The Right Model for the Right Time: Applying Mixture of Experts for Electricity Price Forecasting

Nick Frowerk, Julian Sauerbier, Alexander Szimayer

University of Hamburg

26.03.2026

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Motivation

- Economic Relevance:
  - Electricity price forecasts drive operational decisions in generation scheduling, bidding, and storage optimization.
  - We target the German day-ahead auction market and evaluate forecasts through a Battery Energy Storage System (BESS) simulation.
- Renewable Integration:
  - The increasing share of wind and solar generation supports decarbonization but introduces supply-side stochasticity.
  - Weather dependence $\rightarrow$ supply fluctuations $\rightarrow$ higher price volatility.
- Hypothesis:
  - A Mixture of Experts (MoE) framework can improve forecast accuracy by assigning specialized models to distinct market states.

Related Literature & Research Gap

- Existing hybrid approaches combine clustering or sequential models to handle market heterogeneity:
  - **Panapakidis & Dagoumas (2016):** Partition training data via K-Means, then train cluster-specific ANNs. No explicit temporal feature engineering.
  - **Alkawaz et al. (2022):** Sequential hybrid where ARD captures trend/seasonality and Extra Tree Regression learns residual patterns. Single global architecture, no regime-specific experts.
- **Gap:** Neither framework combines time-series feature engineering with regime-specific expert selection.
- **Our approach:** Use a global SARIMA model as a feature generator, then assign cluster-specific ML experts via a Mixture of Experts architecture.

Data & Preprocessing

- **Source:** Energy Charts API (hourly frequency).
- **Period:** October 2018 – September 2025.
  - Training: Oct 2018 – Dec 2023.
  - Out-of-sample testing: Jan 2024 – Sep 2025.
- **Features:**
  - Exogenous forecasts: residual load, renewable share, and their temporal lags (24h, 48h, 168h).
  - Historical price lags: 24h, 48h, 168h (to prevent look-ahead bias).
- **Target:** 24 hourly day-ahead products, shifted into the positive domain and log-transformed.

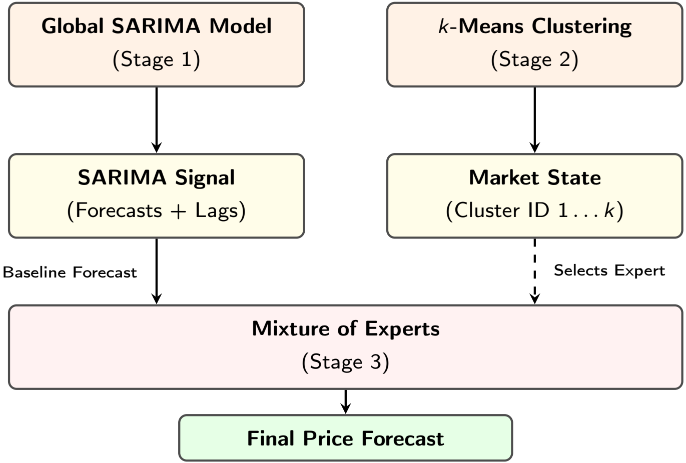
Proposed Methodology: Mixture of Experts (1/3)

- **Core Concept:** Use specialized experts for distinct market regimes instead of a single global model.
- **Stage 1: SARIMA Feature Engineering**
  - Fit a global SARIMA model on the full training period to generate baseline forecasts and temporal features
- **Stage 2: Regime Identification (K-Means Clustering)**
  - Cluster the dataset based on features like residual load, wind forecast, SARIMA forecast
  - Identify distinct market states

Proposed Methodology: Mixture of Experts (2/3)

- **Stage 3: Expert Training (Competition)**
  - For each cluster, train candidate models using the SARIMA forecast and its temporal lags as additional input features
  - Select the best-performing model as the *Cluster Expert*
- **Online Testing (Daily Recalibration)**
  - Assign new observations to a cluster
  - Predict using the associated expert, which incorporates the SARIMA signal alongside the initial feature set
  - Retrain experts daily as realized prices expand the training window

Proposed Methodology: Mixture of Experts (3/3)



Cluster Evaluation

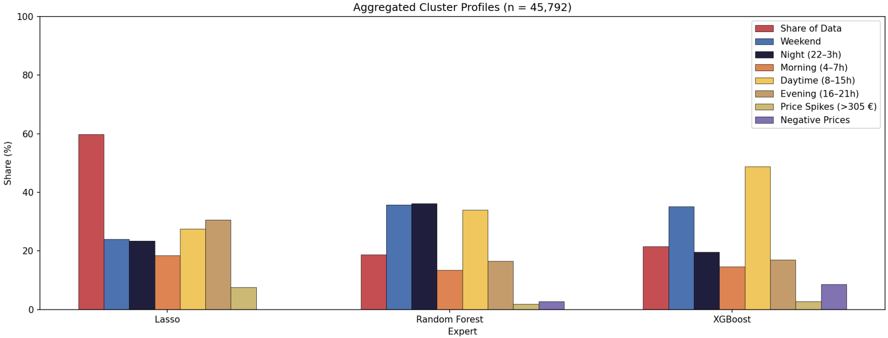


Figure 1: 32 *k*-means clusters aggregated by assigned expert. Lasso covers the majority of observations, Random Forest handles low-volatility off-peak periods and XGBoost captures high-volatility states including negative prices.

Evaluation Framework: Statistical vs. Economic

- Statistical Accuracy:
  - Mean Absolute Error (MAE).
- Economic Value:
  - Feed predicted prices into a BESS optimization model to generate a day-ahead bidding schedule.
  - Compare the profit based on our forecasts ($\pi_{Forecast}$) against the theoretical maximum profit under perfect foresight ($\pi_{Perfect}$).
  - Metric: The Forecasting Efficiency Ratio (FER):

$$\text{FER} = \frac{\pi_{Forecast}}{\pi_{Perfect}} \quad \in [0, 1]$$

Results (1/3)

Model	MAE (EUR/MWh)	FER
MoE (Proposed)	**14.60**	**94.49%**
XGBoost	16.81	93.20%
Random Forest	13.66	95.86%
Lasso	16.16	95.44%

Table 1: Out-of-sample comparison (Jan 2024 – Sep 2025). Benchmarks use the same SARIMA feature engineering but a single global expert (k=1)

Results (2/3)

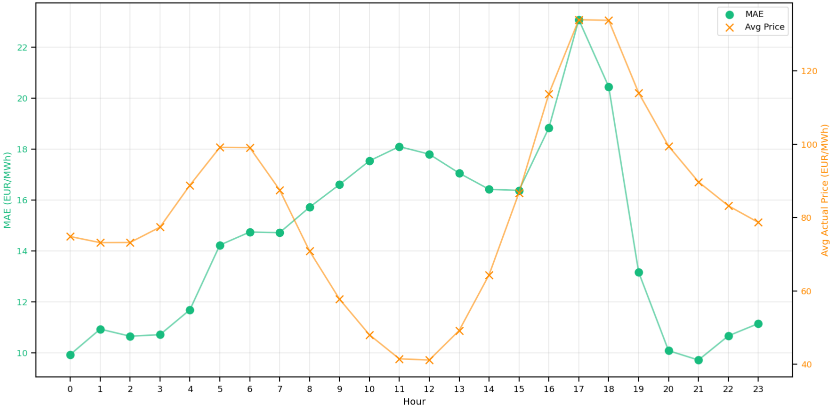


Figure 2: Hourly MAE of the MoE alongside average day-ahead prices.

Results (3/3)

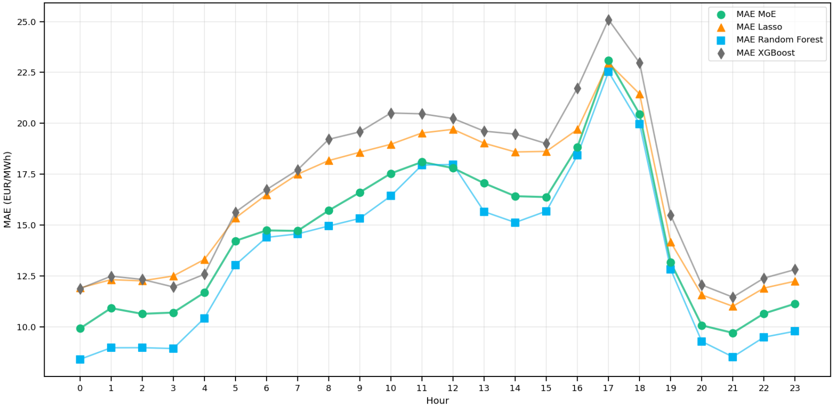


Figure 3: Hourly MAE of the MoE and the three benchmark models.

Conclusion

- The MoE framework achieves a **FER of 94.49%** and an **MAE of 14.60 EUR/MWh** on the German day-ahead market.
- Random Forest (k=1) achieves a lower MAE (13.66) and higher FER (95.86%) without clustering, outperforming the MoE on both metrics.
- MAE and FER do not rank models identically: Lasso achieves a higher FER than the MoE despite a worse MAE. For BESS scheduling, the temporal structure of forecast errors matters beyond their magnitude.
- **Limitation:** SARIMA parameters remain fixed at training-window estimates (Oct 2018 – Dec 2023). Structural shifts after this period are not captured.
- **Next step:** Periodic re-estimation of SARIMA parameters during the out-of-sample period.

References

Panapakidis, I. P., & Dagoumas, A. S. (2016). Day-ahead electricity price forecasting via the application of artificial neural network based models. *Applied Energy*, 172, 132–151. https://doi.org/10.1016/j.apenergy.2016.03.089

Alkawaz, A. N., Abdellatif, A., Kanesan, J., Khairuddin, A. S. M., & Gheni, H. M. (2022). Day-ahead electricity price forecasting based on hybrid regression model. *IEEE Access*, 10. https://doi.org/10.1109/ACCESS.2022.3213081