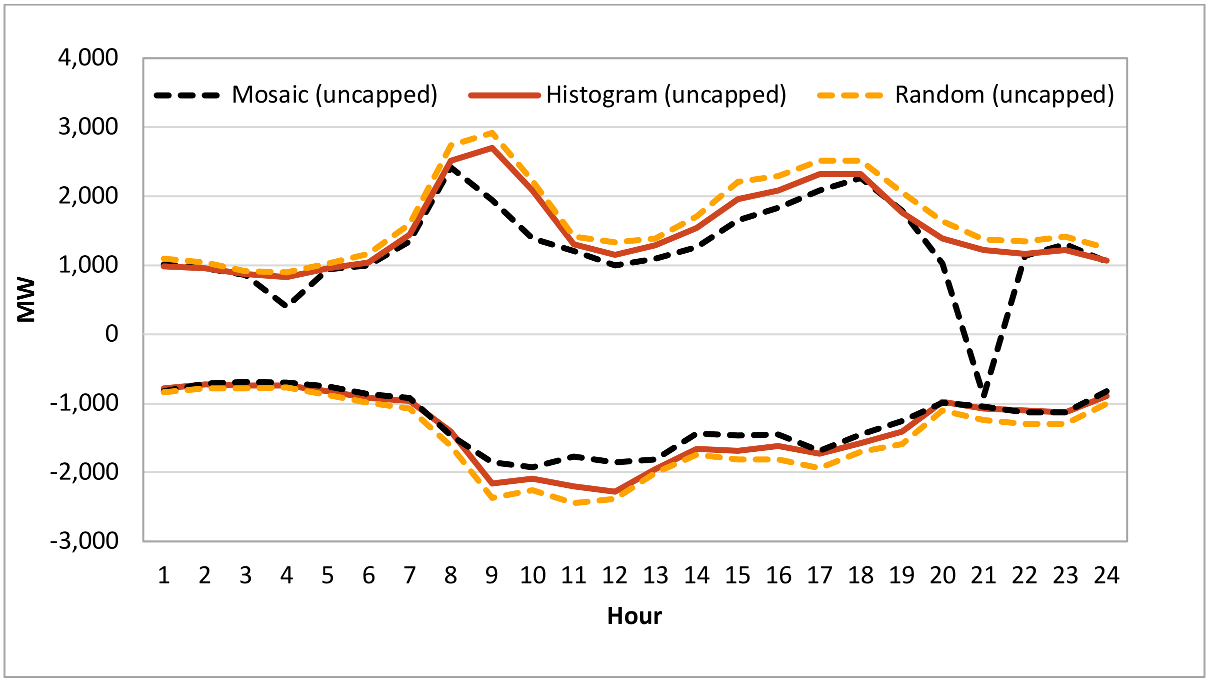


It is also possible that the similarities among mosaic quantile regression, histogram, and random numbers can be attributed to the use of thresholds. In Figure 5.22, each requirement is estimated from these models without the application of thresholds. The black dashed line represents the ISO mosaic quantile regression outcome, the red line represents the histogram, and the yellow dashed lines represent the random number forecast.

Overall, the mosaic regression outcomes without thresholds continue to exhibit close alignment to histogram and random regression outcomes, with the exception of two specific hours: hour-ending 4 and 21. The hour-ending 21 outcome shows that the average requirement for FRU is -900 MW. This negative is a requirement attributed to the mosaic quantile regression generating exceptionally extreme values. Instances where the absolute value of the requirement exceeds 10,000 MW are infrequent, occurring roughly 0.8 percent of the time. However, the magnitude ranges from -870,000 MW to 400,000 MW. These extreme magnitudes can result in a negative requirement for FRU on average.

Figure 5.22 indicates that the threshold may not greatly influence the similarities of results across the mosaic regression, histogram, and random number regression.

Figure 5.22 Uncapped hourly average flexible ramping requirement across different models (February-September 2023)



6 Technical errors in mosaic quantile regression

Misusing constants as an independent variable

The ISO constructs its main predictor, the mosaic variable, by combining a historical percentile of net load, load, wind, and solar uncertainty with predicted values (fitted values) from three quantile regressions. The specific formula is detailed in Equation 8 below. However, this method overcomplicates the estimation by incorporating uncertainty values from the histogram for net load, load, solar, and wind into the mosaic variable construction. Within a sample for each regression, these histogram values remain the same throughout that sample.

Adding these constant histogram values to the mosaic variable does not affect the predicted values from the model. This is because, while the coefficients and scale of the mosaic variable change, the underlying relationship between the mosaic variable and net load uncertainty remains the same.[26]

Coefficient A and B in the mosaic quantile regression are designed to capture the marginal effect of the mosaic variable on net load uncertainty. This marginal effect describes how a one-unit change in mosaic variable impacts net load uncertainty, which is the gradient in this case. The gradient in a regression captures the fundamental relationship between variables. Despite changes in unit measurements, this relationship remains consistent, even if the coefficient's numerical value varies.

Adding a constant term to the mosaic variable via histogram values merely shifts the mosaic variables but does not alter the gradient. The point estimate may shift due to the addition or subtraction of histogram values, as this change alters the unit of measurement for the mosaic variable, but the underlying gradient remains the same. For example, the change from 1 to 2 is the same as the change from 1,001 to 1,002, no matter the scale. The impact of a unit change in the independent variable (the marginal effect) remains consistent, even if the variable values are shifted by a constant, such as the histogram values.[27]

To demonstrate this concept, DMM used an example date and showed that including or excluding the histogram values in the mosaic variable resulted in identical forecast values in the quantile regression. DMM modified the mosaic variable, as described in Equation 9, which did not include the histogram values. Then DMM ran the quantile regression and compared the outcome to the original mosaic variable described in Equation 8.

[26] Adding a constant to the independent variable changes its scale. The range of the variable is shifted. The change in scale may cause the estimated coefficients of the model to change. This can even result in a change in the sign of the coefficients, indicating a reversal in the direction of the relationship between the independent and dependent variables. Despite these changes, adding a constant does not affect the predicted values from the model. For example, add 1,000 MW to the 15-minute solar forecast. If the original forecast was 2,000 MW, in the new model this is considered 3,000 MW. However, the predicted solar uncertainty stays the same in both models because the relationship between the solar forecast and solar uncertainty has not changed, only the scale of the 15-minute forecast variable has shifted.

[27] The coefficient B with and without histogram values can be different from each other. This coefficient indicates the change in net load uncertainty for a unit change in mosaic variable, holding the quadratic mosaic term constant. The coefficient A (the coefficient of quadratic term) remains the same with or without histogram values in the mosaic variable, because the shape of quadratic relationship between uncertainty and mosaic variable does not change with the shift in mosaic variable.

Equation 8 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

Equation 9 Modified mosaic variable for upward net load uncertainty

$$mosaic^{97.5} = \hat{L}_Q^{97.5} - \hat{S}_Q^{2.5} - \widehat{W}_Q^{2.5}$$

Upward mosaic variable: intermediate variable for next step

Predicted values: predicted load, solar, and wind uncertainty from the initial quantile regressions

Figure 6.1 illustrates that the forecasted outcomes (FRU requirement) derived from models using the original mosaic variable and the modified mosaic variable (where histogram values have been removed) were identical. The data sampled for this illustration was from March 31, 2023 during interval 1, for the group of balancing authority areas that have passed the resource sufficiency evaluation. The yellow bars represent the FRU requirement from the original quantile regression, while the green bars signify outcomes from the quantile regression using the modified mosaic variable. The process used for both sets of regression were identical, with the exception of different mosaic variables utilized.

Figure 6.1 Comparison of quantile regression outcomes for upward pass-group uncertainty: original mosaic variable vs. modified mosaic variable (interval 1, March 31, 2023)

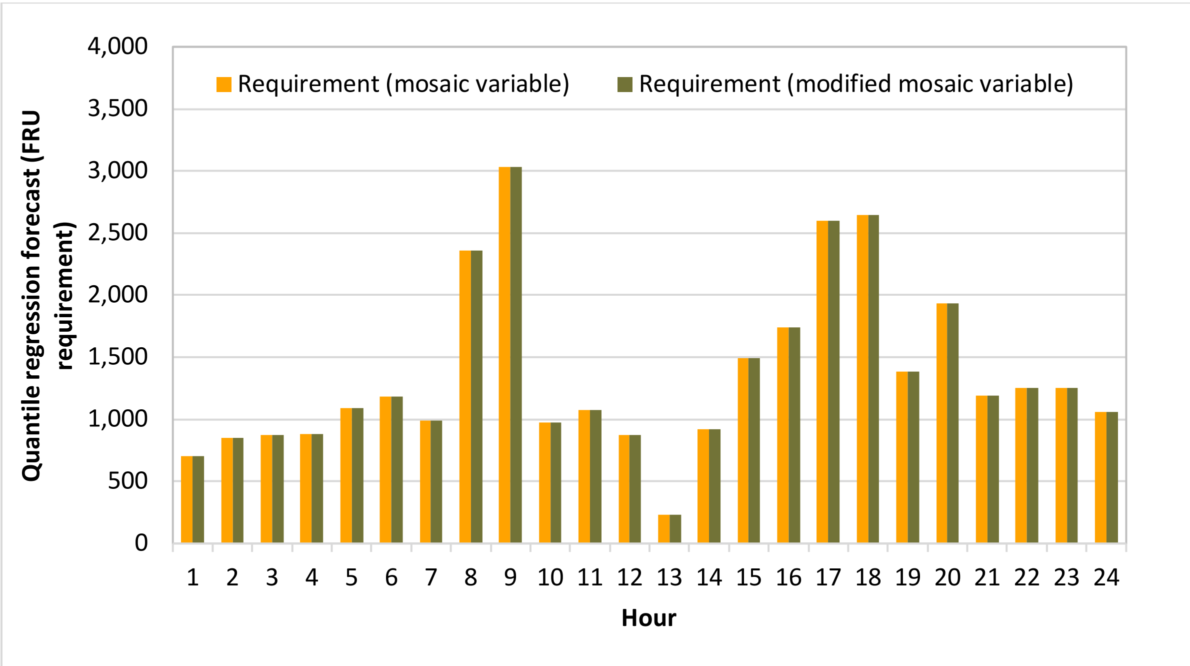


Figure 6.2 and Figure 6.3 demonstrate the impact of excluding histogram values from the mosaic variable on quantile regression. Figure 6.2 plots the original mosaic variable on the horizontal axis and net load uncertainty on the vertical axis, for hour-ending 13 on March 31, 2023 for the group that passed the resource sufficiency evaluation. The green line represents DMM's fitted values from the quantile regression. Figure 6.3 presents the same information but uses the modified mosaic variable (excluding histogram values).

These two figures show that adding a constant number to the independent (mosaic) variable results in a horizontal shift of the parabolic relationship between the mosaic variable and uncertainty. The shape of the parabola did not change, just its position along the horizontal axis. This means that the predicted values from the model are the same before and after the shift, resulting in the same forecasts.

Figure 6.2 Original mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 13, March 31, 2023)

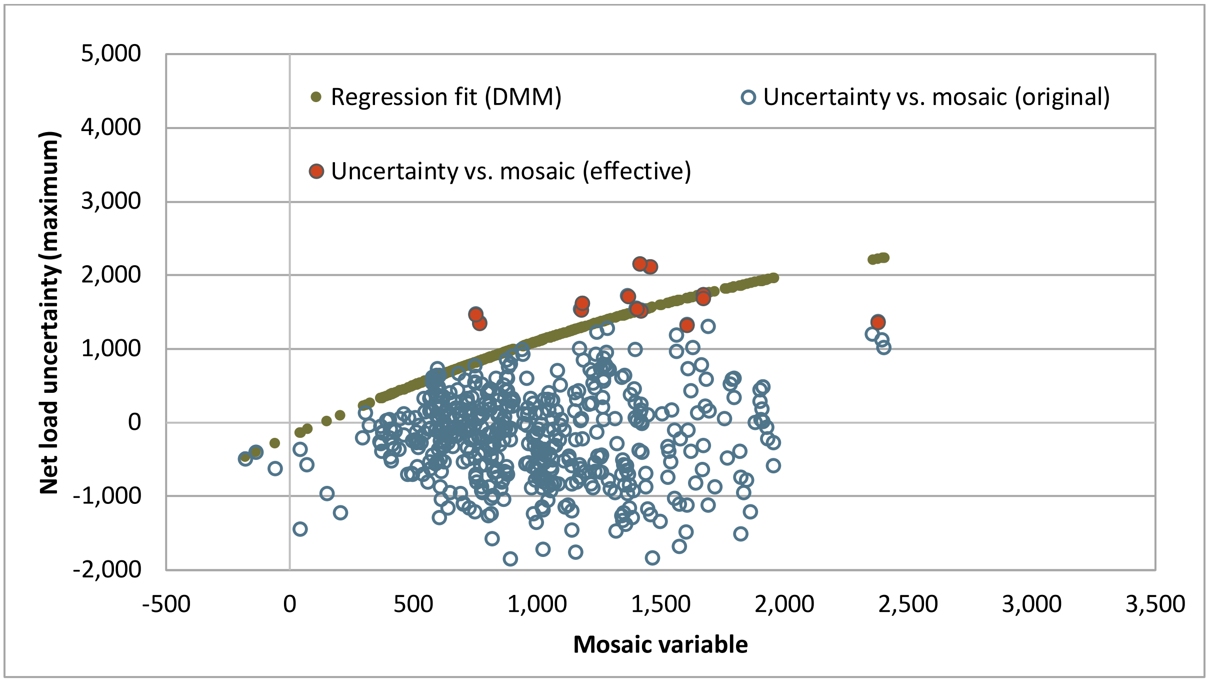
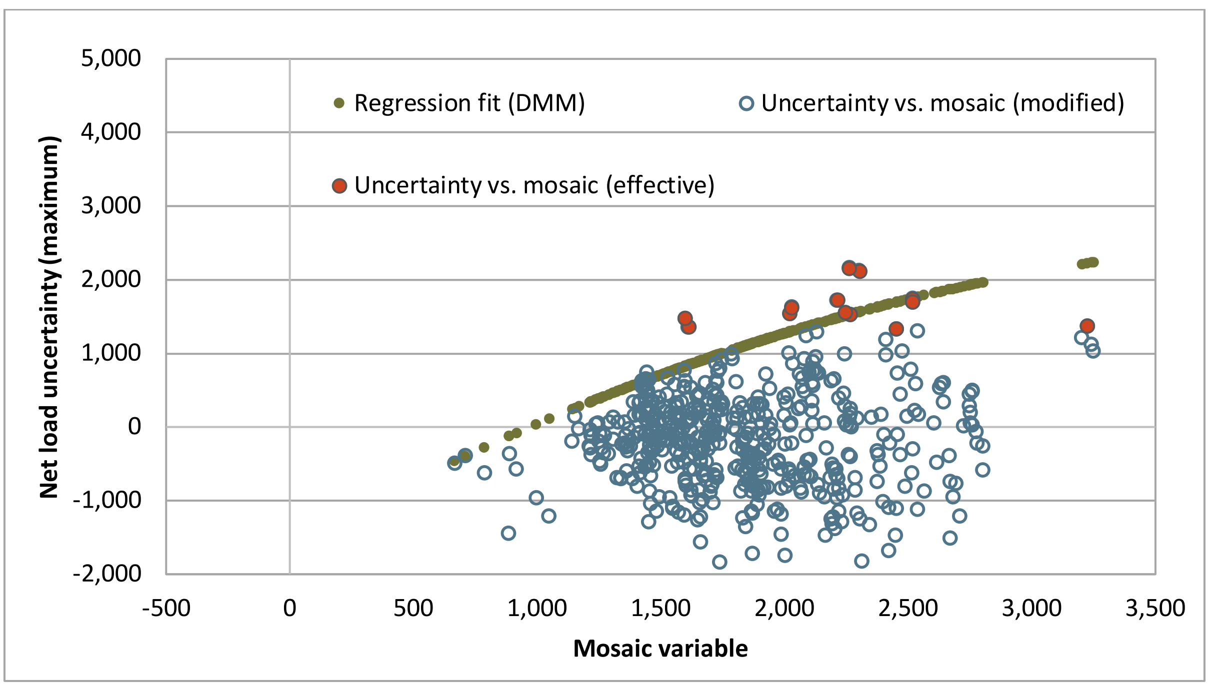


Figure 6.3 Modified mosaic quantile regression fit for upward pass-group uncertainty (hour-ending 13, March 31, 2023)



While the modified mosaic variable does not enhance forecasting power, it does simplify the construction process by eliminating unnecessary complexity. If the values that the ISO added to the mosaic variable exhibited variation within the sample, such an adjustment might affect the forecast outcomes. However, because these histogram values remain constant, adding them to the mosaic variable does not change the result, which is a basic characteristic of linear regression.

Erroneous linear transformation of independent variable in dependent variable construction

In this section, DMM identifies an error in the design of the initial quantile regression equation, which could potentially influence the forecasting power. Equation 10 illustrates the initial quantile regression for load, wind, and solar, which is utilized to construct the mosaic variable. The dependent variable is defined as load, solar, and wind uncertainty, measured by the difference between the binding 5-minute market forecast and the advisory 15-minute market forecast. The independent variables, on the other hand, are the advisory 15-minute market forecasts for load, solar, and wind.

Equation 11, which is a simplified version, illustrates that the independent variable X (the advisory 15-minute forecast) is present on both the left and right sides of the equation. This means that the dependent variable in the quantile regression is a linear transformation of the independent variable, an approach that is unconventional and could lead to inaccurate predictions or misleading conclusions.

These initial quantile regressions for load, solar, and wind included a misspecification issue due to the overlap of 15-minute forecasts in both dependent and independent variables. While the initial goal was to examine the relationship between load, solar, and wind uncertainties, and the corresponding 15-minute forecasts, the model's setup led to a misspecification. The problem stemmed from the inclusion of 15-minute forecasts in both the dependent (uncertainty) and independent variables.

Consequently, the coefficient "b", which was meant to capture the intended relationship, ended up representing only the correlation between the 5-minute and 15-minute forecasts, deviating from the ISO intended relationship between actual outcomes and the 15-minute forecast. Simply put, coefficient b, in this scenario, is estimating the impact of the independent variable (X) on the difference between Y and X. However, this setup means it is actually estimating the change in Y for a unit change in X, minus 1, as any change in X will directly impact X itself by an equal amount.[28]

Equation 12 shows the discrepancy between the ISO initial intention and the actual outcome in the load uncertainty equations, and the same logic applies to solar and wind equations. The ISO pulls out the predicted values from this equation and uses it in construction of the mosaic variable. The predicted value is the portion of load uncertainty variation that can be attributed to the 15-minute forecast (denoted as X). Essentially, the load uncertainty fluctuates daily for various reasons. The objective of the ISO analysis is to determine what proportion of this variation in load uncertainty is explained by the variation in the 15-minute load forecast.

However, due to the misspecification, the model does not extract information about load uncertainty variation, as initially intended. Instead, it ends up revealing the variation in the 5-minute forecast as explained by the 15-minute forecast.

[28] The mathematical approach:

$$coefficient\ b = \frac{\partial(Y - X)}{\partial X} = \frac{\partial Y}{\partial X} - \frac{\partial X}{\partial X} = \frac{\partial Y}{\partial X} - 1$$

The usefulness of this mistakenly extracted 5-minute forecast information for net load forecasting is a subject for further discussion. A basic perspective is that net load uncertainty represents the discrepancy between the 5-minute and 15-minute net load forecasts. In this context, two types of information can be considered: one that is common to both 5-minute and 15-minute forecast, and another that is unique to the 5-minute forecast but absent in the 15-minute forecasts.

The first type of information is derived from the predicted values of the misspecified equation. The second type of information is not captured in the misspecified equation. In fact, this information is precisely what is eliminated by the equation. To obtain this information, one needs to look at the residuals of the regression.

Equation 10 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: 97.5 percentile, low: 2.5 percentile)

Equation 11 Simplified initial quantile regression

$$Y - X = aX^2 + bX + c + \varepsilon$$

Y: Binding 5-minute forecast

X: Advisory 15-minute forecast

Equation 12 Initial intentions vs. actual applications of quantile regression in coefficient b estimation

The ISO intention: Load uncertainty $= X\beta + e$

What is actually running: 5-minute load forecast $= X\gamma + \epsilon$

where Load uncertainty = 5-minute load forecast – 15-minute load forecast

The consequence of this design error is non-negligible. Instead of capturing the linear correlation between the extreme end of uncertainty and 15-minute forecast as intended, the model primarily estimates the correlation between the extreme end of the 5-minute and 15-minute forecast and subtracts one from this estimation.

Figure 6.4 provides empirical evidence that the coefficient b is derived from the correlation between the 5-minute and 15-minute forecasts, rather than from uncertainty. The figure presents the coefficients from two distinct regression models. The two models in discussion are derived from Equation 10 with an exception of excluding the quadratic term of the independent variable. The data set used for these regressions include groups of balancing authority areas which passed resource sufficiency evaluation on March 31, 2023.

In Figure 6.4, the green line represents coefficient b in the model with load uncertainty as the dependent variable. The blue line denotes the coefficient b from the same model, but with binding 5-minute market load forecast as the dependent variable. The independent variables are advisory 15-minute market load forecast for both models.

The figure shows that coefficient b in the model with load uncertainty is equivalent to coefficient b in the 5-minute forecast model, minus one. During this example day, across all hours, the difference between the two coefficients consistently maintains a value of 1, in accordance with the mathematical framework explained earlier. In other words, the primary source of variation in coefficient b is from the model represented by the blue line (5-minute forecast as the dependent variable). The model represented by the green line (load uncertainty as the dependent variable) merely shows this variation with a constant of 1 subtracted.[29]

[29] The coefficient b in Equation 11 captures the change in the dependent variable (Y-X) for a unit change in the independent variable (X). However, it is important to note that in this unique setup, X is not only the independent variable but also a part of the dependent variable (Y-X). When X changes, Y-X also changes, but not because X has an inherent effect on Y-X. Instead, Y-X changes because Y, which is not a part of the independent variables, is changing. This is what the regression model is capturing. The coefficient b in this context captures the variation in Y for a unit change in X, minus the variation in X itself. The variation in X is accounted for by subtracting 1, which is why the difference in the coefficients between the model with dependent variable as 'load uncertainty' and the model with dependent variable as '5-minute forecast' consistently comes out to be 1.

Figure 6.4 Comparing coefficient b for models without quadratic terms including different dependent variables: load uncertainty vs. binding 5-minute load forecast (March 31, 2023, RSE pass-group)

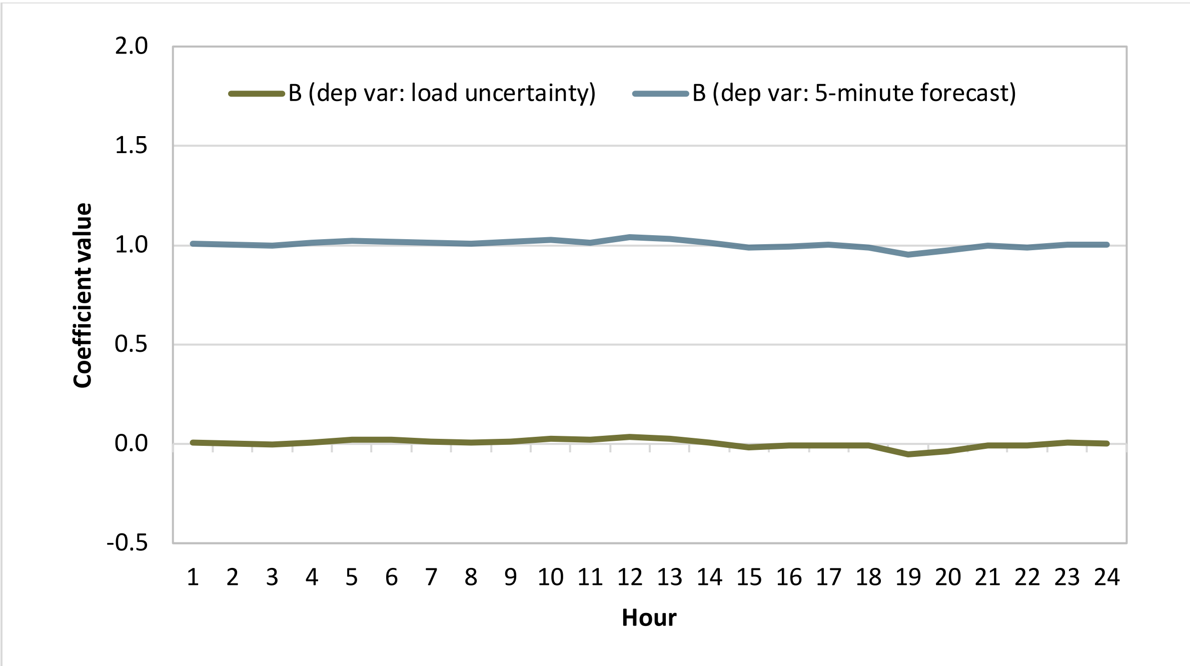
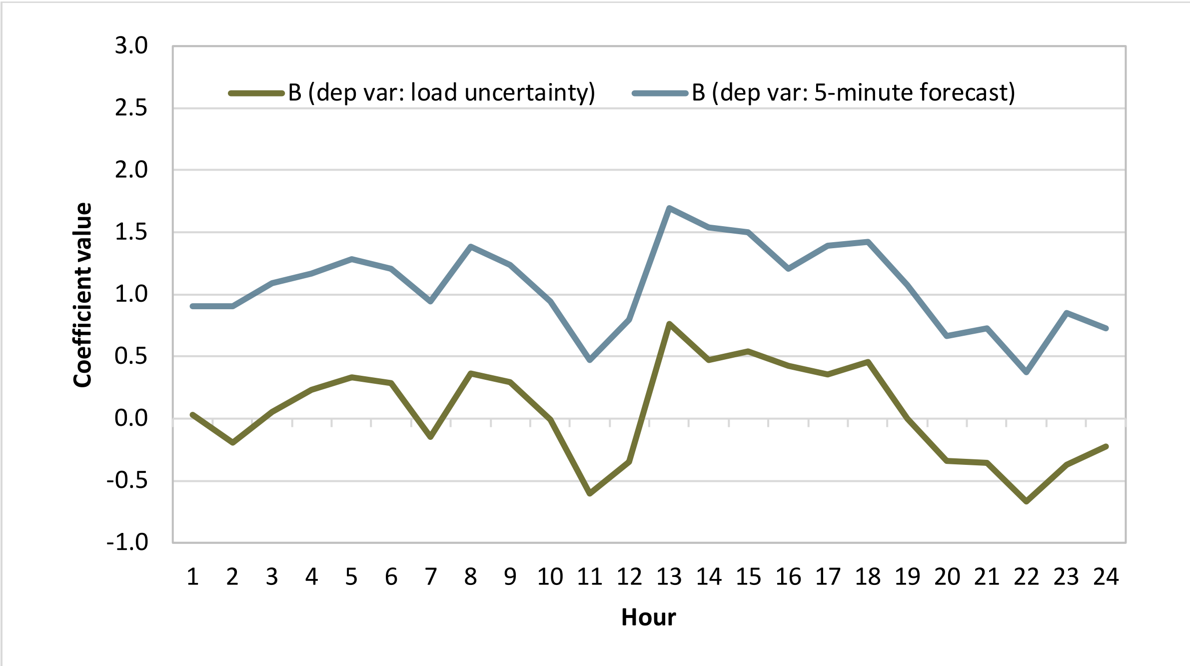


Figure 6.5 presents the comparison of coefficient b across two different models. In this case, both models include the quadratic term in Equation 10. The sample remains consistent with Figure 6.4, using data from March 31 for resource sufficiency passed groups. The blue line signifies coefficient b derived from the model using the binding 5-minute advisory forecast as the dependent variable, whereas the green line corresponds to the model where the dependent variable is load uncertainty.

This figure also shows that a consistent difference in coefficient b between the two models is around 1. The reason why the difference between the coefficients is not always exactly 1 is due to the inclusion of the quadratic terms in both models. The quadratic terms introduce a non-linear effect, which adds complexity to the relationship between the dependent and independent variables. This added complexity can create some deviations from the expected difference of 1 in the coefficients.

However, it is important to note that even though these deviations exist, the primary source of variation in the model with load uncertainty as the dependent variable (green line) still comes from the model with the 5-minute forecast as the dependent variable (blue line). The consistent difference between the coefficients of these two models implies a strong link between these dependent variables, namely the binding 5-minute load forecast and load uncertainty. These variables inherently exhibit distinct trends. However, due to the specifics of the regression model, the trend appears aligned. This consistently close gap reinforces the evidence that coefficient b in both models primarily reflects the marginal effect of the independent variable on the 5-minute forecast, rather than the uncertainty.

Figure 6.5 Comparing coefficient b for models with quadratic terms including different dependent variables: load uncertainty vs. binding 5-minute load forecast (March 31, 2023 RSE pass-group)



The same design flaw exists not only for load uncertainty, but also for solar and wind uncertainties, since the same methodology was used to construct these variables.

This method, which involves the independent variable in creating the dependent variable through a linear transformation, diverges from basic linear regression principles. It is a methodological issue that should be resolved for accurate forecasting. A simple correction would be to offset the independent variable by an interval or hour. The crux of the issue is that the 15-minute market forecast data should not be identical on both the independent and dependent variable sides.

It is unclear to what extent this error influences the ultimate forecasting power, particularly when given the overall low performance of the mosaic quantile regression. However, the coefficient b, by design, is unrelated to uncertainty, which likely undermines the effectiveness of uncertainty forecasting.

7 Forecasting across multiple definitions of net load uncertainty

This section presents the distribution of various forms of uncertainty across different markets. The term 'uncertainty' refers to the discrepancies in net load forecasts employed in the day-ahead and real-time markets. Despite being collectively termed 'uncertainty', these variabilities do not necessarily possess similar properties.

The analysis reveals distinct characteristics in the distribution shapes of these uncertainties. These differing characteristics may necessitate unique forecasting strategies tailored to each market.

This section discusses four illustrative type of uncertainties using net load forecasts for the ISO balancing area. These include: 1) variations between day-ahead and binding 5-minute interval forecasts, 2) differences between day-ahead and advisory 15-minute forecasts, 3) differences between binding 5-minute and advisory 15-minute forecasts, which is the current 15-minute FRP uncertainty, and 4) differences between binding 5-minute and advisory 5-minute forecasts, representing the current 5-minute FRP uncertainty.

Figure 7.1 and Figure 7.2 display the distribution of uncertainty based on the four different definitions. The vertical axis represents the probability density function (PDF), which shows the likelihood of different values occurring within the dataset. The horizontal axis shows the values of net load uncertainty, normalized using the mean and standard deviation.[30] The 5-minute uncertainty PDF (yellow line) displays an unusual pattern with its high peak and pronounced presence of extreme values (long tails). This high peak indicates a strong concentration of data around the average, yet there is a noticeable occurrence of values significantly deviating from the norm. This unusual combination suggests that while the majority of 5-minute uncertainty are closely clustered, there is also substantial likelihood of encountering extremely high or low values, pointing to potential spikes or drops in the 5-minute uncertainty.

On the other hand, the rest of uncertainty PDFs exhibit a low peak, along with a more even and balanced distribution of values. This indicates a more uniform spread of data without a strong tendency to cluster around the mean. The relatively flat appearance of these PDFs is an unusual pattern, reflecting a dataset with fewer extremes or outliers and suggesting a level of predictability and stability in the values, in stark contrast to the more volatile and unpredictable 5-minute uncertainty.

[30] This normalization process converts the data into z-scores. A z-score indicates how many standard deviations a value is from the mean. It helps compare data points by showing how much each value deviates from the average, considering the spread of the entire dataset.

Figure 7.1 Distribution of uncertainty (California ISO, upward, February-September 2023)

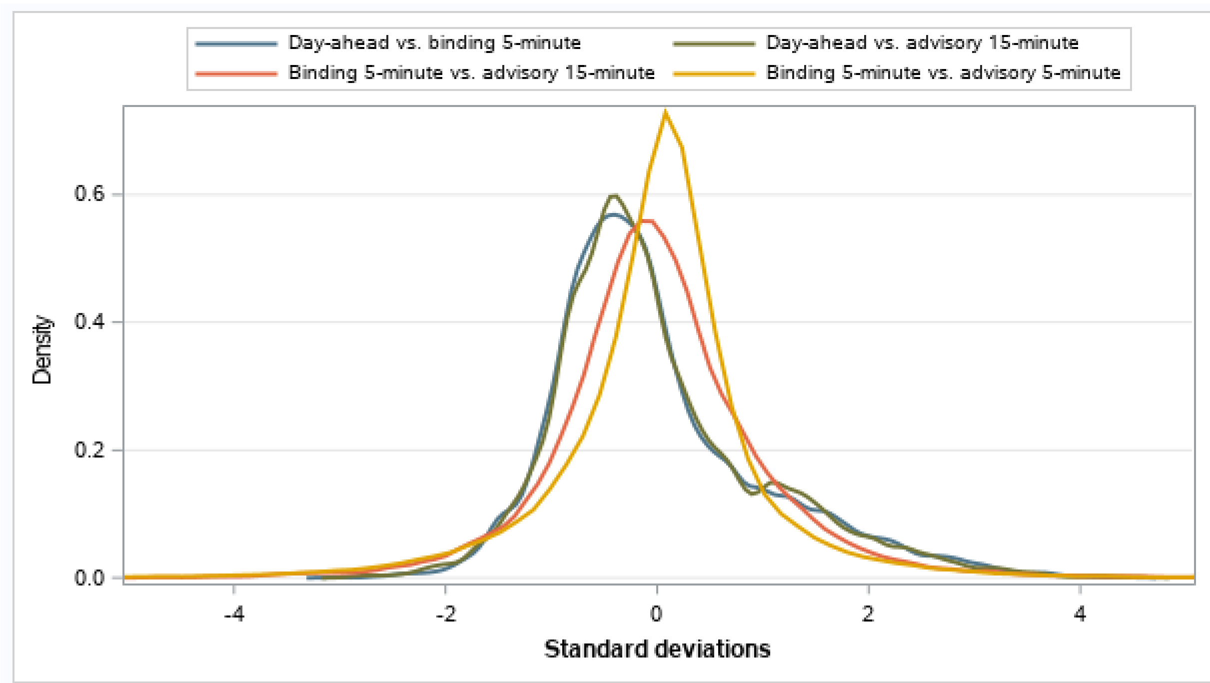


Figure 7.2 Distribution of uncertainty (California ISO, downward, February-September 2023)

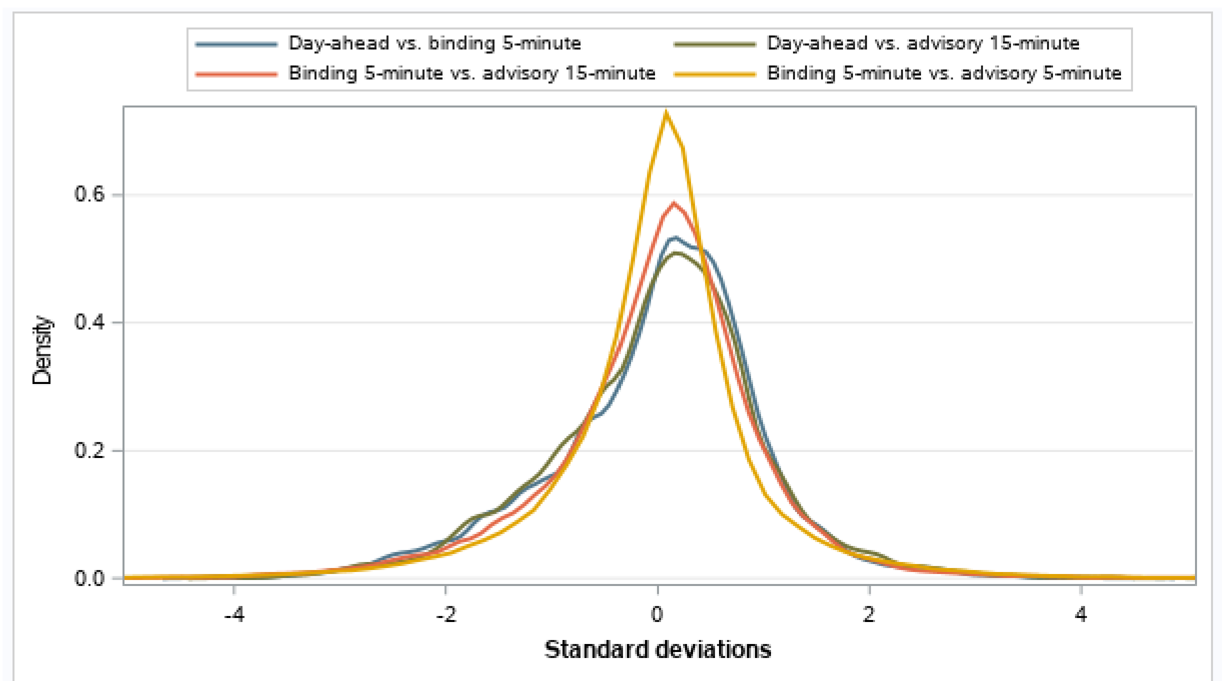


Table 7.1 Characteristics of the distribution of uncertainty (California ISO, February-September 2023)

Direction	Uncertainty type	Mean	STD	Skewness	Kurtosis
Upward	Day-ahead vs. binding 5-minute	1,063	1,395	1.0	1.1
	Day-ahead vs. advisory 15-minute	1,034	1,330	1.0	1.2
	Binding 5-minute vs. advisory 15-minute	24	509	0.3	3.5
	Binding 5-minute vs. advisory 5-minute	-8	105	-0.3	6.4
Downward	Day-ahead vs. binding 5-minute	-752	1,241	-0.4	1.3
	Day-ahead vs. advisory 15-minute	-437	1,074	-0.2	1.1
	Binding 5-minute vs. advisory 15-minute	-314	551	-0.3	3.1
	Binding 5-minute vs. advisory 5-minute	-8	105	-0.3	6.4

The main distinction between real-time and day-ahead uncertainty lies in the distribution's tail thinness. A normal distribution has a kurtosis of 3. In contrast, day-ahead uncertainty, including comparisons between day-ahead and 5-minute, as well as day-ahead and 15-minute net load uncertainty, shows a kurtosis below 3, as shown in Table 7.1. This indicates that the distribution lacks thin tails, meaning there are fewer extreme values in net load uncertainty. This suggests that day-ahead net load uncertainty is more stable than real-time uncertainty.

On the other hand, the 5-minute uncertainty, which compares binding versus advisory 5-minute net load, has a kurtosis of 6.4. This suggests a higher likelihood of observing extreme values in the 5-minute net load data, with more frequent and significant deviations from the average.

The skewness of a distribution is a measure of its asymmetry. A positive skewness indicates that the tail on the right side of the distribution is longer or thicker than the left side. The upward day-ahead uncertainty skewness of 1 indicates a moderately positively skewed distribution. This suggests that most values are concentrated on the lower end, but there are some relatively high values stretching out to the right. It is possible that there are occasional instances of significant overestimation.

In summary, the four net load uncertainties exhibit distinct distribution characteristics. The day-ahead uncertainties display a thin-tail distribution, suggesting the use of models that emphasize recent trends and anomalies for accurate forecasting. The current 15-minute uncertainty, showing more normal distribution, calls for more traditional statistical methods. Lastly, the 5-minute forecast, characterized by long tails, requires robust models that can account for extreme values and sudden shifts.

8 DMM Recommendations

DMM included a recommendation on the quantile regression in its 2022 annual report:

> In February 2023, the ISO implemented a new method of net load uncertainty calculation based on quantile regression for the flexible ramping product. DMM's review of the performance of this new methodology indicates that it is not a clear improvement over the prior method. Although uncertainty values calculated with this method are generally lower while covering uncertainty (an improvement), they fluctuate more significantly and are likely to be more difficult for balancing areas to reproduce or predict in advance.
>
> Therefore, DMM continues to recommend that the ISO and stakeholders consider developing much simpler and more transparent uncertainty adders in the next phase of this initiative, and consider adopting uncertainty calculations customized to the resource sufficiency evaluation, rather than using the flexible ramping product uncertainty calculation.

Based on the analysis included in this report, DMM offers the following additional recommendations for refinement of the net load uncertainty calculation:

- **DMM recommends that the ISO develop separate models of net load uncertainty for each of the five current and potential future market features in which the ISO incorporates uncertainty.** These include: 1) the flexible ramping product, 2) the resource sufficiency evaluation, 3) the extended flexible ramping product, 4) imbalance reserve up/down, and 5) the uncertainty captured in operator adjustment to the residual unit commitment demand. Each of these market features involve different applications, forecasting horizons and timelines of net load uncertainty.

- **The infrastructure developed to support the mosaic quantile regression can be adapted to support other model formulations.** A thorough review of other options would include an evaluation of different sampling methods, independent variable selection and functional form. DMM has begun evaluating potential refinements including: 1) pooling estimation across hours, 2) using a more conventional forecasting approach than the quantile regression, and 3) replacing the multi-stage mosaic variable regression with a regression on net load uncertainty directly. These results will be published in a future report.

- **DMM recommends that the ISO consider reverting to the histogram approach for the current flexible ramping product and the resource sufficiency evaluation** while an improved model or approach is developed. This would provide requirements that capture a high level of uncertainty equivalent to that of the quantile regression model, but in a much more predictable way for current market participants. This could also reduce the additional processing time that the quantile regression approach adds to the real-time market.

- **If the ISO retains the existing mosaic quantile regression, the limited sample size and technical errors in the formulation should be resolved.** DMM suggests that the histogram component of net load, load, solar, and wind could be removed from the construction of the mosaic variable. Model results remained the same without these components. To address the misspecification issue in the initial quantile regression, the dependent variables could remain load, solar, and wind uncertainty, but the independent variable should be adjusted by using one-interval or one-hour lagged 15-minute forecasts. This modification prevents the overlap of the same 15-minute forecast in both the right- and left-hand sides of the regression equation. DMM recommends extending the sample period beyond 180 days, but with appropriate seasonal controls for load forecasts. Unlike renewable forecasts, which are currently adjusted based on installed capacity, load forecasts lack an equivalent method. DMM observed that during the summer months, the current load forecast stands out as an outlier compared to the load forecasts in the sample, resulting in extremely large forecasts.

- **There is value in having a good approach for estimating the different probabilities of different net load realizations for market products that are procured based on a demand curve.** These three products include: 1) the current flexible ramping product, 2) the proposed day-ahead imbalance reserve up, and 3) the extended flexible ramping products. The demand curve for each of these products is explicitly based on the probabilities of different net load realizations, combined with the estimated costs of different levels of procurement.

- **DMM recommends that a much more simplified approach be considered for incorporating uncertainty into resource sufficiency evaluations**. This includes the current WEIM resource sufficiency test, as well as the resource sufficiency tests for the Extended Day-Ahead Market (EDAM). These requirements used in these tests are not based on a demand curve or any other specific reliability standard. Instead, the uncertainty component of these tests is more akin to a reserve capacity margin that is agreed upon by all balancing areas. Thus, DMM has recommended that a different – and much more simplified – approach be considered for setting the uncertainty component of these tests.

- **For resource sufficiency evaluations, DMM suggests that requirements could be based on a very simplified approach based on the amount of load, wind, and solar each hour within each balancing area.** This type of simplified approach would allow balancing areas to know what their resource sufficiency evaluation requirements would be, and plan accordingly to meet these requirements. This would avoid the problems created by the very high variability and uncertainty about resource sufficiency evaluation requirements that participants face under the current quantile regression approach.