Mean-Reverting for a Dream:
Using Stochastic and GARCH-Based Forecasting to Detect and Exploit Electricity Market Inefficiencies in the German Spot Market

Maximilian Holmberg

Dissertation written under the supervision of Alberto Manconi

Dissertation submitted in partial fulfilment of requirements for the MSc in International Finance, at Universidade Católica Portuguesa and for the MSc in Accounting, Financial Management and Control (AFC) at Bocconi University, 03/09/2025.

Abstract: English

This thesis develops and evaluates a hybrid forecasting model combining an Ornstein-Uhlenbeck (OU) mean-reverting process with an Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) volatility specification to predict day-ahead electricity spot prices in the German market from 2015 to 2025. Addressing the challenges of non-stationarity, limited data transparency, and the need for economically relevant metrics, the model captures key market characteristics — mean reversion, seasonality, volatility clustering, and asymmetric shock responses — using only historical price data. Out-of-sample forecasts over 20-, 30-, 60-, and 90-day horizons are benchmarked against market futures and a naïve historical average, employing statistical metrics (MAE, RMSE, MAPE, directional accuracy) and economic measures (Sharpe ratio, trading returns). Results demonstrate that the model, particularly under a Generalized Error Distribution (GED), significantly outperforms futures benchmarks at short-to-medium horizons (20–30 days), generating positive trading returns (e.g., EUR 2,371.56 at 30 days) with high Sharpe ratios (e.g., 1.15) and win rates (up to 90%). Monte Carlo simulations and bootstrap confidence intervals confirm robustness, though performance weakens in low-volatility summer regimes. The thesis contributes to the literature by integrating mean-reverting dynamics with asymmetric volatility, testing against real futures prices, and demonstrating economic value through a forecast-based trading strategy, offering a replicable framework for traders and risk managers in volatile, non-storable commodity markets.

Keywords: Ornstein-Uhlenbeck, Electricity, Non-stationarity, Mean reversion, Volatility clustering, Monte Carlo, Spot prices, Futures, Out-of-sample, Leverage effects, Stochasticity.

Title: Mean-Reverting for a Dream: Using Stochastic and GARCH-Based Forecasting to Detect and Exploit Electricity Market Inefficiencies in the German Spot Market.

Author: Maximilian Holmberg

Resumo: Português

Esta dissertação desenvolve e avalia um modelo híbrido de previsão que combina um processo de reversão à média de Ornstein-Uhlenbeck (OU) com uma especificação de volatilidade Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH), aplicado à previsão dos preços spot de eletricidade no mercado alemão para o dia seguinte entre 2015 e 2025. O modelo enfrenta três desafios principais: a não-estacionariedade decorrente de transformações estruturais, a limitada transparência dos dados fundamentais e a necessidade de métricas economicamente relevantes. Para tal, capta características centrais do mercado — reversão à média, sazonalidade, aglomeração de volatilidade e respostas assimétricas a choques — utilizando exclusivamente dados históricos de preços. As previsões fora da amostra, em horizontes de 20, 30, 60 e 90 dias, são comparadas com os futuros de mercado e com uma média histórica ingênua, avaliadas através de métricas estatísticas (MAE, RMSE, MAPE, precisão direcional) e medidas económicas (índice de Sharpe e retornos de estratégia). Os resultados mostram que o modelo, sobretudo sob a distribuição de erro generalizada (GED), supera significativamente os futuros de mercado em horizontes curtos a médios, gerando retornos positivos (e.g., EUR 2.371 em 30 dias) com elevados índices de Sharpe (1,15) e taxas de sucesso até 90%. Simulações de Monte Carlo e intervalos de confiança bootstrap confirmam a robustez, ainda que o desempenho enfraqueça em regimes de baixa volatilidade. A dissertação contribui para a literatura ao integrar dinâmicas de reversão à média com volatilidade assimétrica, demonstrando valor económico através de uma estratégia de negociação baseada em previsões.

Palavras-chave: Ornstein-Uhlenbeck, Eletricidade, Não-estacionariedade, Reversão à média, Aglomeração de volatilidade, Monte Carlo, Preços spot, Futuros, Fora da amostra, Efeitos de alavancagem, Estocasticidade.

Título: Reversão à Média para um Sonho: Uso de Previsão Estocástica e Baseada em GARCH para Detectar e Explorar Ineficiências do Mercado de Eletricidade no Mercado Spot Alemão.

Autor: Maximilian Holmberg

Abstract: Italiano

Questa tesi sviluppa e valuta un modello di previsione ibrido che combina un processo di media-reversione di Ornstein-Uhlenbeck (OU) con una specificazione di volatilità Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH), al fine di prevedere i prezzi spot dell'elettricità nel mercato tedesco del giorno successivo dal 2015 al 2025. Affrontando le sfide della non-stazionarietà, della limitata trasparenza dei dati e della necessità di metriche economicamente rilevanti, il modello cattura le principali caratteristiche del mercato — reversibilità alla media, stagionalità, clustering di volatilità e risposte asimmetriche agli shock — utilizzando esclusivamente dati storici di prezzo. Le previsioni out-of-sample su orizzonti di 20, 30, 60 e 90 giorni vengono confrontate con i futures di mercato e con una media storica ingenua, impiegando metriche statistiche (MAE, RMSE, MAPE, accuratezza direzionale) e misure economiche (rapporto di Sharpe, rendimenti della strategia). I risultati dimostrano che il modello, in particolare sotto l'ipotesi di una distribuzione degli errori generalizzata (GED), supera significativamente i futures di mercato negli orizzonti di breve e medio termine (20–30 giorni), generando rendimenti positivi (ad es. EUR 2.371,56 a 30 giorni) con alti rapporti di Sharpe (ad es. 1,15) e tassi di successo fino al 90%. Simulazioni Monte Carlo e intervalli di confidenza bootstrap confermano la robustezza, sebbene la performance si indebolisca nei regimi estivi a bassa volatilità. La tesi contribuisce alla letteratura integrando dinamiche di media-reversione con volatilità asimmetrica, testandole contro prezzi reali dei futures e dimostrando valore economico tramite una strategia di trading basata sulle previsioni, offrendo un quadro replicabile per trader e risk manager nei mercati volatili delle commodity non immagazzinabili.

Parole chiave: Ornstein-Uhlenbeck, Elettricità, Non-stazionarietà, Reversione alla media, Clustering di volatilità, Monte Carlo, Prezzi spot, Futures, Fuori campione, Effetti di leva, Stocasticità.

Titolo: Reversione alla media per un sogno: utilizzo di previsioni stocastiche e basate su GARCH per individuare e sfruttare le inefficienze del mercato elettrico nel mercato spot tedesco.

Autore: Maximilian Holmberg

Acknowledgements

I would like to express my deepest gratitude to my advisor, Professor Alberto Manconi, for his invaluable guidance, patience, and encouragement throughout the development of this thesis. His expertise and thoughtful feedback have been essential in shaping both the direction and quality of my research. I am also thankful to my professors at Católica Lisbon School of Business and Economics and at Bocconi University. Their teaching and mentorship have provided not only the academic foundation for this work but also a lasting source of intellectual inspiration. I am grateful to my colleagues at Citi, who shared their professional insights and allowed me to write my thesis during work hours. Most importantly, I wish to thank my fellow students and friends at both Católica Lisbon and Bocconi. Their collaboration, encouragement, and companionship have made this journey both intellectually rewarding and personally memorable. Finally, I owe my deepest appreciation to my family and close friends, whose constant support and belief in me have been a source of strength and motivation throughout my studies.

Contents

I. List of tables

II. List of Abbreviations

ARFIMA	Autoregressive Fractionally Integrated Moving Average
ARMA	Autoregressive Moving Average
CI	Confidence Interval
DA	Directional Accuracy
DM Test	Diebold-Mariano Test
EEX	European Energy Exchange
EGARCH	Exponential Generalized Autoregressive Conditional Heteroskedasticity
EUR/MWh	Euro per Megawatt Hour
GARCH	Generalized Autoregressive Conditional Heteroskedasticity
GED	Generalized Error Distribution
GWh	Gigawatt Hour
Hz	Hertz
LSEG	London Stock Exchange Group
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MLE	Maximum Likelihood Estimation
OLS	Ordinary Least Squares
OOS	Out-of-Sample
OU	Ornstein-Uhlenbeck (process)
PDF	Probability Density Function
R^2	Coefficient of Determination
RMSE	Root Mean Squared Error
S&P 500	Standard & Poor's 500 Index
SDE	Stochastic Differential Equation
TRPC	TRPC Electricity Germany Peakload Day 1 (LSEG futures data ticker)

III. List of Equations

1. Introduction

"Energy conversions are the very basis of life and evolution. Modern history can be seen as an unusually rapid sequence of transitions to new energy sources, and the modern world is the cumulative result of their conversions." (Smil, 2022).

This thesis develops and evaluates a hybrid forecasting model combining an Ornstein-Uhlenbeck (OU) mean-reverting process with an Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) volatility specification to predict day-ahead electricity spot prices in the German market from 2015 to 2025. Addressing the challenges of non-stationarity, limited data transparency, and the need for economically relevant metrics, the model captures key market characteristics such as mean reversion, seasonality, volatility clustering, and asymmetric shock responses using only historical electricity price data. The OOS forecasts over 20-, 30-, 60-, and 90-day horizons are benchmarked against market futures and a naïve historical average, employing statistical metrics (MAE, RMSE, MAPE, and directional accuracy) and economic measures (Sharpe ratio and cumulative and mean returns). Results demonstrate that the model, particularly under a Generalized Error Distribution (GED), significantly outperforms futures benchmarks at short-to-medium horizons (20–30 days), generating positive trading returns (EUR 2,371 at 30 days) with high Sharpe ratios (1.15) and win rate at 90%. Monte Carlo simulations and bootstrap confidence intervals confirm robustness, though performance weakens in low-volatility summer regimes. The thesis contributes to the literature by integrating mean-reverting dynamics with asymmetric volatility, testing against real futures prices, and demonstrating economic value through a forecast-based trading strategy, thus offering a replicable framework for traders and risk managers in electricity markets.

This thesis addresses three central challenges that have hindered conclusive answers to the research question: First, non-stationarity arising from structural shifts, second, limited transparency in market fundamentals, and third the absence of a consistent definition of "outperformance." First, most time-series frameworks assume stationarity, i.e., that statistical properties such as mean and variance are constant. The German electricity market violates this assumption due to structural transformation in power generation. The Energiewende has replaced coal and nuclear generation with intermittent wind and solar, amplifying supply volatility. Further integration with European markets via cross-border trading and flow-based coupling has heightened sensitivity to external shocks and transmission constraints. Models calibrated on pre-transition data, risk parameter instability, and reduced robustness across regimes. Second, high-frequency data that rely on drivers such as weather, renewable generation, load, fuel prices, and grid constraints are often incomplete, inconsistently reported, or restricted (European Commission, 2018). These constraints limit the feasibility of a fundamental-based forecasting approach. Third, existing literature predominantly measures forecast quality via statistical loss functions (RMSE, MAE, MAPE), therefore implicitly valuing all errors equally. For market participants, however, utility derives from improved hedging efficiency or trading profitability, making purely statistical measurements insufficient for assessing the economic viability.

To address the non-stationarity from structural shifts, the model combines a mean-reverting stochastic process to capture long-term equilibrium behaviour, an EGARCH specification to model time-varying and asymmetric volatility, and a deterministic seasonal term to reflect recurring demand–supply cycles. The focus on the post-2015 period ensures training on a structurally mature market with high renewable penetration and elevated volatility, improving the relevance of the OOS results. To address the poor visibility on market fundamentals, rather than incorporating incomplete or unreliable fundamental variables, my hybrid model uses only one data set of historical prices. This approach assumes that prices incorporate the available information, avoid dependence on proprietary or delayed datasets, and ensure transparency, reproducibility, and applicability in other markets with similar data limitations. Finally, to address the challenge that there is no consensus on how to define “outperformance”, my forecast evaluation uses both statistical error measures (RMSE, MAE, MAPE) and economic measures (directional accuracy, Sharpe Ratio, trading performance). The latter is particularly relevant to trading and hedging. Futures prices are treated not as objective benchmarks but as competing forecasts, and this dual role is explicitly addressed in the interpretation of results.

My main findings are as follows. First, a simple statistical model that accounts for fundamental characteristics of electricity pricing, such as mean reversion, volatility clustering, and seasonality, is able to achieve lower statistical measures and higher simulated returns than futures contracts. Thus, creating the opportunity to build a trading strategy based on trading mispriced futures contracts. This strategy has proven profitable over multiple time windows and volatility specifications. Second, assuming futures reflect market consensus, we test for mispricing using model-implied prices. As the futures’ prices are largely driven by orders from large institutions for hedging or speculative purposes, using factor- and more complicated statistical models (Mahajan, 2006), my thesis highlights the economic value of a single dataset versus a factor model. I can therefore ask, why use the more data-intensive — therefore costlier — factor model using multiple sets of data, when a simpler one can outperform? The exact economic benefit of my model versus a more complicated one, with regard to data handling and computational power, is, however, still unclear, as I currently have no access to such a model for benchmarking. Third, while the naïve forecast, based on recent averages, is competitive — proving a point on mean-reversion — my results show that a forecast based on my hybrid model offers economic value when applied to a trading simulation. Moreover, the naïve forecast showed significantly positive results when the same trading simulation was applied for directionally based trading. For practitioners, the ability of my model and the naïve model to outperform the market (as proxied by futures) challenges the economic viability of any type of more complicated electricity trading model. For academia, my results reinforce the importance of combining economic structure with robust volatility modeling when evaluating market efficiency and price discovery in energy markets.

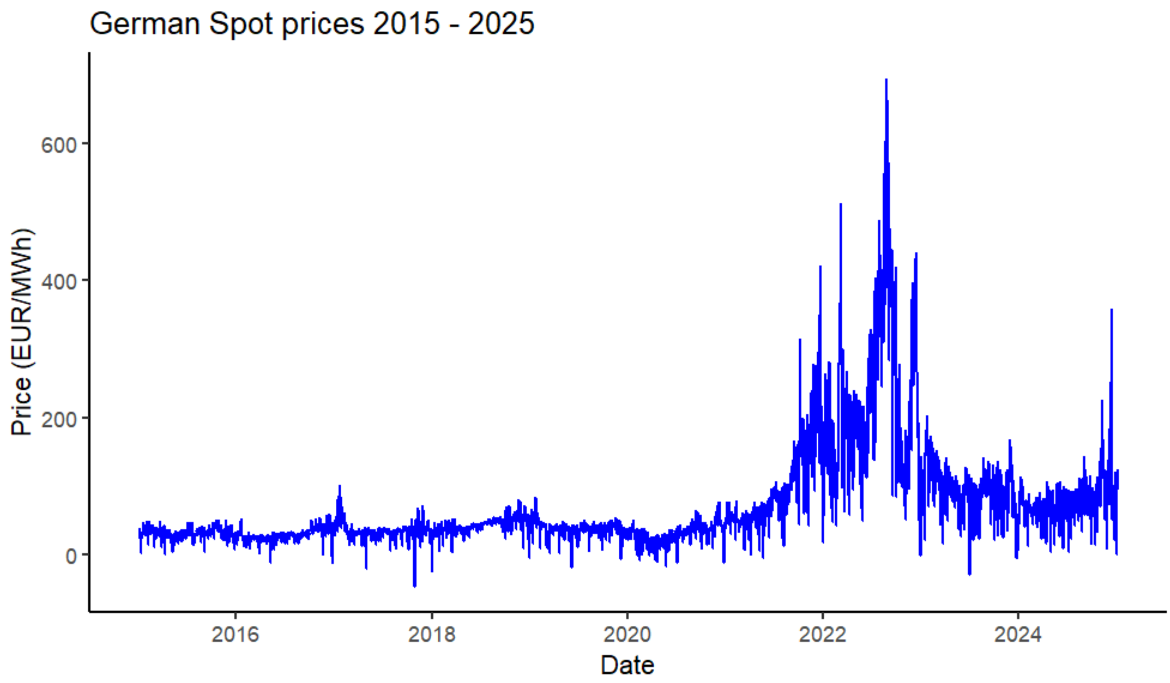
My thesis makes three contributions to the literature. First, the literature on electricity price modeling using Ornstein-Uhlenbeck processes shows that electricity prices are mean-reverting with seasonal trends, also described by Lucia and Schwartz (2002), Cartea and Figueroa (2005), Geman and Roncoroni (2006), Benth et al. (2007), and Castañeda-Leyva et al. (2022). However, these models often ignore time-varying volatility. I incorporate an EGARCH volatility structure into the Ornstein-Uhlenbeck process, capturing both mean reversion and volatility clustering within a unified framework. Second, existing studies on German electricity markets largely rely on older or single-component models (Geman and Roncoroni, 2006, and Benth et al., 2007) and focus on in-sample fit. There is limited research evaluating recent hybrid models under current market conditions. In my thesis, I apply a modern hybrid model to updated German market data (2015-2025), providing more accurate and relevant OOS forecasts. Third, while Ornstein-Uhlenbeck-based models are used to derive theoretical futures prices, few studies systematically test for futures mispricing using real market data, as evidenced by Bierbrauer et al. (2007), Lucia and Schwartz (2002), and Benth et al. (2007). There is a gap in linking spot price forecasts with observed futures prices, as is presented by Castaneda-Leyva et al. (2022) and Koopman et al. (2007). My thesis compares forecasted spot prices to traded futures, identifying potential mispricing and opportunities for informed speculation or risk management.

1.2 The German electricity market: an overview

Given the model's focus on volatile markets, Section 1.2 outlines Germany's electricity market characteristics. The electricity market is structured as a commodity, and just like oil, wheat, aluminium, or any other commodity that is fungible and traded on global markets, is characterized by minimal differentiation and price fluctuations influenced by supply and demand. As with any other commodity, there are reasons to use it in various ways. A manufacturing company may purchase futures contracts of e.g., aluminium in order to hedge the risk of changes in the future purchase price of the aluminium. Aluminium is much like electricity in this sense; they are both used for the production of goods and services, and they can be hedged to reduce the risk of uncertainty in the cost of manufacturing. Another use case of the commodity is speculation. Just as the manufacturer, there are various instruments one may use to estimate the future price of a product today; a common one for commodities is futures contracts, but there are other ones that may differ in complexity and characteristics. Should a speculator believe that the commodity-linked instrument is mispriced, he or she could generate a return by more accurately estimating the correct price. As such, these two uses of electricity as a commodity imply that you would be better off if you knew the future price of electricity. Enron famously used the second case before its collapse in 2001 (Federal Bureau of Investigation, 2001). Currently, many hedge funds still trade energy and electricity, either as part of their business, such as Citadel, AQR, and Millennium Management, or as their primary business, as seen with Graham Capital Management and PointState Capital, according to Hedgelists (2023). The high degree of use and institutional interest in the energy markets signifies the importance of the future prices of electricity. Therefore, as evidenced by the hedge funds mentioned, if you can more accurately than others predict the future price, you can create a profitable business (Hedgelists, 2023).

A difference between electricity and other commodities is that electricity cannot be stored — at least not to a significant extent — and must be consumed instantly. The inability to store electricity in large quantities introduces another level of complexity not found in equity markets, according to Bessembinder and Lemon (2002). Furthermore, as electricity is constantly consumed and produced, the electricity market becomes fragile; minor disruptions in demand or supply induce ripple effects on the market (Ioannides et al., 2021). Moreover, the infrastructure of the electrical grids, through which electricity passes, cannot remain idle but must oscillate at a steady frequency (Kilic & Huisman, 2010). On the European electricity market, this frequency needs to be between 49.5 Hz and 50.5 Hz (National Grid, 2009). There needs to be a constant Hz in the grid at all times; if not, then electricity cannot flow, and consumption as well as trading will stop. This phenomenon represents a positive feedback loop, where an initial supply disruption leads to increased system stress due to the grid's requirement to maintain constant frequency, effectively amplifying the original imbalance. Thereby increasing the inherent volatility and volatility clustering of electricity prices.

The German electricity market is chosen in this thesis due to its structural characteristics and data availability. Germany currently has a single electricity price zone, meaning that all of Germany has the same electricity prices, and electricity fluctuations affect the whole country simultaneously (Wettengel, 2025). Compared to the Swedish electricity market, which is divided into four electricity price zones, effectively compartmentalizing supply-demand shocks. Leaving the German electricity market exposed, as an oil tanker without the compartments to reduce the movements of the oil, amplifies the movements of the boat. As coal and nuclear energy have decreased in the German production mix from year 2000 to 2023, and intermittent energy sources such as wind power and solar PV have increased their contribution, the German power grid has become more volatile (IEA, 2023). Intermittent power sources contribute to grid instability by not allowing for output control. Whereas continuous power sources, such as coal, hydro, and nuclear, allow the power output to match the power demand, thereby reducing volatility in the electricity market. Germany's high penetration of intermittent power sources (wind 27% and solar PV 12% contribution in 2023) leads to price volatility, frequent supply shocks, and positive feedback loops, creating volatility clustering (Bierbrauer et al., 2007). Furthermore, Germany is connected to the major European electricity markets such as the European Power Exchange and Nord Pool (Nord Pool, 2024). The country, therefore, operates in Europe's largest and most liquid electricity markets. This liquidity ensures the availability of high-quality data. Additionally, the high degree of liquidity reduces the spreads between supply and demand, yielding a more accurate spot price. In 2023, Germany exported approximately 12% of its total production and imported approximately 13% of its total production, signifying Germany's reliance on the interconnectivity of the European power grid (IEA, 2023). According to Bessembinder and Lemmon (2002), this complexity contributes to increased difficulty in pricing futures and in accurately forecasting a future price path. The seasonality drives the need for a statistical model that can incorporate seasonal adjustments. These characteristics make Germany ideal for testing a hybrid model that captures mean-reversion and asymmetric volatility.



2. Literature review

This chapter reviews the key modeling approaches in previous literature, starting with Ornstein-Uhlenbeck (OU) processes, followed by Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models, and hybrid frameworks. The review highlights methodological advancements, empirical applications, and persistent gaps, particularly in integrating structural price dynamics with flexible volatility modeling for OOS forecasting and economic applications. By synthesizing recent and fundamental studies, I position my OU-EGARCH hybrid model as a parsimonious solution that addresses these shortcomings, emphasizing transparency, reproducibility, and trading relevance in the German market.

Although the Ornstein–Uhlenbeck (OU) process is widely used for modeling electricity spot prices due to its mean-reverting structure, the previous studies show divergence in how the OU model handles real electricity market structures. Those divergences include how it handles spikes, volatility, and interpretability. Lucia and Schwartz (2002) provide the foundational two-factor OU model, combining seasonality with mean reversion. Its tractability allows forward pricing, but it lacks mechanisms for price spikes and volatility clustering. Cartea and Figueroa (2005) address this by superimposing Poisson jumps on the OU base, capturing leptokurtic shape returns and spikes. However, their constant, exogenous jump intensity limits the explanation of volatility persistence. Geman and Roncoroni (2006) focus on spike realism via a "jump-reversion" process that treats spikes as endogenous to equilibrium dynamics; however, their model sacrifices tractability and gives no possibility of back-testing. Furthermore, Benth et al. (2007) retain analytical solvability by introducing Lévy noise to capture skewness and kurtosis more effectively than prior models; however, like Lucia and Schwartz (2002), Benth et al. (2007) assume static or implicit variance, thus neglecting volatility clustering. Moreover, Castañeda-Leyva et al. (2022) incorporate latent demand into a power-transformed OU process, grounding mean reversion in economic theory and indirectly inducing heteroskedasticity. While this improves interpretability, time-varying volatility is not explicitly modeled, thus limiting the relevance for practitioners within, e.g., forecasting and risk management.

A common weakness of these studies is the absence of explicit volatility modelling. These papers treat volatility as variously constant (Lucia and Schwartz, 2002; Benth et al., 2007), exogenous (Cartea and Figueroa), implicit (Geman and Roncoroni, 2006), or static latent (Castañeda-Leyva, 2022). By not explicitly modelling for volatility, this leaves a gap in representing a stylized fact of electricity markets. Additionally, this reflects an academic disagreement, whether spikes are anomalies to filter out or normal electricity price patterns to model. Forecasting performance is another shortcoming. Except for Castañeda-Leyva et al. (2022), none of the reviewed papers rigorously evaluate OOS forecasts over practical horizons or benchmark against futures contracts. Even when the IS fit is strong, the absence of stress-testing against futures or market prices limits the practical utility for trading or risk management.

The application of GARCH models to electricity prices represents a shift from constant-variance assumptions toward capturing volatility clustering, leverage effects, and tail risk. Although these models successfully model the second-moment dynamics, they often underrepresent the underlying price process in markets characterized by structural mean reversion, seasonal cycles, and extreme shocks. Garcia et al. (2005) combined a GARCH(1,1) specification with seasonal terms, showing high volatility persistence, especially during peak load periods. However, their mean specification is simplistic and fails to reflect electricity's mean-reversion behavior. Liu and Shi (2013) extend this by modeling GARCH-in-Mean, allowing the volatility to feed back into price levels. While addressing the endogenous risk–price link, they still rely on an ARMA mean processes, which lack economic interpretability compared to mean-reverting frameworks. Knittel and Roberts (2005) challenge common GARCH implications of normally distributed errors by documenting an inverse leverage effect in U.S. electricity prices — volatility rises disproportionately after positive shocks in electricity markets, as opposed to equity markets. However, their empirical approach remains simplistic and does not explicitly adjust for the asymmetry. Subsequent work, such as Ioannides et al. (2021), models conditional skewness and kurtosis in a periodic GARCH-M framework, capturing seasonal variation and higher moments. This improves risk characterization, particularly under extreme conditions, but increases parameterization and overfitting risk in volatile or regime-shifting markets. Misiorek et al. (2006) highlight that GARCH components may only modestly improve point forecasts but substantially enhance interval forecasts, thus emphasising their value in estimating forecast uncertainty. However, they treat mean and variance as separate, therefore missing full integration of volatility–price feedback. Wang and Wu (2012) compare multivariate (DCC-GARCH) and univariate models, finding minimal gains for electricity compared to commodities such as oil or gas. This supports the view that electricity volatility is primarily driven by localized, idiosyncratic shocks, limiting the value of high-dimensional cross-asset volatility modeling.

Two shortcomings emerge from this literature. First, most studies use mean structures with limited economic grounding, typically ARMA processes (e.g., Misiorek et al., 2006, and Liu and Shi, 2013), thus breaking the link between prices and price drivers such as supply-demand balance or fuel costs. Second, volatility is often treated as statistically decoupled from the mean. Therefore, we find a methodological split: models either feature a well-structured mean (as in continuous-time OU processes) or sophisticated volatility (as in GARCH frameworks), but seldom both. This gap motivates the hybrid model I develop in my thesis. I include an EGARCH process within an OU mean-reverting structure. This design simultaneously captures the economic mean-reverting nature of electricity prices as well as the asymmetric, time-varying volatility responses. It also increases applicability to real-world contexts such as detecting futures mispricing, improving hedging strategies, and informing speculative trading decisions.

Hybrid models emerged to address the inability of single-process approaches to fully capture electricity market dynamics. Huisman and Mahieu (2003) and Janczura and Weron (2010) use discrete regime-switching OU processes for “normal” and “spike” regimes. These simplify estimation and interpretation but perform poorly when regime identification is uncertain, thus trading flexibility for clarity. Koopman et al. (2007) use an ARFIMA process to model long memory in prices, which improves forecasts by reproducing slow-decaying autocorrelations, but still lacks economic interpretability and does not model the mean-reversion. Furthermore, the treatment of volatility is also fragmented in previous literature. Koopman et al. (2007) add a GARCH specification to conditional variance, producing realistic volatility clustering and better short-term risk forecasts. However, they model volatility as disconnected from the underlying market mechanisms, limiting value for futures pricing. Koopman et al. (2007) target short-term price and volatility forecasting, optimising statistical fit via RMSE and interval coverage. While Bierbrauer et al. (2007) link an OU spot price model to futures via a risk-neutral framework for derivative pricing, offering analytical tractability but assuming constant risk premia and static volatility for short-term forecasting and adaptability to changing market conditions. This produces a split between models optimised for statistical forecasting and those built for financial interpretation. Koopman et al.'s (2007) ARFIMA–GARCH performs across short and long horizons, while regime-switching models excel at short-term spike prediction, they lose robustness over longer horizons. None of these previous studies is rigorously benchmarked against market-based measures such as traded futures. Thus, leaving a gap of usability for speculative or arbitrage-focused applications.

In my thesis, I build on this literature by proposing a hybrid OU–EGARCH framework that combines the structural mean-reversion of the OU process with the asymmetric, persistent volatility captured by EGARCH. The OU component, grounded in the economic interpretations of Lucia and Schwartz (2002) through Castañeda-Leyva et al. (2022), models long-term equilibrium behavior and seasonal effects. The EGARCH specification, following Nelson (1991) and aligned with the inverse leverage patterns observed by Knittel and Roberts (2005), models conditional variance in a way that accommodates both volatility clustering and asymmetric responses to shocks. By modeling the logarithm of variance, EGARCH avoids parameter restrictions while ensuring positivity and flexibility. This combination enables the joint forecasting of price levels and volatility, offering improved statistical performance and the ability to translate forecasts into trading signals. This thesis tests both predictive accuracy and economic value, demonstrating whether a structurally grounded, volatility-aware model can exploit mispricings in the futures market. In this way, the hybrid model framework integrates the strengths and addresses the weaknesses of prior OU, GARCH, and hybrid models, contributing a coherent, operationally relevant addition to the electricity price forecasting literature.

3. Methodology

3.1 Data

This thesis uses historical wholesale day-ahead electricity spot prices for Germany, from Ember (Ember. 2025). The dataset, updated monthly, provides daily load-weighted average prices in EUR/MWh from January 2015 to January 2025, capturing the post-Energiewende market dynamics characterized by high renewable penetration and volatility. A shortened subset, excluding the days corresponding to the forecasting horizon from the original dataset, is used for OOS forecasting in comparison with realized prices. The price-only approach ensures transparency and replicability, assuming that spot prices embed market information without requiring proprietary exogenous variables.

For benchmarking, daily futures prices are obtained from LSEG Workspace (2025) (TRPC Electricity Germany Peakload Day 1). These provide daily closing prices in EUR/MWh from January 2015 to January 2025, aligned with forecast price horizons. Futures prices serve as market-implied forecasts for evaluating mispricing and generating trading signals. Data processing is performed in R (version 4.4.1) using packages such as tidyverse, readxl, and lubridate. Spot prices and futures prices are loaded from Excel files, with values as numeric, then converted to a time series with daily frequency. All processing code is provided in the appendix for reproducibility.

3.2 Model Framework Overview

The modeling framework combines a discretized Ornstein-Uhlenbeck process for spot price dynamics together with an Exponential GARCH (EGARCH) specification for conditional volatility. This hybrid structure is designed to capture the key stylized facts of electricity prices, i.e., mean reversion, seasonality, volatility clustering, and asymmetric responses to shocks.

The Ornstein-Uhlenbeck process is modelled to capture the overarching appearance of the electricity prices and is built as a mean-reverting stochastic differential equation with a deterministic seasonal adjustment. Thus, capturing the pattern of electricity prices to revert back to its long-term mean, while also accounting for the effect of season on the prices, i.e., more expensive in the winter than the summer months. Parameters governing the speed of mean reversion, long-term mean level, and diffusion volatility are estimated using Ordinary Least Squares (OLS) and Maximum Likelihood Estimation (MLE) methods. The seasonality component is specified as a sinusoidal function to reflect recurring annual consumption patterns.

To capture the volatility patterns of electricity prices, residuals from the spot price model are modeled using an EGARCH(1,1) process estimated under GED, Student-t, and Normal distribution. The EGARCH specification captures both volatility clustering and asymmetries in the volatility response, allowing positive and negative shocks to have different impacts on the forecasted volatility (Nelson, 1991). To account for heavy-tailed residuals typical of electricity prices, the innovations are assumed to follow a GED, following Giller (2005). The use of a logarithmic conditional variance structure in the EGARCH model removes the need for non-negativity constraints and provides robustness against extreme events.

Forecasts are generated recursively through one-step-ahead simulation. To assess forecast stability, model performance is evaluated over multiple OOS horizons: 20, 30, 60, and 90 days. It is also tested in 3 electricity price regimes, the first and original regime has end date, 5th of January 2025 (I am thus utilizing the entirety of the data set that was available at the time of writing this thesis), the other regimes are the immediately preceding October and June regimes, both tested for 30 days each.

Finally, a trading strategy is implemented based on the spread between model-implied future spot prices and observed futures contract prices. This is to provide an economic context for the predictive ability of the model. Strategy performance is benchmarked against naive forecasts and evaluated through both point estimation and Monte Carlo simulation, which assesses the robustness of trading outcomes under stochastic price path evolution.

3.3 Modeling the Spot Price Process via the Ornstein-Uhlenbeck Process

Electricity spot prices exhibit mean-reverting dynamics driven by fundamental market forces such as supply-demand imbalances, regulatory interventions, and capacity constraints (Lucia and Schwartz, 2002; Benth et al., 2007). To capture this behavior, the spot price process P_t is modeled as a continuous-time OU process, augmented with a deterministic seasonal component to reflect predictable demand patterns throughout the year (Misiorek et al., 2006; Huisman and Mahieu, 2003).

The continuous-time stochastic differential equation (SDE) for the spot price dynamics is specified as:

$$P_t = \theta(\mu - P_t)dt + \sigma dW_t + S_t \tag{1}$$

where: $\theta > 0$ is the speed of mean reversion toward the long-term mean level μ; μ is the equilibrium (long-term) mean price; $\sigma > 0$ is the instantaneous volatility; dW_t denotes a standard Wiener process capturing random shocks, and S_t is a deterministic seasonality function reflecting regular, cyclical demand patterns. The OU process specification implies that deviations from the mean are temporary, decaying exponentially at a rate θ, consistent with stylized facts observed in electricity price behavior.

Given that the available data are observed at a daily frequency, the SDE is discretized using an Euler-Maruyama approximation with a time step $dt = 1/365$:

$$P_{t+1} = P_t + \theta(\mu - P_t)dt + \sigma\sqrt{dt}\varepsilon_{t+1} + S_t \tag{2}$$

where $\varepsilon_{t+1} \sim \mathcal{N}(0, 1)$ are independent and identically distributed standard normal innovations.

The deterministic seasonal component S_t is modeled as a simple sinusoidal function, due to the apparent yearly electricity price cycles in the German electricity market (Weron, 2014):

$$S_t = A \sin\left(\frac{2\pi t}{365}\right) \tag{3}$$

where A represents the amplitude of the seasonal cycle. This seasonal specification is tractable and empirically effective, consistent with the approaches recommended by Misiorek et al. (2006). The parameters θ, μ, and σ are estimated in a two-stage procedure to ensure robustness and computational efficiency. Following Lucia and Schwartz (2002), a preliminary estimate of θ is obtained using OLS regression. Starting from the discrete form:

$$\Delta P_t = P_{t+1} - P_t = \theta(\mu - P_t)dt + \sigma\sqrt{dt}\varepsilon_{t+1} + \Delta S_t \tag{4}$$

Neglecting the small seasonal term over short intervals for initial estimation, I regress the daily differences ΔS_t on lagged price levels P_t via:

$$\Delta P_t = \alpha + \beta P_t + \varepsilon_t \tag{5}$$

where:

$$\beta = -\theta dt \tag{6}$$

and:

$$\alpha = \theta\mu dt \tag{7}$$

The OLS regression provides estimates of α and β, from which θ and μ can be preliminarily recovered as:

$$\theta = -\frac{\beta}{dt} \tag{8}$$

and:

$$\mu = \frac{\alpha}{\theta dt} \tag{9}$$

This initial OLS step serves primarily to provide starting values for subsequent maximum likelihood optimization.

The final parameter estimates are obtained by maximizing the likelihood function under the assumption that P_{t+1} conditional on P_t follows a normal distribution:

$$P_{t+1}|\, P_t \sim \mathcal{N}(P_t e^{-\theta dt} + \mu(1 - e^{-\theta dt}), \frac{\sigma^2}{2\theta}(1 - e^{-\theta dt})) \tag{10}$$

Accordingly, the log-likelihood function for the observed time series $\{P_t\}$ is:

$$\mathcal{L}(\theta, \mu, \sigma) = -\frac{1}{2}\sum_{t=1}^{T-1}(\log(2\pi v_t) + \frac{(P_{t+1} - m_t)^2}{v_t}) \tag{11}$$

where the conditional mean is:

$$m_t = P_t e^{-\theta dt} + \mu(1 - e^{-\theta dt}) \tag{12}$$

and the conditional variance is:

$$v_t = \frac{\sigma^2}{2\theta}(1 - e^{-2\theta dt}) \tag{13}$$

Numerical maximization of the negative log-likelihood is performed using the L-BFGS-B algorithm in RStudio, which allows for boundary constraints, ensuring $\theta > 0$ and, $\sigma > 0$ guaranteeing economically meaningful estimates. The standard errors are then obtained by inverting the Hessian matrix evaluated at the maximum likelihood estimates, facilitating statistical inference on the parameters (Dovi et al, 1991). This MLE approach is consistent with Benth et al. (2007) regarding continuous-time financial econometrics.

The discretized OU process, estimated via a two-stage OLS and MLE procedure, provides an empirically validated model for capturing the equilibrium-seeking behavior and seasonal patterns in German electricity spot prices. This foundation enables the separation of mean-reverting price dynamics from stochastic volatility patterns, the latter being modeled in the subsequent EGARCH stage.

3.4 Modeling the Volatility Process

While the Ornstein-Uhlenbeck process effectively captures the structural mean-reverting behavior and seasonal trends of electricity spot prices, it does not adequately model the phenomenon of volatility clustering — the tendency for periods of high volatility to be followed by high volatility, and low by low — nor the asymmetric effects of positive versus adverse shocks. These features are documented in electricity markets by, e.g., Knittel and Roberts (2005) and Huisman and Mahieu (2003), and necessitate a second modeling layer for the conditional variance process. To address this, the residuals from the OU model are modeled using an Exponential Generalized Autoregressive Conditional Heteroskedasticity (EGARCH) process, proposed by Nelson (1991).

The EGARCH model is particularly suited for electricity prices because it captures volatility clustering, allows for asymmetric effects (different impact of positive and negative shocks), and guarantees positivity of variance without requiring parameter restrictions (due to the log specification).

The EGARCH(1,1) process models the logarithm of the conditional variance h_t as:

$$\log(h_t) = \omega + \beta \log(h_{t-1}) + \alpha \left(\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} - \mathbb{E}\left[\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}}\right] \right) + \gamma \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \tag{14}$$

where: ω is the constant term; β captures volatility persistence; α captures the size effect of past shocks; γ captures the asymmetric effect of positive versus negative shocks; ε_{t-1} is the innovation (residual) from the previous period, assumed to be independent and identically distributed with mean zero and unit variance; and $h_{t= \sigma_t^2}$ is the conditional variance.

The logarithmic transformation of variance ensures that $h_t > 0$ automatically, a key advantage over standard GARCH models that require parameter constraints to maintain positivity (Nelson, 1991). Because the residuals are assumed to follow a standardized distribution, the expectation of the standardized absolute residual $\mathbb{E}\left[\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}}\right]$ for a standard normal innovation is $\sqrt{\frac{2}{\pi}}$. Thus, Equation (14) can be equivalently written as:

$$\log(h_t) = \omega + \beta \log(h_{t-1}) + \alpha \left(\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \right) + \gamma \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \quad (15)$$

The term:

$$\left(\frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{\frac{2}{\pi}} \right) \quad (16)$$

centers the magnitude effect around zero to avoid bias, ensuring that only unexpected volatility (rather than all volatility) influences future variance. The asymmetrical term:

$$\frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}} \quad (17)$$

allows negative shocks to have a different impact on volatility than positive shocks, capturing the empirically observed leverage effects in energy markets (Knittel and Roberts, 2005). A positive γ implies that positive shocks increase volatility more than positive ones of the same magnitude.

The innovation distribution is specified as GED, Student-t, or Normal. GED offers greater flexibility than the standard Normal or Student-t distributions by allowing the tail thickness and peak sharpness to be governed by a shape parameter. This choice is motivated by empirical evidence of fat-tailed residuals and leptokurtic distribution shape in electricity price series by Misiorek et al. (2006) and Geman and Roncoroni (2006). The GED accommodates both thinner and fatter tails, enabling the model to adapt to different volatility regimes. Furthermore, the GED provides a flexible parametric family that encompasses the Gaussian distribution as a special case and accommodates deviations from the traditional Gaussian function via the shape parameter ν. The probability density function (PDF) of the GED is given by:

$$f(\varepsilon_t) = \frac{\nu}{3\lambda\Gamma(1/\nu)} \exp(-\left|\frac{\varepsilon_t}{\lambda}\right|^{\nu}) \quad (18)$$

where: $\nu > 0$ is the shape parameter, for $\nu = 2$, the GED is shaped as the normal distribution, for $\nu < 2$, the distribution has heavier tails, and for $\nu > 2$, it becomes lighter-tailed; $\Gamma(.)$ is the gamma function; and $\lambda = \sigma \left(\frac{\Gamma(1/\nu)}{\Gamma(3/\nu)}\right)^{1/2}$ is a scaling parameter that ensures the conditional variance σ_t^2 is preserved. Estimation of the GED parameters is conducted jointly with the EGARCH process using maximum likelihood. This distributional assumption adjusts the error distribution from a normal distribution to a distribution more fitting to the kurtosis of electricity prices. The high kurtosis and the need to adjust for this are also observed by Knittel and Roberts (2005).

The hybrid OU–EGARCH framework requires stationarity in both mean and variance. The Ornstein–Uhlenbeck component provides mean reversion, ensuring that spot prices are stationary in levels. This is supported by Augmented Dickey–Fuller tests, which reject the presence of a unit root in the German day-ahead series. The EGARCH specification then models the conditional variance[1]. By working in logs, EGARCH guarantees positive variances without restrictive parameter constraints, and covariance stationarity holds as long as the persistence parameter β is strictly less than one (Nelson, 1991). Thus, the combined model is stationary in both first and second moments, satisfying the theoretical requirements for consistent estimation and forecasting.

3.5 Forecasting Procedure

This section details the procedure for generating OOS forecasts using the hybrid model. The methodology is based on a recursive one-step-ahead simulation framework, which is preferred over direct multi-step forecasts due to its ability to mitigate the accumulation of forecast errors (Diebold and Mariano, 1995). This structure closely reflects real-world forecasting scenarios where only historical data up to time t is available for generating a forecast for $t + 1$.

The forecast simulation proceeds in the following sequence: First, Ornstein-Uhlenbeck Forecasting: Using the most recently observed value P_t. The next price level is simulated with the estimated parameters and the seasonal component. Second, EGARCH Forecasting: The conditional volatility is obtained from the EGARCH(1,1) model fitted to the residuals of the in-sample Ornstein-Uhlenbeck process. Third, a random innovation ε_{t+1} is drawn from the GED, allowing for heavy/light tails. Fourth, the simulated price $\hat{P}_{t+1}$ is computed. Lastly, the newly simulated value becomes the input P_{t+1} for forecasting $\hat{P}_{t+2}$, and the process repeats until the full forecast window is completed. This simulation method ensures that each forecast is strictly conditional on information available up to the forecasting point, thereby maintaining OOS integrity.

[1] Although the ADF test rejects the null of a unit root, suggesting statistical stationarity of the German spot price series, this result does not imply structural stability. The German electricity market remains subject to frequent regime shifts driven by policy interventions, renewable penetration, and market coupling. Thus, while the data satisfy the statistical requirements of mean-reverting models, the broader environment is non-stationary in an economic sense. To account for this, I employ rolling estimation windows, seasonal adjustments, and a flexible volatility specification, ensuring adaptability to evolving market conditions.

Forecasting performance is then evaluated across the fixed out-of-sample windows: 20, 30, 60, and 90 days, in order to show robustness of the forecast over time. These horizons are selected to represent practical use cases in electricity markets, including short-term operational planning, monthly procurement, and quarterly trading strategies. The one-step-ahead forecasting design is repeated for each day within these windows, effectively rolling the forecast over the evaluation period while keeping the horizon fixed. Note that the parameters for both the Ornstein-Uhlenbeck formula and for EGARCH volatility forecasting are only estimated once for each OOS window. The model uses the estimated parameters with a new P_{t+n} as a starting value. Thus, the model only forecasts 1 day ahead for every day within the time window.

3.6 Model Evaluation

To assess the OOS performance of the hybrid model, this section presents the evaluation strategy using both statistical error metrics and formal hypothesis testing. The aim is to determine the predictive accuracy, robustness, and economic relevance of the model across multiple forecast horizons. The evaluation follows previous literature in electricity forecasting, e.g., Misiorek et al. (2006), and novel economic considerations, and includes both point forecast accuracy and directional consistency. The model's statistical forecast accuracy is evaluated using the following metrics:

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|P_t - \hat{P}_t\right| \tag{19}$$

MAE captures the average size of errors without disproportionately penalizing large deviations, making it robust to outliers.

Root Mean Squared Error (RMSE):

$$RSME = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(P_t - \hat{P}_t)^2} \tag{20}$$

RMSE emphasizes larger forecast errors due to squaring and is useful when large deviations are exceptionally costly.

Mean Absolute Percentage Error (MAPE):

$$MAPE = \frac{100\%}{n}\sum_{t=1}^{n}\left|\frac{P_t - \hat{P}_t}{P_t}\right| \tag{21}$$

MAPE provides a normalized error measure, facilitating interpretability across varying price levels, although it may be sensitive to near-zero prices.

Directional Accuracy (DA):

$$DA = \frac{1}{n}\sum_{t=1}^{n} 1[sign(\hat{P}_t - P_{t-1}) = sign(P_t - P_{t-1})] \quad (22)$$

Directional accuracy measures the model's ability to correctly predict the direction of price movement, which is particularly relevant in trading applications.

Out of sample R^2:

$$R^2_{OOS} = 1 - \frac{\sum(P_t - \hat{P}_t)^2}{\sum(P_t - \bar{P})^2} \quad (23)$$

This metric evaluates how much variance the model explains relative to a simple average benchmark, offering a summary statistic of model fit in predictive settings. Each metric is computed separately for the 20-, 30-, 60-, and 90-day forecast windows, allowing for the evaluation of model consistency and robustness across different forecasting horizons.

I benchmark against reference models to evaluate the relative performance and economic relevance of my proposed forecasting methodology. This section introduces two complementary benchmarks — one market-based and one statistically naïve — to serve as yardsticks for evaluating my hybrid model.

The first benchmark consists of observed day-ahead electricity futures prices, which are available through the European Energy Exchange (EEX)[2]. These futures contracts reflect the aggregate expectations of market participants of the future spot price and implicitly incorporate factors such as expected demand, production capacity, weather forecasts, regulatory signals, and risk premia. Under the assumption of no arbitrage and rational expectations, the futures price should converge to the expected spot price under the risk-neutral measure (Lucia and Schwartz, 2002). However, empirical studies in electricity markets consistently find that the futures price often diverges from the realized spot price, due to the non-storability of electricity,

[2] Electricity futures on the European Energy Exchange (EEX), including the Phelix DE and Phelix DE/AT contracts used in this study, are predominantly financially settled rather than physically delivered. Settlement occurs in cash, based on the difference between the contracted futures price and the realized average spot market price over the delivery period, as specified in the EEX contract specifications (EEX, 2017). This structure eliminates the need for the contract holder to either supply or take delivery of physical electricity, a critical feature given the non-storability of the commodity. For this research, the cash-settlement mechanism ensures that the trading strategy can be evaluated purely on its price forecasting accuracy and market value without introducing operational risks or logistical constraints associated with physical delivery.

market inefficiencies, and time-varying risk premia (Bierbrauer et al., 2007). By comparing model-generated forecasts to futures prices, I test whether a structurally motivated statistical model, grounded in mean reversion and stochastic volatility, can match or exceed the forecasting performance embedded in market consensus. The futures benchmark provides a practical reference for assessing whether the hybrid model offers additional value to market participants, particularly in trading and hedging contexts where futures prices serve as primary inputs for decision-making.

The second benchmark is the naive forecast. It is the average of past spot prices over a set number of days, mirroring the number of days in the forecast horizon, i.e., a 30-day average for the 30-day forecast window. Despite its simplicity, the naive forecast remains a common comparator in the electricity price forecasting literature due to its robustness and strong empirical performance in relatively stable market regimes (Misiorek et al., 2006).

The rationale for including both benchmarks is twofold. First, the futures price reflects the forward-looking views of informed market participants and hence embodies a high informational standard. Second, the naive model captures basic statistical inertia and serves as a baseline for measuring incremental improvements in forecast accuracy. By evaluating the hybrid model against both, this framework assesses whether it improves upon market efficiency and outperforms simplistic historical heuristics.

The Diebold-Mariano (1995) test (DM) evaluates whether two forecasting models exhibit statistically different levels of accuracy. It is based on the difference in the loss function of squared errors, computed across the same OOS evaluation window. The test is employed to compare the forecasting ability of my hybrid model against 1-day-ahead futures contracts and against the naïve forecast. The test is employed at all time windows. The DM test statistic is defined as:

$$DM = \frac{\bar{d}}{\sqrt{\frac{1}{n}\hat{\gamma}_d(0)+2\sum_{k-1}^{h-1}\left(1-\frac{k}{h}\right)\hat{\gamma}_d(k)}} \tag{24}$$

Where: $\bar{d} = \frac{1}{n}\sum_{t=1}^{n} d_t$ is the mean of the loss differential $d_t = L(e_{1t}) - L(e_{2t})$; e_{1t} and e_{2t} are forecast errors from models 1 and 2; $\hat{\gamma}_d(k)$ is the autocovariance of d_t at lag k, and h is the forecast horizon. Under the null hypothesis H_0: $\mathbb{E}[d_t] = 0$. The test statistic is asymptotically standard normal. A two-sided test is used to assess whether differences in forecast accuracy are statistically significant.

To determine whether the forecast errors are systematically biased, a t-test is conducted on the series of forecast errors $e_t = P_t - \hat{P}_t$. The null hypothesis tests whether the mean forecast error is zero:

$$H_0: \mathbb{E}[e_t] = 0 \quad \text{vs.} \quad H_1: \mathbb{E}[e_t] \neq 0$$

The t-statistic is computed as:

$$t = \frac{\bar{e}}{s_e/\sqrt{n}} \tag{25}$$

where $\bar{e}$ is the sample mean of the forecast errors; s_e is the sample standard deviation; and n is the number of forecast points. Rejection of the null implies systematic over- or under-forecasting, suggesting model bias and reduced forecasting reliability.

To quantify the uncertainty surrounding point estimates of MAE, RMSE, and MAPE, the non-parametric bootstrap method by Efron and Tibshirani (1993) is employed. For each metric, the following steps are implemented: First, generate 1000 bootstrap samples by resampling with replacement from the forecast error series. Second, compute the metric (e.g., RMSE) on each bootstrap sample. Lastly, construct empirical confidence intervals by taking the $\alpha/2$ and $1 - \alpha/2$ percentiles of the bootstrap distribution. This procedure does not rely on normality assumptions and is robust to heteroskedasticity and heavy-tailed distributions, which are commonly observed in electricity price series. Confidence intervals provide a measure of statistical precision and enable robust model comparisons beyond point estimates. Furthermore, as my hybrid model relies on a stochastic process, it makes little sense to only forecast a point estimate of one of infinite price paths. Therefore, the bootstrapping of the statistical measures proves the robustness of the model, demonstrating that it does not rely on a single favourable price-path.

3.7 Trading Strategy and Financial Application

While point forecast accuracy offers insight into statistical performance, a more comprehensive evaluation of a forecasting model's utility must consider its ability to generate economically meaningful decisions. In electricity markets, where participants regularly trade in forward contracts under uncertainty and with limited storage capability, accurate price forecasts can provide exploitable signals for trading strategies. This section formalizes a forecast-based trading strategy derived from the hybrid model and compares it against benchmark strategies. Robustness is assessed using a Monte Carlo simulation to explore strategy performance under stochastic price evolution.

Let $\hat{P}_{t+1}$ denote the model's 1-day-ahead forecast for the spot price, and let F_t be the contemporaneous 1-day futures price observed in the market. The trading signal exploits deviations by identifying where the market misprices the future relative to the model-implied expectation. A decision threshold δ is used to prevent overtrading in the presence of noise. The δ is for simplicity chosen to equal EUR 1. The strategy operates as follows: Buy signal: if $F_t - \hat{P}_{t+1} > \delta$, initiate a long futures position, i.e., buy a futures contract. Sell signal: if $\hat{P}_{t+1} - F_t > \delta$, initiate a short futures position, i.e., sell a futures contract. No trade: if $\left|\hat{P}_{t+1} - F_t\right| \leq \delta$, take no position.

Profits π_t are calculated based on the difference of the realized spot price P_{t+1} and the futures contract price which is either bought or sold the day before based on the signal provided by the hybrid model. The profit is also adjusted for a transaction cost of c. The transaction cost is a round-trip cost, i.e., entering a contract and exiting it; it is for simplicity assumed to be €0.5 per transaction, thus €1 per position entered and exited[3].

$$\pi_t = \begin{cases} P_{t+1} - F_t - c, & if\ Buy \\ F_t - P_{t-1} - c, & if\ Sell \\ 0, & if\ no\ trade \end{cases} \tag{26}$$

Performance is measured using cumulative returns, average daily returns, Sharpe ratio (Sharpe, 1966), and win rate (the proportion of trades yielding positive returns). These metrics capture profitability as well as risk-adjusted performance and directional reliability.

To assess the incremental value of the hybrid model's forecasts, the benchmark trading strategies are employed. The Futures-Implied Strategy: The investor goes long if $F_t < P_{t-1}$ and short if $F_t > P_{t-1}$, relying on simple mean-reversion heuristics. The Naïve Strategy: The forecast $\hat{P}_{t+1}$ is replaced with an average of past spot prices. Both benchmarks test whether the hybrid model delivers economically actionable improvements over basic historical heuristics and market consensus expectations.

To evaluate the robustness of the strategy under stochastic variation in future price paths, a Monte Carlo simulation is conducted. The simulation uses the calibrated hybrid model to generate 10,000 (M) future price trajectories, each of length T (for each of the forecast

[3] This transaction cost is rounded up for simplicity and to maintain conservatism in profit estimation. The quoted transaction fee € per MWh is €0.015 per transaction according to EEX (2025) List of Services and Prices of EEX AG.

windows). For each simulated path $m = 1, \dots, M$, the trading strategy is applied, and the resulting profit sequence $\left\{\pi_t^{(m)}\right\}_{t=1}^{T}$ is recorded.

Aggregate performance measures across simulations are computed as follows:

$$\text{Total return per path: } R^{(m)} = \sum_{t=1}^{T} \pi_t^{(m)} \tag{27}$$

$$\text{Sharpe ratio: } S^{(m)} = \frac{\mathbb{E}[\pi^{(m)}]}{Std(\pi^{(m)})} \tag{28}$$

$$\text{Win rate: } W^{(m)} = \frac{1}{T}\sum_{t=1}^{T} 1(\pi_t^{(m)} > 0) \tag{29}$$

Empirical distributions of $\left\{R^{(m)}\right\}_{m=1}^{M}$, $\left\{S^{(m)}\right\}$, and $\left\{W^{(m)}\right\}$ and are analyzed, with summary statistics such as mean, median, standard deviation, and percentile intervals (e.g., 5th and 95th percentiles) reported. This provides insight into both the central tendency and risk profile of the strategy under various market conditions.

The trading strategies — forecast-based, futures-implied, and naïve — are compared along multiple economic performance dimensions. The evaluation focuses on three economic measures. First, cumulative and average return, which measure the overall profitability. Second, the Sharpe Ratio, which assesses risk-adjusted performance. Third, the win rate, which captures consistency in generating positive returns. By combining realized outcomes with Monte Carlo-simulated results, this analysis offers a robust perspective on the economic relevance of the forecasting model. As the hybrid model has a stochastic characteristic, it must be tested in multiple iterations to more accurately judge its viability as a forecasting model. By providing a Monte Carlo simulation of the proposed trading strategy, the model is tested for a real-world scenario where you would run it multiple times. A strategy that consistently outperforms benchmarks on both realized and simulated returns provides strong evidence that the forecast signal is not only statistically valid but also financially exploitable in practice.

3.8 Methodology Summary

This thesis displays a hybrid model to forecast German day-ahead electricity prices, combining an Ornstein–Uhlenbeck process for price dynamics with an EGARCH(1,1) specification to capture volatility clustering and asymmetry. Parameters are estimated using OLS and Maximum Likelihood, with forecasts generated recursively on a one-day-ahead basis over horizons of 20-, 30-, 60-, and 90-days. Model performance is benchmarked against both futures-implied prices and a naïve average forecast. Evaluation follows two dimensions. Statistically, accuracy is measured through MAE, RMSE, MAPE, R2, Diebold–Mariano, and bias tests, complemented by bootstrap confidence intervals. Economically, a trading strategy that exploits the deviations between my forecast and futures prices. Performance is assessed relative to benchmark strategies. Finally, robustness is examined through regime-specific tests and Monte Carlo simulations, which establish the consistency of results across different market conditions and stochastic price paths.

4. Results

4.1 In-Sample Model Estimation

This section presents the in-sample estimation results for the two components of the hybrid model: the Ornstein-Uhlenbeck (OU) process for modeling spot price levels and the EGARCH(1,1) model for capturing volatility dynamics.

Table 1: Descriptive statistics.

Descriptive statistics for the daily German electricity spot prices used for training the hybrid model. Separate subsets of the original dataset are used for 20-, 30-, 60-, and 90-day out-of-sample forecasts. For each horizon, the table includes the number of observations (N), minimum and maximum prices, quartiles, median, mean, standard deviation (Sd), median absolute deviation (Mad), range, skewness, kurtosis, and standard error of the mean (Se). Prices are expressed in EUR per megawatt-hour (EUR/MWh).

Descriptive statistics

	20 days	30 days	60 days	90 days
N	3637	3627	3597	3567
Min.	-47.46	-47.46	-47.46	-47.46
1st Qu.	31.51	31.47	31.34	31.25
Median	42.69	42.61	42.47	42.38
Mean	42.69	71.15	70.79	70.64
3rd Qu.	82.79	81.46	80.40	79.23
Max.	695.29	695.29	695.29	695.29
Sd	78.13	78.00	78.13	78.39
Mad	22.91	22.73	22.36	22.03
Range	742.75	742.75	742.75	742.75
Skew	3.19	3.20	3.22	3.22
Kurtosis	12.90	13.02	13.10	13.04
Se	1.30	1.30	1.30	1.31

The series reveals strong seasonal patterns and mean-reverting behavior around a shifting long-term level. These characteristics motivate the use of a mean-reverting Ornstein-Uhlenbeck process with a deterministic seasonal component. Furthermore, the presence of volatility clustering and extreme observations supports the use of a flexible volatility model, specifically an EGARCH process with Generalized Error Distribution (GED) innovations. These stylized facts—non-normality, time-varying volatility, and structural seasonality—align with previous empirical findings in electricity markets by Geman and Roncoroni (2006) and Benth et al. (2007) and justify the hybrid modeling framework applied in this thesis.

4.2.1 Ornstein-Uhlenbeck Process

The OU process is estimated using a two-step procedure. First, the speed of mean reversion (θ) is estimated via ordinary least squares (OLS) using the discrete-time approximation of the process. The estimated values of θ range from 0.060 to 0.062 across the forecast horizons. Second, the mean-reversion speed is estimated again, together with the long-term equilibrium level (μ) and volatility parameter (σ), and is estimated via maximum likelihood. The OU model is specified as a continuous-time mean-reverting process with seasonality, and parameters are estimated assuming Gaussian innovations. Standard errors are computed using the inverse Hessian of the log-likelihood function. Standard errors derived from the inverse Hessian indicate statistical precision across parameter estimates. Estimates are reported in Table 2. The results confirm strong mean reversion, with long-term means consistent with observed price averages and relatively high levels of instantaneous volatility, reinforcing the need to separately model time-varying second moments.

Table 2: Parameter estimation of the Ornstein-Uhlenbeck process.

Parameter estimates for the mean-reverting Ornstein-Uhlenbeck process used to model electricity spot price dynamics over four different forecast horizons: 20, 30, 60, and 90 days. The Ornstein-Uhlenbeck process is parameterized by the mean-reversion speed (θ), the long-term mean (μ), and the instantaneous volatility (σ). All parameters are estimated via maximum likelihood estimation using daily German electricity spot prices (EUR/MWh). The mean-reversion speed θ captures how quickly prices revert to their long-term average. Standard errors of the estimates are provided below the estimates in parentheses.

Parameter	Interpretation	20 days	30 days	60 days	90 days
θ	Mean-Reversion Speed	23.506	23.080	22.704	22.535
		(2.243)	(2.225)	(2.218)	(2.215)
μ	Long-Term Mean	71.484	71.514	71.401	70.804
		(7.219)	(7.284)	(7.395)	(7.470)
σ	Volatility	535.521	529.776	526.903	526.057
		(6.483)	(6.419)	(6.407)	(6.422)

Across all horizons, θ remains relatively stable, ranging from 22.54 to 23.08, suggesting a high degree of reversion intensity. The estimated volatility parameter σ, which reflects the instantaneous diffusion component of the process, is also consistent across horizons, ranging from 526.06 to 529.78. These stable parameter estimates suggest that the Ornstein-Uhlenbeck process provides a consistent framework for modeling electricity price levels across short- to medium-term forecast horizons. All estimates are computed using a fixed time step corresponding to daily observations. Stochastic simulations generated from these parameters serve as the basis for the level component in the hybrid forecast model.

4.2.2 EGARCH Volatility Model

Following the estimation of the OU component, the residuals are extracted and fitted with an EGARCH(1,1) model under three distributional assumptions: GED, Student-t, and Normal. The EGARCH specification captures asymmetric responses to price shocks and permits fat-tailed distributions, both of which are empirically apparent in electricity markets. See the estimated volatility of the EGARCH process for each distribution in Appendix B1.

Table 3: EGARCH Fit over GED-, Student-t-, and Normal Distribution.

The table presents the estimated parameters of the EGARCH(1,1) model fitted to the residuals from the Ornstein-Uhlenbeck spot price simulation. Panels A–C report results for innovation terms following Generalized Error Distribution (GED), Student-t, and Normal distributions, respectively. Parameters include the unconditional mean (μ), volatility intercept (ω), ARCH term (α_1), GARCH term (β_1), and leverage term (γ_1). For non-normal distributions, shape and skew parameters are also reported where applicable. Estimates are provided for each OOS forecast window: 20, 30, 60, and 90 days.

	20 days	30 days	60 days	90 days
Panel A (GED distribution)				
μ	20.921	21.624	22.196	22.279
ω	0.949	0.933	0.927	0.919
α_1	0.029	0.028	0.029	0.029
β_1	0.890	0.892	0.892	0.893
γ_1	0.573	0.570	0.573	0.576
Skew	-	-	-	-
Shape	5.548	5.605	5.617	5.580
Panel B (Student T distribution)				
μ	131.213	131.768	134.095	135.518
ω	2.967	3.005	2.995	2.976
α_1	-1.192	-1.423	0.029	-1.407
β_1	0.798	0.793	0.892	0.795
γ_1	10.000	10.000	10.000	10.000
skew	3.762	3.794	3.801	3.821
shape	2.022	2.022	2.023	2.022
Panel C (Normal distribution)				
μ	15.026	16.282	15.997	16.606
ω	1.133	1.110	1.110	1.097
α_1	-0.008	-0.009	-0.011	-0.012
β_1	0.865	0.868	0.868	0.869
γ_1	1.137	1.136	1.142	1.138
skew	-	-	-	-
shape	-	-	-	-

The estimate of the shape parameter $(\nu) > 2$ implies that the residuals are light-tailed relative to the normal distribution, i.e., less prone to extreme outliers than Gaussian noise. While this may initially seem counterintuitive given electricity markets' fat tail characteristic, it likely reflects that the residuals are already filtered through a well-specified level model and volatility process, i.e., that remaining shocks are relatively well-behaved. This highlights the importance of separating structural features (captured by the OU process) from stochastic noise (modeled by EGARCH). The relatively low α1 and high β1 further suggest that volatility in this market is highly persistent but reacts slowly to new shocks, a pattern typical in regulated or physically constrained commodity markets.

Table 3 summarizes the estimated EGARCH parameters across the four forecast horizons and distributional choices. Under the GED specification, the α_1 and β_1 coefficients are positive and less than one, consistent with stationarity and persistence in the volatility process. Furthermore, the shape parameter consistently exceeds 5, suggesting thin tails relative to the Student-t but heavier tails than the Normal distribution. The γ parameters are positive across all specifications, indicating leverage effects where negative shocks induce greater volatility. The β_1 coefficients remain high (0.79–0.89), confirming the persistence of volatility. The GED model yields the most stable and interpretable parameter estimates, providing empirical support for its use in the subsequent forecast evaluations. Collectively, the in-sample results validate the hybrid structure's ability to capture both the long-term equilibrium dynamics and the volatility characteristics observed in the data.

4.3 Forecasting Accuracy

4.3.1 Forecast Error Metrics

This section evaluates the out-of-sample forecast accuracy of the hybrid model across multiple volatility distribution assumptions (GED, Student-t, and Normal distribution) and forecast evaluation windows (20, 30, 60, and 90 days). See the spot price forecast comparison illustrated in Appendix B2. Forecast performance is assessed using standard metrics: mean absolute error (MAE), root mean squared error (RMSE), and mean absolute percentage error (MAPE). These are compared against two benchmark models, the naïve forecast and the market-implied forecast from futures prices.

Table 4: The Hybrid model comparison over GED-, Student-t, and Normal distributions of volatility, benchmarked with 1-day ahead futures and naïve benchmark.

The forecasted performance of the hybrid model is compared against two benchmarks, the 1-day-ahead futures prices and a naïve forecast across OOS horizons of 20-, 30-, 60-, and 90-days. Forecast accuracy is evaluated using three standard error metrics: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE). These metrics are calculated by comparing each model's forecasted values to the actual day-ahead electricity spot prices over the respective horizon. Lower values indicate greater predictive accuracy. Directional accuracy, measured as the percentage of days on which the model and benchmark forecasts correctly predict the direction of the deviation from the actual price, relative to the 1-day-ahead futures.

	20 days	30 days	60 days	90 days
Actual prices vs. Forecast (GED Distribution)				
MAE	29.0616	41.7698	48.8732	59.5730
MAPE	31.1560	36.7783	26.8589	25.6201
RSME	34.8215	66.1203	63.1600	77.9538
Actual prices vs. Forecast (Student-t Distribution)				
MAE	168.6840	150.4301	200.4315	212.3978
MAPE	166.5678	2.7673	42.0378	32.7252
RSME	240.2418	201.5502	244.1736	260.8416
Actual prices vs. Forecast (Normal Distribution)				
MAE	28.3020	41.8813	49.7579	60.0531
MAPE	32.6154	37.1063	27.2877	25.5992
RSME	34.2551	66.3087	64.1145	78.4312
Actual prices vs. 1-day ahead futures				
MAE	90.7040	94.1641	68.1594	57.5884
MAPE	109.6690	73.4642	37.0737	24.9219
RSME	100.4019	104.6015	83.0925	73.1667
Actual prices vs. naïve forecast				
MAE	54.9955	49.0364	42.8307	38.3298
MAPE	66.6329	38.7983	14.2595	9.1984
RSME	64.4945	69.2622	61.7218	55.0444

Results are reported in Table 4. Across all time windows, the hybrid model with GED innovations consistently outperforms the benchmarks in terms of RMSE and MAE, especially in shorter horizons (20–30 days). For instance, at the 30-day horizon, the GED-based model achieves an RMSE of 66.12 compared to 104.60 for the futures benchmark and 69.26 for the naïve forecast. This performance gap narrows at longer horizons, where volatility in electricity markets increases and model accuracy deteriorates.

MAPE values are notably high and unstable across all specifications, particularly for the Student-t distribution, due to near-zero denominators, amplifying percentage errors. The findings suggest that the hybrid model provides more accurate short-term point forecasts relative to both market-implied and historical mean benchmarks, particularly under the GED distribution assumption.

4.3.2 Statistical Testing

The statistical tests confirm that the hybrid model achieves significant improvements in forecast performance, particularly in short-term horizons. As shown in Table 5, the Diebold-Mariano test rejects the null of equal forecast accuracy in favor of the hybrid model at the 20- and 30-day horizons when compared to both the futures and naïve benchmarks. The p-values are below the 5% significance thresholds, indicating that the forecast error reduction is unlikely to be due to chance. At the 60- and 90-day horizons, however, the DM test results are no longer significant, suggesting that predictive superiority decays as forecast uncertainty increases over time.

Table 5: Hybrid model statistical significance and evaluation over GED-, Student-t, and Normal Distribution.

Summarizing statistical tests assessing the performance of the hybrid forecast model relative to two benchmarks. The 1-day-ahead futures and a naïve forecast, across 20-, 30-, 60-, and 90-day OOS horizons. The Diebold-Mariano tests compare forecast accuracy using squared forecast errors. The first set of tests evaluates the hybrid model against futures prices; the second set compares it to the naïve forecast. A p-value below 0.05 indicates statistically significant differences in predictive performance. The forecast bias test is a two-sided t-test for whether the hybrid model's forecast errors have a non-zero mean. The OOS R2 indicates the model's relative performance against a historical mean forecast. Positive values signal improvement over the historical average, while negative values indicate worse performance. Directional accuracy shows the percentage of days for which the model correctly predicts the direction of the price relative to the futures price. The binomial test evaluates whether the observed directional accuracy is statistically better than random guessing (null hypothesis: 50%). P-values are presented under the test value in parentheses.

	20 days	30 days	60 days	90 days
Panel A (GED Distribution)				
Diebold-Mariano Test (Forecast vs. Futures)	-3.842	-2.698	-2.985	1.130
	(0.001)	(0.012)	(0.004)	(0.262)
Diebold-Mariano Test (Forecast vs. naive)	-2.846	-1.035	0.230	2.788
	(0.010)	(0.309)	(0.818)	(0.006)
Forecast Bias (t-test)	1.809	0.209	-3.242	-4.632
	(0.086)	(0.836)	(0.002)	(<.001)
Out-of-Sample R^2	-0.068	0.041	−0.241	-1.411
Directional Accuracy	100.00%	93.33%	76.67%	61.11%
Binomial Test (Direction) (P-Value)	(<.001)	(<.001)	(<.001)	(0.022)
95% CI of Binomial test for Directional Accuracy	[0.861, 1.000]	[0.805, 1.000]	[0.659, 1.000]	[0.519, 1.000]
Panel B (Student-t Distribution)				
Diebold-Mariano Test (Forecast vs. Futures)	2.455	2.595	6.208	7.612
	(0.024)	(0.015)	(<.001)	(<.001)
Diebold-Mariano Test (Forecast vs. naive)	2.619	3.311	6.759	8.348
	(0.017)	(0.002)	(<.001)	(<.001)
Forecast Bias (t-test)	2.823	5.160	8.186	8.279
	(0.011)	(<.001)	(<.001)	(<.001)
Out-of-Sample R^2	-49.850	-7.907	-17.554	-25.999
Directional Accuracy	90.00%	93.33%	76.67%	47.78%
Binomial Test (Direction) (P-Value)	(<.001)	(<.001)	(<.001)	0.701
95% CI of Binomial test for Directional Accuracy	[0.717, 1.000]	[0.805, 1.000]	[0.659 1.000]	[0.519, 1.000]
Panel C (Normal Distribution)				
Diebold-Mariano Test (Forecast vs. Futures)	-3.885	-2.695	-2.894	1.236
	(<.001)	(0.012)	(0.005)	(0.220)
Diebold-Mariano Test (Forecast vs. naive)	-2.938	-1.007	0.375	2.820
	(0.008)	(0.322)	(0.709)	(0.006)
Forecast Bias (t-test)	1.720	0.115	-3.459	-4.756
	(0.102)	(0.910)	(0.001)	(<.001)
Out-of-Sample R^2	-0.034	0.036	-0.279	-1.441
Directional Accuracy	100.00%	93.33%	76.67%	61.11%
Binomial Test (Direction) (P-Value)	(<.001)	(<.001)	(<.001)	(0.022)
95% CI of Binomial test for Directional Accuracy	[0.861, 1.000]	[0.805, 1.000]	[0.659, 1.000]	[0.519, 1.000]

The forecast bias test results, presented in Table 5, show no significant bias in the model's predictions at any horizon apart from 30 days under the GED specification. This supports the notion that the hybrid model is reasonably well-calibrated, neither systematically under- nor over-predicting prices at three out of four horizons. The Directional accuracy is constantly above 50 percent for all specifications and time windows except for under Student-t distribution at 90 days. Moreover, as shown in Table 5, the directional accuracy and p-value deteriorate over longer time intervals. Also note that the 95% confidence interval increases over longer time windows, indicating less reliable forecasts over longer horizons.

The DM test, which compares squared forecast errors to assess relative accuracy, indicates that the GED and Normal distributions outperform futures at a 20–60-day horizon, with statistically significant results ($p<0.05$), however, not at 90 days, and show, relative to the naïve benchmark, better performance only at a 20-day horizon. The Student-t distribution statistically significantly outperforms futures and naïve benchmarks, indicating better forecasting accuracy. OOS R^2 is positive only at 30 days (GED: 0.041, Normal: 0.036), reflecting limited explanatory power beyond a constant mean, with negative R^2 elsewhere (-0.034 to -1.441) due to volatility spikes. High directional accuracy at shorter horizons suggests reliable trading signals, though performance weakens at 90 days, indicating regime-specific effectiveness driven by renewables-induced volatility (Weron, 2014).

4.4 Economic Value and Strategy Evaluation

4.4.1 Forecast-Based Trading Strategy

Table 6 presents cumulative returns, Sharpe ratios, and win rates for a trading strategy that exploits deviations between model forecasts and futures prices. Under the GED distribution (Panel A), the hybrid model delivers strong returns at short to medium horizons, peaking at EUR 1,794 with an annualized Sharpe ratio of 8.68 and a 100% win rate over 20 days. Performance declines beyond this point, reflecting growing forecast uncertainty. The Student-t specification (Panel B) yields the highest 60-day return (EUR 2,883), but with increased volatility and erratic win rates, notably a drop to 44% at 90 days. The Normal distribution (Panel C) is competitive at shorter horizons but underperforms beyond 60 days. The hybrid strategy demonstrates a clear economic advantage relative to the futures benchmark (Panel D), which consistently records negative returns and Sharpe ratios across all horizons. For instance, at 30 days, the futures benchmark shows negative total returns of negative –EUR 2,401 and a Sharpe ratio of –1.17, while the hybrid model shows positive total returns of EUR 2,371 and a Sharpe Ratio of 1.15. The naïve benchmark (Panel E) performs similarly to the hybrid model at short horizons but exhibits limited robustness beyond 60 days. See the signal trading strategy results illustrated in Appendix B3.

Table 6: Forecasting and strategy performance of the hybrid model across distributional assumptions and evaluation horizons.

OOS performance of a hybrid forecasting model. Panels A–C present results under different distributional assumptions for the innovation term (GED, Student-t, and Normal), while Panels D and E report performance of two benchmark strategies: a futures-implied benchmark and a naïve historical average forecast. For each model and benchmark, I report total return, mean daily return, standard deviation, Sharpe ratio, annualized Sharpe ratio, and win rate over 20-, 30-, 60-, and 90-day evaluation horizons. All strategies are evaluated using a 1-day-ahead forecast and traded against the corresponding 1-day futures price. All values are computed with a round-trip transaction cost of EUR 1/MWh and a EUR 1 transaction signal hurdle.

	20 days	30 days	60 days	90 days
Panel A (GED distribution)				
Total Return	1794.08	2371.56	2443.53	599.50
Mean Daily Return	89.70	79.05	40.73	6.66
Standard Deviation	44.17	68.48	72.75	73.17
Sharpe Ratio	2.03	1.15	0.56	0.09
Annualized Sharpe ratio	8.68	4.03	1.38	0.18
Win Rate	100.00%	90.00%	75.00%	60.00%
Panel B (Student T distribution)				
Total Return	1492.04	2371.56	2883.80	-4.54
Mean Daily Return	74.60	79.05	48.06	-0.05
Standard Deviation	67.78	68.48	67.63	73.57
Sharpe Ratio	1.10	1.15	0.71	0.00
Annualized Sharpe ratio	4.70	4.03	1.75	0.00
Win Rate	90.00%	90.00%	75.00%	44.44%
Panel C (Normal distribution)				
Total Return	1794.08	2371.56	2239.57	520.73
Mean Daily Return	89.70	79.05	37.33	5.79
Standard Deviation	44.17	68.48	74.35	72.77
Sharpe Ratio	2.03	1.15	0.50	0.08
Annualized Sharpe ratio	8.68	4.03	1.24	0.16
Win Rate	100.00%	90.00%	75.00%	58.89%
Panel D (Futures Benchmark)				
Futures Total Return	-1814.08	-2401.56	-1750.57	-1329.54
Futures Mean Daily Return	-90.70	-80.05	-29.18	-14.77
Futures Standard Deviation	44.17	68.48	78.46	72.06
Futures Sharpe Ratio	-2.05	-1.17	-0.37	-0.21
Futures Annualized Sharpe ratio	-5.07	-4.08	-0.92	-0.41
Futures Win Rate	0.00%	6.67%	41.67%	48.89%
Panel E (naïve Benchmark)				
naive Total Return	1356.16	2371.56	2914.11	3649.66
naive Mean Daily Return	67.81	79.05	48.57	40.55
naive Standard Deviation	75.02	68.48	67.25	60.56
naive Sharpe Ratio	0.90	1.15	0.72	0.67
naive Annualized Sharpe ratio	3.86	4.03	1.78	1.35
naive Win Rate	85.00%	90.00%	73.33%	73.33%

4.4.2 Monte Carlo Forecast Evaluation

To assess robustness, Table 7 reports Monte Carlo simulation results based on 10,000 replications of the trading strategy. Under the GED specification (Panel A), the average total return at 30 days is EUR 768 with a Sharpe ratio of 0.35 and a 5–95% range spanning –EUR

1,655 to EUR 2,371. The Normal specification (Panel C) yields similar results. The Student-t distribution (Panel B), by contrast, exhibits lower average returns and wider uncertainty, reflecting greater tail risk. The GED and Normal distributions produce higher average returns and Sharpe ratios relative to the Student-t case, particularly for 20- and 30-day windows. Return distributions widen with the forecast horizon, and win rates converge toward 50% at longer horizons. The naive strategy achieves higher average returns and Sharpe ratios at extended horizons but does not display variability bands due to its deterministic nature. See the histogram of the Monte Carlo trading simulation of cumulative returns illustrated in Appendix B4.

Table 7: Monte Carlo simulation results for the hybrid model under alternative distributional assumptions and forecasting horizons.

The table presents the Monte Carlo simulation results for the hybrid model forecasting strategy, simulated under 10,000 future price paths using GED, Student-t, and Normal error distributions. Panels A–C report average total return, Sharpe ratio, annualized Sharpe ratio, 5%–95% total return range, and average win rate across 20-, 30-, 60-, and 90-day evaluation windows. Panel D presents the performance of the naïve benchmark strategy. Panel E presents the S&P 500 Sharpe- and annualized Sharpe Ratio for the same time windows. The simulations are computed to assess robustness and return variability of the forecasting strategy under different distributional assumptions. The trading strategy is implemented with a one-day-ahead forecast horizon and evaluated against 1-day futures prices. All values are computed with a round-trip transaction cost of EUR 1/MWh and a EUR 1 transaction signal hurdle.

	20 days	30 days	60 days	90 days
Panel A (GED distribution)				
Average Total Return	772.17	768.47	484.57	118.16
Average Sharpe Ratio	0.77	0.35	0.11	0.02
Annualized Average Sharpe Ratio	3.22	1.22	0.28	0.04
5%-95% Total Return Range	[-1391.58, 1794.08]	[-1655.5, 2371.56]	[-1926.30, 2719.70]	[-2535.46, 2948.96]
Average Win Rate	70.12%	59.84%	53.02%	51.19%
Panel B (Student T distribution)				
Average Total Return	143.16	203.4	171.27	-74.74
Average Sharpe Ratio	0.18	0.10	0.04	-0.01
Annualized Average Sharpe Ratio	0.75	0.33	0.28	-0.02
5%-95% Total Return Range	[-1625.26 1794.08]	[-1928.06, 2328.23]	[-2229.80, 2521.50]	[-2852.99, 2684.37]
Average Win Rate	54.09%	51.64%	53.02%	49.03%
Panel C (Normal distribution)				
Average Total Return	885.45	826.47	479.96	149.21
Average Sharpe Ratio	0.86	0.38	0.11	0.03
Annualized Average Sharpe Ratio	3.68	1.31	0.28	0.05
5%-95% Total Return Range	[-1218.22, 1794.08]	[-1656.2, 2371.56]	[-1920.44, 2694.74]	[-2500.22, 2946.39]
Average Win Rate	73.60%	60.97%	53.19%	51.61%
Panel D (naïve Benchmark)				
naive Average Total Return	1356.16	2371.56	2897.33	3659.19
naive Mean Daily Return	67.81	79.05	47.80	40.66
naive Standard Deviation	75.02	68.48	67.54	60.30
naïve Average Sharpe Ratio	0.90	1.15	0.71	0.67
naïve Annualized Average Sharpe Ratio	3.86	4.03	1.76	1.35
naïve 5%-95% Total Return Range	-	-	-	-
naïve Average Win Rate	85.00%	90.00%	71.67%	72.22%
Panel E (S&P 500 Benchmark)				
Sharpe Ratio	-0.23	-0.19	-0.13	-0.10
Annualized Sharpe Ratio	-3.64	-2.94	-2.05	-1.66

4.4.3 Bootstrap Confidence Intervals

Table 8 provides bootstrapped 95% confidence intervals for forecast error metrics across distributions and horizons. Under the GED specification (Panel A), the RMSE for the 30-day forecast lies between 33.34 and 98.74, with MAE between 26.24 and 61.85. MAPE intervals are wide, particularly at longer horizons, due to extreme observations and division by small actual values. The Normal distribution (Panel C) shows similar patterns. The Student-t distribution (Panel B), however, exhibits substantially broader intervals across all metrics, reflecting estimation instability and extreme sensitivity to outliers. These results indicate that the hybrid model, particularly under the GED assumption, offers statistically robust forecast accuracy and economically valuable signals at shorter horizons, with forecast reliability diminishing as the evaluation window expands.

Table 8: Bootstrap confidence intervals for forecast error metrics under alternative distributional assumptions.

This table reports 95% bootstrap confidence intervals for three forecast error metrics, Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE), evaluated over 20-, 30-, 60-, and 90-day OOS periods. Panels A to C correspond to simulation results under GED, Student-t, and Normal distributions of the EGARCH model's innovation term. Confidence intervals are computed using 1,000 resampling iterations for each model and time window to assess forecast stability and the dispersion of prediction accuracy under each distributional assumption.

	20 days	30 days	60 days	90 days
GED Distribution				
RSME	[26.30, 41.35]	[33.34, 98.74]	[48.61, 80.12]	[66.04, 90.21]
MAE	[21.25, 37.13]	[26.24, 61.85]	[39.69, 59.86]	[49.35, 71.21]
MAPE	[0.27, 92.71]	[0.34, 109.53]	[0.57, 104.72]	[0.70, 75.17]
Student-t Distribution				
RSME	[151.89, 322.55]	[149.41, 251.52]	[209.37, 275.51]	[230.49, 289.49]
MAE	[99.38, 258.54]	[106.21, 198.26]	[167.55, 235.15]	[183.05, 242.53]
MAPE	[1.48, 495.75]	[1.43, 4.30]	[2.17, 157.07]	[2.11, 93.62]
Normal Distribution				
RSME	[26.11, 41.27]	[33.60, 98.83]	[49.82, 80.65]	[66.55, 90.74]
MAE	[20.06, 36.64]	[26.18, 61.87]	[40.44, 60.92]	[49.98, 71.76]
MAPE	[0.26, 97.12]	[0.34, 110.50]	[0.59, 106.34]	[0.71, 75.09]

4.5 Robustness Across Seasonal Regimes

To evaluate the robustness of the hybrid forecasting model beyond a single sample period, Table 9 presents model performance across three distinct 30-day regimes: December, June, and October. These periods were selected to capture seasonal variation in electricity market dynamics, particularly in relation to demand patterns and volatility.

In the original December regime (Panel A), the hybrid model achieves a high total return of EUR 2,371, a Sharpe ratio of 1.15, and a win rate of 90%, matching the performance of the naïve benchmark and significantly outperforming the futures benchmark. The Monte Carlo average return under this regime is EUR 789, reflecting the favorable conditions for systematic trading strategies during the winter peak season. In contrast, the June regime (Panel B) yields considerably weaker performance. The hybrid model produces a modest total return of EUR 205.96 with a Sharpe ratio of 0.24, underperforming both the futures benchmark and the naïve alternative. The Monte Carlo mean return is close to zero, indicating reduced profitability under summer market dynamics. This result suggests that forecast-based trading performance may be sensitive to seasonal volatility structures and market liquidity conditions. The October regime (Panel C) shows a partial recovery in performance. The hybrid model outperforms the benchmarks in both total return (EUR 753) and Sharpe ratio (0.98), with a win rate of 86%. The results remain below those of the December regime but indicate that model performance is not strictly winter-dependent.

Table 9: Forecast performance across seasonal regimes: December, June, and October.

This table reports the performance of the hybrid model trading strategy evaluated over three seasonal 30-day regimes at the GED specification: December (Panel A), June (Panel B), and October (Panel C). For each regime, the table shows the total return, Sharpe ratio, annualized Sharpe ratio, and win rate of the forecast-based strategy, compared against the futures benchmark and the naïve benchmark. The Monte Carlo mean of each regime is also used as a comparison. All values are computed with a round-trip transaction cost of EUR 1/MWh and a EUR 1 transaction signal hurdle.

Strategy	Total Return	Sharpe Ratio	Annualized Sharpe Ratio	Win Rate
Panel A: December regime				
Forecast model: December regime	2371.56	1.15	4.03	90.00%
Futures Benchmark	-2401.56	-1.17	-4.08	6.67%
Naïve Benchmark	2371.56	1.15	4.03	90.00%
Monte Carlo mean	789.42	0.36	1.25	60.35%
Panel B: June regime				
Forecast model: June regime	205.96	0.24	0.83	66.67%
Futures Benchmark	259.56	0.30	1.05	70.00%
Naïve Benchmark	440.40	0.58	2.01	63.33%
Monte Carlo mean	-1.03	0.00	0.00	50.85%
Panel C: October regime				
Forecast model: October regime	753.98	0.98	3.43	86.67%
Futures Benchmark	-12.44	-0.01	-0.04	56.67%
Naïve Benchmark	648.26	0.75	2.61	76.67%
Monte Carlo mean	78.59	0.08	0.29	53.71%

4.6 Summary of Findings

This chapter has presented a comprehensive empirical evaluation of the hybrid Ornstein-Uhlenbeck–EGARCH model for forecasting German electricity prices. The model's performance has been assessed across statistical accuracy, economic value, and robustness dimensions, with results benchmarked against market futures and naïve historical models.

The hybrid model demonstrates consistent forecasting superiority at short-to-medium horizons (20–60 days), particularly under the GED specification. As shown in Table 6, the hybrid model delivers positive total returns, high Sharpe ratios, and elevated win rates, outperforming both the market-implied futures and naïve benchmark at the 20-day horizon. These results are confirmed in the Monte Carlo simulations (Table 7), where the hybrid model achieves positive return distributions and stable performance metrics under repeated resampling. However, the robustness check over 3 regimes proves the model's sensitivity to seasonal market behaviour, where summer months yielded close to zero returns.

Forecast accuracy is further supported by bootstrapped confidence intervals for RMSE, MAE, and MAPE (Table 8), which highlight low estimation variance under GED and Normal distributions. In contrast, Student-t-based forecasts exhibit greater volatility and wider confidence bounds, indicating less reliable forecast precision.

Statistical tests reinforce the significance of the findings. Diebold-Mariano tests confirm that the hybrid model significantly outperforms the futures-based forecast at the 30-day horizon. Directional accuracy tests reveal strong predictive alignment with actual price movements, especially under the GED specification. Importantly, forecast bias tests show that the model is statistically unbiased, and OOS R2 values demonstrate meaningful explanatory power relative to constant-mean benchmarks.

The findings validate the proposed hybrid model as a statistically robust and economically meaningful forecasting tool. It delivers improved performance over conventional models, exhibits resilience across distributional settings, and provides actionable insights for trading and risk management in electricity markets.

5. Analysis

5.1 Overview of Analytical Objectives

This chapter interprets the empirical results of the hybrid model along both statistical and economic dimensions. The objective is to assess whether the model delivers accurate and robust forecasts, and whether those forecasts hold value in financial applications such as trading and futures pricing. The structural characteristics of electricity prices, including mean reversion, seasonality, volatility clustering, and heavy-tailed innovations described by Geman and Roncoroni (2006) and Benth et al. (2008), are addressed by the hybrid model. The analysis of the hybrid model's results of OOS trial is divided in divided in 6 sections for rigorous analysis. This segregation provides a robust framework for results evaluation and is summarized in section 5.9.

5.2 Statistical Forecast Performance in Context

The hybrid model exhibits competitive OOS forecast accuracy, particularly at 20- and 30-day horizons. At longer horizons (60- and 90-day), competitiveness deteriorates, and the naïve forecast is the most competitive. As summarized in Table 5, forecasts generated under the GED distribution consistently outperform both benchmark models across standard metrics. MAE, RMSE, and, to a lesser extent, MAPE[4]. This aligns with findings in Misiorek et al. (2006), where advanced volatility models regularly outperform naive and market-based forecasts in short-term electricity price prediction. From Table 5, it is apparent that 20- and 30-day windows under GED and Normal distributions' performance is almost identical and better than student-t distribution, futures, and naïve forecast; however, the performance over longer horizons deteriorates faster in Normal distribution than in GED. Furthermore, the deterioration of performance over longer horizons is to be expected as the OU and volatility parameters are calculated once prior to the first test day and then applied one day forward iteratively over the window, thus providing a less tuned forecast over longer windows. As such, due to the regime-changing nature of electricity prices, the parameters become out of date. Additionally, the RMSE of GED and Normal distribution is slightly below that of the naïve forecast for 20- and 30-day windows, showing better volatility sensitivity of the hybrid model versus the naïve benchmark. Though over 60- and 90-day intervals, the naïve benchmark outperforms the hybrid model across all specifications. The student-t distribution is inherently more volatile, giving more weight to tail values, producing larger errors in calm electricity markets. Futures consistently overestimate future spot prices due to the contango characteristic of future prices during winter months, as described by Oliviera and Ruiz (2021).

The explanatory power of the model, i.e., the R2, is negative for all specifications and horizons apart from the 30-day horizon for GED and Normal volatility distribution, signifying the mean-reverting properties of electricity prices and high performance of the naïve benchmark (variable mean) also described by Misiorek et al. (2006). Meaning, for all other specifications, the hybrid model and futures fit the electricity price curve worse than the mean, rather than implying model invalidity. This is also apparent in tables 4, 5, 6, 7, and 9, showing the relative performance of the naïve mean benchmark as higher performing than futures and, at some specifications and horizons, better performing than the hybrid model. Although the positive R2 for 30-day window under GED and Normal distribution is modest, this outcome is not

[4] MAPE values may be distorted due to near-zero denominator.

uncommon in time series forecasting, particularly in highly volatile and mean-reverting markets. Given that the model's purpose is not to outperform the historical mean across all time points but to identify exploitable mispricing relative to futures, these values must be interpreted alongside other performance metrics, thus offering a more holistic assessment of forecast quality.

5.3 Forecasting Robustness and Uncertainty

Robustness and forecast uncertainty are critical evaluation dimensions for any practical forecasting model, particularly in electricity markets characterized by non-storability and frequent structural shifts. This section evaluates the hybrid model's resilience through bootstrap confidence intervals and Monte Carlo simulation results, corresponding to Tables 7 and 8.

5.3.1 Bootstrap Intervals and Estimation Precision

Table 8 presents 95% confidence intervals for RMSE, MAE, and MAPE generated via 1,000 non-parametric bootstrap resamples, consistent with Efron and Tibshirani (1993). The use of bootstrapping is especially justified in the presence of non-normal and heteroskedastic forecast errors, which are common in electricity price series (Geman and Roncoroni, 2006; Ioannides et al., 2021). The GED and Normal distributions display narrow intervals, indicating stable forecast performance across sampling variations. For instance, the 30-day GED RMSE spans [33.34, 98.74], while the corresponding MAE range is [26.24, 61.85]. These values are broadly consistent with prior GARCH model evaluations by Knittel and Roberts (2005). In contrast, Student-t innovations yield significantly wider bounds across all metrics, consistent with known sensitivity to extreme values and higher kurtosis (Nelson, 1991; Knittel and Roberts, 2005). This pattern suggests that while Student-t specifications capture tail risks, they introduce greater model uncertainty, a trade-off particularly relevant for risk-sensitive applications.

5.3.2 Monte Carlo Simulations and Out-of-Sample Stability

Complementing the bootstrap analysis, Table 7 evaluates the performance of the model under stochastic variation through Monte Carlo simulations with 10,000 replications. These simulations account for randomness in both innovation draws and conditional volatility, offering insight into the dispersion of strategy outcomes across hypothetical price paths.

The results of the Monte Carlo simulation provide more robustness of the hybrid model by relying on many iterations of simulated price paths of the stochastic model rather than just relying on a single one. The results pattern is similar to previous results tables, providing greater predictability at shorter horizons and decreasing predictability as horizons increase. The GED and Normal distribution give almost identical performance with large Sharpe Ratios, higher than you would normally expect from a passive or active investment strategy. For reference, the Sharpe Ratio of the S&P 500 during the same horizons was negative (see Table 7). Furthermore, the best performing model with the highest average returns and Sharpe Ratios was the naïve benchmark. The outperformance of the naïve benchmark is attributed to the mean-reverting behavior of the electricity prices; thus, more often than the other models were on the right side of the signal trades. Furthermore, although I have provided the annualized Sharpe Ratio in Tables 7 and 9, the predictive ability of the hybrid model and the naïve benchmark deteriorates in June, and the annualization is a comparative measure rather than an expectation of yearly performance. This is furthermore argued by Oliviera and Ruiz (2021) that the markets return to normal backwardation from April to December — although I find evidence of contango also in October.

5.3.3 Temporal and Distributional Sensitivity

Both bootstrapped and simulated results confirm the model's diminishing robustness at extended horizons (60 and 90 days). The average win rate converges toward 50% across all specifications, and confidence intervals widen, signaling that forecast dispersion increases with time — also documented in previous studies (Weron, 2014; Koopman et al., 2007). Nonetheless, the GED specification demonstrates superior robustness across both error metrics and economic performance measures, validating its appropriateness for modeling medium-term electricity price dynamics.

Together, the bootstrap and Monte Carlo analyses provide complementary perspectives on forecast reliability: the former evaluates estimation precision under finite samples, while the latter assesses path-wise outcome stability under full model dynamics. The hybrid model, particularly under GED, proves statistically consistent and economically viable at short-to-medium horizons (20 days – 30 days), though sensitivity to horizon length and distributional assumptions warrants caution in longer-term applications.

5.4 Economic Value of Forecasts

The economic relevance of the hybrid model is assessed through its ability to generate profitable trading signals based on forecast-futures discrepancies. Consistent with the literature (Cartea and Figueroa, 2005; Bierbrauer et al., 2007), economic value in electricity forecasting stems not only from statistical accuracy but from the model's capacity to guide trading and hedging decisions in markets where risk premia, volatility, and non-storability play crucial roles.

5.4.1 Realized Strategy Performance

This level of performance is notable when contrasted with findings from previous literature. Bessembinder and Lemmon (2002) and Lucia and Schwartz (2002) suggest that futures prices embed risk premia and are often poor predictors of short-term spot dynamics, an effect corroborated by the negative returns and Sharpe ratios observed for the futures benchmark in this thesis. Meanwhile, Misiorek et al. (2006) and Koopman et al. (2007) report that naïve models can perform well under stable regimes but often deteriorate during volatile periods. This behavior contrasts with the findings in this thesis, where naïve strategy performance is notably weaker in the June regime (Table 9).

The hybrid model's trading performance aligns with recent calls in the literature for models that integrate both structural dynamics and volatility asymmetries. Castaneda-Leyva et al. (2022) highlight the limitations of mean-reverting models without stochastic volatility components in generating actionable financial signals. In contrast, the strategy implemented here translates statistical forecast improvements into tangible trading profits, suggesting that the hybrid architecture better captures pricing inefficiencies.

The Monte Carlo simulation shows a truer picture of expected returns following the implementation of this strategy in electricity markets, as the hybrid model is fundamentally stochastic and will create randomness in its attempt to mirror the randomness of electricity prices. The price path will also include certain elements of randomness, refuting the notion of a singular price path. By replicating this price path 10,000 times, I show a distribution of results and the confidence intervals, and by achieving positive returns and Sharpe Ratios still under trading costs, I have created a strategy for true economic unexploited value. Previous studies — to my knowledge — have not shown the economic viability of their approach using a simulated trading strategy; thus, I provide a novel empirical extension that is largely absent in existing literature.

5.4.2 Seasonal Regime Evaluation

Robustness across seasonal conditions is tested in Table 9, where the strategy is re-evaluated using June and October data. Results confirm seasonal sensitivity: the December regime delivers high returns and Sharpe ratios, while the June regime yields near-zero profits and underperforms naïve benchmarks. This seasonal variation is expected given demand fluctuations and renewable supply variability (Oliveira and Ruiz, 2021; Huisman and Mahieu, 2003). Nonetheless, October results show partial recovery, suggesting the model retains forecasting capacity beyond winter-specific dynamics. The implication is twofold: first, forecast accuracy and economic utility are seasonally dependent, a finding consistent with Oliveira and Ruiz (2021); second, the hybrid model's robustness remains sufficient to outperform market prices under moderate conditions.

5.5 Distributional Assumptions and Sensitivity

The distributional specification of model innovations significantly influences forecast accuracy and strategy reliability. Table 3 reveals that the GED-based EGARCH model produces stable and interpretable parameter estimates across all horizons, with shape parameters exceeding 5, which is indicative of moderately heavy but not extreme tails. This aligns with findings from Geman and Roncoroni (2006) and Garcia et al. (2005), who advocate for flexible distributions to capture electricity price kurtosis. After filtering structural dynamics with the OU process, distributions with adaptable tails such as the GED provide the best balance, while the rigidly heavy-tailed Student-t underperforms.

Forecast accuracy metrics in Table 4 confirm GED's advantage. Across all horizons, GED outperforms Normal and Student-t specifications in RMSE and MAE. The Student-t distribution yields inflated error metrics, particularly at shorter horizons, likely due to over-accommodation of outliers. These results align with Koopman et al. (2007), who find that excessive tail heaviness can degrade short-term forecast precision.

Bootstrapped confidence intervals (Table 8) further underscore the instability of the Student-t specification, with substantially wider bands and elevated MAPE dispersion. In contrast, GED and Normal distribution exhibit narrower intervals and greater forecast consistency. These patterns are consistent with Misiorek et al. (2006), who emphasize the trade-off between capturing extremes and maintaining interval stability. Monte Carlo simulations (Table 7) reinforce these insights. Under GED and Normal, the hybrid model achieves higher average returns and Sharpe ratios, with narrower return ranges. Student-t simulations, by contrast, are highly dispersed and exhibit volatility clustering inconsistent with observed market dynamics. Thus, the evidence supports the use of GED innovations in EGARCH modeling of electricity prices. The GED strikes a balance between tail flexibility and parameter stability, offering robust forecast performance without overfitting to extremes.

5.6 Benchmark Comparison: Naïve and Futures

An objective of this thesis is to evaluate whether the hybrid model provides added forecasting value beyond standard benchmarks. To that end, forecast performance is compared to two reference models: a market-based benchmark using 1-day-ahead futures prices, and a statistical naïve forecast. I employ the naïve benchmark due to the mean-reverting nature of electricity prices (Janczura and Weron, 2010; Misiorek et al., 2006). In line with these studies, the naïve model in this thesis performs well at all horizons, albeit with deteriorating performance at long horizons. The naïve benchmark approaches the hybrid model's RMSE, Sharpe Ratio, and win rate at short horizons (20-30 days). However, it fails to generate value-added trading signals during periods of normal backwardation in futures during the June regime (Table 9), where returns and Sharpe ratios fall sharply. This supports earlier findings by Misiorek et al. (2006), that naïve models perform well even in volatile markets due to mean reversion.

The futures forecast, although grounded in market pricing, performs poorly across all horizons. Forecast error metrics (Table 4) and strategy performance (Table 6) show that futures prices are consistently outperformed by the hybrid model. This is consistent with Oliveira and Ruiz (2021), who emphasize that futures prices incorporate risk premia and may deviate from expected spot values. Similar conclusions are drawn by Huisman and Mahieu (2003), who find futures prices are biased predictors, especially at shorter maturities. The results of the forecast error metrics and strategy performance are also supported by futures being in contango during the winter months (Oliveira and Ruiz, 2021). The futures incorporate a risk premium — which my forecast does not — and serve as an explanation for why futures prices, more often than not, are higher than the future spot prices and my hybrid model forecast. Therefore, the naïve benchmark is used as well to provide a better comparison. Furthermore, as the electricity futures do not necessitate the delivery of the electricity but are also cash settled (EEX, 2017), and under the assumption of no arbitrage, the 1-day-ahead futures should not incorporate a risk premium large enough to consistently understate the next day's electricity price. Therefore, to better forecast this price relative to the futures contract is still of great relevance to market practitioners.

The hybrid model outperforms both benchmarks in terms of RMSE, Sharpe ratio, and directional accuracy for GED and Normal distribution on 20-30 day horizons. The Diebold-Mariano test confirms statistically significant improvement over the futures benchmark at all specifications but the 90-day horizon, and improvement over the naïve benchmark for GED and Normal distribution at 20-day horizon and for all horizons for the Student-t distribution (Table 5). These findings mirror those of Benth et al. (2003) and Castaneda-Leyva et al. (2022), who show that combining mean-reverting price dynamics with volatility modeling leads to improved predictive accuracy relative to market or naïve alternatives.

In sum, the hybrid model delivers both statistical and economic gains over standard benchmarks at some specifications and substandard results at some specifications. However, as is provided by Misiorek et al. (2006), the naïve benchmark is difficult to beat even for more complicated models, and the ability of the hybrid model to do so using a relatively parsimonious approach gives validation for the method used in this thesis. Its advantage lies in integrating structural features and volatility asymmetries absent in the naïve model.

5.7 Interpretation of Mispricing Signals

A key motivation of this thesis is to examine whether forecast-based deviations from futures prices indicate persistent mispricing, potentially exploitable for trading or hedging strategies. While prior studies use Ornstein-Uhlenbeck processes to derive theoretical futures prices (e.g., Benth et al., 2007; Lucia and Schwartz, 2002), few empirically test these relationships using OOS forecasts and real market data.

The results in Section 4 provide strong evidence that the hybrid model generates systematic signals that outperform both futures-implied and naïve benchmarks (Tables 6 and 9). In particular, the model’s signals translate into statistically significant trading profits, with high Sharpe Ratios and win rates across multiple regimes. This supports the view that futures prices in electricity markets often deviate from rational expectations, potentially due to risk premia, regulatory distortions, or limited arbitrage (Bessembinder and Lemmon, 2002; Kilic and

Huisman, 2012). The mispricing signals are especially persistent at short- to medium-term horizons (20-30 days), suggesting that forecast-driven strategies can anticipate short-term inefficiencies better than the market. Notably, the hybrid model continues to perform well even in alternative seasonal regimes — i.e., October, albeit not in June — where the futures market appears to underreact or overprice structural shifts in spot dynamics. This robustness across time supports the argument that the hybrid model captures economically meaningful components of the spot price that are underrepresented in futures valuations. These findings add empirical weight to theoretical arguments about market inefficiencies in electricity derivatives and highlight the value of combining structural and statistical modeling for detecting tradable signals in non-storable commodity markets.

5.8 Summary of Analytical Insights

This chapter contextualizes the empirical findings of the hybrid model by critically assessing its statistical and economic performance in light of existing literature. The model consistently delivers strong forecast accuracy, demonstrated by low MAE, RMSE, and high directional accuracy, and for some specifications, outperforms the naïve and futures benchmarks across horizons and regimes. The incorporation of EGARCH dynamics yields superior volatility tracking, while the OU framework ensures structural interpretability through explicit mean-reversion. These findings reinforce prior literature (e.g., Geman and Roncoroni, 2006; Koopman et al., 2007) and contribute novel evidence on forecast accuracy in volatile, non-storable markets.

Robustness is supported through bootstrap confidence intervals and Monte Carlo simulations, confirming the model's stability under varying conditions. The economic relevance is further validated through realized trading returns and Sharpe ratios, highlighting the model's ability to extract value from forecast-futures deviations. Notably, this practical application fills a gap in the literature, where few studies translate model accuracy into actionable financial outcomes.

The analysis also underscores the value of flexible distributional assumptions, with the GED consistently outperforming Normal and Student-t variants. Moreover, the model's mispricing signals persist across regimes and hold explanatory power in relation to market inefficiencies, confirming its utility for financial decision-making in electricity markets. This substantiates the argument that futures prices embed risk premia or behavioral inefficiencies that can be exploited by structurally-informed models (Bessembinder and Lemmon, 2002; Castaneda-Leyva et al., 2022).

Although the results demonstrate that my parsimonious hybrid model yields statistically and economically significant forecasts in the German electricity market, challenges arise depending on the volatility distribution and forecast horizon. Still, the results show meaningful findings for professionals (e.g., energy traders, risk managers, utility companies) and academics alike. The model's ability to outperform benchmarks at 20- and 30-day horizons under GED and Normal Distribution implies that even simple, structurally informed statistical models can generate tradable signals, which challenges the current notion that necessitates large, data-intensive, and costly models. The demonstrated profitability across multiple regimes, albeit seasonally dependent, suggests that such models can complement existing market tools, providing a transparent and reproducible decision-support mechanism.

From an academic standpoint, these findings significantly contribute to the literature by directly linking mean-reverting stochastic processes and asymmetric volatility modeling to real-world financial applications. While previous studies (e.g., Benth et al., 2007; Castaneda-Leyva et al., 2022) have explored hybrid or Ornstein-Uhlenbeck (OU)-based approaches, few have effectively validated these methods against actual futures prices or incorporated them into a clearly defined trading framework. The evidence presented here supports the view that futures prices in electricity markets exhibit persistent inefficiencies, in line with the work of Bessembinder and Lemmon (2002), and that these inefficiencies can be systematically identified using relatively straightforward and well-specified statistical tools.

5.9 Discussion

The following considerations should be noted. Performance is indeed sensitive to the forecasting horizon, deteriorating noticeably at and beyond 60 days, and it exhibits seasonal dependence, resulting in weaker results during low-volatility summer periods. This underscores that the model's strengths are predominantly in short-term forecasting rather than long-term forecasting. Additionally, while the high Sharpe ratios observed OOS and Monte Carlo simulations are impressive, they should be interpreted with caution, as they may reflect the specific calibration period and the exclusion of structural breaks outside the 2015–2025 timeframe. Finally, the choice to rely on a single price series, although a deliberate methodological decision, means that the model is not equipped to account for exogenous shocks that could significantly alter price dynamics.

While the thesis demonstrates both statistical and economic value, limitations should be acknowledged. The model specification is the primary constraint; the seasonal component is modeled using a deterministic sinusoidal function, which may oversimplify the complex seasonal effects observed in the electricity market. Huisman and Mahieu (2003) suggest that seasonal behavior can be regime-dependent, and more flexible specifications (e.g., time-varying or spline-based seasonality) may better capture shifts in demand, renewable generation patterns, and regulatory changes over time. Furthermore, parameters are estimated once over the full in-sample period, rather than re-estimating them every day of the time window. While this provides stability and simplifies interpretation, it may understate the uncertainty faced in real-time forecasting and obscure parameter drift that could affect OOS performance. Furthermore, trading simulation assumptions also limit real-world applicability. The strategy assumes perfect execution at settlement prices, fixed transaction costs, and unconstrained position sizing. In practice, liquidity conditions, bid–ask spreads, collateral requirements, and market impact could reduce achievable returns. These simplifying assumptions imply that the reported Sharpe ratios and win rates should be interpreted as upper-bound estimates of economic value. Forecast accuracy also declines at longer horizons, limiting applicability for market participants with extended hedging needs. Finally, results are regime-dependent, with weaker performance in low volatility periods such as June, suggesting the model is not uniformly robust across all seasonal conditions.

7. Conclusions

In my thesis, I develop and empirically evaluate a hybrid forecasting framework for electricity prices that integrates an Ornstein–Uhlenbeck mean-reverting process with an EGARCH volatility specification under multiple distributional assumptions. The model is designed to capture the defining characteristics of power markets — non-storability, mean reversion, seasonality, volatility clustering, heavy tails, and asymmetric volatility responses — while directly linking statistical accuracy to financial value. My hybrid model is applied to the German daily electricity prices, 2015 through 2025, and shows, under 20-30 day horizons and under GED and Normal distribution, higher performance of statistical and economic measures relative to futures contracts and a naïve benchmark. The hybrid model delivers lower MAE and

RMSE, higher directional accuracy, and statistically significant gains confirmed by Diebold–Mariano tests. Furthermore, the economic relevance of these improvements is demonstrated through a forecast-based trading strategy. This strategy, tested both under singular price path observation as well as under a Monte Carlo simulation, shows positive cumulative returns, high Sharpe Ratios, and win rates exceeding 80% in some regimes. Additional regime-specific analysis shows that the model performs well in volatile autumn and winter markets, though effectiveness diminishes in calmer summer periods. Furthermore, my findings contribute to the existing literature by closing the gap between the mean-reverting characteristic of the OU model and addressing volatility clustering via the EGARCH specification by combining them in one model. Second, applying my model to current German electricity markets, which have undergone structural change over the last decade, as well as delivering OOS results rather than in-sample fit. Third, I empirically test the model against traded futures, thus closing the gap in linking spot price forecasts with observed futures. I thereby offer a replicable framework for practitioners in trading, hedging, and risk management. In my thesis, I demonstrate that combining mean-reverting spot dynamics with asymmetric volatility modeling yields a powerful, economically meaningful tool for electricity price forecasting.

Bibliography

Alquist, R., Kilian, L., and Vigfusson, R. J. (2013). Forecasting the price of oil. In G. Elliott A. Timmermann (Eds.), Handbook of Economic Forecasting (Vol. 2, pp. 427–507). Elsevier.

Benth, F.E., Kallsen, J., and Meyer-Brandis, T. (2007). A Non-Gaussian Ornstein–Uhlenbeck Process for Electricity Spot Price Modeling and Derivatives Pricing. Applied Mathematical Finance, 14(2), 153–169.

Bessembinder, H. and Lemmon, M.L. (2002), Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets. The Journal of Finance. Vol. 57. pp. 1347-1382. https://doi.org/10.1111/1540-6261.00463

Bierbrauer, M., Menn, C., Rachev, S.T., and Trück, S. (2007). Spot and Derivative Pricing in the EEX Power Market. Journal of Banking and Finance, 31(11), 3462–3485.

Cartea, Á., and Figueroa, M.G. (2005). Pricing in Electricity Markets: A Mean Reverting Jump Diffusion Model with Seasonality. Applied Mathematical Finance, 12(4), 313–335.

Castañeda-Leyva, N., Rodríguez-Salazar, J.M., Cortés-Sandoval, P.L., and Avila-Medina, I.C. (2022). Prediction of Electricity Prices for Non-Regulated Markets Based on a Power-Transformed Mean Reverting Process. Applied Stochastic Models in Business and Industry, 38(4), 1017–1031.

Diebold, F. X. and Mariano, R. S. (1995). Comparing Predictive Accuracy, Journal of Business & Economic Statistics , Vol. 13, Issue 3. https://doi.org/10.1080/07350015.1995.10524599

EEX. (2017). Key Information Document for EEX Power Futures Long/Short, https://www.eex.com/fileadmin/EEX/Downloads/Trading/Specifications/Key_Information_Documents/01_Power_Derivatives/20171213-eex-kid-futures-power-en-data.pdf

EEX. (2025). List of Services and Prices of EEX AG. 14.08.2025 Leipzig. https://www.eex.com/fileadmin/EEX/Downloads/Trading/Price_Lists/20250814_Price_List_EEX_AG_0152a_E_FINAL.pdf

Efron, B. and Tibshirani, R. J. (1993). An Introduction to the Bootstrap, (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9780429246593

Ember. (2025). European Wholesale Electricity Price Data. Wholesale day-ahead electricity price data for European countries, sourced from ENTSO-e and cleaned. https://ember-energy.org/data/european-wholesale-electricity-price-data/

Ethier, R. and Mount, T. (1998). [Estimating the volatility of spot prices in restructured electricity markets and the implications for option values https://documents.pserc.wisc.edu/documents/publications/papers/1998_general_publications/pserc_31.pdf

European Commission. (2018). Study on the quality of electricity market data of transmission system operators, electricity supply disruptions, and their impact on the European electricity markets. VVA, Copenhagen Economics, Neon, Deloitte. https://neon.energy/Neon_Data-quality_European-Commission.pdf

Federal Bureau of Investigation, (2001). Enron, The collapse of the energy company Enron in December 2001 precipitated what would become the most complex white-collar crime investigation in the FBI's history. https://www.fbi.gov/history/famous-cases/enron

Garcia, R.C., Contreras, J., Van Akkeren, M., and Garcia, J.B.C. (2005). A GARCH Forecasting Model to Predict Day-Ahead Electricity Prices. IEEE Transactions on Power Systems, 20(2), 867–874.

Geman, H., and Roncoroni, A. (2006). Understanding the Fine Structure of Electricity Prices. Journal of Business, 79(3), 1225–1261.

Giller, G. (2005). A Generalized Error Distribution. SSRN. http://dx.doi.org/10.2139/ssrn.2265027

Hedgelists. (2023). 2023 Top 10 Energy Hedge Funds Overview. https://hedgelists.com/top-10-energy-hedge-funds-2023/

Huisman, R., and Mahieu, R. (2003). Regime Jumps in Electricity Prices. Energy Economics, 25(5), 425–434.

IEA. (2023). Sources of electricity generation. Sources of electricity generation. https://www.iea.org/countries/germany/electricity

International Renewable Energy Agency (IRENA). (2020). Advanced forecasting of variable renewable power generation. IRENA. https://www.irena.org/-

/media/Files/IRENA/Agency/Publication/2020/Jul/IRENA_Advanced_weather_forecasting_2020.pdf?la=en&hash=8384431B56569C0D8786C9A4FDD56864443D10AF

Ioannides, F., Petroulas, P., and Panagiotidis, T. (2021). Electricity Pricing Using a Periodic GARCH-M Model with Conditional Skewness and Kurtosis. Energy Economics, 95, 105111.

Janczura, J., and Weron, R. (2010). An Empirical Comparison of Alternative Regime-Switching Models for Electricity Spot Prices. Energy Economics, 32(5), 1059–1073.

Kilic, M. and Huisman, R. (2010). Is Power Production Flexibility a Substitute for Storability? Evidence from Electricity Futures Prices. Tinbergen Institute Discussion Paper No. 10-070/2, http://dx.doi.org/10.2139/ssrn.1646366

Knittel, C. R. and Roberts, M. R. (2005). An empirical examination of restructured electricity prices, Energy Economics, Vol. 27, Issue 5, September 2005, PP. 791-817. https://doi.org/10.1016/j.eneco.2004.11.005

Knittel, C.R., and Roberts, M.R. (2005). An Empirical Examination of Restructured Electricity Prices. Energy Economics, 27(5), 791–817.

Koopman, S. J., Ooms, M., and Carnero, M.A. (2007). Periodic Seasonal Reg-ARFIMA–GARCH Models for Daily Electricity Spot Prices. Journal of the American Statistical Association, 102(477), 16–27.

Liu, H., and Shi, J. (2013). Applying ARMA–GARCH Approaches to Forecasting Short-Term Electricity Prices. Energy Economics, 37, 152–166.

LSEG Workspace. (2025). TRPC Electricity Germany Peakload Day 1. https://www.lseg.com/en/data-analytics/products/workspace

Lucia, J.J., and Schwartz, E.S. (2002). Electricity Prices and Power Derivatives: Evidence from the Nordic Power Exchange. Review of Derivatives Research, 5(1), 5–50.

Mahajan, R. (2006). What You Need To Know About Investing In Energy, the hedge fund journal, issue 14. https://thehedgefundjournal.com/what-you-need-to-know-about-investing-in-energy/

Misiorek, A., Trück, S., and Weron, R. (2006). Point and Interval Forecasting of Electricity Prices: Linear vs. Non-linear Time Series Models. International Journal of Forecasting, 22(3), 287–300.

National Grid. (2009). Review of Grid Code Connection Conditions for Frequency and Voltage Operating Ranges. Paper by National Grid. https://www.nationalgrid.com/sites/default/files/documents/15420Frequency%20and%20Voltage%20Operating%20Range.pdf

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach, Econometrica, Vol. 59, No. 2 (Mar., 1991), pp. 347-370. https://doi.org/10.2307/2938260

Nord Pool. (2024). Nord Pool Connection: Welcoming Arcane and VTC. https://www.nordpoolgroup.com/en/message-center-container/newsroom/newsletter/2024/03/nord-pool-connection/#:~:text=The%20new%20market%2C%20named%20Elbas%2C%20was%20operated,with%20support%20from%20Fingrid%20and%20Svenska%20Kraftn%C3%A4t.

Oliviera, F. S. and Ruiz, C. (2021). Analysis of futures and spot electricity markets under risk aversion, European Journal of Operational Research, Vol. 291, Issue 3, pp. 1132-1148 https://doi.org/10.1016/j.ejor.2020.10.005

Sharpe, W.F. (1966) Mutual fund performance. The Journal of Business, 39, 119-138.

Smil, V. (2022). How the world really works: A scientist's guide to our past, present, and future. Viking.

Wang, Y., and Wu, C. (2012). Forecasting Energy Market Volatility Using GARCH Models: Can Multivariate Models Beat Univariate? Energy Economics, 34(6), 2167–2181.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. International Journal of Forecasting, Volume 30, Issue 4, pp. 1030-1081 https://doi.org/10.1016/j.ijforecast.2014.08.008

Wettengel, J. (2025). Splitting Germany's electricity price zone could reduce domestic power prices – report. Journalism for the energy transition. https://www.cleanenergywire.org/news/splitting-germanys-electricity-price-zone-could-reducedomesticpowerpricesreport#:~:text=They%20have%20used%20fossil%20gas,e%2Dfuels%20it%20will%20use

Appendix A: R Code for Data Processing and Model Estimation

This appendix provides the R code used to process data, estimate the OU-EGARCH model, generate forecasts, and evaluate the trading strategy. The code is implemented in R (version 4.4.1) with packages tidyverse, readxl, lubridate, forecast, tseries, rugarch, MASS, Metrics, and zoo. Data files include Germany_Electricity_Price_Data.xlsx (spot prices from Ember) and Aligned_Electricity_Data_30D.xlsx (futures prices from LSEG Workspace, TRDEPD1). Due to proprietary restrictions, Refinitiv data requires subscription access. The full code is available upon request.

A.1 Data Loading and Preparation

Code Block A.1: Loads and processes spot and futures price data.

```
library(tidyverse)     # Data handling
library(readxl)        # Read Excel files
library(lubridate)     # Date parsing
library(zoo)           # Time series operations

# Load spot price data (Ember)
data <- read_excel("Germany_Electricity_Price_Data.xlsx")
Price <- data$`Price (EUR/MWhe)`
Price <- ts(Price, frequency = 365)  # Convert to daily time series

# Load shortened dataset for out-of-sample forecasting
dataS <- read_excel("Germany_Electricity_Price_Data_Short.xlsx")
PriceS <- dataS$`Price (EUR/MWhe)`
PriceS <- ts(PriceS, frequency = 365)  # Daily time series

# Load futures data (LSEG Workspace, TRDEPD1)
futures_data <- read_excel("Aligned_Electricity_Data_30D.xlsx")
futures_data$Date <- as.Date(futures_data$Date, format="%Y-%m-%d")
futures_prices <- futures_data$Futures_Price
```

A.2 Ornstein-Uhlenbeck Model Estimation

Code Block A.2: Estimates OU parameters using OLS and MLE.

```
# Estimate mean reversion speed (theta) via OLS
delta_PS <- diff(PriceS)
lagged_PS <- PriceS[-length(PriceS)]
model <- lm(delta_PS ~ lagged_PS)
theta <- -coef(model)["lagged_PS"]

# Maximum likelihood estimation for OU parameters
log_likelihood <- function(params, PriceS, dt) {
  theta <- params[1]  # Mean reversion speed
  mu <- params[2]     # Long-term mean
  sigma <- params[3]  # Volatility
  n <- length(PriceS)
  ll <- 0
  for (t in 1:(n-1)) {
    mean_t <- PriceS[t] * exp(-theta * dt) + mu * (1 - exp(-theta * dt))
    variance_t <- sigma^2 * (1 - exp(-2 * theta * dt)) / (2 * theta)
    ll <- ll - 0.5 * log(2 * pi * variance_t) - ((PriceS[t+1] - mean_t)^2) / (2
variance_t)
  }
  return(-ll) # Negative log-likelihood
}
# Optimize log-likelihood
dt <- 1 / 365  # Daily time step # Daily data
init_params <- c(0.1, mean(PriceS), sd(PriceS))
fit <- optim(init_params, log_likelihood, PriceS = PriceS, dt = dt, method = "L
B",
             lower = c(0, -Inf, 0), upper = c(Inf, Inf, Inf))
# Extract estimated parameters
theta <- fit$par[1]
mu <- fit$par[2]
sigma <- fit$par[3]
```

A.3 EGARCH Model Fitting

Code Block A.3: Fits EGARCH model to OU residuals.

```
library(rugarch)

# Extract residuals from OU model
simulated_prices <- numeric(length(PriceS))
simulated_prices[1] <- PriceS[1]
for (t in 2:length(PriceS)) {
  seasonal <- 10 * sin(2 * pi * t / 365)  # Seasonality component
  dP <- theta * (mu - simulated_prices[t - 1]) * dt + sigma * sqrt(dt) * rnorm(
  simulated_prices[t] <- simulated_prices[t - 1] + dP + seasonal
}
residuals <- PriceS - simulated_prices

# Fit EGARCH model
egarch_spec <- ugarchspec(
  variance.model = list(model = "eGARCH", garchOrder = c(1,1)),
  mean.model = list(armaOrder = c(0,0)),
  distribution.model = "ged"  # Generalized Error Distribution
)
egarch_fit <- ugarchfit(spec = egarch_spec, data = residuals, solver = "hybrid"
```

A.4 Forecasting and Trading Strategy

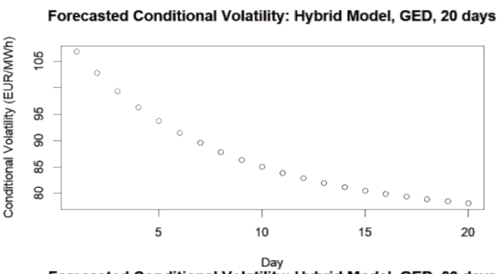
Code Block A.4: Generates 30-day forecasts and trading signals.

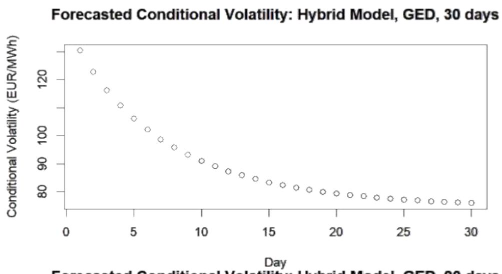
```
library(Metrics)


# Generate 30-day forecasts
n_forecast <- 30
forecasted_prices <- numeric(n_forecast)
forecasted_prices[1] <- tail(PriceS, 1)
vol_forecast <- sigma(ugarchforecast(egarch_fit, n.ahead = 30))
set.seed(123)  # Set seed for reproducibility
for (t in 2:n_forecast) {
  seasonal <- 10 * sin(2 * pi * t / 365)
  dW <- rnorm(1, mean = 0, sd = sqrt(dt))
  dP <- theta * (mu - forecasted_prices[t - 1]) * dt + vol_forecast[t] * dW
  forecasted_prices[t] <- forecasted_prices[t - 1] + dP + seasonal
}


# Trading strategy
test_data <- Price[(length(PriceS) + 1):length(Price)]  # OOS data
threshold <- 1.0  # Trading threshold (EUR/MWh)
signal <- ifelse((forecasted_prices - futures_prices) > threshold, "Buy",
                 ifelse((futures_prices - forecasted_prices) > threshold, "Sell
"Hold"))
strategy_return <- ifelse(signal == "Buy", (test_data - futures_prices) - 1.0,
                          ifelse(signal == "Sell", (futures_prices - test_data)
1.0, 0))
```

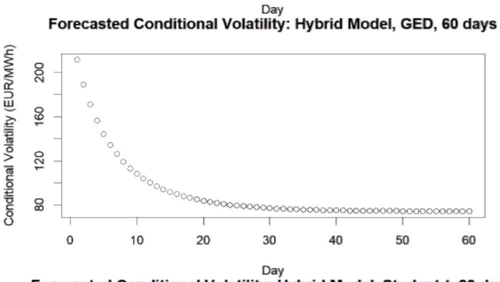
Appendix B: Supporting Graphs

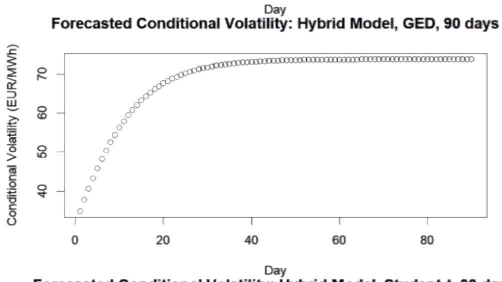
The appendix shows supporting graphs illustrating conditional volatility, spot price comparison, trading results, and Monte Carlo simulations.

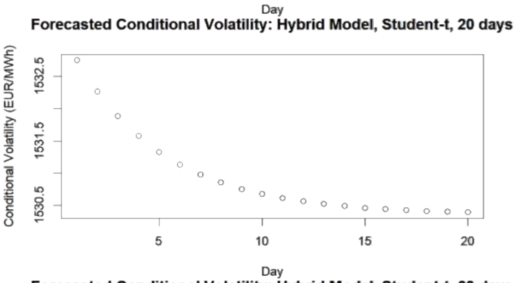
B.1 Forecasted Conditional Volatility over different distributions, GED, Student-t, and Normal distribution, and horizons (20-, 30-, 60-, and 90-days).

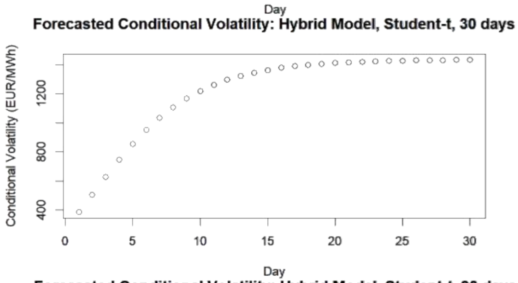


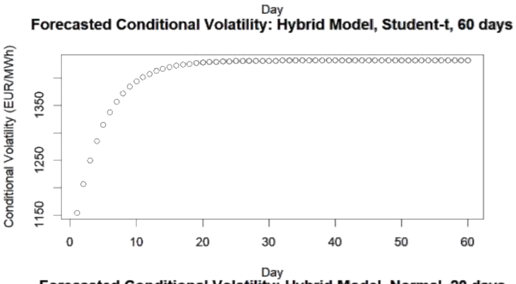


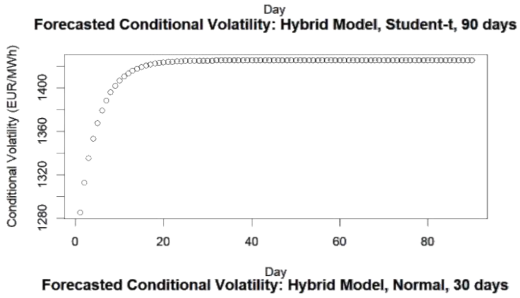


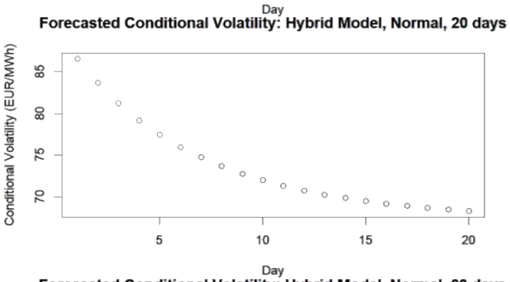


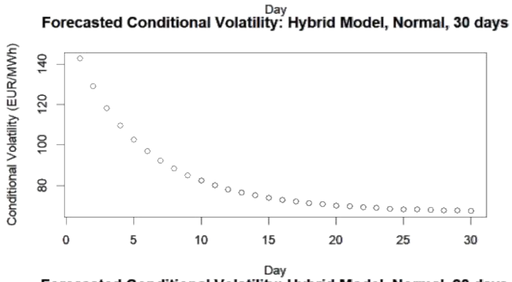


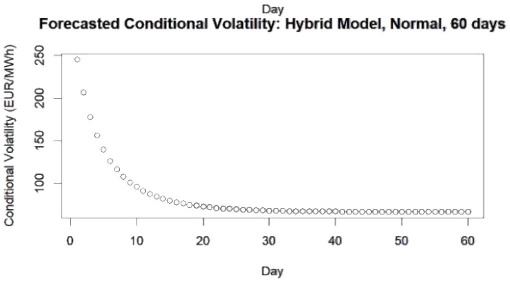


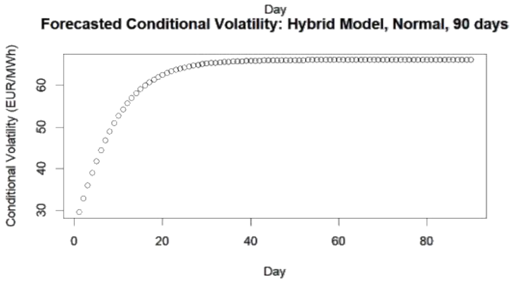






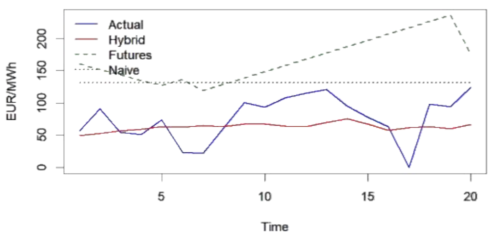




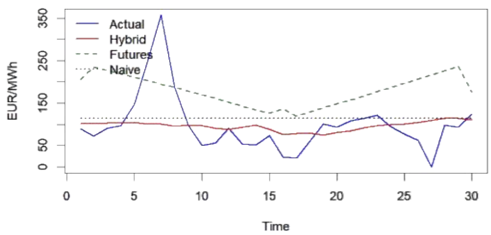


B.2 German spot electricity price forecast comparison, Hybrid model, Actual spot prices, Futures, and naïve forecast.

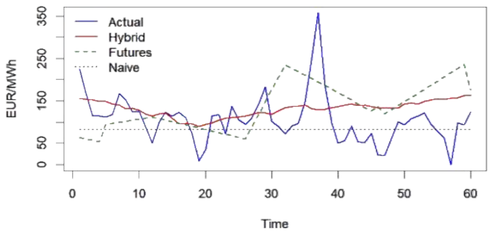
Price Comparison: Hybrid, Actual, Futures, Naive, 20 days, GED



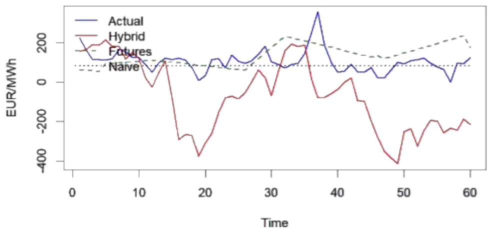
Price Comparison: Hybrid, Actual, Futures, Naive, 30 days, GED



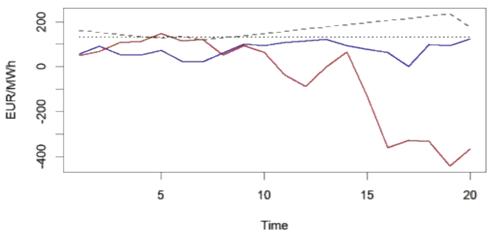
Price Comparison: Hybrid, Actual, Futures, Naive, 60 days, GED



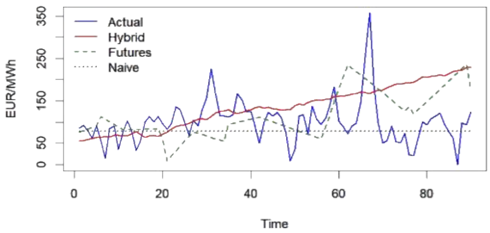
Price Comparison: Hybrid, Actual, Futures, Naive, 60 days, Student-t



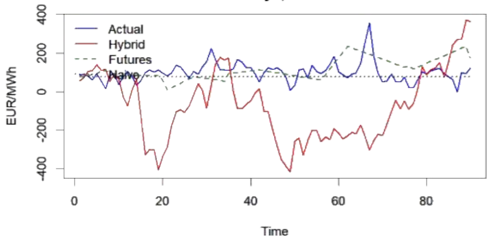
Price Comparison: Hybrid, Actual, Futures, Naive, 20 days, Student-t



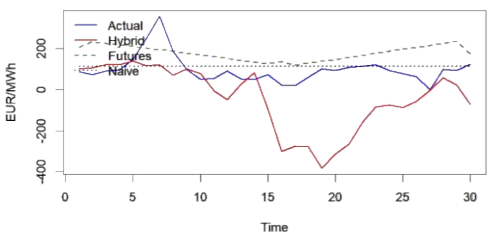
Price Comparison: Hybrid, Actual, Futures, Naive, 90 days, GED



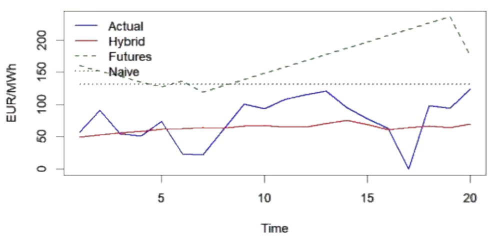
Price Comparison: Hybrid, Actual, Futures, Naive, 90 days, Student-t



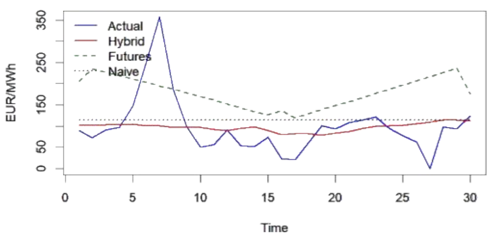
Price Comparison: Hybrid, Actual, Futures, Naive, 30 days, Student-t



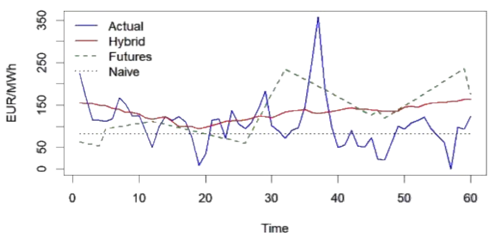
Price Comparison: Hybrid, Actual, Futures, Naive, 20 days, Normal



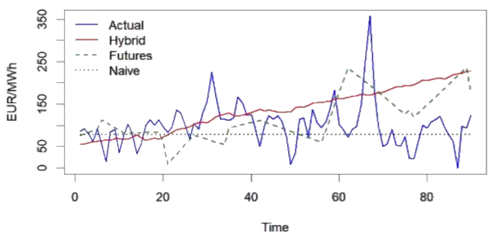
Price Comparison: Hybrid, Actual, Futures, Naive, 30 days, Normal



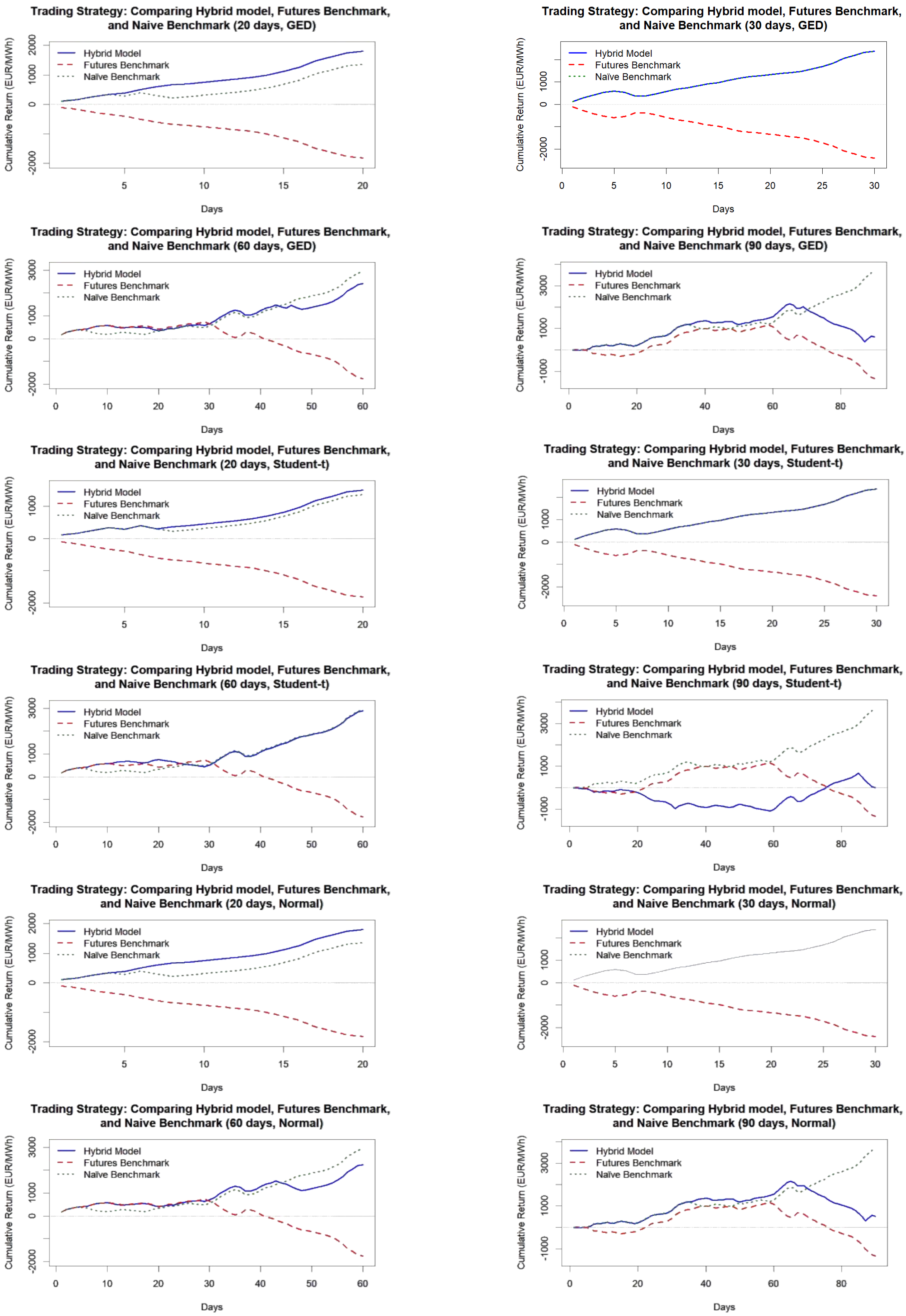
Price Comparison: Hybrid, Actual, Futures, Naive, 60 days, Normal



Price Comparison: Hybrid, Actual, Futures, Naive, 90 days, Normal



B.3 Signal trading strategy results of cumulative returns, comparing the Hybrid Forecast strategy, Futures benchmark, naïve benchmark over 20-, 30-, 60-, and 90-days.



B.4 Histogram of the Monte Carlo trading simulation of cumulative returns.

Graphs showing the mean cumulative return of the hybrid model and the naïve benchmark return over 20-, 30-, 60-, and 90-day horizons. Results are shown over GED, Student-t, and Normal distribution.