LOAD FORECASTING

Forecasting Data Center Load for Operations

ESIG 2025 Forecasting & Markets Workshop

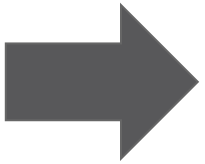
Marija Markovic, PhD
Power System Research Engineer, EPRI
MMarkovic@epri.com

Special thank you to my forecasting team – Cheng Lyu and David Larson

Data Center Loads Are Growing Fast

Current Industry Focus

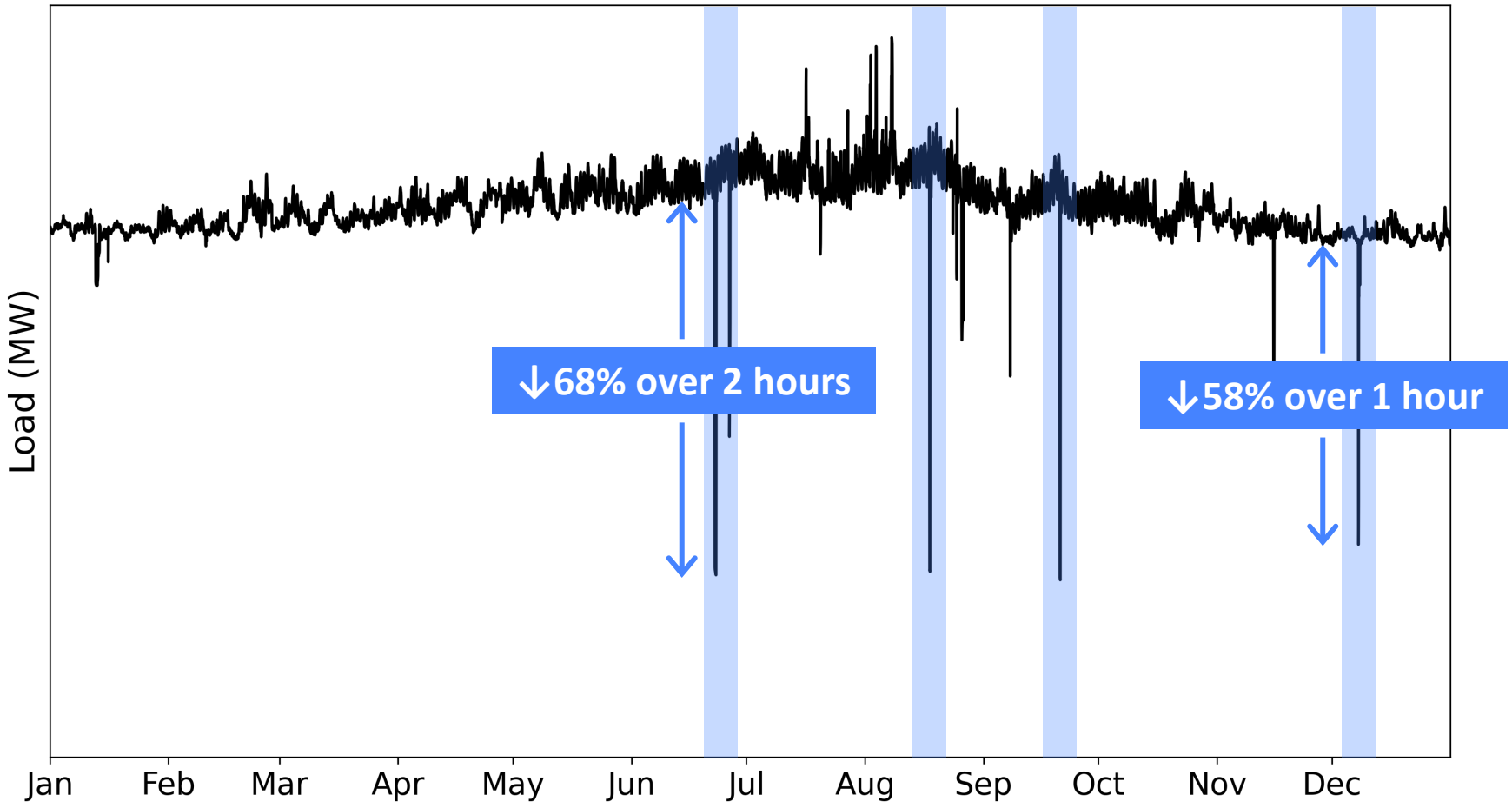
- Long-term growth projections for **planning**
  - Forecast horizon: **5 – 10 years** ahead
  - Spatial resolution: **system-level**
  - Temporal resolution: **annual peak**
- Guides **capacity expansion** and infrastructure investments



Next Industry Focus

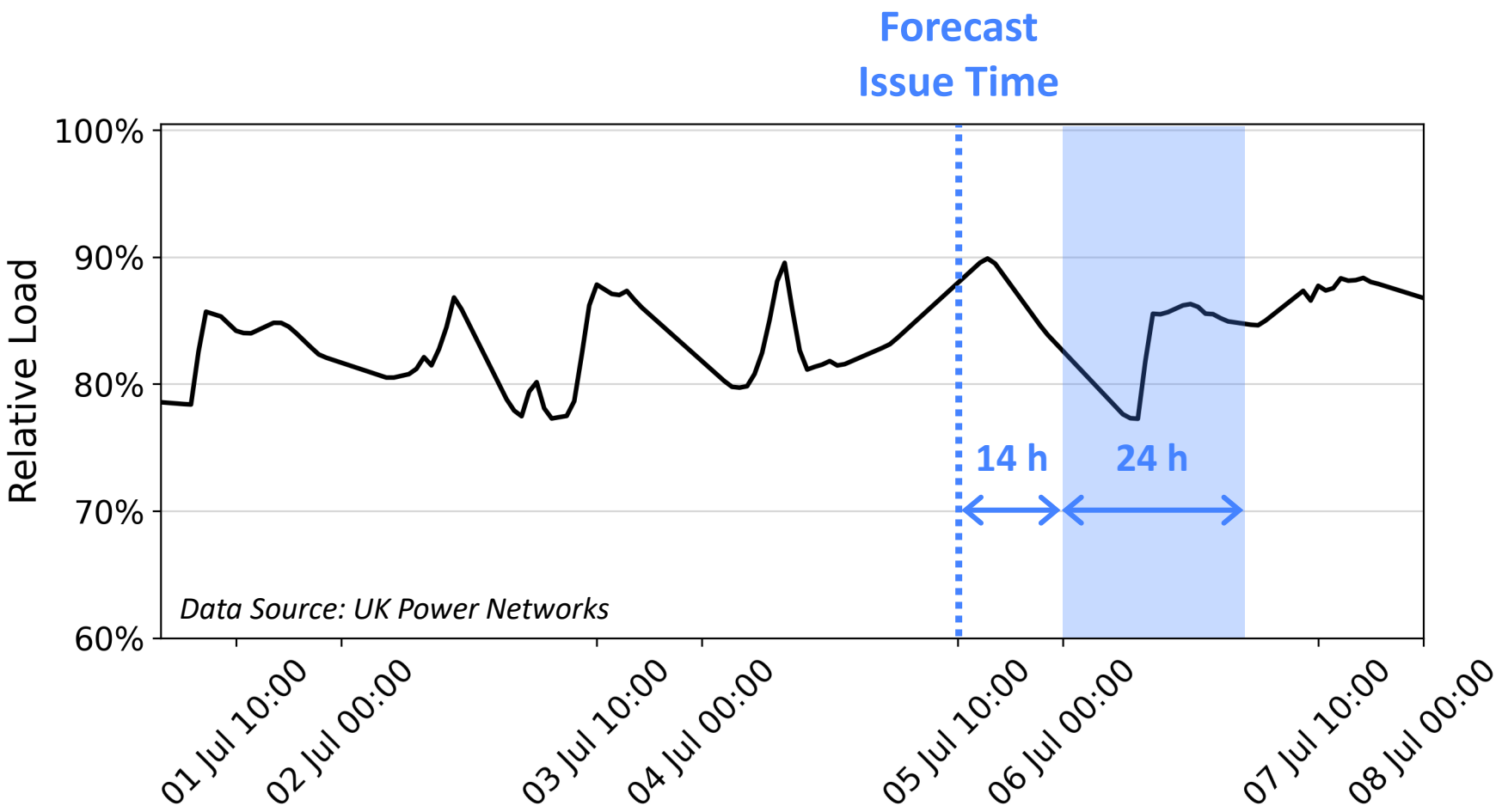
- Short-term forecasting for **operations**
  - Forecast horizon: **hours** to **days** ahead
  - Spatial resolution: **site-level** or **zonal**
  - Temporal resolution: **hourly power**
- Informs **operational decision-making**

Data Centers Are Not Like Other Loads



High load factor but *not* flat + sudden *MW-scale* spikes

Our Focus: Day-Ahead Forecasting of Data Center Load*



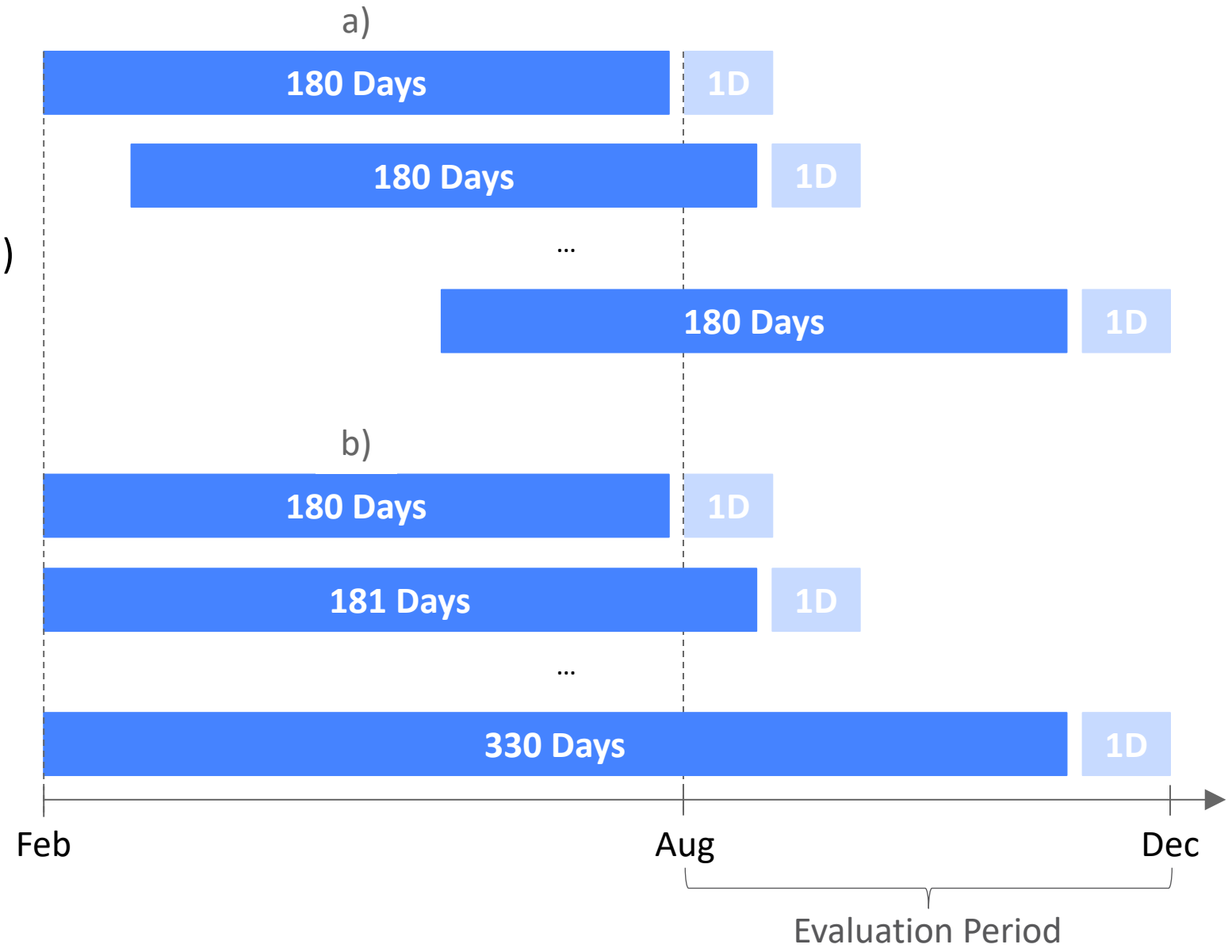
** Total power demand (kW or MW) measured at facility level*

Methodology

- **Day-ahead forecasting with walk-forward retraining**
  - a) Sliding window (fixed size)
  - b) Expanding window (cumulative)

train | test

- **Features**
  - Hourly lags
  - Calendar/time features
  - Rolling stats (mean/min/max) over past 1–2 days
  - Weather features ***not*** used



Beginning with simple persistence forecasts

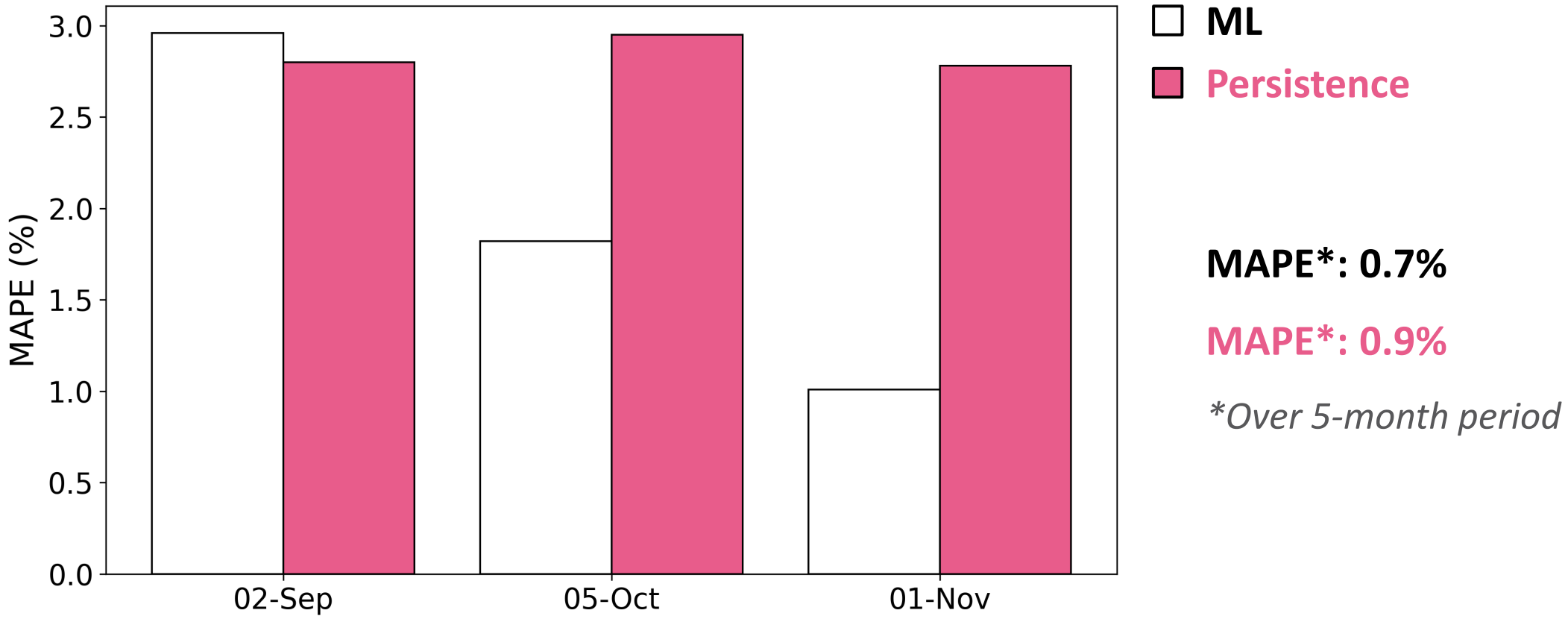
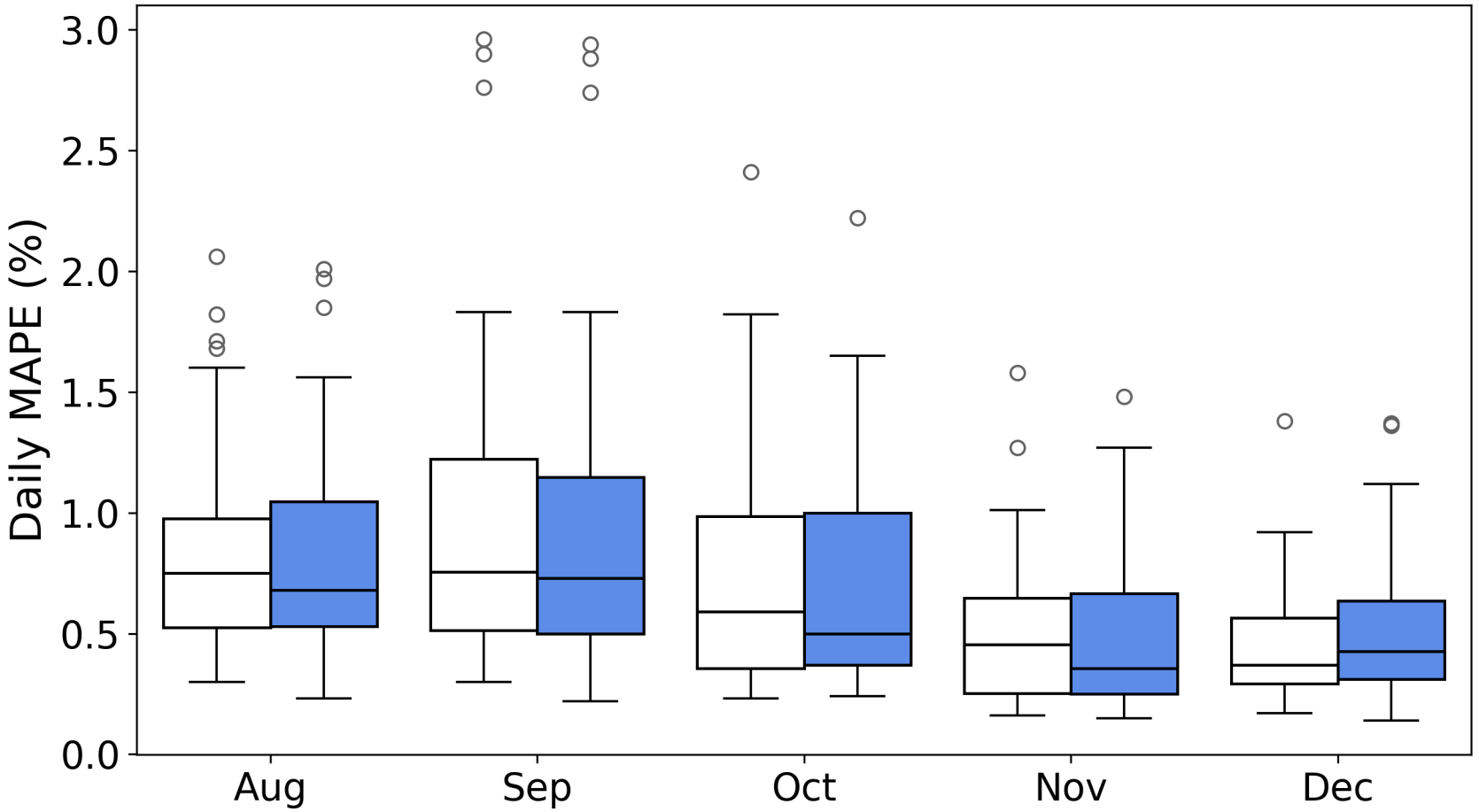


Figure: Comparing ML and Persistence on Difficult Forecast Days

ML performs better, but persistence sets a solid starting point

No need to wait a year to start forecasting



Fixed-size training window

Expanding training window

Total MAPE: 0.7%

Total MAPE: 0.7%

Comparable forecast accuracy with less data

All models show comparable MAPE – no clear best choice

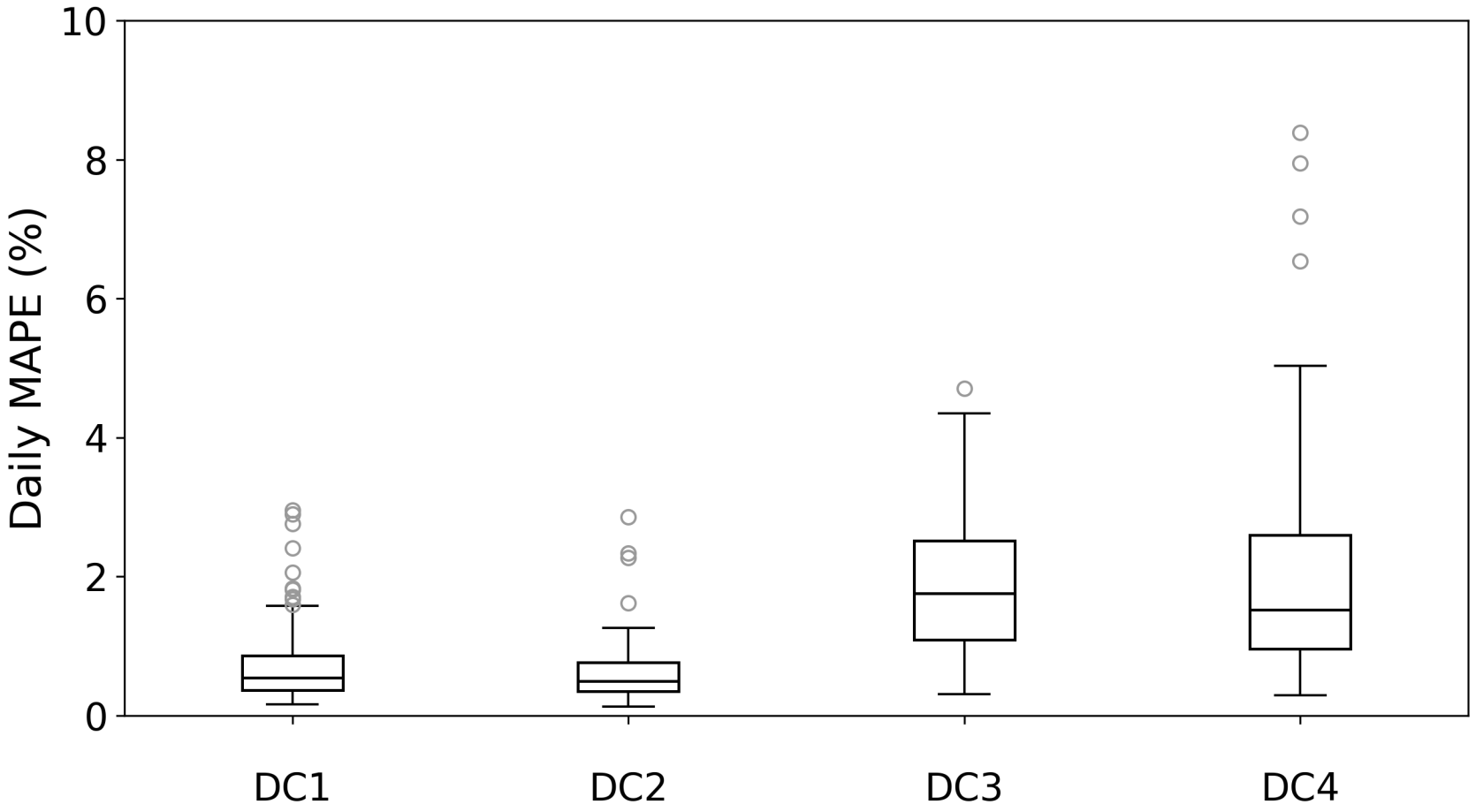
	DC1	DC2	DC3	DC4
XGBoost	0.8%	0.6%	2.1%	2.9%
Random Forest	**0.7%**	0.6%	**1.9%**	**2.4%**
Gradient Boosting	0.7%	**0.5%**	2.0%	2.6%

Datasets: 1 year of load data from four data centers (DC1–DC4) with diverse characteristics

Methodology: Sliding 1D window with training size fixed at 180 days

Evaluation Metric: MAPE (%) averaged over August–December

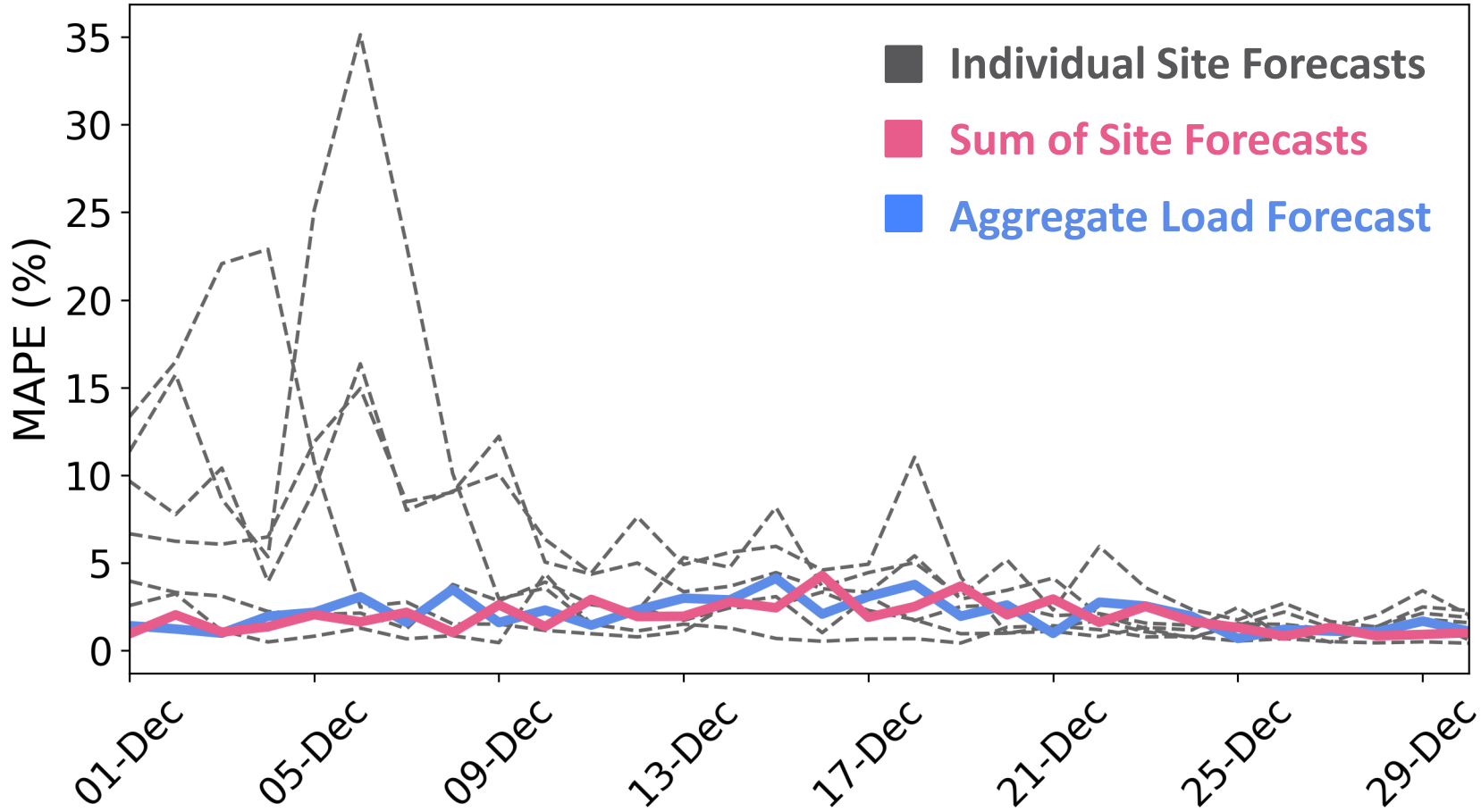
Data center characteristics matter more



Data Center	Load Factor (LF)
DC1	LF > 90%
DC2	LF > 90%
DC3	80% < LF < 90%
DC4	70% < LF < 80%

One model doesn't fit all – adapt to data centers accordingly

From site-level to zonal forecasting



Forecast accuracy improves with aggregation

Summary

1. Limited data yields accuracy comparable to larger sets
2. Data center characteristics impact performance more than model choice
3. Aggregate-level forecasting is more accurate than site-level

A deliverable on this topic is coming later this year — stay tuned!

Thank you!

TOGETHER…SHAPING THE FUTURE OF ENERGY®

www.epri.com