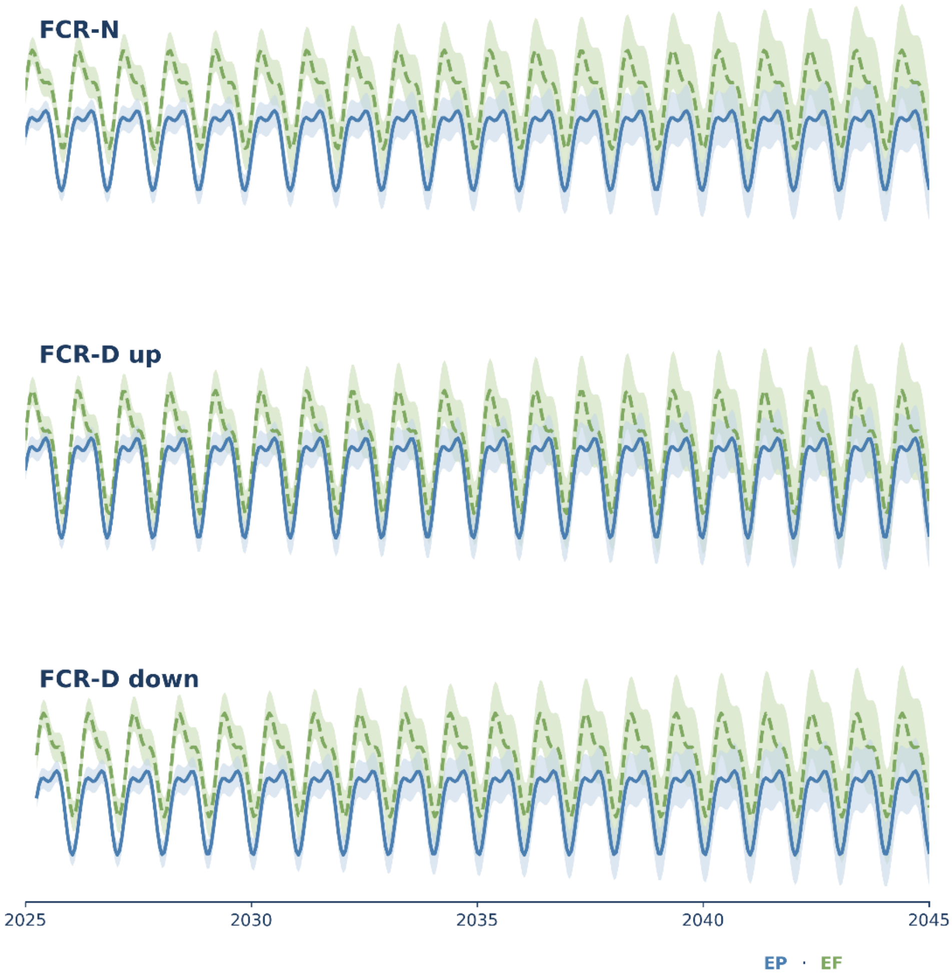
DIVISION BUILT ENVIRONMENT
ENERGY SYSTEMS ANALYSIS



Long-Term Forecasts of Frequency Containment Reserve

Gustaf Bengtsson, Erik Weihs

RISE Report 2026:49

Long-Term Forecasts of Frequency Containment Reserve

Gustaf Bengtsson, Erik Weihs

Abstract

Long-Term Forecasts of Frequency Containment Reserve

This study develops a machine-learning framework for long-term forecasting of prices in the Nordic Frequency Containment Reserve (FCR) markets. Accurate long-term price expectations are increasingly important for investors and system planners as the energy transition introduces higher variability in power systems and increases the need for flexibility resources.

The proposed approach applies a Temporal Fusion Transformer (TFT) model to historical balancing market data together with scenario-based electricity price trajectories from Svenska kraftnät's Long-term Market Analysis 2021 (LMA2021). Electricity spot price scenarios are treated as known future covariates, enabling the model to translate projected energy market conditions into long-term forecasts of ancillary service prices.

Historical data covering 2019–2025 were used for training and validation, including reserve prices, reserve volumes, generation mix, hydro storage levels, and electricity spot prices. The trained models were then applied to multiple weather-year realizations under two system development scenarios: Electrification with Renewables (EF) and Electrification with Plannable Generation (EP).

Results indicate that the model successfully captures broad seasonal patterns in reserve prices but struggles to reproduce extreme price spikes due to limited explanatory variables and short historical data series. Long-term forecasts suggest relatively stable price ranges with moderate variability between scenarios, although results remain highly dependent on the underlying electricity price assumptions.

The study demonstrates that scenario-driven machine-learning models can provide useful long-term insights into balancing markets, while also highlighting the importance of incorporating additional system drivers and longer historical datasets to improve forecasting robustness.

Key words: FCR-N, FCR-D, TFT, ML, Long-term, Forecasts, Scenarios

RISE Research Institutes of Sweden AB

RISE Report 2026:49
ISBN:978-91-90109-78-6
Göteborg 2025

Content

Abstract......1
Content......2
Preface......4
Summary......5
1 Background......6
1.1 Frequency Regulation Services in the Nordic Power System......6
1.2 Outlook to 2030......7
1.3 Long-term Electricity Price Scenarios......8
1.4 Why a Long-Term Forecasting Model for FCR?......9
2 Methodology......10
3 Input data......11
3.1 Data selection......11
3.1.1 Methodological Basis......11
3.1.2 Data sources......11
3.1.3 Dataset overview......12
3.2 Data handling......13
3.2.1 Data cleaning......14
3.2.2 Data adjustments......14
4 Model......16
4.1 State of the Art Time-Series Forecasting......16
4.2 Temporal Fusion Transformer......17
4.3 Model setup......17
4.3.1 Architecture......17
4.3.2 Feature integration......18
4.3.3 Target......19
4.3.4 Parameter tuning......19
4.4 Training process......20
4.5 Evaluation......20
4.6 Forecasting......21
4.7 Limitations......22
5 Results......23
5.1 FCR-N......24
5.1.1 Evaluation......24
5.1.2 Forecasts......26
5.2 FCR-D up......28
5.2.1 Evaluation......28

5.2.2 Forecasts 30
5.3 FCR-D down 32
5.3.1 Evaluation 32
5.3.2 Forecasts 34
6 Discussion 37
6.1 Input data and Methodological Considerations 37
6.2 Behaviour of Forecasts 37
6.3 Comparison with Historical Prices 38
6.4 Limitations and Directions for Future Work 39
Appendix 40

Preface

This report presents the work carried out within the project aimed at developing long-term forecasting models for the Swedish balancing market, specifically for Frequency Containment Reserve services (FCR-N, FCR-D up, and FCR-D down). The project was initiated in response to the growing challenges posed by an evolving energy system with increasing shares of variable renewable generation. These changes introduce greater volatility and flexibility requirements, which in turn affect the design and operation of ancillary service markets.

The work described here was funded by Energimyndigheten and supported by Svenska kraftnät, Vattenfall, Flower, Kungliga Tekniska Högskolan (KTH), Chalmers Tekniska högskolan, Skellefteå Kraft and Bixia elbolag. The report summarises the modelling approach, input data selection, and training process, as well as the results and limitations of the developed models. It also outlines future work and potential improvements.

The forecasting approach combines advanced machine learning techniques with established long-term electricity price scenarios provided by Svenska Kraftnät. By using these price prognoses as known variables, the model translates expected electricity spot price developments into balancing market price forecasts. This methodological choice ensures consistency with official system development scenarios but also introduces certain constraints, which are discussed later in the report.

We hope this report will serve as a useful reference for researchers and stakeholders interested in long-term market design, investment planning, and frequency stability.

Summary

This report presents the development and evaluation of machine-learning models for long-term price forecasting of frequency containment reserve (FCR) services in the Nordic balancing market, specifically FCR-N, FCR-D up, and FCR-D down. The work was carried out within the research project *Long-term price scenarios of frequency regulation services* and funded by Energimyndighetens project number P2022-00752.

The modelling approach is based on the Temporal Fusion Transformer (TFT), a deep learning architecture well-suited for multi-horizon time-series forecasting with mixed input types. A central methodological choice was to treat electricity spot price trajectories from Svenska kraftnät's Långsiktig Marknadsanalys 2021 (LMA2021) as known future covariates, allowing the model to translate official long-term energy market scenarios directly into balancing market price forecasts. Two scenarios were used — Electrification with Plannable Generation (EP) and Electrification with Renewables (EF) — applied across 21 weather years for the target year 2045.

Historical data covering 2019–2025 was used for training and validation, including reserve prices and volumes, electricity spot prices by bidding zone, generation mix, hydro storage, and prequalified ancillary service capacities. To support long-horizon forecasting and reduce model complexity, hourly data was resampled to daily aggregates (minimum, mean, and maximum). Seventeen model configurations were trained per target service, with hyperparameters tuned using Optuna. Model selection was based on a composite evaluation score combining MAPE, RMSE, MAE, MBE, skewness, and kurtosis, computed over a validation period spanning April 2024 to March 2025.

Results show that the models successfully capture broad seasonal patterns in all three reserve markets. FCR-N predictions follow a sinusoidal annual trend with summer peaks and winter troughs reasonably well. FCR-D up and FCR-D down results are more mixed, with most models overshooting during summer and struggling to reproduce sharp autumn price spikes. Across all services, the models consistently underestimate extreme price events, a known limitation of the available explanatory features and the short training history.

Long-term forecasts for 2045 suggest price ranges substantially lower than recent historical levels — generally 3–30 EUR/MW — reflecting the underlying electricity price assumptions in LMA2021 rather than recent market conditions. The EF scenario produces slightly higher average prices and greater variability than EP, consistent with expectations for a system with higher shares of variable renewable generation.

The study demonstrates that scenario-driven machine-learning models can provide useful long-term structural insights into balancing markets. Key directions for future work include expanding the set of input variables, exploring intermediate temporal resolutions, incorporating longer historical datasets as the FCR market matures, and benchmarking the TFT against alternative model architectures.

1 Background

The Swedish electricity system is undergoing a profound transformation driven by increasing electrification and the rapid deployment of variable renewable energy sources such as wind and solar power. This development introduces greater volatility and reduces system inertia, creating new challenges for maintaining frequency stability and ensuring reliable system operation. Ancillary services, particularly frequency containment reserves (FCR), play a critical role in balancing short-term deviations and safeguarding system security.

Traditionally, balancing markets have been designed for short-term operational needs, with limited consideration of long-term effects. However, as the energy transition accelerates, the demand for flexibility and frequency regulation is expected to grow significantly. Without robust long-term perspectives, there is a risk of inefficient resource allocation and increased societal costs for balancing. This project addresses the challenge of having a long-term perspective by developing forecasting models for scenarios in the Swedish balancing market, focusing on FCR-N and FCR-D (up and down).

Svenska kraftnät (SvK), the transmission system operator (TSO), has the statutory mission to balance the power system safely and cost-efficiently.

1.1 Frequency Regulation Services in the Nordic Power System

Frequency regulation services are essential for maintaining system frequency at 50 Hz. While all serve the same purpose, they operate on different time scales, with various volumes (power and energy) and activation principles. Following is an overview of the roles of each service:

- **Fast Frequency Reserve (FFR)**
  FFR is a very fast service designed for situations with low inertia in the grid, which are becoming more common as renewable generation increases. It is activated rapidly to support the system during sudden and deep frequency changes that can occur following faults when rotational energy in the Nordic power system is low.

- **Frequency Containment Reserves (FCR-N and FCR-D)**
  - **FCR-N (Normal operation):** Activated within seconds for small frequency deviations within the range 49.90–50.10 Hz, maintaining balance under normal operating conditions.
  - **FCR-D (Disturbances):** Activated equally fast for larger deviations outside this range, such as frequency drops to 49.50 Hz or rises to 50.50 Hz, to limit operational disturbances.

- **Automatic Frequency Restoration Reserve (aFRR)**
  aFRR takes over after FCR and operates automatically to restore frequency to 50 Hz within a few minutes. It is used for longer-term balancing of supply and

demand and is particularly important for managing medium-sized imbalances.

- **Manual Frequency Restoration Reserve (mFRR)**
mFRR is activated by system operators to relieve FCR and aFRR, handling longer-lasting and larger imbalances with activation times that can extend up to one hour. Note that automatic activation of mFRR was introduced in the Nordic markets in 2025.[1]

These services complement each other to ensure frequency stability across different time horizons. Table ITable I. contains technical information of requirements for each service [2].

Table I. Technical requirements of ancillary services on Svenska Kraftnät's balancing market.

Reserve	Min bid size	Activation	Activation time	Endurance / Repeatability
FFR	0.5 MW	Automatic at frequency changes, low rotational energy	Three options for 100%: 0.7s (49.50 Hz), 1.0s (49.60 Hz), 1.3s (49.70 Hz)	30s or 5s; Repeatability: 15 min
FCR-N	0.1 MW	Automatic linear within 49.90–50.10 Hz		30s or 5s; Repeatability: 15 min
FCR-D up	0.1 MW	Automatic linear within 49.50–49.90 Hz		At least 20 min
FCR-D down	0.1 MW	Automatic linear within 50.10–50.50 Hz		At least 20 min
aFRR	0.1 MW	Automatic at frequency deviation from 50.00 Hz	100% within 5 minutes	1 hour
mFRR	1 MW	Automatic based on forecasted imbalances per bidding zone	Full activation within 12.5 min, prep 2.5 min, ramp 10 min	15 min (scheduled), 30 min (direct)

1.2 Outlook to 2030

SvK's report *Balancing Market Outlook 2030*[3] provides projections for reserve requirements and market evolution. Future reserve needs are collected from SvK's published outlook *Framtida Volymbehov*[4]. Summarized in following list, an outline of future trends and needs looks like, beginning with current situation:

- **Historical trends:**
Balancing costs have increased significantly since 2019, driven by rising prices and volumes, partly due to new markets (e.g., mFRR capacity market) and increased FCR demand. Data up to September 2024 shows a 300% increase in procured capacity, with prices strongly correlated to spot prices and

[1] https://www.svk.se/aktorsportalen/bidra-med-reserver/om-olika-reserver/
[2] https://www.svk.se/siteassets/aktorsportalen/bidra-med-reserver/om-olika-reserver/fcr/fcr-technical-requirements-may-23.pdf
[3] https://www.svk.se/aktorsportalen/bidra-med-reserver/behov-av-reserver-nu-och-i-framtiden/framtidsrapport-om-balansmarknaderna/
[4] https://www.svk.se/aktorsportalen/bidra-med-reserver/behov-av-reserver-nu-och-i-framtiden/framtida-volymbehov/

competition.

- **Projected reserve needs:**
    - **FFR:** Highly energy system scenario-dependent, ranging from near zero to 2.5 times current levels. Estimated demand ranges from 1–48 GWh (2035).
    - **FCR-N:** ~224 MW, assumed stable but subject to Nordic reassessment.
    - **FCR-D up:** ~542 MW; FCR-D down: ~524 MW, based on reference incidents (Oskarshamn 3 and Nordlink/North Sea Link).
    - **aFRR:** Increase from 150 MW (2025) to 300 MW (2030), uncertainty range 160–400 MW.
    - **mFRR:** Upward capacity from 800 MW (2025) to 1400 MW (2030); downward from 990 MW to 1150 MW.
    -

- **Market evolution:**
  Introduction of 15-minute market time units, new IT platforms, and Nordic integration via PICASSO and MARI will reshape reserve activation and pricing mechanisms, reflecting ACE-based balancing.

- **Uncertainties:**
  Linked to production mix, electrification pace, and technology shifts. Market integration and competition are key to ensuring a robust system.

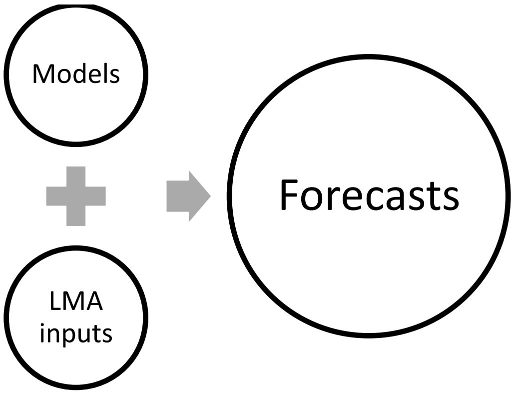
These projections highlight the need for long-term forecasting tools to support investment decisions in flexible resources such as battery storage and demand-side flexibility.

1.3 Long-term Electricity Price Scenarios

SvK's *Långsiktig marknadsanalys 2021*[5] (LMA 2021) sets out four scenarios for the Nordic and North-European power system to 2050—SF, FM, EP, EF—and publishes underlying Swedish input datasets (for 2030/2035/2040/2045/2050) for transparency and reuse. The scenarios explore different combinations of consumption growth and production mix, with annual additions of electricity production in the range 2–7.5 TWh per year (2025–2050). LMA2021 is explicitly not a point forecast, it is a scenario framework to analyse challenges along multiple pathways.

The LMA scenarios are produced with established market modelling and include, among other elements, electricity spot price trajectories for multiple weather years for each scenario year (e.g., 2035 and 2045).

[5] https://www.svk.se/4ab626/siteassets/om-oss/rapporter/langsiktig-marknadsanalys---lma/langsiktig-marknadsanalys-2021.pdf



A central methodological choice in this study is to treat SvK's scenario spot prices as "known" future covariates in our long-term forecasting of balancing service prices (FCR-N, FCR-D up/down). For each scenario (i.e., "Elektrifiering planerbart" and "Elektrifiering förnybart") and target year, we use the multi-weather-year price series to represent a distribution of plausible spot price paths, which the model then maps to ancillary service price outcomes.

1.4 Why a Long-Term Forecasting Model for FCR?

The increasing complexity of balancing markets and their interaction with energy markets requires advanced modelling approaches. This project adopts a scenario-based methodology aligned with SvK's long-term electricity spot price prognoses for 2045. By using these prognoses as known variables, the model translates expected spot price developments into forecasts for balancing market prices under different future scenarios. This ensures consistency with official system development assumptions while enabling robust long-term price trajectories for FCR-N, FCR-D up/ down.

2 Methodology

This study relies on timeseries data available from SvK's *Långsiktig Marknadsanalys 2021*[6] (LMA2021) as the guiding framework for long-term electricity market development. The model design and input feature selection were aligned with the assumptions and data structure provided in the LMA2024 to ensure consistency with official planning scenarios. It should be noted that during the project, a newer *Långsiktig Marknadsanalys 2024*[7] *(LMA2024)* was released. However, as this was not available at the start of the project, the electricity price scenarios from the earlier *LMA2021* where used.

We use scenario data for the year 2045, specifically the two scenarios "Elektrifiering planerbart" (EP) and "Elektrifiering förnybart" (EF), as defined in LMA2021. Each scenario includes multiple weather years, resulting in a distribution of electricity spot prices across weather years rather than a single deterministic price path — each weather year produces one hourly price series, and together they represent the range of plausible spot price outcomes under that scenario. These multi-weather-year datasets provide a confidence interval for future price expectations.

The electricity spot price trajectories from LMA2021 act as known covariates in the applied ML model. By anchoring the model to these scenario-based price paths, we ensure that ancillary service price forecasts (FCR-N, FCR-D up, FCR-D down) reflect the same underlying assumptions as SvK's long-term market analysis.

Input features were limited to variables available in the LMA2021 dataset and corresponding historical time-series data (e.g., spot prices, production mix, hydro storage). This constraint ensures methodological transparency and alignment with the scenario framework.

The results and assumptions in LMA2021 serve as guidance for our modelling approach, influencing both the choice of scenarios and the interpretation of forecast outputs. This alignment strengthens the relevance of our forecasts for strategic planning and investment decisions.

Technical specifications and model applications are further described in section 3 and 4.

[6]https://www.svk.se/4ab626/siteassets/om-oss/rapporter/langsiktig-marknadsanalys---lma/langsiktig-marknadsanalys-2021.pdf
[7] https://www.svk.se/siteassets/om-oss/rapporter/2024/lma_2024.pdf

3 Input data

The input data forms the foundation of the Forecasting model and plays a critical role in ensuring consistency with SvK's Long-term Marked Analysis. Historical datasets were used for model training and validation, while future projections were guided by scenario data from LMA2021. The following subsections describe the selection criteria, methodological basis, and data sources.

3.1 Data selection

The primary criterion for selecting input data was alignment with the LMA2021 dataset. To maintain methodological consistency, the model was restricted to variables available in LMA2021, except for a few essential market-specific variables related to frequency containment reserves FCR. Variables inherited from LMA2021:

- Electricity Spot Prices per bidding zone
- Production data (Hydro, Wind, Solar, Nuclear, Other)
- Stored Hydro reserves

Essential market specific variables:

- Balancing service prices and volumes
- Prequalified volumes for ancillary services

This approach ensures that the model reflects the same structural assumptions as LMA2021 while incorporating key variables for forecasting ancillary service prices.

3.1.1 Methodological Basis

A central methodological decision was to use electricity price prognoses from LMA2021 as known variables in the model. Specifically:

- **Scenario Year:** 2045
- **Scenarios Used:** EP and EF
- **Weather Years:** Multiple weather-year datasets were included to capture variability and provide a confidence interval for future price expectations.

By anchoring the model to these scenario-based trajectories, we ensure consistency with SvK's assumptions and enable scenario-driven forecasts for ancillary service prices. This alignment strengthens the relevance of the results for strategic planning and investment decisions.

3.1.2 Data sources

Historical data were collected from the open sources presented in Table II below.

Table II. Input data sources.

Source	*Data types*
ENTSO-E Transparency Platform[8]	Electricity spot prices, generation mix, and hydro storage data.
Svenska Kraftnät (Mimer)[9]	Frequency reserve market prices and volumes.
Svenska Kraftnät[10]	Prequalified capacities for ancillary services
Statnett[11]	Reference for Nordic reserve requirements

Collected data plotted as a graph per data type can be found in Figure 1 below.

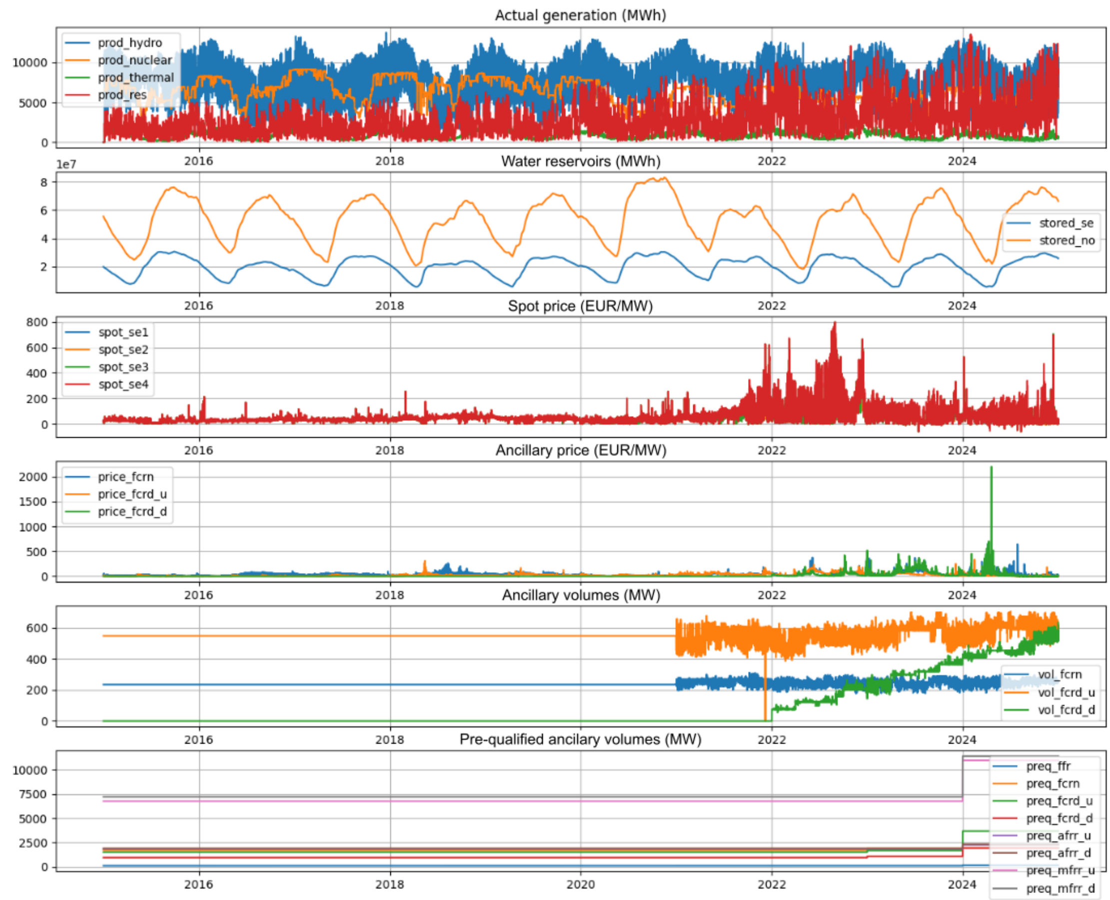


Figure 1. Input data collected for the project (note that data prior 2021 is missing for ancillary volumes and prequalified ancillary volumes).

3.1.3 Dataset overview

The training dataset with included variables were structured as follows:

---

[8] https://transparency.entsoe.eu
[9] https://mimer.svk.se/PrimaryRegulation/PrimaryRegulationIndex
[10] https://www.svk.se/aktorsportalen/bidra-med-reserver/behov-av-reserver-nu-och-i-framtiden/utbud-pa-marknaderna-for-reserver/
[11] https://www.statnett.no/globalassets/for-aktorer-i-kraftsystemet/systemansvaret/reservemarkeder/overview-of-frequency-control-in-the-nordic-power-system.pdf

Table III. Overview of dataset with temporal resolution and covered period. Note that it is training data used.

Type	Variable name	Resolution	Period
Balancing service price data	price_fcrn	1h	2019-01-01 – 2025-03-31
	price_fcrd_u	1h	2019-01-01 – 2025-03-31
	price_fcrd_d	1h	2019-01-01 – 2025-03-31
Balancing service volume data	vol_fcrn	1h	2019-01-01 – 2025-03-31
	vol_fcrd_u	1h	2019-01-01 – 2025-03-31
	vol_fcrd_d	1h	2019-01-01 – 2025-03-31
Stored energy hydro reserves	stored_se	1h	2019-01-01 – 2025-03-31
	stored_no	1h	2019-01-01 – 2025-03-31
Electricity spot price data	spot_se1	1h	2019-01-01 – 2024-12-31
	spot_se2	1h	2019-01-01 – 2025-03-31
	spot_se3	1h	2019-01-01 – 2025-03-31
	spot_se4	1h	2019-01-01 – 2025-03-31
Production data	prod_hydro	1h	2019-01-01 – 2025-03-31
	prod_nuclear	1h	2019-01-01 – 2025-03-31
	prod_other	1h	2019-01-01 – 2025-03-31
	prod_solar	1h	2019-01-01 – 2025-03-31
	prod_wind	1h	2019-01-01 – 2025-03-31
	prod_total	1h	2019-01-01 – 2025-03-31
Prequalified volume	preq_ffr	1h	2021-01-01 – 2025-03-31
	preq_fcrn	1h	2021-01-01 – 2025-03-31
	preq_fcrd_u	1h	2021-01-01 – 2025-03-31
	preq_fcrd_d	1h	2021-01-01 – 2025-03-31
	preq_afrr_u	1h	2021-01-01 – 2025-03-31
	preq_afrr_d	1h	2021-01-01 – 2025-03-31
	preq_mfrr_u	1h	2021-01-01 – 2025-03-31
	preq_mfrr_d	1h	2021-01-01 – 2025-03-31
Datetime	datetime	1h	2019-01-01 – 2025-03-31

Historical data available before 2019 was available for the project but deemed unusable, since ancillary service market data showed limited imitation to latter behaviour.

Data from 2025 was kept separate and used for evaluation.

3.2 Data handling

This section describes the steps taken to prepare and transform the input data for model training and forecasting. The process included cleaning raw data, handling missing values, and applying feature engineering techniques to enhance model performance.

3.2.1 Data cleaning

To ensure data quality and consistency, the following procedures were applied:

- **Duplicate Removal:**
  Rows containing identical values across all columns were deleted.
- **Missing Value Handling:**
  Missing values were addressed using a two-step approach:
  1. **Forward propagation:** The last valid observation was carried forward to fill gaps.
  2. **Backward propagation:** Remaining gaps were filled by propagating the next valid observation backward.

This approach ensured continuity within the dataset and preserved temporal consistency.

3.2.2 Data adjustments

Several adjustments and feature engineering steps were performed to improve model interpretability and forecasting accuracy.

3.2.2.1 Calendar-Based Features and Cyclical Feature Encoding

Making use of variable Datetime, the dataset was enriched with calendar components as new features. The new features were generated from datetime data by extracting information from the existing datetime variable.

- **Month of the Year**: Captures seasonal patterns.
- **Week of the Year**: Identifies yearly weekly cycles.
- **Day of the Year**: Highlights long-term yearly trends.
- **Weekday**: Detects behavioural shifts based on specific days.
- **Workday or Weekend**: Differentiates patterns related to business days and weekends.

Since time-related features exhibit periodic behaviour, sine and cosine transformations were applied to encode cyclical components. These transformations preserve continuity, ensuring smooth transitions between time periods. Given a cyclical feature x with a period max(x) (e.g., 12 months for months of the year), two new features can be generated:

1. **Sine Transformation**: $x_{sin} = sin(\frac{2\pi \times x}{max(x)})$
2. **Cosine Transformation**: $x_{cos} = cos(\frac{2\pi \times x}{max(x)})$

These transformations map the cyclical feature on a unit circle, preserving the cyclical nature of the data. Both cyclical features and calendar components added to the dataset are listed in Table IV below.

Table IV. Added calendar based and cyclical features to the input dataset.

Type	Variable name	Resolution	Period
Time index range	time_idx	1h	2019-01-01 – 2025-03-31

Type	Variable name	Resolution	Period
Calendar data	month	1h	2019-01-01 – 2025-03-31
	week	1h	2019-01-01 – 2025-03-31
	day	1h	2019-01-01 – 2025-03-31
	weekday	1h	2019-01-01 – 2025-03-31
	workday	1h	2019-01-01 – 2025-03-31
	weekend	1h	2019-01-01 – 2025-03-31
Periodic encoded	sin_month	1h	2019-01-01 – 2025-03-31
temporal data	cos_month	1h	2019-01-01 – 2025-03-31
	sin_week	1h	2019-01-01 – 2025-03-31
	cos_week	1h	2019-01-01 – 2025-03-31
	sin_weekday	1h	2019-01-01 – 2025-03-31
	cos_weekday	1h	2019-01-01 – 2025-03-31

3.2.2.2 Decomposed target

To enhance model performance, the target was decomposed into three components using a LOESS smoothing model[12]:

- **Trend:** Captured the long-term directional movement of price data.
- **Seasonality:** Identified repeating patterns at monthly and weekly intervals.
- **Residuals:** Represented unexplained variations and random fluctuations.

The new decomposed variables, generated by applying the LOESS model, adjusted for annual trend, monthly and weekly seasonality was added to the dataset to support pattern recognition. The generated decomposed targets, added to the dataset, are listed in Table V below.

Table V. Added decomposed targets to input dataset.

Type	Variable name	Resolution	Period
Decomposed target	annual_trend	1h	2019-01-01 – 2025-03-31
	weekly_seasonality	1h	2019-01-01 – 2025-03-31
	monthly_seasonality	1h	2019-01-01 – 2025-03-31
	residuals	1h	2019-01-01 – 2025-03-31

3.2.2.3 Resampling input data

The project's objective to predict future price of balancing services beyond 2030, towards 2050, reduces the requirement of high-resolution time series. In combination with model limitations of preserving accuracy over multiple timesteps, the decision to resample input data to daily aggregates was made.

12 https://www.statsmodels.org/v0.14.4/examples/notebooks/generated/mstl_decomposition.html

The original hourly data was resampled into daily aggregates, specifically the minimum, mean, and maximum values for each day. This down sampling was necessary to align with the forecasting model's objective: predicting balancing service prices over long horizons, potentially spanning an entire year.

By reducing the dataset size, computational load was decreased, minimizing the number of training iterations required and reducing the risk of error feedback accumulation. This approach also filters short-term noise, allowing the model to focus on broader trends rather than volatile hourly fluctuations, thereby improving generalization.

After down sampling, all numerical values were aggregated, concatenated, and sorted based on type of aggregate (minimum, mean, maximum), ensuring a structured dataset ready for model input. Values were identified, by using grouping by aggregate and used for down sampling in a dataset of lower resolution. Added variable is listed in Table VI below.

Table VI. Aggregates added to the dataset, sorted on minimum, mean, and maximum aggregates day.

Type	Variable name	Resolution	Period
Aggregates	aggregate	1D	2019-01-01 – 2025-03-31

4 Model

This section describes the modelling approach used to forecast long-term prices for ancillary services on the Swedish balancing market. It begins with an overview of state-of-the-art methods in time-series forecasting and explains why the Temporal Fusion Transformer (TFT) was chosen for this study. The section then details the model setup, how scenario-based inputs from SvK were integrated, and the training and evaluation process. Finally, it outlines the limitations of the approach and considerations for future improvements.

4.1 State of the Art Time-Series Forecasting

Forecasting time-series data has evolved rapidly with deep learning, enabling forecasts of complex phenomena such as electricity spot prices. Traditional statistical models like ARIMAX have long supported forecasting with external factors, known as exogenous variables. However, these models often struggle with complex patterns and larger datasets.

Modern approaches use neural networks to capture relationships between multiple variables, a technique called multivariate forecasting. Among these, Transformer-based models, originally developed for language processing, are now widely used because they can learn long-term dependencies in data[13].

Recent Transformer (among others) variants include:

[13] https://arxiv.org/abs/2402.19072

- **Informer and Autoformer:** Designed for efficiency when handling long sequences.
- **PatchTST:** Groups time steps into patches to capture local patterns.
- **iTransformer:** Focuses on relationships between different variables.

While these models improve accuracy, they often lack interpretability and can be complex to deploy. Newer methods such as TFT and TimeXer aim to better integrate external information but remain primarily research-focused.

4.2 Temporal Fusion Transformer

The Temporal Fusion Transformer (TFT) was selected because it combines strong predictive performance over longer periods with interpretability, which is essential for practical applications like electricity market forecasting.

TFT is designed for multi-horizon forecasting, meaning it predicts several future time steps at once[14]. It can work with three types of inputs:

- **Past observations**, which are historical data points,
- **Future known inputs**, such as SvK's electricity long-term market analysis scenarios,
- **Static features**, which are values that do not change over time, like aggregate (min, mean, and max).

Unlike many advanced models that act as "black boxes," TFT provides interpretability through so called attention weights, revealing some information of which variables influence any prediction the most. This helps understanding how factors like spot prices or production mix affect reserve prices.

The architecture combines feature selection and attention mechanisms to capture long-term patterns while filtering out irrelevant data. It also supports probabilistic forecasts, which means it can provide a range of possible outcomes rather than a single prediction.

4.3 Model setup

This section outlines the design and configuration of the forecasting model, including its underlying architecture, the integration of historical and scenario-based features, the definition of target variables, and the strategy for hyperparameter tuning. Together, these elements describe how the Temporal Fusion Transformer was adapted to predict long-term prices for frequency containment reserves under SvK's scenario framework.

4.3.1 Architecture

The forecasting model is based on the TFT architecture as implemented in PyTorch Forecasting[15]. The TFT relies on recurrent layers, applying a Long Short-Term Memory

---

[14] https://pytorch-forecasting.readthedocs.io/en/v1.4.0/api/pytorch_forecasting.models.temporal_fusion_transformer._tft.TemporalFusionTransformer.html
[15] https://pytorch-forecasting.readthedocs.io/en/stable/

(LSTM) encoder-decoder, to capture short-term patterns in timeseries. This is combined with multi-head attention that identifies long-range dependencies across the entire historical window. Simply, this design allows the model to learn from both local variations in the frequency containment reserve's market prices and broader seasonal or scenario-driven trends. This means that the model can build on known future covariates such as the SvK long-term market analysis data alongside historical variations for more accurate forecasting.

The LSTM encoder-decoder process sequences of past observations, such as daily or weekly variations in ancillary service's market prices. While the multi-head attention enables the model to focus on the most relevant historical points and features such as winter price spikes or hydro reservoir levels that influence reserve prices months later.

The forecasts are probabilistic, optimized using Quantile Loss across seven quantiles (split between percentiles withing the range: p10–p90). This means that the model provides prediction intervals rather than point estimates per forecast. Though all probabilistic forecasts are stored, all results are presented for the 50th percentile, reducing the amount of data points presented.

4.3.2 Feature integration

The model integrates three categories of features to ensure realistic forecasting and alignment with scenario-based assumptions.

Static features include scenario information, which remains constant throughout the forecast horizon. In this project, those are the aggregates: min, mean, and max.

Known time-varying features comprise calendar encodings (month, week, weekday), historical data and scenario-driven covariates such as electricity spot price, and production data from SvK's long-term market analysis. These are available for both historical and future periods, enabling the model to learn from and forecast on available scenario paths (EF and EP).

The **Unknown time-varying** features include historical balancing service prices and volumes, hydro storage, and production data, which are only observed in the past and serve as predictive signals during training.

A complete list of model feature integration is presented in Table VII below.

Table VII. Feature setup in the model.

Type	*State*	*Awareness*	*Value type*
Aggregate	Static		Categorical
Calendar data	Time varying	Known	Categorical
Time index range	Time varying	Known	Reals
Electricity spot price data	Time varying	Known	Reals
Periodic encoded temporal data	Time varying	Known	Reals
Balancing service price data	Time varying	Unknown	Reals
Balancing service volume data	Time varying	Unknown	Reals
Stored energy hydro reserves	Time varying	Unknown	Reals
Production data	Time varying	Unknown	Reals
Prequalified volume	Time varying	Unknown	Reals

State and Awareness define how the model handles each variable during training, evaluation, and forecasting. Value type specifies how the data is interpreted: categorical values represent discrete states, while real values are treated as numerical inputs processed by the encoder-decoder

4.3.3 Target

The model is set up for prediction of a target variable. The project wishes to generate a picture of future prices on the frequency containment reserve market, thus targeting prices of FCR-N, FCR-D down, and FCR-D up when training the model. The models are limited to one target per setup, meaning that one ancillary service is target and the process of training is done three times.

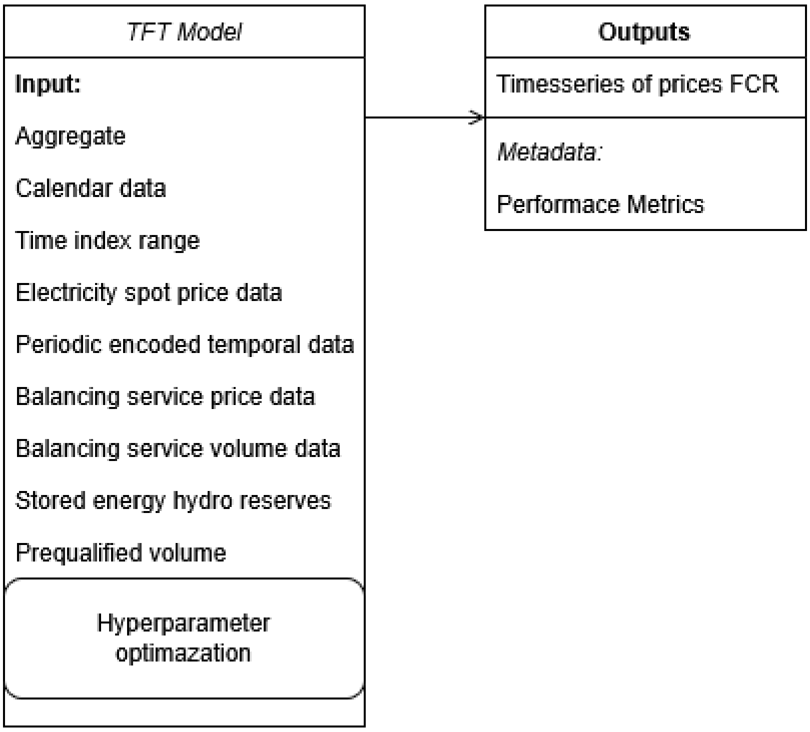


Figure 2. Simple overview of the model setup, including features and the parameter settings from the optimization.

4.3.4 Parameter tuning

Initially, all models were tuned with by a hyperparameter search using Optuna[16], optimizing against a minimum Quantile Loss, identifying the lowest validation metric per optimization trail. Each study was configurated as following:

- Number of trails: Each study was iterated at least twice (per target), ensuring robustness by capturing different optimal parameter combinations and limiting impacts from early converging.
- Optimization duration: 100 trials per target. A maximum of 100 epochs per trial.
- Early stopping: Training halts if validation loss stagnates for 10 consecutive epochs, monitored at every epoch.

[16] https://optuna.org/

- Hardware: Executed on a CPU (Intel Core Ultra 7) due to computational constraints.

To improve reproducibility, results from all optimization trials were stored and analysed, selecting the best-performing configurations based on quantile loss reduction. While quantile loss was the primary optimization metric, additional metrics such as mean absolute error (MAE), root mean squared error (RMSE) and mean absolute percentage error (MAPE) were considered for post-study validation.

4.4 Training process

Per target, each model was trained to generate probabilistic forecasts with a 28-day horizon. Training was conducted using PyTorch Lighting's Trainer[17], ensuring structured optimization, storing and logging. Training Configuration:

- Training duration: Limited to 200 epochs for convergence.
- Hardware: Executed on a CPU (Intel Core Ultra 7) due to computational constraints.
- Validation: Quantile Loss
- Early stopping: Training halts if validation loss stagnates for 10 consecutive epochs, monitored at every epoch.
- Encoder length: 112
- Decoder length: 28 (same as the forecast horizon)
- Training split: 2024-03-31 (limiting training to dates prior)

A total of 17 models for each target was carried out with various hyper parameter settings, using the results from the top performers from optimization stage, as well as parameter presets from the model designer as benchmarks. All hyper parameter settings used in each training session are presented in Appendix Table IX.

4.5 Evaluation

Model predictions of balancing service prices for a period spanning over 2024-03-31 to 2025-03-31 was used for validation. Each trained model's generated frequency containment reserve price forecast was evaluated against the actual price, which were resampled and aggregated over the same period.

To identify the best-performing models, a comprehensive evaluation was conducted using multiple metrics quantifying model fit against actual data of corresponding prices over the validation period, scoring the fit's ability to respond to various market behaviours. The included metrics are listed and described in Table VIII below.

All metric scores were normalized using min-max scaling across models, grouped by the price aggregates (a), to ensure comparability and mitigate the influence of scale differences due to market variation.

Each normalized Metric scores were weighted to reflect their relative importance in assessing model performance, with a focus on penalizing large errors. In other words, normalized metric score were multiplied by its corresponding weight (ω). A composite

[17] https://lightning.ai/docs/pytorch/stable/common/trainer.html

score per aggregate was calculated as the product of every normalized and weighted metric. A final composite score for each model was then calculated by multiplying the composite scores across the three aggregates (daily min, mean, and max price) according to Equation 1, ensuring balanced performance across all price dimensions. This multiplicative approach penalizes models that perform poorly on any single aggregate, promoting consistency across daily price fluctuations. Selected weights are presented in Table VIII below.

Table VIII. Metrics used to evaluate models, along with descriptions and selected weight.

Metric (M)	*Abbreviation*	*Description*	*Weight* (ω)
Mean Absolute Percentage Error	MAPE	Emphasizes interpretability and relative error	0.3
Root Mean Square Error	RMSE	Penalizes large errors	0.25
Mean Absolute Error	MAE	Measures overall accuracy	0.2
MBE	MBE	Detects systematic bias	0.15
Skewness	-	Assesses residual distribution asymmetry	0.05
Kurtosis	-	Evaluates the shape of the residual distribution	0.05

$$(1)\ Composite\ score_a = \prod_{m \in M} (\omega_m \times Metric\ score_{\,a,m})\ \ where\ a \in \{min, mean, \max\}$$

The weighting significantly influences the final rankings. Models that consistently minimize relative errors (MAPE) and large deviations (RMSE) are favoured due to their higher weights (0.30 and 0.25, respectively). This benefits models that are robust across varying price levels and avoid extreme prediction errors. Conversely, models that may perform well on average (low MAE or MBE) but occasionally produce large outliers are penalized more heavily. The inclusion of Skewness and Kurtosis, though lightly weighted, ensures that models with poorly distributed residuals (e.g., heavy tails or asymmetry) are slightly downgraded, promoting models with more stable and symmetric error distributions. The full composite score was calculated as:

$$(2)\ Final\ Composite\ score = \prod_{a \in \{min, mean, max\}} Composite\ Score_a$$

Models were ranked based on their final composite scores, calculated as the product of each aggregates composite score. The top three models, i.e. the models with the lowest final composite score, for each balancing service were selected for future use in forecasting prices beyond year 2025.

4.6 Forecasting

After model selection, the best-performing three models per target were applied to SvK's long-term market analysis scenarios. Known future covariates such as spot price trajectories and scenario metadata were injected for each weather year, enabling multi-horizon forecasts of FCR prices under different system development pathways.

The forecast horizon of 28 days limits each forecast to the same number of days meaning that for a full year each model needs to be forecasted 13 times per weather year, and the results concatenated into a year's full time series. This was done for every weather year in the trajected SvK's scenario data, for both scenarios EF and EP. A final weekly average of the aggregate's minimum, mean and maximum was generated to create intervals of possible future price ranges per model.

4.7 Limitations

Forecasts rely on SvK's long-term market analysis for future spot price trajectories and system conditions. This ensures consistency with official scenarios but introduces dependency on their accuracy and assumptions.

The model treats electricity spot prices as fixed known covariates rather than predicting them jointly with reserve prices. This simplifies the approach but omits potential feedback effects between energy and balancing markets.

Historical input data of frequency containment reserves are limited to a few years containing both price, volumes and prequalified volumes, resulting in a dataset useful for training spanning over a few years. This impacts the TFT models ability learn from broader seasonal or scenario-driven trends. Market maturity also play a crucial role in providing reliable results, where training data showed great variations per year and none to little resemblance of annual reoccurring trend per target, see Figure 3,Figure 4,Figure 5 below.

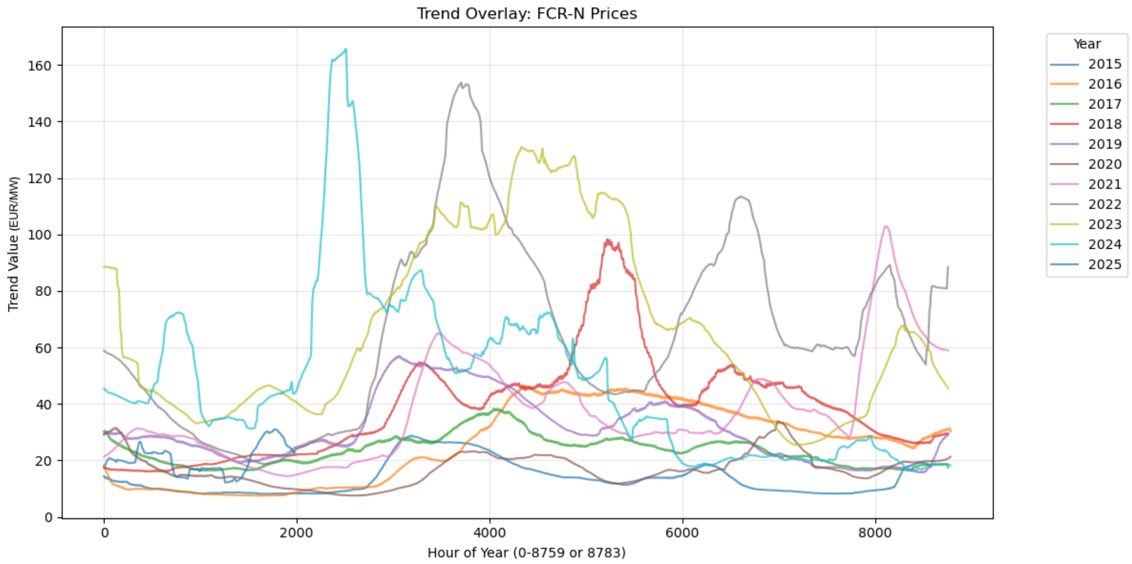


Figure 3. Annual FCR-N Price trend (2015-2025). Decomposed using LOESS, filtering for trend.

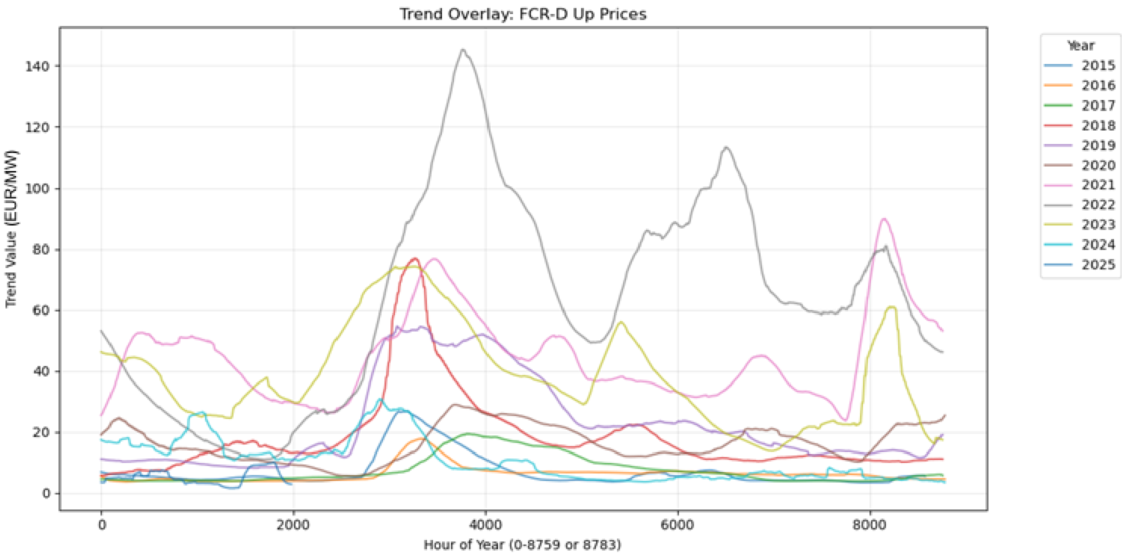


Figure 4. Annual FCR-D up Price trend (2015-2025). Decomposed using LOESS, filtering for trend.

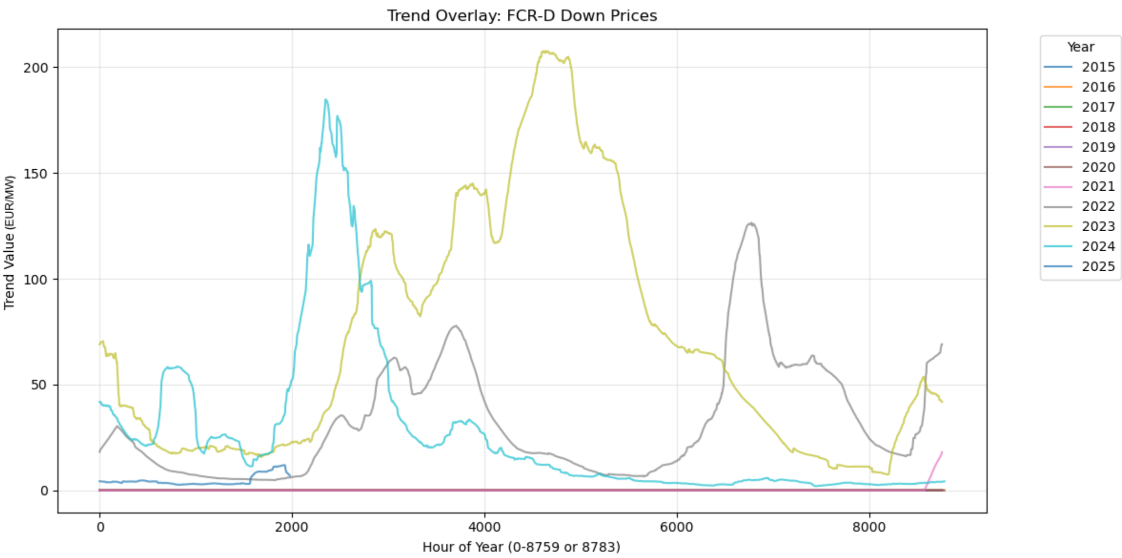


Figure 5. Annual FCR-D down Price trend (2021-2025). Decomposed using LOESS, filtering for trend.

Number of years with historical data vary, with only three years of available data exists for FCR-D market data. This limits the number of data points for training, since one year is allocated for evaluation.

Input data was down sampled to daily aggregates to reduce complexity and improve long-horizon stability. While this supports scenario-based forecasting, it limits the model's ability to capture short-term volatility and intraday variations.

5 Results

This section presents the results of the forecasting models developed for long-term price prediction of frequency containment reserves (FCR-N, FCR-D up, and FCR-D down). The results include both model evaluation on historical data and scenario-based forecasts for 2045 under SvK's long-term market analysis. For each service, we first assess model performance during the validation period using composite scores and visual comparisons of predicted versus actual prices. We then present scenario-based forecasts, highlighting weekly price trajectories across multiple weather years, under two system development pathways: Electrification plannable (EP) and Electrification renewables (EF).

5.1 FCR-N

The 17 models trained to target FCR-N weekly prices, plotted together with actual values from the period 2024-03-31 to 2025-03-31 are shown in Figure 6 below. The figure contains the results of each aggregate for both predictions and actuals. The results show that the main price trend is captured by the predictions, except for any high spikes in prices. The price forms a sinusoidal with a peak in July and a bottom in February

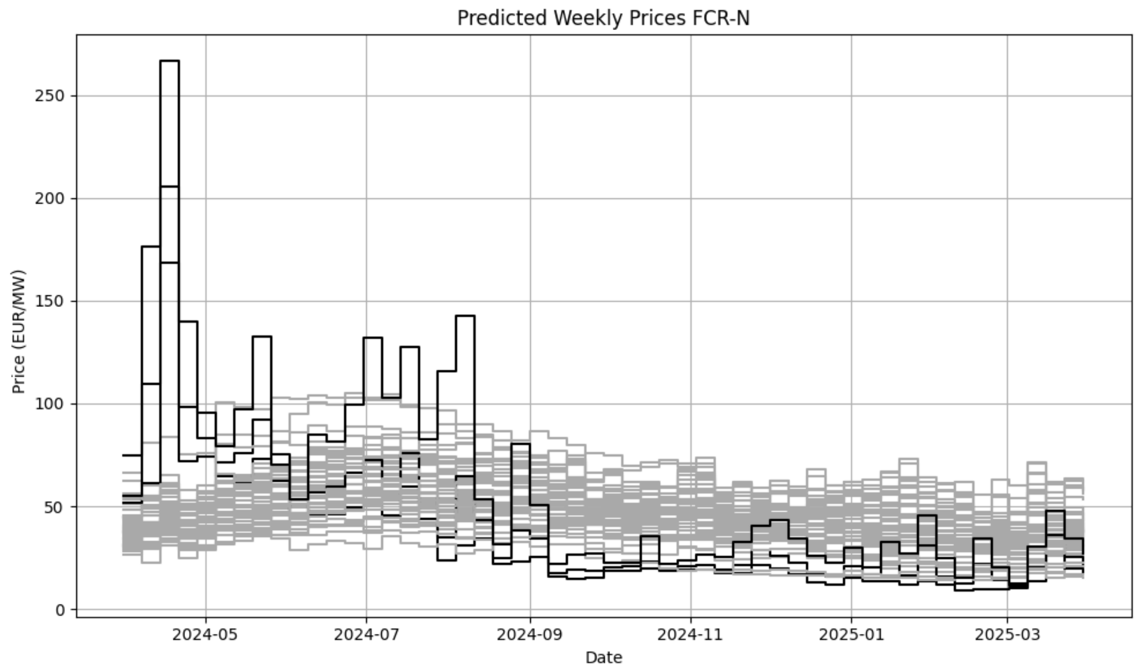


Figure 6. Predicted weekly prices for FCR-N for all 17 models (grey) and actual (black) over the validation period. Notice that label on the y-axis is in MWh, when it should be MW.

5.1.1 Evaluation

To limit the results, three top-ranked models based on final composite scores were selected. This approach ensures focus on models with balanced and consistent performance across all aggregates. The best-performing models are fcrm_6, fcrm_2, and fcrm_8.

Figure 7 presents composite and final scores for a subset of trained models evaluated with actual prices of FCR-N. Each evaluation includes:

- Composite scores of aggregates' (min, mean, and max) price predictions (blue, orange, green)
- A final score (grey), calculated as the product of the three

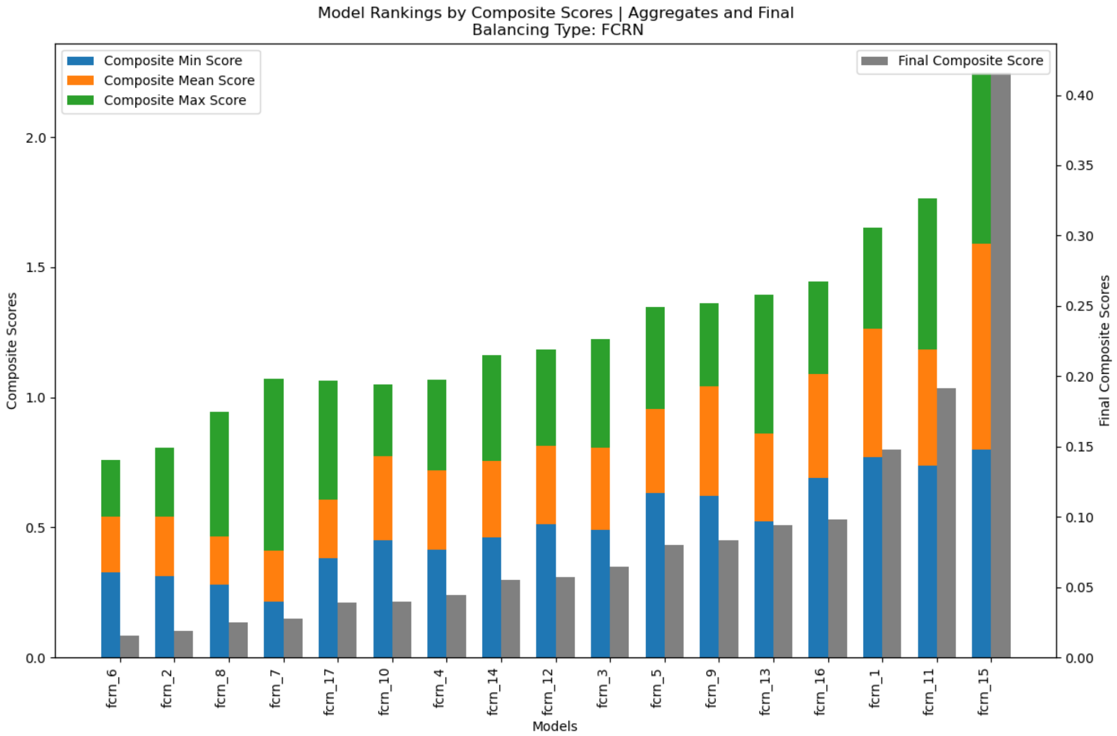


Figure 7. Composite scores per aggregate, with a final score in grey (product of each aggregate's score).

Figure 8 shows predicted weekly mean prices of FCR-N from 2024-03-31 to 2025-03-31, filtered to the three best-performing models (fcrn_6, fcrn_2, fcrn_8) based on final composite scores. All models capture the overall trend, with actual prices dropping sharply early in the period and stabilizing later. The models vary slightly in how closely they follow price changes.

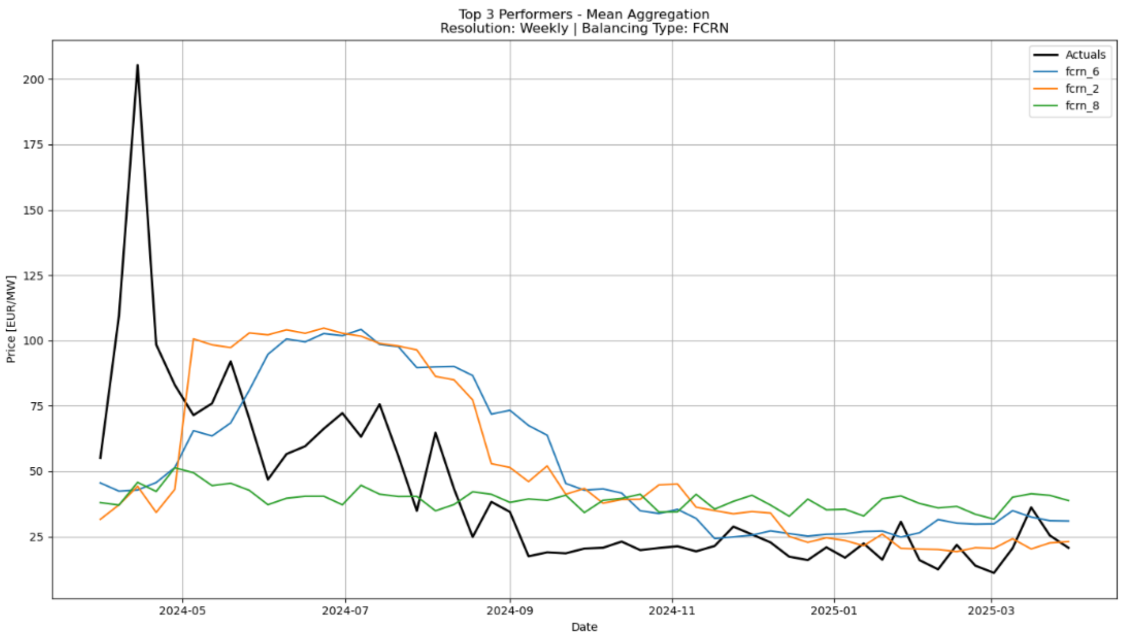


Figure 8. Top three performing models' prediction of weekly mean prices over the validation period.

While the top-performing models (fcrn_6, fcrn_2, fcrn_8) capture the overall downward trend in weekly mean prices, they struggle to adapt to sharp spikes and rapid changes, such as the surge observed in May 2024. This indicates a limited responsiveness to short-term volatility, likely due to weak correlation between the input features and sudden price movements. These results suggest that external or unmodeled factors may be driving price spikes, which are not captured in the training data.

5.1.2 Forecasts

Figure 9 illustrates forecasted weekly mean prices for FCR-N across 21 weather years under both scenarios: Electrification Renewables (EF) and Electrification Plannable (EP). Each subplot represents one scenario, showing predictions from the three best-performing models (fcrn_2, fcrn_6, fcrn_8). Seasonal patterns are evident in both cases, with recurring peaks and troughs aligned with weather-year cycles. Prices generally fluctuate between 20–100 EUR/MW, though the magnitude and variability differ by model and scenario.

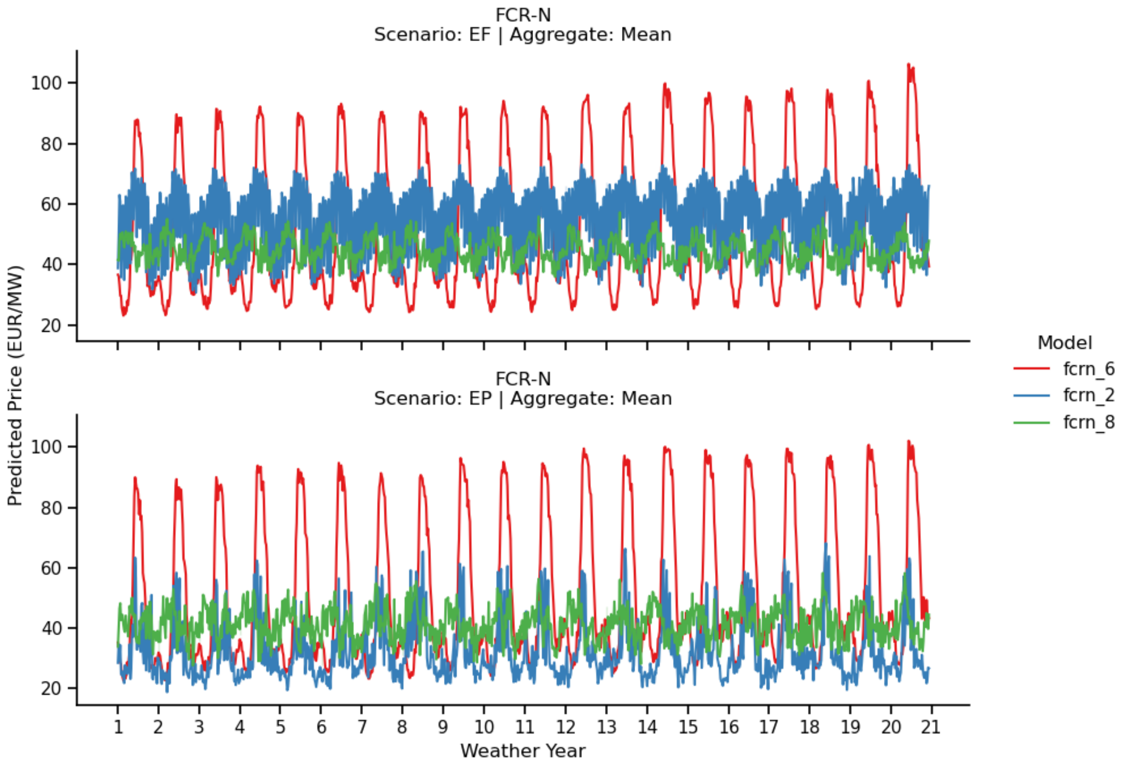


Figure 9. Forecasted weekly mean FCR-N prices for EF and EP scenarios over 21 weather years.

To provide a closer look at intra-year variations, Figure 10. zooms in on Weather Year 11, highlighting short-term variations and uncertainty bands, generated from the min and max aggregates, for each model. While long-term trends remain consistent, short-term volatility is more pronounced in EF, where price spikes might occur during high renewable penetration periods. This suggests that system variability under EF could lead to higher balancing costs compared to EP. Model fcrn_8 response to higher penetration from renewables during the summer months and can be ruled out as a reliable forecast.

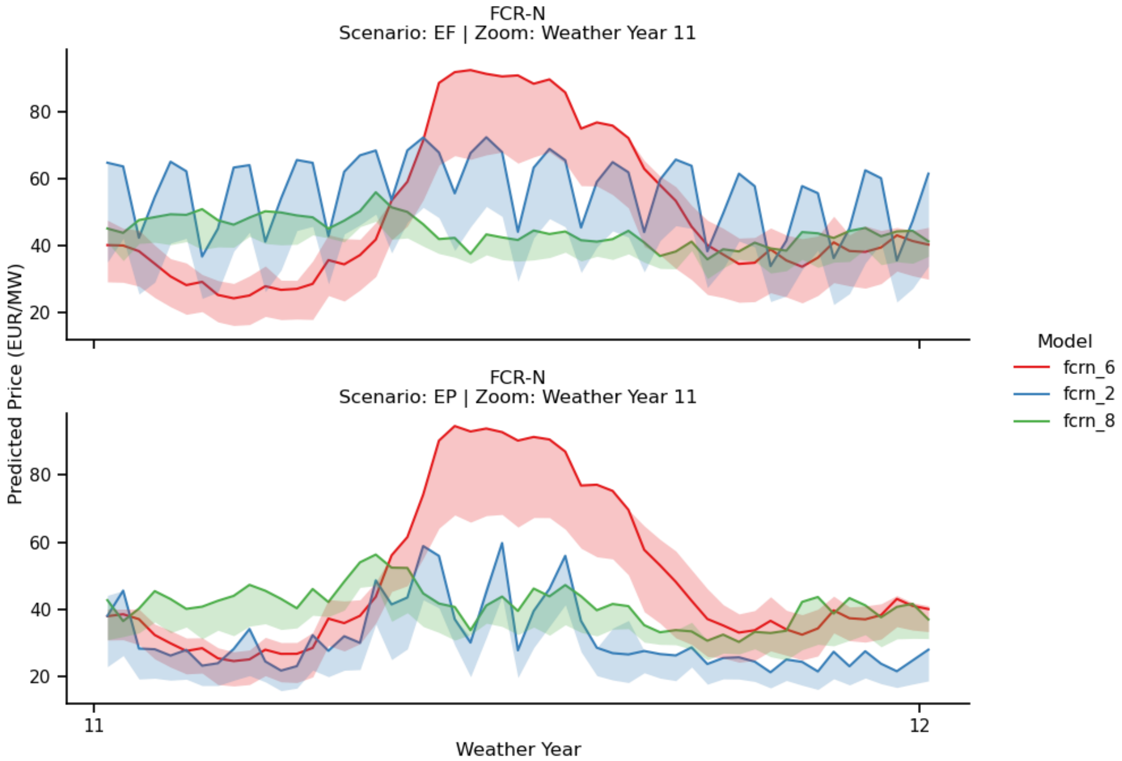


Figure 10. Zoomed view of Weather Year 11 with uncertainty bands (min–max) for each model.

Figure 11. summarizes the distribution of predicted prices across all weather years using violin plots, split by scenario. The EF scenario exhibits slightly higher median prices and broader tails, reinforcing the observation of increased variability and potential for extreme values. Conversely, EP shows a more concentrated distribution, indicating relatively stable reserve costs under a plannable electrification pathway.

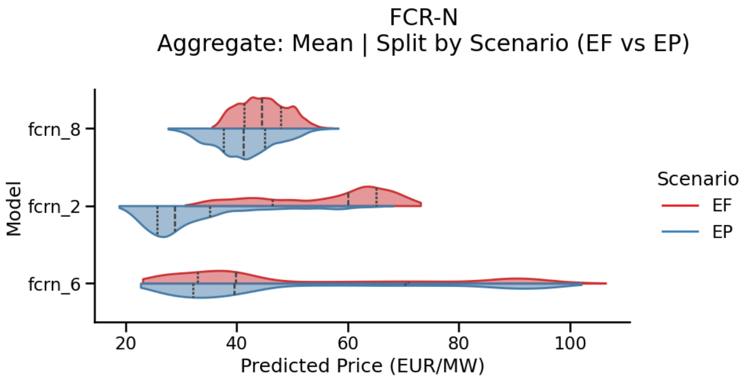


Figure 11. Violin plot showing price distributions and quartiles for EF vs EP scenarios. Dotted lines mark the 1st and 3rd quartiles, while the dashed lines mark the 2nd quartile (the median price).

Overall, both scenarios predict comparable average price ranges, but EF tends toward higher mean prices and greater variability and could reflect operational challenges of integrating renewables.

5.2 FCR-D up

The 17 models trained to target FCR-D up weekly prices, plotted together with actual values from the period 2024-03-31 to 2025-03-31 are shown in Figure 12 below. The figure contains the results of each aggregate for both predictions and actuals. The over all results show that the main price trend is not captured by a majority of predictions, where most models overshoot during the summer and with few responding to the price spikes in the autumn and following winter. The price forms a sinusoidal with a peak in June and a bottom in December.

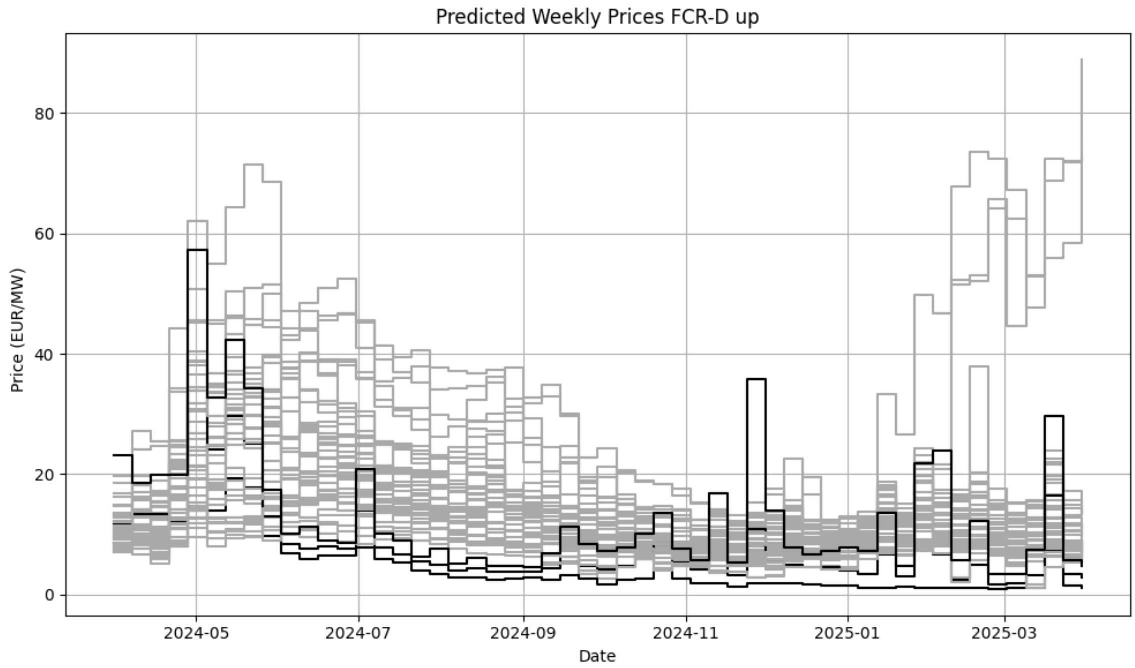


Figure 12. Predicted weekly prices for FCR-D up for all 17 models (grey) and actual (black) over the validation period.

5.2.1 Evaluation

To limit the results, three top-ranked models based on final composite scores were selected. This approach ensures focus on models with balanced and consistent performance across all aggregates. The best-performing models are fcrd_u_11, fcrd_u_14, and fcrd_u_3.

Figure 13 presents composite and final scores for a subset of trained models evaluated with actual prices of FCR-N. Each evaluation includes:

- Composite scores of aggregates' (min, mean, and max) price predictions (blue, orange, green)
- A final score (grey), calculated as the product of the three

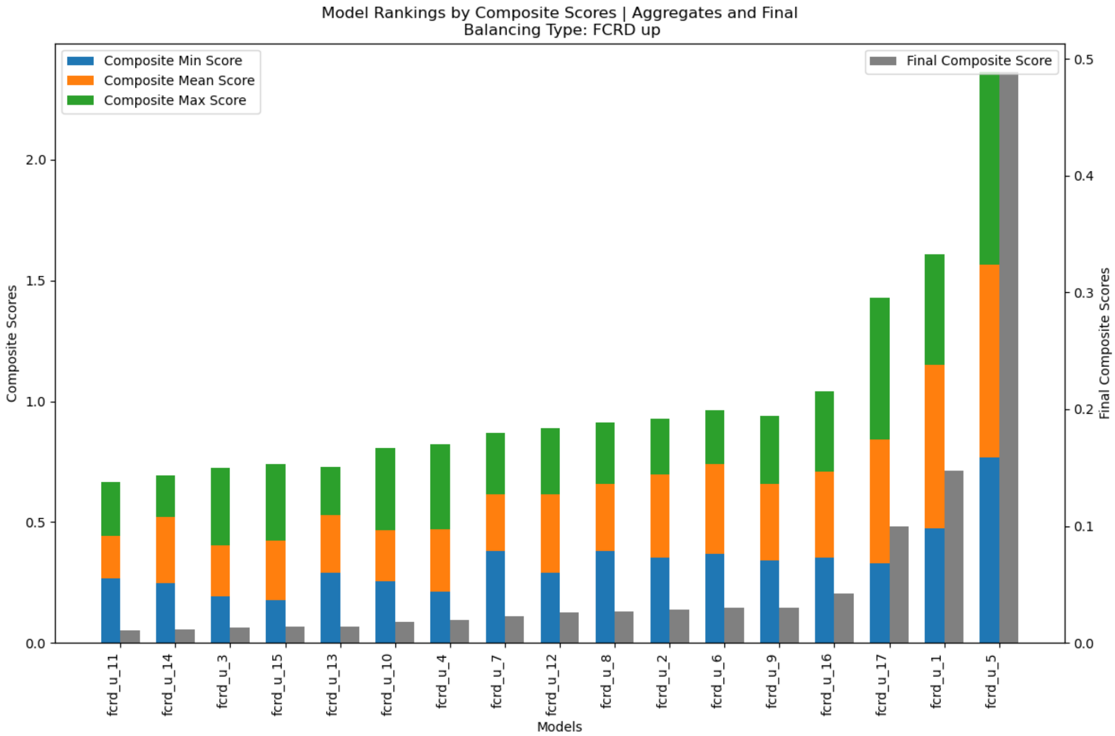


Figure 13. Composite scores per aggregate, with a final score in grey (product of each aggregate's score).

Figure 14 shows predicted weekly mean prices of FCR-D up from 2024-03-31 to 2025-03-31, filtered to the three best-performing models (fcrd_u_11, fcrd_u_14, fcrd_u_3) based on final composite scores.

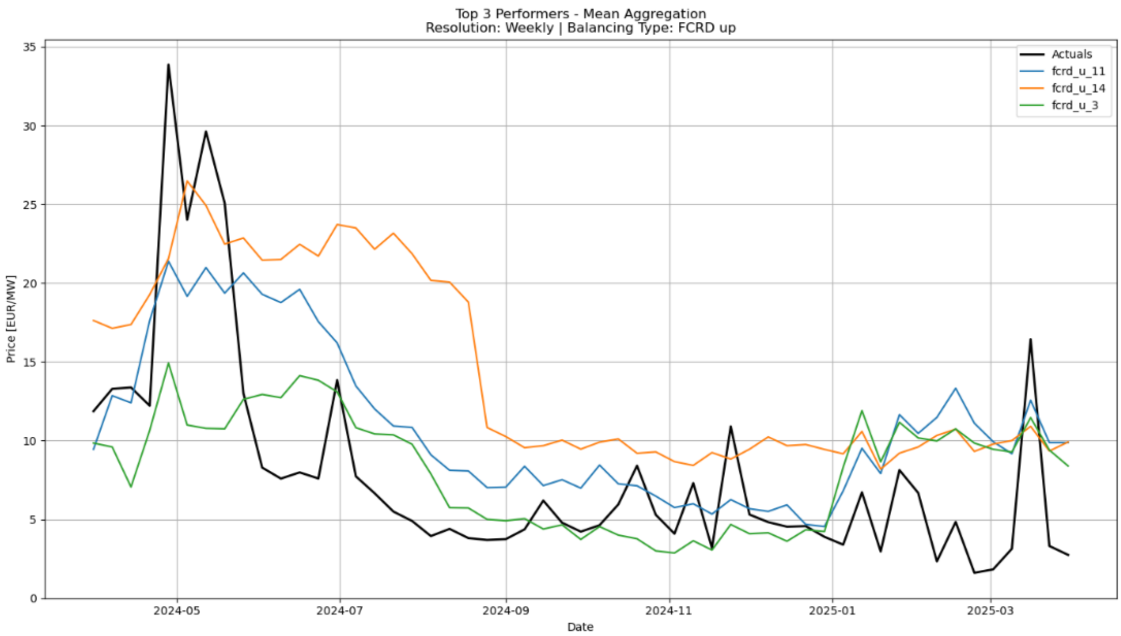


Figure 14. Top three performing models' prediction of weekly mean prices over the validation period.

All models capture the initial price spike but struggles a bit to follow the immediate surge. All models miss reacting to the spikes in November and overshoot after the latter but seems to reflect the fluctuations occurring late winter. Overall, model fcrd_u_11 handles the large weekly price variations confidently, together with capturing the long-term trend of higher prices during summer months and lower prices winter months. Model fcrd_u_3 is better at following the long-term trend.

The results indicate a limited, but not excluded, responsiveness to short-term volatility, likely due to weaker or delayed correlation between the input features and sudden price movements. The result also suggests an external or unmodeled factor or market behaviour that may be driving price spikes, which is not captured in the training data. This can be observed, as price increments but not with a corresponding magnitude.

5.2.2 Forecasts

Figure 15 shows forecasted weekly mean prices for FCR-D up across 21 weather years under both scenarios: Electrification Renewables (EF) and Electrification Plannable (EP). Each subplot represents one scenario and includes predictions from the three best-performing models (fcrd_u_3, fcrd_u_11, fcrd_u_14). All models exhibit clear seasonal patterns, with recurring peaks and troughs that align with renewable generation impacts. However, a notable divergence appears over time: prices for model fcrd_u_14 increase in magnitude in later weather years, suggesting a feedback effect caused by iterating predictions beyond the model's forecast horizon. This behaviour is not observed for the other two models, which remain relatively stable.

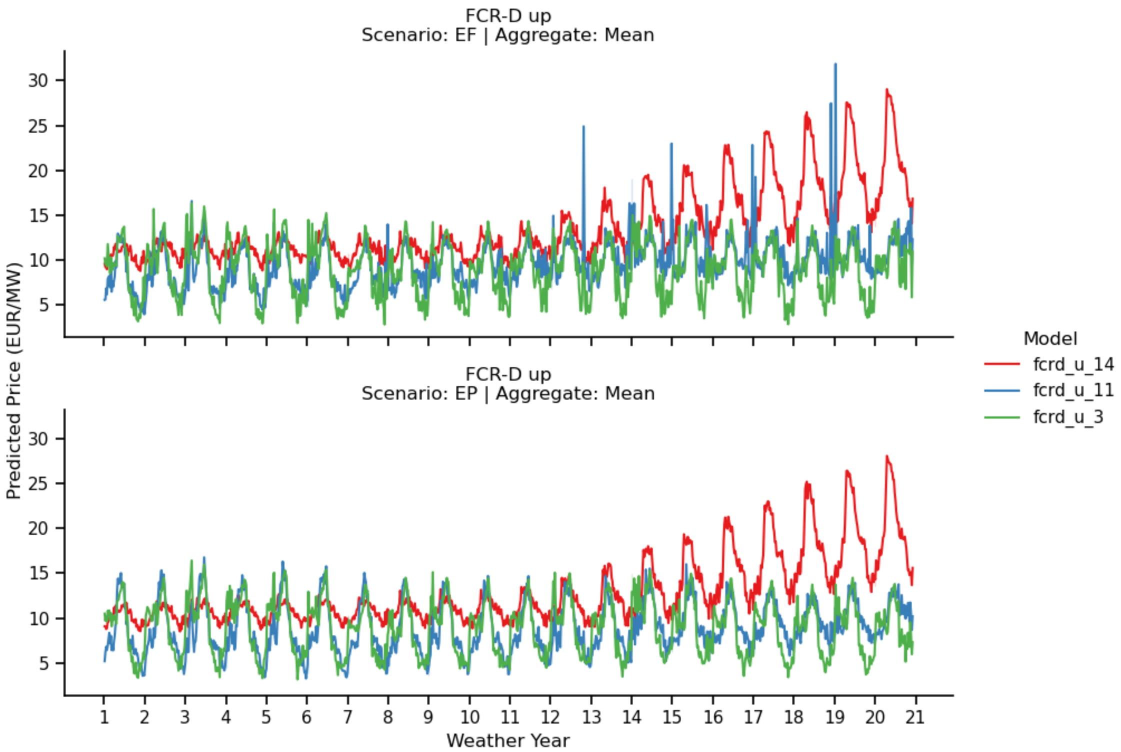


Figure 15. Forecasted weekly mean FCR-D up prices for EF and EP scenarios over 21 weather years.

To examine short-term variations, Figure 16 zooms in on Weather Year 11, displaying weekly variations and uncertainty bands (min–max ranges) for each model. Seasonal fluctuations remain evident, but the EF scenario demonstrates slightly higher volatility compared to EP. The uncertainty bands confirm that most price variation occurs within a narrow range, except for occasional spikes in EF. All three models follow an annual trend with higher demand during summer months as a result from increased penetration from renewables.

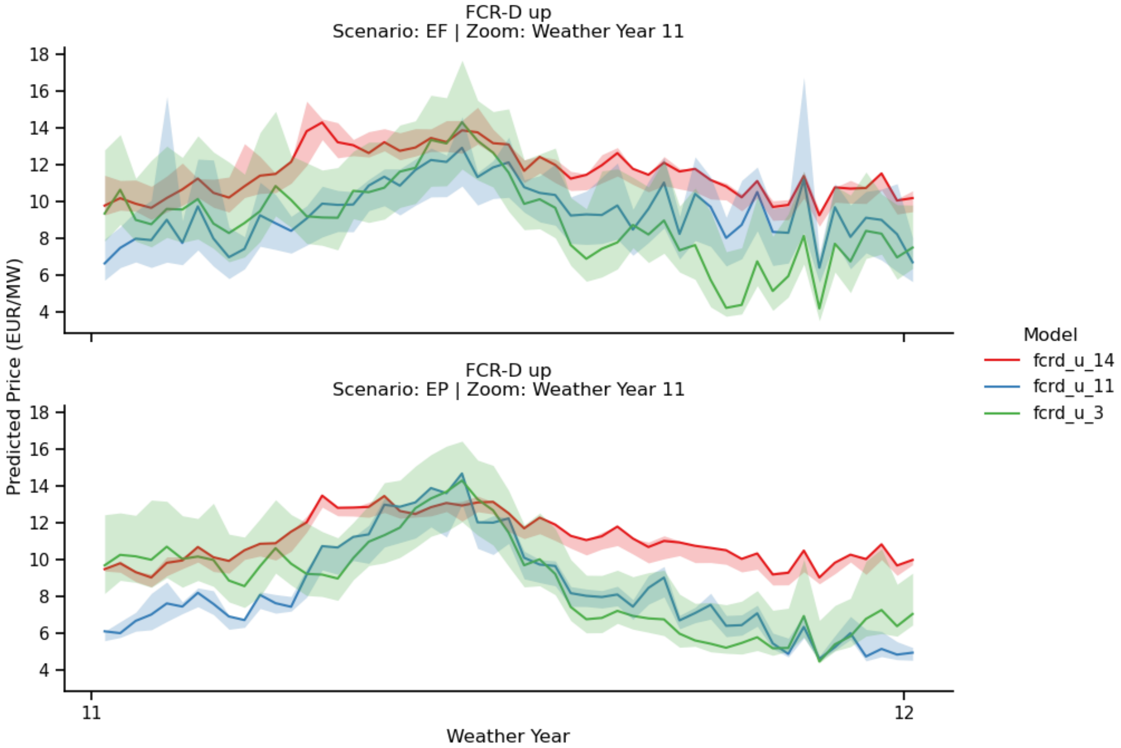


Figure 16. Zoomed view of Weather Year 11 with uncertainty bands (min–max) for each model.

Figure 17. Violin plot showing price distributions and quartiles for EF vs EP scenarios summarizes the distribution of predicted prices across all weather years using violin plots, split by scenario. The EF scenario exhibits broader tails and higher extremes, driven primarily by the feedback effect in model fcrd_u_14. In contrast, EP shows a more concentrated distribution, indicating relatively stable reserve costs under a plannable electrification pathway.

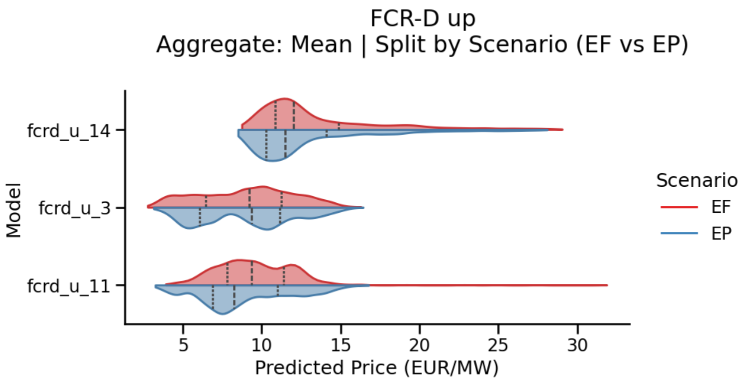


Figure 17. Violin plot showing price distributions and quartiles for EF vs EP scenarios

Overall, prices range between 3 EUR/MW and 30 EUR/MW (due to divergence), with only a few spikes above 15 EUR/MW occurring in the EF scenario during four weather years. Scenario EP generally shows slightly lower prices, except for model fcrd_u_3, where the median price is marginally higher than in EF.

In summary, while both scenarios predict similar average price levels for FCR-D up, EF introduces greater variability and occasional extreme values, highlighting the importance of addressing feedback effects in iterative forecasting and ensuring robust flexibility measures to manage volatility.

5.3 FCR-D down

The 17 models trained to target FCR-D down weekly prices, plotted together with actual values from the period 2024-03-31 to 2025-03-31 are shown in Figure 18 below. The figure contains the results of each aggregate for both predictions and actuals. The overall results show that the main price trend is not captured by many predictions, where most models overshoot during the summer and with few responding to the price spikes in the autumn and following winter. The price forms a sinusoidal with a peak in June and a bottom in December.

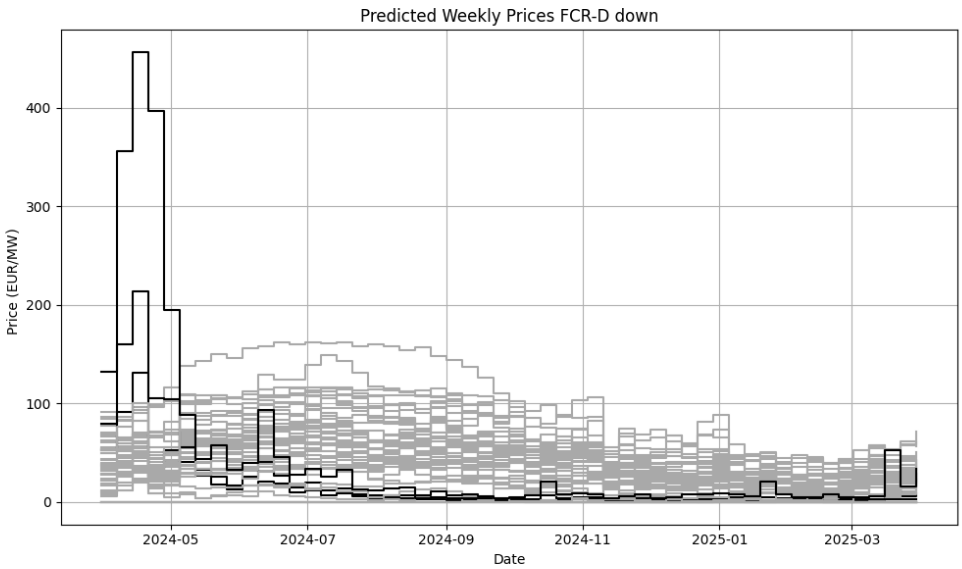


Figure 18. Predicted weekly prices for FCR-D down for all 17 models (grey) and actual (black) over the validation period.

5.3.1 Evaluation

To limit the results, three top-ranked models based on final composite scores were selected. This approach ensures focus on models with balanced and consistent performance across all aggregates. The best-performing models for FCR-D down are fcrd_d_1, fcrd_d_6, and fcrd_d_7.

Figure 19 presents composite and final scores for a subset of trained models evaluated against actual prices of FCR-D down. Each evaluation includes:

- Composite scores for each aggregate (min, mean, and max) price predictions (blue, orange, green)
- A final score (grey), calculated as the product of the three

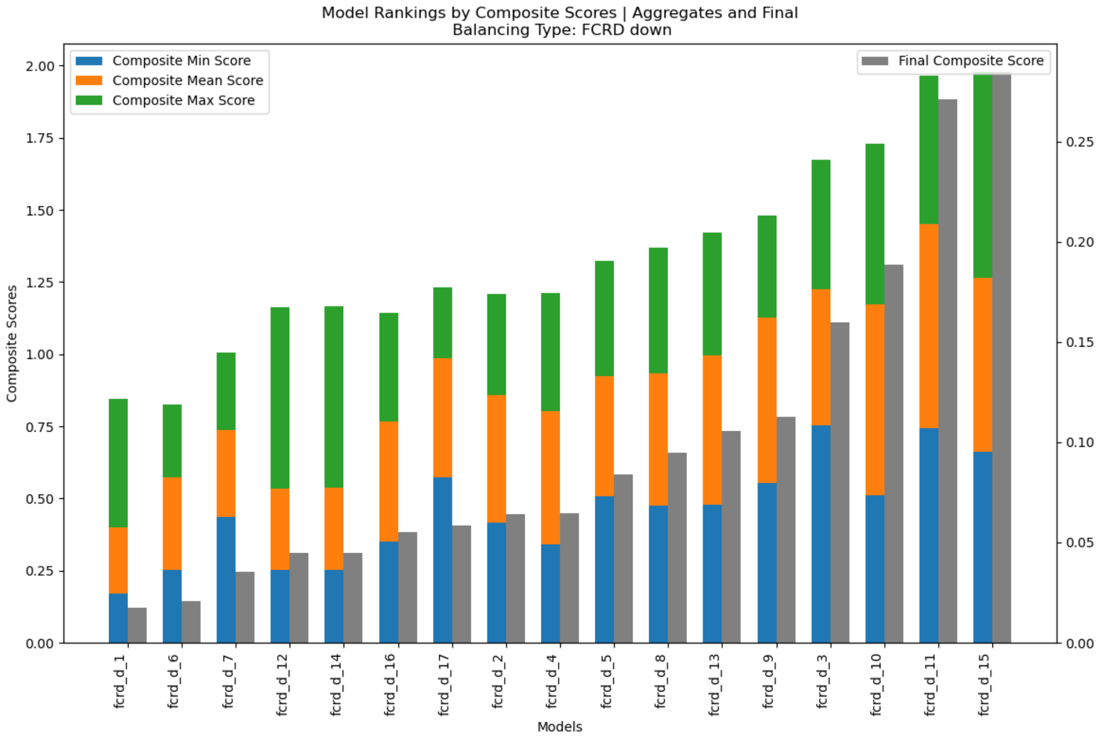


Figure 19. Composite scores per aggregate, with a final score in grey (product of each aggregate's score).

Figure 20 shows predicted weekly mean prices of FCR-D down from 2024-03-31 to 2025-03-31, filtered to the three best-performing models based on final composite scores. All models capture the overall downward trend in prices but exhibit notable limitations in responsiveness to short-term volatility. None of the models accurately predict the sharp price spike observed in April 2024; instead, they either delay the response or overestimate prices during the summer months. Later in the period, predictions flatten and converge toward actual prices during autumn and winter.

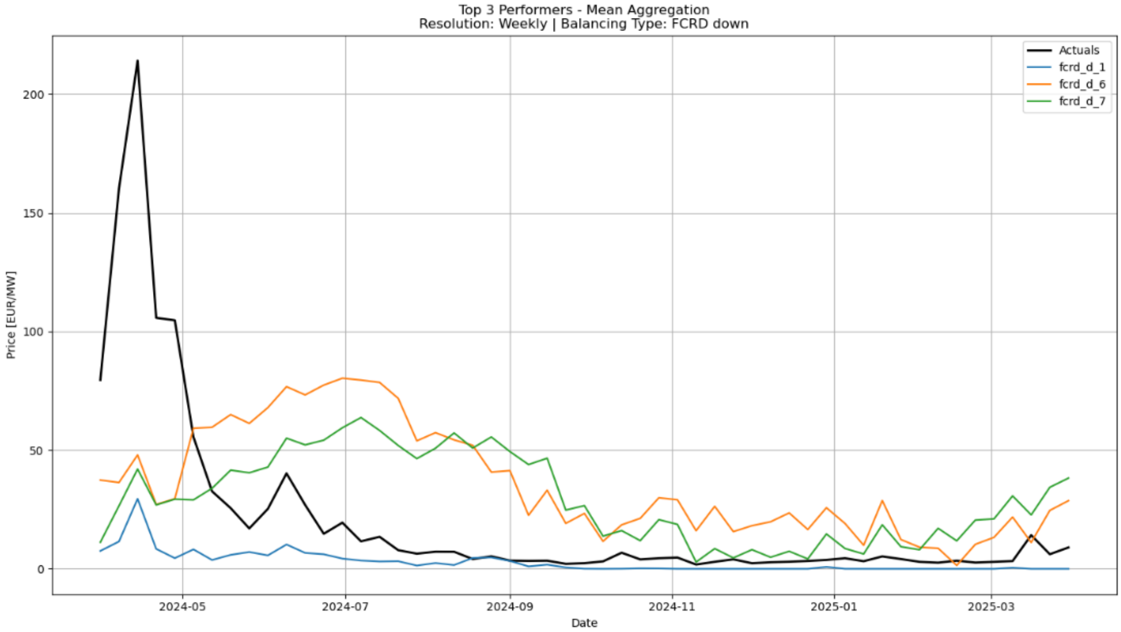


Figure 20. Top three performing models' prediction of weekly mean prices over the validation period.

These results highlight two key challenges:

- Delayed adaptation to abrupt changes. The inability to capture the April spike suggests that the models lack predictive signals for sudden market stress events.
- Bias toward seasonal smoothing. Predictions tend to overestimate prices during summer and underestimate volatility, reflecting the model's reliance on historical seasonal patterns rather than rare anomalies.

The training dataset was limited to three years, which constrains the model's ability to learn from diverse market conditions. Recent price behaviour, including sharp spikes and rapid declines, represents new variations that were not present in earlier data. This explains why the models struggle with extreme events despite performing well on long-term trends. Future improvements could include expanding the historical dataset and incorporating additional explanatory features that capture scarcity conditions or regulatory interventions.

5.3.2 Forecasts

Figure 21 presents forecasted weekly mean prices for FCR-D down across 21 weather years under both scenarios: Electrification Renewables (EF) and Electrification Plannable (EP). Each subplot shows predictions from the three best-performing models (fcrd_d_6, fcrd_d_1, fcrd_d_7). All models exhibit strong seasonal patterns, with recurring peaks and troughs that align with renewable generation cycles. Prices fluctuate significantly, ranging from near 0 EUR/MW to peaks around 80 EUR/MW. These high spikes occur intermittently and are concentrated in a few weather years, suggesting sensitivity to extreme system conditions.

Scenario EP generally shows slightly lower price volatility compared to EF, though both scenarios share similar seasonal trends. Model fcrd_d_1 stands out with extended periods of very low prices, occasionally dropping close to zero, indicating a potential bias toward overestimating downward flexibility availability.

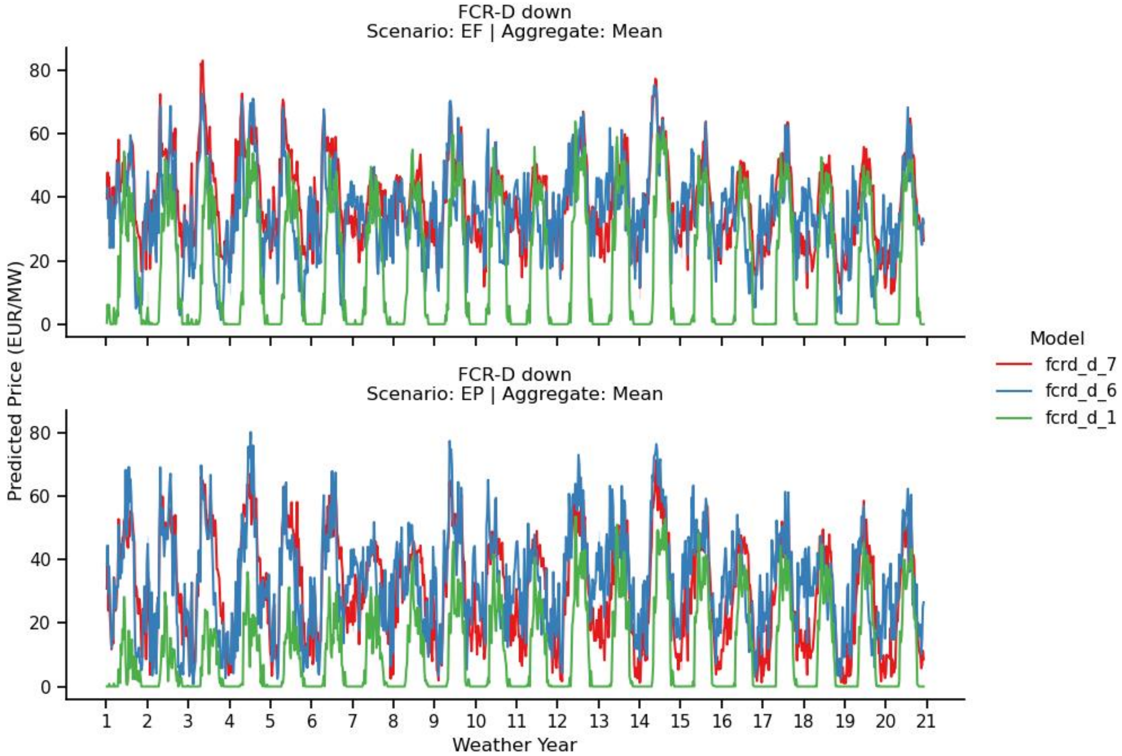


Figure 21. Forecasted weekly mean FCR-D down prices for EF and EP scenarios over 21 weather years.

Figure 22 zooms in on Weather Year 11, highlighting short-term variations and uncertainty bands (min–max ranges) for each model. While the overall seasonal pattern persists, EF demonstrates broader uncertainty bands and higher peaks than EP, reinforcing the observation of greater variability under high renewable penetration. The differences between models are pronounced: fcrd_d_6 and fcrd_d_7 maintain moderate price ranges, while fcrd_d_1 exhibits sharp drops to zero, which could distort aggregated cost estimates.

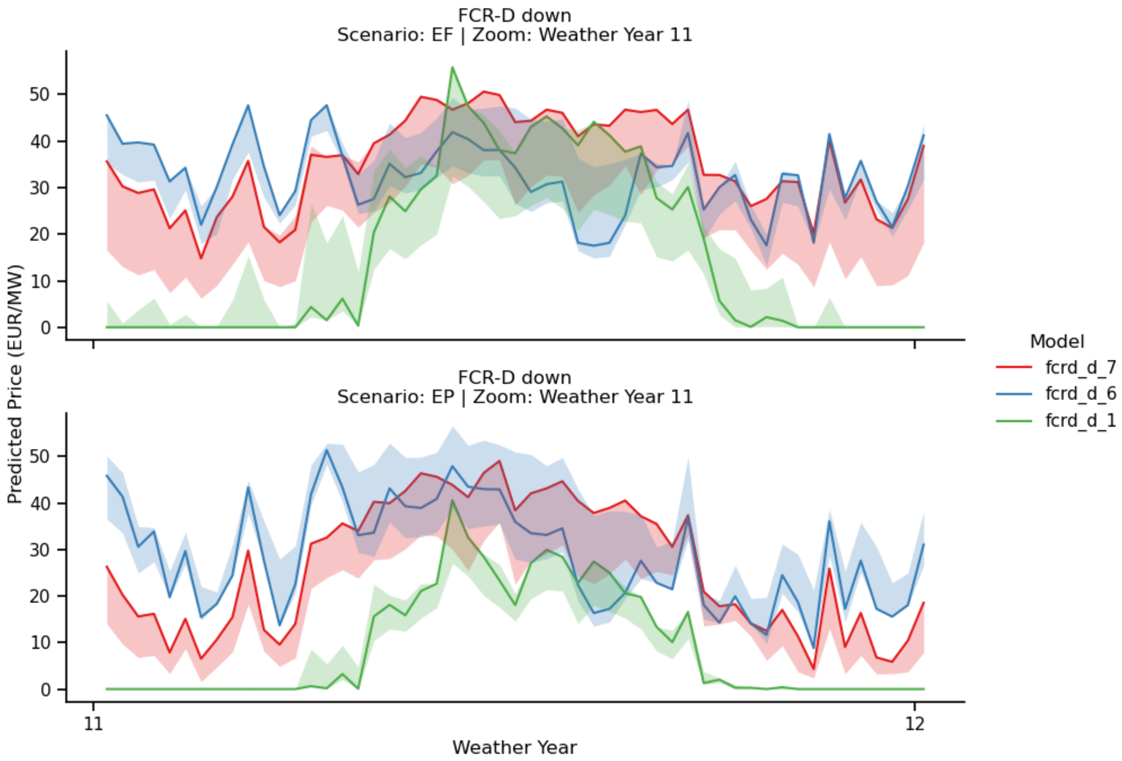


Figure 22. Zoomed view of Weather Year 11 with uncertainty bands (min–max) for each model.

Figure 23 summarizes the distribution of predicted prices across all weather years using violin plots, split by scenario. EF shows wider tails and higher extremes, while EP distributions are more concentrated. The violin plot for fcrd_d_1 confirms its tendency toward very low prices, contrasting with the broader distributions of the other two models, but resembling historical prices from the market according to the annual trends presented in Limitations.

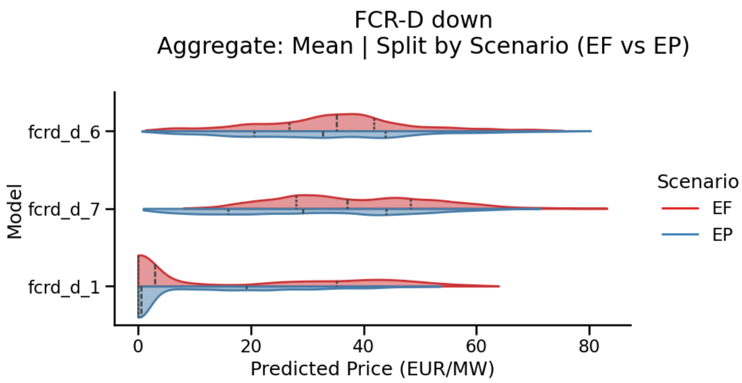


Figure 23. Violin plot showing price distributions and quartiles for EF vs EP scenarios.

In summary, FCR-D down forecasts reveal substantial variability across both EF and EP scenarios, with models predicting extended periods of prices close to 0 EUR/MW.

6 Discussion

This discussion synthesises the main findings from the modelling approach, the characteristics of the available input data, and the behaviour of the forecasting results. The objective is to reflect on the strengths and limitations of the study, explain how methodological choices shaped the outcomes, and identify opportunities for future improvements. The results should be interpreted in light of the constraints imposed by both the historical data and the structure of the long-term scenarios from Svenska kraftnät.

6.1 Input data and Methodological Considerations

The quality and character of the input data influenced the performance of the model more strongly than any other factor. The historical datasets exhibit clear seasonal patterns, which the TFT model managed to capture reliably. Prices tended to decline during winter months and rise during summer, a behaviour that appeared consistently in both the decomposed time series analysis and the model outputs. At the same time, historical data contain short-lived anomalies caused by operational disturbances or market fluctuations. These events are not cyclical and are not well represented by the available explanatory features, which explains why the model consistently underestimates the magnitude of sudden price spikes.

Another important constraint is the limited maturity of the FCR markets. Only a few years of consistent and representative market data are available, and several of these years display different structural characteristics. Regulatory adjustments, evolving procurement strategies, and changes in competition have shaped market behaviour in ways that reduce the availability of stable long-term patterns. This makes it difficult for the model to learn robust relationships between prices and the explanatory variables, particularly for events that occur rarely or under specific system conditions.

A central methodological choice was to treat the long-term spot price trajectories from SvK's scenario framework as fixed known covariates rather than modelling them jointly with FCR prices. This approach ensures full consistency with the official planning scenarios and makes the results directly interpretable in the context of the 2045 EP and EF storylines. The drawback is that the forecasts become dependent on the accuracy of the electricity price prognoses. Any structural deviation between the anticipated and actual spot price development will be reflected in the model output. Furthermore, by using fixed price trajectories, the model does not capture any possible feedback effects between balancing services and energy markets.

6.2 Behaviour of Forecasts

Despite the limitations described above, the model succeeds in reproducing broad long-term patterns. Weekly averages smooth out hourly fluctuations, which makes the forecasts more stable and suitable for long-horizon analysis. The weekly representation inevitably reduces insight into short-term operational behaviour, yet for planning purposes this abstraction is acceptable. Both the EP and EF scenarios produce coherent seasonal forecasts that reflect underlying weather-year variations and general system development. The EF scenario consistently yields higher average prices across all models, which is consistent with expectations for an energy system characterised by higher shares of non-dispatchable generation and greater variability in net load.

The forecasts reveal that the TFT architecture is suited for learning medium- and long-range dependencies but less capable of capturing extreme events or abrupt price movements. Since these spikes are often caused by factors that are not part of the scenario dataset, such as unexpected outages or abrupt cross-border changes, the model does not have sufficient information to reproduce them. This suggests that the predictive signals required for accurate short-term extremity are absent rather than poorly learned.

6.3 Comparison with Historical Prices

An important observation from the results is that the forecasted price ranges for all three balancing services (FCR-N, FCR-D up, and FCR-D down) are significantly lower than historical market prices, if compared to the annual trends of the last three years. This discrepancy is evident in the violin plots and continuous forecasts, where predicted prices rarely exceed the upper bounds observed in recent years. For example:

- FCR-N exhibited even higher historical volatility, with maximum prices of 160–120 EUR/MW in 2023–2024, dropping rapidly to 30–20 EUR/MW by the start of year 2025.
- FCR-D up historically reached around 70 EUR/MW at its peak in recent years and has a lower price around 5 EUR/MW.
- FCR-D down showed highs of 200–180 EUR/MW in 2023–2024, falling to approximately 5 EUR/MW in 2025.

In contrast, the forecasts for 2045 under both EF and EP scenarios suggest price ranges that are substantially lower, typically between 3 EUR/MW and 30 EUR/MW, with only occasional spikes in EF for FCR-D up and FCR-D down. These results indicate that the long-term scenario assumptions from SvK, which underpin the electricity price trajectories used as covariates, strongly influence the predicted balancing service prices. If the underlying energy price projections imply a stable or declining cost environment, the model will propagate these trends into reserve markets.

This raises two important considerations for future work:

1 **Scenario Dependency**: Forecast accuracy depends on the validity of SvK’s long-term electricity price assumptions. Any structural deviation in energy prices will directly affect reserve price forecasts.

2 **Historical Comparison and Calibration**: Given that historical prices are roughly twice as high as forecasted values, further investigation is needed to determine whether this gap reflects genuine expectations for future system conditions or limitations in the modelling approach. Comparing annual trends in SvK’s market analysis with historical balancing prices could help identify whether the observed decline is plausible or whether additional explanatory features should be incorporated.

Overall, while the forecasts provide insights into long-term trends, they should be interpreted cautiously. The relatively low predicted prices may underestimate future reserve costs if volatility drivers are not fully captured in the scenario framework.

6.4 Limitations and Directions for Future Work

The limitations identified in this study suggest several priorities for future research. First, the explanatory power of the model could likely be enhanced by broadening the set of input variables. Factors such as cross-border flows, production outages, or short-term imbalance indicators may help the model capture both volatility and more abrupt behavioural changes. Quarterly prices, introduced in fall of year 2025 impacts spot price markets whit new market behaviour, which should be considered in future modeling. Care must be taken when selecting additional features to ensure that the reproducibility of the scenario-driven methodology is preserved. An alternative path would be to develop a hybrid modelling approach in which spot and reserve prices are estimated jointly, providing a more integrated representation of market dynamics.

Second, the coarseness introduced by daily resampling improves long-term stability but removes fluctuations that may contain useful information. Investigating intermediate temporal resolutions, for example three- or six-hour aggregates, could improve the model's ability to capture dynamic behaviour while still supporting long-horizon forecasting.

Third, further work comparing the TFT model to other architectures would be beneficial. Although this was outside the scope of the present project, benchmarking against more recent transformer variants such as TimeXer, PatchTST or iTransformer, as well as classical lightweight models, would strengthen confidence in the selected approach and clarify where the benefits of deep learning are most pronounced.

Taken together, the results demonstrate that the chosen modelling framework can generate plausible long-term forecasts under different future scenarios, while also highlighting the areas where additional methodological development would improve the robustness and interpretability of the outcomes. The model is best suited for capturing broad structural trends rather than short-term variations, and its usefulness is strongest when applied in conjunction with scenario-based planning where long-term trajectories take precedence over high-frequency accuracy.

Appendix

Following Table IX contains a full summary of model configurations used for each training session. The table includes Hyper parameters imported from Optima studies or presets recommended by model developer. The summary also contain training periods and selected number of Epoch and Step used for training.

Table IX. Summary of model setup, including parameter and hyper parameters settings.

target_variable	model_name	prediction_horizon	training_period	epoch	step	val_loss	hparams_hidden_size	hparams_dropout	hparams_hidden_continu	hparams_attention_head	hparams_learning_rate
price_fcrd_d	fcrd_d_1	T-28	2019-03-31-2024-03-30	23	120	26.96	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_2	T-28	2019-03-31-2024-03-30	19	100	27.22	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_3	T-28	2019-03-31-2024-03-30	28	145	25.32	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_4	T-28	2019-03-31-2024-03-30	27	140	26.87	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_5	T-28	2019-03-31-2024-03-30	18	95	27.83	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_6	T-28	2019-03-31-2024-03-30	21	110	26.81	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_7	T-28	2019-03-31-2024-03-30	32	165	26.07	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_8	T-28	2019-03-31-2024-03-30	29	150	27.44	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_9	T-28	2019-03-31-2024-03-30	28	145	26.58	64	0.1	34	2	0.000231

target_variable	model_name	prediction_horizon	training_period	epoch	step	val_loss	hparams_hidden_size	hparams_dropout	hparams_hidden_continuous	hparams_attention_head	hparams_learning_rate
price_fcrd_d	fcrd_d_10	T-28	2019-03-31-2024-03-30	12	65	28.17	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_11	T-28	2019-03-31-2024-03-30	3	20	27.83	16	0.1	8	2	0.01
price_fcrd_d	fcrd_d_12	T-28	2019-03-31-2024-03-30	3	20	29.91	16	0.1	8	2	0.01
price_fcrd_d	fcrd_d_13	T-28	2019-03-31-2024-03-30	9	50	27.33	16	0.1	8	2	0.01
price_fcrd_d	fcrd_d_14	T-28	2019-03-31-2024-03-30	6	35	27.07	16	0.1	8	2	0.01
price_fcrd_d	fcrd_d_15	T-28	2019-03-31-2024-03-30	30	155	27.36	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_16	T-28	2019-03-31-2024-03-30	18	95	27.63	64	0.1	34	2	0.000231
price_fcrd_d	fcrd_d_17	T-28	2019-03-31-2024-03-30	11	60	26.96	64	0.1	34	2	0.000231
price_fcrd_u	fcrd_u_1	T-28	2018-03-31-2024-03-30	10	55	2.77	9	0.1	9	4	0.02352
price_fcrd_u	fcrd_u_2	T-28	2018-03-31-2024-03-30	8	45	2.42	9	0.1	9	4	0.02352
price_fcrd_u	fcrd_u_3	T-28	2018-03-31-2024-03-30	12	65	2.08	9	0.1	9	4	0.02352

target_variable	model_name	prediction_horizon	training_period	epoch	step	val_loss	hparams_hidden_size	hparams_dropout	hparams_hidden_continu	hparams_attention_head	hparams_learning_rate
price_fcrd_u	fcrd_u_4	T-28	2018-03-31-2024-03-30	13	70	2.05	9	0.1	9	4	0.02352
price_fcrd_u	fcrd_u_5	T-28	2018-03-31-2024-03-30	21	110	2.03	9	0.1	9	4	0.02352
price_fcrd_u	fcrd_u_6	T-28	2018-03-31-2024-03-30	13	70	2.06	9	0.1	9	4	0.02352
price_fcrd_u	fcrd_u_7	T-28	2018-03-31-2024-03-30	9	50	2.29	9	0.1	9	4	0.02352
price_fcrd_u	fcrd_u_8	T-28	2018-03-31-2024-03-30	12	65	2.3	9	0.1	9	4	0.02352
price_fcrd_u	fcrd_u_9	T-28	2019-03-31-2024-03-30	10	55	2.03	16	0.1	8	2	0.01
price_fcrd_u	fcrd_u_10	T-28	2019-03-31-2024-03-30	21	110	2.2	16	0.1	8	2	0.01
price_fcrd_u	fcrd_u_11	T-28	2019-03-31-2024-03-30	25	130	2.11	16	0.1	8	2	0.01
price_fcrd_u	fcrd_u_12	T-28	2019-03-31-2024-03-30	13	70	2.64	16	0.1	8	2	0.01
price_fcrd_u	fcrd_u_13	T-28	2019-03-31-2024-03-30	17	90	2.38	16	0.1	8	2	0.01
price_fcrd_u	fcrd_u_14	T-28	2019-03-31-2024-03-30	17	90	2.56	16	0.1	8	2	0.01

target_variable	model_name	prediction_horizon	training_period	epoch	step	val_loss	hparams_hidden_size	hparams_dropout	hparams_hidden_continuous	hparams_attention_head	hparams_learning_rate
price_fcrd_u	fcrd_u_15	T-28	2019-03-31- 2024-03-30	21	110	2.24	14	0.1	9	4	0.039221
price_fcrd_u	fcrd_u_16	T-28	2019-03-31- 2024-03-30	30	155	2.06	14	0.1	9	4	0.039221
price_fcrd_u	fcrd_u_17	T-28	2019-03-31- 2024-03-30	12	65	2.15	14	0.1	9	4	0.039221
price_fcrn	fcrn_1	T-28	2018-03-31- 2024-03-30	17	90	12.66	18	0.1	13	1	0.004862
price_fcrn	fcrn_2	T-28	2018-03-31- 2024-03-30	10	55	12.12	18	0.1	13	1	0.004862
price_fcrn	fcrn_3	T-28	2018-03-31- 2024-03-30	3	20	12.34	18	0.1	13	1	0.004862
price_fcrn	fcrn_4	T-28	2018-03-31- 2024-03-30	6	35	12.89	18	0.1	13	1	0.004862
price_fcrn	fcrn_5	T-28	2018-03-31- 2024-03-30	9	50	12.13	18	0.1	13	1	0.004862
price_fcrn	fcrn_6	T-28	2018-03-31- 2024-03-30	8	45	11.84	18	0.1	13	1	0.004862
price_fcrn	fcrn_7	T-28	2018-03-31- 2024-03-30	6	35	12.12	18	0.1	13	1	0.004862
price_fcrn	fcrn_8	T-28	2019-01-01- 2024-03-30	4	25	13.74	21	0.1	14	3	0.000612

target_variable	model_name	prediction_horizon	training_period	epoch	step	val_loss	hparams_hidden_size	hparams_dropout	hparams_hidden_continu	hparams_attention_head	hparams_learning_rate
price_fcrn	fcrn_9	T-28	2019-01-01-2024-03-30	18	95	12.24	21	0.1	14	3	0.000612
price_fcrn	fcrn_10	T-28	2019-01-01-2024-03-30	11	60	12.91	21	0.1	14	3	0.000612
price_fcrn	fcrn_11	T-28	2019-01-01-2024-03-30	24	125	11.82	21	0.1	14	3	0.000612
price_fcrn	fcrn_12	T-28	2019-01-01-2024-03-30	18	95	11.62	21	0.1	14	3	0.000612
price_fcrn	fcrn_13	T-28	2019-01-01-2024-03-30	12	65	12.58	21	0.1	14	3	0.000612
price_fcrn	fcrn_14	T-28	2019-01-01-2024-03-30	14	75	11.86	21	0.1	14	3	0.000612
price_fcrn	fcrn_15	T-28	2019-01-01-2024-03-30	18	95	11.83	21	0.1	14	3	0.000612
price_fcrn	fcrn_16	T-28	2019-01-01-2024-03-30	10	55	11.77	16	0.1	8	2	0.01
price_fcrn	fcrn_17	T-28	2019-01-01-2024-03-30	4	25	12.79	16	0.1	8	2	0.01

Table X. Complete list of composite scores from evaluation.

Model	Final Composite Score	Target
fcrn_6	0.015374	price_fcrn
fcrn_2	0.019055	price_fcrn
fcrn_8	0.024811	price_fcrn
fcrn_7	0.027858	price_fcrn
fcrn_17	0.039289	price_fcrn
fcrn_10	0.039857	price_fcrn

Model	Final Composite Score	Target
fcrn_4	0.044180	price_fcrn
fcrn_14	0.054896	price_fcrn
fcrn_12	0.057276	price_fcrn
fcrn_3	0.064494	price_fcrn
fcrn_5	0.080213	price_fcrn
fcrn_9	0.083680	price_fcrn
fcrn_13	0.094423	price_fcrn
fcrn_16	0.098245	price_fcrn
fcrn_1	0.147763	price_fcrn
fcrn_11	0.191454	price_fcrn
fcrn_15	0.415997	price_fcrn

Through our international collaboration programmes with academia, industry, and the public sector, we ensure the competitiveness of the Swedish business community on an international level and contribute to a sustainable society. Our 2,800 employees support and promote all manner of innovative processes, and our roughly 100 testbeds and demonstration facilities are instrumental in developing the future-proofing of products, technologies, and services. RISE Research Institutes of Sweden is fully owned by the Swedish state.

I internationell samverkan med akademi, näringsliv och offentlig sektor bidrar vi till ett konkurrenskraftigt näringsliv och ett hållbart samhälle. RISE 2 800 medarbetare driver och stöder alla typer av innovationsprocesser. Vi erbjuder ett 100-tal test- och demonstrationsmiljöer för framtidssäkra produkter, tekniker och tjänster. RISE Research Institutes of Sweden ägs av svenska staten.

RISE Research Institutes of Sweden AB
Box 857, 501 15 BORÅS, SWEDEN
Telephone: +46 10-516 50 00
E-mail: info@ri.se, Internet: www.ri.se

Energy systems analysis
RISE Report 2026:49
ISBN:978-91-90109-78-6