

What Does It Take to Generalize in Electricity Price Forecasting?

Dr. Jochen L. Cremer
Associate Professor

www.Jochen-Cremer.com

<https://www.tudelft.nl/ai/delft-ai-energy-lab>

DELFT AI ENERGY LAB

Research

- Synthetic data for grid operating conditions
- Power flow solvers and theory
- ML for dynamic stability & steady state use cases
- Graph-based ML for grids & markets

Education

- EE4C12 ML for Electrical Engineering
- SET 3125 Machine Learning Workflows for Digital Energy Systems
- SC42150 Statistical Signal Processing
- SC42110 Dynamic Programming and Stochastic Control
- Industry training

Team



Jochen Cremer



Peyman Mohajerin Esfahani



Haiwei Xie



Olayiwola Arowolo



Daniel Wieser



Mert Karaçelebi



Selin Özcan



Basel Morsy



Benoît Jeanson



Ali Rajaei



Luca Hofstadler



Jochen Stiasny



Viktor Zobernig



Maosheng Yang



Paul Bannmüller



Betül Mamudi



Runyao Yu

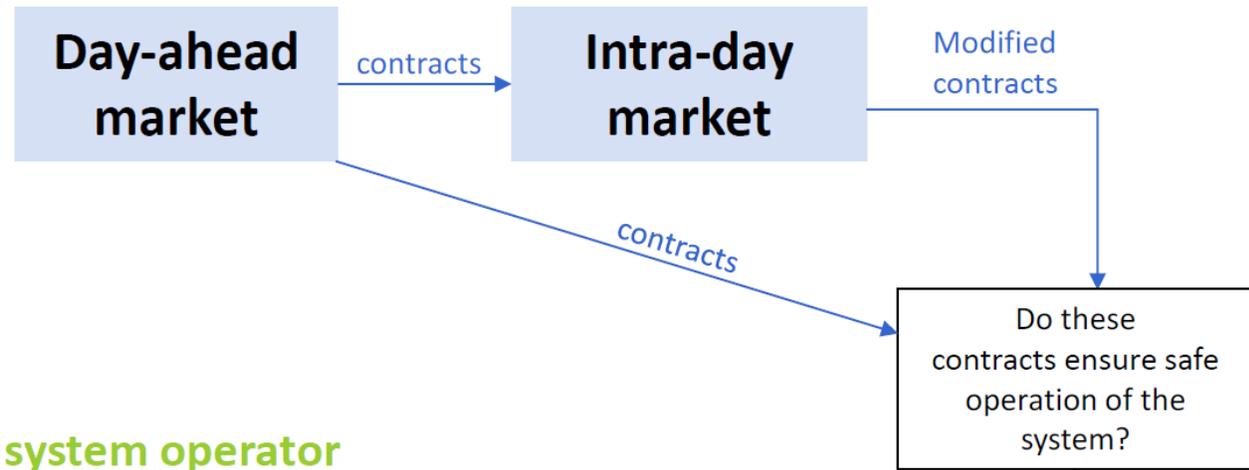


Kutay Bölat



Example of market mechanisms

By the market operator



By the transmission system operator



*Slide from Jalal Kazempour, DTU

Challenges: Energy transition & electrification

The grid is partly at capacity. (No new connections to the grid possible)



Limits Dutch economic growth
Getting worse in the future

Grid challenge

- Electrification: power peaks challenge the grid -> higher asset maintenance
- Renewable integration: (Price) uncertainty increases

Market solution

- Flatten peaks by trading demand and generation shifts ('flexibility')
- Increase predictability

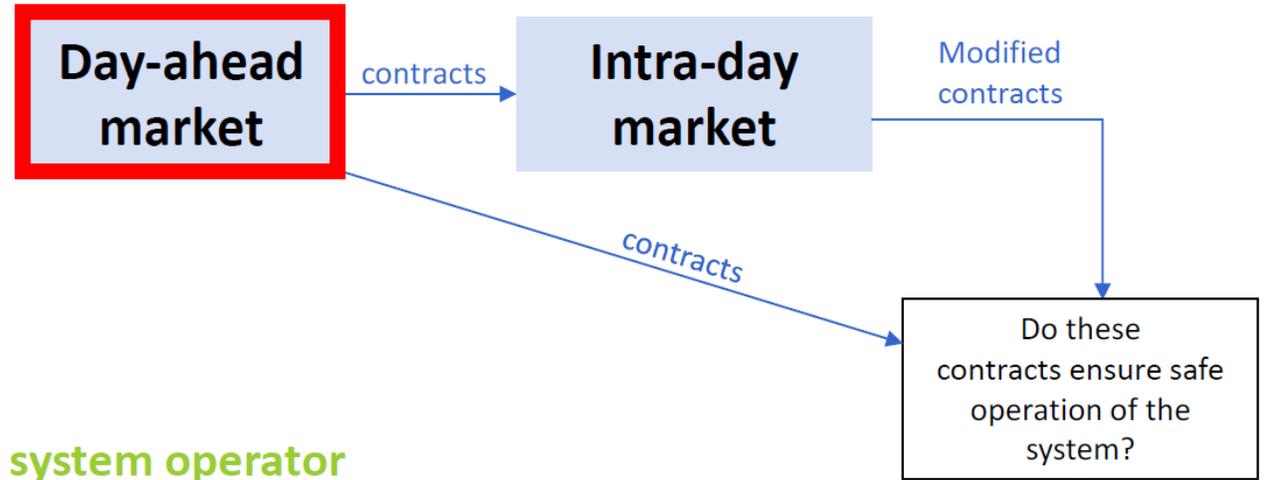


Runyao Yu

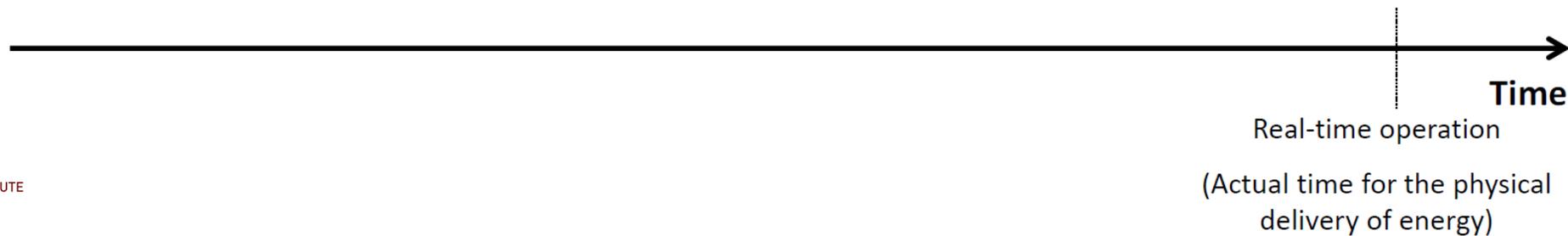
Example of market mechanisms

By the market operator

Can we increase predictability with a Foundation Model?

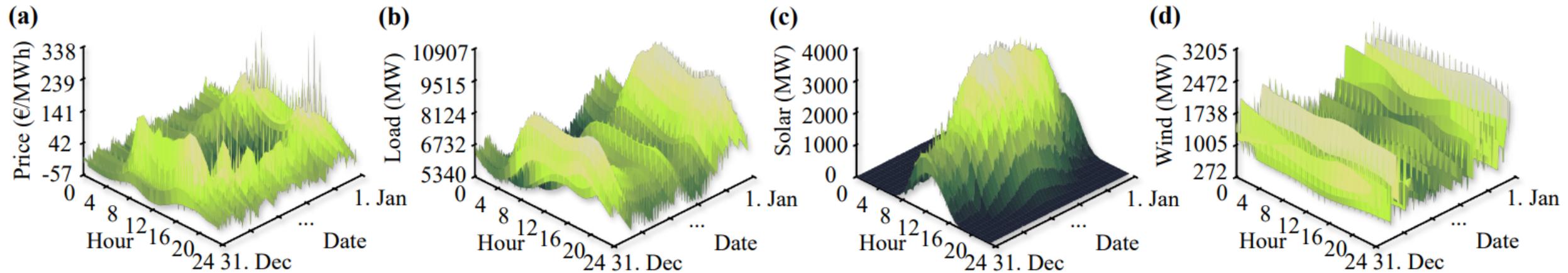


By the transmission system operator



Day-ahead price forecasting

- Auction-style market
- Data are **regular** time grids (e.g., EU switching to 15-min trajectories)
- Price strongly driven by **fundamentals** (load, solar, wind)
- Prediction targets are day-ahead prices for the next day



European-level energy data in 2025, averaged across regions. **(a)** Electricity price. **(b)** Forecasted load. **(c)** Forecasted solar power generation. **(d)** Forecasted wind power generation (onshore and offshore).

Foundational Model

Objective: learn reusable representations across scenarios

Assessment: Zero-Shot and Full-Shot

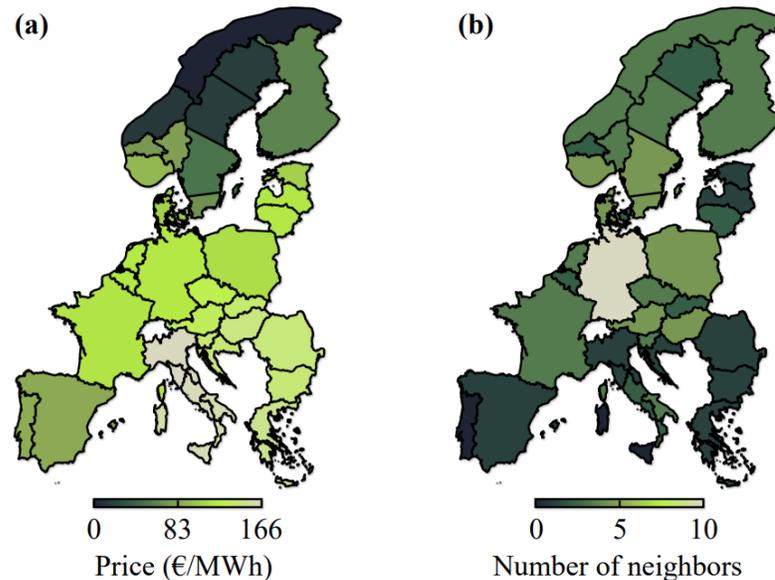
Existing Foundation Models: Chronos, Moirai, TimeMoE, TimesFM, etc.

Model	Multivariate Input	Probabilistic Output
Chronos		✓
Chronos [2.0]	✓	✓
Moirai [S]	✓	✓
Moirai [M]	✓	✓
Moirai [L]	✓	✓
TimeMoE		
TimesFM [2.0]		✓
TimesFM [2.5]	✓	✓

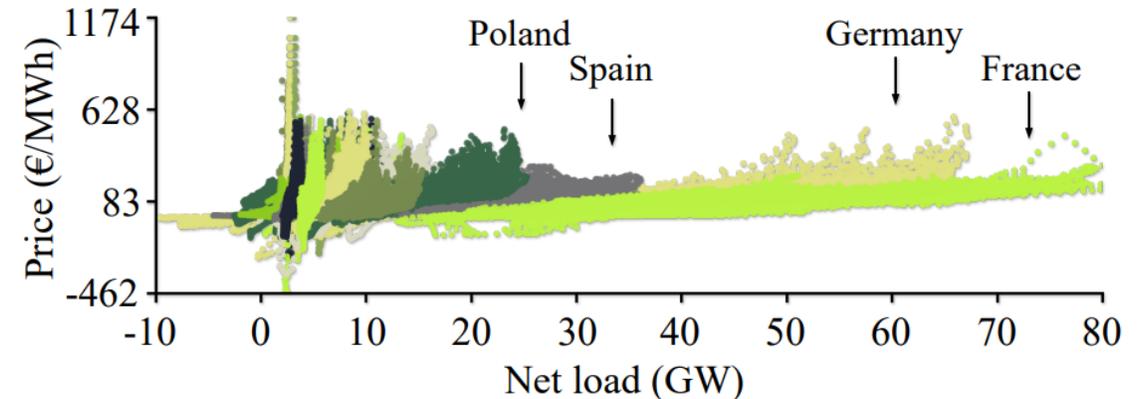
- These models are pretrained on multiple datasets for general time-series forecasting tasks
- Day-ahead & intraday have special characteristics, which calls for domain-specific foundation models

Day-Ahead: Scale + Topology

- Pretrain on 38 European regions (2022–2026) to capture diverse regimes and cross-region structure
- The grid induces a natural graph: each region has neighbors via transmission lines
- Regions exhibit distinct price–netload behaviors → motivates a model with region-specific embeddings



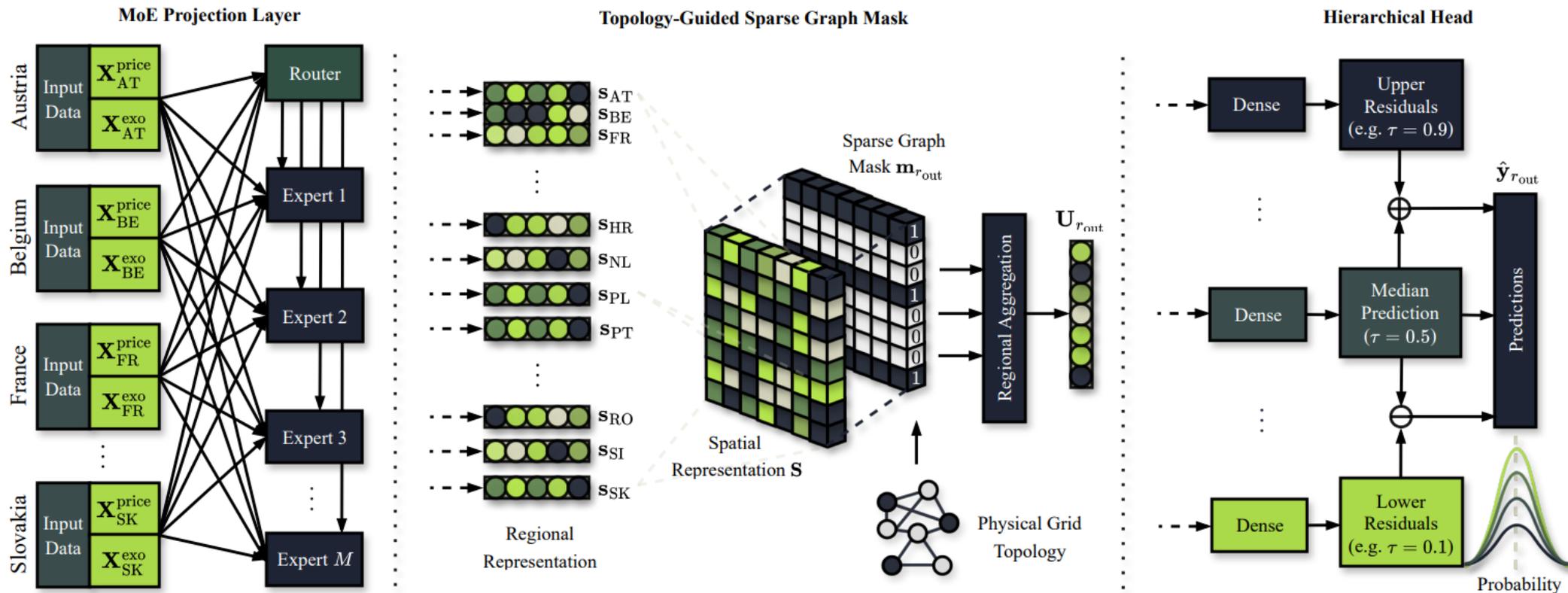
(a) Electricity prices for 38 European regions averaged from 2022 to 2026. (b) Number of directly connected neighboring regions via transmission lines



Price–net load relationship across European regions (Net load = load – solar – wind)

Day-Ahead Model: PriceFM

- MoE projection: map each region's inputs into a shared latent space while preserving heterogeneity
- Sparse graph mask: restrict message passing using grid topology to control noisy long-range transfer
- Hierarchical quantile heads: generate multiple quantiles with stable, non-crossing behavior



Results: Zero-shot accuracy

- Consistently strong calibration: low Calibration Error (**AQCE**) with narrow interval width (**AIW**)

Model	AQL ↓ $\frac{\text{€}}{\text{MWh}}$	AQCR ↓ [%]	AQCE ↓ [%]	AIW ↓ $\frac{\text{€}}{\text{MWh}}$	RMSE ↓ $\frac{\text{€}}{\text{MWh}}$	MAE ↓ $\frac{\text{€}}{\text{MWh}}$	R ² ↑
CHRONOS	11.14	0.00	9.20	33.79	42.19	25.98	0.12
CHRONOS [2.0]	<u>8.03</u>	0.00	7.59	<u>26.01</u>	<u>30.93</u>	<u>19.44</u>	<u>0.48</u>
MOIRAI [S]	11.24	0.00	7.67	38.00	43.66	27.22	0.07
MOIRAI [M]	12.07	0.00	<u>7.03</u>	42.00	47.94	30.47	-0.12
MOIRAI [L]	11.94	0.00	9.14	37.15	46.66	29.13	-0.07
TIMEMOE	-	-	-	-	40.83	25.54	0.16
TIMESFM [2.0]*	10.50	0.00	<u>5.43</u>	40.83	41.91	26.01	0.15
TIMESFM [2.5]*	<u>7.97</u>	0.00	7.62	<u>25.98</u>	<u>30.83</u>	<u>19.48</u>	<u>0.48</u>
PRICEFM	6.85	0.00	5.30	25.69	26.13	16.83	0.55
PRICEFM*	6.91	0.00	5.39	25.88	26.24	16.90	0.55

PRICEFM* is evaluated using the same quantile set as TimesFM*

Results: Full-shot accuracy

- Consistently strong calibration: low Calibration Error (**AQCE**) with narrow interval width (**AIW**)

Model	AQL $\downarrow \frac{\text{€}}{MWh}$	AQCR $\downarrow [\%]$	AQCE $\downarrow [\%]$	AIW $\downarrow \frac{\text{€}}{MWh}$	RMSE $\downarrow \frac{\text{€}}{MWh}$	MAE $\downarrow \frac{\text{€}}{MWh}$	R ² \uparrow
NAIVE ¹	15.29	0.00	11.34	108.02	34.68	22.06	0.33
NAIVE ²	15.35	0.00	11.80	103.56	34.31	23.31	0.35
NAIVE ³	15.46	0.00	12.26	101.26	32.61	22.64	0.39
FEDFORMER	8.22	15.33	8.34	25.20	31.75	20.15	0.38
PATCHTST	8.06	18.21	7.92	24.99	31.59	20.20	0.39
ITRANSFORMER	8.24	13.96	8.11	24.65	32.11	21.03	0.38
TIMESNET	7.98	13.42	8.03	23.99	30.94	19.48	0.40
TIMEXER	8.30	14.77	9.02	25.23	31.88	21.94	0.39
GCN	<u>6.61</u>	6.88	7.91	23.76	25.97	16.81	0.53
GAT	7.13	10.33	8.44	24.96	26.11	17.00	0.51
GRAPHSAGE	6.78	6.01	<u>7.44</u>	24.18	26.03	17.56	0.53
GRAPHDIFFUSION	6.69	<u>5.72</u>	<u>7.94</u>	<u>23.50</u>	25.93	<u>16.44</u>	0.54
GRAPHARMA	6.72	6.03	8.00	23.55	<u>25.84</u>	16.56	<u>0.55</u>
PRICEFM	5.80	0.00	5.25	21.27	22.39	14.28	0.60

NAIVE¹: Uses 96 prices from the previous day

NAIVE²: Uses 96 prices averaged over the past three days

NAIVE³: Uses 96 prices averaged over the past seven days

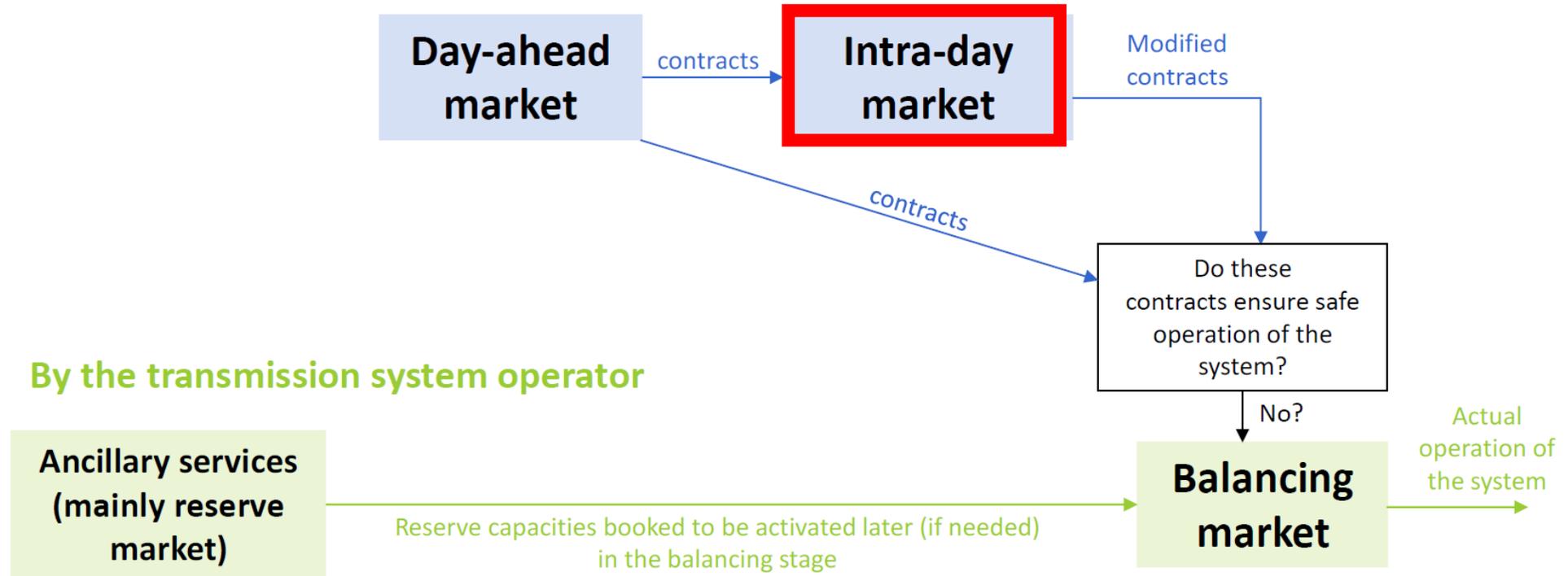


Runyao Yu

How about Intra-day markets?

By the market operator

Can we increase predictability with a Foundation Model?



By the transmission system operator



Time

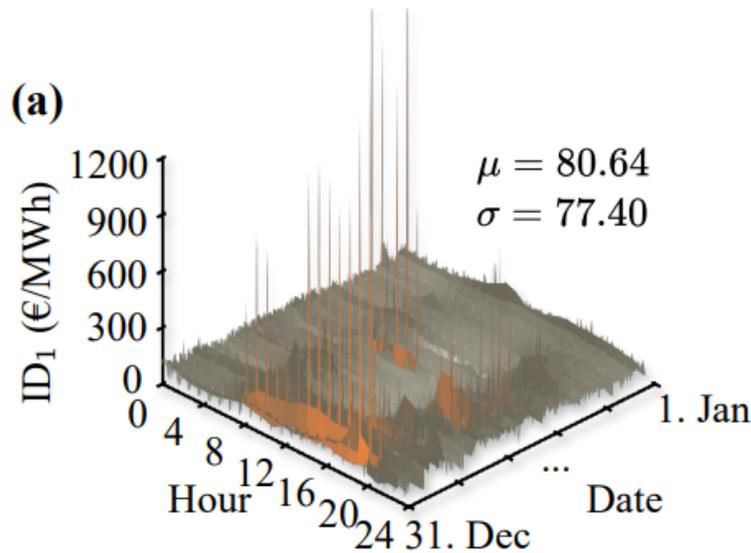
Real-time operation

(Actual time for the physical delivery of energy)

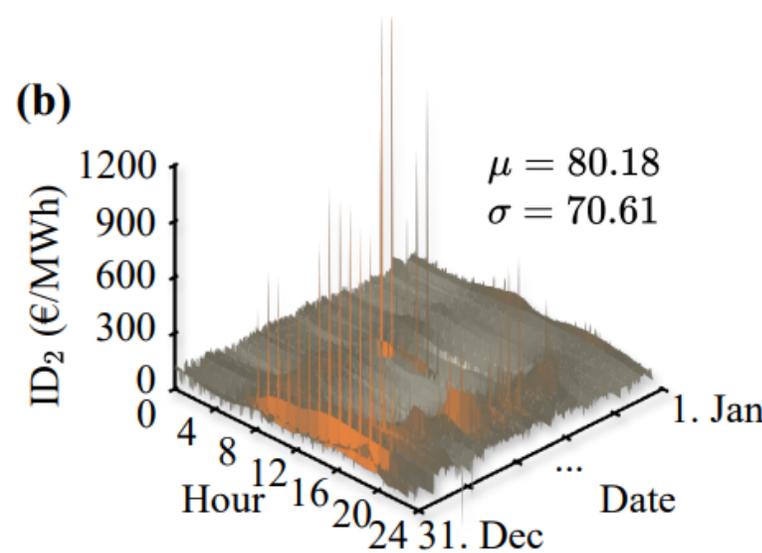
(Continuous) Intraday price forecasting

- Continuous trading
- Strong predictive power from **irregular microstructure orderbook**
- Prediction targets are often intraday indices defined by EPEX Spot: **ID1, ID2, ID3**

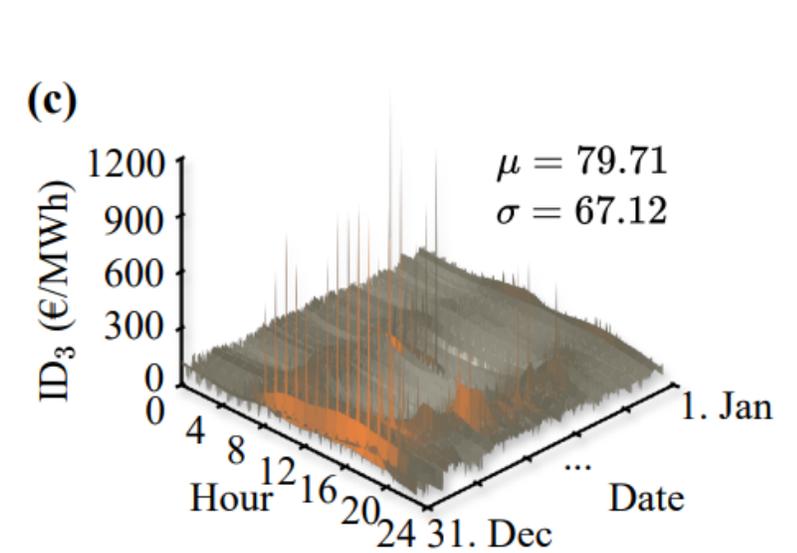
$$ID_x = \frac{\sum_{s \in S} \sum_{t \in \mathcal{T}_f} P_t^{(s)} V_t^{(s)}}{\sum_{s \in S} \sum_{t \in \mathcal{T}_f} V_t^{(s)}}$$



ID1 Submission: 15:00-15:30 (D-1)



ID2 22:00-22:30 (D-1)



ID3 10:00-10:30 (D)

Intraday price indices in the German market (2024). Orange areas indicate extreme electricity price. **(a)**, ID1 displays frequent price spikes, reflecting last-minute trading under imbalance pressure. **(b)**, ID2 reflects mid-session adjustments. **(c)**, ID3 corresponds to the most liquid trading window

Order book snapshot example

Product: **18:00–18:15 (delivery interval)**

Time: **09:42 (same day, before ID3 finishes)**

$$X_i^{(s)} = \begin{bmatrix} P_{t_1}^{(s)} & V_{t_1}^{(s)} & \nabla t_1 \\ P_{t_2}^{(s)} & V_{t_2}^{(s)} & \nabla t_2 \\ \vdots & \vdots & \vdots \\ P_{t_j}^{(s)} & V_{t_j}^{(s)} & \nabla t_j \\ \vdots & \vdots & \vdots \\ P_{T_i}^{(s)} & V_{T_i}^{(s)} & \nabla t_{T_i}^{(s)} \end{bmatrix}$$

Buy

<u>Price (€/MWh)</u>	<u>Volume (MW)</u>	<u>Cum. Volume</u>
84.5	6	6
83.0	10	16
82.5	8	24
81.0	15	39
80.0	20	59

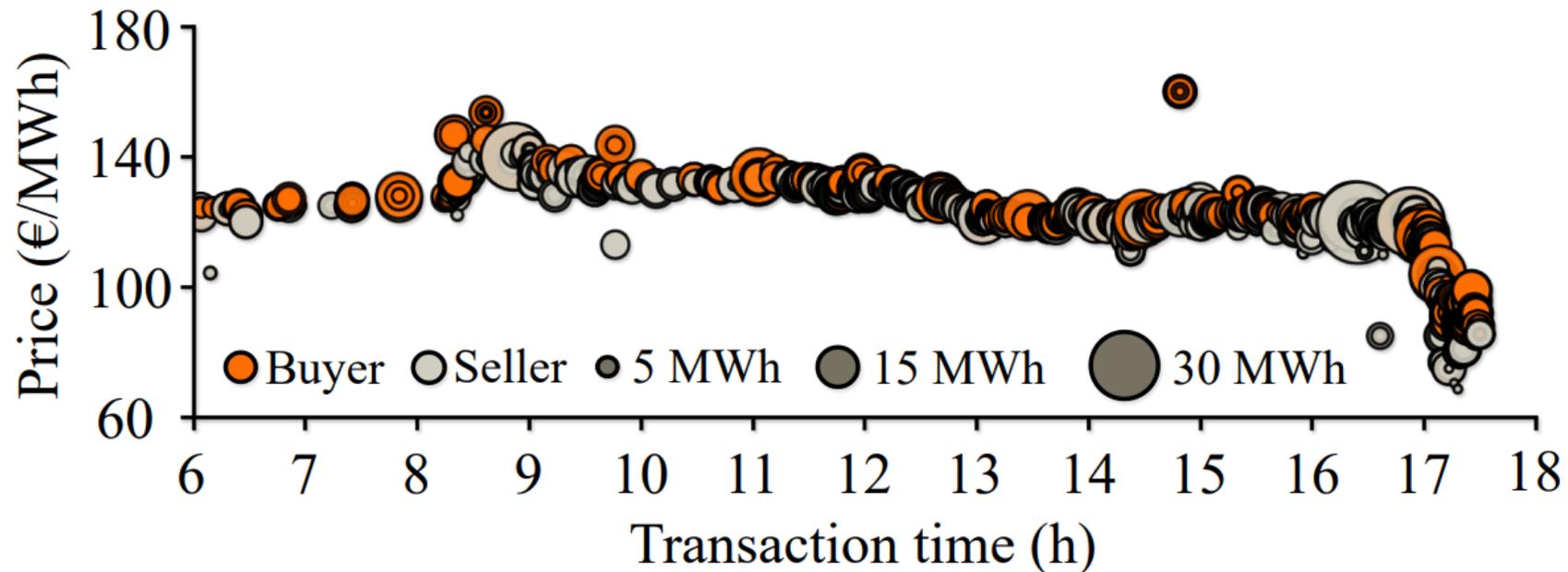
Sell

<u>Price (€/MWh)</u>	<u>Volume (MW)</u>	<u>Cum. Volume</u>
85.5	5	5
86.0	12	17
87.5	10	27
89.0	18	45
92.0	25	70

Good liquidity as we see tight spread (e.g. 84.5 / 85.5)

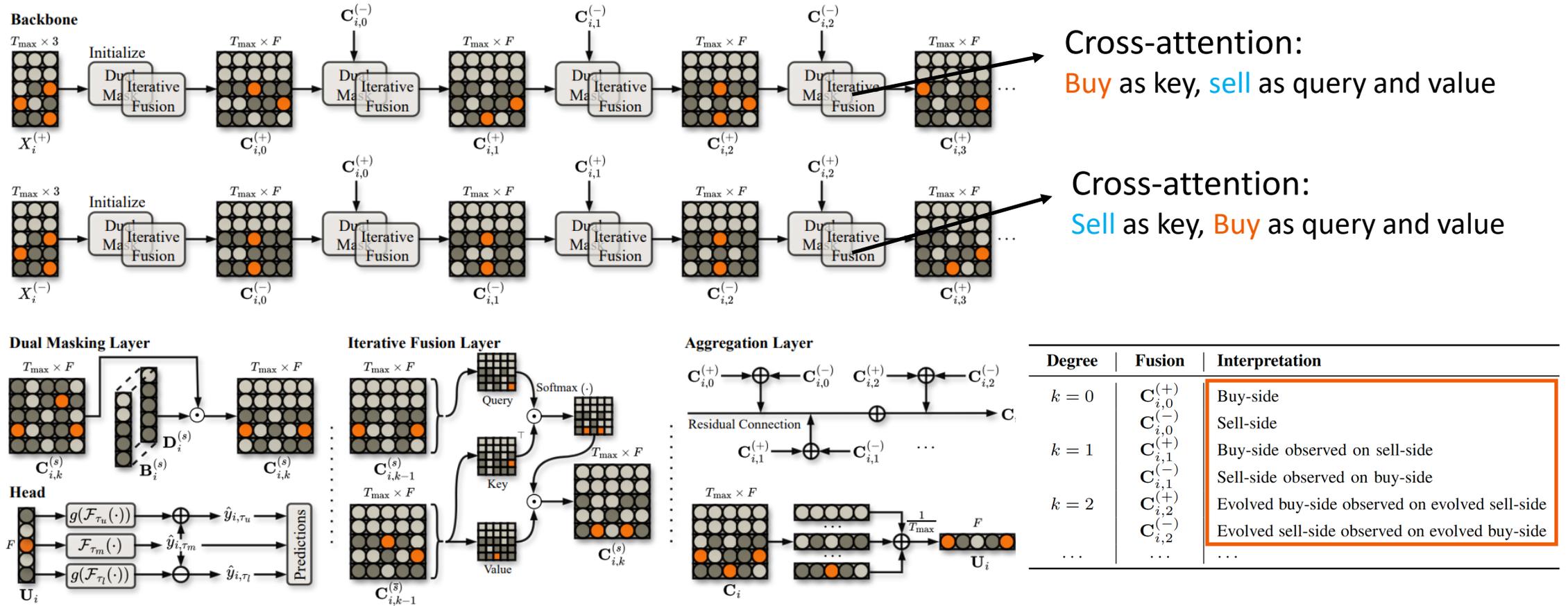
Interaction is Key

- Intraday prices emerge from strategic **buy–sell interactions**
- **Buy-sell interactions** for delivery at 18:00 on 2024-07-23
- Buyers and sellers adjust their bids and offers iteratively based on the opposite side, reflecting strategic buy-sell interactions



Model: OrderFusion

- Dual masking: handle variable-length trading streams (padding validity + temporal structure)
- Iterative fusion: repeatedly fuse buy-side and sell-side representations to model interaction degrees



Results

- Germany: outperforming strong baselines (e.g., TIMESNET 4.38±0.06)

$$AQL = \frac{1}{N|Q|} \sum_{i=1}^N \sum_{\tau \in Q} L_{\tau}(y_i, \hat{y}_{i,\tau}),$$

$$L_{\tau}(y_i, \hat{y}_{i,\tau}) = \begin{cases} \tau \cdot (y_i - \hat{y}_{i,\tau}), & \text{if } y_i \geq \hat{y}_{i,\tau}, \\ (1 - \tau) \cdot (\hat{y}_{i,\tau} - y_i), & \text{otherwise.} \end{cases}$$

Model	AQL ↓ $\frac{\text{€}}{MWh}$	AQCR ↓ [%]	AQCE ↓ [%]	AIW ↓ $\frac{\text{€}}{MWh}$	RMSE ↓ $\frac{\text{€}}{MWh}$	MAE ↓ $\frac{\text{€}}{MWh}$	R ² ↑
NAIVE ¹	6.34	0.00	8.12	29.00	40.11	14.75	0.70
NAIVE ²	16.17	0.00	10.10	79.27	85.66	36.62	-0.34
NAIVE ³	15.83	0.00	11.98	83.45	76.63	36.62	-0.06
MLP LQR ¹	4.72	<u>0.19</u>	<u>8.11</u>	18.82	34.21	11.43	0.78
MLP LQR ²	4.87	0.46	8.16	18.95	45.46	11.89	0.59
MLP LQR ³	4.61	0.54	8.13	17.11	32.23	<u>10.96</u>	<u>0.79</u>
FEDFORMER	4.70	3.96	8.94	17.02	34.49	11.52	0.77
ITRANSFORMER	4.66	3.77	8.72	16.94	33.99	11.55	0.77
PATCHTST	4.40	2.99	8.81	16.22	32.34	11.27	0.78
TIMESNET	<u>4.38</u>	2.57	8.77	<u>16.05</u>	<u>32.20</u>	10.98	0.78
TIMEXER	4.53	2.87	8.86	16.66	33.01	11.32	0.77
OrderFusion	3.81	0.00	8.09	13.34	26.84	9.06	0.82

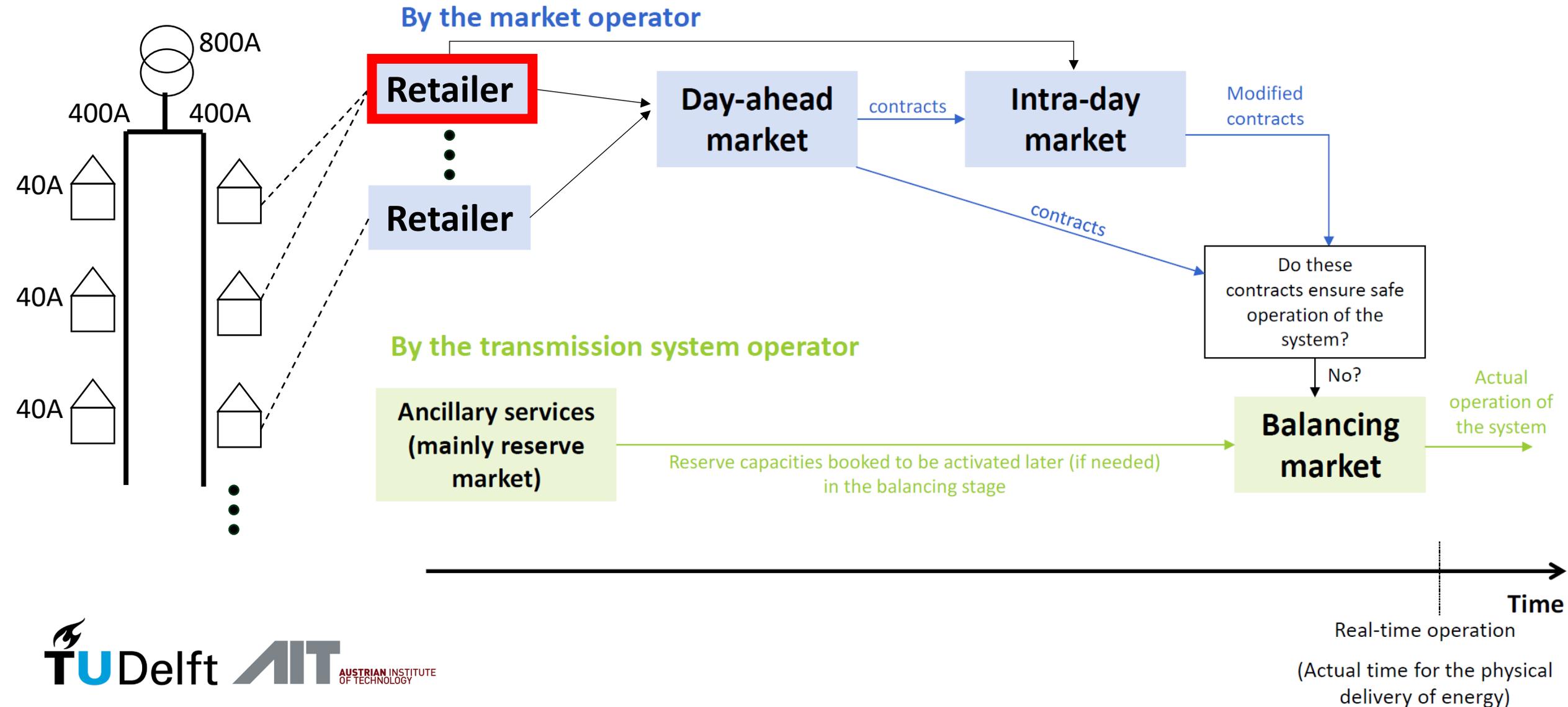
Superscript ^{1,2,3} denoted the inclusion of domain features

NAIVE¹: The price index from the most recent delivery hour is used as the NAIVE point forecast

NAIVE²: The price index from the same delivery hour on the previous day is used as the NAIVE point forecast

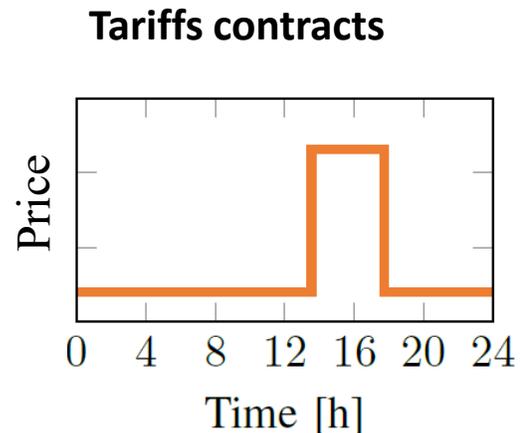
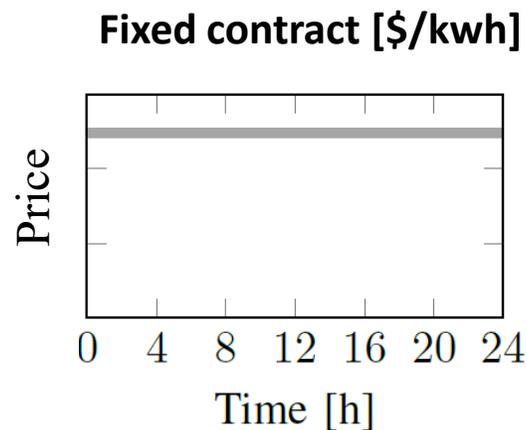
NAIVE³: The price index from the same delivery hour, averaged over the past 3 days, is used as the NAIVE point forecast

How about retail markets?

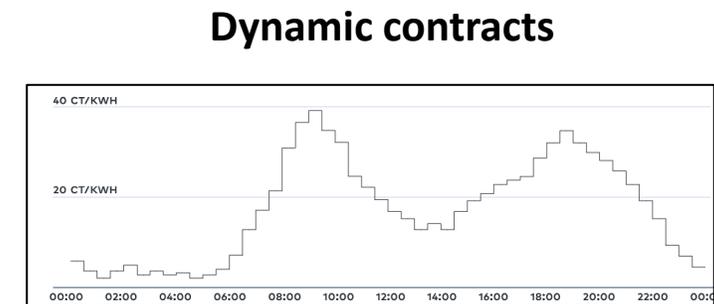


Demand Side Flexibility currently limited

1. Standard contracts don't incentive shifting energy usage behaviours



On-peak and off-peak

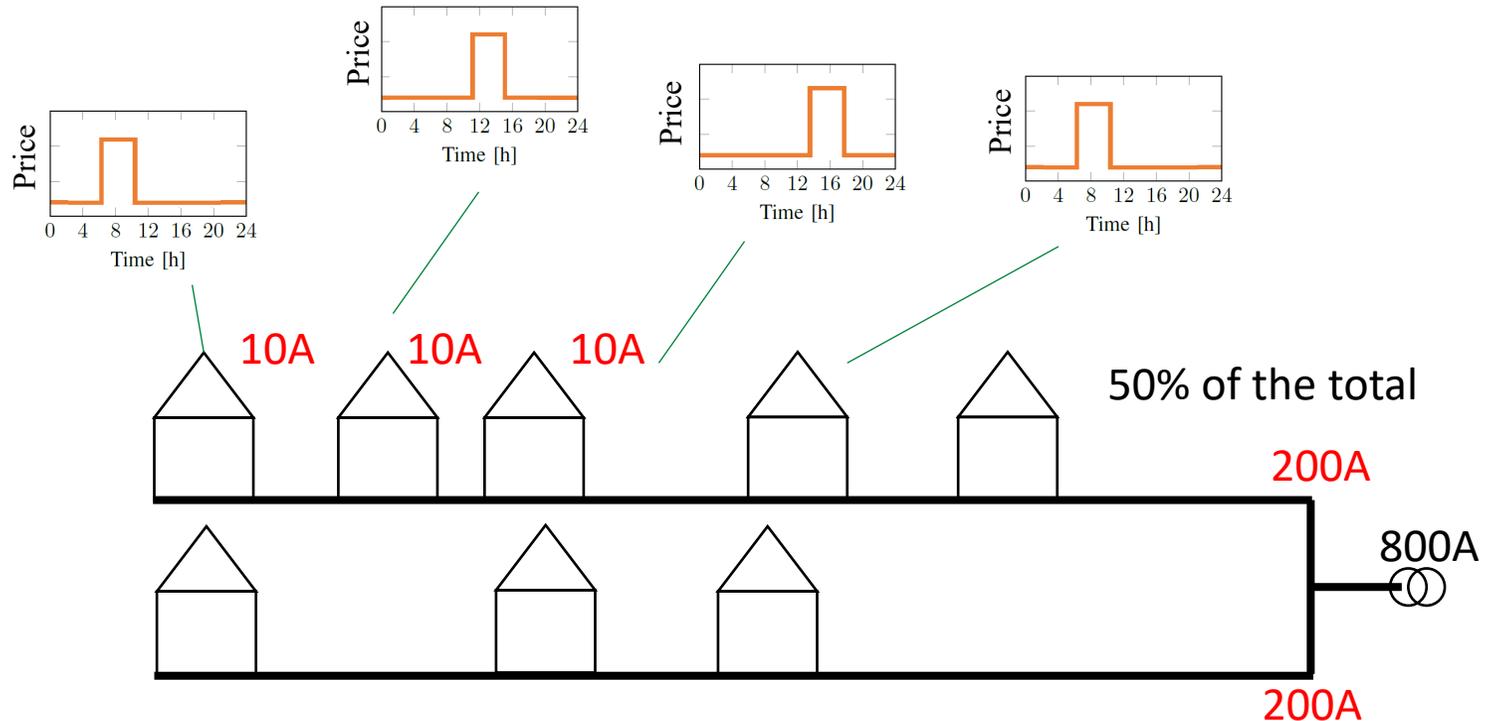


Does my coffee machine or washing machine consume more energy?

2. Lack of knowledge about own energy behaviour in the “everybody” person



Solution: Customised tariffs contracts



- + sell more energy with the same capacity
- + reduce congestion
- + allow for higher connection capacity

- But how to find individual contracts **at scale**?
- How to educate users **at scale**?

Idea: Chat to customise electricity contracts



Dont know much about energy

Hi

Hello! How can I assist you with planning your daily electricity consumption today? Let's start by gathering some information about your typical daily usage of various electrical appliances and activities. Here are a few questions to begin with:

1. **Washing Machine:** How many hours do you use your washing machine per day?
2. **Dishwasher:** How many hours per day is your dishwasher running?
3. **Television:** How many hours do you watch TV per day?
4. **Computer:** How many hours do you use a computer per day?
5. **Lights:** How many hours are the lights on in your home per day?
6. **Electric Vehicle:** Do you have an electric vehicle? If so, how many hours do you charge it per day?
7. **Electric Vehicle Charging Time:** If you have an electric vehicle, what time do you usually start charging it?

Please provide as much detail as you can for each question. If you don't use an appliance or if it's not applicable, you can simply say "none" or "0" for that question.

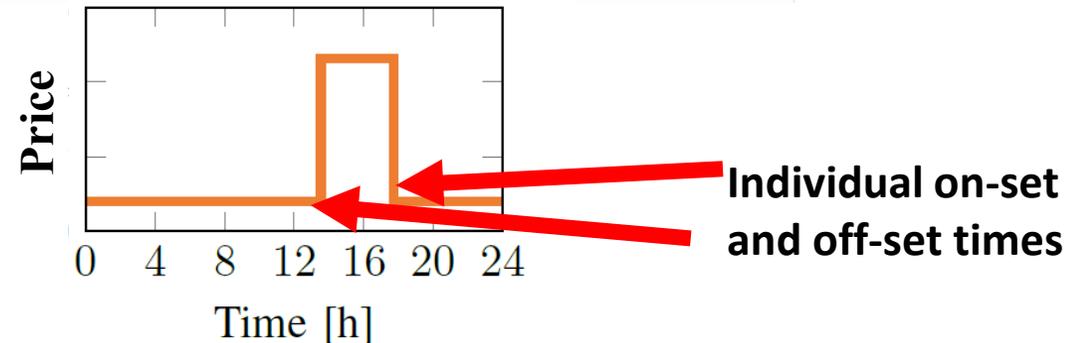
I am using my washing machine 2h per day in the evening.

Great, thank you for sharing that information. Let's continue gathering the rest of the details needed to calculate your daily electricity consumption:

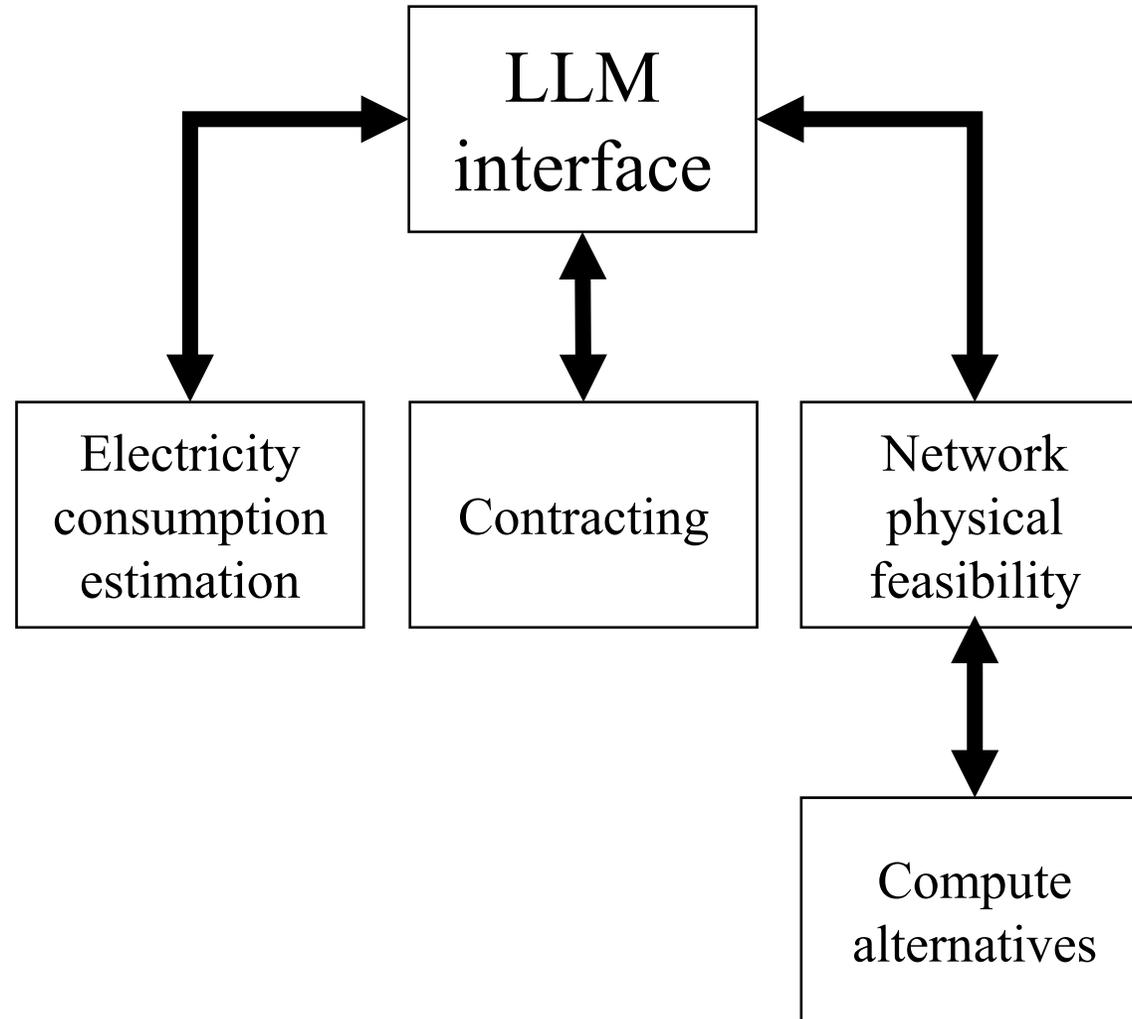
1. **Dishwasher:** How many hours per day is your dishwasher running?
2. **Television:** How many hours do you watch TV per day?
3. **Computer:** How many hours do you use a computer per day?
4. **Lights:** How many hours are the lights on in your home per day?



Customise contracts

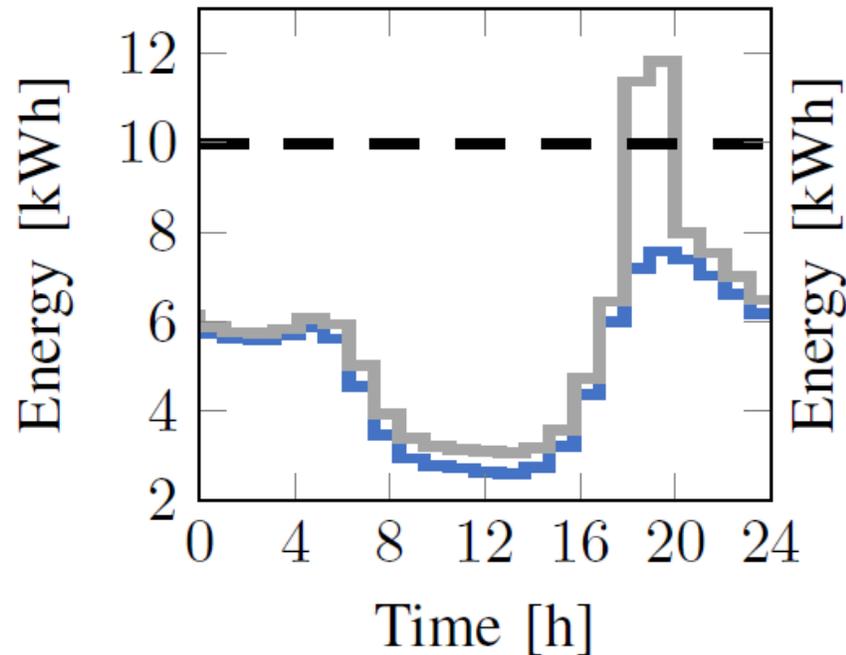


Algorithm technical details



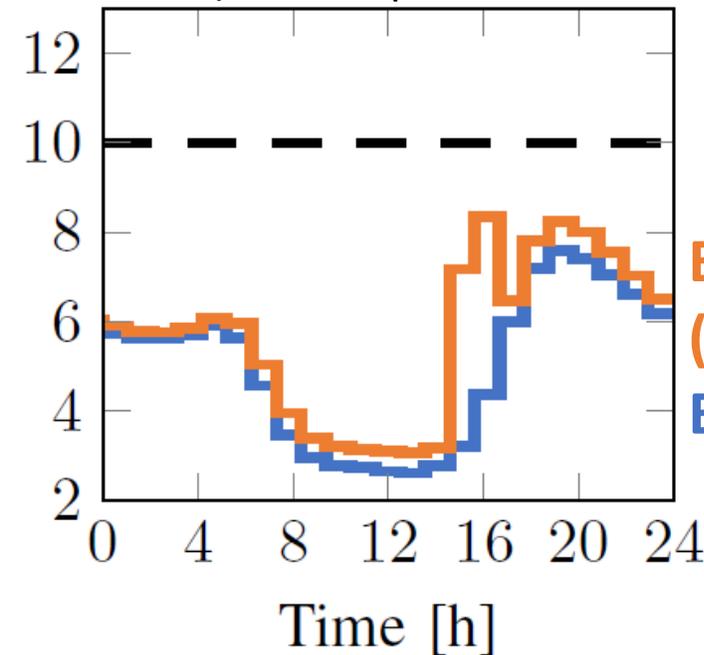
Benefits example: two neighbours

Scenario 1: No customised contracts



Scenario 2: Customised contract

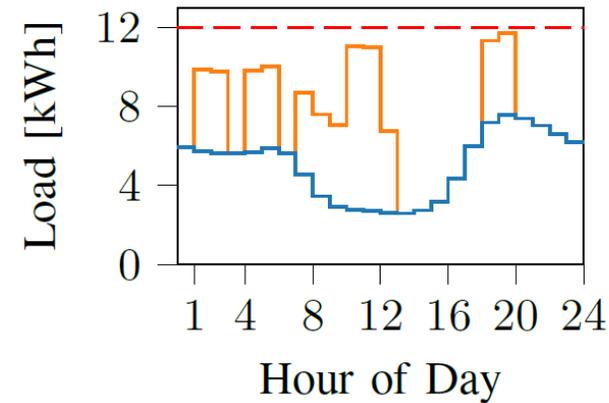
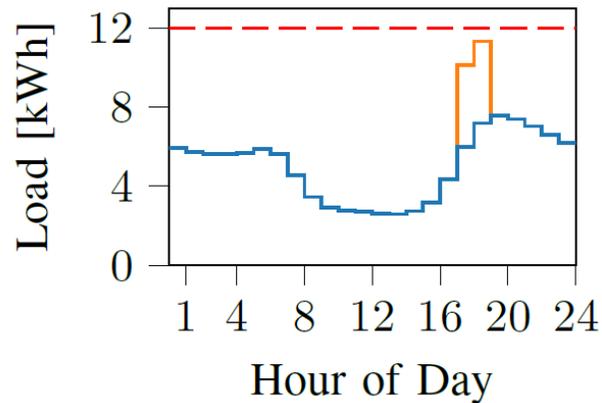
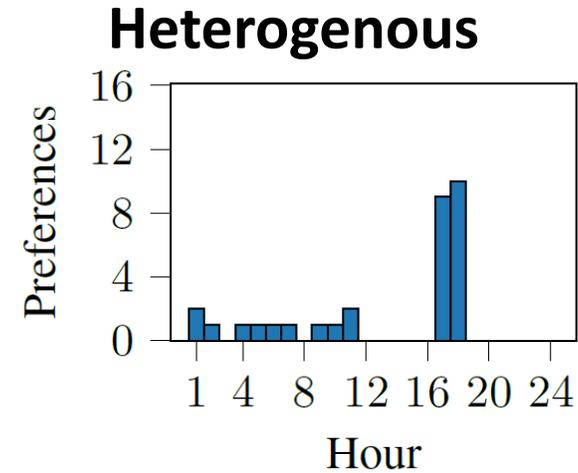
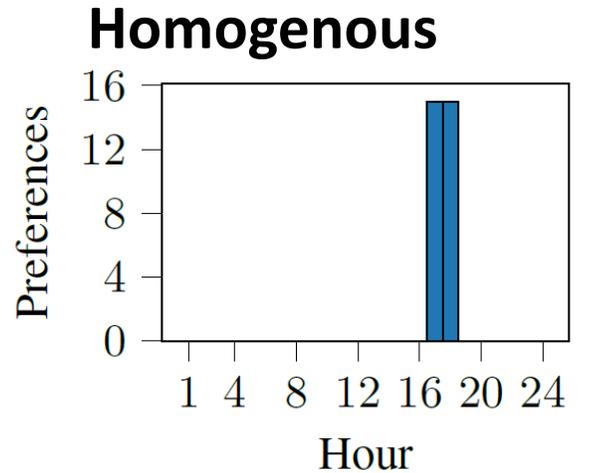
(with off-peak tariff before 16:00)



Energy aggregated
(household 1+ household 2)
Energy household 1

- + Lower total energy capacity
- + Realised residential demand shifts
- + Enable new assets to be connected like EVs

Impact on neighbourhood of 15 households



➔ **With standard contract only
1 EV can be connected**

➔ **With customised contracts
7EVs can be connected**

Conclusions

- Day-ahead FM benefits from prior knowledge of grid topology
- Pretraining with large samples from many EU regions brings better generalization
- Intraday FM design favors liquidity, not just “more samples” due to asymmetric generalization
- Accurate Intraday prediction requires explicit modeling of buy-sell interactions, which are fundamental formation mechanism of intraday prices

Future work

- Develop application to trading (and tailored strategies)
- Combine with grid models for interconnector power flows

Thank you

Speaker

Jochen Cremer

Associate Professor IEPG, TU Delft

Web: <https://www.tudelft.nl/ai/delft-ai-energy-lab>

Personal www.jochen-cremer.com

Email: j.l.cremer@tudelft.nl

Code: <https://github.com/TU-Delft-AI-Energy-Lab>



Lead PhD researcher

Runyao Yu



Project page, code, and data:

<https://runyao-yu.github.io/PriceFM/>

<https://runyao-yu.github.io/AsymGen/>

<https://runyao-yu.github.io/OrderFusion/>

References

- [1] Yu, R., Bunn, D.W., Lin, J., Stiasny, J., Leimgruber, F., Esterl, T., Tao, Y., Qi, L., Chen, Y., Wang, W. and Cremer, J.L., "Deep Learning for Electricity Price Forecasting: A Review of Day-Ahead, Intraday, and Balancing Electricity Markets", *arXiv preprint arXiv:2602.10071*, (2026)
- [2] Yu, Runyao, Chenhui Gu, Jochen Stiasny, Qingsong Wen, Wasim Sarwar Dilov, Lianlian Qi, and Jochen L. Cremer. "PriceFM: Foundation model for probabilistic electricity price forecasting." *arXiv preprint arXiv:2508.04875* (2025)
- [3] Runyao Yu, Ruochen Wu, Yongsheng Han, Jochen L. Cremer, "Orderbook Feature Learning and Asymmetric Generalization in Intraday Electricity Markets", *to be presented at 2026 IEEE Power System Computation Conference, Cyprus*
- [4] Yu, Runyao, Yuchen Tao, Fabian Leimgruber, Tara Esterl, Jochen Stiasny, Derek W. Bunn, Qingsong Wen, Hongye Guo, and Jochen L. Cremer. "OrderFusion: Encoding Orderbook for End-to-End Probabilistic Intraday Electricity Price Forecasting." *arXiv preprint arXiv:2502.06830* (2025)
- [5] Cremer, Jochen L. "Customising electricity contracts at scale with large language models." *IEEE Transactions on Power Systems* (2026)