DEPARTMENT OF MATHEMATICS AND STATISTICS

Congestion Modelling and Locational Spread Trading in European Power Markets

by

Patrick Richard James Montgomery
202476449

MSc Applied Statistics

2024/25

Statement of work in project

The work contained in this project is that of the authors and where material from other sources has been incorporated full acknowledgement has been given.

Signed	Patrick Montgomery
Print Name	Patrick Montgomery
Date	29th September 2025

Supervised by Alison Gray

Table of Contents

List of Figures

List of Tables

Chapter 1 – Introduction

1.1 Background & Context

1.1.1 Electricity Markets in Europe: Liberalisation, Market Coupling, Congestion.

Since the late 1990s, the European Union has undertaken a multi-phase liberalisation of its electricity markets, aiming to dismantle vertically integrated monopolies and foster competition. Key milestones include the Third Energy Package (2009), enforce the unbundling of generation and transmission activities, strengthened regulatory oversight via national regulators, and promoted cross-border trade and investment. (Fact Sheets on the European Union, 2025) More recently, the Electricity Directive 2019 built upon earlier reforms by reinforcing rules for non-discriminatory market access, greater market transparency, and legal separation of system operators from suppliers. (Official Journal of the European Union, 2019)

1.1.2 Market Coupling and Integration

Complementing liberalisation, the EU has implemented market coupling mechanisms to harmonise electricity trading across national borders. Market coupling aligns price formation with available cross-border capacities through implicit auctions, reducing inefficiencies stemming from separate capacity and energy allocations. (Pontenagel, 2025)

- Price Coupling of Regions (PCR) – Used for day-ahead markets via the EUPHEMIA algorithm, enabling simultaneous matching of energy and capacity across bidding zones.
- Flow-Based Market Coupling (FBMC) – Particularly in the Central Western Europe Core Region, integrates cross-border capacity restrictions during the market clearing process to reflect actual grid constraints.

These mechanisms have enhanced market liquidity, narrowed price differentials, and optimised interconnector utilisation. Industry coordination bodies like the European Market Coupling Company (EMCC) has played instrumental roles in implementing these methods, managing cross-border allocations and settlement procedures between exchanges and TSOs. (Entsoe, 2025)

1.1.3 Congestion in Electricity Markets

Transmission congestion arises when electricity flows cannot meet scheduled dispatch due to physical constraints on the grid (e.g., thermal limits, n-1 security rules). This results in localised price divergence and necessitates remedial actions like redispatch by TSOs. Europe has traditionally managed congestion using zonal pricing, where inter-zone constraints are accounted at market clearing, while intra-zone congestion is handled post-clearing through redispatch. (Entsoe, 2025)

Congestion not only impacts system efficiency and welfare but also creates opportunities for price arbitrage—particularly for quantitative traders exploiting inter-zonal price differentials. Germany, with its high penetration of wind/solar generation and numerous interconnectors, is especially prone to congestion and therefore a prime focus for this type of analysis. Research is increasingly leveraging explainable AI and statistical models to understand congestion drivers and redispatch patterns. One example identifies wind generation, cross-border exchanges, and hydropower as key factors shaping congestion in Germany. (Titz, Putz, & Witthaut, 2023)

1.2 Problem Statement

Despite significant liberalisation and the introduction of market coupling mechanisms, congestion remains a persistent feature of European electricity markets. Transmission bottlenecks arise from the physical limitations of interconnectors and intra-zonal networks, particularly in systems with high renewable penetration. In Germany, for example, strong growth in wind and solar capacity has led to frequent redispatching needs, as generation in the north often exceeds transmission capacity to demand centres in the south. (Entsoe, 2025)

From a policy perspective, congestion reduces overall system efficiency, raises costs for transmission system operators (TSOs), and can undermine the effectiveness of market coupling. From a trading perspective, congestion creates price divergence across bidding zones, generating both risk and potential arbitrage opportunities. While electricity price forecasting has received extensive academic attention (Weron, 2014), there is comparatively less research focused specifically on forecasting congestion events and assessing their implications for trading strategies.

Furthermore, statistical modelling of congestion is challenging because congestion outcomes are jointly determined by multiple interacting processes — including load patterns, renewable generation forecasts, cross-border flows, and interconnector availability. These high-dimensional and interdependent drivers require rigorous statistical and machine learning methods for effective analysis.

This research project addresses this gap by combining a robust data infrastructure with advanced statistical modelling. It seeks to understand whether congestion in the German market, and on its interconnectors, can be forecast with sufficient accuracy to inform quantitative trading strategies.

1.3 Aims and Objectives

1.3.1 Aims

The primary aim of this dissertation is to develop and evaluate statistical models for forecasting congestion events in European electricity markets, with a focus on Germany, and to assess the profitability of trading strategies based on such forecasts.

1.3.2 Objectives

1. Data Infrastructure – Construct a scalable data pipeline using the ENTSO-E Transparency Platform and TimescaleDB to collect and store historical data on prices, load, generation, cross-border flows, and congestion indicators.
2. Exploratory Analysis – Conduct descriptive and exploratory analysis of German congestion patterns and their drivers, including renewable generation variability and interconnector utilisation.
3. Model Development – Implement a range of statistical and machine learning models (e.g., econometric time-series models, classification models, and ensemble methods) to forecast congestion events.
4. Model Evaluation – Assess model performance using appropriate statistical measures (e.g., ROC/AUC, forecast accuracy, out-of-sample testing).

5. Trading Strategy Design – Translate model forecasts into trading strategies, back-testing their profitability and robustness under realistic market assumptions.
6. Comparison with Literature – Benchmark findings against existing academic and industry research on congestion and price forecasting.
7. Critical Reflection – Identify strengths and limitations of the modelling approach, and outline opportunities for future work.

1.4 Contributions and Significance

This project makes three distinct contributions:

- Academic Contributions - It extends the statistical literature on electricity markets by focusing explicitly on congestion forecasting, an area that has received less attention than price forecasting. By applying a combination of classical econometric methods and modern machine learning approaches, the dissertation demonstrates the relative strengths and limitations of these methods in predicting congestion events.
- Practical Contribution - The research has direct relevance for market participants such as utilities, traders, and system operators. Congestion drives both risks and opportunities in cross-border trading; being able to anticipate congestion can inform spread trades, hedging strategies, and operational planning. The results also provide insights into the drivers of congestion in Germany and its neighbouring markets, which may be valuable for policymakers and TSOs seeking to improve market design and integration.
- Technical Contribution - Beyond the modelling, the dissertation develops a reproducible data infrastructure for power market research. The data pipeline integrates ENTSO-E Transparency Platform data into a scalable time-series database (TimescaleDB), with automated collection, cleaning, and storage.

Together, these contributions highlight the significance of the research both within the field of applied statistics and in its application to real-world energy markets.

1.5Structure of the Report

- Chapter 1 – Introduction. This chapter outlines the background and context of European electricity markets, states the problem addressed in this project, sets out the aims and objectives, and highlights the contributions and overall structure of the work.
- Chapter 2 – Literature Review. This chapter surveys relevant academic and industry literature on electricity price forecasting, congestion modelling, and statistical approaches to analysing market dynamics, identifying the gaps that this research aims to address.
- Chapter 3 – Data & Infrastructure. This chapter describes the ENTSO-E Transparency Platform, the types of data collected (prices, load, generation, flows, capacities), and the design of the data pipeline and database used in this study.
- Chapter 4 – Data & Exploratory Analysis. This chapter audits data quality, profiles distributions and temporal patterns, examines cross-border spreads and congestion, tests stationarity, and records missing/ outlier handling.
- Chapter 5 – Methodology. This chapter sets out the statistical and machine learning methods employed, including model formulations, assumptions, and evaluation metrics.
- Chapter 6 – Results. This chapter presents the findings from the empirical analysis, including the performance of forecasting models and the evaluation of congestion prediction accuracy.
- Chapter 7 – Discussion. This chapter interprets the results in the context of the research aims, comparing them with findings from the literature and discussing their implications for congestion analysis.
- Chapter 8 – Strengths & Limitations. This chapter evaluates the robustness of the research design, data, and methods, and acknowledges the limitations of the study.
- Chapter 9 – Conclusion and Future Work. The project report concludes by summarising the key findings, reflecting on their significance, and proposing directions for future research.

Chapter 2 – Literature Review

2.1 Electricity Price Forecasting

Electricity price forecasting (EPF) has evolved through successive methodological waves. Early literature is dominated by classical econometrics—ARIMA/ARIMAX, exponential smoothing, and regression with calendar/temperature effects—augmented by volatility models to handle the pronounced spikes and time-varying variance characteristic of electricity prices. Foundational surveys e.g. (Weron, 2014) documents how mean–variance dynamics, non-storability, and market design features distinguish EPF from other commodity forecasting tasks, motivating regime-switching and heavy-tailed error models.

A second wave emphasised non-linear and regime-aware models. Markov-switching specifications, threshold autoregressions and heterogeneous autoregressive volatility structures were used to accommodate spikes, structural breaks and intraday seasonality. As cross-border integration deepened and renewable penetration rose, exogenous regressors (ARX/VARX) such as load, wind/solar forecasts and fuel prices became standard, reflecting the increasing role of fundamentals in price formation. Comprehensive reviews culminating in (Lago, Marcjasz, De Schutter, & Weron, 2021) compare these families systematically across multiple markets and years, highlighting that well-tuned statistical baselines remain competitive for day-ahead horizons when appropriately enriched with exogenous inputs.

In parallel, machine learning (ML) and deep learning approaches—support vector regression, random forests, gradient-boosted trees, convolutional/recurrent neural networks—gained prominence. These methods can capture complex, non-linear relationships between prices and a high-dimensional set of predictors (e.g., weather ensembles, interconnector variables), with modern work reporting competitive or superior accuracy to classical benchmarks in some settings. Review papers and benchmarks (Lago, Marcjasz, De Schutter, & Weron, 2021) caution, however, that performance gains depend on careful feature engineering, robust cross-validation (market-aware time splits), and attention to non-stationarity; they also stress the importance of transparent baselines and open datasets for fair comparison. Recent surveys continue to catalogue both classical and ML/DL techniques, noting the growing use of probabilistic and quantile forecasts for risk-aware decision-making.

A notable methodological trend is the turn towards probabilistic forecasting and rigorous evaluation practice. Beyond point metrics (MAE/RMSE), studies increasingly report interval/quantile scores and calibration diagnostics, reflecting users' needs for full predictive distributions. Comparative studies emphasise rolling-origin evaluation, market-specific cross-validation and the perils of information leakage. Benchmarks in Lago et al. (2021) and subsequent replications reinforce that consistent gains often arise from improved data curation (e.g., accurate calendar and auction-timing alignment), judicious inclusion of exogenous signals (load/RES forecasts), and model ensembling rather than from any single architecture. (Lago, Marcjasz, De Schutter, & Weron, 2021)

For European day-ahead markets, institutional design affects forecast ability. The PCR/EUPHEMIA algorithm clears coupled day-ahead auctions by jointly maximising social welfare subject to cross-border capacity constraints. Understanding this mechanism, and the evolution to Flow-Based Market Coupling (FBMC) in the Core region, helps explain observed cross-zonal price dynamics: when constraints bind, zonal prices diverge; when not, prices converge. Public descriptions of EUPHEMIA and Core FBMC (by exchanges/JAO) codify how network constraints enter the clearing problem, providing context for feature selection (e.g., net positions, PTDF-related indicators) in forecasting models. (Nord Pool, 2025)

Although EPF is a mature field, congestion-aware forecasting—explicitly linking prices to grid constraints and remedial actions—remains relatively under-represented compared with price-only modelling. Recent work has begun to bridge this gap. For Germany, Titz, Pütz and Witthaut (2023/2024) use explainable ML to predict redispatch volumes, identifying wind generation, cross-border exchanges and hydropower as salient drivers; such findings motivate incorporating congestion proxies alongside standard fundamentals in European EPF. (Titz, Putz, & Witthaut, 2023)

Public disclosures are sparse due to commercial sensitivity, but available signals indicate widespread adoption of algorithmic and ML-driven forecasting across European market participants. Reporting on specialist trading houses and utilities highlights the use of advanced data-driven methods to cope with increased volatility from renewables and cross-border integration; while models are proprietary, accounts consistently describe heavy reliance on high-frequency fundamentals and automated systems to anticipate auction outcomes. (Financial Times, 2025)

The literature suggests: (i) strong baselines are attainable with well-specified statistical models (ARX/VARX) enriched by fundamentals; (ii) ML/DL can add value where non-linearity and interactions are pronounced; (iii) in a coupled European setting, integrating congestion-related features (e.g., flow/capacity indicators, net positions) is justified both institutionally (PCR/FBMC) and empirically (redispatch studies). These insights inform the subsequent modelling choices and evaluation design in this project.

2.2 Congestion and Redispatch

2.2.1 Definitions and institutional context

In European power systems, transmission congestion occurs when flows that would be optimal under unconstrained economic dispatch are infeasible due to network security limits (thermal, voltage or stability constraints). In Europe's zonal market design, cross-border (inter-zonal) constraints are represented in the day-ahead clearing (PCR/EUPHEMIA), while most intra-zonal constraints are addressed by remedial actions (redispatch and countertrading) undertaken by transmission system operators (TSOs). The Clean Energy Package strengthened the obligation to make minimum cross-zonal capacity available to markets (commonly expressed as the "70% rule"), implemented either as a percentage of NTC or, under Flow-Based Market Coupling (FBMC), as a minimum remaining available margin (minRAM) on critical network elements. These provisions directly shape congestion patterns and their market manifestation in zonal prices. (Nord Pool, 2025)

2.2.2 Measurement and indicators

Empirically, congestion is proxied through several observable indicators: (i) congestion income (also called "congestion rents") accruing from price differentials under capacity constraints in coupled auctions; (ii) redispatch and countertrading volumes/costs reported by TSOs and regulators; and (iii) the frequency of binding constraints in flow-based domains. ENTSO-E documents the publication of congestion income on the Transparency Platform and provides metadata for its calculation; daily Core-region figures are visible via its market transparency portal. These series, together with TSO/regulator monitoring of redispatch activity, offer complementary lenses on congestion: the former captures the market value of constraints, the latter the system-operation response to internal bottlenecks. (Entsoe, 2025)

2.2.3 Germany: Redispatch 2.0

Germany provides a rich case study because of pronounced north–south transfer needs driven by wind generation, and extensive interconnections with neighbouring bidding zones. Since 1 October 2021, the regime known as Redispatch 2.0 consolidated previous “feed-in management” for renewables and CHP into the general redispatch framework, expanding the set of controllable resources (including RES/CHP plants and storage), formalising data exchange processes, and harmonising curtailment/compensation rules. The reform sought to improve coordination across grid levels and enhance pre-emptive congestion management through better planning data. (Bundesnetzagentur, 2025)

Recent empirical work employs explainable machine learning to identify the drivers of German redispatch volumes. Using hourly data, (Titz, Putz, & Witthaut, 2023) show that wind generation is the dominant driver of redispatch, with cross-border exchanges and hydropower also materially influencing outcomes; these findings align with the intuition that spatial generation–load imbalances and interconnector usage interact with internal bottlenecks. Their results—available as a 2023 preprint and a 2024 journal article—underscore the value of feature-rich, interpretable models for congestion analysis.

At the statistical reporting level, the Bundesnetzagentur Monitoring Report provides annual overviews of redispatch activity, costs and volumes, enabling trend analysis around the Redispatch 2.0 introduction and subsequent years (with data primarily for 2022 and developments in 2023–24). These official series are useful for calibration and validation of congestion proxies derived from open data. (Bundesnetzagentur, 2025)

2.2.4 Flow-Based Market Coupling (FBMC) and minRAM

Under PCR/EUPHEMIA, the day-ahead auction internalises inter-zonal constraints; the Core region’s shift to FBMC altered how constraints map into prices and scheduled exchanges by representing them via PTDFs and RAMs on critical network elements. Empirical evaluations indicate that the introduction of FBMC in Central Western/Core Europe increased price convergence and affected cross-border exchange volumes, though impacts vary by period and border; regulatory parameters such as minRAM materially influence the frequency of binding constraints and price spreads. These results matter for congestion modelling because they justify including flow-based margin indicators and net position signals as predictors when available. (Ovaere, Kenis, Von Den Bergh, Bruninx, & Delarue, 2023)

Regulatory and operator documentation (ENTSO-E, JAO) details the evolution of Core FBMC, stakeholder feedback on capacity-calculation amendments, and publication tooling for FB domains and constraints. Complementary analyses discuss how enforcement of the 70% minRAM and methodology choices shape available capacity and thus congestion outcomes. Together, these sources provide institutional grounding for modelling choices that incorporate cross-zonal capacity context. (Entsoe, 2025)

2.2.5 Implications for statistical modelling

For a statistics-centred project, the literature and institutional context suggest several design choices:

- Event definition - Congestion can be framed as an event (e.g., redispatch presence/volume above threshold; or market-side proxies such as price divergence conditional on limited RAM/NTC). Clear, operational definitions aligned with available data will affect label quality and evaluation. (Entsoe, 2025)
- Predictors. Feature sets should combine fundamentals (load, wind/solar forecasts, actual generation), cross-border flows/capacities, and where possible flow-based indicators (e.g., RAM/minRAM proxies, net positions). Evidence from Germany points to wind generation and cross-border exchanges as salient predictors of remedial actions.
- Spatial-temporal structure – Because congestion propagates across interfaces rather than within isolated zones, models that respect spatial relations (e.g., multi-zone VARX, graph-based learners, or feature engineering with lagged neighbour variables) are well-motivated by FBMC’s network representation. (Finck, 2021)

- Evaluation – Beyond point accuracy, event-based metrics (precision/recall, ROC-AUC), and calibration of probabilistic outputs are appropriate; comparisons should control for market design changes (e.g., FBMC go-lives, minRAM adjustments) and use rolling-origin back tests aligned with auction timing. (Entsoe, 2025)

In sum, European congestion is simultaneously a market artefact (visible through price differentials and congestion income) and a system-operation outcome (redispatch/countertrading). Germany's Redispatch 2.0 and the Core region's FBMC provide both the institutional rationale and the data signals needed to formulate rigorous, testable statistical models of congestion incidence and intensity.

2.3 Market coupling and flow-based capacity

2.3.1 Single Day-Ahead Coupling (SDAC) and PCR

Europe's day-ahead electricity markets are integrated through Single Day-Ahead Coupling (SDAC). In SDAC, local power exchanges collect orders and submit them to a common welfare-maximising optimisation that respects transmission constraints; the result is a coupled price and schedule across bidding zones. The operational implementation used by the Price Coupling of Regions (PCR) initiative is the EUPHEMIA algorithm. (Nord Pool, 2025)

2.3.2 The EUPHEMIA algorithm

EUPHEMIA solves a large, mixed-integer welfare-maximisation problem over a full 24-hour horizon, accommodating a rich order book (hourly, block and complex orders) while enforcing network constraints. Public descriptions from the exchanges detail the modelling approach (combinatorial optimisation with branch-and-bound and standard solvers), the objective (social welfare), admissible order types, and how interzonal transmission limits enter the clearing problem. For your literature review, these documents are the canonical references for the market-clearing mechanism your data reflect. (EPEX, 2025)

2.3.3 From NTC to Flow-Based Market Coupling

Historically, cross-zonal constraints in the auction were represented via Net Transfer Capacities (NTC) per border. In the Core (formerly CWE) region, this has been replaced by Flow-

Based Market Coupling (FBMC), which models constraints on critical network elements with contingency (CNECs) using power transfer distribution factors (PTDFs) and a remaining available margin (RAM). The feasible set of simultaneous zonal exchanges forms a flow-based domain; market coupling selects the point in this domain that maximises welfare. Operator documentation and academic treatments explain the construction of PTDFs, RAM and the FB domain, and how this richer network representation can alter price convergence and net positions versus NTC. (jao, 2020)

2.3.4 Go-live and scope

Core day-ahead FBMC went live on 8 June 2022 (for delivery day 9 June 2022). ENTSO-E and the Joint Allocation Office (JAO) provide official status notes and user-facing information on the coupling processes. These sources are useful for pinning down structural breaks in your time series around mid-2022. (jao, 2020)

2.3.5 Minimum Capacity (minRAM) and the 70% rule

The EU Electricity Regulation (EU) 2019/943 introduced a minimum cross-zonal capacity requirement, commonly referred to as the 70% rule. In FBMC terms, this is operationalised through minimum RAM (minRAM) parameters on CNECs. ACER monitors compliance and publishes annual assessments and methodological notes; recent updates summarise progress, action plans and derogations across Member States. When you model post-2019 data—especially Core after 2022—these policy parameters provide an institutional rationale for including capacity-related features and for segmenting evaluation periods. (EUR-Lex, 2019)

2.3.6 Empirical evidence on FBMC impacts

The academic evidence has grown:

- Price convergence and exchanges - (Ovaere, Kenis, Von Den Bergh, Bruninx, & Delarue, 2023) estimate the effect of FBMC introduction in CWE/Core on price convergence and cross-border exchange volumes, finding an initial rise in both immediately after go-live, with dynamics evolving thereafter. Their working paper (2022) and journal article (2023) are widely cited empirical evaluations.
- Mechanism and modelling. (Schönheit, et al., 2021) provides a rigorous, didactic treatment of FBMC fundamentals, including sensitivity to Fref/FRM choices and other

parameters that shift the FB domain—highly relevant when interpreting regime shifts in the data.

- Policy parameters (minRAM). Several studies examine welfare and operational trade-offs of raising minRAM (linked to the 70% rule). Results indicate that higher minRAM increases commercial exchange and can improve market integration but may also raise the need for remedial actions if internal bottlenecks bind more frequently. This trade-off is a live research and regulatory topic. (Finck, Impact of Flow Based Market Coupling on the European Electricity Markets, 2025)
- Recent advances. (Bucksteeg, Voswinkel, & Blumberg, 2025)discuss improvements to flow-based coupling design, linking capacity calculation, coupling and post-clearing aspects; Nagy (2025) provides a contemporary chapter comparing NTC and FB outcomes (social welfare, net positions, price convergence) with a focus on active constraints and shadow prices. These are helpful for anchoring a modern literature review.

2.3.7 Current Developments beyond day-ahead

Flow-based capacity calculation is propagating intraday: Core FB-IDCC processes went live in phases during 2024–2025 (D-1 15:00/22:00, then D 04:30). For high-frequency studies or when aligning day-ahead with subsequent intraday adjustments, these go-lives mark further structural changes in available capacities and, potentially, congestion signals. (Entsoe, 2025)

2.3.8 Implications for project

1. Feature engineering. FBMC justifies the inclusion of capacity-context variables (e.g., proxies for RAM/minRAM, net positions, historical binding frequency) alongside fundamentals (load/RES forecasts, realised generation).
2. Regime segmentation. Model estimation and evaluation should account for design changes (FBMC go-live, minRAM revisions) via structural-break flags or period-specific models.
3. Interpretability. Given the policy salience of minRAM and CNEC selection, methods supporting explainability (e.g., SHAP over constrained feature sets) offer value beyond raw accuracy.

4. External validity. Cross-border heterogeneity in CCR design (FB vs NTC) invites difference-in-differences style comparisons or multi-task learning across borders. These design choices are defensible in light of the institutional record and the emerging empirical literature

2.4 Data quality, transparency, and reproducibility

Empirical work on European electricity markets is unusually data-rich but methodologically fragile: many published results hinge on subtle choices about data provenance, time alignment, and evaluation design. Surveys in EPF and congestion modelling consistently stress that careful curation and transparent workflows explain a large share of reported gains, often more than choice of model class. (Lago, Marcjasz, De Schutter, & Weron, 2021)

2.4.1 Provenance and coverage

The ENTSO-E Transparency Platform is the canonical source for market fundamentals (prices, load, generation by technology, forecasts) and system operation (outages, interconnector capacity/flows, congestion income). Coverage is heterogeneous across zones, borders, and time: specific series (e.g., wind/solar intraday forecasts; forecasted transfer capacities) are missing for some areas or periods; "data item" definitions also evolve. Studies therefore document exact endpoints, schema versions, and any national-TSO supplements to mitigate gaps.

2.4.2 Timing, clocks and publication lags

European datasets interleave CET/CEST timestamps, local time labels, and UTC; daylight-saving transitions create duplicate/absent hours. Publication lags differ by series (e.g., forecasts vs actuals), and several fields are revised ex post. The literature warns that misaligned clocks and post-gate-closure data can induce look-ahead bias in day-ahead studies; robust work aligns all predictors to information available at or before gate closure and documents any lead/lag conventions.

2.4.3 Harmonisation and unit conventions

Researchers routinely standardise units (MW vs MWh), naming (bidding zones/borders that change over time), and calendars (public holidays, daylight-saving flags). For flow-based regions, reproducible analyses also state how they map CNE(C)s to hours (e.g., selecting RAM at base case vs.

contingency) and how they aggregate from constraint-level artefacts to zonal or border-level indicators.

2.4.4 Revisions, backfills, and missingness

ENTSO-E series may be revised days to months after first publication, and some TSOs backfill outages or generation. Best practice records both the first-seen ("as-observed") and the latest-available values, pins a dataset snapshot for modelling, and reports sensitivity to late revisions. Missingness is often structured (border- or technology-specific); the literature recommends explicit imputation rules (e.g., short-gap interpolation with safeguards; model-based imputation for forecasts) and, where feasible, border-subset analyses to avoid bias.

2.4.5 Label construction and ground truth

"Congestion" lacks a single observable label. Common proxies include (i) congestion income and price divergence across coupled zones; (ii) redispatch/countertrading volumes and costs; and (iii) flow-based binding indicators (RAM saturations, limiting CNECs). Each proxy captures a different facet—market vs. system operation—and each has blind spots (e.g., internal constraints without price divergence). High-quality studies justify the label choice, report alternative labels in robustness checks, and acknowledge measurement error. (Entsoe, 2025)

2.4.6 Structural breaks and policy regime shifts

Design changes—Core FBMC go-live, subsequent parameter updates (e.g., minRAM related to the 70% rule), bidding-zone splits/mergers, or balancing-market reforms—create regime shifts that alter both the data-generating process and the meaning of features. The literature typically (a) introduces regime dummies, (b) re-estimates models by period, or (c) applies difference-in-differences style comparisons across affected/unaffected borders.

2.4.7 Evaluation protocols and leakage control

Methodological reviews emphasise rolling-origin back-testing, market-aware folds, and strict separation of training/validation/test periods to prevent temporal leakage (Lago, Marcjasz, De Schutter, & Weron, 2021)For classification or event prediction (e.g., "redispatch day"), class imbalance is handled with calibrated thresholds, cost-sensitive metrics, and precision–recall reporting. Probabilistic assessments rely on CRPS/quantile scores and calibration diagnostics; where predictions drive trading or operations, studies increasingly report decision-relevant metrics alongside statistical error.

2.4.8 Transparency of feature engineering

Given the sensitivity of results to engineered predictors (residual load, ramps, utilisation/headroom, proximity-to-FB boundary), replicable work specifies exact formulas, windows, outlier policies (e.g., negative prices, price caps), and treatment of zero/negative RAM or curtailment events. For features derived from FB data, authors clarify whether they use cleared net positions, pre-coupling domains, or ex post realised constraints.

2.4.9 Reproducible research practices

The field increasingly expects (i) version-controlled code and environment manifests, (ii) immutable data snapshots with checksums/DOIs, (iii) seeded randomness and deterministic inference where possible, and (iv) experiment logs that map figures/tables to runs. When full data release is restricted, papers share synthetic or down-sampled replicas plus scripts that rebuild all figures from public sources.

2.4.10 Ethical and compliance considerations

Studies drawing on near-real-time or high-frequency operational data note market-abuse and confidentiality constraints. Reproducible workflows therefore avoid mixing non-public data with public benchmarks without a clear disclosure and, where necessary, delay publication or aggregate sensitive series.

2.5 Synthesis, gaps, and implications for the study design

2.5.1 What the literature establishes

Taken together, Sections 2.1–2.3 indicate a fairly mature state of electricity price forecasting (EPF) and a clearer institutional understanding of how European market design shapes prices and congestion. On the modelling side, well-specified statistical baselines (ARX/VARX families, often with volatility or regime components) remain competitive when enriched with exogenous fundamentals and evaluated with leakage-safe, rolling procedures (Lago, Marcjasz, De Schutter, & Weron, 2021) Machine-learning and deep-learning methods can add value where non-linear interactions and high-dimensional inputs matter, but gains are contingent on careful feature engineering and robust evaluation practice. Probabilistic forecasting—quantiles, intervals and calibration diagnostics—has moved from "nice-to-have" to expected practice in operational settings.

On the institutional side, the Single Day-Ahead Coupling (SDAC) with PCR/EUPHEMIA embeds cross-zonal transmission constraints directly into day-ahead clearing; the Core region's transition from NTC to Flow-Based Market Coupling (FBMC) represents these constraints with PTDF-based critical network elements and remaining available margins (RAM). Policy enforcement of minimum capacity (the "70% rule," often operationalised via minRAM) further conditions observed price convergence, net positions and the frequency/severity of congestion. The German Redispatch 2.0 regime has, in parallel, expanded and systematised remedial actions for intra-zonal constraints. Empirical work for Germany highlights wind generation, cross-border exchanges and hydropower as salient drivers of redispatch volumes, reinforcing the value of congestion-aware features in EPF (Titz, Putz, & Witthaut, 2023).

2.5.2 Gaps and opportunities

Despite the breadth of the literature, several under-served areas align directly with the aims of this project:

1. Congestion-aware EPF remains patchy. Many price-only studies do not explicitly incorporate capacity context (RAM/minRAM indicators, binding-frequency proxies, or net-position dynamics), even though SDAC/FBMC makes these first-order drivers when constraints bind.
2. Regime segmentation is often insufficient. Comparisons that straddle major design changes (FBMC go-lives, minRAM parameterisation, Redispatch 2.0 rollout, intraday FB milestones) risk conflating structural effects with model performance.
3. Evaluation beyond point accuracy is uneven. Distributional scoring (e.g., CRPS, pinball loss), calibration checks, and event-based metrics for congestion are not universally reported, despite their relevance for trading and system-operation use-cases.
4. Data engineering is under-documented. Studies frequently reference ENTSO-E Transparency data, TSO reports and exchange publications, but provide limited detail on curation steps

(auction-time alignment, DST handling, API edge cases, backfills/revisions), even though these steps materially affect outcomes.

5. Spatio-temporal dependence is under-exploited at the zonal level. Graph-based or multi-task approaches that respect cross-border coupling are still relatively sparse for European zonal markets compared with nodal/LMP contexts.

2.5.3 Implications for this project

Across EPF and congestion studies, several recurrent practices and operationalisations emerge.

- Outcome definitions commonly adopted - Studies typically target (a) day-ahead zonal prices (point and, increasingly, probabilistic outputs) and (b) congestion indicators captured either as events (e.g., redispatch occurrence or price divergence conditional on limited capacity) or intensity measures (e.g., redispatch volumes/costs, congestion income, or frequency of binding RAM/NTC constraints).
- Predictor Families Routinely report:
    - Fundamentals: load and RES forecasts, realised generation by technology, and weather covariates.
    - Cross-border context: scheduled flows/NTC (or RAM/minRAM proxies where available), net positions, and historical binding information.
    - Spatial structure: neighbour-zone prices/flows and lags to encode market coupling.
    - Calendar/auction alignment: delivery-hour effects, weekday/holiday dummies, seasonality, and precise auction-time alignment.
    - Policy/ regime markers: FBMC go-live, minRAM phases, and redispatch framework changes.
- Model classes and reporting conventions - Transparent statistical baselines (ARX/VARX, with volatility/regime elements) are widely retained alongside non-linear learners (e.g., boosted trees) where interactions are salient. Probabilistic outputs are evaluated with

pinball loss/CRPS and calibration checks. Rolling-origin evaluation and leakage-aware splits are emphasised, particularly when samples span design changes.

- Interpretability and auditability - Where capacity context is included, studies increasingly report variable importance or post-hoc explanations (e.g., SHAP) over constrained feature sets, and provide at least high-level documentation of data curation (calendar alignment, revisions/DST handling).

2.5.4 Limitations of the literature

Several structural limitations recur across the corpus and condition how findings should be read.

- Measurement and observability - Congestion proxies (congestion income, redispatch statistics, RAM/minRAM indicators) are imperfect and vary in coverage and granularity across periods and borders.
- Regime heterogeneity - Comparisons that straddle FBMC/NTC regimes or evolving minRAM settings can conflate institutional shifts with model performance; many studies only partially segment samples.
- Data curation opacity - Exact ingestion and cleaning steps (API revisions, backfills, DST quirks) are often under-documented, complicating replication and cross-study comparison.
- External Validity - Results can be market- and period-specific given differences in interconnection topology, RES mix, and national implementations of EU rules.
- Proprietary blind spots - Practitioner methods and high-frequency datasets are frequently undisclosed, so academic benchmarks may under- or over-state attainable accuracy in production contexts.

2.5.5 Chapter summary

The reviewed literature establishes that (i) well-specified statistical models enriched with fundamentals remain competitive for day-ahead horizons; (ii) non-linear methods can add value when capacity context and RES-driven interactions matter; and (iii) Europe's market design—SDAC with

FBMC and minimum-capacity enforcement—creates identifiable congestion mechanisms that are visible in prices, flows, and remedial actions. At the same time, gaps persist around explicit congestion-aware modelling, regime-sensitive evaluation, and transparent data engineering. These conclusions delineate the empirical choices typically seen in the field and the constraints under which results should be interpreted.

Chapter 3 – Data Infrastructure

This chapter describes the data infrastructure that supports the analysis. Its purpose is to ensure that all datasets used for signal construction and back-testing are acquired, transformed, and stored in a manner that guarantees temporal correctness, reproducibility, and auditability. The system provides controlled ingestion from the ENTSO-E Transparency Platform, normalisation to a common event-time convention, and persistence in a time-partitioned relational store. These design choices prioritise methodological rigour over convenience and allow any result in later chapters to be regenerated from first principles.

The platform comprises three layers: (i) an asynchronous client that interacts with the source API under explicit rate limits; (ii) a data pipeline that validates, normalises, and writes idempotently to storage; and (iii) a repository layer that exposes analysis-ready extracts to notebooks and dashboards. All components are containerised and driven by a small set of declarative commands, enabling independent reproduction of data cuts without manual intervention.

3.1 Design Goals

3.1.1 Rationale for a data platform

A self-contained platform was adopted to support long-horizon back-tests and congestion-aware feature engineering. Three properties are essential for credible empirical results:

1. Event-time correctness across all series.
2. Repeatable and fast query results
3. Operability so that an independent reviewer can stand up the system and regenerate the data without bespoke configuration.

3.1.2 Requirements

- Rate-limited ingestion. Enforce the provider's published limits (requests per window and concurrency) with back-off, jitter, and resumable jobs.
- Idempotent loads and revisions. Re-runs and late provider updates converge to a single truth via natural keys; duplicates and drift are excluded.

- Event-time correctness. Reconstruct timestamps from period start, resolution and position; store in UTC; handle daylight-saving transitions explicitly. Features reflect only information available at or before gate closure.
- Large time-series persistence. Support full backfills, point-in-time snapshots, and multi-year retention without sacrificing query performance.
- Notebook-first analysis. Provide tidy extracts for interactive analysis that can operate offline once data are local.
- One-command operation. Start/stop services, collect recent history or full backfills, and reset the environment with a small, memorable command set.
- Observability. Emit structured logs (dataset, zone/border, window, attempts, rows, latency) and expose a lightweight dashboard for freshness, throughput, error rates, and table growth.
- Easy to Install. Allow a clean installation on common operating systems in a small number of commands.
- Practical scope. Provide an easy path to retrieve the previous five years of ENTSO-E data for selected countries; batch rather than real-time operation is sufficient for this study, but queries and batch jobs should be as fast as possible.

3.2 High Level Architecture

3.2.1 Containers

The system runs as two docker containers on a single host:

- Application container (Python 3.13) – Performs data collection from the ENTSO-E Transparency API, applies batch processing (parsing, normalisation, validation), and writes results to the database. It also exposes simple functions that notebooks can call to retrieve analysis-ready tables.
- Time-series database container (PostgreSQL/TimescaleDB) - Stores all series in time-partitioned tables with composite indices for fast reads. The database uses PostgreSQL 16.

3.2.2 User Facing Tools

- Jupyter notebooks (Python 3.13) - Used for exploratory analysis and back-tests. Notebooks read from the database and may call the repository functions in the application container.
- Grafana dashboards - Used to monitor data freshness, throughput and errors; panels read from the database (and/or exposed metrics).

3.2.3 Data Flow

1. Python application calls the ENTSO-E API under explicit rate limits.
2. Responses are parsed, normalised to event time in UTC, and validated.
3. Records are written in batches to the time-series database (idempotent upserts).
4. Notebooks query the database for analysis.
5. Grafana queries the database for operational views.

This layout keeps responsibilities clear (ingestion/processing vs storage), is easy to reproduce on any machine with Docker, and supports quick iteration in notebooks while maintaining an auditable store. ENTSOE-E was contacted to obtain an API key; these were stored as environment variables.

3.2.4 Architecture diagram

Figure 1 summarises the runtime layout: the Python application ingests ENTSO-E data, performs batch processing, and writes to a PostgreSQL/TimescaleDB container; Grafana queries the database for monitoring.

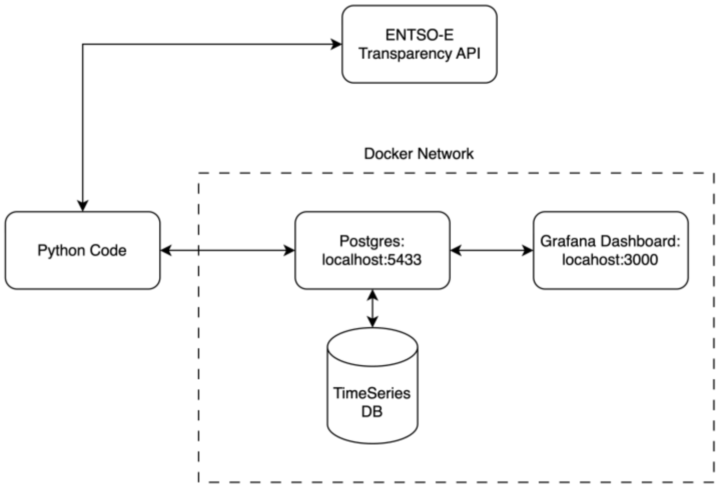

Figure 1 - High Level Architecture Diagram

3.3 Data Sources

3.3.1 Source

All data are obtained from the ENTSO-E Transparency Platform via the project's asynchronous client and stored in a time-partitioned relational store. The study restricts itself to this single, authoritative source to maximise consistency across series and simplify auditing of provenance.

3.3.2 Coverage

The platform ingests a focused set of series sufficient for congestion-aware analysis across selected European bidding zones (DE, FR, NL, BE, AT, CH, PL, CZ, DK1, DK2). The supported series exposed through the repository layer are: day-ahead prices; actual load; load forecasts; actual generation by PSR type; generation forecasts; cross-border flows; transmission capacity; allocated capacity; net positions; and congestion income.

3.3.3 Storage Schema

Tables are defined with SQLAlchemy and backed by natural uniqueness constraints to guarantee idempotent loads. Representative tables and keys are:

- **day_ahead_prices** — columns include timestamp, value, currency, country, eic_code, resolution; unique on (timestamp, country, eic_code, resolution); composite index on (timestamp, country).
- **actual_load** — timestamp, value, unit, country, eic_code; unique on (timestamp, country, eic_code).
- **load_forecasts** — timestamp, value, unit, country, eic_code, forecast_type; unique on (timestamp, country, eic_code, forecast_type).
- **actual_generation** — timestamp, value, unit, country, eic_code, psr_type; upserts keyed by (timestamp, country, eic_code, psr_type).
- **generation_forecasts** — timestamp, value, unit, country, eic_code, psr_type, forecast_type; upserts keyed by (timestamp, country, eic_code, psr_type, forecast_type).
- **cross_border_flows** — timestamp, value, unit, from_country, to_country, from_eic, to_eic; upserts keyed by (timestamp, from_eic, to_eic).
- **transmission_capacity** — schema analogous to flows; upserts keyed by (timestamp, from_eic, to_eic).
- **allocated_capacity** — schema analogous to flows; upserts keyed by (timestamp, from_eic, to_eic).
- **net_positions** — timestamp, value, unit, currency, country, eic_code; upserts keyed by (timestamp, country, eic_code).
- **congestion_income** — timestamp, value, unit, currency, country, eic_code; upserts keyed by (timestamp, country, eic_code).

3.3.4 Event time, units and resolution

All series are stored with a "timestamp" column in UTC and appropriate unit/currency metadata, enabling consistent joins across sources and precise filtering by period. The repository methods coerce timestamp to time zone-aware pandas "datetime64[ns, UTC]" and numeric types for values at read time.

3.3.5 Scope and limitations

The study used day-ahead prices and does not include intraday trades or order-book data. This was deliberate, to get live or historic intraday prices for various countries would require access to a paid 3rd party resource. All data used in this project is open source and free to access.

3.4 Gap compared to industry standard

This section contrasts the current platform with common practice in production data teams and sets out pragmatic extensions. The aim is to show that the research system already follows sound engineering patterns while identifying the limited steps required for an operational deployment.

3.4.1 Aligned with Industry Practice

- Time-series storage. Use of a relational time-series engine with time partitioning and composite indices is standard for operational analytics.
- Idempotent ingestion. Upserts keyed by natural identifiers remove duplicates and allow safe re-runs, which is a typical control in regulated environments.
- Containerisation. Services are isolated and reproducible via Docker; local bring-up is scripted.
- Configuration over code. Source parameters and limits are declared in configuration, making behaviour auditable.
- Basic observability. Structured logs and a Grafana dashboard provide visibility of freshness, throughput and errors.
- Notebook access. A repository layer exposes tidy, well-typed extracts for analysis, reducing one-off script work.

3.4.2 Deliberate scope limits in this study

- Coverage. The system relies on the ENTSO-E Transparency Platform and focuses on day-ahead series; intraday and commercial feeds are out of scope.
- Operating model. Ingestion is batch-oriented and triggered on demand. Given ENTSO-E's update cadence and the historical focus of the analysis, real-time streaming is unnecessary for the research questions addressed here.

3.4.3 Gaps to an industry production deployment

- Scheduling & SLAs: trigger collection shortly after provider publication; record and alert on success/failure.
- Data-quality checks: completeness, duplicates, DST edges, value-range sanity, schema drift; block downstream steps on failure.
- Alerting: thresholds for missing updates, high retry rates, or stalled jobs; notify via e-mail/chat.
- Secrets & access control: replace local .env with a secrets manager; least-privilege DB roles; key rotation.
- Backups & retention: regular backups with restore tests; retention policies; optional point-in-time snapshots.
- CI/CD & testing: unit tests for parsers, integration tests on a staging DB, schema migrations; automated build and vulnerability scanning.
- Documentation & runbooks: concise guides for bring-up, recovery, and routine operations.
- Intraday or third-party data: integrate a paid vendor for intraday prices; if streaming is provided, ingest via Kafka (or a managed equivalent) to buffer/partition events and write idempotently into the existing tables.
- Deployment options: package the two containers for a managed cloud or on-prem host; enable compression and read replicas for scale.
- Metadata & lineage: generate dataset documentation/lineage (e.g., dbt-style docs) so joins and assumptions are explicit.

3.4.4 Summary

The items above outline what might be required for an operational deployment and are out of scope for this MSc thesis. The current platform already provides time-indexed storage, idempotent ingestion, containerised services, and basic monitoring. Moving to production would mainly involve scheduling, automated data-quality checks, secrets management, and routine operational safeguards—incremental rather than architectural changes that would not alter the analytical interface used elsewhere in this thesis.

3.5 Software Environment and External Dependencies

All modelling, analysis, data processing, and orchestration were built using open-source software tools, managed in a reproducible Python environment via Poetry, and deployed locally using Docker containers. The table below lists all packages and tools used throughout the project.

3.5.1 Core Python Environment:

- Python 3.12 – Primary programming language (Python, 2023)
- Poetry 1.7.1 – Python dependency manager (Poetry, 2023)
- Jupyter 1.1.1 – Interactive Notebook (Jupyter, 2024)
- Notebook 7.4.5 – Backend server for running Jupyter Notebooks (Notebook, 2025)
- Ipykernal 6.25.0 – Kernal interface for Jupyter integration (Ipykernal, 2025)

3.5.2 Data Processing and Numerical Computing:

- Pandas 2.1.0 – Data manipulation and analysis (Pandas, 2023)
- Numpy 1.24.0 – Array operations and numerical computing (Numpy, 2022)
- Scipy 1.11.0 – Scientific and statistical utilities (SciPy, 2025)
- Tqdm 4.67.1 – Progress bars for iterative tasks (Tqdm, 2024)

3.5.3 Machine Learning and Forecasting:

- Sckit-Learn 1.3.0 – Model training, evaluation, and cross-validation (Sckit-Learn, 2023)
- Pydantic 2.4.0 – Data validation and structured type enforcement (Pydantic, 2023)
- Pydantic-settings 2.0.0 – Configuration management via pydantic (Pydantic-settings, 2023)
- Aiohttp 3.8.0 – Asynchronous HTTP client for external data (Aiohttp, 2025)

3.5.4 Database, ORM and HTTP Utilities:

- SQLAlchemy 2.0.0 – ORM and SQL query abstraction. (SqlAlchemy, 2023)

- Python-dotenv 1.0.0 – Environment variable management from ".env" (Python-Dotenv, 2023)
- Requests 2.31.0 – Synchronous HTTP client (Requests, 2023)

3.5.5 Visualisation and Plotting

- Plotly 5.17.0 – Interactive charts and dashboards (Plotly, 2023)
- Matplotlib 3.7.0 – Static, publication-quality plots (Matplotlib, 2023)
- Seaborn 0.12.0 – Statistical visualisation built on matplotlib (Seaborn, 2024)

3.5.6 Code Quality and Linting

- Black 23.9.0 – Code formatter (PEP8-compliant) (Black, 2023)
- Loguru 0.7.0 – Structured logging library (Loguru, 2023)

3.5.7 Deployment, Visualisation and Time Series Infrastructure

- Docker – Containerisation platform (Docker, 2025)
- Docker-Compose 3.8 – Multi-service orchestration (Docker-Compose, 2025)
- Timescaledb:latest-pg16 – PostgreSQL extension for time-series data (Timescale, 2025)
- Grafana/grafana:latest – Monitoring Dashboard Platform (Grafana, 2025)

Chapter 4 – Data & Exploratory Analysis

4.1 Overview

This chapter presents the exploratory statistical analysis conducted on the key variables underpinning the modelling and forecasting strategies developed in later stages of the project. The primary objectives are to validate data quality, summarise the behaviour of relevant series, and identify patterns that inform model design. These series were retrieved from the ENTSO-E Transparency Platform, processed through a custom data pipeline (see Chapter 3), and normalised to a consistent timestamp format.

A particular emphasis is placed on location price spreads, which serve as proxies for cross-border transmission constraints, offering insight into market separation events and the operational limits of interconnectors.

By characterising the statistical properties of these spreads—including their volatility, temporal structure, and frequency of extreme values—this chapter provides empirical motivation for the modelling approaches adopted in Chapter 5. The analysis also supports feature selection by highlighting potential explanatory variables and candidate congestion signals.

4.2 Data Acquisition and Preprocessing

The datasets used in this project were obtained from the ENTSO-E Transparency Platform, covering the period from 1st January 2021 to mid-2025. All data was retrieved via an automated, containerised pipeline described in Chapter 3, and stored in a time-series relational database for ease of access and reproducibility.

The exploratory analysis focused on day-ahead hourly electricity prices for Germany (DE) and its immediate neighbours: France (FR), the Netherlands (NL), Austria (AT), Switzerland (CH), Belgium (BE), Poland (PL), and the Czech Republic (CZ). These zones were selected due to their direct interconnection with Germany and their relevance to cross-border congestion dynamics.

Each time series underwent a basic validation routine to ensure:

- Timestamps were in UTC and aligned at hourly resolution

- Missing or malformed entries were filtered
- Duplicate timestamps were removed
- Series were sorted chronologically

A summary of data coverage and resolution across zones is presented in Table 1.

Table 1 - Summary Day-Ahead Price Data

Zone	Rows	Duplicate Timestamps	Resolution(s)	Start Date	End Date
DE	202,941	40,564	PT15M, PT60M	01/01/2021 00:00	27/08/2025 21:45
FR	40,411	24	PT60M, PT15M	01/01/2021 00:00	27/08/2025 21:00
NL	40,553	0	PT60M	01/01/2021 00:00	27/08/2025 21:00
AT	197,152	40,063	PT15M, PT60M	01/01/2021 00:00	27/08/2025 21:45
CH	40,692	0	PT60M	01/01/2021 00:00	27/08/2025 21:00
BE	40,637	0	PT60M	01/01/2021 00:00	27/08/2025 21:00
PL	40,348	0	PT60M	01/01/2021 00:00	27/08/2025 21:00
CZ	40,553	0	PT60M	01/01/2021 00:00	27/08/2025 21:00

The pre-processed data was then assembled into a multi-zone hourly price panel, from which pairwise price spreads were constructed for further analysis.

4.3 Price Spread Construction and Characteristics

4.3.1 Price Spreads

Pairwise price spreads between Germany (DE) and each of its neighbouring bidding zones were constructed to capture relative market values and potential congestion signals. Each spread is defined as:

$$Spread_{DE-X}(t) = P_{DE}(t) - P_X(t)$$

Where $P_{DE}(t)$ and $P_x(t)$ denote the hourly day-ahead prices for Germany and bidding zone X, respectively, at timestamp t. These spreads serve as both potential predictors and targets in later modelling tasks, depending on how congestion is operationalised.

The spreads were computed from a harmonised hourly panel and cleaned to remove any hours with missing or duplicate values. Table 2 summarises key distributional properties of each spread, while Figure 2 shows day-ahead price spreads.

Table 2 – Summarises key distributional properties of each spread

	Start	End	N	Mean	Std	P05	P50	P95
DE-FR	2021-01-01 00:00:00+00:00	2025-08-27 21:00:00+00:00	40387	-3.59141	50.56443	-86.528	0	60.99675
DE-NL	2021-01-01 00:00:00+00:00	2025-08-27 21:00:00+00:00	40553	-2.04381	28.68474	-47.461	0	29.547
DE-AT	2021-01-01 00:00:00+00:00	2025-08-27 21:00:00+00:00	40798	-10.9207	32.16565	-71.6453	-1.9225	16.46325
DE-CH	2021-01-01 00:00:00+00:00	2025-08-27 21:00:00+00:00	40692	-16.9328	44.95165	-106.855	-4.07875	24.648
DE-BE	2021-01-01 00:00:00+00:00	2025-08-27 21:00:00+00:00	40637	-1.37344	30.09783	-51.2755	0	35.531
DE-PL	2021-01-01 00:00:00+00:00	2025-08-27 21:00:00+00:00	40348	8.265377	65.42405	-64.4956	-1.39	138.2265
DE-CZ	2021-01-01 00:00:00+00:00	2025-08-27 21:00:00+00:00	40553	-6.43526	24.70644	-50.245	-0.36	16.7985

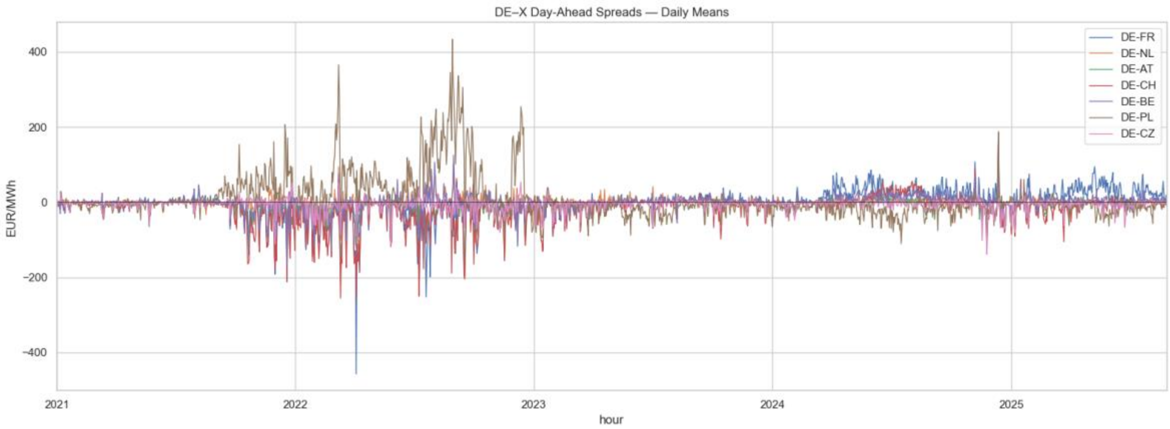

Figure 2 – Day-ahead Price spreads since 2021

4.3.2 Key Characteristics

In Table 2 several notable patterns emerge:

- Day-ahead prices in Germany are typically lower than those in its neighbours — particularly Austria (DE–AT: –10.92 €/MWh) and Switzerland (DE–CH: –16.93 €/MWh). Poland is an exception, with a positive mean of +8.27 €/MWh, suggesting on mean Germany often prices higher than Poland.
- Volatility and Dispersion: The standard deviation of spreads varies significantly across borders. DE–PL shows the highest spread volatility (65.42 €/MWh), followed by DE–FR (50.56 €/MWh) and DE–CH (44.95 €/MWh). In contrast, DE–CZ and DE–NL are more stable, with standard deviations of 24.71 and 28.68 €/MWh, respectively.

- Tails and Extreme Values: Several spreads exhibit large tails, as seen in their 5th and 95th percentiles. For example, 5% of DE–FR spreads fall below –86.53 €/MWh and 5% of DE–PL spreads exceed +138.23 €/MWh. These extreme values are potential indicators of congestion or temporary market separation.
- Central Tendency: The median spread is close to zero for most borders (e.g. DE–FR, DE–NL, DE–BE), implying symmetric distribution around the mean in many cases. However, DE–CH and DE–AT exhibit skewed distributions with more consistently negative values.

These characteristics suggest that certain spreads, particularly those involving CH, PL, and FR, are more likely to capture congestion-related signals and are prime candidates for both forecasting and trading strategies.

4.4 Visual Exploration

To complement the descriptive statistics in Section 4.3, this section presents two key visualisations of the DE–X price spreads: a distributional summary via boxplots, and the average intraday spread profile across borders.

4.4.1 Distribution of Hourly Spreads

Figure 3 shows the distribution of hourly price spreads between Germany and its neighbouring zones. Several patterns are evident:

- Most spreads are centred around zero, but with considerable dispersion.
- The DE–PL spread exhibits a wide, positively skewed distribution, indicating frequent episodes of Germany pricing significantly above Poland.
- Conversely, DE–CH and DE–AT are skewed negatively, reflecting persistent periods where German prices are below those of Switzerland and Austria.
- Outliers are excluded to better highlight the interquartile range.

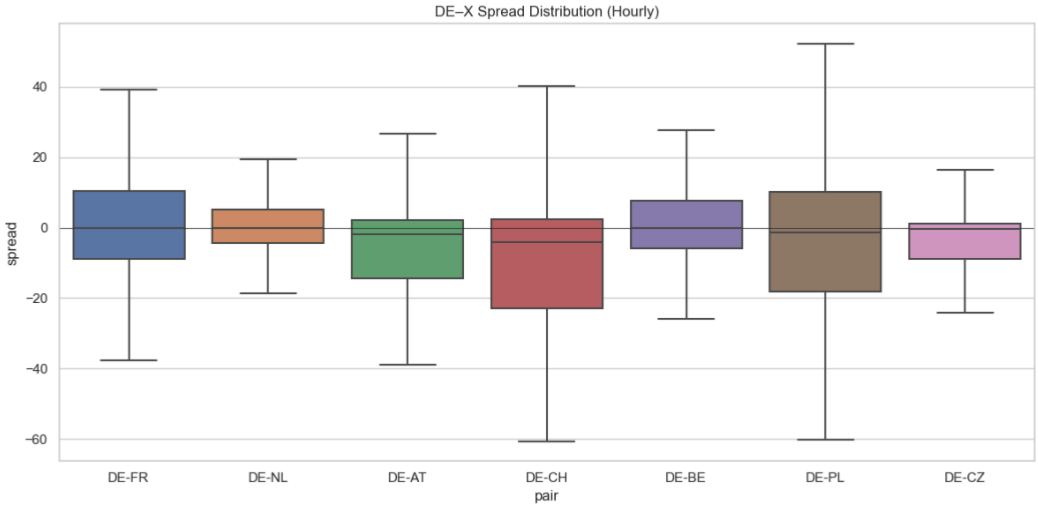

Figure 3 - Distribution of DE-X Price Spreads

4.4.2 Intraday Spread Profile

Figure 4 - Average Intraday Spread Profile illustrates the average intraday profile of each DE–X spread, computed across all hours in the dataset. The plot reveals clear diurnal variation:

- Most spreads widen during daytime peak hours (typically 06:00 to 18:00).
- DE–FR, DE–NL, and DE–PL show pronounced morning and evening peaks, likely reflecting load-driven price dynamics.
- DE–CH and DE–AT maintain consistently negative values throughout the day, suggesting structural differences rather than time-of-day effects.

This variation indicates the presence of daily cyclical effects in spread behaviour, which may justify the use of time-of-day features in forecasting models.

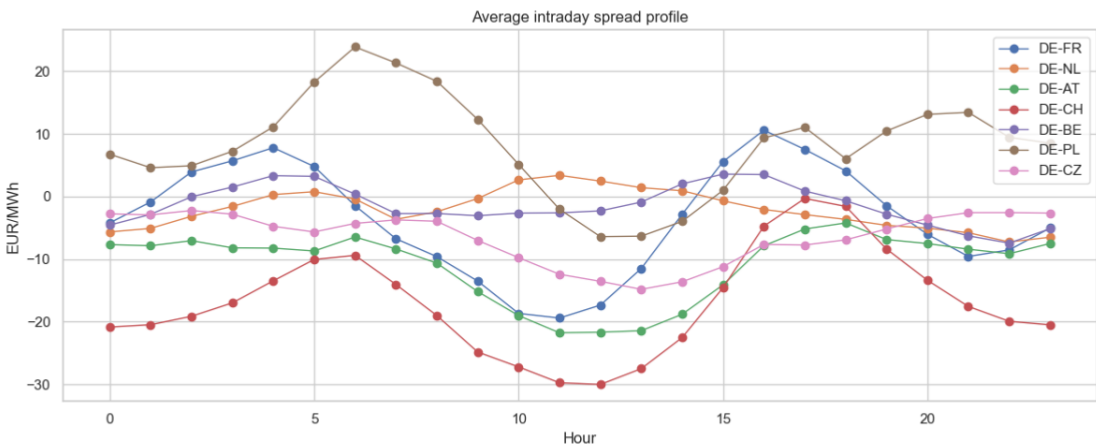

Figure 4 - Average Intraday Spread Profile

4.5 Spread Structure and Congestion Proxies

In this section a series of empirical diagnostics are presented to assess how spread behaviour relates to observable fundamentals and physical network constraints. These diagnostics offer DA-safe congestion proxies that inform the feature set used in Chapter 5. The analysis focuses on four key areas: relationships with DA renewable forecasts, physical flow misalignments, correlation with congestion income, and preliminary cross-checks with actual generation.

4.5.1 Spread Sensitivity to DA Renewable Forecasts

Table 3 shows the Pearson correlations between DE–X spreads and four DA forecast features: the levels and first differences (ramps) of solar and wind generation forecasts.

Table 3 - Correlation of Spreads with DA Renewable Forecast Features

Pair	Solar DA	Solar DA Ramp	Wind DA	Wind DA Ramp
DE–AT	–0.044	0.053	–0.445	–0.090
DE–BE	–0.060	–0.110	–0.316	0.06
DE–CH	–0.092	–0.049	–0.497	–0.043
DE–CZ	–0.011	0.104	–0.393	–0.071
DE–FR	–0.152	–0.125	–0.331	0.013
DE–NL	0.102	–0.045	–0.263	0.092
DE–PL	–0.121	0.206	–0.205	–0.139

Negative correlations dominate, especially between wind forecasts and spreads. For example, the DE–CH and DE–AT spreads show strong negative correlations with DA wind forecasts (–0.497 and –0.445, respectively), suggesting that high wind expectations tend to widen negative spreads, consistent with internal congestion due to renewable surpluses. Solar effects are weaker and more variable across borders.

4.5.2 Flow-Spread Misalignment as a Congestion Signal

Another DA-safe congestion proxy is the degree to which cross-border physical flows contradict economic logic—i.e., power flowing from the higher-priced to the lower-priced zone. Table 4 reports the share of hours where this occurs, along with the correlation between spread and flow direction.

Table 4 - Flow Spread Diagnostics

Pair	% Wrong Direction	Corr(Spread, Flow)	% Blocked (Spread High, Flow Low)	n_obs
DE–AT	61.00%	–0.25	3.60%	336
DE–BE	16.80%	–0.42	0.00%	333
DE–CH	14.00%	–0.34	0.00%	336
DE–FR	3.00%	–0.30	0.00%	234
DE–NL	19.80%	–0.34	0.00%	237
DE–PL	28.80%	–0.30	2.40%	333

The DE–AT border exhibits a very high rate of directional inconsistency (61%), possibly due to non-market-based redispatch. Other borders show more economically intuitive flow behaviour, though still with measurable inefficiencies.

4.5.3 Spread Magnitudes and Congestion Incomes

Table 5 shows the correlation between absolute spread values and German congestion income. While not causative, this serves as a validation check for spread-based proxies.

Table 5 - Correlation of spread with DE Congestion Income

pair	corr(|spread|, CI_DE)	95% CI	n_obs
DE-AT	-0.068	(-0.151, 0.006)	34031
DE-BE	0.219	(0.164, 0.275)	33916
DE-CH	0.082	(0.003, 0.152)	33934
DE-CZ	-0.094	(-0.151, -0.044)	33838
DE-FR	0.188	(0.094, 0.274)	33666
DE-NL	0.173	(0.121, 0.22)	33833
DE-PL	0.098	(0.062, 0.135)	33673

Strongest positive correlations are observed on the BE, FR, and NL borders, reinforcing that large spreads here are aligned with physical congestion rents.

4.5.4 Cross-Checks Using Actual Generation

For completeness, Table 6 presents correlation results using actual hourly generation rather than forecasts. While not DA-safe, this analysis validates that the same structural patterns persist even with perfect foresight.

Table 6 - Correlation of Spread with Actual Wind and Solar Generation

Pair	Corr(Wind, Spread)	95% CI	N
DE–AT	–0.333	(–0.465, –0.184)	240
DE–BE	–0.174	(–0.387, –0.125)	237
DE–NL	–0.240	(–0.368, 0.047)	240
DE–FR	0.097	(–0.122, 0.313)	234
DE–CH	–0.240	(–0.368, –0.047)	240
DE–CZ	–0.224	(–0.385, –0.078)	237
DE–PL	–0.092	(–0.347, –0.096)	237

4.6 Regime Shifts and Structural Breaks

Congestion patterns and price spreads in European power markets are not stationary. They reflect both underlying physical system dynamics and policy or geopolitical shocks. This section investigates key structural regimes observed over the sample period and their implications for modelling.

4.6.1 Historical Regimes and Policy Makers

Based on observable shifts in spread behaviour and known policy events, the following regime boundaries are defined:

Table 7 - Regime Boundaries

Regime	Period	Description
Pre-2022	up to 31 Dec 2021	Stable pre-crisis phase under normal market operation
Crisis 2022	1 Jan 2022 – 31 Dec 2022	Energy crisis: Russia–Ukraine war, gas price spikes
Post-2023	from 1 Jan 2023 onward	Stabilisation phase, updated market arrangements

Notable internal events include:

- Redispatch 2.0 (Oct 2021): Shift in congestion management
- FBMC Go-Live (June 2022): Flow-based market coupling in Core region

These events likely contributed to the shifts in observed spread dynamics.

4.6.2 Share of High-Spread Hours by Regime

A key congestion proxy is the share of hours where the absolute spread exceeds the 90th percentile. Table 8 presents this indicator by border and regime.

Table 8 - % of Hours with High Absolute Spread

Pair	Crisis 2022	Post-2023	Pre-2022
DE–FR	22.50%	6.30%	6.90%
DE–NL	25.90%	4.40%	8.80%
DE–AT	24.40%	5.40%	7.80%
DE–CH	25.70%	4.60%	8.60%
DE–BE	24.00%	5.00%	9.10%
DE–PL	33.80%	2.70%	5.10%
DE–CZ	20.70%	7.70%	5.00%

During 2022, all borders experienced a surge in high-spread hours, consistent with extreme fuel prices, transmission bottlenecks, and security-of-supply concerns. The post-2023 period shows a dramatic return to normality, with most borders under 8% — indicating increased stability or reduced market stress.

4.6.3 Monthly Evolution and Annotated Events

Figure 5 plots the monthly average share of high-spread hours across all DE–X pairs, highlighting structural shifts and policy events.

- A sustained spike is visible through most of 2022
- The drop-off begins in early 2023 and stabilises by Q2
- Key vertical markers (Redispatch 2.0, WAR, FBMC) show coinciding inflection points

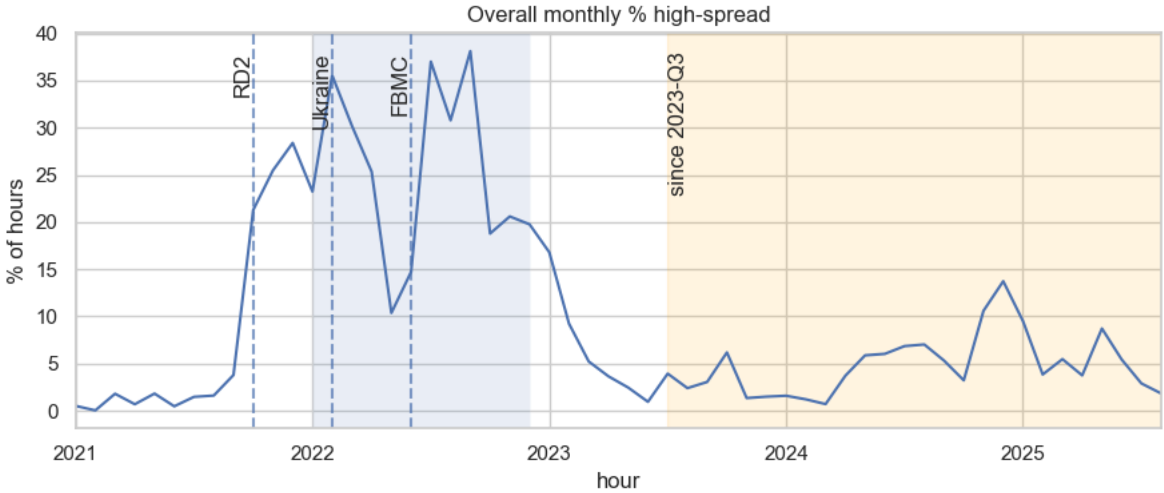

Figure 5 - Monthly share of High-Spread Hours (%)

4.6.4 Implications for Modelling

. These findings have clear modelling implications:

- Model stability and transferability may be compromised across regimes
- Feature behaviour (e.g. wind impact on spreads) may vary by regime
- Evaluating predictive performance across out-of-regime periods is essential

Therefore, in Chapter 5, the modelling framework accounts for regime sensitivity through:

- Feature scaling or selection tailored to stable periods
- Rolling-window validation with crisis-excluded training sets
- Explicit tagging of events for potential use in regime-aware models

Chapter 5 – Methodology

5.1 Overview

This chapter outlines the modelling approaches used to forecast congestion-related price spreads in the German power market and to evaluate their application in day-ahead trading strategies. All models are designed to be *day-ahead safe*, relying strictly on information available before gate closure to ensure realistic, operationally feasible signals.

Three main strategies are developed. First, a persistence-based baseline uses the previous day's same-hour spread, applying seasonal thresholds and volatility filters to trigger trades. This provides a simple, interpretable benchmark.

Second, quantile regression models are fitted to predict conditional spread quantiles (Q25, Q50, Q75) using calendar features, lagged spreads, and z-scored day-ahead forecasts for wind and solar. Trades are placed only when zero lies outside the predicted interquartile range. A seasonal variant fits separate models for winter and summer, reflecting spread asymmetries observed in the exploratory analysis.

Finally, a residual-load gated strategy targets only the top 5% of days by forecasted residual load—where system stress, and thus congestion, is most likely. Quantile models are trained and applied within this regime to focus on high-opportunity periods.

All strategies are evaluated using out-of-sample testing, with leakage control and realistic trading assumptions, including transaction costs and size caps. The remainder of this chapter details each method, including data preparation, model fitting, validation, and trading logic.

5.2 Persistence-Based Strategy

This strategy serves as a deliberately simple baseline. It trades on the assumption that large price spreads in each hour are likely to persist into the same hour the following day. While it lacks any formal statistical modelling, it provides a useful reference point for evaluating the added value of more complex methods.

5.2.1 Strategy Logic

The signal is based solely on the previous day's same-hour spread. A trade is entered if:

- The absolute value of the lagged spread exceeds a season-specific threshold.
- The recent volatility (30-day rolling standard deviation) is below a cap.

If both conditions are met, a position is entered in the same direction as the lagged spread. This is applied per hour, per border.

The strategy uses:

- Fixed or volatility-adjusted sizing, capped at 2.0 MW;
- Seasonal allow-lists to restrict trading to borders where persistence appears stronger;
- An optional peak-hour filter (08:00–20:00 UTC) to avoid low-activity periods.

5.2.2 Role in Project

There is nothing sophisticated about this approach. It requires no modelling, no forecasts, and minimal computation. Its purpose is purely comparative — to act as a lower bound on performance for more advanced strategies.

If a quantile-based model cannot meaningfully outperform this rule-of-thumb strategy (on out-of-sample PnL or risk-adjusted return), then its added complexity may not be justified.

5.3 Quantile Regression Strategy

The quantile regression (QR) strategy forms the core modelling approach in this project. It aims to estimate conditional quantiles of the day-ahead spread between Germany and its neighbouring markets using only inputs available at the time of day-ahead market closure. By generating probabilistic forecasts rather than single-point estimates, the model provides both directional insight and a measure of forecast uncertainty. This structure naturally supports the construction of disciplined, threshold-based trading rules.

All models in this family are trained per border and per hour, and share the same architecture, input features, and trading logic. Two versions are implemented: a rolling walk-forward model, which adapts to evolving dynamics over time, and a seasonal model, which isolates winter and summer patterns using fixed train/test windows.

5.3.1 General Modelling Setup

The target variable is the hourly spread between the German day-ahead price and that of a given neighbouring country (e.g., DE–FR, DE–AT). The model predicts the 25th, 50th, and 75th percentiles of the conditional distribution, denoted Q25, Q50, and Q75, respectively.

All inputs are strictly day-ahead safe, with no use of intraday prices or realised generation. The feature set includes:

- Lagged Spread: Same-hour spread from the previous day.
- Forecasted RES Generation:
  - Wind and solar forecasts (z-scored over a rolling 30 day window)
  - Hourly ramps (first differences)
- Calendar Effects:
  - Hour of day
  - Day of week
  - Month

Categorical features are one-hot encoded, and the feature table is assembled per (border, hour) pair. Data is filtered to ensure label timestamps align with delivery hour, maintaining leakage control.

The core model is an ElasticNet quantile regression, selected for its simplicity, speed, and regularisation control. In all cases, quantile crossing is avoided via post-prediction checks.

Forecast accuracy is evaluated using:

- Interquartile Coverage (proportion of realised spreads within [Q25, Q75])
- CRPS (continuous ranked probability score)
- Directional hit rate (optional, where applicable)

5.3.2 Rolling Walk-Forward Models

The first version employs a walk-forward validation scheme, designed to test performance under evolving market conditions. For each border-hour combination, models are trained using a trailing window and evaluated on a holdout period, with periodic refitting:

- Target window: 630 days

- Test Window: 60 days
- Step: 120 days
- Minimum training size: 2,160 hours (= 90 days)

This setup ensures all predictions are made out-of-sample and that the model regularly adapts to changing spread dynamics. Each fold is evaluated independently, and diagnostics are aggregated across borders and time.

This version is used as the default quantile strategy in the results chapter and serves as the main comparison point for the persistence baseline and the residual-load gated model.

5.3.3 Seasonal Models

The second variant uses fixed seasonal windows, motivated by patterns observed during the exploratory analysis. Spreads and congestion dynamics differ meaningfully between summer and winter, suggesting that season-specific models may generalise better.

Separate models are trained for each season using the following windows:

- Training: 1st July 2023 – 30th June 2024 (covers one summer and one winter)
- Testing: 1st July 2024 – 30th June 2025 (same seasonal cycle)

A separate quantile model is trained per season and per border. Inputs, model architecture, and trading rules are identical to the rolling version. Evaluation is performed on a per-season basis, enabling comparisons of forecast quality and trading performance under different seasonal regimes.

While the seasonal model sacrifices temporal adaptiveness, it offers improved interpretability and may help capture more stable seasonal signals in spreads and renewable output.

5.3.4 Trading Logic

Both quantile strategies use the same carry-trade rule, applied per hour and per border. A trade is entered only when:

- Zero is outside the predicted interquartile range, i.e. $0 \notin [Q25, Q75]$
- IQR $\geq$ minimum threshold (default: 0.25)

When these conditions are met:

- Trade direction is given by the sign of Q50

- Trade size is scaled as |Q50| / IQR, capped at a maximum of 2.0 MW
- A round-trip transaction cost of €0.75/MWh is deducted

This rule aims to avoid low-conviction trades and concentrate exposure on scenarios where the model forecasts a clear and significant directional spread.

5.4 Residual-Load Gated Strategy

This final strategy narrows its focus to periods of likely transmission congestion by conditioning both model training and execution on extreme system load conditions. The approach is based on the hypothesis that congestion-induced price spreads are more predictable—and more tradable—during periods of high residual load, when the system is under stress and flexibility is constrained.

The methodology uses a quantile regression framework similar to Section 5.3, but is selectively gated to a small subset of high-risk hours. This structure limits model complexity while increasing the relevance of training data.

5.4.1 Motivation and Gating Mechanism

Residual load, defined as:

$$Residual\ Load\ (DA) = Load\ Forecast\ (DA) - Wind\ Forecast\ (DA) - Solar\ Forecast\ (DA)$$

Serves as a proxy for system tightness. High residual load reflects a situation where demand must be met largely from dispatchable (often thermal) sources, increasing the likelihood of internal congestion and price separation between bidding zones.

5.4.2 Data and Features

The model is applied to three interconnectors where congestion effects are known to be more pronounced: DE–AT, DE–CZ, and DE–PL.

Features are limited to day-ahead safe inputs:

- Lagged Spread: Previous day's same hour spread.
- Residual Load (DA): Level, hourly ramp, and 30-day z-score.
- Calendar Features: Hour of day, day of week, month (one-hot encoded)

Only observations flagged as "extreme residual load hours" are included. This ensures the model is trained specifically on the regime it is intended to operate in.

5.4.3 Model and Evaluation

Each border is modelled independently using ElasticNet quantile regression, following the same procedure as in Section 5.3. Fixed seasonal train/ test windows are used:

- Training: 1st July 2023 – 30th June 2024
- Testing: 1st July 2024 – 30th June 2025

The same trading rule is used:

- Enter if $0 \notin [Q25, Q75]$ and IQR $\geqslant$ threshold
- Direction: sign(Q50)
- Size: |Q50| / IQR, capped at 2.0 MW
- Cost: €0.75/MWh round-trip

Performance is assessed only on gated test days, meaning results represent returns from a low-frequency, high-conviction strategy. This setup implicitly reduces exposure while concentrating trades into periods of higher congestion probability.

5.4.4 Role in the Broader Framework

This strategy is not intended to maximise time coverage or trading volume. Rather, it explores the idea that *when* you trade may be as important as *how*. By conditioning the model on a known structural stress indicator, the method aims to produce a sparse but high-quality signal.

Compared to the baseline and general quantile models, this strategy tests whether statistical learning improves when noise is reduced via regime selection. It also provides a framework for potential future extensions using other gating mechanisms (e.g. forecast errors, temperature extremes, or renewable curtailments).

5.5 Forecast Evaluation and Trading Simulation

All models and strategies are evaluated using strictly out-of-sample data and realistic market assumptions. Forecasts are aligned to the day-ahead market timeline, with labels corresponding to the hourly delivery period and all features restricted to data available before market closure.

Forecast accuracy is assessed using two standard metrics:

- Interquartile Coverage: the proportion of realised spreads falling within the predicted [Q25, Q75] interval. This is expected to be close to 50% if the model is well-calibrated.
- Continuous Ranked Probability Score (CRPS): a proper scoring rule that rewards sharp, accurate probabilistic forecasts.

For trading performance, all models are mapped to hourly carry positions using a consistent rule: enter only when zero lies outside the predicted interquartile range and the IQR exceeds a predefined threshold. The trade direction is given by the sign of the median prediction (Q50), and position size is scaled by |Q50| / IQR, capped at 2.0 MW. A round-trip transaction cost of €0.75/MWh is applied to all positions.

Performance is reported in terms of:

- Net PnL: spread minus transaction cost, per MWh
- Hit Rate: percentage of trades with the correct directional sign
- Sharpe Ratio: mean return over realised volatility, annualised where appropriate
- Trade Frequency: number of trades per day or per regime

This framework allows for direct comparison across strategies of varying complexity and coverage. The persistence baseline sets a lower bound on performance, while the quantile models and residual-load gated strategy are evaluated on both predictive and financial grounds. All results are presented in the following chapter.

Chapter 6 – Results

6.1 Evaluation Setup

All results are strictly out-of-sample and aligned to the day-ahead timetable. Models are trained and evaluated on DE–X hourly spreads constructed from coupled day-ahead prices; labels are aligned to delivery hours to avoid look-ahead bias. Inputs for model-based strategies are "DA-safe" (calendar features, previous-day same-hour spread, and DE day-ahead renewables/load forecasts), with no intraday or realised data used as predictors.

Three evaluation regimes are used. First, a rolling walk-forward with refits on a moving window (train ≈ 18 months, test ≈ 1 month, stepped monthly) to reflect operational deployment. Secondly, fixed seasonal models with a one-year train (July 2023–June 2024) and one-year test (July 2024–June 2025) to examine regime differences between winter and summer. Thirdly, a residual-load-gated variant that restricts testing to "extreme" days defined on the training year.

Trading assumptions are common across strategies: an order is entered only when 0 lies outside the interquartile band $[Q_{25}, Q_{75}]$, with a minimum IQR filter, portion size min $(2.0, Q_{50}|/IQR)$, and a round-trip cost of €0.75/MWh. Baseline persistence uses the same cost and size cap. Performance is reported as net €/MWh, hit rate, trade count, and cumulative PnL; probabilistic calibration is assessed via IQR coverage (target 50%) and CRPS. Confidence intervals use heteroskedasticity- and autocorrelation-consistent (Newey–West) errors appropriate for hourly data.

6.2 Baseline: Persistence Carry

6.2.1 Purpose and intuition

This benchmark exploits short-horizon continuation: if the previous day's same-hour DE–X spread was strongly positive (negative), trade the same direction today—*but only* when the prior move was large and recent volatility was modest.

6.2.2 Implementation

Hourly DE-X spreads were constructed and applied for season/ border, restrict to peak hours, and gate on:

i) Absolute lag-24h ≥ threshold (summer: 5€/MWh, winter: 8€/MWh)

ii) 30-day rolling σ ≤ cap (summer: 25 €/MWh; winter: 18 €/MWh)

iii) Sizing is either constant or volatility-scaled

6.2.3 Evaluation

Out-of-sample testing uses 1 July 2023–30 June 2025 day-ahead data. We report net €/MWh, hit rate, trade count, and cumulative PnL by border. Confidence intervals use heteroskedasticity- and autocorrelation-consistent (Newey–West) errors suitable for hourly series (lags ≈ 24).

6.2.4 Results

Portfolio averages €8.03/MWh (95% CI [6.67, 9.39]), hit rate 0.620, n = 9,702 hours as shown in Table 9 below. Across borders, DE–FR shows the highest margin (€10.77/MWh, hit 0.561, *n* = 1,241), i.e., fewer but larger wins. DE–BE is similarly strong (€10.56/MWh) with the highest hit rate (0.650) and the most trades (*n* = 3,970), so it drives much of the portfolio. DE–CH is positive but smaller (€6.90, hit 0.589, *n* = 1,801). DE–NL is the weakest (€3.80, hit 0.624, *n* = 2,690) yet remains statistically above zero. All CIs exclude zero, so every border is profitable after costs; the portfolio average (€8.03/MWh) sits below the best single borders, as expected given inclusion of lower-margin exposures.

Table 9 - Baseline persistence carry: performance by border

Border	avg_eur_mwh	ci_lo	ci_hi	hitrate	n
DE-FR	10.77297	6.496685	15.04926	0.560838	1241
DE-BE	10.55637	8.739435	12.37331	0.649874	3970
PORTFOLIO	8.030739	6.667377	9.3941	0.619872	9702
DE-CH	6.895576	3.158281	10.63287	0.588562	1801
DE-NL	3.798231	2.42878	5.167682	0.623792	2690

As shown from Figure 6 The curve climbs steadily through mid-2024, is broadly flat over winter 2024/25, then resumes modest gains into spring 2025, with only shallow pull-backs—consistent with reduced winter exposure and the volatility/cost gates.

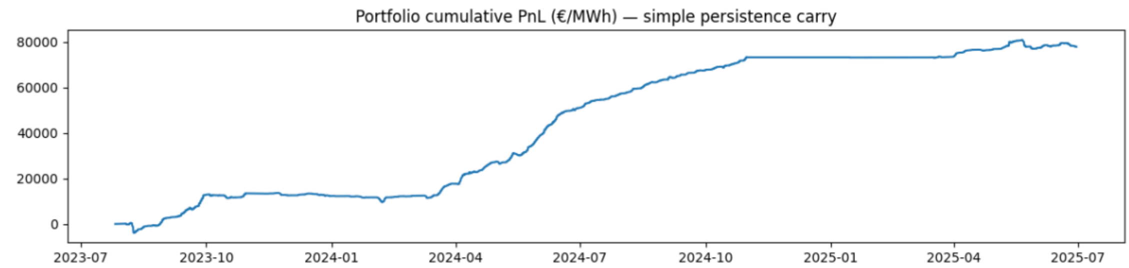

Figure 6 - Portfolio cumulative PnL simple persistence carry

6.3 Quantile Regression – Rolling & Seasonal

6.3.1 Set-up and entry rule

Elastic-net quantile regression per border/hour produces $(, Q_{50}, Q_{75})$ from DA-safe inputs (DE wind/ solar forecasts incl. ramps, previous-day same-hour spread, calendar). Trade only if $0 \notin [Q_{25}, Q_{75}]$ and $IQR \geq IQR_{min}$; $size = \min(2MW, |Q_{50}|\ IQR)$; cost €0.75/MWh.

6.3.2 Rolling walk-forward

Per-border elastic-net quantile models were fitted on a rolling scheme (630-day training, 60-day test, 120-day step), using only day-ahead-safe inputs; labels were aligned to delivery hours. Orders followed the rule in 6.3.1 and results were reported after a €0.75/MWh round-trip cost.

Out-of-sample calibration was close to target, with interquartile coverage of 0.531 and CRPS of 5.799. Forecast performance was strongly border-dependent: DE–FR achieved €30.21/MWh net (95% CI [26.57, 33.85]) with 72% hit rate, while most other borders performed near-zero or negative after costs. The full portfolio averaged €1.55/MWh (CI [0.35, 2.75]) with 12,031 hours traded and a modest overall hit rate of 0.235.

These results suggest that while quantile models produced some well-calibrated and high-return signals—especially on DE–FR—robust performance was limited to a few borders. Figures Figure 7 and Figure 8 show the calibration and distribution of profitability; Table 10 presents detailed border-level results.

Table 10 - Quantile regression rolling results by border

Border	avg_eur_mwh	ci_lo	ci_hi	hitrate	n_hours
DE-FR	30.21251	26.57464	33.85039	0.724526	1688
DE-BE	1.28282	0.634168	1.931471	0.309565	1725
DE-NL	0.150257	-0.042708	0.343223	0.094406	1716
DE-CZ	-0.624987	-1.185751	-0.064222	0.085217	1725
DE-AT	-1.659099	-2.716034	-0.602165	0.10989	1729
DE-PL	-8.642492	-10.65548	-6.629506	0.171793	1723
DE-CH	-9.268206	-13.03981	-5.496603	0.162319	1725
PORTFOLIO	1.549669	0.350261	2.749077	0.235392	12031

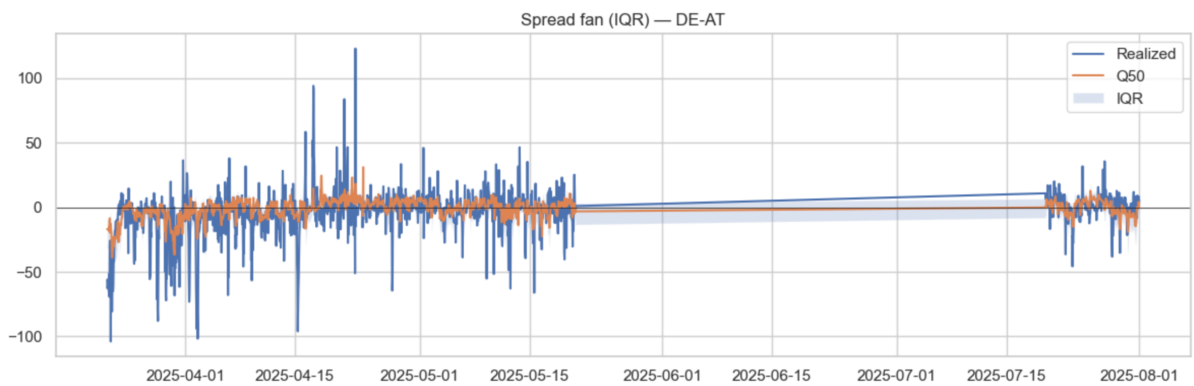

Figure 7 - Fan chart example for DE-AT

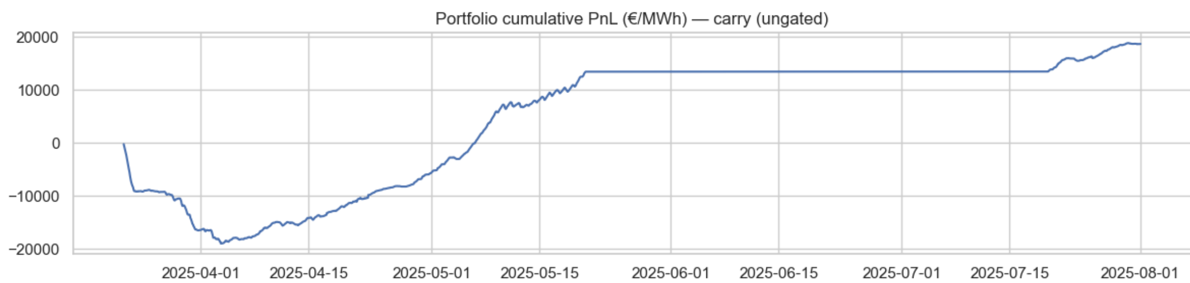

Figure 8 - Cumulative PnL - Quantile Regression Carry

6.3.3 Seasonal Models

Separate quantile regression models were trained for winter (November–March) and summer (April–October) periods, using fixed estimation windows. The training set covered 1 July 2023 to 30

June 2024, and the test set covered 1 July 2024 to 30 June 2025. The same entry and sizing rule described in 6.3.1 was applied, with season-aware sizing caps and thresholds. Certain borders, including DE–PL and DE–CZ, were excluded due to persistent underperformance.

In-sample interquartile coverage was close to the nominal 50% level across borders. Out-of-sample coverage declined to 0.356 in winter and 0.437 in summer, with corresponding CRPS values of 7.805 and 6.793, respectively. Example fan charts for DE–AT are shown Figure 9 and Figure 10. The summer forecasts demonstrated narrower and more directional interquartile ranges, while winter forecasts were less confident and more dispersed.

Over the full 12-month test period, the portfolio generated an average of €6.15/MWh (Newey–West 95% CI: [5.56, 6.75]), with a hit rate of 0.237 across 36,202 traded hours. At the border level:

- DE–FR achieved the highest profitability (€18.40/MWh, CI [16.40, 20.41], hit rate 0.428, *n* = 8,590)
- DE–CH also performed well (€9.21/MWh, CI [6.94, 11.47])
- DE–BE remained positive but modest (€2.27/MWh, CI [1.52, 3.01])
- DE–NL and DE–AT were close to break-even
- DE–PL and DE–CZ were excluded from the seasonal strategy.

When disaggregated by season, results showed a clear regime contrast:

- Summer: €7.75/MWh (CI [7.12, 8.37]; hit rate 0.294; *n* = 25,377)
- Winter: €2.42/MWh (CI [1.23, 3.60]; hit rate 0.102; *n* = 10,825)

Figure 11 shows the cumulative portfolio PnL over the test year. Strong gains were realised during summer, followed by a plateau over the winter period, and resumed performance in spring 2025.

Table 11 - Seasonal Model Performance by Border

border	avg_eur_mwh	ci_lo	ci_hi	hitrate	n_hours
DE-FR	18.40432	16.39621	20.41244	0.427707	8590
DE-CH	9.20695	6.943322	11.47058	0.309762	5091
DE-BE	2.265041	1.518592	3.01149	0.222961	8719

DE-NL	-0.130911	-0.406685	0.144863	0.079047	8691
DE-AT	-0.156162	-0.720375	0.408051	0.133046	5111
PORTFOLIO	6.153768	5.560309	6.747228	0.236506	36202

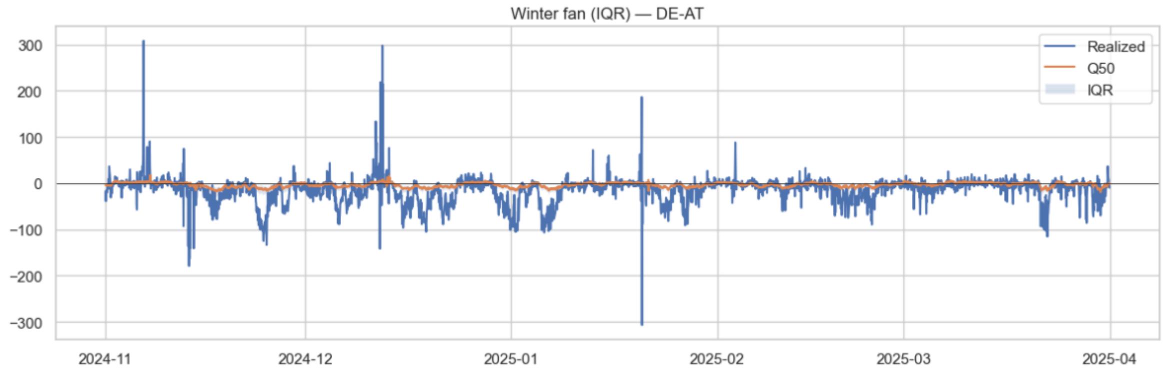

Figure 9 - Winter fan (IQR) - DE-AT

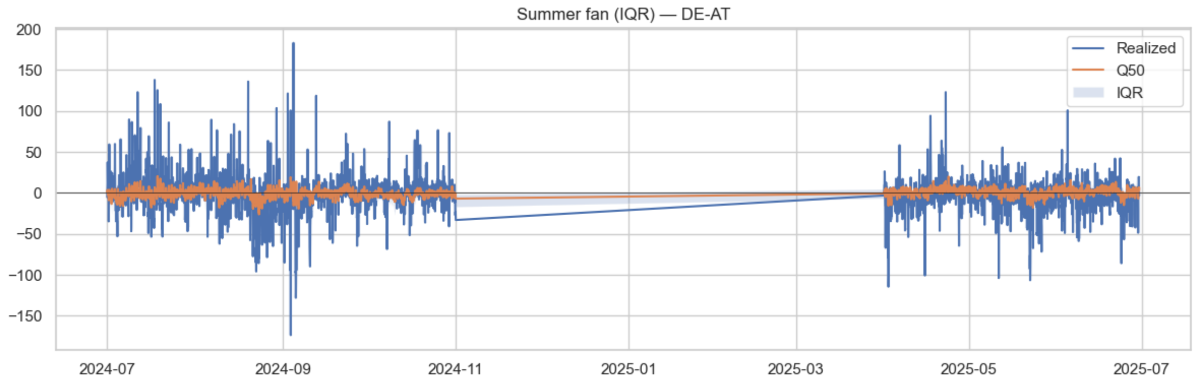

Figure 10 - Summer fan (IQR) - DE - AT

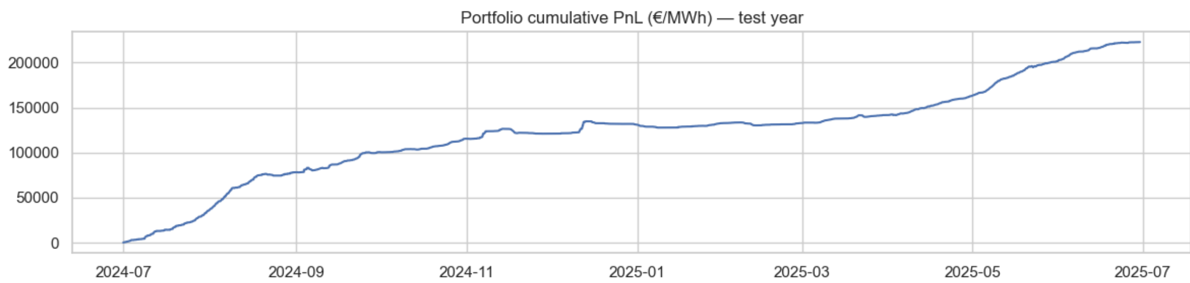

Figure 11 - Cumulative Portfolio PnL (€/MWh)

6.4 Residual Load Gated Strategy

Trading was restricted to test hours falling within the top 5% of residual-load days, as defined over the training period. Out-of-sample performance on this gated subset is summarised in Table 6.

The DE–PL border yielded the strongest performance, with an average return of €12.56/MWh (Newey–West 95% CI: [0.52, 24.61]; hit rate: 0.29) over 1,079 hours. Moderate returns were observed on DE–AT (€2.32/MWh) and DE–CZ (€1.63/MWh), though both included confidence intervals overlapping zero. The portfolio, averaging across all three borders, achieved a net return of €5.50/MWh (CI: [0.96, 10.05]) across 3,238 traded hours, with a hit rate of 0.23.

Cumulative PnL over the gated period is shown in Figure 12. The stepwise nature reflects the sparse exposure inherent in the gating mechanism. Several high-return days account for a disproportionate share of cumulative performance, particularly in Q4 2024 and Q1 2025.

Example fan charts are provided in the figures below. These illustrate high-conviction forecasts coinciding with large, realised spreads under extreme system conditions. In DE–PL, in particular, multiple high-magnitude events were captured where the predictive interquartile range remained directional and narrow.

These results support the hypothesis that gating by residual load can concentrate trading into more predictable, congestion-prone hours—improving risk–return characteristics despite reduced exposure.

Table 12 - Performance on residual-load gated test hours

Border	Avg €/MWh	CI Lo	CI Hi	Hit rate	n
DE–PL	12.56	0.52	24.61	0.29	1,079
DE–AT	2.32	–0.49	5.14	0.219	1,080
DE–CZ	1.63	–0.16	3.42	0.17	1,079
Portfolio	5.5	0.96	10.05	0.226	3,238

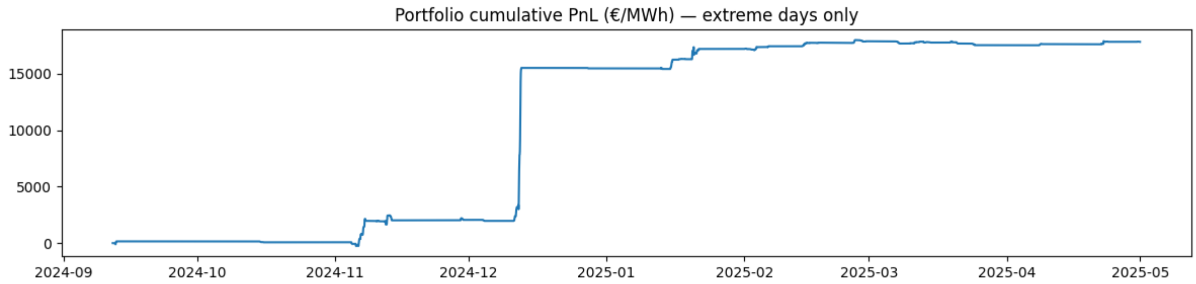

Figure 12 - Portfolio cumulative PnL (€/MWh)

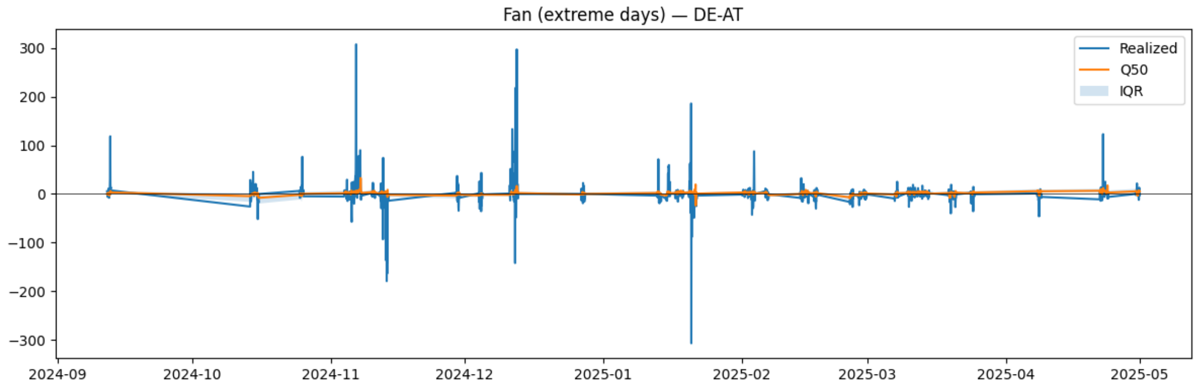

Figure 13 - Fan plot - DE-AT

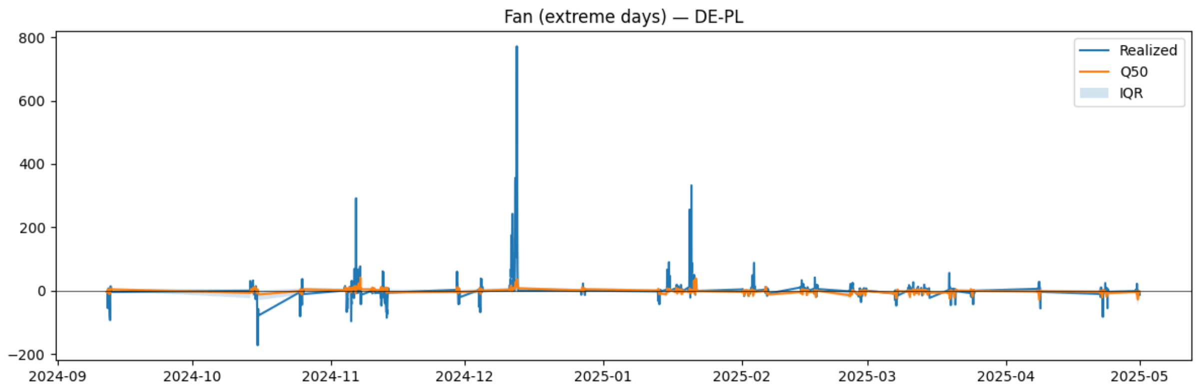

Figure 14 - Fan plot DE-PL

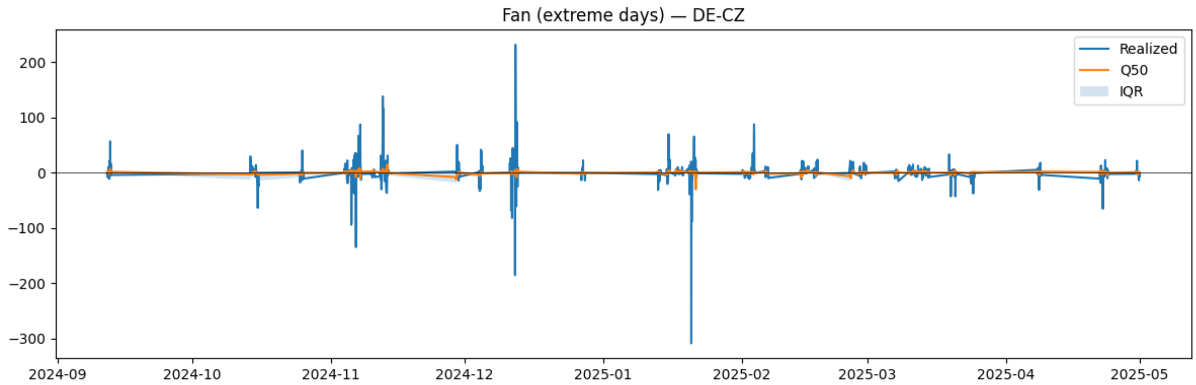

Figure 15 - Fan plot DE-CZ

6.5 Summary of Results

Table 13 provides a consolidated comparison across all strategies evaluated in this chapter. Results reflect strictly out-of-sample performance, using the common trading rule defined in Section 6.1 and a uniform €0.75/MWh round-trip cost.

The simple persistence strategy offered a strong baseline with an average portfolio return of €8.03/MWh, driven by high-frequency, short-horizon continuation signals—particularly on DE–FR and DE–BE.

The rolling quantile model demonstrated selective trading with good probabilistic calibration but inconsistent border-level performance. While DE–FR returned €30.21/MWh, the overall portfolio return was low at €1.55/MWh, due to weak signals on most other borders.

The seasonal quantile model performed more consistently. By isolating summer and winter regimes, it achieved a test-year portfolio return of €6.15/MWh, with stronger results in summer (€7.75/MWh) and a notable contribution from DE–FR and DE–CH.

The residual-load gated strategy further reduced exposure but increased risk-adjusted return concentration. Despite trading only ~3,200 hours, it yielded €5.50/MWh, with DE–PL capturing the clearest congestion signal (€12.56/MWh).

These results support two conclusions:

1. Statistical models can outperform naive baselines on selected borders—but require careful gating to avoid overtrading low-confidence signals.

2. Conditioning on structural regimes (seasonality or residual load) improves both economic performance and forecast sharpness, particularly in constrained system states.

Table 13 - Strategy comparison

Strategy	Avg €/MWh	95% CI	Hit rate	n hours
Persistence (baseline)	8.03	[6.67, 9.39]	0.62	9,702
Rolling quantile	1.55	[0.35, 2.75]	0.235	12,031
Seasonal quantile (all)	6.15	[5.56, 6.75]	0.237	36,202
‣ Summer only	7.75	[7.12, 8.37]	0.294	25,377
‣ Winter only	2.42	[1.23, 3.60]	0.102	10,825
Residual-load gated	5.5	[0.96, 10.05]	0.226	3,238

Chapter 7 – Discussion

7.1 Overview

This chapter reflects on the empirical results presented in Chapter 6, drawing broader conclusions about the design and effectiveness of congestion forecasting methods in the European day-ahead electricity market. Each strategy developed in the study—ranging from simple persistence carry to seasonally segmented and residual-load gated quantile regressions—produced materially different outcomes, despite sharing a common execution rule and evaluation framework. The discussion here is structured to move beyond isolated performance metrics, focusing instead on underlying mechanisms, structural insights, and the interaction between predictive power, regime design, and economic returns.

The chapter begins by synthesising the key findings across all strategies, including observed differences between borders and regimes. It then considers why certain strategies outperformed others, exploring the role of structure, stress, and statistical calibration. These reflections are grounded in both domain-specific intuition (e.g., congestion under high residual load) and relevant literature on electricity forecasting and trading.

The discussion further examines the relative value of complexity in forecasting models, assessing whether additional statistical sophistication yields proportional improvements in economic outcomes. This includes a review of the trade-off between simple, interpretable heuristics and more complex machine learning approaches, particularly in light of deployment speed, robustness, and operational constraints.

Finally, the chapter outlines practical implications for strategy design, before identifying key limitations of the present study and setting the stage for Chapter 8.

7.2 Synthesis of Key Findings

This section draws together the outcomes from Chapter 6 to assess how different strategies performed, identify consistent patterns, and distil broader insights for trading European power spreads.

- Persistence Strategy: Delivered a robust baseline return (approx. €8/MWh), especially on strong interconnectors like DE–FR and DE–BE. This underscores the role of structural inertia and regime persistence in congestion dynamics. For example, (Newbery, 2004) notes that transmission congestion tends to fragment markets in predictable ways, reinforcing short-term autocorrelation in price spreads.
- Rolling quantile regression: Statistically well-calibrated but generated modest average return (€1.55/MWh), with profitability concentrated on DE–FR. This outcome illustrates that even well-calibrated models may underperform when applied indiscriminately without regime control—an insight supported by literature suggesting that probabilistic models must be combined with selective execution filters to yield economic value.
- Seasonal Quantile Models: Improved economic performance (approx. €6.15/MWh), particularly in summer. Regime-based segmentation enhanced signal sharpness and economic returns—affirming that temporal context (such as seasonality) is critical in energy markets. Forecasting research consistently endorses regime-aware models for seasonal markets.
- Residual-load Gated Strategy: Applied to a limited set of hours, achieved high returns (€5.50/MWh), with standout performance on DE–PL (12.56 €/MWh). This validates the hypothesis that structural stress indicators (e.g. high residual load) provide a useful lens to isolate high-confidence trading windows.

7.3 Complexity vs Simplicity in Trading Strategies

This subsection examines the broader trade-off between sophisticated machine learning models and simpler, more interpretable strategies—particularly relevant for commodity traders and quant practitioners.

7.3.1 Academic Perspectives on Simplicity

Industrial practitioners often invoke Ockham's razor in quantitative modelling: they avoid adding unnecessary complexity unless it demonstrably adds value. Kristian Bondo Hansen's study found that quant traders frequently scale back machine learning models to preserve interpretability and avoid over-parameterisation. (Hansen, 2020)

Moreover, (Cliff & Rollins, 2020) demonstrated that some simple heuristic strategies outperform public-domain AI/ML strategies in trading benchmarks, suggesting that complexity is not a guarantee of performance.

7.3.2 Industry Practice and Trade-offs

Leading commodity trading firms like Vitol, Trafigura, and Citadel are investing heavily in data infrastructure and AI, but with a dual focus: improving both operational efficiency and trading edge. Russell Hardy of Vitol described this as an "arms race" requiring infrastructure investment, not purely model sophistication. (Kumar, 2025)

High-frequency trading (HFT) firms focus heavily on speed and execution quality rather than model complexity. The "mundane economics" of HFT emphasise that many edges come from reacting faster to well-known signals, not uncovering exotic patterns. (Steer, 2025)

7.3.3 Implications for Power Spread Strategies

Applying these insights:

- Simple persistence rules: Captured meaningful value with minimal complexity and zero overfitting risk. They provide a strong baseline.
- Quantile models: Added probabilistic nuance but only paid off when disciplined by regime gating (seasonal or residual-load). Blind complexity is ineffective.
- Speed and operational readiness: Simple heuristics are faster to deploy, easier to monitor, and more transparent—raising confidence in live trading.
- Structural filters (like extreme residual load): Are powerful, interpretable levers that can be embedded alongside models to boost PnL without massive complexity increases.

7.4 Literature Integration

This section situated the empirical insights of Chapter 6 within the broader academic and practitioner literature, highlighting points of convergence, divergence, and novel contribution.

7.4.1 Simplicity over Complexity

A recurrent theme in quantitative trading is the preference for simpler models, especially when operational robustness is critical. (Hansen, 2020) argues that simpler models are less prone to errors and more interpretable – a principle he terms "the virtue of simplicity" in algorithmic trading. Similarly, a recent practitioner-oriented review emphasises that "simple trading strategies with few variables or parameters trump complex strategies," citing risks of overfitting and fragility. (Groette, 2025)

7.4.2 Complexity with Discipline

However, academic debate around the benefit of complexity remains nuanced. (Financial Times, 2025) found that highly complex machine learning models could outperform simpler ones under certain conditions, particularly through the phenomenon of double descent—a state where adding parameters improves generalisation beyond traditional bias–variance trade-off limits. Critics have argued, however, that such performance may reflect momentum proxies rather than true forecast innovation.

7.4.3 Interpretability and Explainability

In energy applications, model interpretability is especially important for stakeholders. A 2025 review of machine learning in energy systems highlights key barriers around model interpretability, real-time decision-making, and scalability. (Harbourfront Technologies, 2025)

The gated strategies used—particularly residual-load gating—enhanced transparency and decision-makers' confidence, aligning with calls for explainable AI (XAI) frameworks.

7.4.4 Summary

- Simplicity yields robustness as seen in the persistence strategy.
- Complexity can add value, but only when paired with regime-specific decision logic; complexity without structure underperformed.
- Explainability and a modular design offer transparency and deployment advantage.
- Industry trends align with the insights found in this project, power traders demand speed, scale and interpretability – not complexity.

7.5 Practical Implications

The results presented in this study offered several practical implications for the design, deployment, and management of congestion-based trading strategies in the European power market. Although the focus was methodological, the performance outcomes and regime-specific behaviours provided valuable insights for both commercial trading desks and system analysts.

7.5.1 Value of Simplicity in Operational Settings

The persistence carry strategy delivered consistent returns using only lagged spreads and simple gating conditions. This reinforced the practical advantage of using low-complexity, transparent models in operational contexts. Such strategies required minimal computation, were easy to audit, and could be implemented rapidly—factors that were essential in time-sensitive environments such as day-ahead trading. These findings were particularly relevant for firms with limited data infrastructure or where regulatory environments demanded explainable decision processes.

Furthermore, the robustness of the persistence benchmark on certain borders (e.g., DE–FR) suggested that a baseline of edge existed structurally, without requiring advanced learning models. As such, simple heuristics remained a useful starting point or fallback strategy in live operations.

7.5.2 Risk of Overtrading in Model-Driven Strategies

The rolling quantile regression models, despite being statistically well-calibrated, underperformed at the portfolio level. This highlighted a critical operational risk: without regime-specific gating, statistical models could generate false positives or low-confidence signals that diluted profitability. Traders implementing such models without sufficient filtering would have faced unnecessary transaction costs, degraded risk-adjusted returns, and reduced conviction in the strategy.

This underlined the importance of model governance frameworks that included economic back-testing, structural validation, and well-defined intervention rules—particularly when models were deployed in automated execution systems.

7.5.3 Integration of Structural Regime Filters

The two most effective strategies—seasonal segmentation and residual-load gating—achieved higher economic performance not through more complex model architecture, but through selective exposure to known structural regimes. This suggested that when to trade was often more important than how to trade. Regime gating acted as a risk management layer, aligning statistical forecasts with physical system constraints (e.g., seasonal load patterns, internal congestion during high residual load).

For market participants, these findings supported the integration of domain-informed structural filters into trading models, particularly in contexts where congestion events were episodic and regime-dependent.

Chapter 8 – Strength & Limitations

This chapter critically assesses the methodological and empirical strengths of the study, while also acknowledging key limitations in scope, data, and evaluation. The aim is to provide a balanced view of the reliability and applicability of the results presented, as well as to guide appropriate interpretation and future development.

8.1 Strengths

8.1.1 Probabilistic Framework with Regime Control

A central strength of the study was the use of probabilistic forecasting methods—specifically quantile regression—to generate spread predictions with embedded uncertainty estimates. This approach allowed for explicit control over risk exposure via entry and sizing rules based on predictive distributions, rather than relying on point forecasts alone. By gating trades using interquartile coverage, the strategy filtered out low-confidence signals in a principled manner.

In addition, the introduction of regime-specific filters—such as seasonal segmentation and residual-load gating—enhanced the signal quality without increasing model complexity. These filters were derived from interpretable, system-relevant variables, providing a transparent and justifiable mechanism to improve economic outcomes. This modularity in strategy design allowed for structural interpretability alongside statistical learning.

8.1.2 Uniform Evaluation and Execution Framework

All strategies were implemented within a consistent execution and evaluation framework. Entry rules, trade sizing, and transaction cost assumptions were held constant across models, ensuring fair and directly comparable performance. This uniformity allowed differences in economic results to be attributed to model structure and gating mechanisms, rather than implementation artefacts.

Moreover, the inclusion of Newey–West standard errors accounted for heteroskedasticity and autocorrelation in hourly spread data, improving the robustness of statistical inference around average returns. Confidence intervals were reported for all performance metrics, improving the transparency and interpretability of reported results.

8.1.3 Realism and Deployability

The study used only day-ahead-safe inputs, including official forecast values and lagged spreads, ensuring that all models could be implemented in practice without hindsight bias. Execution rules were simple, rule-based, and risk-controlled, aligning with real-world constraints on trade sizing and transaction costs. The framework was therefore deployment-ready in principle and could be adapted to operational systems with minimal adjustment.

8.2 Limitations

8.2.1 Market Scope and Temporal Generalisability

The study was limited to the day-ahead market and focused exclusively on DE–X borders. Although this scope was sufficient to demonstrate the viability of congestion forecasting strategies, the results may not generalise to other temporal scales (e.g., intraday or real-time markets) or interconnectors with different congestion characteristics.

Furthermore, the back test was conducted over a two-year period (2023–2025), which may not capture longer-term structural changes in grid topology, market design (e.g. flow-based allocation rules), or renewable penetration. As a result, the observed profitability and model effectiveness may not persist under changing system conditions.

8.2.2 Data Access and Coverage Constraints

The study was limited to publicly available day-ahead data, including a single daily forecast and hourly price spreads. While this ensured the approach was transparent and reproducible, it excluded higher-frequency and commercial data sources that would likely have improved model accuracy and responsiveness.

Intraday market data—such as continuous trades, updated forecasts, and imbalance prices—was not used, largely due to API access costs (e.g., via EPEX SPOT). This prevented exploration of shorter-term strategies or real-time signal refinement.

Only a single forecast run per day was available, whereas many operational systems make use of multiple forecast updates, including ensemble runs and revised weather inputs. These could have been used to track forecast deltas and model uncertainty more dynamically, especially in volatile conditions.

Finally, several key data series from ENTSO-E were incomplete. Notably, historical cross-border capacity data was often missing, which limited the ability to model transmission constraints directly or account for flow-based allocation effects. As discussed in Chapter 4, this restricted both the feature set and the interpretability of spread behaviour in some regions.

8.3 Summary

The study presented a transparent and practical framework for forecasting cross-border electricity price spreads using probabilistic models with structural regime control. Its key strengths included the use of day-ahead-safe inputs, interpretable modelling techniques, and a consistent execution framework across strategies. However, limitations around market scope and data access, particularly the absence of intraday forecasts capacity data, should be considered in not allowing as much as exploration of potential strategies as possible.

Chapter 9 – Conclusion & Future Work

9.1 Conclusion

This study investigated the design and evaluation of trading strategies based on congestion forecasting in the European day-ahead electricity market. The analysis focused on the performance of both simple and probabilistic models for predicting cross-border spreads, using only day-ahead-safe inputs and a consistent rule-based execution framework.

Four strategies were developed and tested: a simple persistence benchmark, a rolling quantile regression model, a seasonally segmented quantile model, and a residual-load gated variant. While each strategy followed the same entry rule and cost assumptions, the results varied significantly depending on how well structural regimes were captured.

The persistence strategy produced surprisingly strong results, particularly on borders such as DE–FR and DE–BE, suggesting that short-term autocorrelation in spreads and structural constraints can be exploited even without complex modelling. In contrast, the rolling quantile regression model, though well-calibrated, generated only modest portfolio-level returns—highlighting the limitations of applying statistical models uniformly without regime control.

Seasonal segmentation improved forecast sharpness and profitability, especially in the summer months. The residual-load gated strategy, despite trading fewer hours, achieved high returns per trade, validating the hypothesis that congestion becomes more predictable during system stress. These findings collectively reinforced the importance of regime-specific filtering, which often added more value than increasing model complexity.

The study demonstrated that it is possible to build congestion-sensitive trading strategies that are both interpretable and economically viable. More broadly, it showed that combining domain knowledge with probabilistic forecasting can produce robust results using public data, even in structurally complex markets such as European electricity.

9.2 Future Work

Several avenues for future work emerged from this project. These fall into four main categories: data, models, implementation, and system relevance.

9.2.1 Data and Feature Enhancements

The study was constrained to a single day-ahead forecast per day and lacked access to intraday data streams. Future research could incorporate multiple daily forecast runs, ensemble weather models, and forecast deltas to build more responsive and uncertainty-aware signals. Access to intraday prices, imbalance data, or near real-time system indicators would also enable shorter-horizon strategies and adaptive updates. Historical capacity data, where complete, could improve modelling of transmission constraints and flow-based coupling effects.

9.2.2 Model Development

While quantile regression was chosen for its interpretability, other probabilistic methods—such as quantile gradient boosting or Bayesian forecasting approaches—could be explored for improved accuracy. There is also scope for unsupervised regime detection, where clustering or change-point detection identifies congestion regimes directly from data, rather than relying on predefined filters.

9.2.3 Real-Time Deployment and Trading Infrastructure

This study focused on historical back-testing. Future work should implement the most promising strategies in a live environment, with real-time data feeds, execution tracking, and post-trade analysis. This would allow the evaluation of operational risks, latency, and slippage, which are abstracted away in offline simulations.

9.2.4 System and Policy Applications

Beyond trading performance, the forecasts developed here could be extended to support TSO operations, market monitoring, or regulatory benchmarking. For example, accurate spread predictions could help identify recurring congestion patterns, assess the effectiveness of cross-border capacity allocation, or inform future revisions to the market design. There is also potential to link economic spread predictions with welfare metrics, bridging quantitative finance with power system planning.

References

Aiohttp. (2025, September 4). Retrieved from https://docs.aiohttp.org/en/v3.8.0/

Black. (2023, September 11). Retrieved from https://pypi.org/project/black/

Bucksteeg, M., Voswinkel, S., & Blumberg, G. (2025, August 25). *ScienceDirect*. Retrieved from Improving flow-based market coupling by integrating redispatch potential—Evidence from a large-scale model☆: https://www.sciencedirect.com/science/article/pii/S0301421524001137

Bundesnetzagentur. (2025, August 25). *Ruling Chamber - Chamber 8 - Generation Issues*. Retrieved from bundesnetzagentur: https://www.bundesnetzagentur.de/EN/RulingChambers/Chamber8/RC8_07_Generation%20issues/RC8_07_Generation%20issues.html

Cliff, D., & Rollins, M. (2020). *Methods Matter: A Trading Agent with No Intelligence Routinely Outperforms AI-Based Traders.* Ithaca: Cornell University.

Docker. (2025, September 1). Retrieved from https://www.docker.com/

Docker-Compose. (2025, July 8). Retrieved from https://docs.docker.com/compose/releases/release-notes/

Entsoe. (2025, August 25). *Detailed Data Descriptions*. Retrieved from eepublicdownloads: https://eepublicdownloads.entsoe.eu/clean-documents/Transparency/MoP_Ref2_DDD_v3r4.pdf

Entsoe. (2025, August 25). *ENTSO-E Market Report 2024*. Retrieved from eepublicdownloads: https://eepublicdownloads.entsoe.eu/clean-documents/nc-tasks/240628_ENTSO-E_Market_Report_2024.pdf

EPEX. (2025, August 25). *EUPHEMIA: Description and functioning*. Retrieved from Epexspot: https://www.epexspot.com/sites/default/files/2020-02/Euphemia_Description%20and%20functioning_1812.pdf

EUR-Lex. (2019). *Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the internal market for electricity.* Brussels: European Union.

Fact Sheets on the European Union. (2025, August 25). *Internal energy market* . Retrieved from Fact Sheets on the European Union - European Parliament: https://www.europarl.europa.eu/factsheets/en/sheet/45/internal-energy-market

Financial Times. (2025, September 4). *Are bigger AI models better than stock pickers? Maybe, but probably not*. Retrieved from Financial Times: https://www.ft.com/content/89d88cbf-a92c-43d2-b8af-88ae26529be0

Financial Times. (2025, August 25). *Silicon Valley of energy trading': the Danish companies seeking an edge in renewables*. Retrieved from Financial Times: https://www.ft.com/content/ab992f40-c6c2-45be-aa82-c6bf5b38df29?utm_source=chatgpt.com

Finck, R. (2021). Impact of Flow Based Market Coupling on the European Electricity Markets. In R. Finck, *Impact of Flow Based Market Coupling on the European Electricity Markets* (pp. 173-186).

Finck, R. (2025, August 25). *Impact of Flow Based Market Coupling on the European Electricity Markets*. Retrieved from econstor.eu: https://www.econstor.eu/bitstream/10419/286843/1/s00550-021-00520-w.pdf

Grafana. (2025, September 4). *grafana/grafana* . Retrieved from Docker: https://hub.docker.com/r/grafana/grafana

Groette, O. (2025, April 1). *Simple Vs. Complex Trading Strategies: Keep It Simple* . Retrieved from Quantified Strategies: https://www.quantifiedstrategies.com/simple-vs-complex-trading-strategies/?utm_source=chatgpt.com

Hansen, K. B. (2020). *The Virtue of Simplicity On Machine Learning Models in Algorithmic Trading.* Copenhagen: Copenhagen Business School.

Harbourfront Technologies. (2025, May 27). *Simplicity or Complexity? Rethinking Trading Models in the Age of AI and Machine Learning* . Retrieved from Medium: https://derivvaluation.medium.com/simplicity-or-complexity-rethinking-trading-models-in-the-age-of-ai-and-machine-learning-e04ed7090f38

Ipykernal. (2025, August 4). Retrieved from https://pypi.org/project/ipykernel/

jao. (2020). *Documentation of the CWE FB MC solution* . Brussels: jao.eu.

Jupyter. (2024, August 30). *Jupyter*. Retrieved from https://pypi.org/project/jupyter/

Kumar, D. K. (2025, September 4). *'Arms Race' for Ai-Driven Oil Trading is Underday, Vitol Says*. Retrieved from Bloomberg UK: https://www.bloomberg.com/news/articles/2025-03-10/-arms-race-for-ai-driven-oil-trading-is-underway-vitol-ceo?embedded-checkout=true

Lago, J., Marcjasz, G., De Schutter, B., & Weron, R. (2021). *Forecasting day-ahead electricity prices: A review of state-of-the-art algorithms, best practices and an open-access benchmark.* Delft: Delft University of Technology.

Loguru. (2023, April 10). Retrieved from https://pypi.org/project/loguru/

Matplotlib. (2023, February 13). Retrieved from https://pypi.org/project/matplotlib/3.7.0/

Newbery, D. (2004). *Deterministic frequency deviations – root causes and solutions for Europe.* Brussels: ETSO.

Nord Pool. (2025, August 25). *Single Price Coupling Algorithm*. Retrieved from nordpoolgroup: https://www.nordpoolgroup.com/globalassets/download-center/single-day-ahead-coupling/euphemia-public-description.pdf

Notebook. (2025, August 5). Retrieved from https://pypi.org/project/notebook/#description

Numpy. (2022, December 18). Retrieved from https://pypi.org/project/numpy/1.24.0/

Official Journal of the European Union. (2019). *DIRECTIVE (EU) 2019/944 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU.* Brussels: European Union.

Ovaere, M., Kenis, M., Von Den Bergh, K., Bruninx, K., & Delarue, E. (2023). *Marten Ovaere, Michiel Kenis, Kenneth Van den Bergh, Kenneth Bruninx, Erik Delarue, The effect of flow-based market coupling on cross-border exchange volumes and price convergence in Central Western European electricity markets.* Ghent: Energy Economics.

Pandas. (2023, August 30). Retrieved from https://pypi.org/project/pandas/

Plotly. (2023, September 15). Retrieved from https://pypi.org/project/plotly/

Poetry. (2023, 11 16). Retrieved from https://python-poetry.org/

Pontenagel, P. (2025, August 25). *All you need to know about market coupling in European power markets*. Retrieved from time2market: https://www.time2market.net/blog/market-coupling-in-european-power-markets?

Pydantic. (2023, September 25). Retrieved from https://pypi.org/project/pydantic/2.4.0/

Pydantic-settings. (2023, June 30). Retrieved from https://pypi.org/project/pydantic-settings/

Python. (2023, October 2). *Python 3.12*. Retrieved from https://www.python.org/downloads/release/python-3120/

Python-Dotenv. (2023, February 24). Retrieved from https://pypi.org/project/python-dotenv

Requests. (2023, May 22). Retrieved from https://pypi.org/project/requests/

Schönheit, D., Kenis, M., Lorenz, L., Most, D., Delarue, E., & Bruninx, K. (2021). *Toward a fundamental understanding of flow-based market coupling for cross-border electricity trading* . Dresden: Technische Universität Dresden.

SciPy. (2025, September 4). Retrieved from https://docs.scipy.org/doc/scipy/release/1.11.0-notes.html

Sckit-Learn. (2023, June 30). Retrieved from https://pypi.org/project/scikit-learn/1.3.0/

Seaborn. (2024, January 25). Retrieved from https://pypi.org/project/seaborn/

SqlAlchemy. (2023, January 23). Retrieved from https://pypi.org/project/SQLAlchemy/

Steer, G. (2025, September 4). *The 'mundane economics' of trading at the speed of light*. Retrieved from Financial Times: https://www.ft.com/content/4bedfc86-df01-44d8-ae31-6d7fa37520c9

Timescale. (2025, September 2). *TimescaleDB - Docker*. Retrieved from Docker: https://hub.docker.com/r/timescale/timescaledb

Titz, M., Putz, S., & Witthaut, D. (2023). *Identifying drivers and mitigators for congestion and redispatch in the German electric power system with explainable AI* . Ithaca: Cornell University.

Tqdm. (2024, November 24). Retrieved from https://pypi.org/project/tqdm/

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting*, 1030-1081.

Appendix