

# Hourly Electricity Load Forecasting Using Machine Learning Algorithms

Yinghua Wu, PJM

Laura Walter, PJM

Anthony Giacomoni, PJM

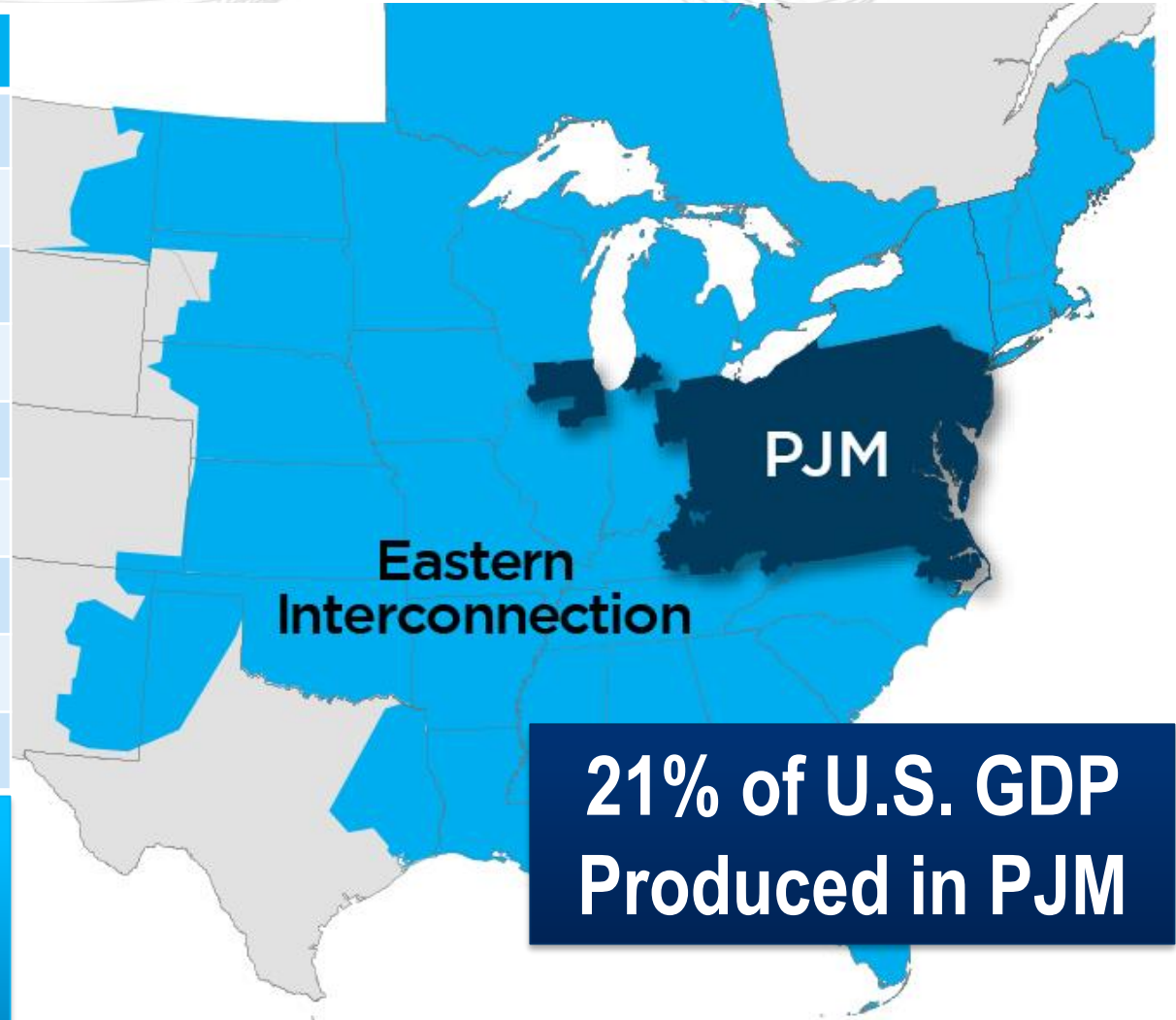
FERC Technical Conference

July 9, 2024

## Key Statistics

Member companies	1,090
Millions of people served	65+
Peak load in megawatts	165,563
Megawatts of generating capacity	180,785
Miles of transmission lines	88,185
Gigawatt hours of annual energy	770
Generation sources	1,439
Square miles of territory	368,906
States served	13 + DC

- 26% of generation in Eastern Interconnection
- 25% of load in Eastern Interconnection
- 20% of transmission assets in Eastern Interconnection



As of 2/2024

---

## Load Forecast at PJM and New Challenges

---

### Manipulate Time Series Data

- Training and testing
  - Back test
  - Challenge
- 

### Methods

- XGBoost
  - Neural network
  - LSTM
  - Transformer
  - Bias correction
- 

### Case Study

- RTO
  - Dominion
- 

### Conclusion

## Short-Term Forecast (Hourly):

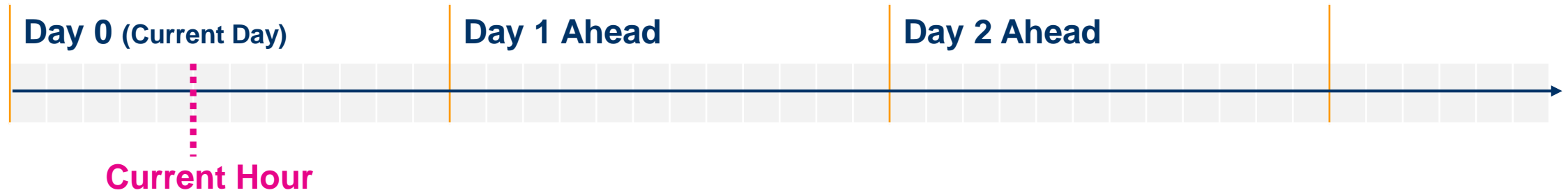
- Looks seven days ahead
- 1100: Day-Ahead Market closes
- 1800: Reliability Assessment and Commitment (RAC) run

## Very Short-Term Forecast (5-minute):

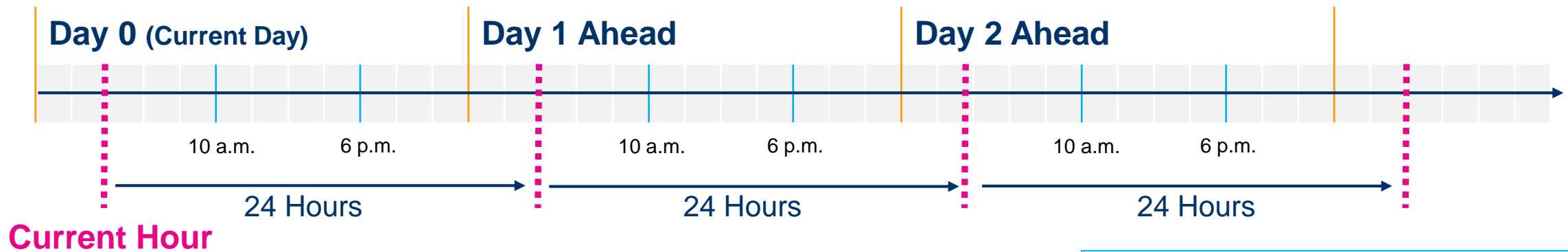
- Looks six hours ahead
- Used by Security Constrained Economic Dispatch (SCED)

Load forecasting is complex, dynamic and an important part of PJM's mission to supply reliable electricity to the 65 million people in the PJM region.

## Operation Timeline:



## Forecast Timeline:



**Important Time Points: 10 a.m., 6 p.m.**

## **XGBoost, Neural Network (NN)**

Tabular type feature ~ target prediction

## **Long Short-Term Memory (LSTM), Transformer**

- Sequence-to-sequence prediction
- Past 24 hours to encode; future 24 hours to forecast

## Experiment Setup

### **Test period:**

April 2023 to May 2024,  
progress monthly

### **Training period:**

Past seven years  
of history

## Megawatts (at FUTURE\_HOUR) ~

Year | Month | Weekday  
Hour | Holiday

Temperature | Dew point  
Wind speed | Cloud cover

Temperature differences  
WRT past and after three  
hours

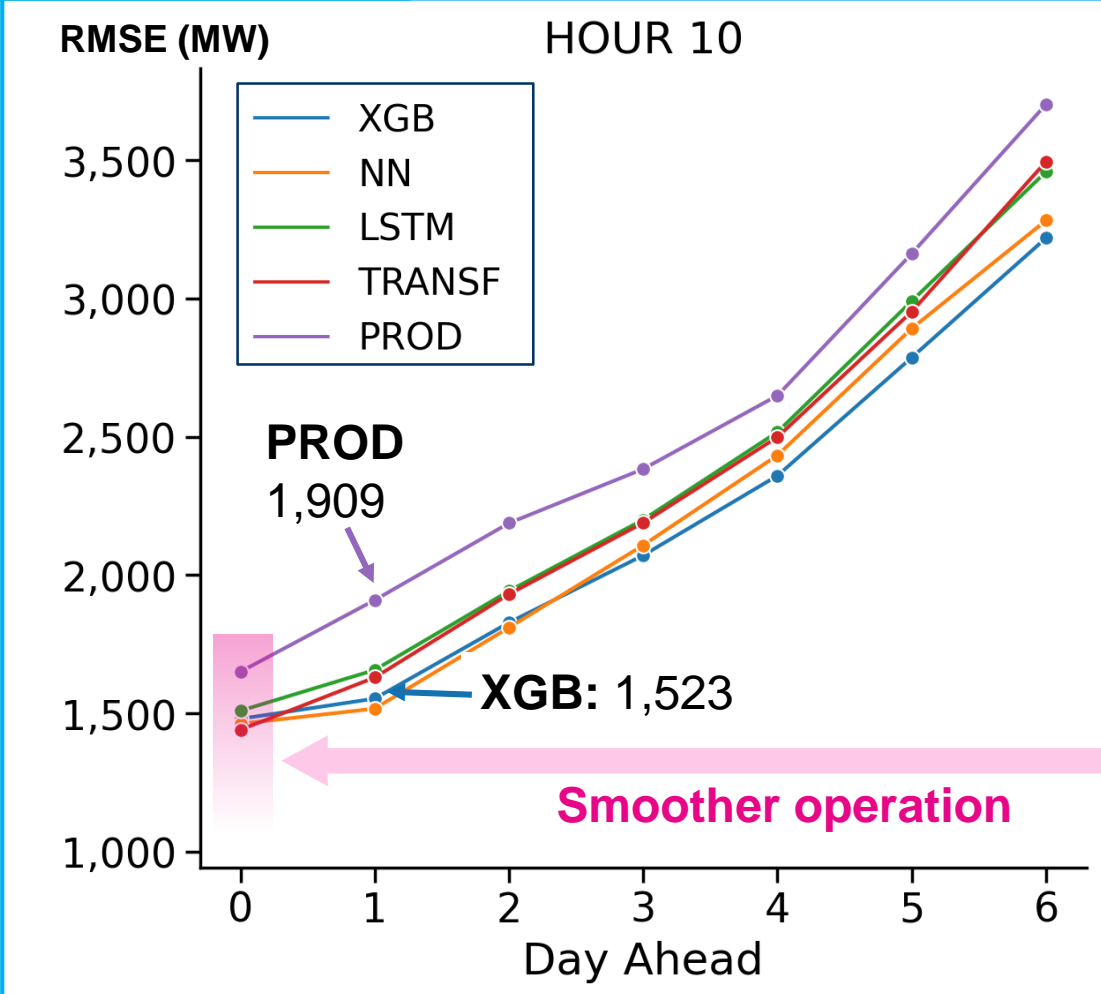
MW | Temperature at  
CURRENT\_HOUR

MW | Temperature during  
same hour yesterday

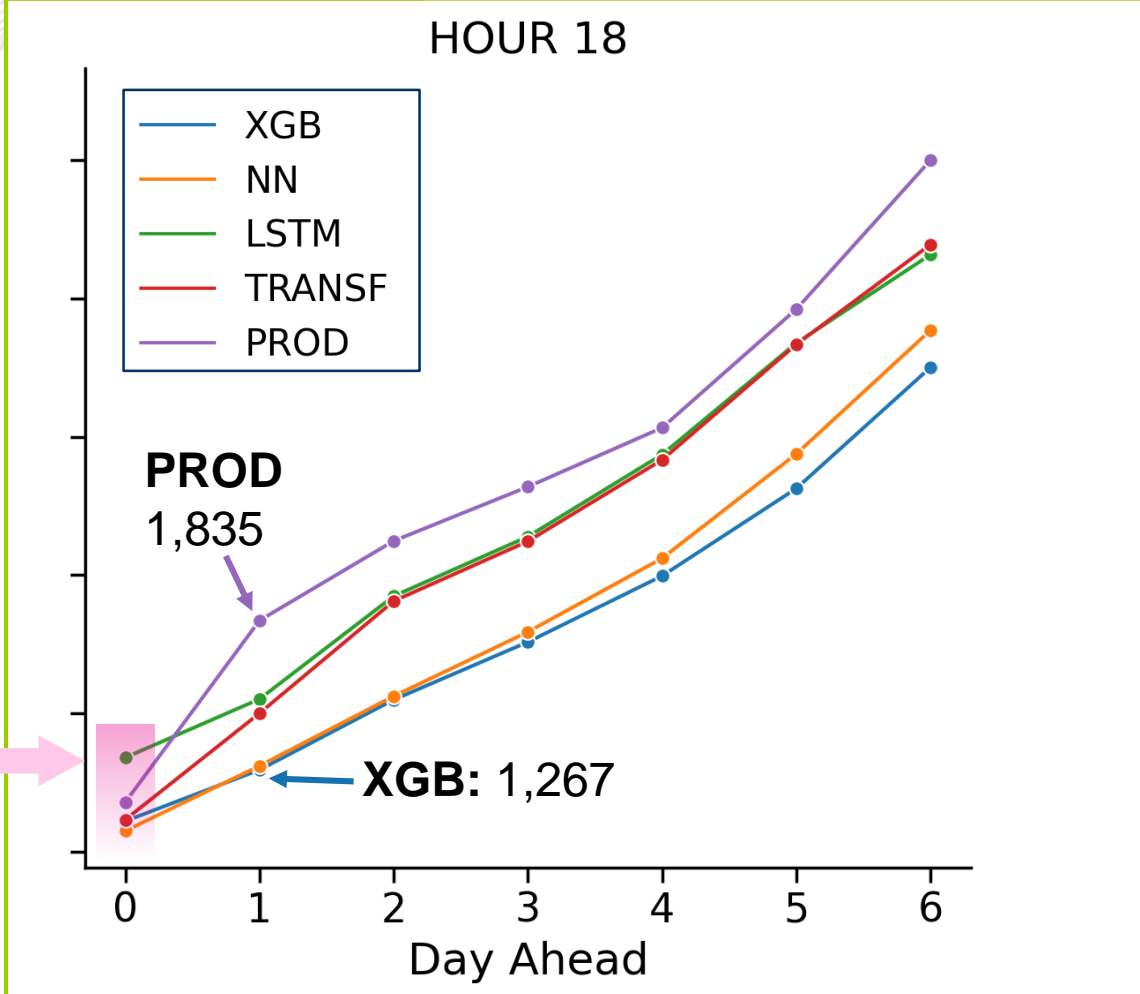
MW | Temperature during  
same hour last weekday

**Future information, current information, past-similar-hour information**

## Economic Benefit

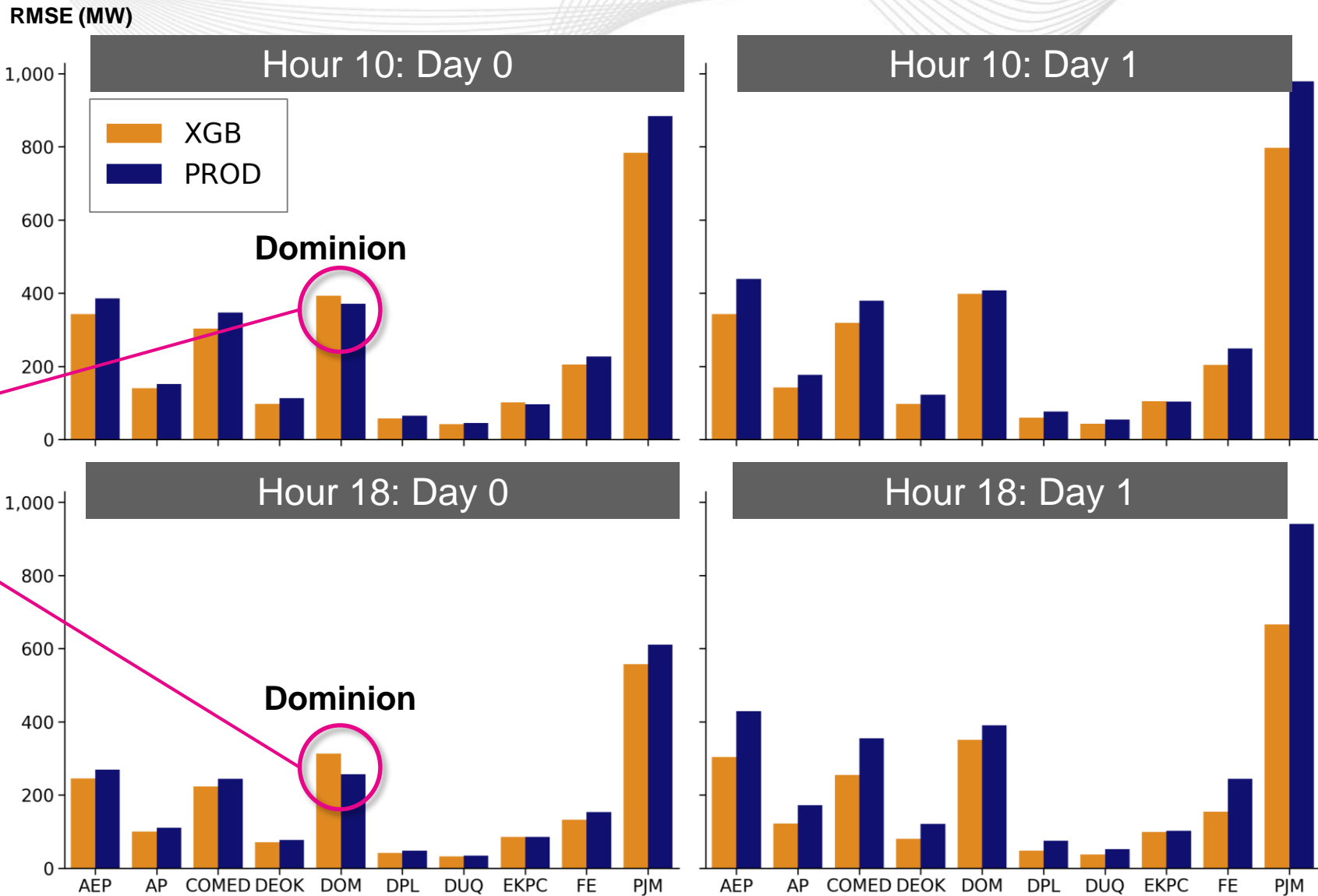


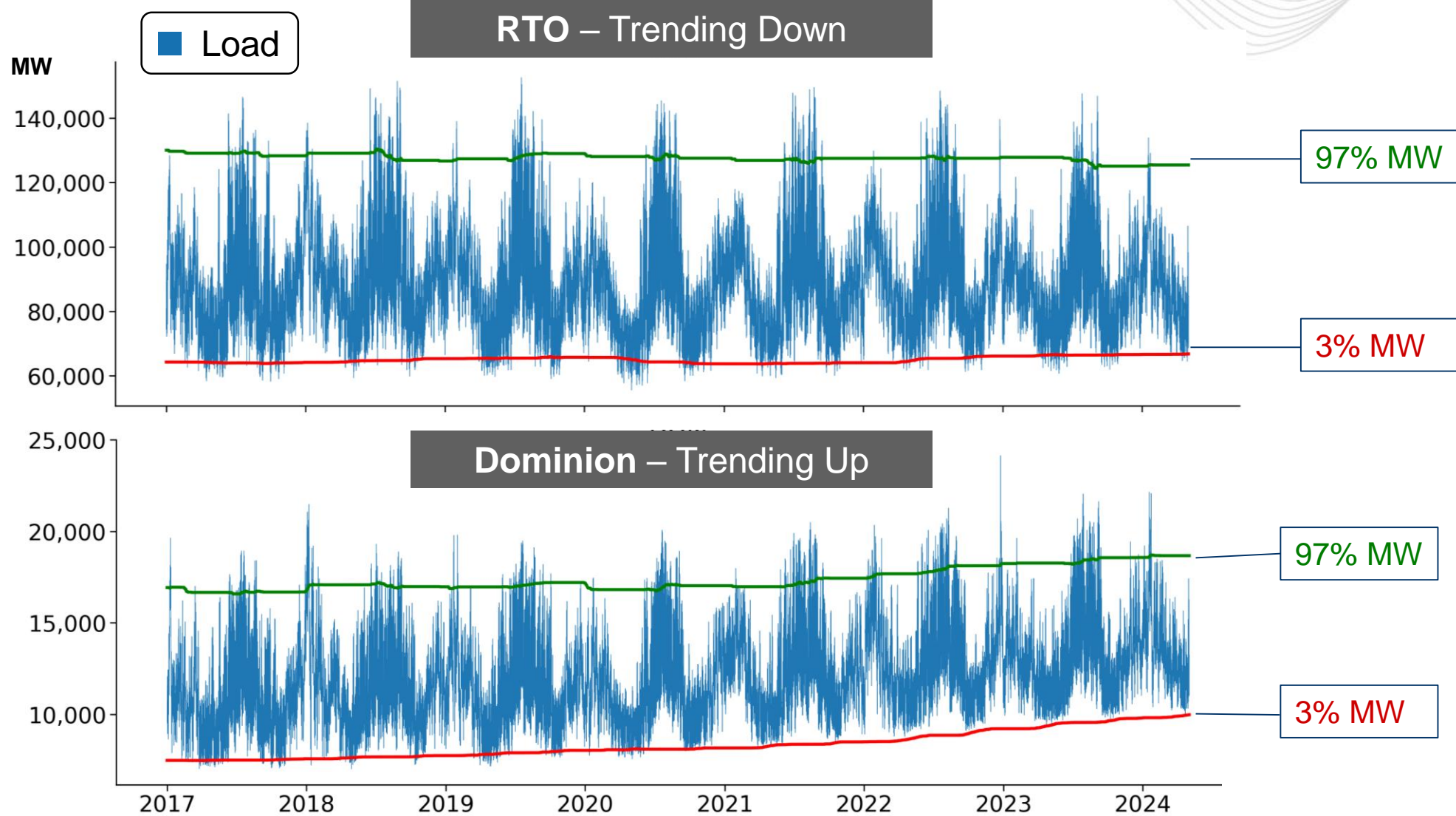
## Reliability Benefit





**XGBoost**  
Improves all large zones except Dominion.







# Dominion Data Centers

## Data Center Contracted Capacity vs. Metered Coincident Demand Dominion Energy Service Territory

**Connected Data Centers:**

**67**

*(35% worldwide)*

15 connected in **2023**

15 more in **2024**

**Hourly Average:**

**In 2022**

2.7 GW

**2023**

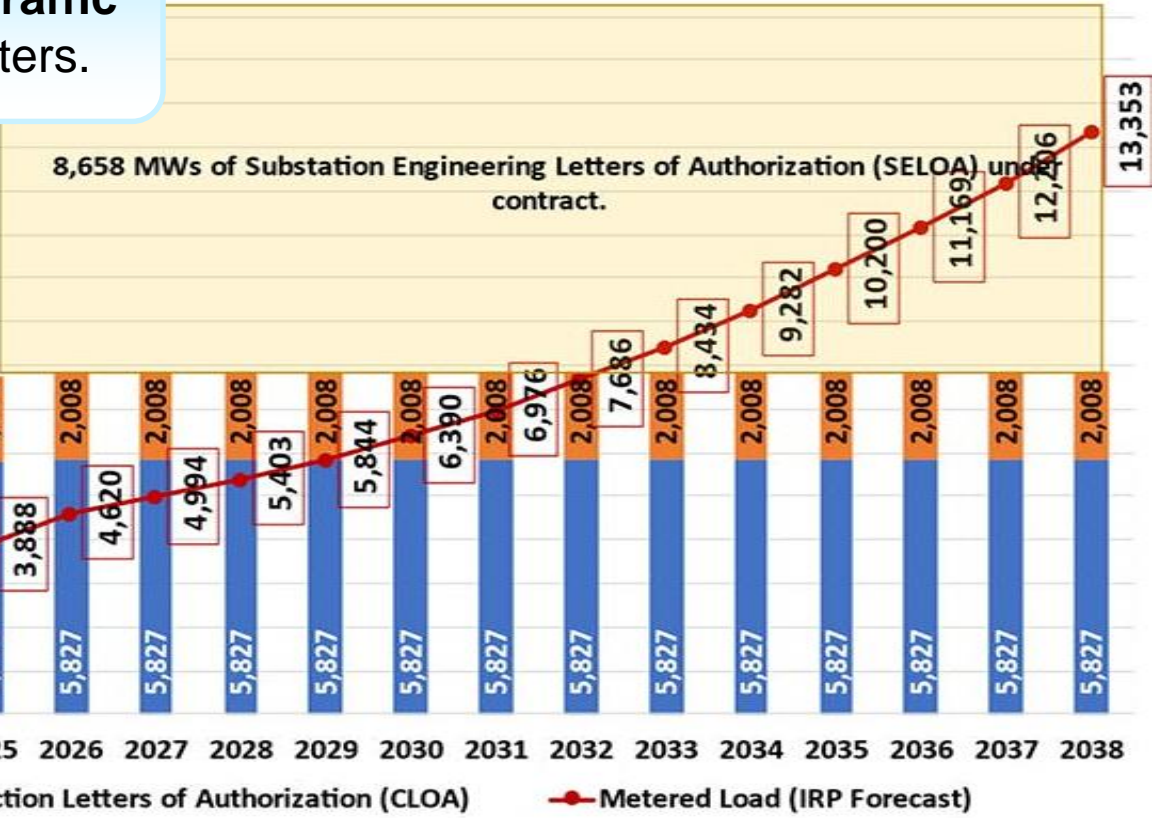
3.2 GW

Coincident Demand (MW) | Demand (MW)

18,000  
17,000  
16,000  
15,000  
14,000  
13,000  
12,000  
11,000  
10,000  
9,000  
8,000  
7,000  
6,000  
5,000  
4,000  
3,000  
2,000  
1,000  
0

**70%** of the world's internet traffic flows through Virginia data centers.

Dominion expects **10 GW** data center load by 2035.

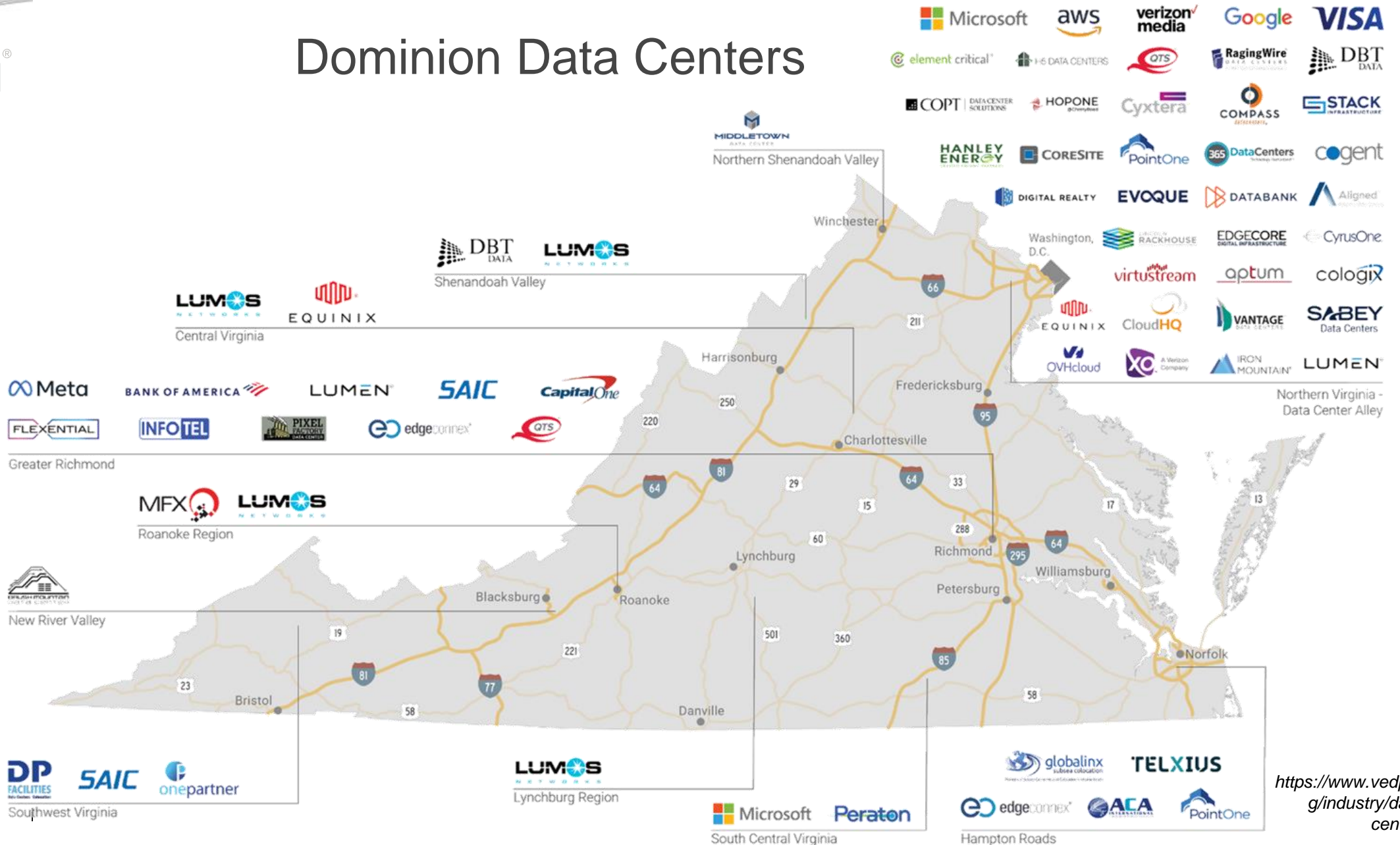


8,658 MWs of Substation Engineering Letters of Authorization (SELOA) under contract.

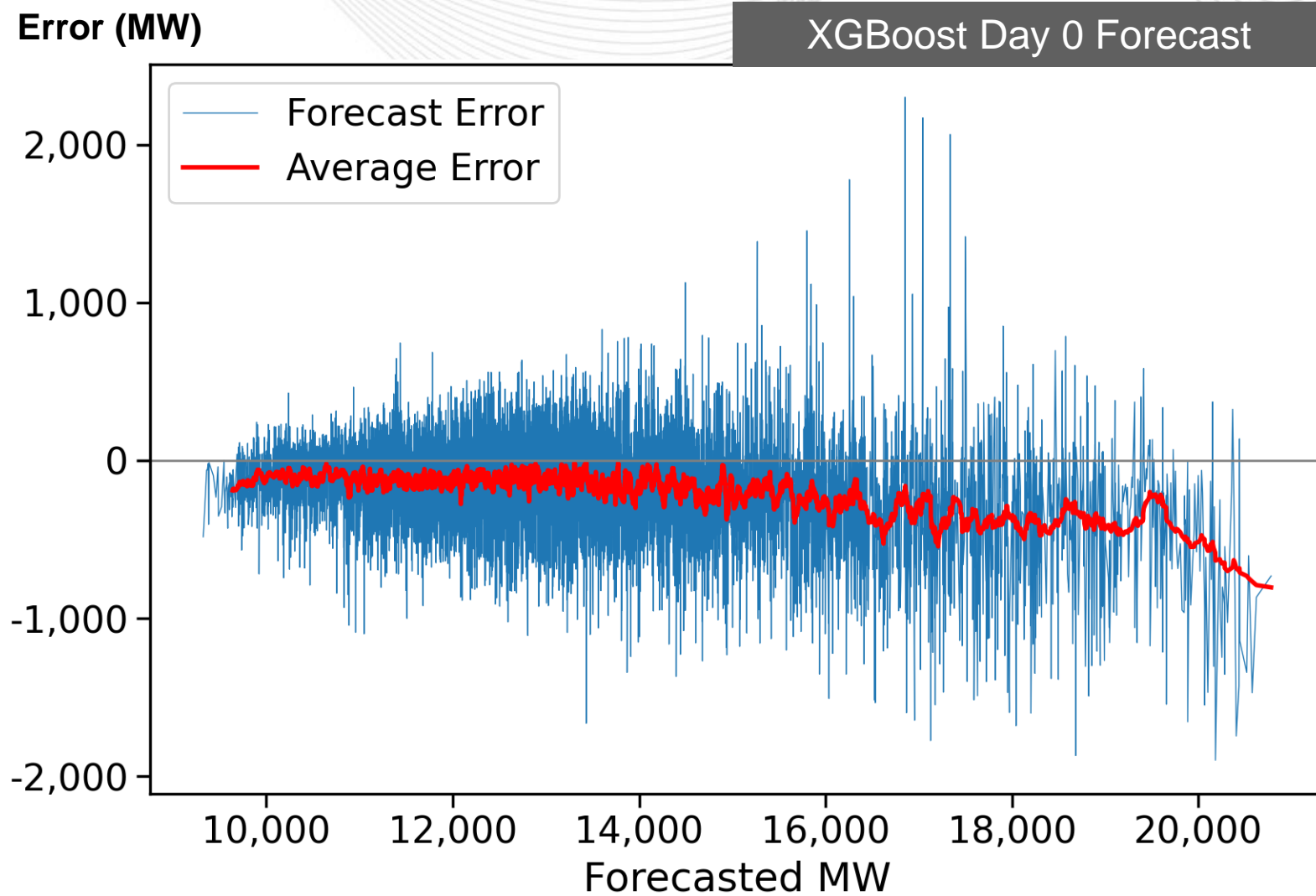
Note: The Company did not review ESAs prior to 2018 and assumed ESAs were equal to actual demand in 2017. Actual ESA totals will be higher than this assumption.

<https://www.datacenterfrontier.com/energy/article/33013010/dominion-virginias-data-center-cluster-could-double-in-size>

# Dominion Data Centers



<https://www.vedp.org/industry/data-centers>



**We use two learners:**



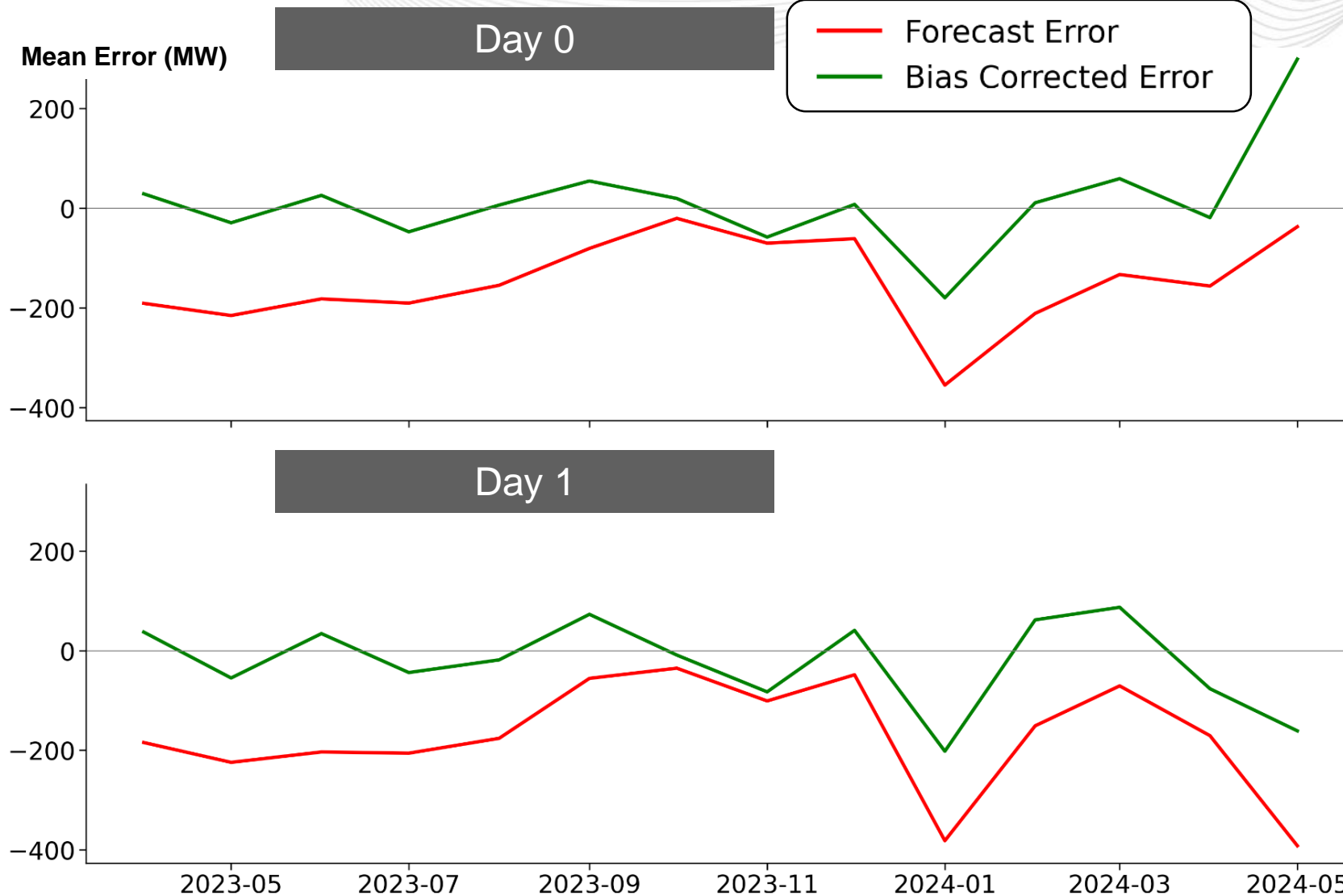
$$MW = MW(\text{Forecast}) + \text{Bias}$$

**MW**  
Learner

MW ~ Features, use efficient and accurate models

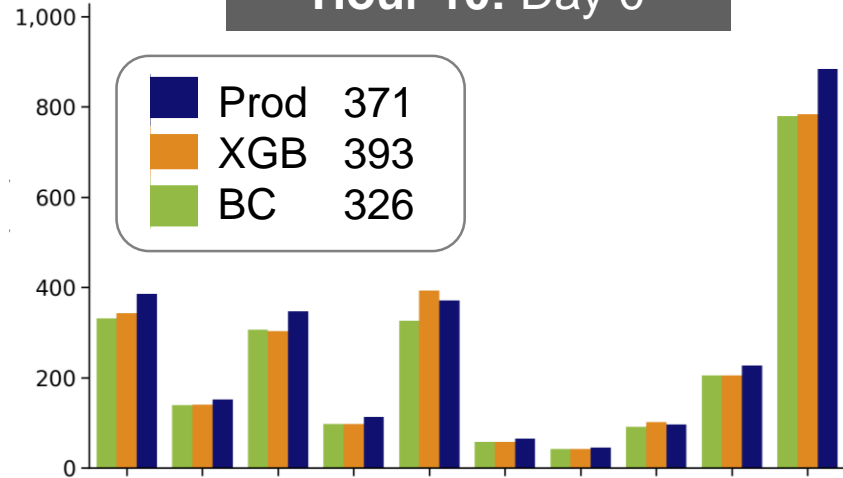
**BIAS**  
Learner

- We don't know true bias.
- Forecast Error ~ MW(Forecast) + Hour
- The idea is to use a weak learner to infer bias, XGBoost again.

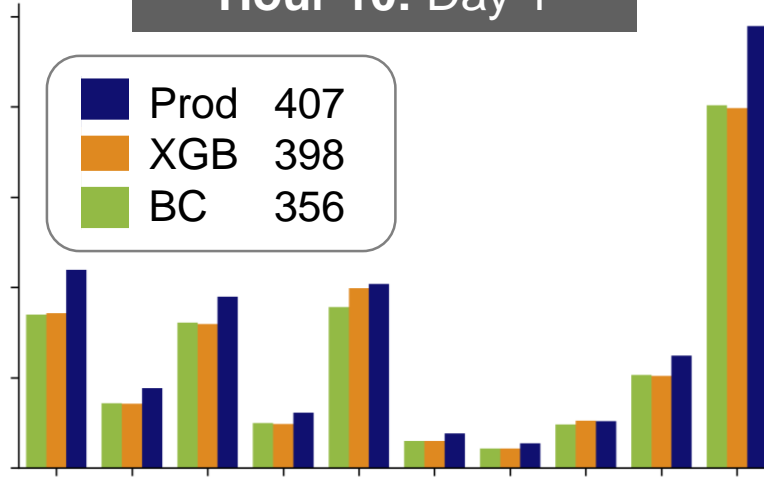


RMSE (MW)

Hour 10: Day 0

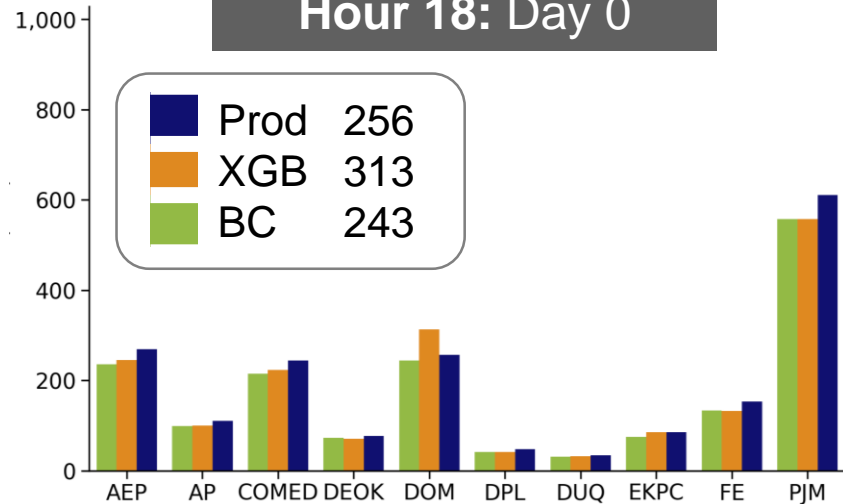


Hour 10: Day 1

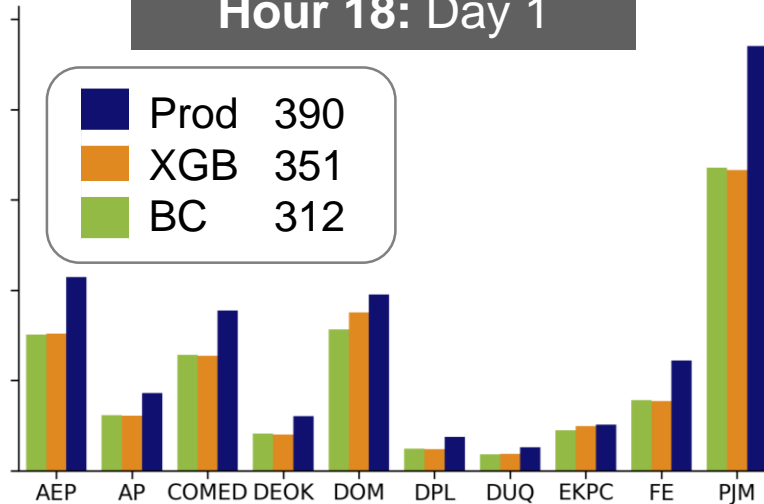


✔ Bias correction significantly improved forecast in Dominion.

Hour 18: Day 0



Hour 18: Day 1



✔ Bias correction did not impact forecast in other zones.

✔ Bias correction is robust.



We tested XGB, NN, LSTM, Transformer on hourly forecast.

Dominion data center load is challenging and will be more so in next few years.

Practically, XGB is significantly faster and slightly more accurate than others.

PJM is prepared to deal with the challenges economically.

Acknowledgement: Ms. Kexin Xie, PhD candidate, Department of Statistics, Virginia Tech