Flexible Ramping Product Requirements: Performance review, challenges, and potential enhancement analysis/paths discussion

Amber Motley
Short Term Forecasting

Market Surveillance Committee Meeting
General Session
November 29, 2023
Recent concerns on FRP requirements

- Limited Sample Size
- Historical Data Utilized
- Future Awareness of Requirements
- Movement in requirement between time periods
- Mosaic requirement outcomes weekends
- Threshold Activations

How to analyze requirement performance
Performance metrics mapped to key items

Current Performance Metrics:
- Directional Coverage
- Average Requirement
- *Inter-hour movement
- *RSE requirement movement

Key Items:
- Quality of calibration
- Informativeness
- Cost
- Usability

* denotes new requirement for presentation
bold denotes proposed grouping by MSC

How to analyze requirement performance
Presentation Roadmap

- PRODUCTION RESULTS
  - FRP

- SIMULATION RESULTS
  - FRP

- FUTURE CONSIDERATIONS
  - FRP
  - IBR
PRODUCTION RESULTS

• FRP
Informativeness

(1) Perfect predictor

(2) Dynamic predictor

(3) Static predictor

Equivalent coverage
Difference in informativeness
Mosaic requirements are responsive to realized uncertainty

Instances where uncertainty exceeded mosaic and/or histogram requirement (summer 2023, all BAAs)

1) Mosaic carries low requirement at low uncertainty
2) Mosaic shows coverage closer to target at high uncertainty
Mosaic better captures realized uncertainty

This slide pares down the data visualization from the previous slide and shows spread during periods where uncertainty is near requirements.

Performance benefits of mosaic are most clearly seen at the extremes (low and high) of uncertainty.

<table>
<thead>
<tr>
<th>Spread (10th-90th)</th>
<th>Lowest 20% [0-701 MW]</th>
<th>Middle 20% [981–1525 MW]</th>
<th>Top 20% [2026–3460 MW]</th>
</tr>
</thead>
<tbody>
<tr>
<td>mosaic</td>
<td>642.7</td>
<td>753.9</td>
<td>685.9</td>
</tr>
<tr>
<td>histogram</td>
<td>873.5</td>
<td>839.5</td>
<td>631.7</td>
</tr>
</tbody>
</table>

Informativeness

Both plots draw from instances where uncertainty exceeds the requirement provided by either mosaic and/or histogram.

1) Mosaic carries low requirement at low uncertainty
2) Mosaic shows coverage closer to target at high uncertainty

Per quantile, the solid colored line identifies the median and shading from lightest to darkest opacity shows spreads from 10th-90th, 20th-80th, and 30th-70th.
Mosaic leads to lower requirements

Mosaic method has lower average and median requirement magnitudes. Historical production data shows about a 100 MW reduction in FRU and 10 to 50 MW in FRD.

<table>
<thead>
<tr>
<th></th>
<th>Mosaic</th>
<th>Histogram</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FRU</strong> Mean</td>
<td>1040</td>
<td>1141</td>
</tr>
<tr>
<td>Median</td>
<td>874</td>
<td>996</td>
</tr>
<tr>
<td><strong>FRD</strong> Mean</td>
<td>-801</td>
<td>-868</td>
</tr>
<tr>
<td>Median</td>
<td>-694</td>
<td>-707</td>
</tr>
</tbody>
</table>

Distributions shown for production period in CISO BAA.
Mosaic coverage is greater on weekdays than weekends

<table>
<thead>
<tr>
<th>Coverage</th>
<th>Weekday only</th>
<th>Weekend only</th>
<th>Weekday/Weekend combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>Mosaic</td>
<td>0.931</td>
<td>&gt;</td>
</tr>
<tr>
<td></td>
<td>Histogram</td>
<td>0.942</td>
<td></td>
</tr>
<tr>
<td>FRU</td>
<td>Mosaic</td>
<td>0.966</td>
<td>&gt;</td>
</tr>
<tr>
<td></td>
<td>Histogram</td>
<td>0.973</td>
<td></td>
</tr>
<tr>
<td>FRD</td>
<td>Mosaic</td>
<td>0.965</td>
<td>&gt;</td>
</tr>
<tr>
<td></td>
<td>Histogram</td>
<td>0.973</td>
<td></td>
</tr>
</tbody>
</table>

- Coverage is evaluated over the historical period of analysis (2/2/23 to 10/18/23).
- Note that coverage is one characteristic of many and the above metrics offer comparison of day-type performance rather than mosaic vs histogram performance.

Weekend coverage suffers from smaller sample size due to day-type split.
Quantification of inter-hour movement in RTPD

(1) Mosaic movement relative to histogram is greater in FRD than in FRU and (2) movement in observed uncertainty is greater than either method, which helps validate the need for this movement.

<table>
<thead>
<tr>
<th>Inter-hour absolute magnitude changes</th>
<th>Mean</th>
<th>Median</th>
<th>90th percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>FRU</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mosaic</td>
<td>277.8</td>
<td>175.9</td>
<td>660.2</td>
</tr>
<tr>
<td>Histogram</td>
<td>240.0</td>
<td>149.1</td>
<td>587.3</td>
</tr>
<tr>
<td>Histogram \to mosaic % change</td>
<td>15.8%</td>
<td>18.0%</td>
<td>12.4%</td>
</tr>
<tr>
<td>FRD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mosaic</td>
<td>222.1</td>
<td>155.5</td>
<td>508.2</td>
</tr>
<tr>
<td>Histogram</td>
<td>162.2</td>
<td>107.3</td>
<td>396.0</td>
</tr>
<tr>
<td>Histogram \to mosaic % change</td>
<td>36.3%</td>
<td>44.9%</td>
<td>28.5%</td>
</tr>
<tr>
<td>Observed uncertainty</td>
<td>414.2</td>
<td>297.8</td>
<td>931.8</td>
</tr>
</tbody>
</table>

![Graph showing inter-hour requirement changes for different methods](image1)

![Graph showing inter-hour changes in [MW]](image2)
Dynamic thresholds comprise the bulk of threshold activity

<table>
<thead>
<tr>
<th></th>
<th>Summer (6/1 – 9/30)</th>
<th>Overall (2/2 – 10/18)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Zero</td>
<td>Static</td>
</tr>
<tr>
<td>CISO</td>
<td>0.5%</td>
<td>1.7%</td>
</tr>
<tr>
<td>EIM Area</td>
<td>0.7%</td>
<td>0.6%</td>
</tr>
</tbody>
</table>

% of requirements hitting thresholds, evaluated per day and averaged over a defined period (# hits / 288)

FRU (CISO)

FRD (CISO)
Top priorities of cause to symptom: FRP requirements from a forecasting perspective

- Limited Sample Size
- Historical Data Utilized
- Future Awareness of Requirements
- Movement in requirement between time periods
- Mosaic requirement outcomes weekends

Symptom
SIMULATION RESULTS

- FRP
STF Analysis Plan

1) Define Objective [function]
   – From Benchmark, increase calibration towards target coverage and decrease cost
   – From Benchmark, increase calibration towards target coverage and increase informativeness.

2) Select the parameters to modify/add/subtract
   – Pooled suggestions from DMM/MSC and are pursuing in merit order

3) Choose a Search Strategy
   – Expand grid space / results as they increase objective
Simulations focused on two key areas with respect to sample size:
  – Day Type Consideration
  – Historical Data Utilized

CAISO will use a subset of metrics to approximate the Pareto optimization of calibration (through coverage) vs. cost (through average requirement) as primary driver of decisions (objective)

Proposed alternatives are performed in parameter “sets”
### Parameter sets to evaluate alternative methods

Vary the sample scheme and size used to determine mosaic and histogram requirements

<table>
<thead>
<tr>
<th>Sample Scheme</th>
<th>Sample Days</th>
<th>Description</th>
<th>Effective Sample Days Weekday</th>
<th>Effective Sample Days Weekend</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>90</td>
<td>Backwards 90 days w/ no day-type consideration</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>1</td>
<td>180</td>
<td>Backwards 180 days w/ no day-type consideration</td>
<td>180</td>
<td>180</td>
</tr>
<tr>
<td>2</td>
<td>90</td>
<td>Backwards 90 days w/ day-type consideration</td>
<td>~64</td>
<td>~25</td>
</tr>
<tr>
<td>2</td>
<td>180</td>
<td>Backwards 180 days w/ day-type consideration</td>
<td>~128</td>
<td>~51</td>
</tr>
<tr>
<td>3</td>
<td>90</td>
<td>Backwards 45 days / Forward 45 days w/ day-type consideration</td>
<td>~64</td>
<td>~25</td>
</tr>
<tr>
<td>3</td>
<td>180</td>
<td>Backwards 90 days / Forward 90 days w/ day-type consideration</td>
<td>~128</td>
<td>~51</td>
</tr>
<tr>
<td>4</td>
<td>90</td>
<td>Backwards 45 days / Forward 45 days w/ no day-type consideration</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>4</td>
<td>180</td>
<td>Backwards 90 days / Forward 90 days w/ no day-type consideration</td>
<td>180</td>
<td>180</td>
</tr>
</tbody>
</table>

**Current Scheme**

- Sample days (90 vs. 180)
- Sample scheme (backwards-only vs. backwards/forwards)
- Day-type split (y/n)
Interpreting simulation results

More Reliability
+ Less Requirement

Less Reliability
+ More Requirement

<table>
<thead>
<tr>
<th>Requirement [MW]</th>
<th>Percent Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline (current scheme)</td>
<td>97.00%</td>
</tr>
<tr>
<td>Target coverage</td>
<td>97.50%</td>
</tr>
<tr>
<td>More Reliability</td>
<td>98.00%</td>
</tr>
</tbody>
</table>

method
- histogram
- mosaic

sample size
- 90
- 180
Pareto Optimization of Calibration and Cost (FRU/WEEKEND)

FRU Requirement vs. FRU Coverage All Hours Weekend (Period Mean)
mosaic w/threshold (blue) vs. histogram (red) | sample size (ascending)
production benchmark (black-dashed) | target coverage (gold-dashed)

Current

Strong improvement for multiple alternative models

2_backwards_w_daytype 3_forwards_backwards_w_daytype 4_forwards_backwards_all

as.factor(SAMPLE_DAYS)

- 60
- 70
- 80
- 100
- 120
- 140
- 160
- 180
- 190
- 200
- 210

variable

- nh_up
- current_threshold

2023-5-23 - 2023-11-5
SUMMARY AND NEXT STEPS

Future Considerations

• FRP
Summary

Production Results

• Positives
  – Mosaic shows significant benefit over histogram at the tails of realized uncertainty.
  – Mosaic has a smoother distribution of requirements and a lower cost/requirement compared to histogram.

• Negatives
  – Mosaic has worse performance on weekends compared to weekdays.
  – Mosaic performance coverage is slightly below target calibration coverage.

Simulation Results

• Eliminating day-type split will improve weekend overage and mosaic performance overall.
• Forward-backward sample scheme is shown to provide benefit to FRP and IR.
• The next step is to run further simulations across BAAs to ensure the amount of historical data is optimized.
  – Example 90 days historical, 90 days forward
Summary of next step focus areas for FRP requirements

Increase Sample Size:
Day Type Modifications

Historical Data Utilized:
Forward and Backward Looking

Threshold:
Further study on dynamic thresholds

Future Awareness of Requirements:
Posting future requirements for customers
## Next steps with priority order

<table>
<thead>
<tr>
<th>Priority</th>
<th>Impact</th>
<th>Description</th>
<th>Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Medium</td>
<td>Change day-type within the mosaic methodology</td>
<td>Weekend model performance</td>
</tr>
<tr>
<td>2</td>
<td>Low</td>
<td>Change existing model parameters – in relation to historical sample data utilized</td>
<td>Overall model performance especially during seasonal transition</td>
</tr>
<tr>
<td>3</td>
<td>Medium</td>
<td>Posting future requirements for further stakeholder visibility</td>
<td>Customer awareness</td>
</tr>
<tr>
<td>4</td>
<td>TBD</td>
<td>Evaluate key timing utilized in requirement formulations</td>
<td>TBD</td>
</tr>
<tr>
<td>5</td>
<td>TBD</td>
<td>Dynamic threshold simulations and potential modification</td>
<td>TBD</td>
</tr>
</tbody>
</table>
Pareto Optimization of Calibration and Cost (FRU/PEAK)

FRU Requirement vs. FRU Coverage HE 18-21 (Period Mean)
mosaic w/threshold (blue) vs. histogram (red) | sample size (ascending)
production benchmark (black-dashed) | target coverage (gold-dashed)

20230523 to 20231023
Pareto Optimization of Calibration and Cost (FRU/ALL)

FRU Requirement vs. FRU Coverage All Hours (Period Mean)
mosaic w/threshold (blue) vs. histogram (red) | sample size (ascending)
production benchmark (black-dashed) | target coverage (gold-dashed)

- Variable:
  - nh_up
  - current_threshold

as.factor(SAMPLE_DAYS)

- Requirement [MW]:
  - 1050 to 1150
  - 20230523 to 20231023

- Percent Coverage:
  - 96.0% to 98.0%
Simulation results (scheme 4, 160 days) – Informativeness

Removing day-type split and enhancing the sample scheme improves the performance of mosaic as seen in simulated summer 2023 results, especially at low uncertainty values. Histogram changes from production version of this plot as sample scheme is revised, so histogram values reflect simulated histogram requirements.
PRODUCTION RESULTS
APPENDIX
Morning ramp (summer 2023, CISO)

Mosaic:
- Uncertainty exceeding mosaic and/or histogram (6am-9am)

Histogram:
- Uncertainty exceeding mosaic and/or histogram (6am-9am)
Evening ramp (summer 2023, CISO)

**Mosaic**

Uncertainty exceeding mosaic and/or histogram (4pm-7pm)

**Histogram**

Uncertainty exceeding mosaic and/or histogram (4pm-7pm)

summer 2023, CISO

ISO PUBLIC
Evening peak (summer 2023, CISO)
STF BAA-level Replication of Percentage of Significant Inte

Statistical Significance of Mosaic Coefficients All Hours (Period Mean)
BAA NAME (color) | DMM Reported (shape)

Percent Significance SQUARE Term

Percent Significance LINEAR Term

TID, TPWR, SCL (entities w/o wind or solar)
Empirically Three Different Ways to Interpret Proposed Enhancements

1) More Reliability / Less Requirement
   *Mosaic / Sample Scheme 4 / 90 days*

2) Better Model
   *Require weighting criteria*

3) Hard Constraint
   *Differs between peak and all periods*

1) baseline mosaic
   (current parameters)

2) calibration target
   (can exceed)

3) constraint
   (cannot exceed)