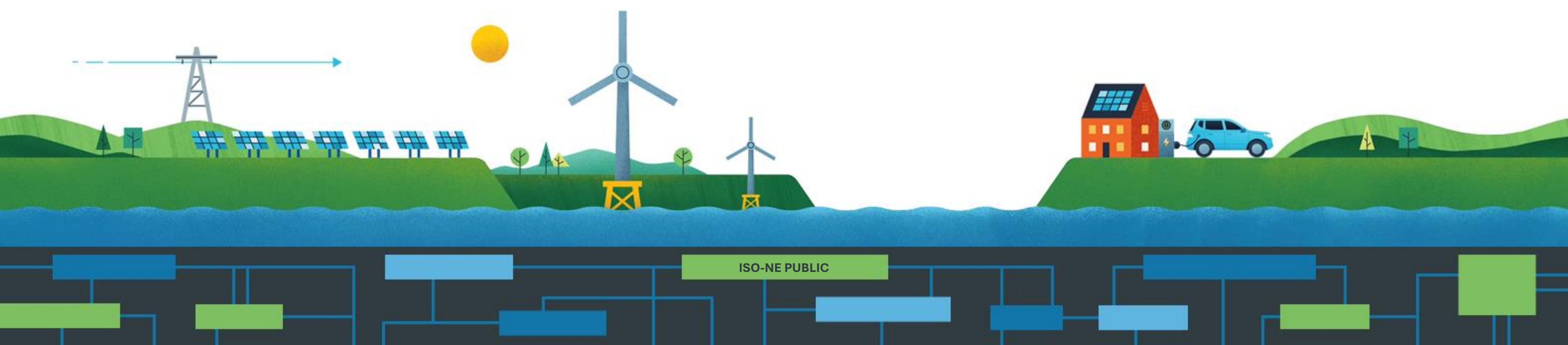
MARCH 2026

Prototype Methodology for 7-Day Probabilistic Load Forecast

Review ISO-NE Developments in Short-Term Probabilistic Forecasting

Joseph Roberts

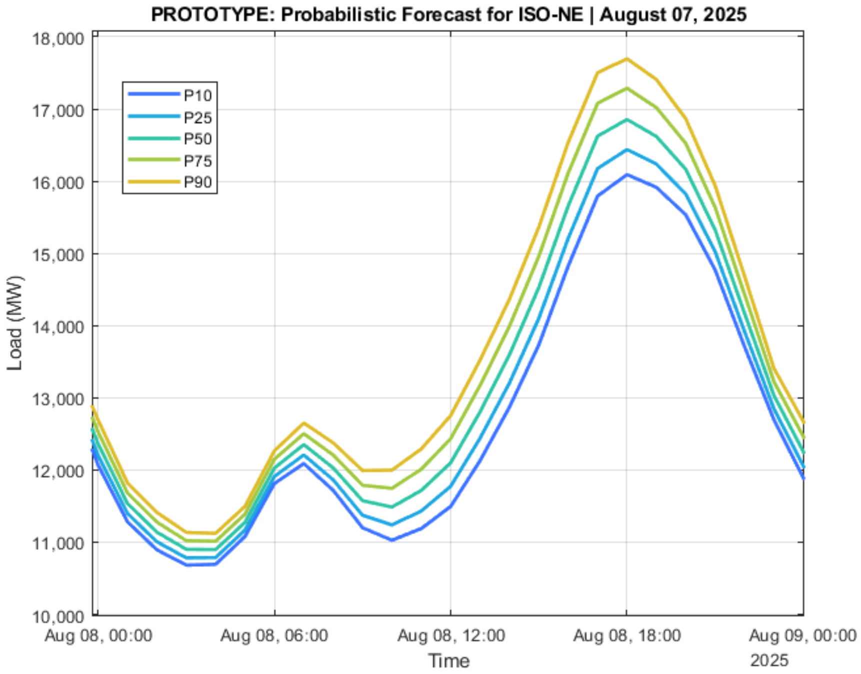
LEAD OPERATIONS FORECASTING ANALYST



Introduction

- Probabilistic load forecasts are meant as a tool to quantify the uncertainty in our load forecast
  - Intention is not to make our forecast better, but to explain our forecast better

- This presentation reviews recent developments on this topic
  - This is intended as a reasonable prototype of 7-day probabilistic load forecasts



Overview – 3 Pillar Components

Reanalysis of Weather Forecast Errors

- Create realistic weather scenarios
- Run weather scenarios through LF models
- Combine results with probabilistic bands

Reanalysis of Load Forecast Errors

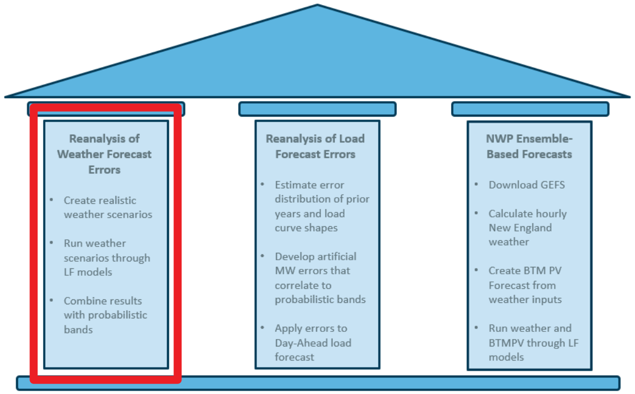
- Estimate error distribution of prior years and load curve shapes
- Develop artificial MW errors that correlate to probabilistic bands
- Apply errors to Day-Ahead load forecast

NWP Ensemble-Based Forecasts

- Download GEFS
- Calculate hourly New England weather
- Create BTM PV Forecast from weather inputs
- Run weather and BTMPV through LF models

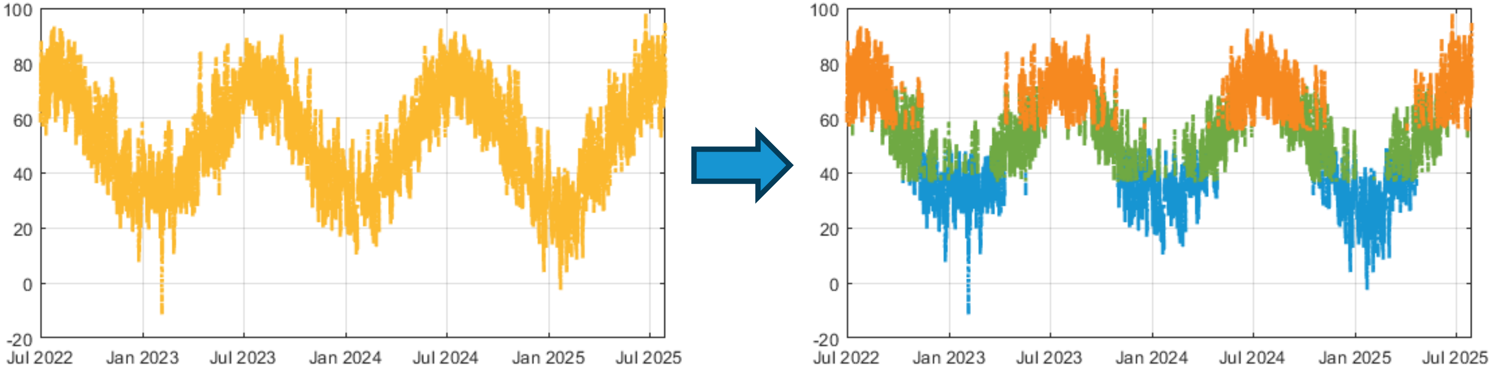
METHODOLOGY 1

Reanalysis of Historical Weather Forecast Error



Split weather data into regimes

- Weather forecast error is not consistent through the year
  - Weather forecasts for high heat events are likely to be skewed low
  - Weather forecasts for severe cold events are likely to be skewed high
- Split weather actuals into three (3) regimes
  - Hot
    - Highest 33% defined by THI (*note: highest 33% is determined for each hour)*
  - Cold
    - Lowest 33% defined by Effective Temperature (*lowest 33% for each hour)*
  - Mild
    - Neither "Hot" or "Cold"



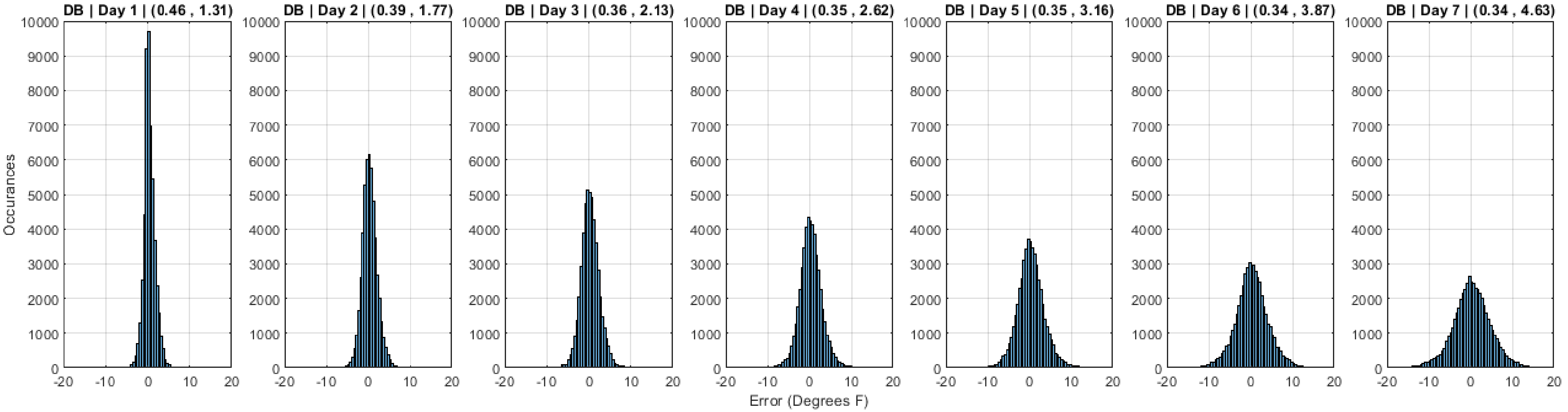
Split weather data into regimes

- Results in an hourly "decision matrix"
- Given a weather forecast for a particular hour, calculate THI and Effective Temperature
  - Then, use the decision matrix to determine forecast regime

Hour	Eff.T. | 33%	THI | 67%
1	37.18	59.61
2	36.57	59.14
3	35.85	58.72
4	35.42	58.33
5	35.10	57.93
6	34.90	57.56
7	34.65	57.46
8	34.66	58.13
9	35.94	59.46
10	38.55	60.97
11	40.94	62.35
12	42.62	63.40
13	44.13	64.44
14	45.44	64.97
15	45.85	65.33
16	46.10	65.57
17	45.35	65.46
18	44.23	65.20
19	42.86	64.55
20	41.39	63.43
21	40.28	62.27
22	39.58	61.30
23	38.74	60.61
24	37.83	60.04

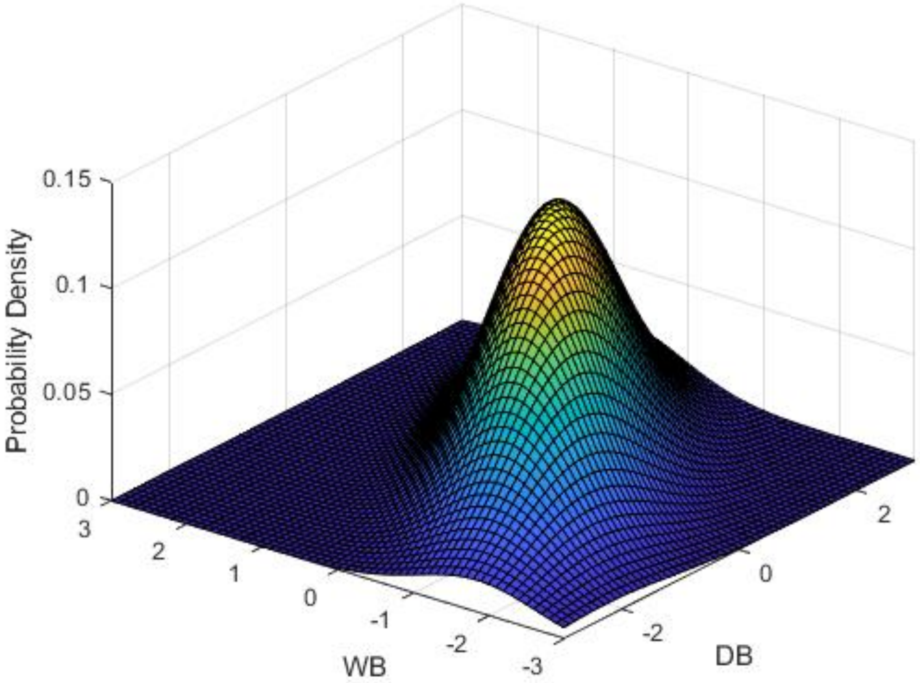
Calculate Weather Forecast Error across 7-Day Horizon

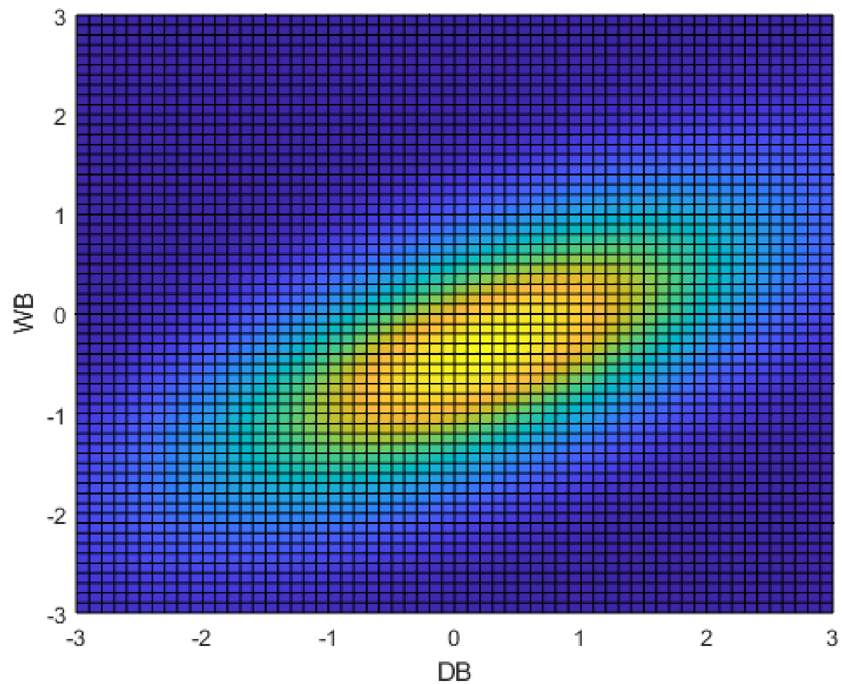
- Each weather forecast extends 7 days
  - Forecasted weather concepts: DB, DP, WB, WS, CC, GHI
- Weather forecast accuracy degrades across the forecast horizon
  - Calculate error for each weather concept for each day of the forecast horizon



Model Distribution of DB and WB error

- DB and WB are significant drivers of load
  - Model the DB and WB error as jointly normal using the means, variances, and covariance matrix

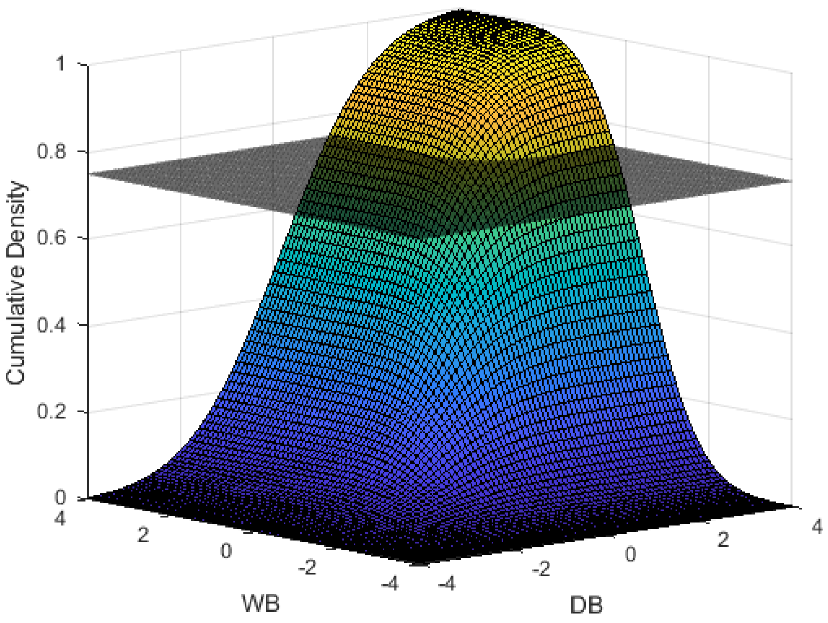


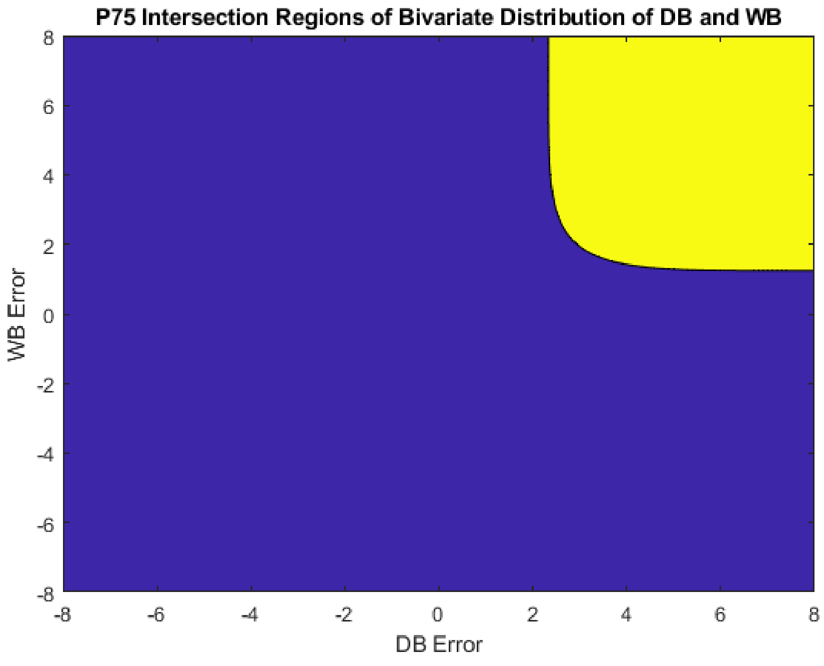


Probability density function of bivariate normal distribution of the DB and WB forecast error (side and top view)

Model Distribution of DB and WB error, continued

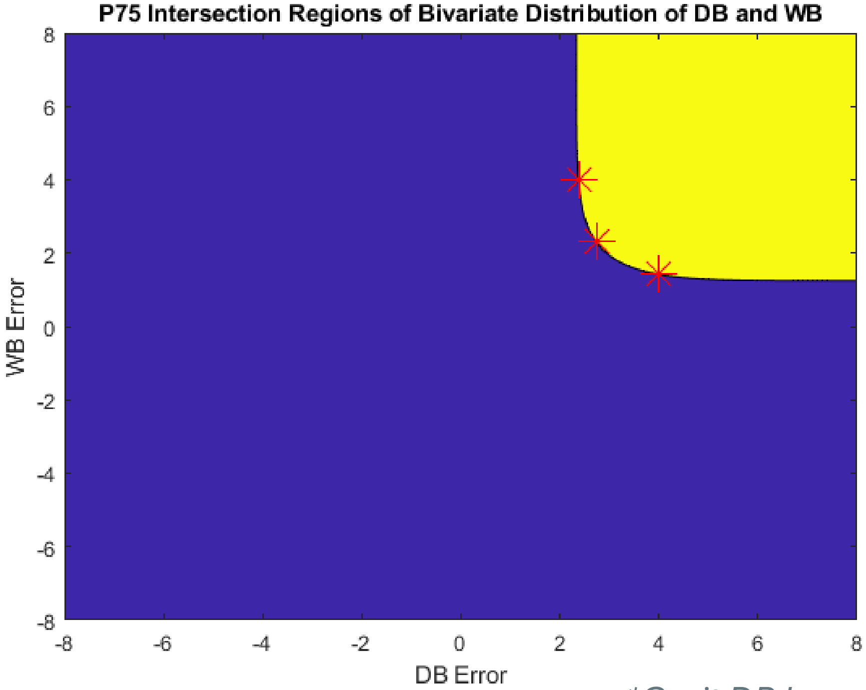
- Intersect the bivariate cumulative distribution function and intersect percentiles – this is what “assigns” the probability to the condition
  - Note: intersection region is a curve





Construct Coincident Weather Scenario

- Sample (3) points of the intersection curve
  - Sample at percentiles: 33, 50, and 67
- Construct conditional distribution of the error for remaining weather concepts (WS, CC, GHI)*



Take multivariate normal distribution:

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

Now, given that we know X2, what is the distribution of X1?

It is multivariate normal with:

Conditional mean = $\mu_1 + \Sigma_{12}\,\Sigma_{22}^{-1}(x_2 - \mu_2)$
Conditional Covariance = $\Sigma_{11} - \Sigma_{12}\,\Sigma_{22}^{-1}\,\Sigma_{21}$

*Omit DP because we included WB

Construct Coincident Weather Scenario

- Draw (10) random samples from the conditional distribution for each of the (3) selected intersection points
  - Result: 30 forecast error scenarios for each percentile
    - Apply errors to current weather forecast

- Recap: We have constructed 30 weather scenarios for each of the 5 percentiles for a total of 150 weather scenarios

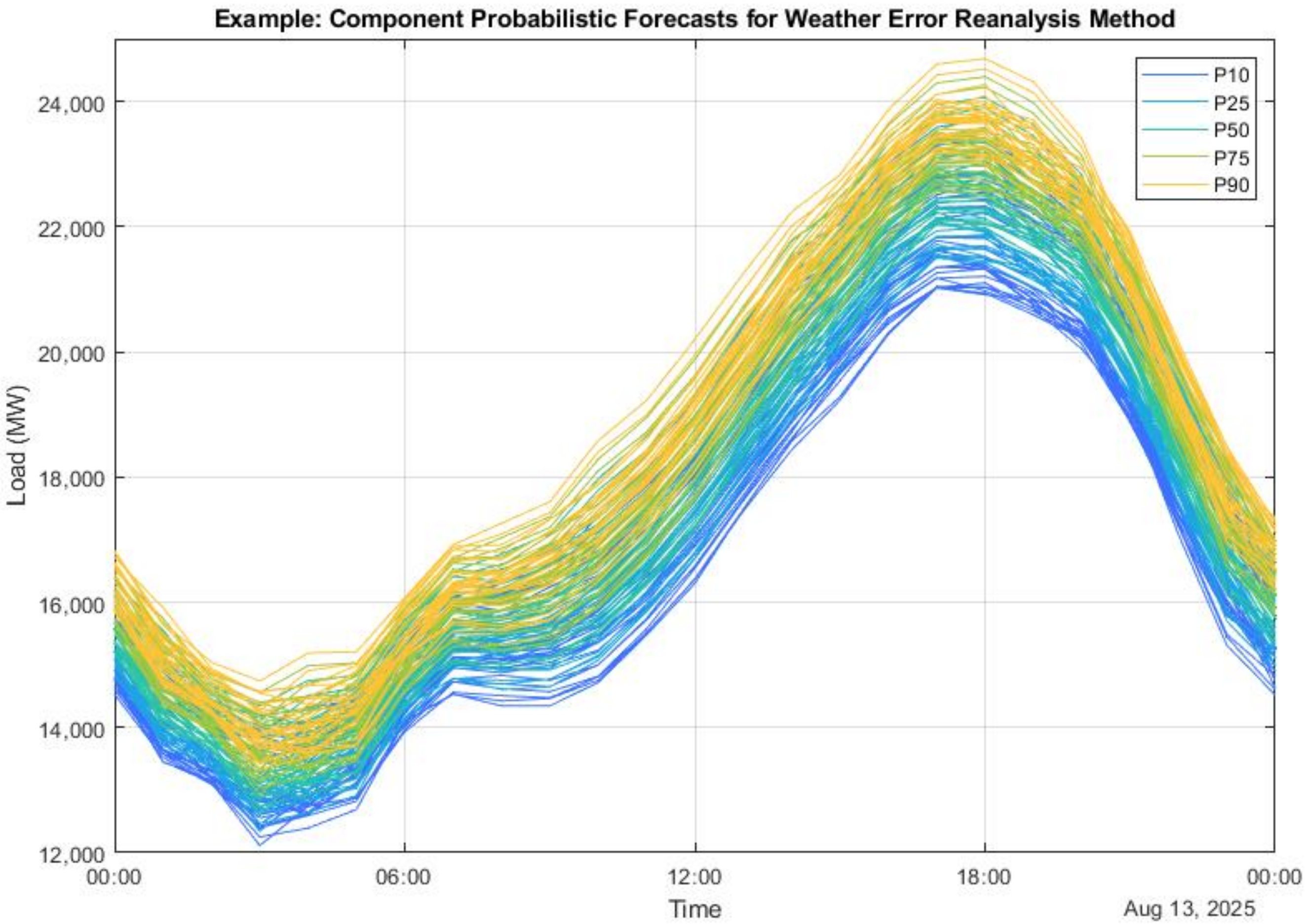
Implement Load Forecast Model

- Use each weather scenario as input into (3) separate load forecast models
  - Each load forecast model consist of 25 neural networks
    - 1 Daily Energy model
    - 24 Hourly Load models

- Result: For each of the 5 percentiles, we ran 30 weather scenarios through 3 load forecast models
  - 90 load forecasts at each of the 5 percentiles
  - 450 component load forecast models

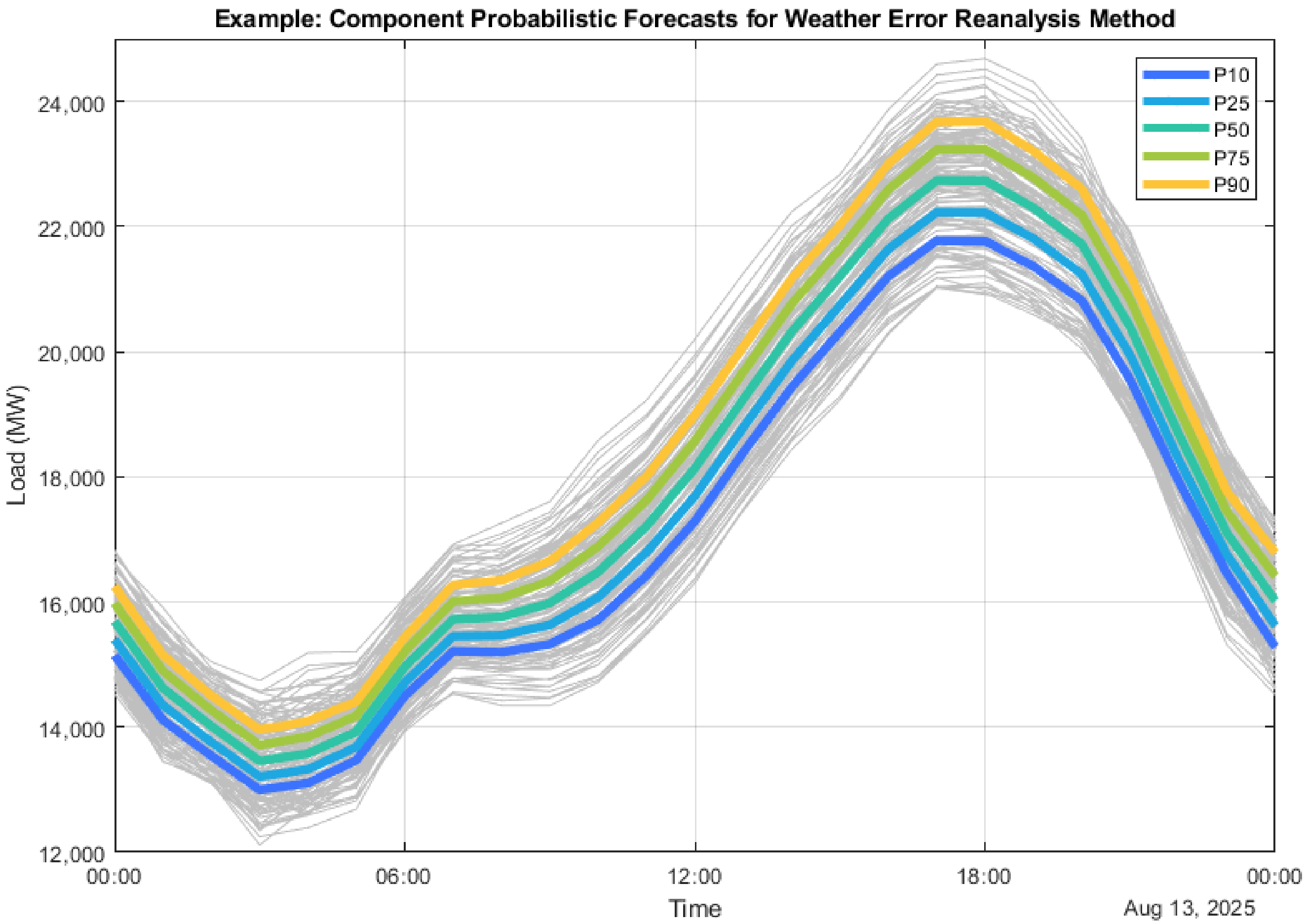
- Finally, at each percentile, collapse the 90 component forecasts into 1 forecast

Implement Load Forecast Model



Note: Overlap of curves exists

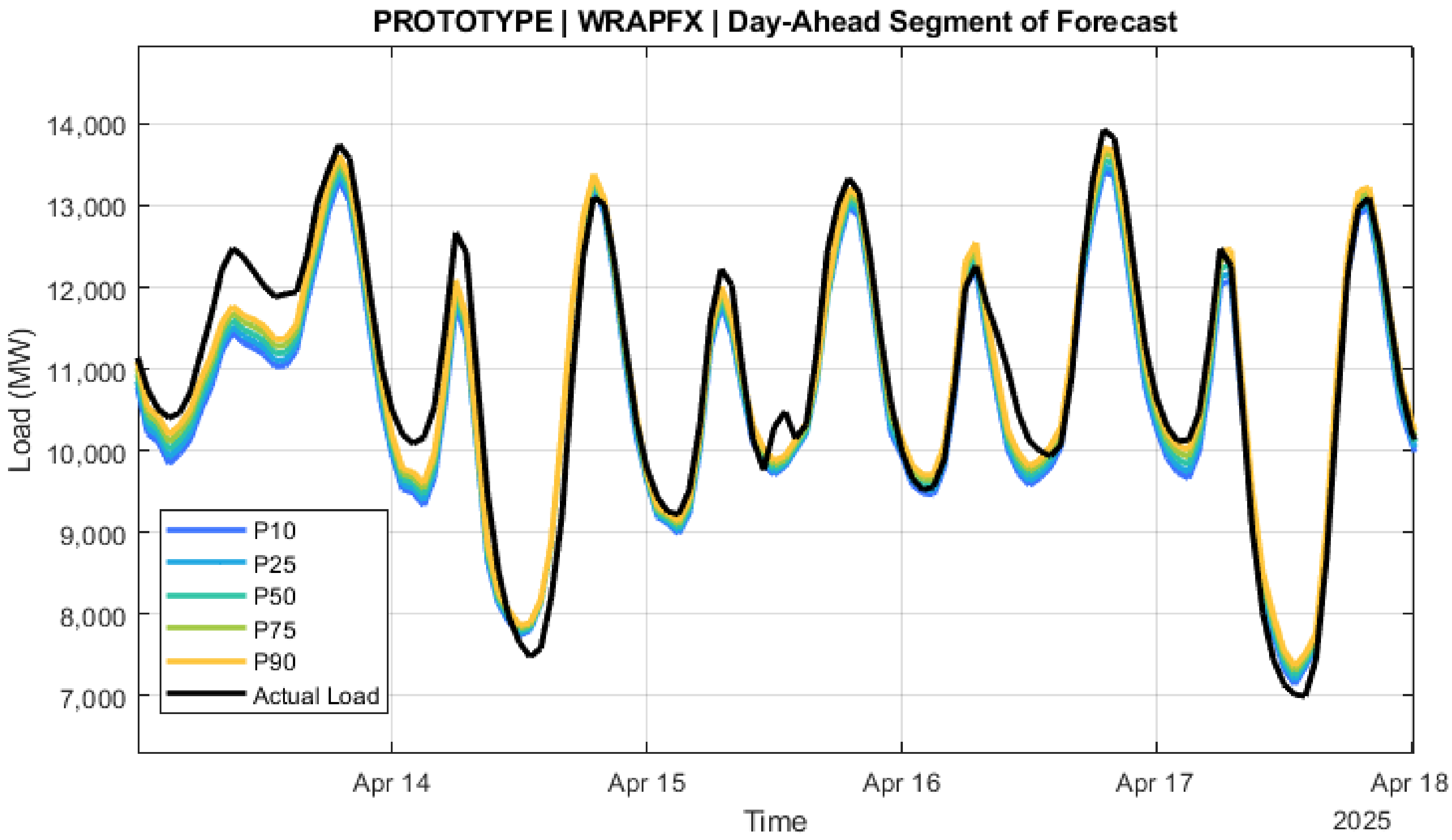
Implement Load Forecast Model



Limitations of Methodology 1

- When weather is not a major driver of load forecast error, the width of the bands shrinks and does not capture *any* of the actual loads
  - The load forecast model error is greater than the weather induced load forecast error
- Low bias in winter
  - Will be addressed with better modeling
- Methodology is tied to the strengths and weaknesses of the load forecast models used
  - Note, this was not an exercise in creating *really good load* forecast models, so improving the models would improve results
- Methodology does not explicitly capture PV-induced forecast error
  - Major improvement underway: Daylight hours will be driven by bivariate distribution of DB/BTMPV instead of DB/WB

Limitations, graphically



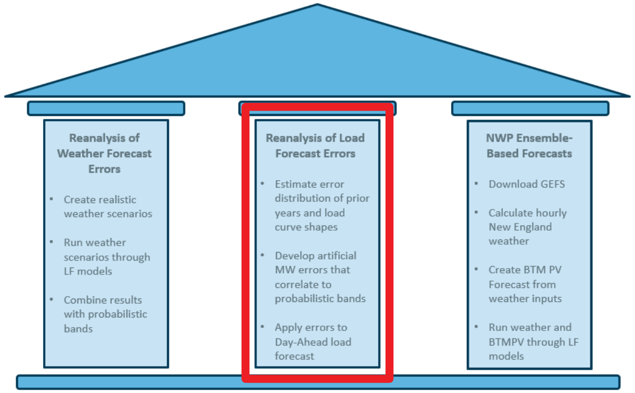
Many of these days represent a decent load forecast, but the probabilistic bands do not capture the actual load!

Next Steps

- Explore use of PV error reanalysis
  - Instead of defining percentiles on DB/WB, define percentiles on DB/PV during mid-day hours
- Improvement of load forecast models
  - Improve specifications and add new modeling techniques to improve model diversity

METHODOLOGY 2

Reanalysis of Historical Load Forecast Error



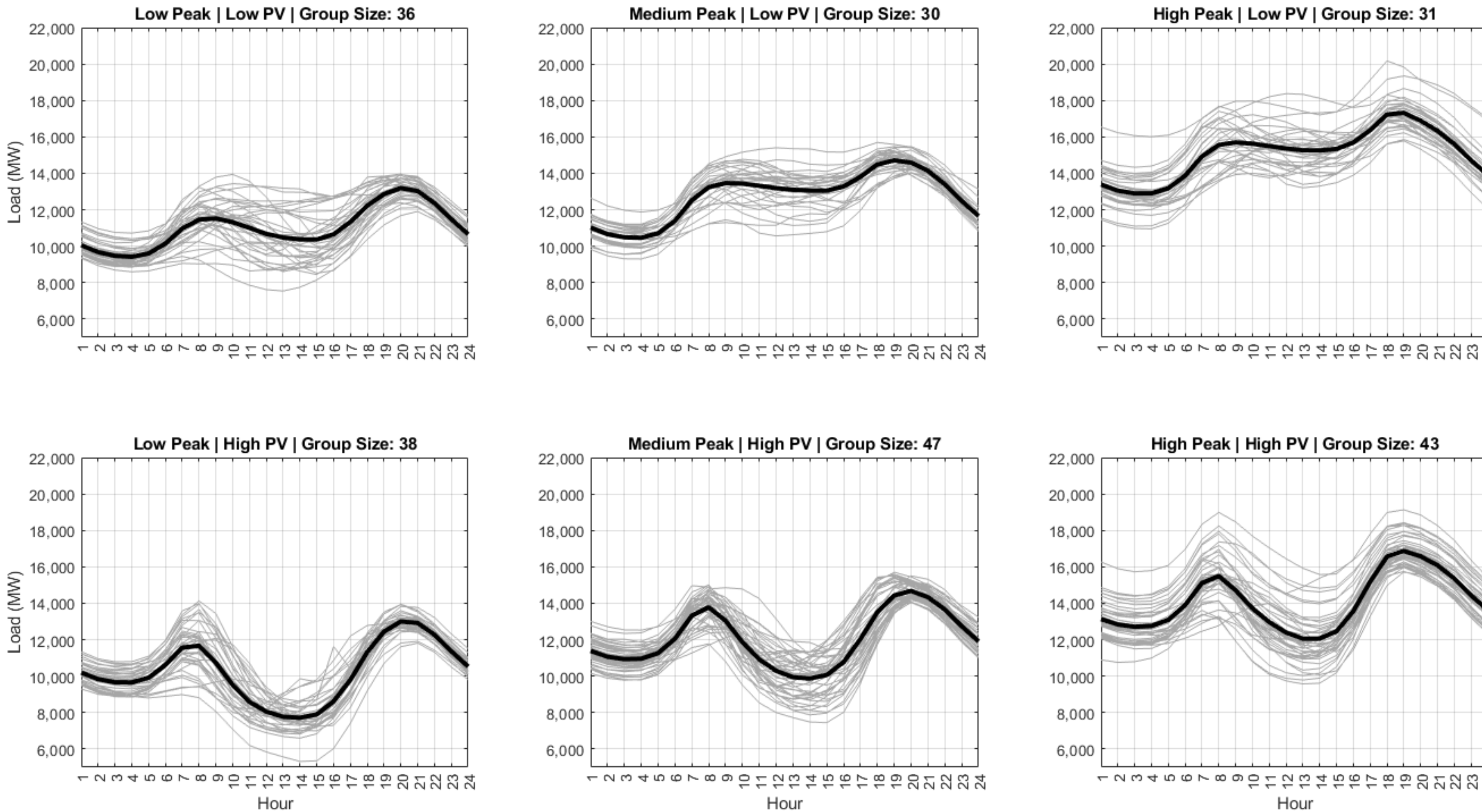
Find All days within +/- 45 days of calendar date

- Example: Use Monday, August 18, 2025
  - Use the dates from the following ranges
    - 7/4/2023-10/2/2023 | 91 days
    - 7/4/2024-10/2/2024 | 91 days
    - 7/4/2025-8/17/2025 | 45 days
- The date range slides with each day of the forecast horizon
  - Day 3 of the 8/18/25 forecast is 8/20/25, which would use the following ranges:
    - 7/6/2023-10/4/2023 | 91 days
    - 7/6/2024-10/4/2024 | 91 days
    - 7/6/2025-8/17/2025 | 43 days

Sort days into bins

- Sort subset of days by evening peak
  - Low Peaks: All days with evening peak < P33
  - High Peaks: All days with evening peak > P67
  - Medium Peaks: All the other days
    - P33 and P67 are the 33rd and 67th percentile of evening peaks in the subset of days
- Sort each “peak bin” into **high** and **low** PV-impacted load shape
  1. Calculate difference between <u>evening peak</u> and <u>mid-day minimum</u> (afternoon ramp)
  2. Calculate difference between <u>morning peak</u> and <u>mid-day minimum</u> (morning ramp)
  3. Complete k-means clustering on 2-D data of both differences
     - Weight evening ramp higher than morning ramp
       - Many days, particularly weekends, have very low morning peaks, making it less reliable

Sort days into bins

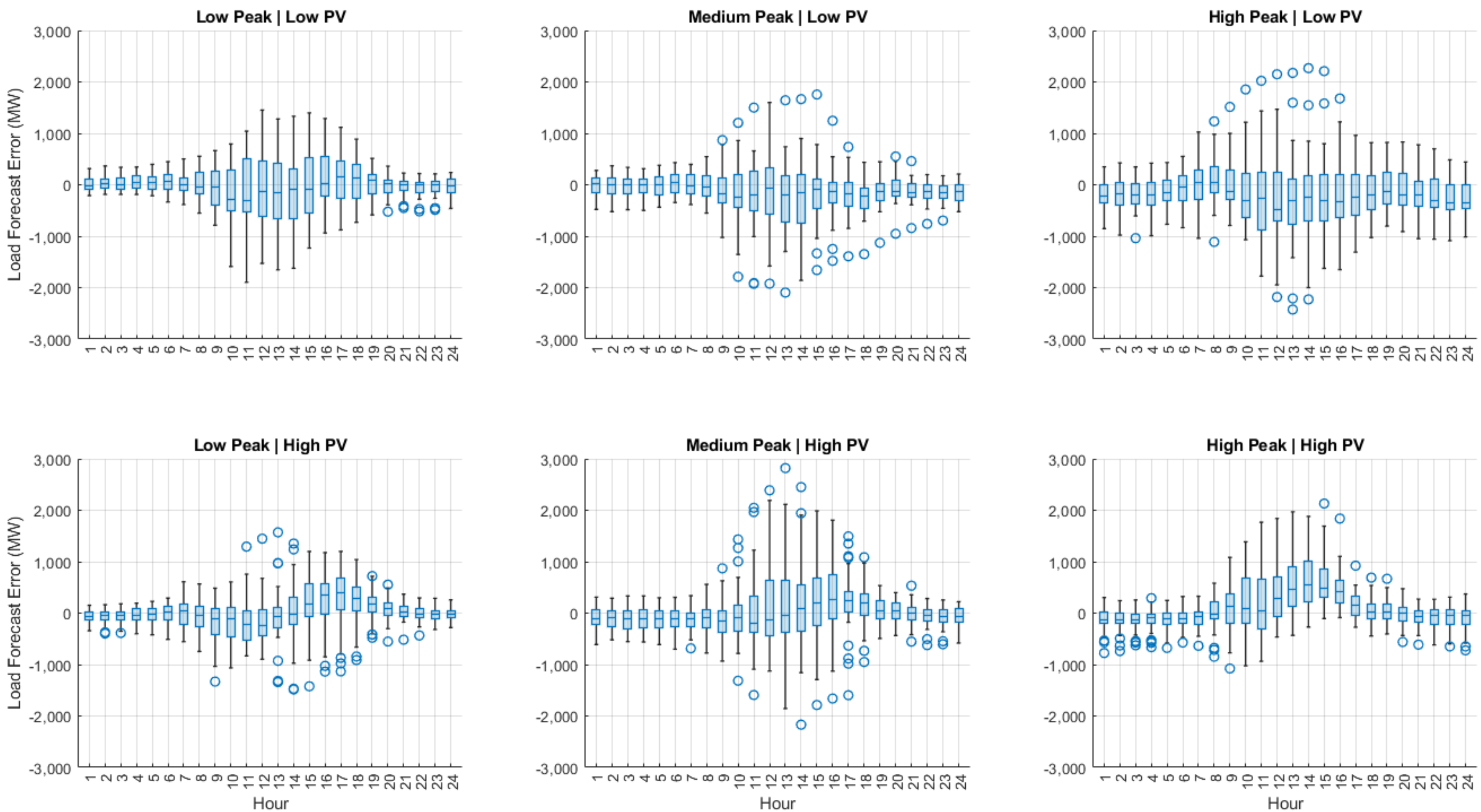


Note the more pronounced “duck” curve on the high-PV days

Calculate Hourly Load Forecast Error by Bin

- For each bin, calculate the hourly load forecast error
  - Error is calculated separately for each day of the 7-day forecast
- Model hourly error as normally distributed

Hourly Day-Ahead Load Forecast Error

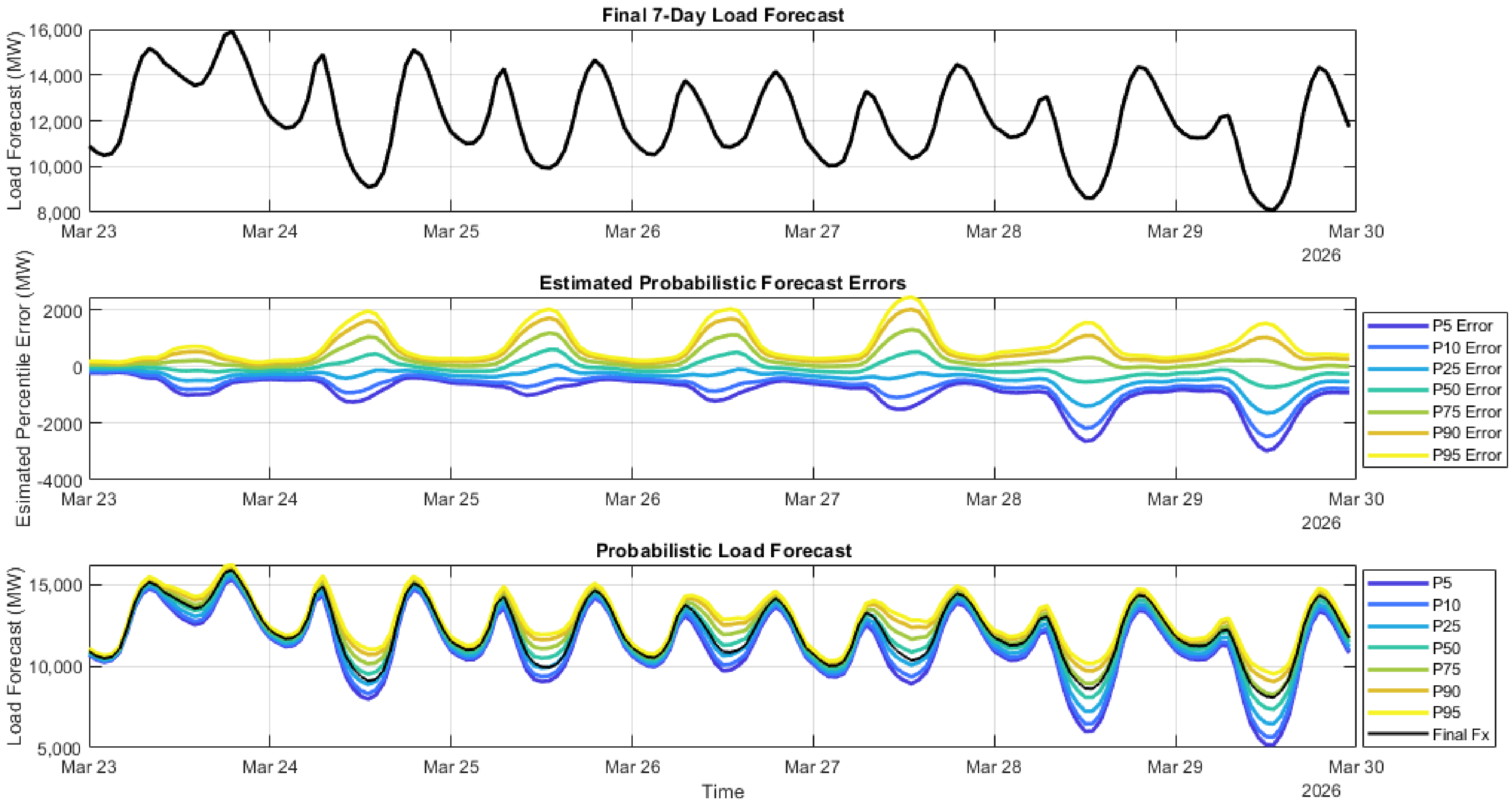


These error distributions will change with each grouping of days, but it is gradual, as the process “drops 1 day, and adds 1 day” each day

Create Load Errors for Given 7-day Forecast

- Query 7-day load forecast
- Loop through each day of forecast and determine to which group the daily forecast belongs
- For each hour in the day, estimate the percentile MW error from the normal distribution for the given group and forecast day
  - Results in 168-hourly MW errors associated with each percentile
- Add matrix of MW errors to the load forecast to create probabilistic forecast

Putting it all together - Example



*Probabilistic Forecast via this method is the deterministic forecast **plus** the estimated probabilistic errors*

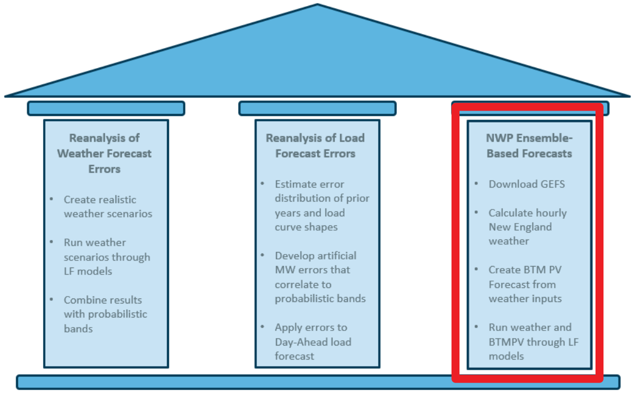
Limitations of Methodology 2

- Potentially too simplistic
- Hourly distribution of load forecast error may not be normal
- No growth of PV installed capacity is considered

- *Note, this methodology performs much better when model forecast error dominates weather forecast error*

METHODOLOGY 3

NWP Ensemble Load Forecasts



Overview

1. Download GEFS 0.25° (31 members, 240 hours each)
2. Find nearest grid point to 23 New England cities
   1. ISO-NE forecasting uses weather data at 23 different cities
3. Apply PCHIP temporal interpolation
   1. Converts 3-hour data to hourly data
      1. PCHIP (piecewise cubic Hermite interpolating polynomial) preserves physical realism by preventing overshooting and artificial extrema
4. Execute BTM PV forecasts for each member
5. Execute Load Forecast models for each member
   1. These models ingest the PV forecast

Fields Used from GEFS

- Currently, we use:
  - Dry bulb (DB)
  - Dew Point (DP)
  - Cloud Cover (CC)
  - Precipitation (PWAT)
  - Wind Speed
  - Downward shortwave radiation flux (DSRF, proxy for GHI)

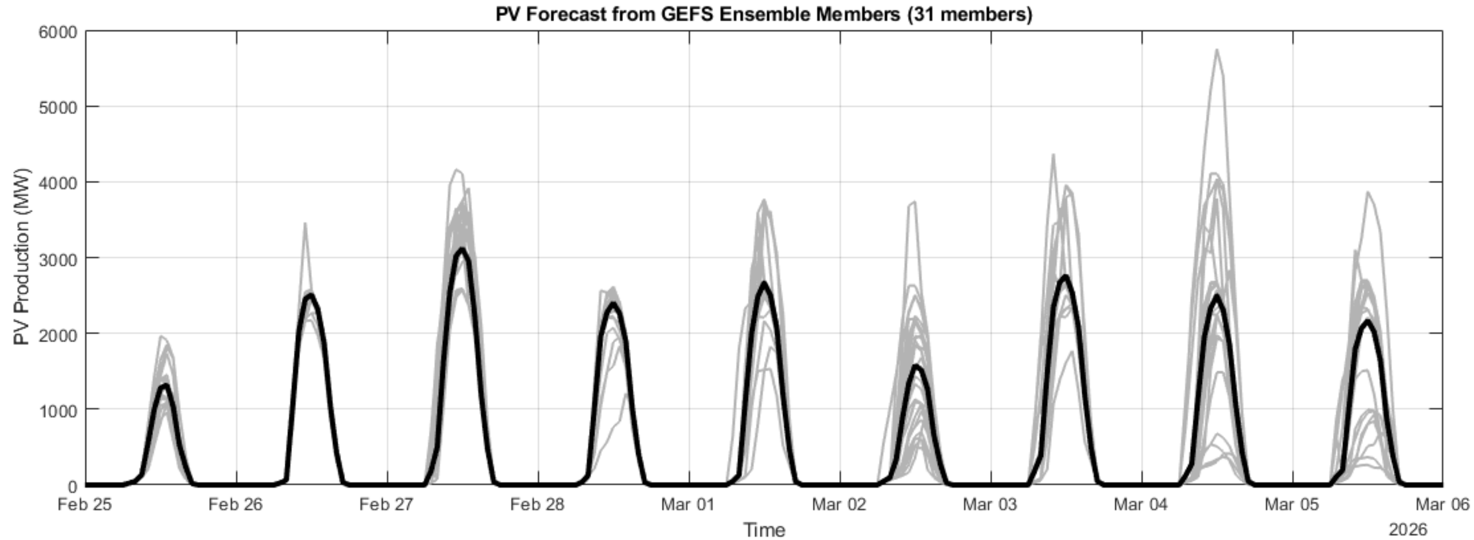
  - These are the variables used for the PV forecast and Load Forecast

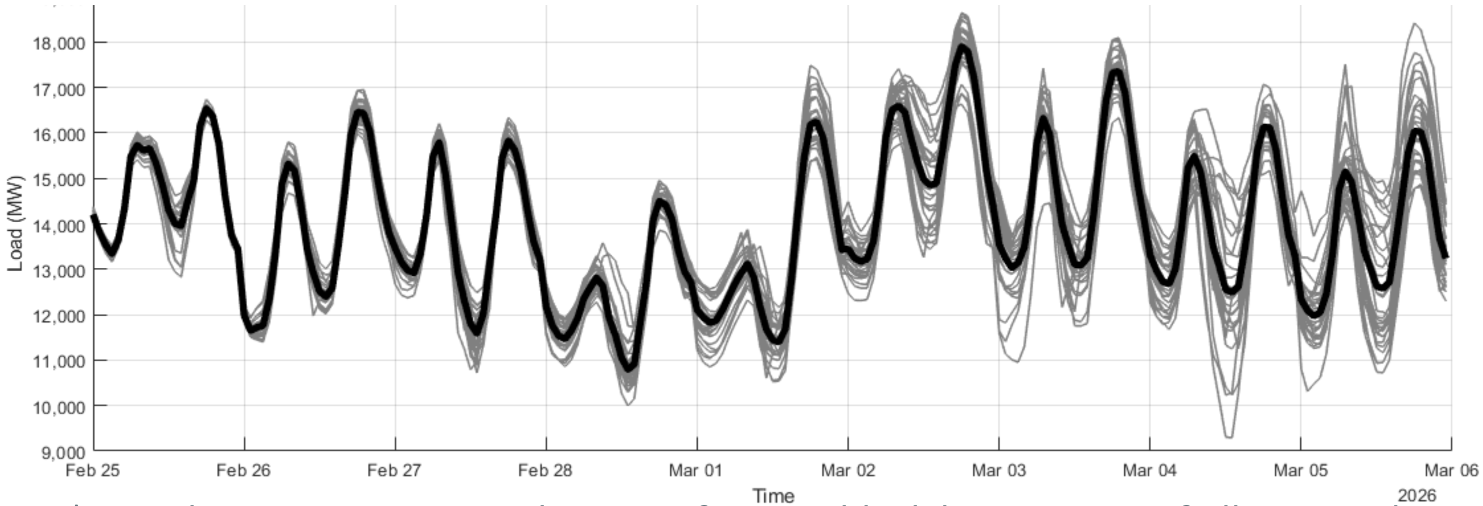
BTM PV and Load Forecast Implementation

- Simple PV forecast has been implemented
  - Bagged tree model
    - Inputs:
      - Clear Sky | Clear Panel index
        - » Theoretical maximum PV output on a <u>perfectly clear day</u> with <u>perfectly clear panels</u>
      - Installed Capacity
      - Weather concepts
  - Designed as an "assume a perfect baseline, then derate total MWs" model

- Load forecasts are same NN models as used in other steps
  - For each member run of the load forecast model, the same member-specific PV forecast is used as input

NWP Ensemble-Based Forecasts





1) Gray lines represent member-specific runs, black line is mean of all 31 members

2) At this stage, ensembles are used for shape and spread, not calibrated probabilities

Limitations of Methodology 3

- Spatial and temporal granularity
  - 0.25 gridded data is approximately 12-17 miles wide
    - This is rather wide for PV forecasts, but it is the best we can do with this model
  - 3 hour granularity is coarser than the hourly models we develop, so the interpolation scheme may not exactly capture maxima and minima
- We do not have crisp or well-defined probabilities for each member's output
  - Each member is not equally likely, and we are uncomfortable assuming this
- Model only runs out 10 days

Next Steps

- Productionalize this process
- Associate member output within the cloud of defined probabilities from other methods
- Improve PV and Load models, and consider additional models to improve model diversity
- Development of Wind and FTM PV models

BLENDING OF METHODOLOGIES

Combine Methods 1 and 2

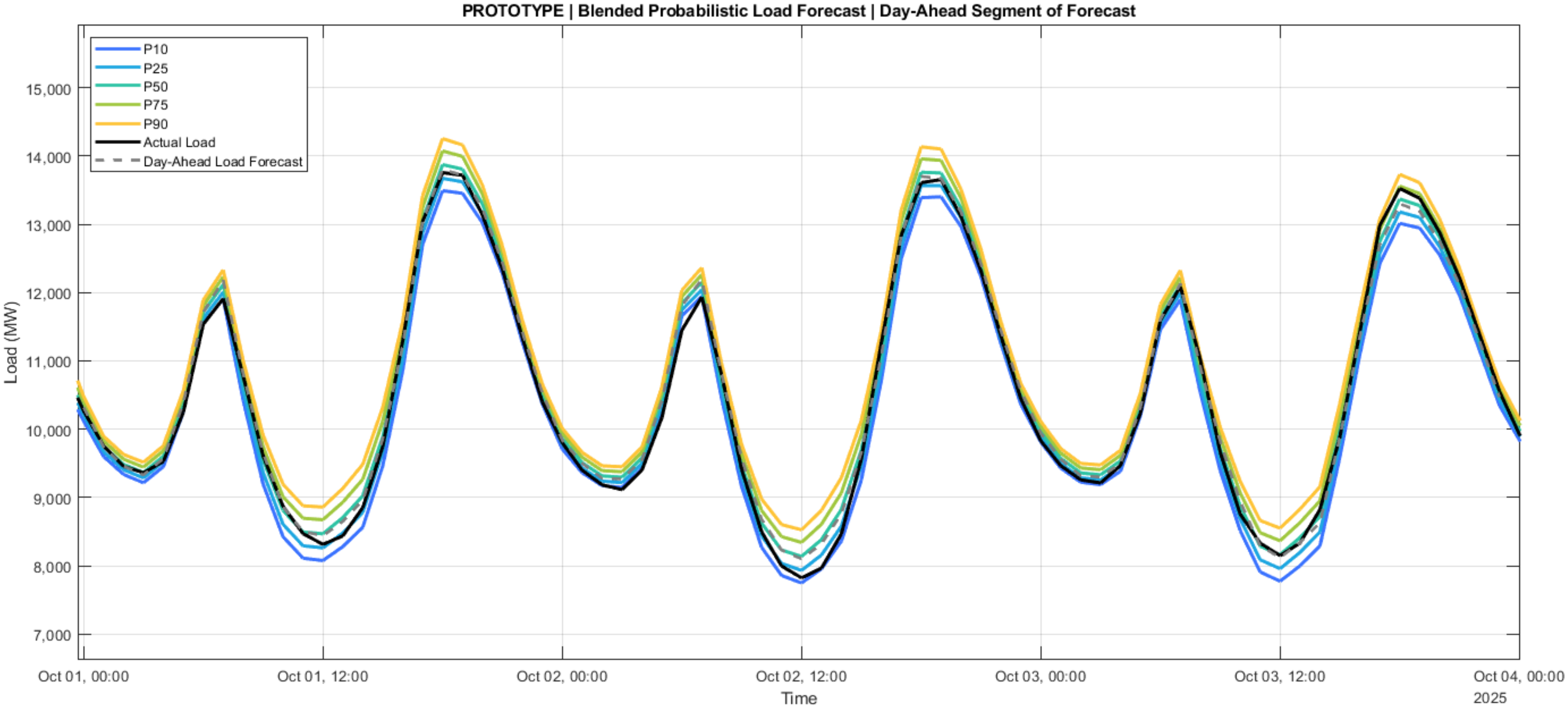
- Blending occurs hourly
  - Weights for hourly blending are proportional to the squared width of the probabilistic forecast (P90 – P10)
    - This favors a more conservative approach when faced with greater uncertainty

  - Example
    - M1: P10 = 15,000 MW, P90 = 16,000 MW → Width = 1000 MW
    - M2: P10 = 15,250 MW, P90 = 15,750 MW → Width = 500 MW

    - Weight Calculation:
      - $W1 = (1000^2) / (1000^2 + 500^2) = 0.8$
      - $W2 = (500^2) / (1000^2 + 500^2) = 0.2$

    - Blended P10 = 0.8*15,000 + 0.2*15,250 = 15,050 MW
    - Blended P90 = 0.8*16,000 + 0.2*15,750 = 15,950 MW

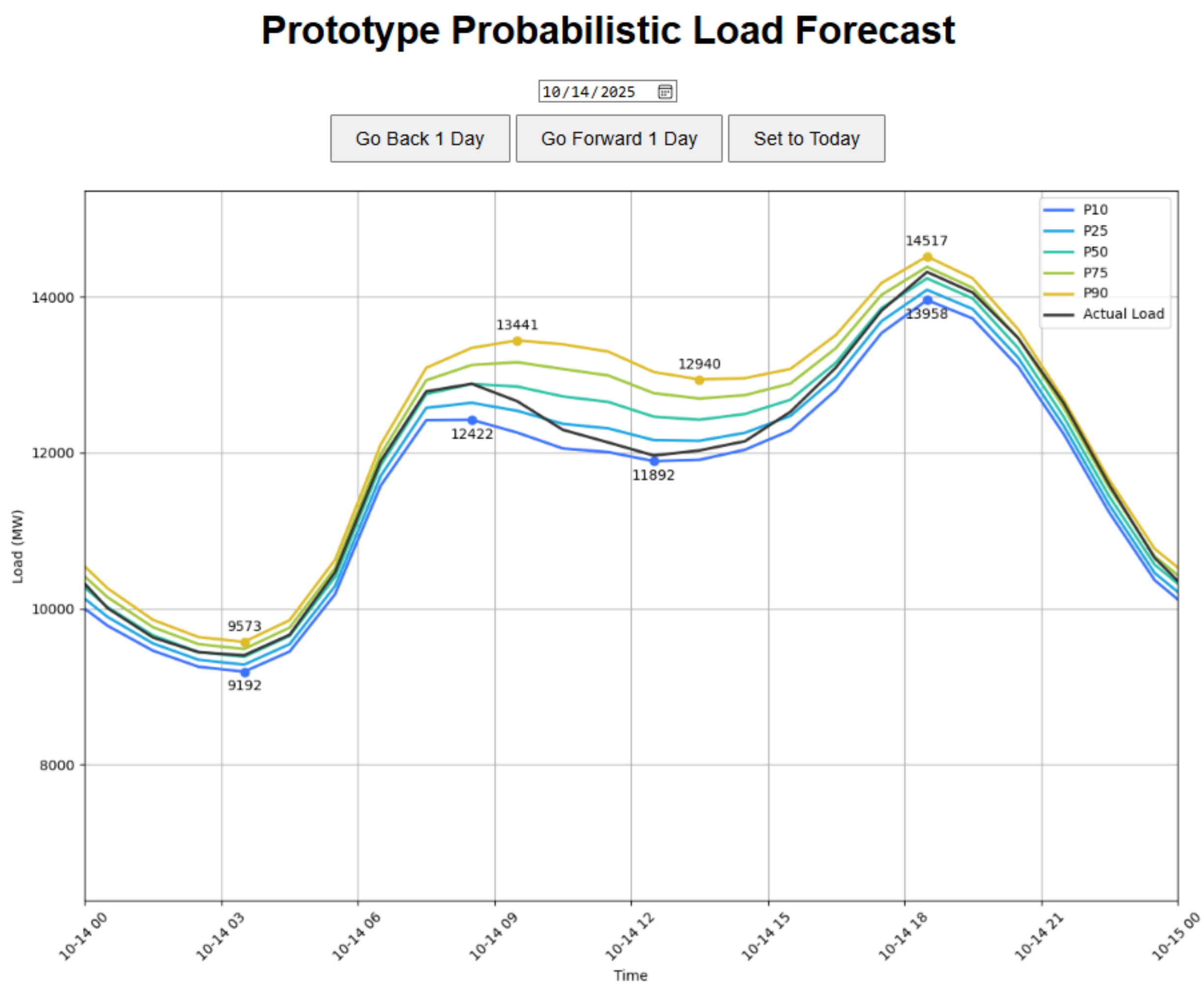
Results of Daily Process



	P10	P25	P50	P75	P90
Day Ahead Blend	11.14%	24.56%	48.45%	76.74%	90.84%

Values through ~3 months of daily running. Winter forecasts experiencing low bias.

Daily Monitor



Private HTML page is running to track hourly load forecast performance for situational awareness

NEXT STEPS

In the next few months

3-Pillar Design

Reanalysis of Weather Forecast Errors

- Create realistic weather scenarios
- Run weather scenarios through LF models
- Combine results within probabilistic bands

Reanalysis of Load Forecast Errors

- Estimate error distribution of prior years and load curve shapes
- Develop artificial MW errors that correlate to probabilistic bands
- Apply errors to Day-Ahead load forecast

NWP Ensemble-Based Forecasts

- Download GEFS
- Calculate hourly New England weather
- Create BTM PV Forecast from weather inputs
- Run weather and BTMPV through LF models

Next Steps

Component-level

Reanalysis of Weather Forecast Errors

- Create realistic weather scenarios
- Run weather scenarios through LF models
- Combine results within probabilistic bands

Reanalysis of Load Forecast Errors

- Estimate error distribution of prior years and load curve shapes
- Develop artificial MW errors that correlate to probabilistic bands
- Apply errors to Day-Ahead load forecast

NWP Ensemble-Based Forecasts

- Download GEFS
- Calculate hourly New England weather
- Create BTM PV Forecast from weather inputs
- Run weather and BTMPV through LF models

Copula-based Weather Scenario Sampling

- *Create realistic weather scenarios from copula*
- *Run scenarios through LF models*
- *Calculate probability of events through accumulation of scenarios*

Implement scenario development from bivariate distribution of DB/BTMPV errors during mid-day hours

Productionalize and correlate to other methodologies

Next Steps

- Redesigning combination of methodologies
  - Aggregate quantile estimates per level and fit a coherent, smooth, quantile function
    - Enables transforming P(10), P(25), P(50), P(75), P(90) into a full continuous distribution P(τ)
- Exploring probabilistic wind / PV forecasts
  - Validate vendor-provided probabilistic products against the quantile levels defined in our load methodology
    - For example: higher PV output typically correlates with lower mid-day load
    - If not possible, we need to develop our own (underway)
- Improving load-forecast model accuracy and diversity
  - Expand and refine the models to improve forecast robustness

NEXT STEPS

Medium to long-term

i.e. 3+ month

Ideas

- Extend GEFS, consider additional NWP ensembles
- Develop FTM Wind and FTM PV
- Implement additional methods of BTM PV forecasts
- Cohesive combination of methodologies that preserves relationship between components
  - Example: distribution of BTM PV should reflect probabilistic load assumptions

Questions