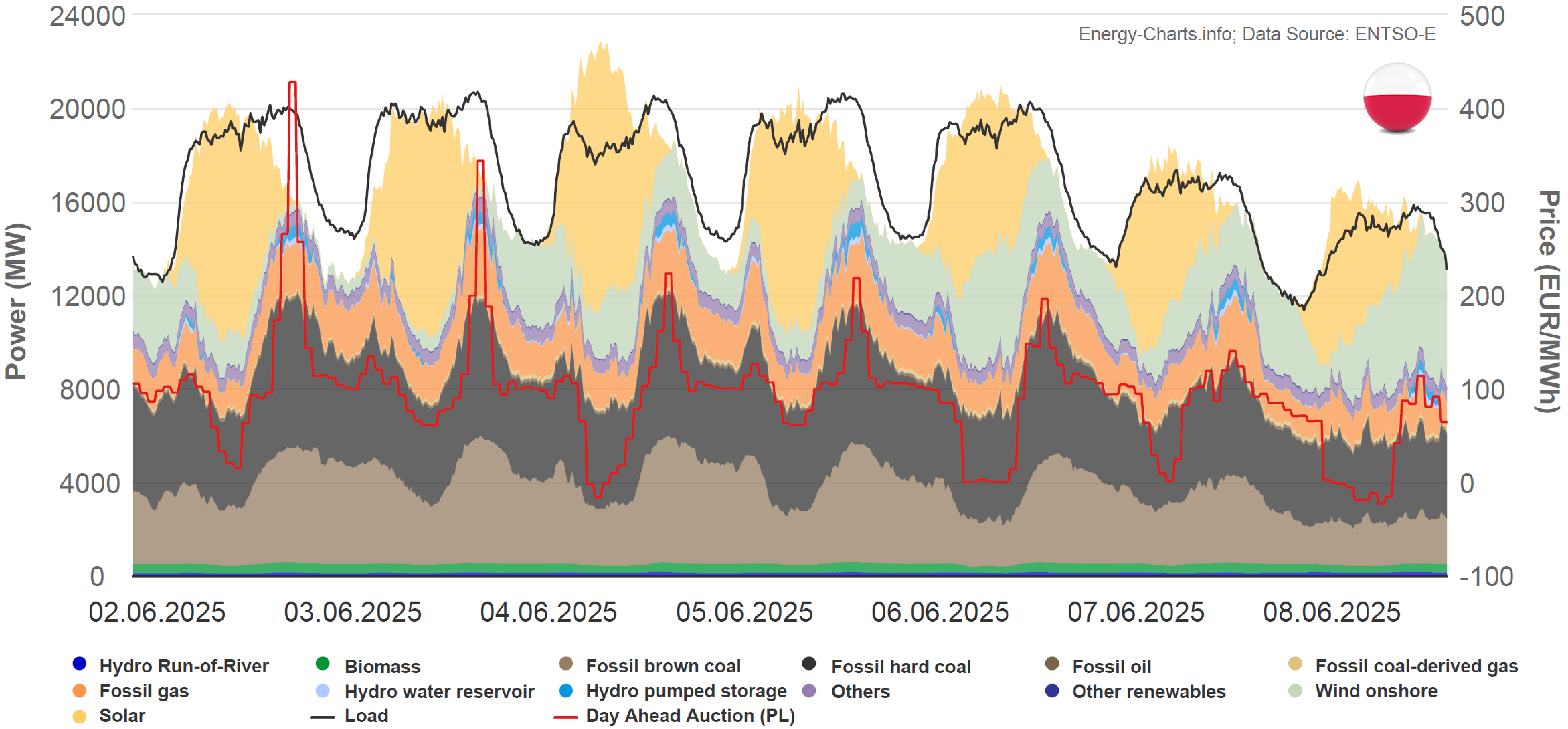
Stealing accuracy: Predicting day-ahead electricity prices with Temporal Hierarchy Forecasting (THieF)

Kaja Bilińska*, Nikolaos Kourentzes**, Arkadiusz Lipiecki*, Rafał Weron*

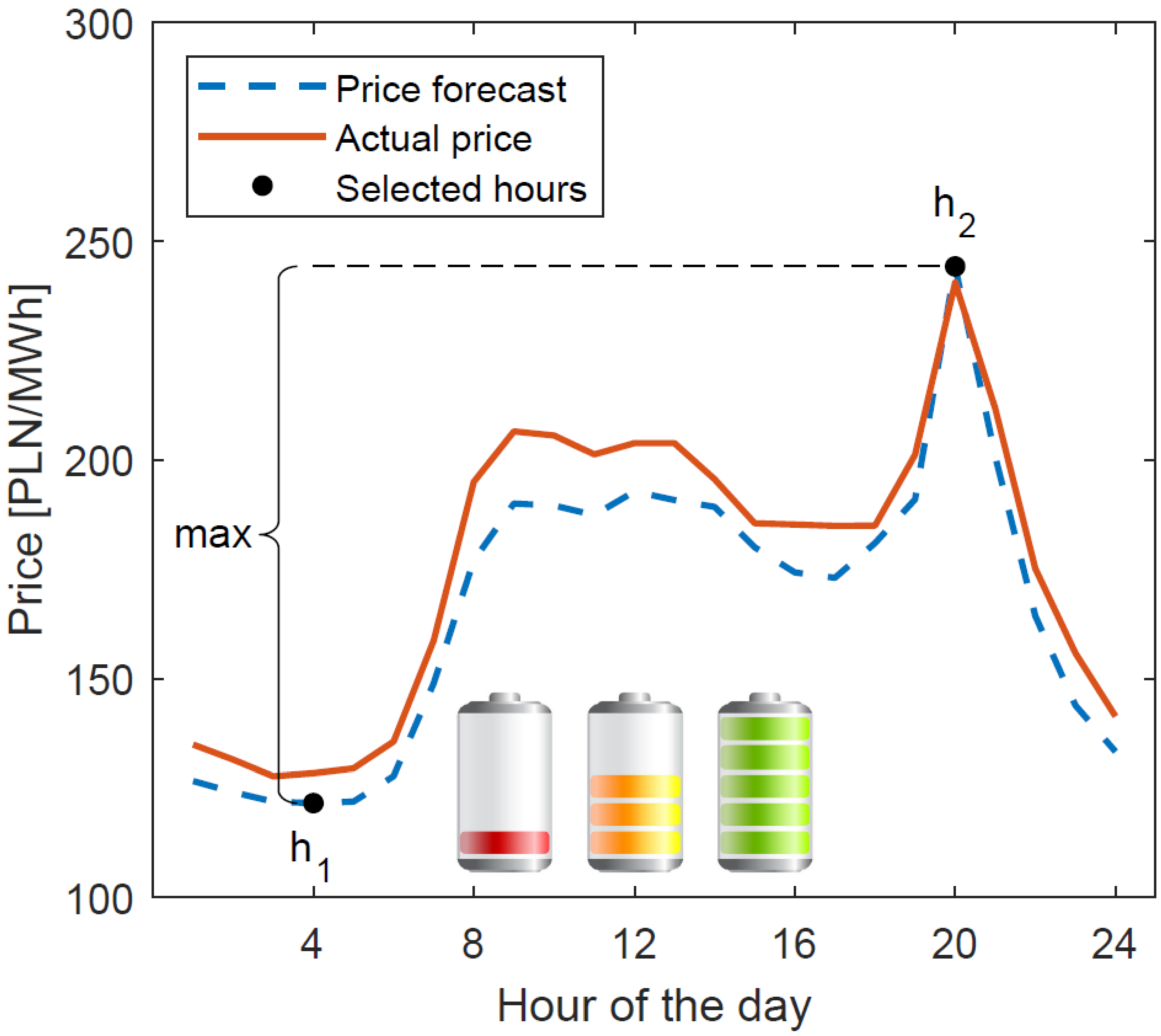
* Wrocław University of Science and Technology

** UNIVERSITY OF SKÖVDE

What are we trying to predict?



Why do we need electricity price forecasts?

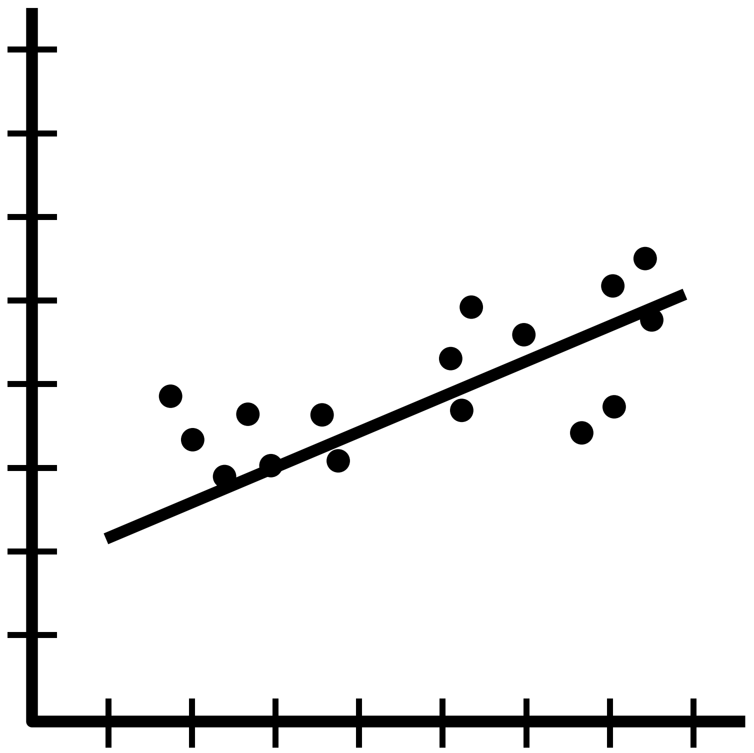


- Buy and charge the battery when the price is low
- Discharge the battery and sell when the price is high

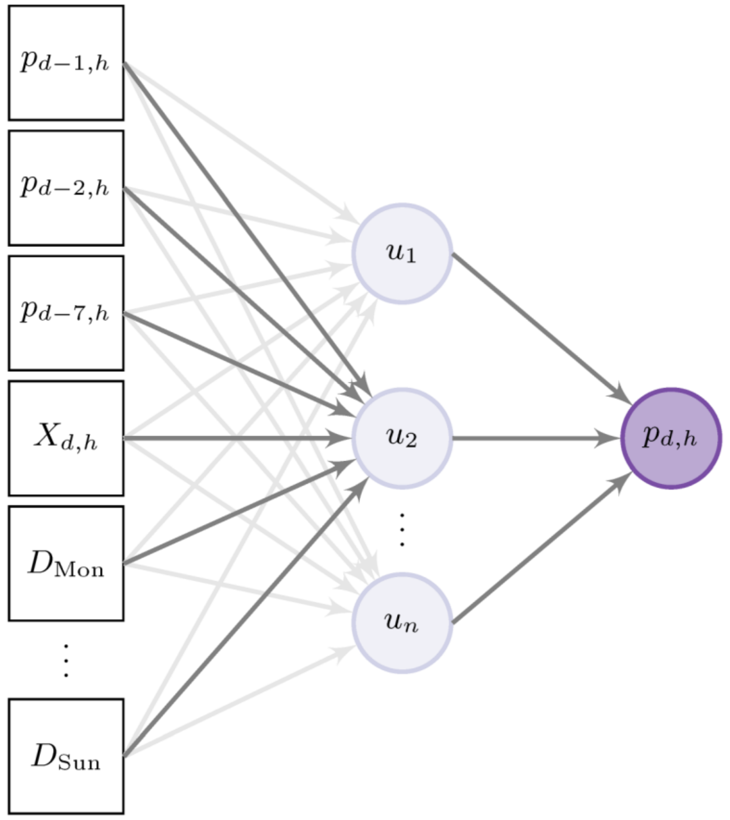
How do we generate forecasts?

$$\hat{p}_{d,h} = f\Big(\underbrace{p_{d-1,h}, \ldots, \; p_{d-7,h}, \; p_{d-1}^{min}, \; p_{d-1}^{max}}_{\text{AutoRegression}}, \; \underbrace{\hat{L}_{d,h}, \; \hat{W}_{d,h}}_{\text{Load, Wind}}, \underbrace{\mathrm{API}_{d-2}, \; \mathrm{TTF}_{d-2}}_{\text{Fuels}}, \; \underbrace{\mathrm{D}_d^{(1)}, \ldots, \; \mathrm{D}_d^{(7)}}_{\text{Dummies}}\Big)$$

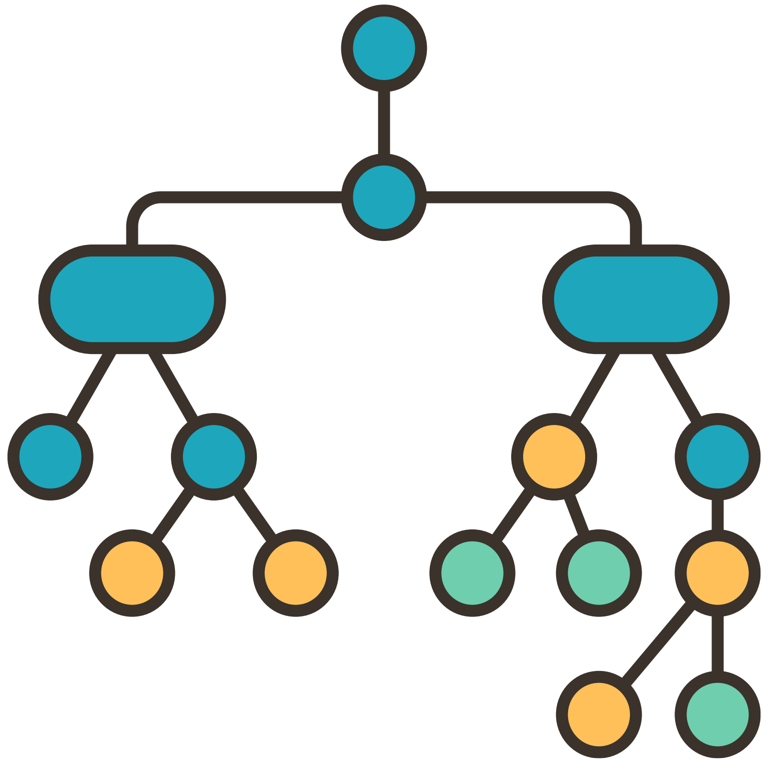
ARX



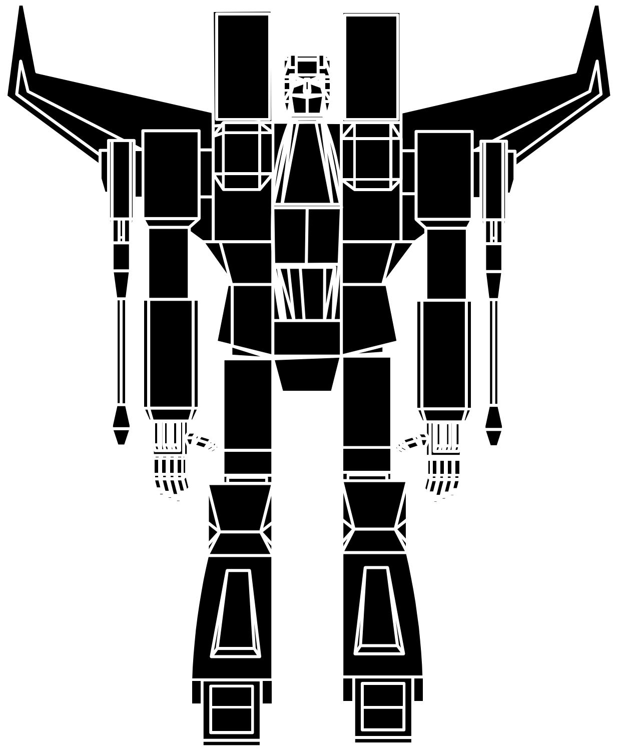
NARX



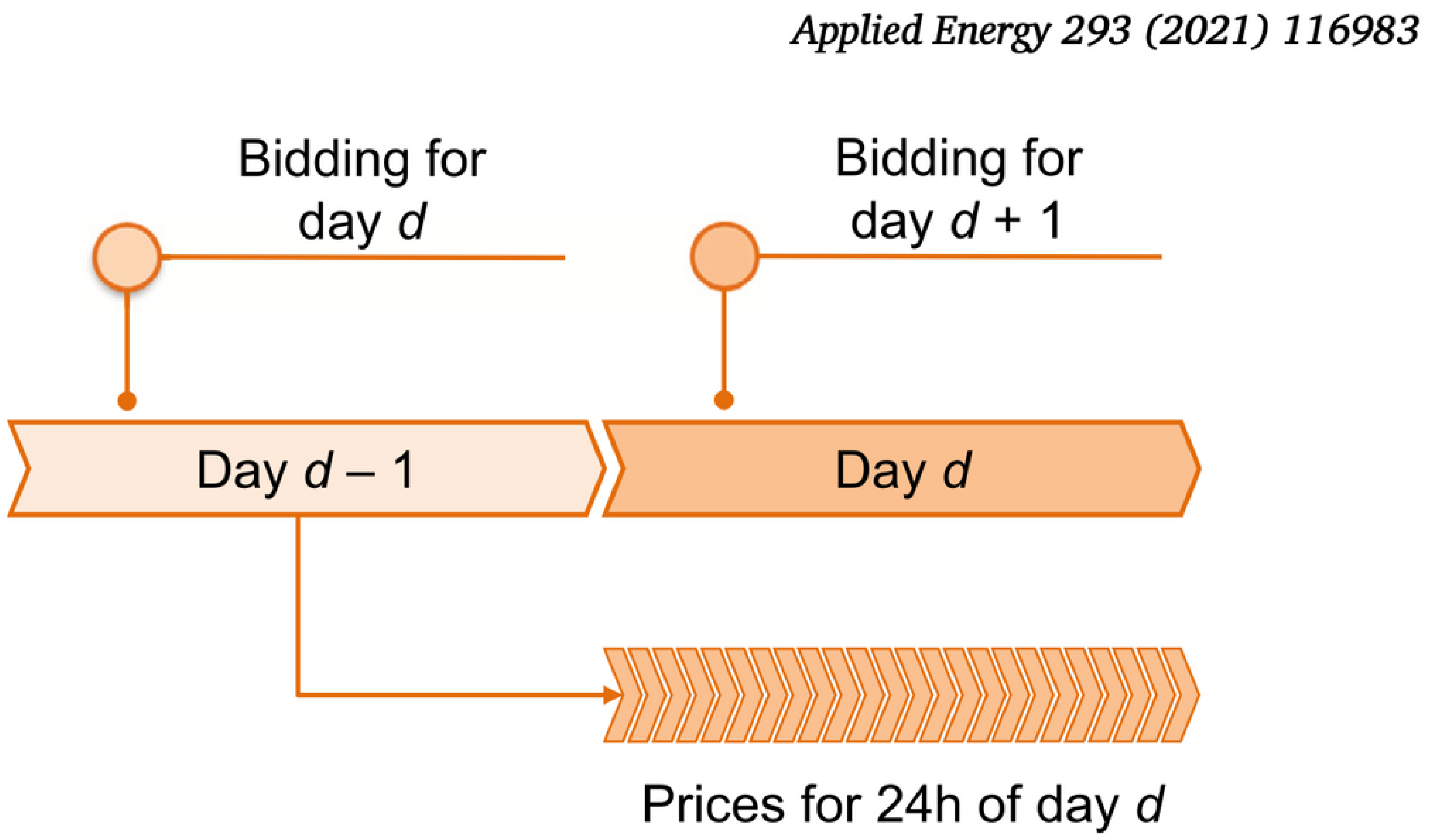
XGB

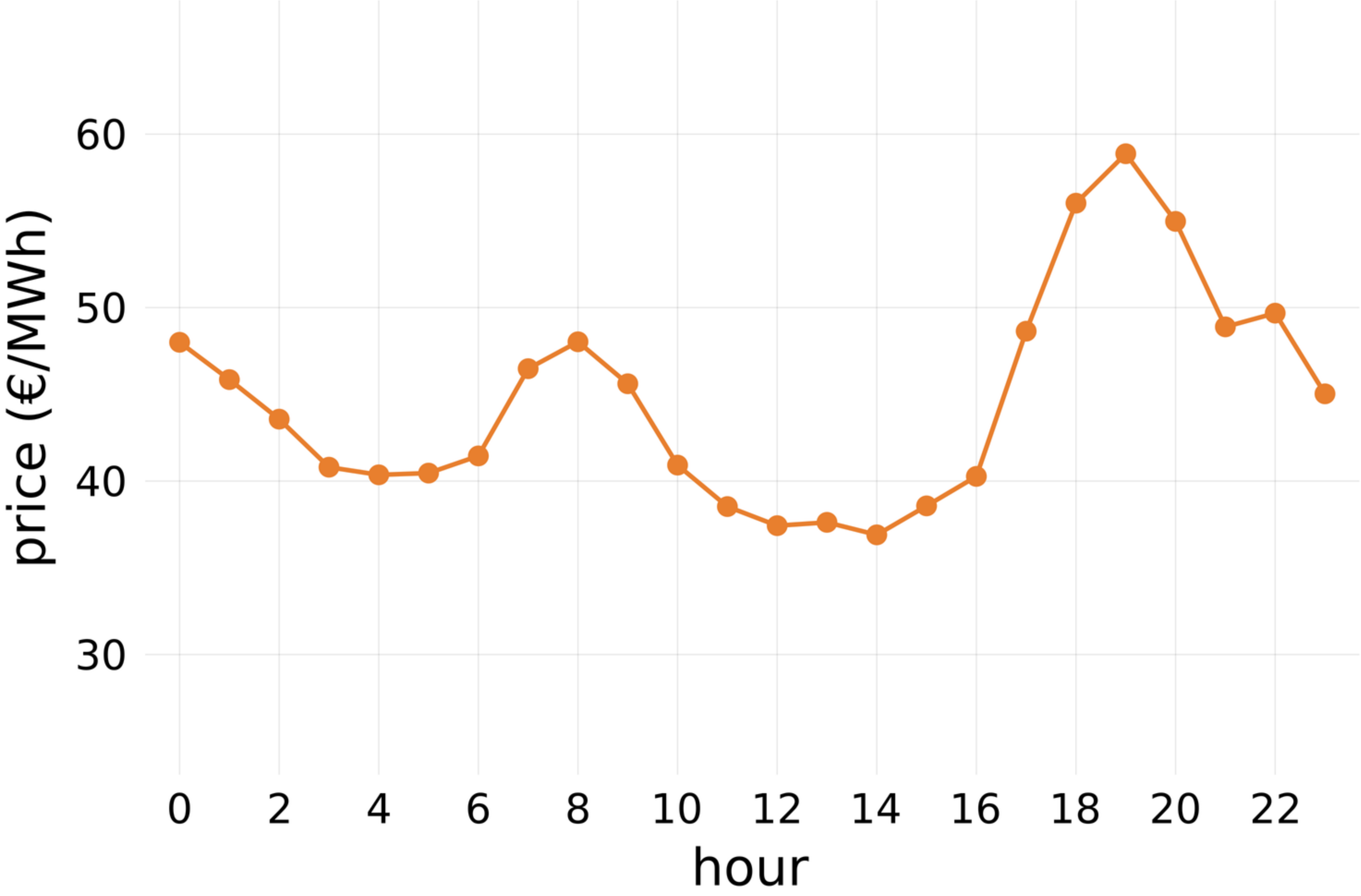


Mitra



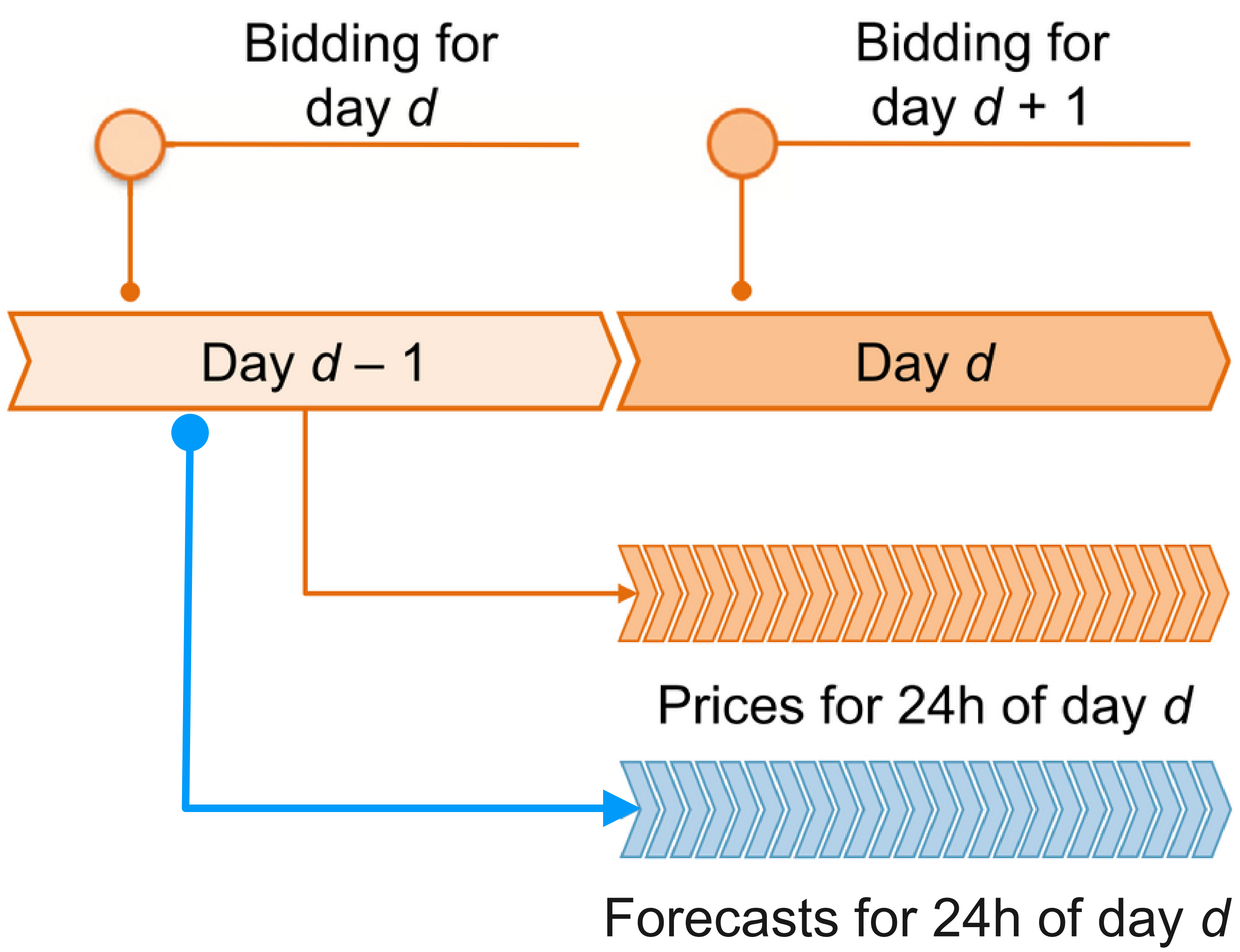
Day-ahead markets for electricity

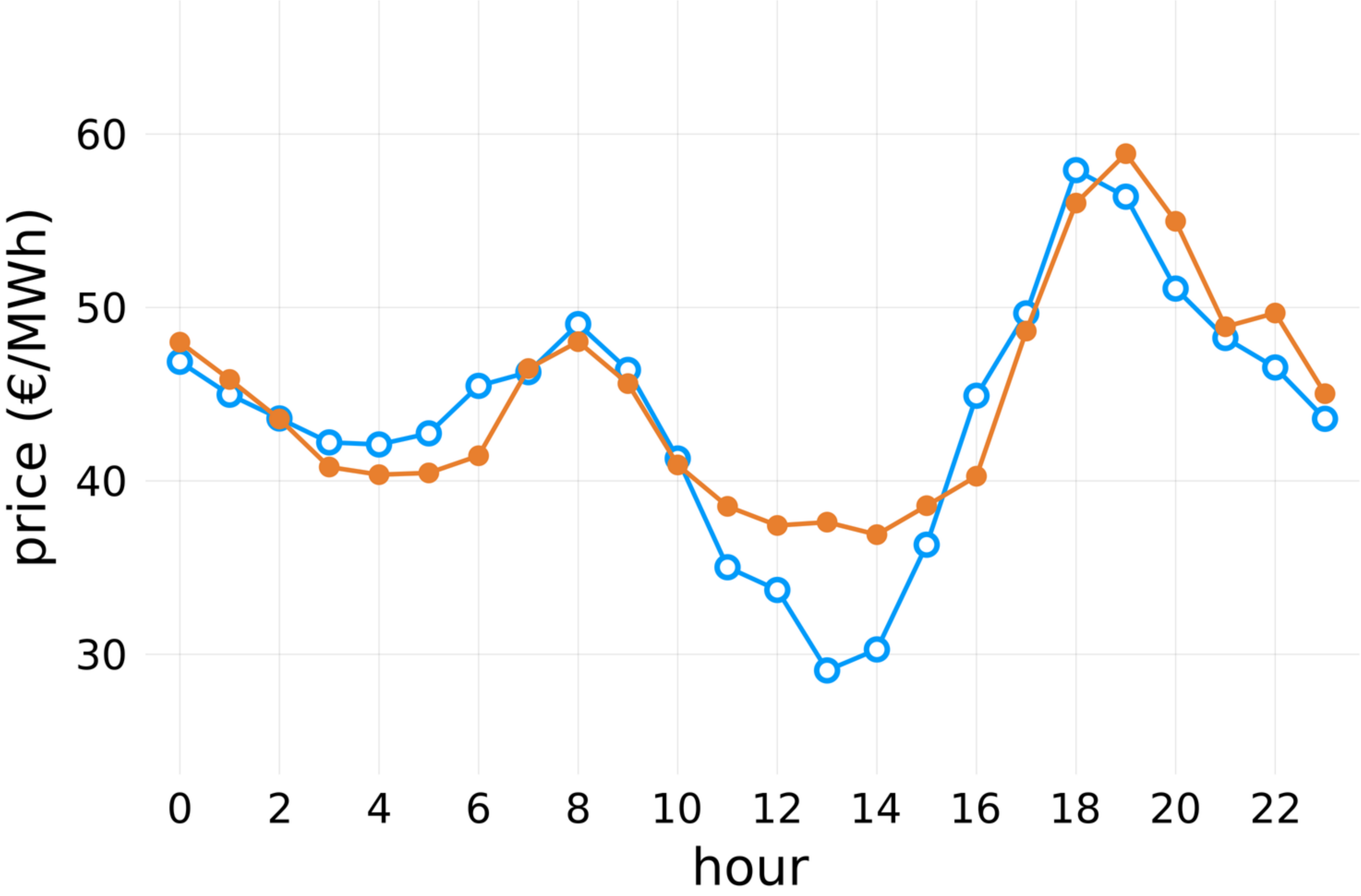




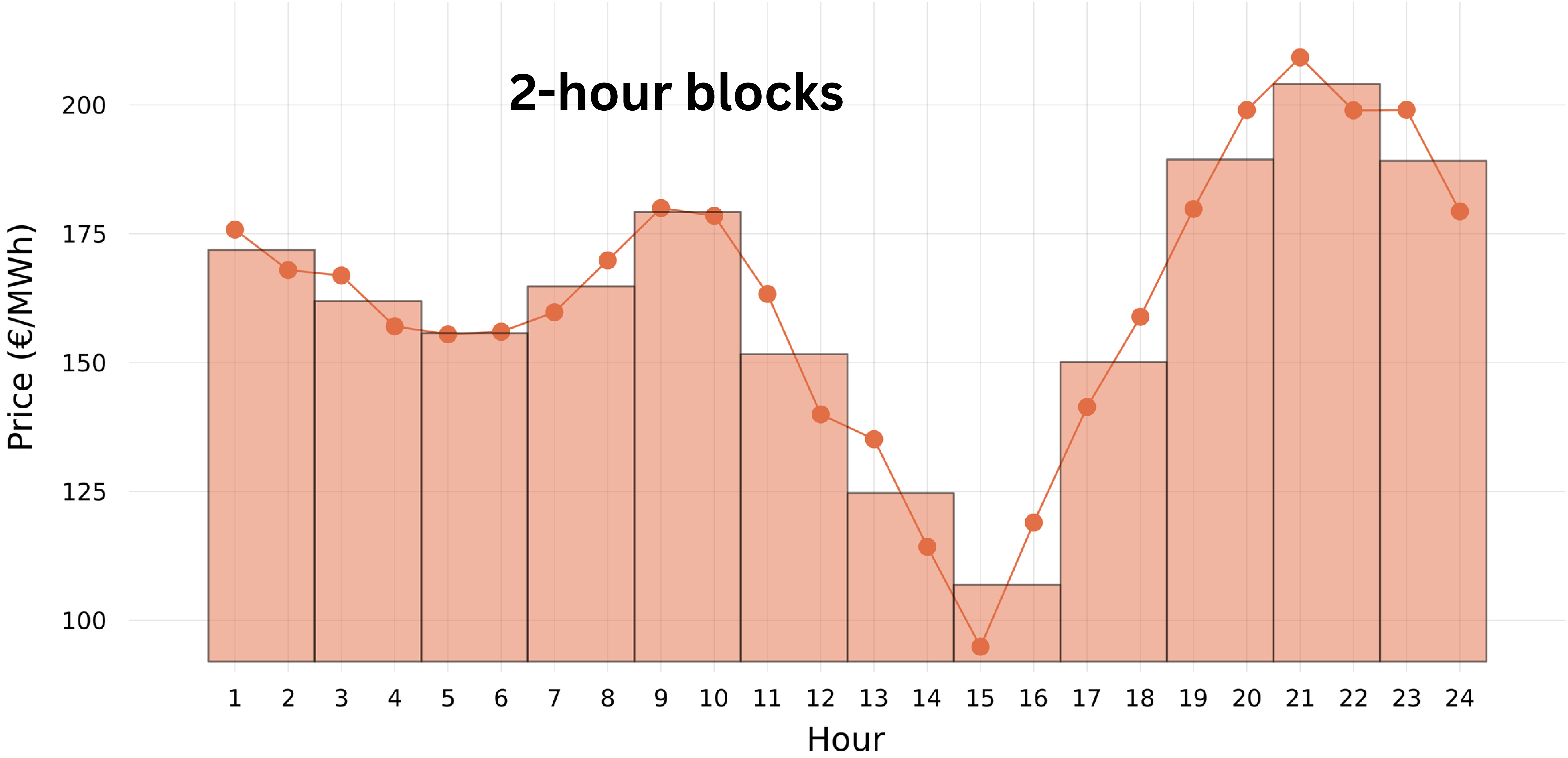
Forecasts of prices for the next day

Applied Energy 293 (2021) 116983

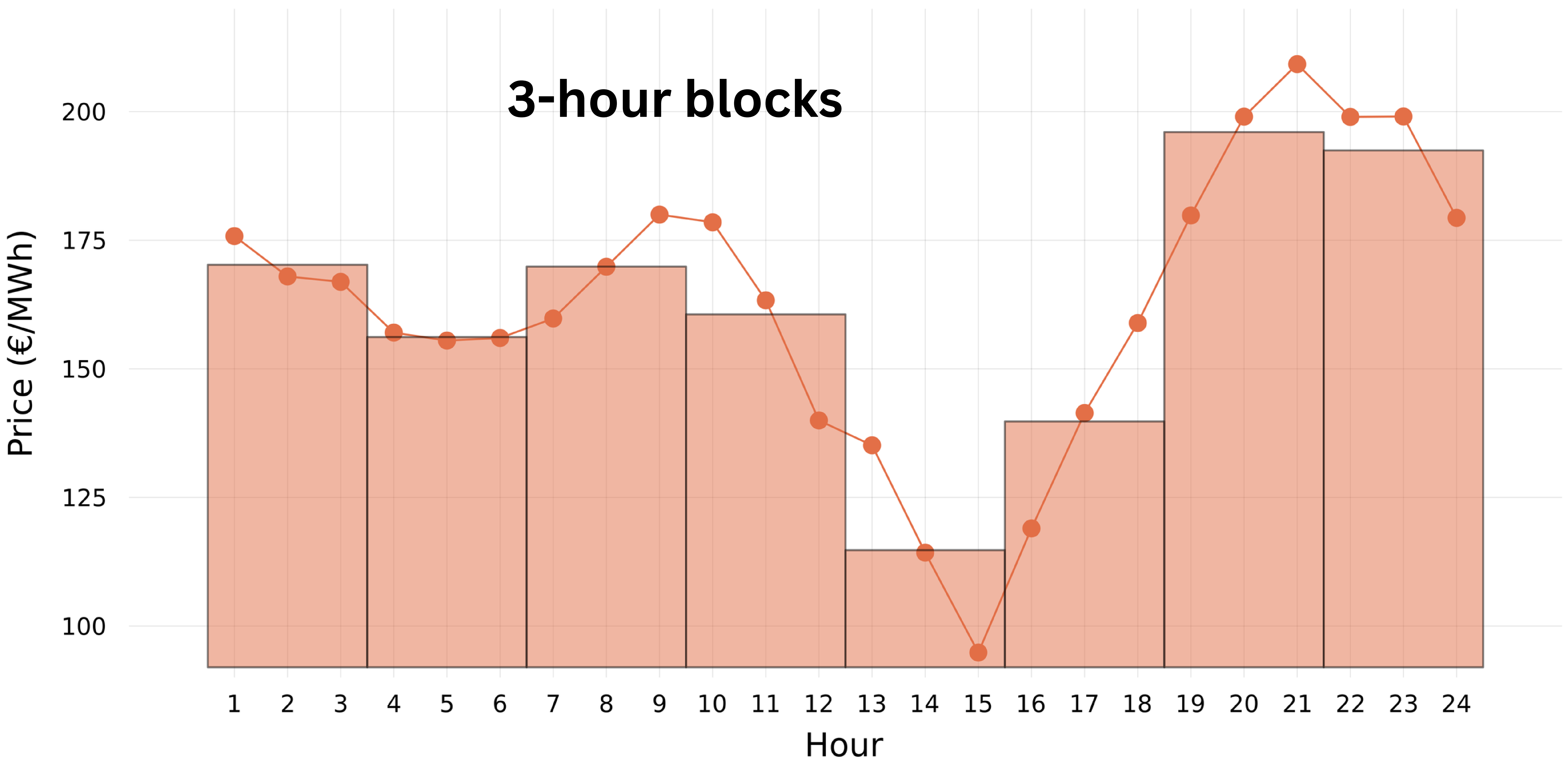




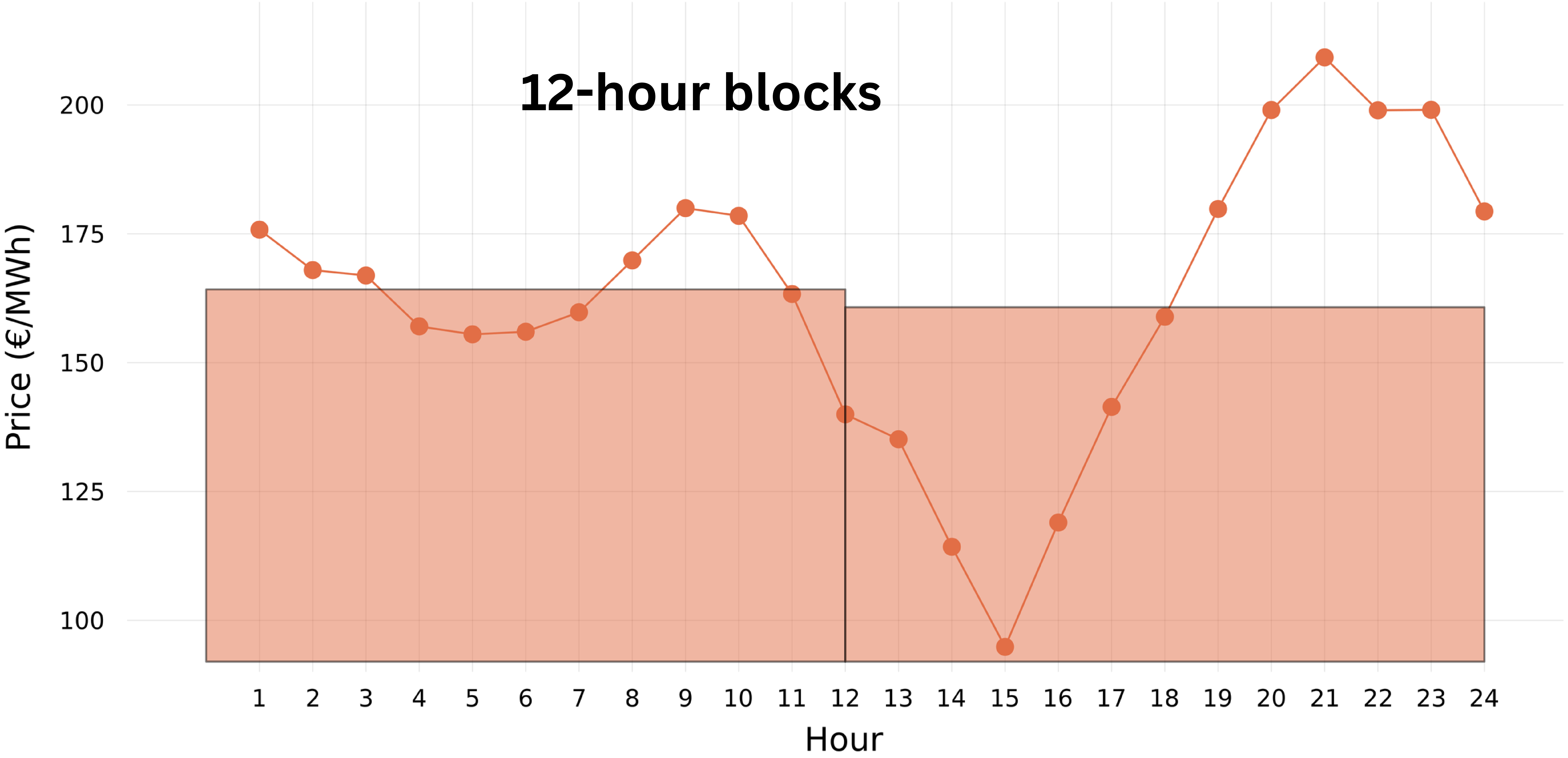
Trading: Block contracts



Trading: Block contracts

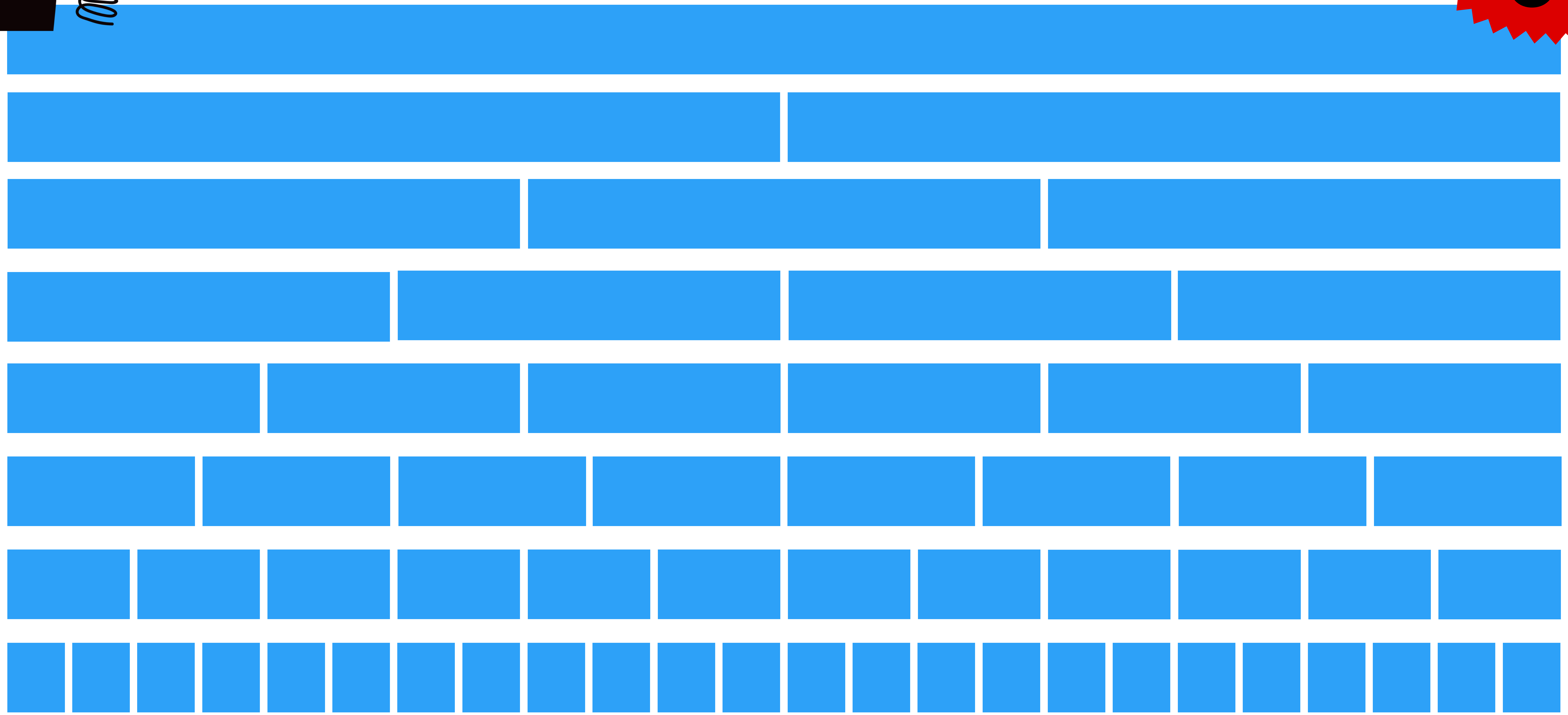


Trading: Block contracts



Temporal hierarchies

one 24-hour block (daily average)

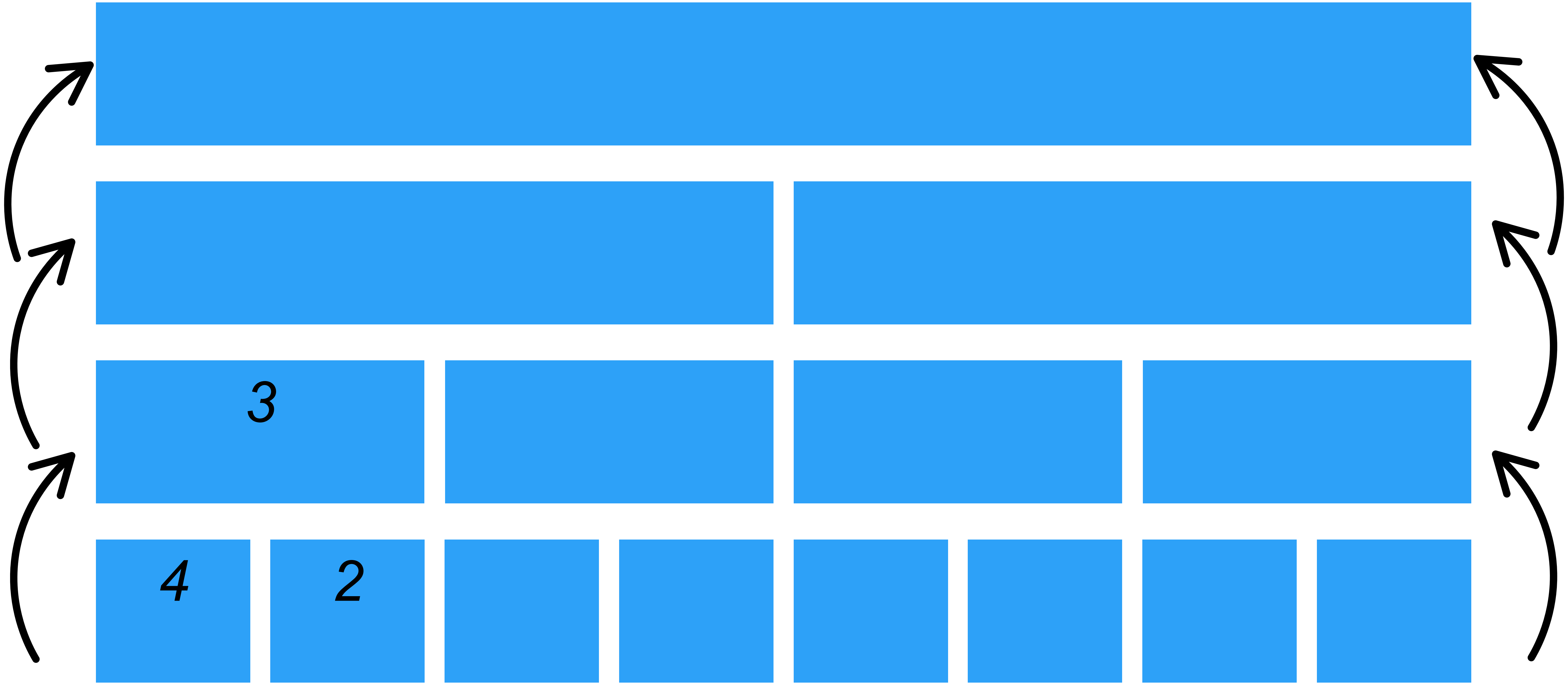


24 hourly blocks

Temporal hierarchies

Bottom up

Temporal hierarchies



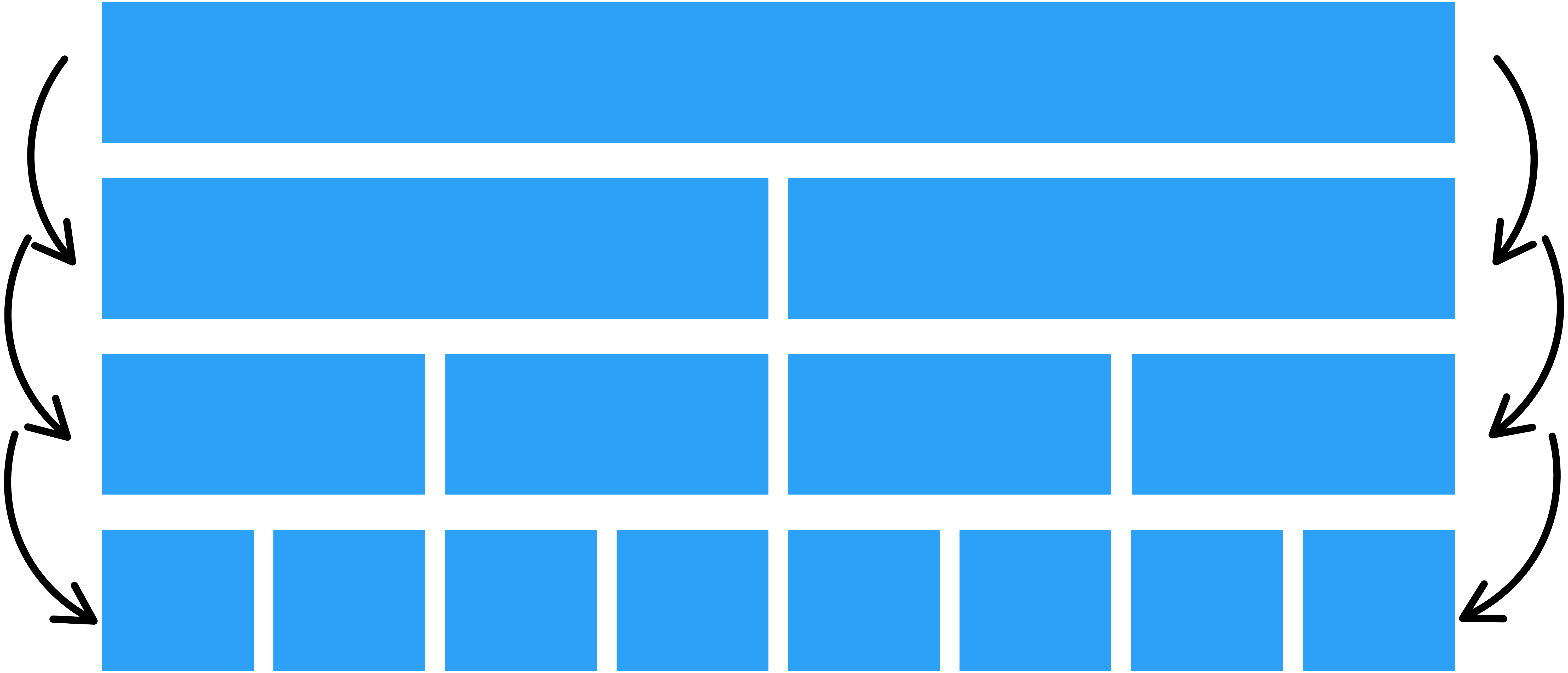
Bottom up

Temporal hierarchies

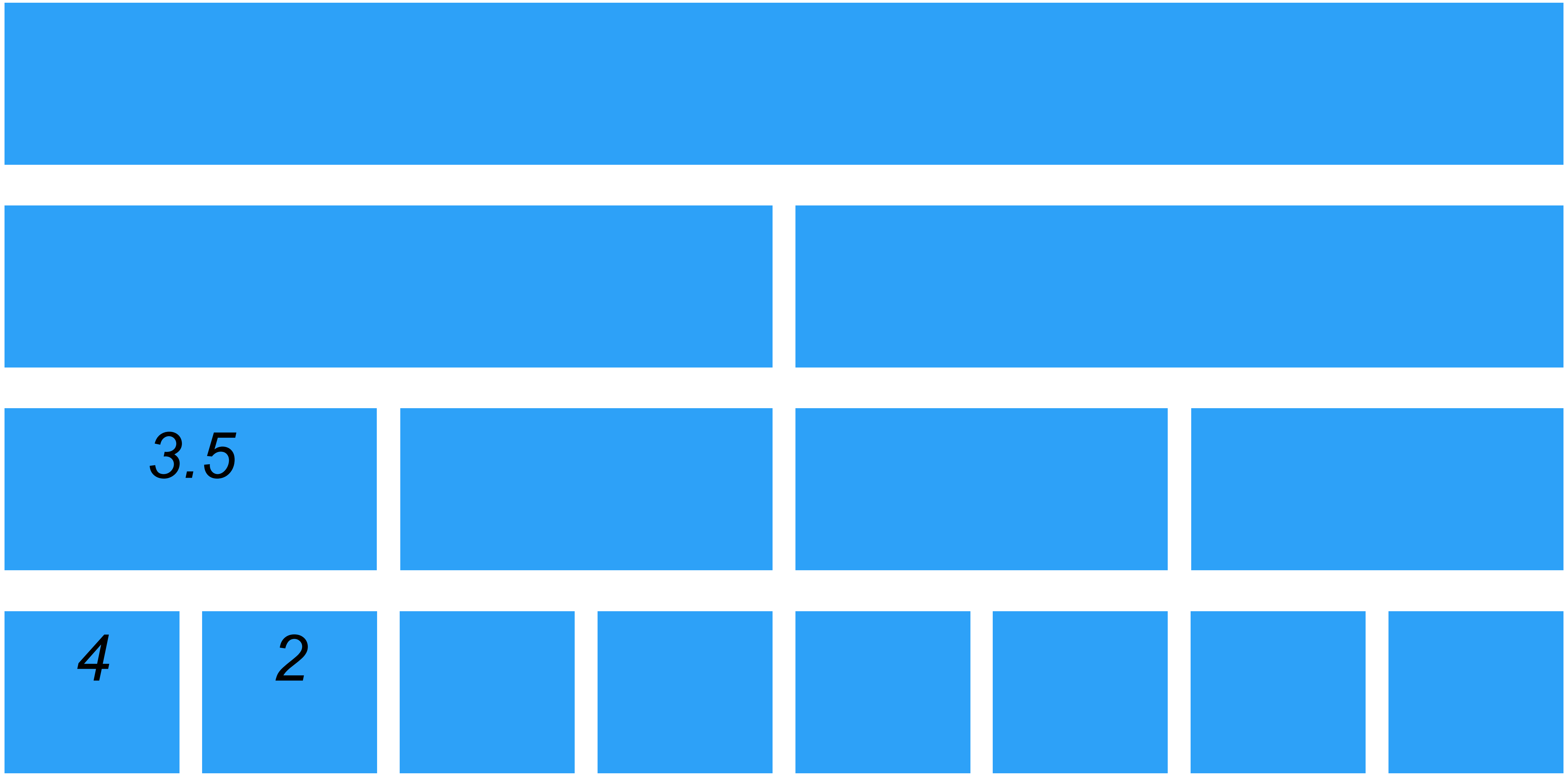
Top down

Temporal hierarchies

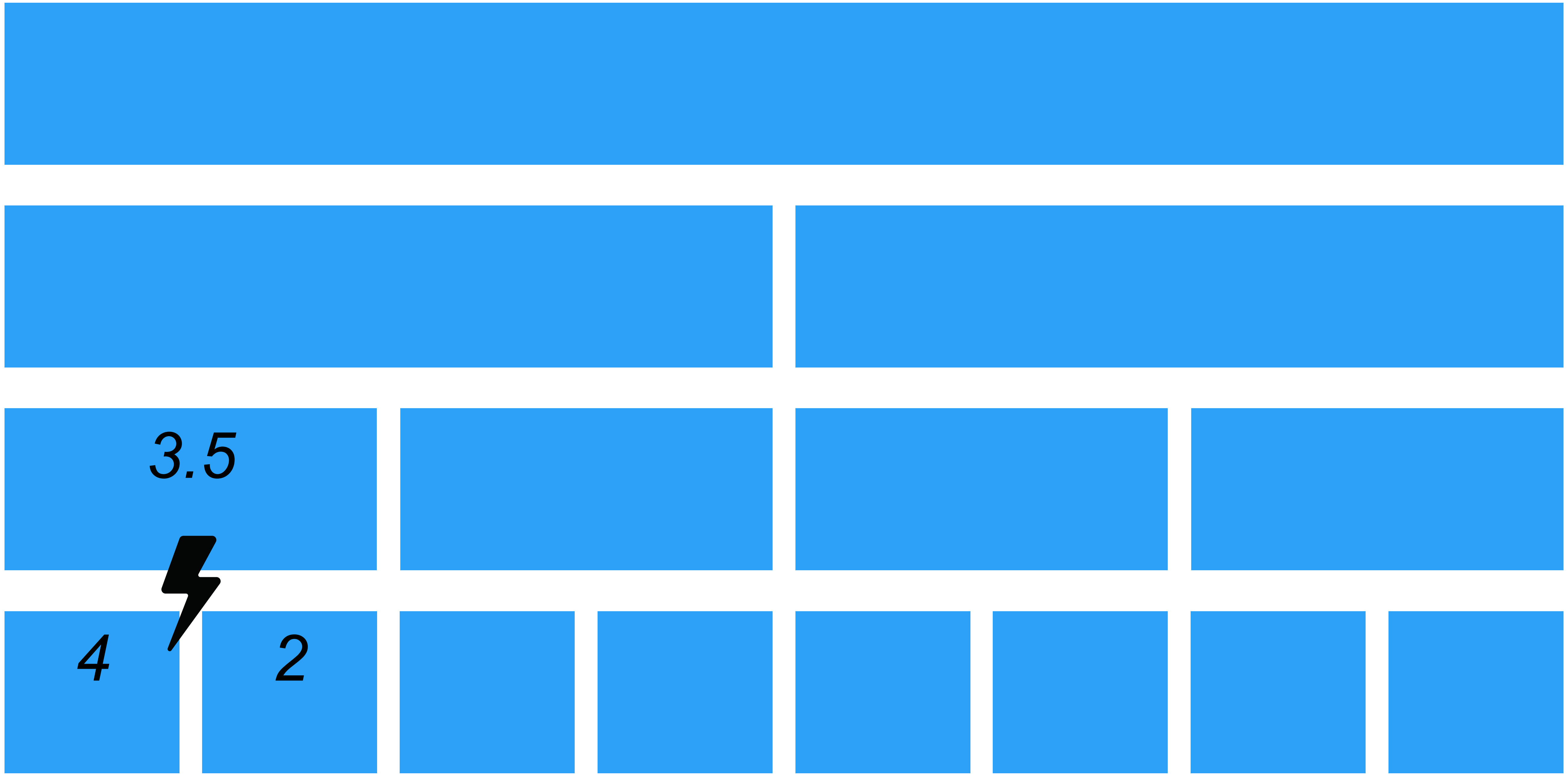
Top down



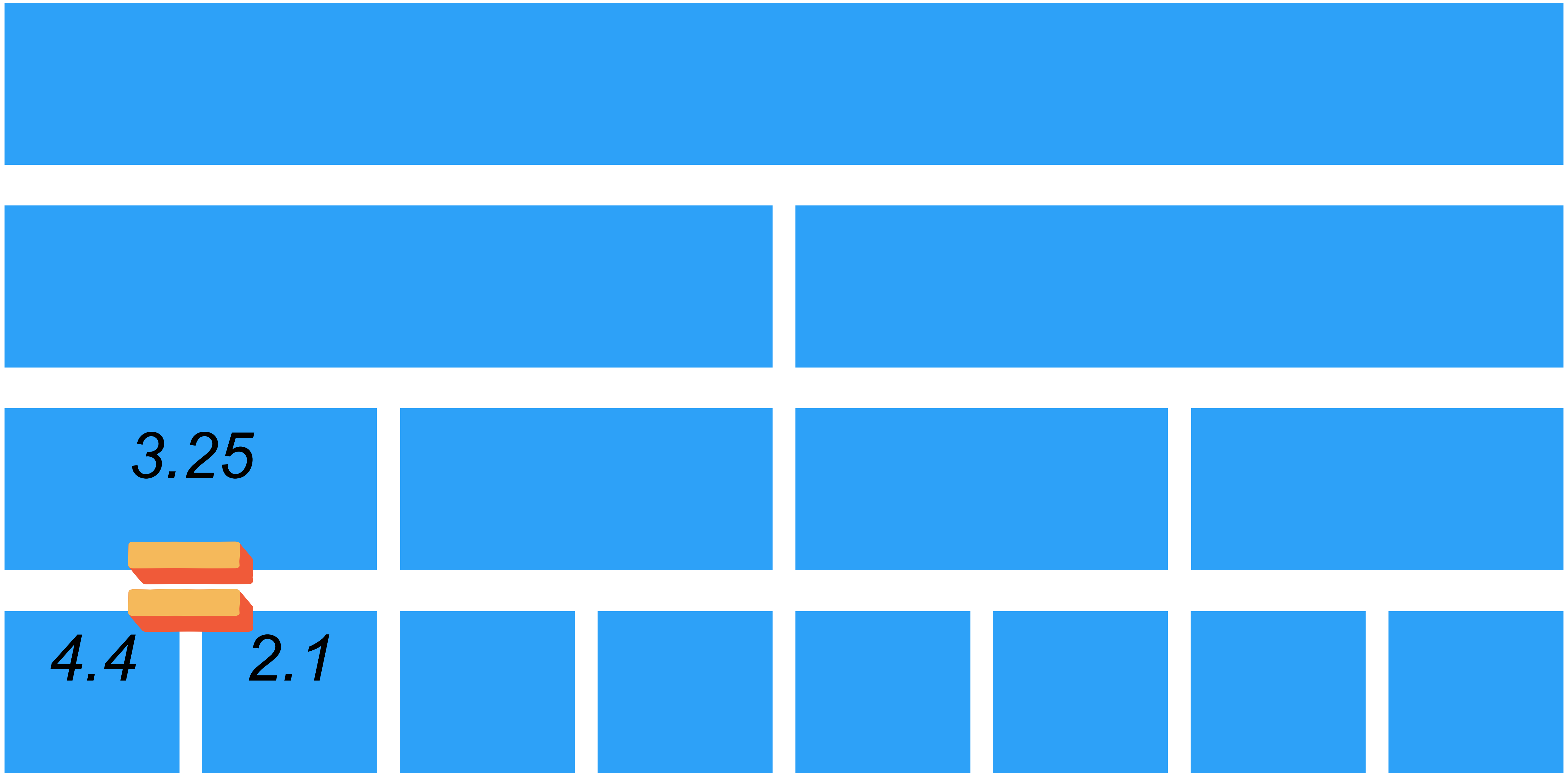
Forecast reconciliation



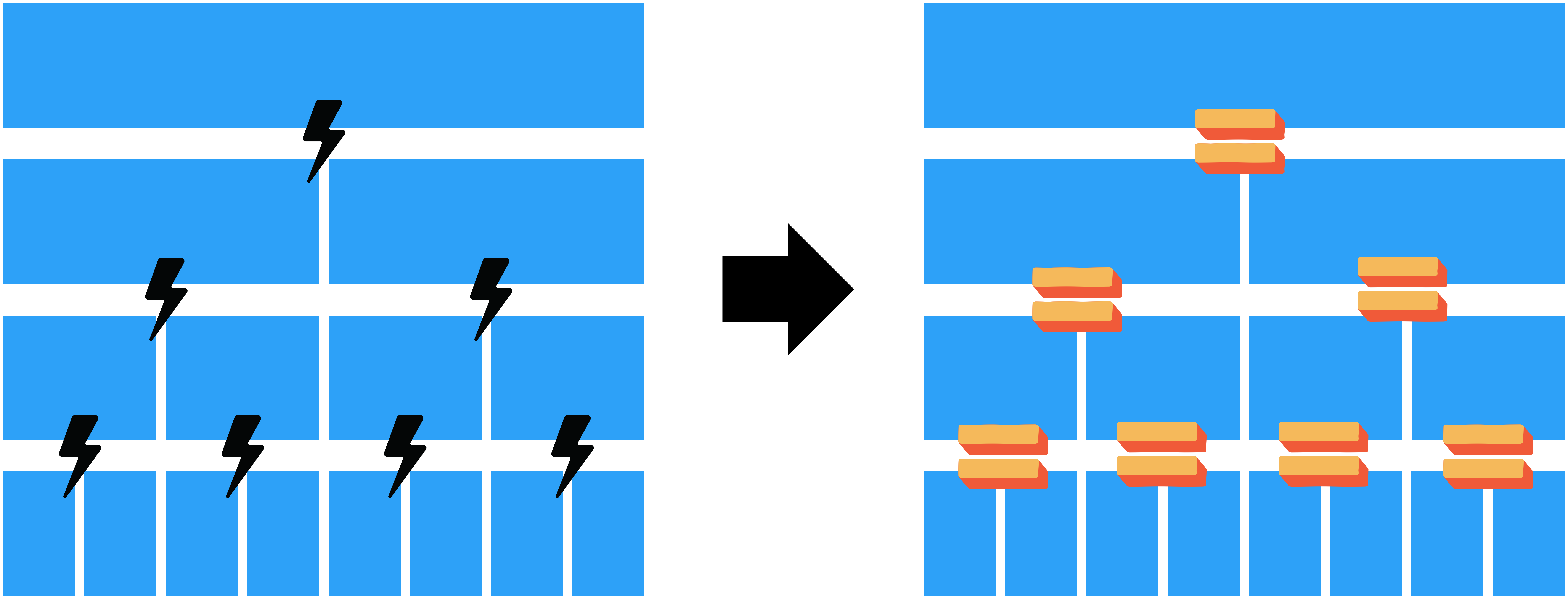
Forecast reconciliation



Forecast reconciliation

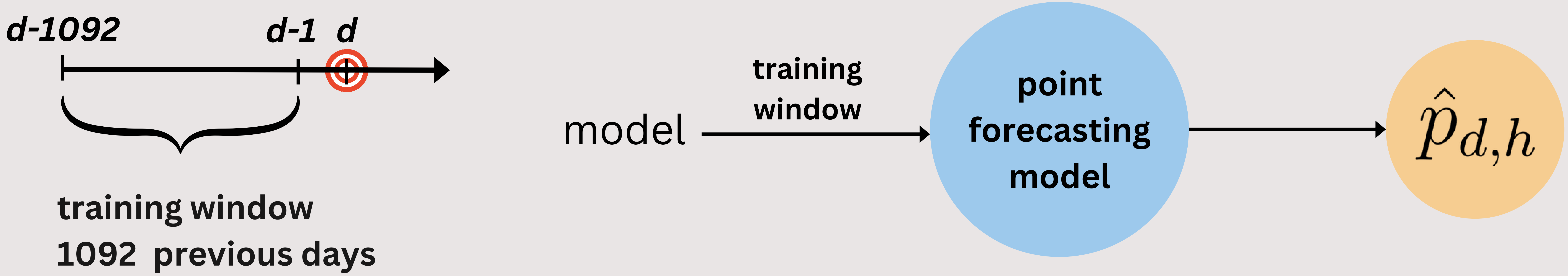


Forecast reconciliation



Step I: Compute base forecasts

Estimate base models for block ***h*** of day ***d***

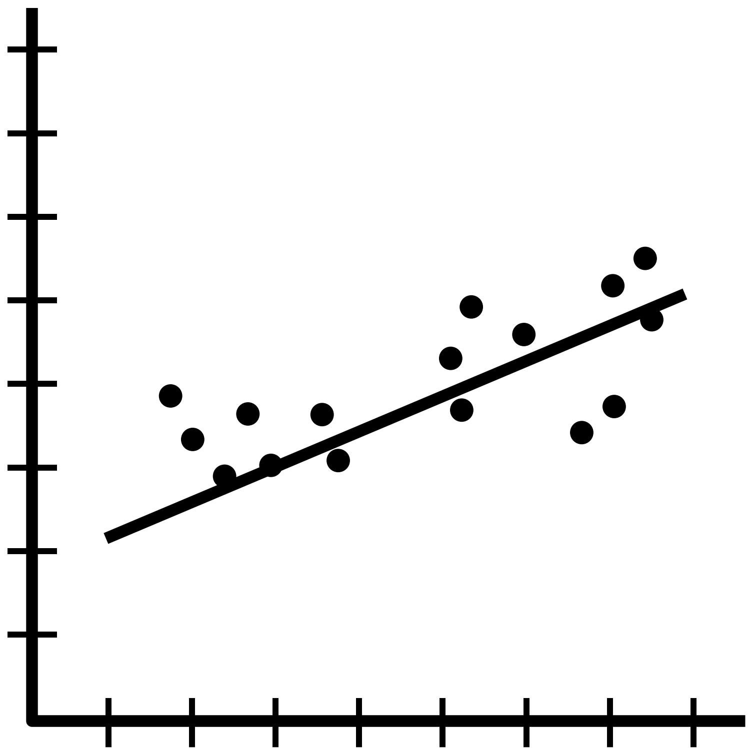


$$\hat{p}_{d,h} = f\left(p_{d-1,h}, \ldots, p_{d-7,h}, p_{d-1}^{min}, p_{d-1}^{max}, \hat{L}_{d,h}, \hat{W}_{d,h}, \mathrm{API}_{d-2}, \mathrm{TTF}_{d-2}, \mathrm{D}_d^{(1)}, \ldots, \mathrm{D}_d^{(7)}\right)$$

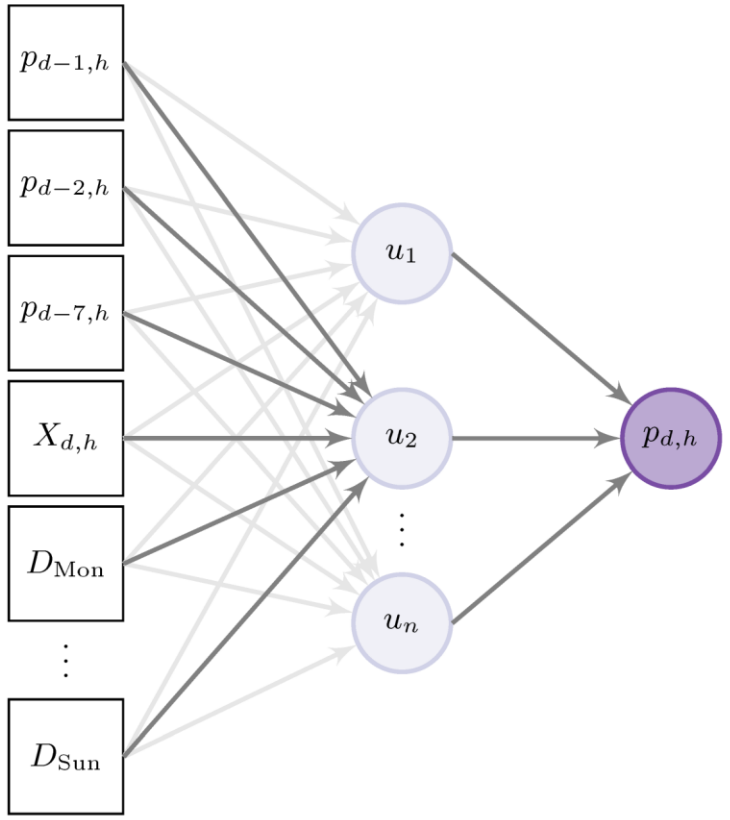
Base forecasts

$$\hat{p}_{d,h} = f\Big(\underbrace{p_{d-1,h}, \dots, p_{d-7,h}, p^{min}_{d-1}, p^{max}_{d-1}}_{\text{AutoRegression}}, \underbrace{\hat{L}_{d,h}, \hat{W}_{d,h}}_{\text{Load, Wind}}, \underbrace{\mathrm{API}_{d-2}, \mathrm{TTF}_{d-2}}_{\text{Fuels}}, \underbrace{\mathrm{D}^{(1)}_{d}, \dots, \mathrm{D}^{(7)}_{d}}_{\text{Dummies}}\Big)$$

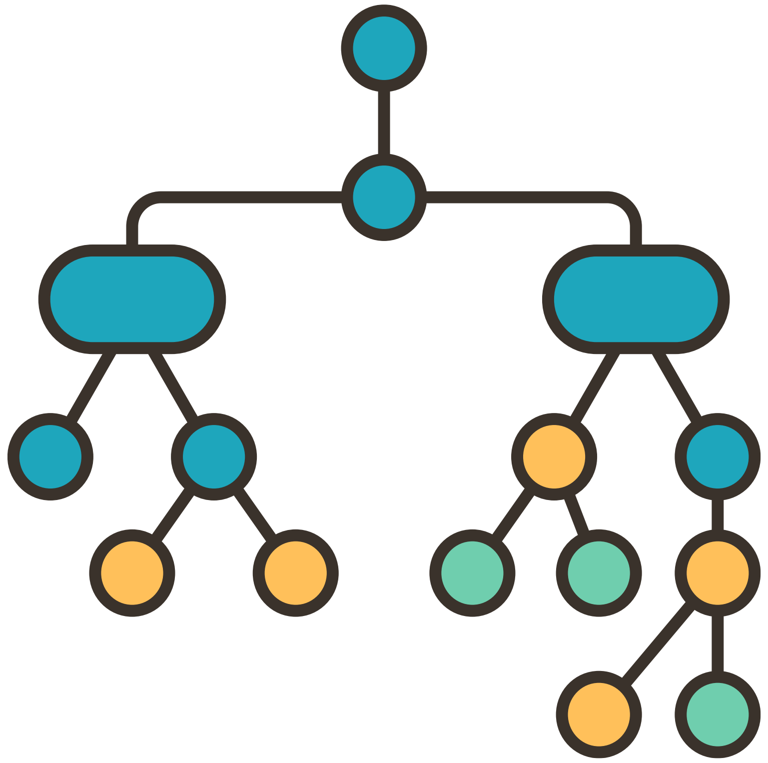
ARX



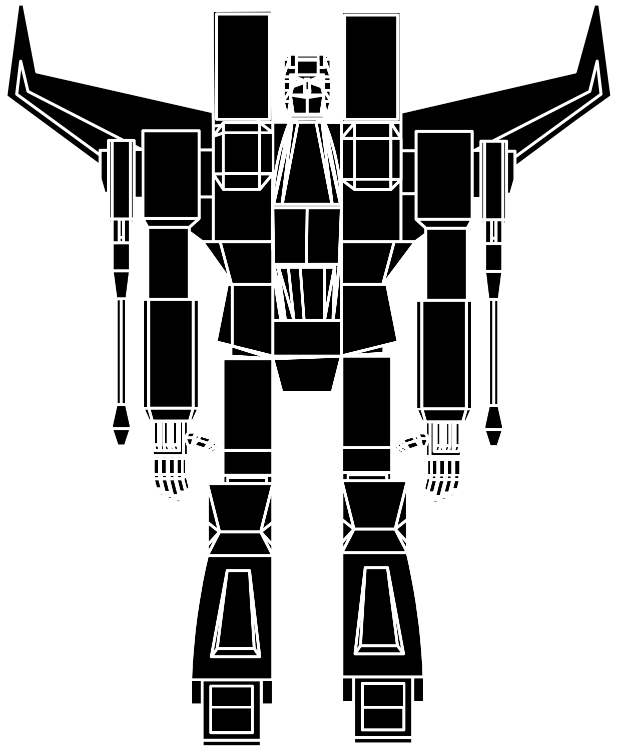
NARX



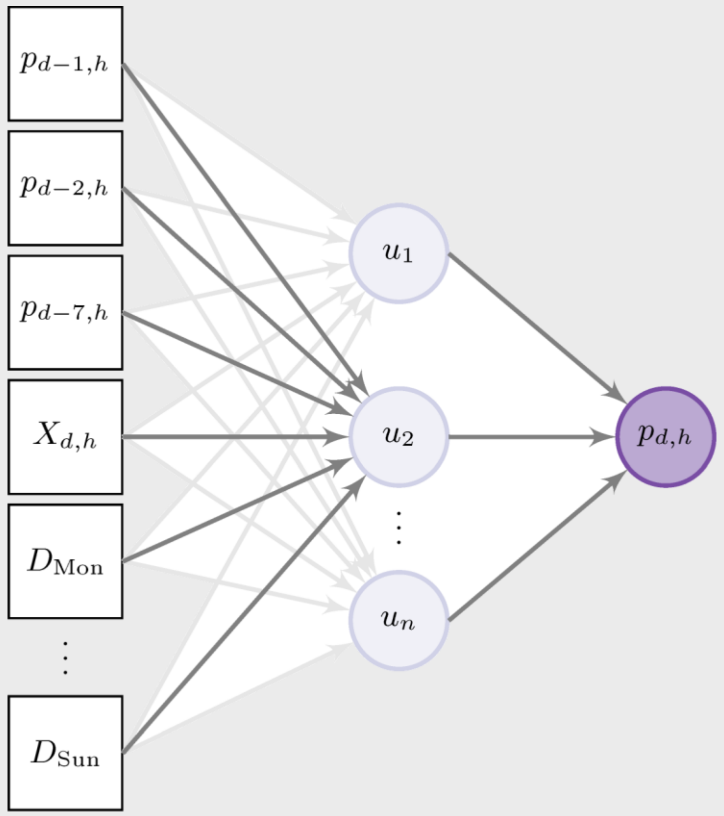
XGB



Mitra

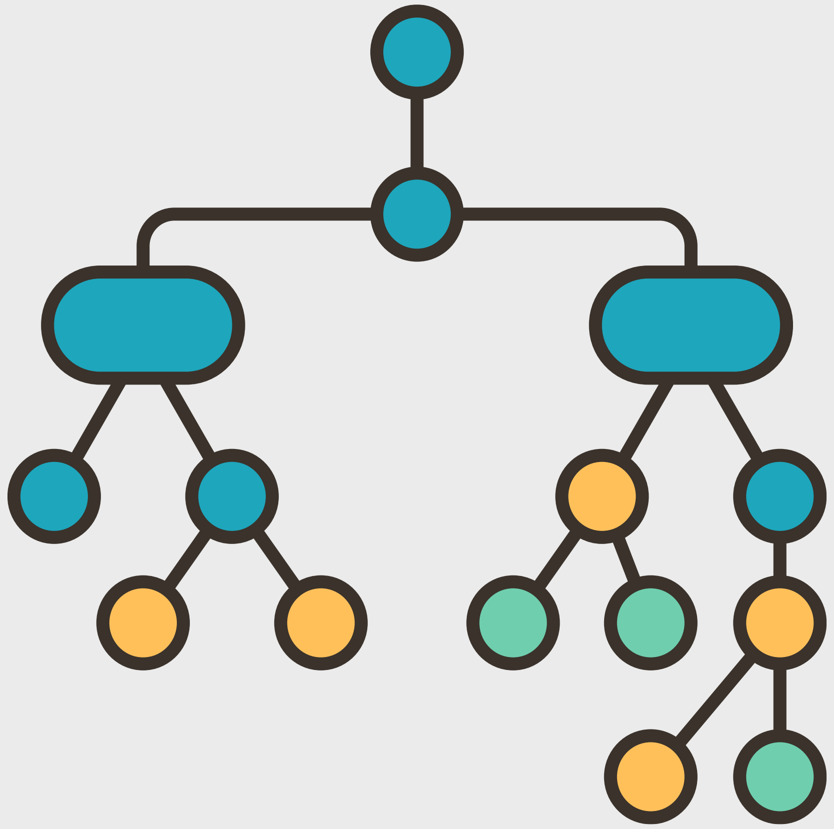


NARX (Non-linear ARX model)



- 5 neurons in the hidden layer
- Hyperbolic tangent activation functions
- Early stopping with 10% validation set
- Weights estimated with Levenberg-Marquadt in Matlab
- Final forecast is an average of 10 independently trained networks

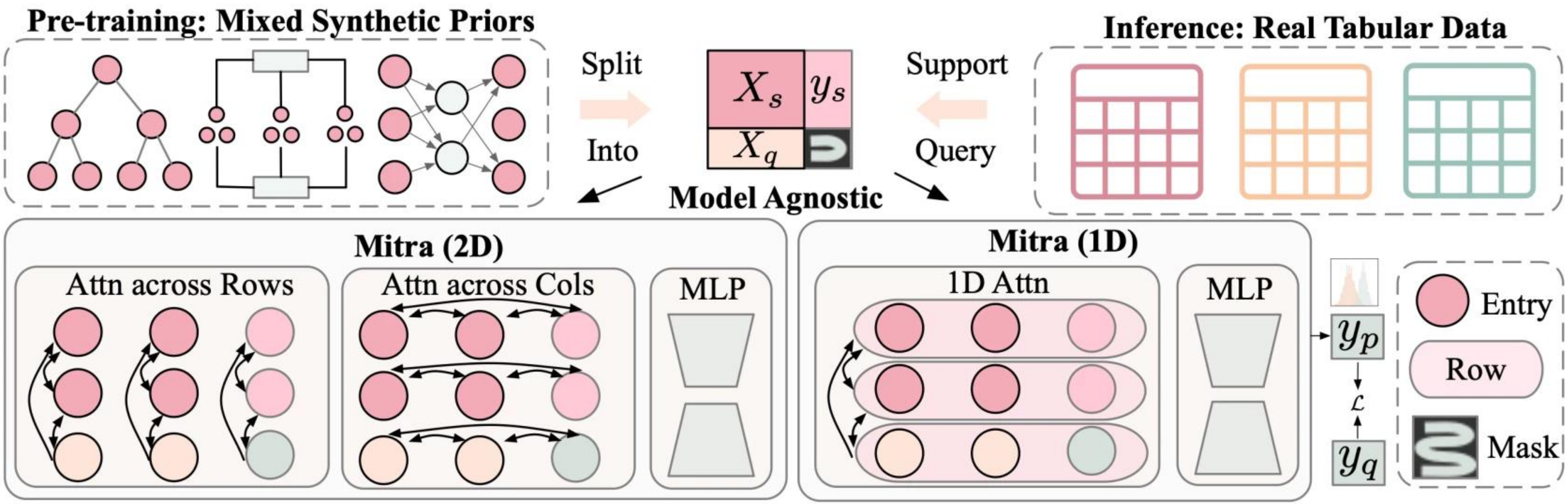
XGB (eXtreme Gradient Boosting)



- Python **XGBoost** package
- Mean squared error loss function
- Optuna-based hyperparameter optimization
- 10 independent runs once a year, separately for each hour/block
- Final forecast is an average of 10 XGB decision trees with different hyperparameters

Mitra

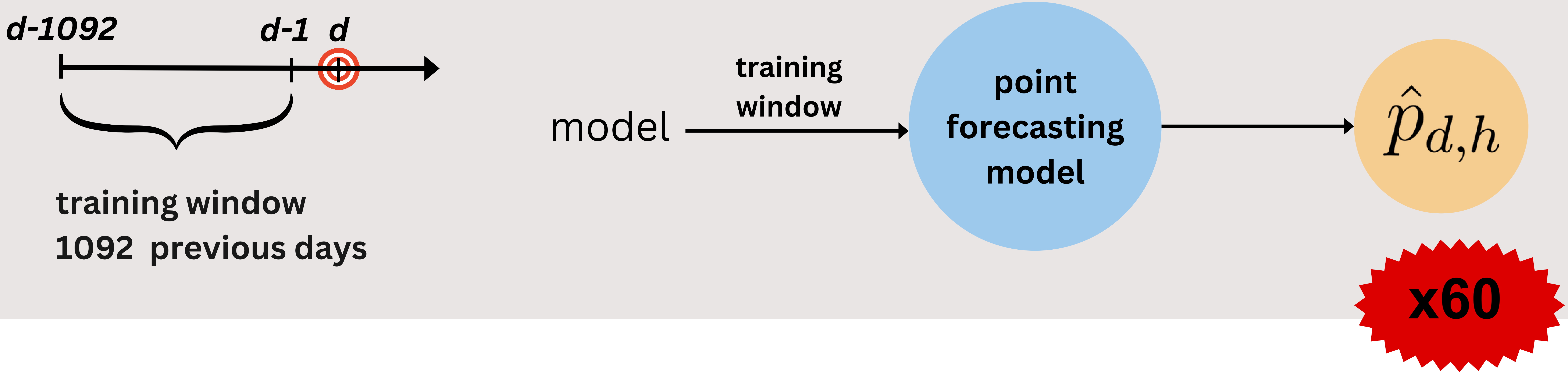
- 12-layer **transformer** with 72 million parameters
- Pretrained on synthetic data
- Zero-shot forecasting with in-context learning



amazon.science/blog/mitra-mixed-synthetic-priors-for-enhancing-tabular-foundation-models

Step I: Compute base forecasts

Estimate base models for block ***h*** of day ***d***



$$\bar{\mathbf{P}}_{\mathbf{d}} = \left[\hat{p}_{d,1}, \ldots, \hat{p}_{d,60}\right]^T$$

Unreconciled forecasts

Minimum trace (MinT) reconciliation

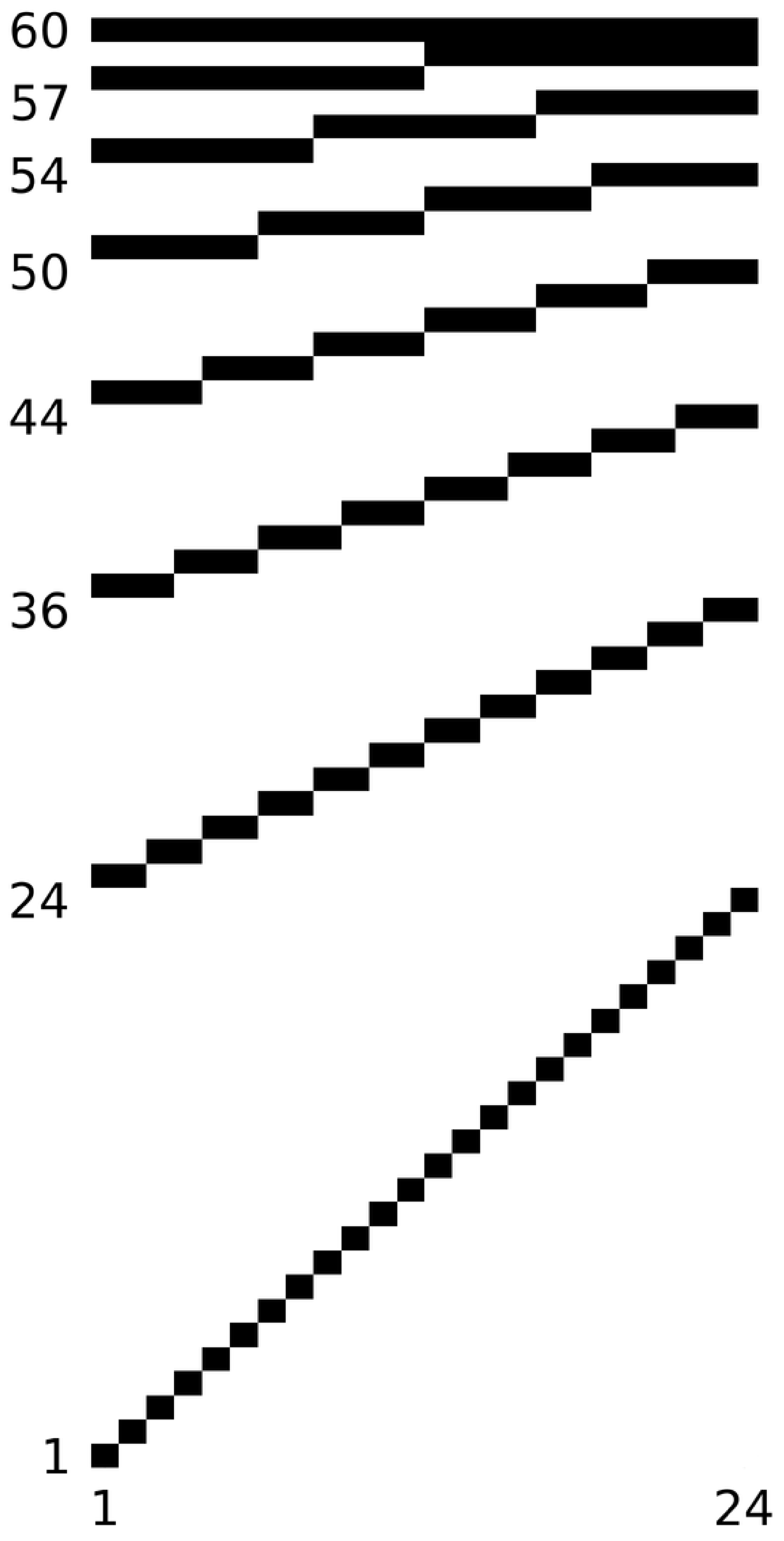
Reconciled forecasts

Error covariance matrix

$$\tilde{\mathbf{P}}_{\mathbf{d}} = \mathbf{S}(\mathbf{S}^T \boldsymbol{\Lambda}^{-1} \mathbf{S})^{-1} \mathbf{S}^T \boldsymbol{\Lambda}^{-1} \bar{\mathbf{P}}_{\mathbf{d}}$$

Summing matrix

Base forecasts

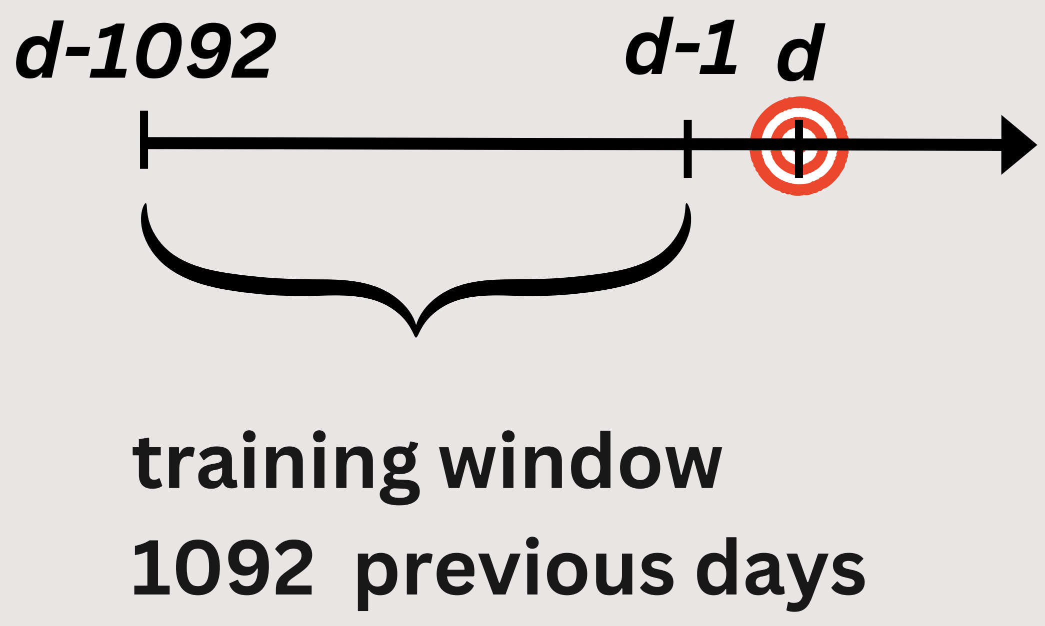


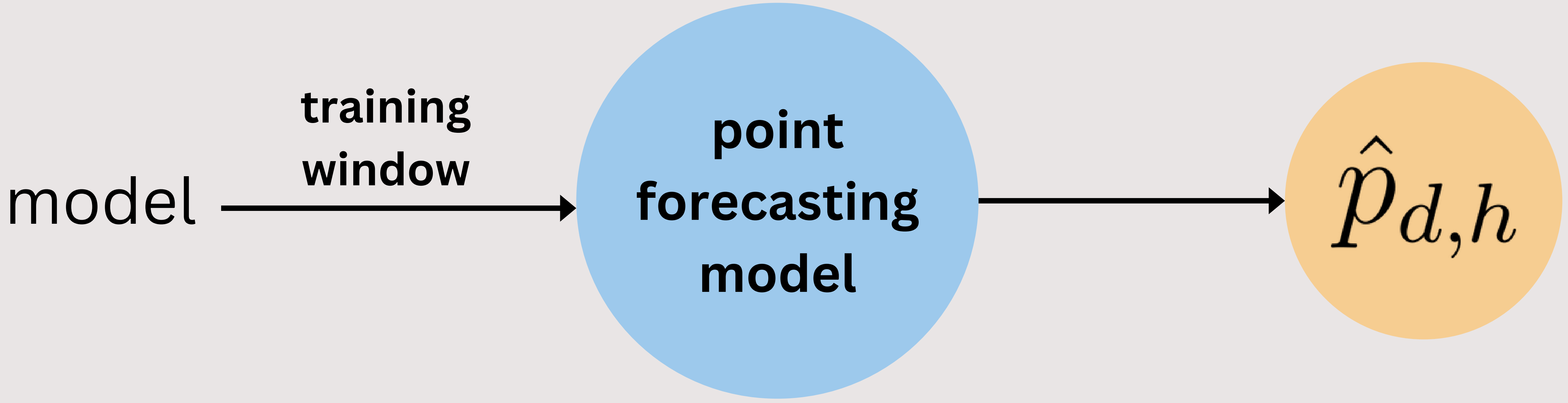
Summing matrix

Athanasopoulos et al. (2017, EJOR) Wickramasuriya et al. (2019, JASA)

Step II: Compute in-sample errors

Having estimated base models for block ***h*** of day ***d***

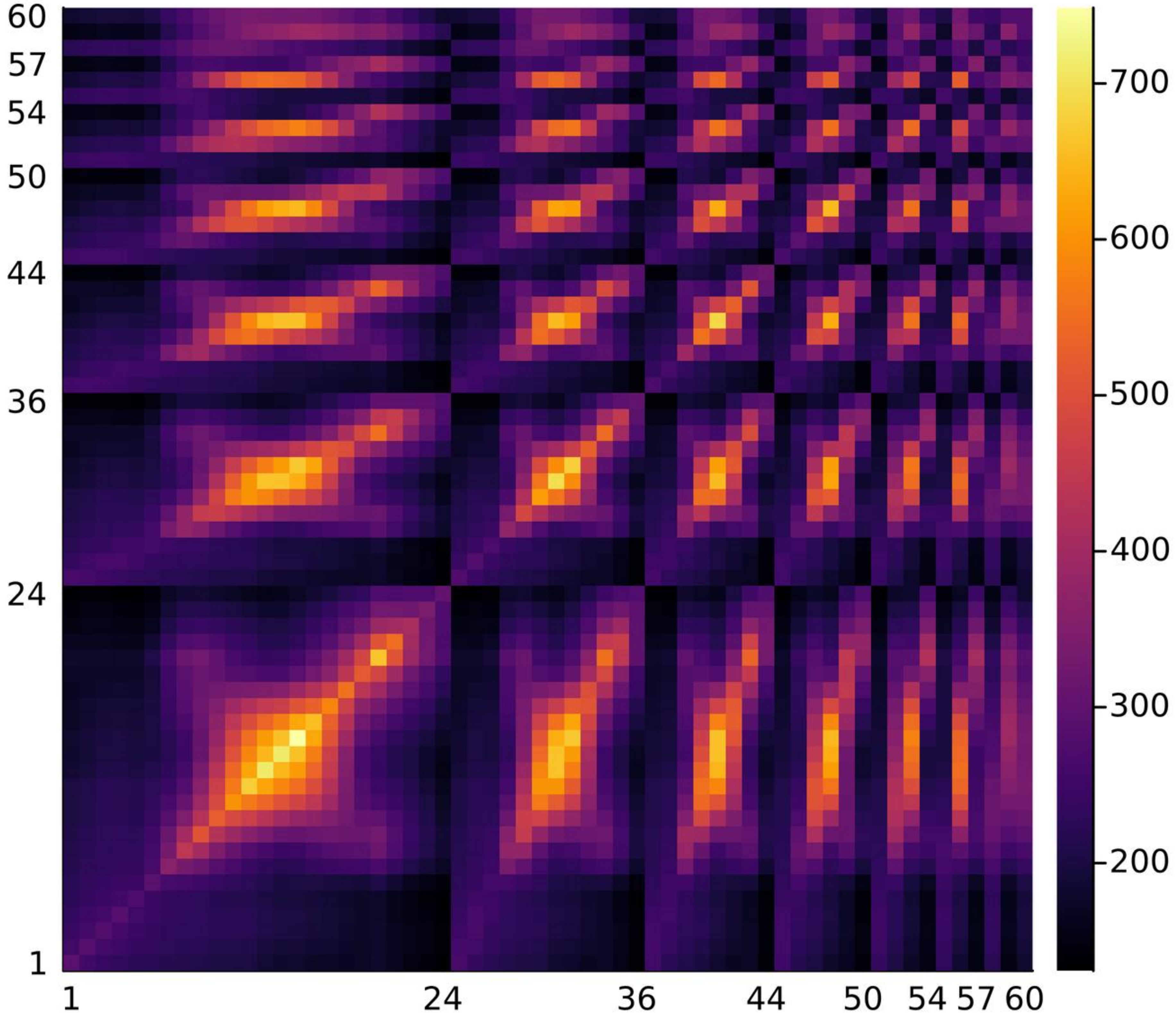




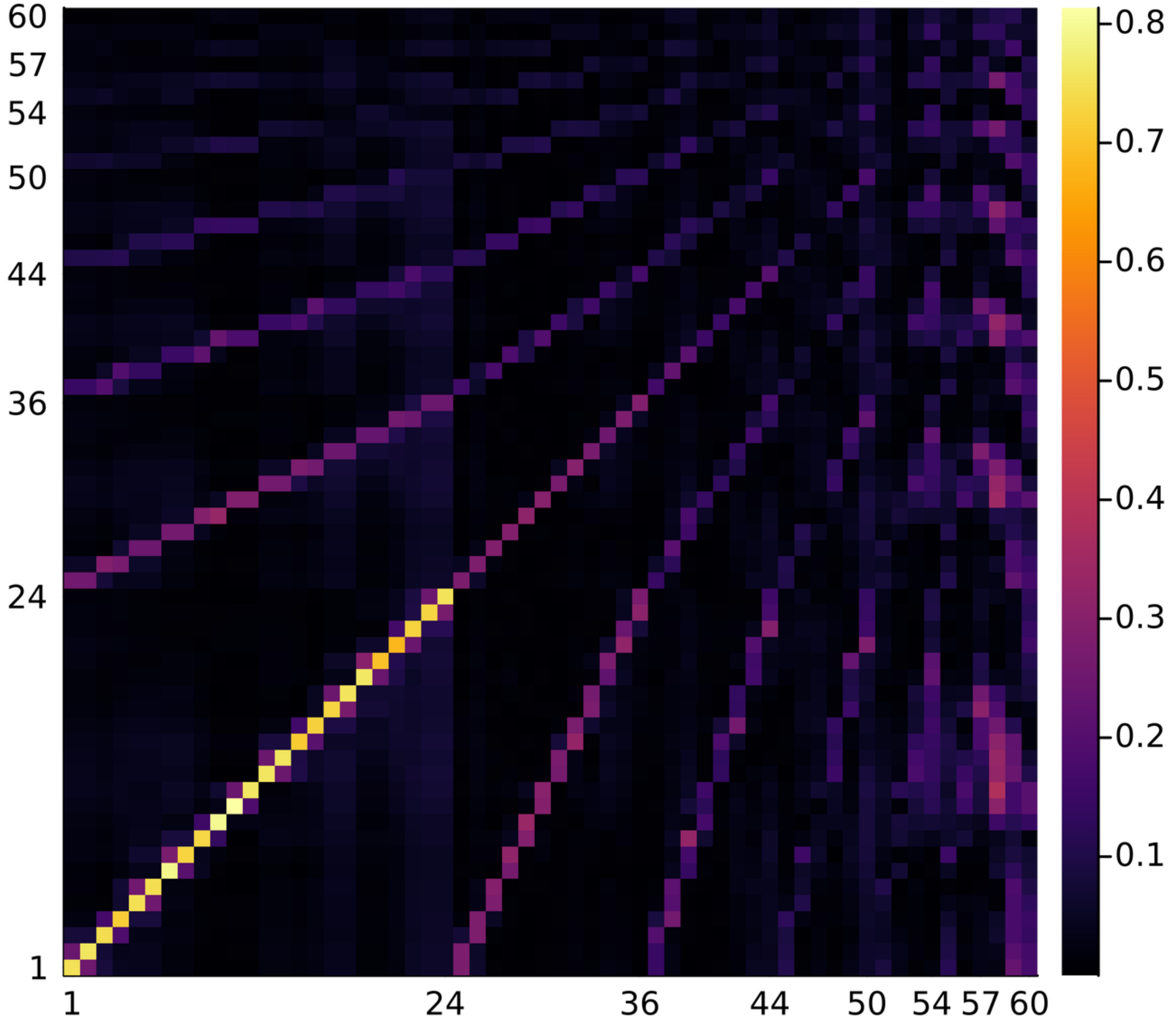
$$\mathbf{E_d} = \mathbf{P_d} - \bar{\mathbf{P}}_\mathbf{d} \xrightarrow[\text{Ledoit \& Wolf (2004, JPM)}]{\text{shrinkage}} \textbf{Forecast error covariance matrix}$$

Step III: Compute weights through MinT

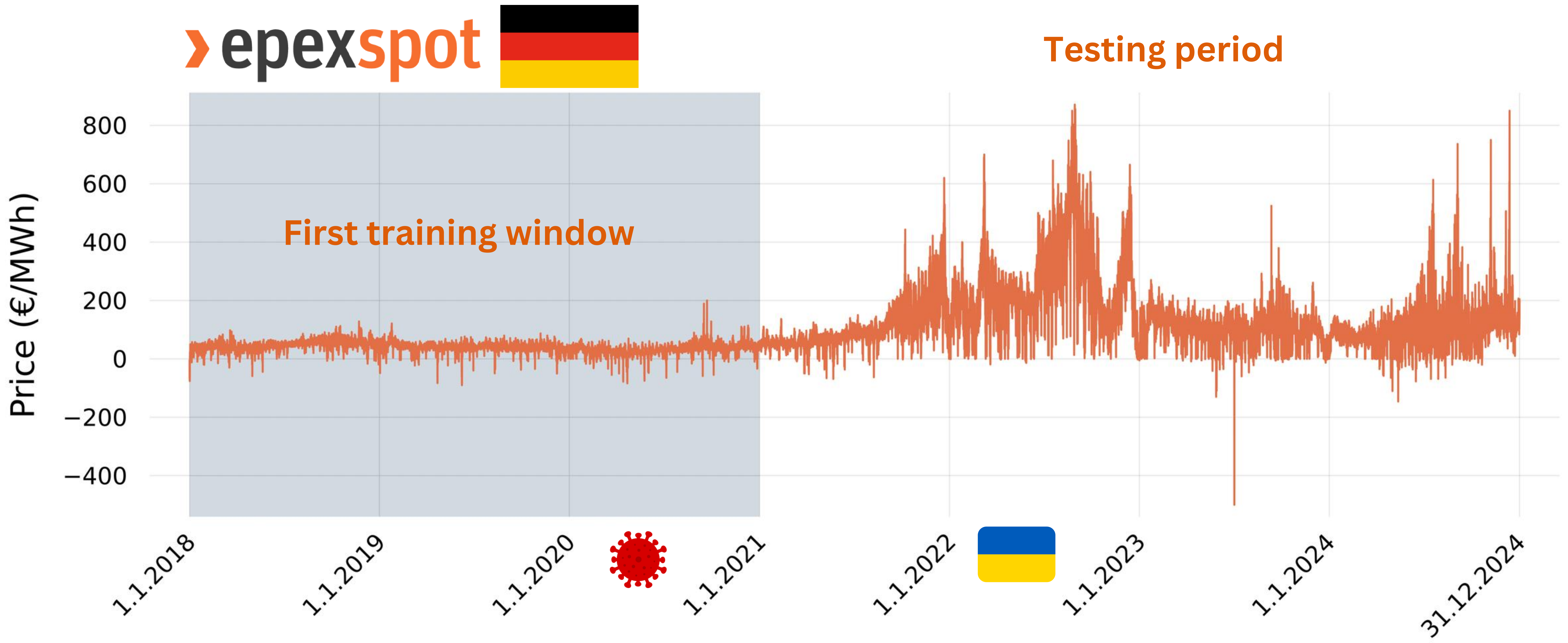
Error covariance



Estimated weights

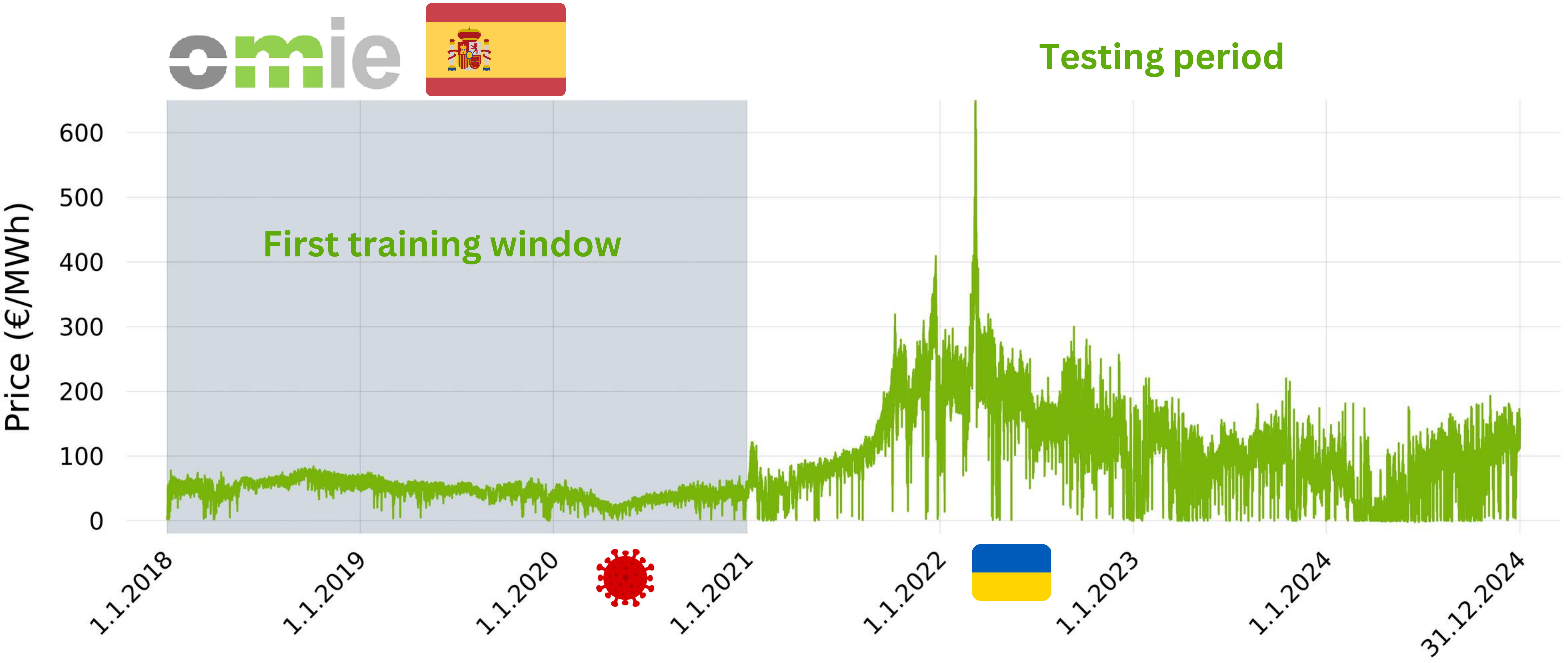


Data



epexspot
Testing period
First training window
Price (€/MWh)
800
600
400
200
0
−200
−400
1.1.2018
1.1.2019
1.1.2020
1.1.2021
1.1.2022
1.1.2023
1.1.2024
31.12.2024

Data

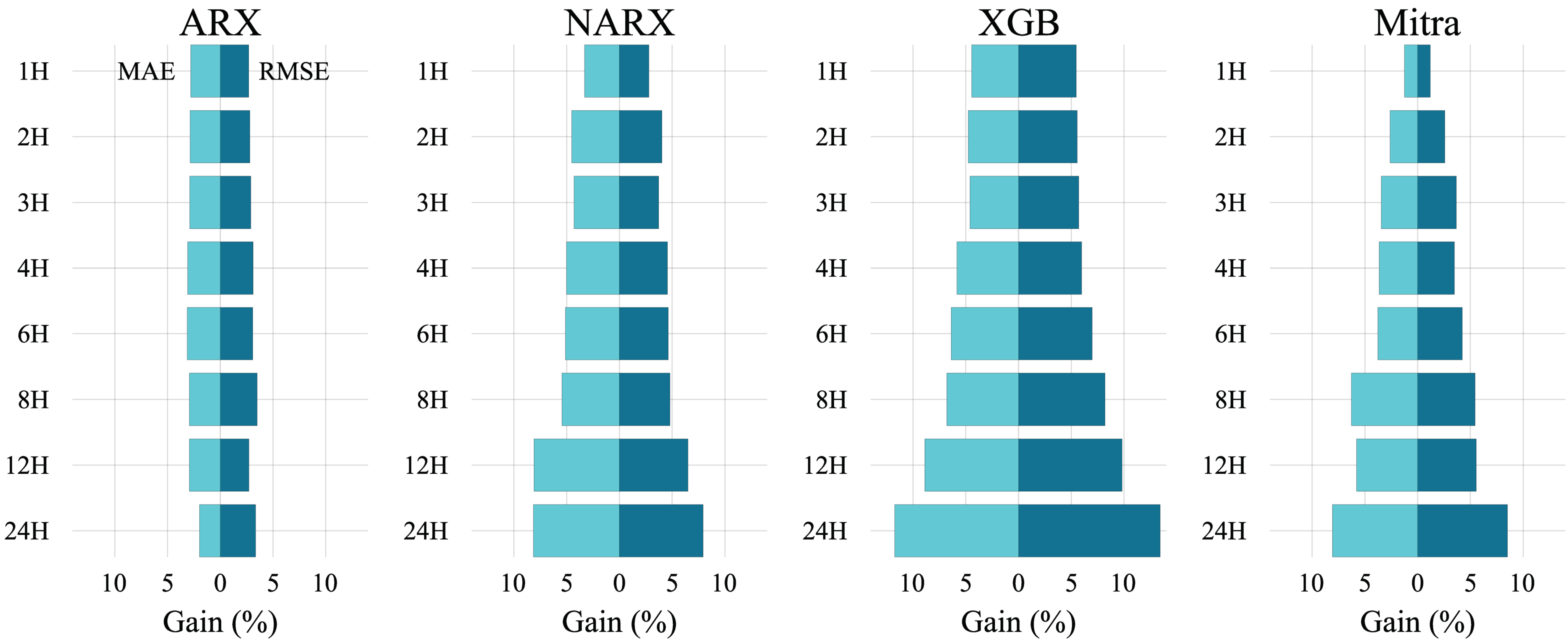


omie
Testing period
First training window
Price (€/MWh)
600
500
400
300
200
100
0
1.1.2018
1.1.2019
1.1.2020
1.1.2021
1.1.2022
1.1.2023
1.1.2024
31.12.2024

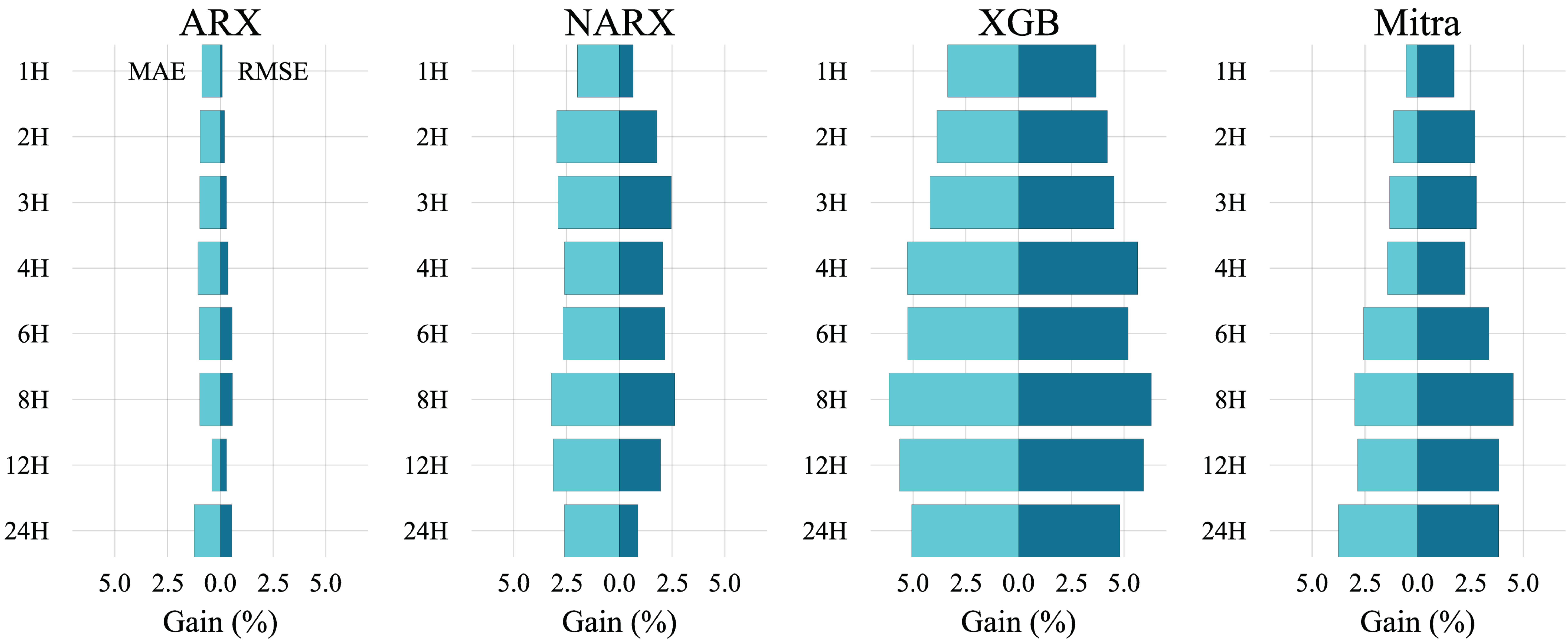
Results

Level	Model	Germany (EPEX-DE)						Spain (OMIE)					
		MAE	MAE_R	%	RMSE	$RMSE_R$	%	MAE	MAE_R	%	RMSE	$RMSE_R$	%
1H	ARX	26.93	26.17	2.8	39.83	38.75	2.7	16.72	16.57	0.9	24.40	24.38	0.1
	NARX	23.06	22.29	3.3	35.62	34.63	2.8	16.57	16.24	2.0	24.07	23.92	0.6
	XGB	23.33	22.30	4.4	37.05	35.02	5.5	16.54	15.98	3.4	24.57	23.67	3.7
	Mitra	21.37	21.11	1.2	33.31	32.91	1.2	15.92	15.84	0.6	23.61	23.20	1.7
2H	ARX	26.44	25.68	2.9	38.95	37.86	2.8	16.41	16.25	1.0	23.96	23.92	0.2
	NARX	22.64	21.61	4.5	34.84	33.44	4.0	16.29	15.80	3.0	23.71	23.29	1.8
	XGB	22.74	21.66	4.8	35.87	33.88	5.6	16.19	15.56	3.9	24.11	23.09	4.2
	Mitra	20.92	20.38	2.6	32.43	31.60	2.6	15.56	15.38	1.1	23.17	22.54	2.7
4H	ARX	25.65	24.85	3.1	37.50	36.33	3.1	15.88	15.71	1.1	23.25	23.17	0.4
	NARX	21.75	20.66	5.0	33.27	31.75	4.6	15.68	15.27	2.6	22.96	22.48	2.1
	XGB	22.01	20.72	5.8	34.24	32.19	6.0	15.78	14.95	5.3	23.60	22.26	5.7
	Mitra	20.13	19.39	3.7	30.86	29.79	3.5	15.01	14.79	1.4	22.19	21.70	2.2
8H	ARX	24.37	23.65	2.9	35.59	34.35	3.5	15.14	14.99	1.0	22.12	21.99	0.6
	NARX	20.54	19.42	5.4	31.20	29.71	4.8	15.01	14.52	3.2	21.92	21.35	2.6
	XGB	21.06	19.63	6.8	32.94	30.24	8.2	15.10	14.18	6.1	22.50	21.08	6.3
	Mitra	19.45	18.22	6.3	29.34	27.74	5.5	14.44	14.01	3.0	21.45	20.47	4.5
24H	ARX	20.92	20.50	2.0	30.28	29.27	3.4	13.57	13.40	1.2	19.50	19.39	0.6
	NARX	17.99	16.53	8.2	27.24	25.08	7.9	13.29	12.94	2.6	19.14	18.98	0.9
	XGB	18.80	16.59	11.7	29.50	25.54	13.4	13.25	12.58	5.1	19.61	18.66	4.8
	Mitra	16.56	15.22	8.1	24.81	22.70	8.5	12.91	12.42	3.8	18.84	18.11	3.8

Results


ARX
NARX
XGB
Mitra
MAE
RMSE
1H
2H
3H
4H
6H
8H
12H
24H
10
5
0
5
10
Gain (%)

Results



ARX
NARX
XGB
Mitra
MAE
RMSE
1H
2H
3H
4H
6H
8H
12H
24H
5.0 2.5 0.0 2.5 5.0
Gain (%)

[Submitted on 15 Aug 2025]

Stealing Accuracy: Predicting Day-ahead Electricity Prices with Temporal Hierarchy Forecasting (THieF)

Arkadiusz Lipiecki, Kaja Bilinska, Nicolaos Kourentzes, Rafal Weron