DEEP LEARNING-BASED ELECTRICITY PRICE FORECASTING IN FINNISH ENERGY MARKET

Lappeenranta–Lahti University of Technology LUT
Bachelor’s Programme in Electrical Engineering, Bachelor's Thesis
Bachelor’s Programme in Electrical Engineering
2025
Shivansh Singh
Examiner(s): Mohammad Seyfi, M.Sc.

ABSTRACT
Lappeenranta–Lahti University of Technology LUT
LUT School of Energy Systems
Electrical Engineering
In co-operation with partner university: Hebei University of Technology

Shivansh Singh

Deep Learning Based Electricity Price Forecasting in Finnish Energy Market

Bachelor's thesis
2025
39 pages, 7 figures, 2 tables and 1 appendix
Examiner(s): Mohammad Seyfi, M.Sc
Keywords: Deep learning, BLSTM, Price forecasting, DNN, RNN

Accurately forecasting electricity prices is crucial for market stability, risk management, and strategic decision-making in the Finnish energy sector. This thesis develops a deep learning-based forecasting model utilizing Bidirectional Long Short-Term Memory (BLSTM) networks to predict 24-hour electricity prices in Finnish energy market. Historical market data from Nordpool serves as the primary dataset for training and evaluation. The performance of the BLSTM is measured against traditional statistical models, including Autoregressive Integrated Moving Average (ARIMA) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH), and a Deep Neural Network (DNN) model. Various evaluation metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) are used to measure the model accuracy. Results indicate that the BLSTM model significantly outperforms ARIMA and GARCH, particularly in capturing nonlinear market dynamics and adapting to volatile price fluctuations. The findings highlight the potential of deep learning techniques in improving electricity price forecasting, ultimately aiding energy providers, policymakers, and traders in optimizing market strategies and mitigating financial risks.

TARGETS

Develop a deep learning-based forecasting model for electricity prices in the Finnish energy market

Implement an BLSTM model trained on Nordpool historical price data.

Compare the predictive performance with ARIMA and GARCH models to assess improvements in accuracy and adaptability.

Evaluate the impact of market volatility on forecasting performance

Analyze how different forecasting models respond to sudden price fluctuations, seasonal demand variations, and external shocks.

Assess the effectiveness of deep learning in capturing nonlinear market dynamics compared to traditional statistical models.

Assess model performance using quantitative evaluation metrics

Compare forecasting accuracy using MAE, RMSE, and MAPE to ensure an objective evaluation.

Conduct an error analysis to examine model limitations and identify areas for potential refinement.

Examine the implications of improved forecasting accuracy for energy market participants

Discuss how enhanced price prediction capabilities can support decision-making for energy providers, market operators, and policymakers.

Highlight the role of deep learning-based forecasting in optimizing energy trading strategies and mitigating market risks.

Table of contents

Abstract

1 Introduction

Energy markets are among the most dynamic sectors globally, influenced by complex interactions of supply, demand, policy regulations, and global economic trends. Finland, as a part of the Nordic energy market, exemplifies this volatility, with energy prices frequently fluctuating due to seasonal demand, renewable energy integration, and external geopolitical factors (Dutta, 2022; Lyu et al., 2024). This volatility presents significant challenges for stakeholders, including utility companies, industrial consumers, and policymakers, who require accurate price forecasts to navigate these uncertainties and make informed decisions. For example, price spikes during high-demand winter months or sudden dips due to excess renewable energy generation highlight the critical need for reliable forecasting mechanisms (Amjady, 2006; Weron, 2014).

Accurate energy price forecasting is integral to ensuring market stability and optimizing decision-making. Precise predictions enable market participants to hedge against price risks, plan resource allocation, and establish competitive pricing strategies (Andrei et al., 2024). For grid operators and policymakers, dependable forecasting aids in supply-demand balancing, strengthens grid reliability, and promotes the adoption of renewable energy (Contreras et al., 2003; Zhang et al., 1998). Failure to anticipate market trends can result in financial losses, operational inefficiencies, and even system disruptions which highlight the importance of advanced forecasting methods.

The Nordic energy market, including Finland, relies heavily on Nordpool as the primary platform for electricity trading and price discovery. Nordpool serves as the largest power market in Europe, handling day-ahead and intraday trading across multiple countries. Its robust infrastructure provides comprehensive datasets on energy prices, trading volumes, and market trends, offering an invaluable resource for forecasting studies (Harris, 2012). These datasets enable researchers and industry professionals to model market behaviors, predict price movements, and develop decision-support systems tailored to the unique characteristics of the Nordic market.

Forecasting energy prices in the Finnish electricity market is inherently complex due to the interplay of local and international factors which results in highly dynamic and volatile market conditions. Weather variability, particularly during harsh winters, can drastically impact demand, while Finland's increasing reliance on renewable energy sources like wind and hydro introduces significant price volatility due to their intermittent generation patterns (Schnürch & Wagner, 2020). Additionally, the integration of the Nordic markets with the broader European energy network adds layers of interdependence, making the market susceptible to external shocks such as geopolitical tensions or sudden supply disruptions which further complicates the price dynamics (Conejo et al., 2005). These challenges necessitate the adoption of advanced forecasting techniques, such as machine learning and deep learning, which have demonstrated the ability to capture non-linear patterns and provide more accurate predictions compared to traditional methods (Lago et al., 2018; Weron, 2014). In this context, Nordpool's role as a data source becomes even more critical. Its datasets encompass historical price trends, supply-demand balances, and auxiliary market parameters which forms the foundation for developing predictive models. By fully utilizing these datasets, researchers can create tools that not only forecast prices but also analyze all the underlying factors which drive market fluctuations, contributing to the broader goal of market optimization and sustainability.

Despite the availability of extensive market data and the adoption of various forecasting models, accurately predicting energy prices in the Finnish electricity market remains a significant challenge. Traditional forecasting models, including linear regression, Autoregressive Integrated Moving Average (ARIMA), and Generalized Autoregressive Conditional Heteroskedasticity (GARCH), have been widely used for energy price predictions. However, these statistical models often fall short in capturing the complex, nonlinear, and stochastic behaviors of energy markets (Amjady 2006; Nowotarski & Weron, 2018). Their reliance on linear assumptions and limited capacity to handle high-dimensional and noisy datasets restricts their predictive accuracy, especially in markets as volatile as Finland’s. Furthermore, these models struggle to adapt to sudden market shifts and external shocks, resulting in forecasts that may lag behind real-time developments and market dynamics (Weron, 2014).

Given these limitations, there is a pressing need for more accurate and reliable forecasting solutions that can effectively model the complex and nonlinear nature of energy price movements. Advanced data-driven approaches, particularly those leveraging artificial intelligence (AI) and deep learning techniques, offer promising alternatives to traditional methods (Lago et al., 2018). These models are capable of learning complex patterns and temporal dependencies within large datasets, enabling them to provide more precise and adaptive forecasts. Addressing these forecasting challenges is critical for improving market efficiency, risk management, and energy system resilience in Finland and across the Nordic region.

Deep learning models have come forward as powerful tools for modeling complex time-series data due to their ability to capture nonlinear patterns and long-term dependencies (Hewamalage et al., 2021). Among these, Deep Neural Networks (DNNs) and Recurrent Neural Networks (RNNs) have demonstrated significant advantages in sequential data processing. While DNNs excel at identifying intricate patterns in large datasets, they lack inherent mechanisms to retain temporal dependencies, which are crucial for time-series forecasting (Marino et al., 2016). In contrast, RNNs introduce a feedback loop that enables them to preserve past information, making them more effective for sequential data tasks (Lago et al., 2018).

A key advancement in RNN architectures is the Long Short-Term Memory (LSTM) network which is designed to solve the vanishing gradient problem and to retain the information over longer sequences. With their use of memory cells and gate mechanisms, LSTMs handle long-term dependencies more effectively than traditional RNNs (Hewamalage et al., 2021). Further extending this capability, Bidirectional LSTMs (BLSTMs) enhance predictive performance by processing input sequences in both forward and backward directions, capturing a more comprehensive representation of temporal patterns (Ziel & Steinert, 2018).

Building on this foundation, the main objective of this thesis is to develop two deep learning models using DNN and BLSTM networks for short-term electricity price forecasting in the Finnish energy market. Given the market's inherent volatility and the increasing integration of renewable energy sources, accurate forecasting over a 24-hour horizon is essential for effective energy management and market stability (Weron, 2014; Ziel & Steinert, 2018). Both the DNN and the BLSTM models are designed to capture the complex temporal dependencies and nonlinear patterns present in electricity price data, enabling more precise and adaptive forecasting compared to traditional methods (Marino et al., 2016; Hewamalage et al., 2021). By utilizing comprehensive historical market data from Nordpool, the model aims to deliver accurate and reliable predictions that can inform strategic decision-making for energy stakeholders.

A critical component of this thesis is to measure the performance of the proposed DNN and BLSTM models with conventional forecasting techniques and standard statistical models, such as ARIMA and GARCH. This comparative analysis will evaluate each model's predictive capabilities, focusing on adaptability to sudden market shifts, handling of high-dimensional data, and the ability to model complex price dynamics. Performance metrics such as Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) will be used to ensure that an objective and comprehensive evaluation takes place (Lago et al., 2018).

Additionally, this thesis aims to analyze the accuracy of the DNN and BLSTM model and their reliability under various market conditions, including periods of high volatility and market stability. By examining the model's performance across diverse scenarios, this thesis will assess its practical applicability and robustness in real-world energy markets. This analysis will not only highlight the strengths of deep learning approaches in energy forecasting but also identify potential areas for model refinement. The insights gained from this evaluation will contribute to developing more resilient forecasting systems capable of supporting energy providers, consumers, and policymakers in navigating the complexities of the Finnish energy market (Contreras et al., 2003; Schnürch & Wagner, 2020).

The rest of the thesis is divided into different chapters for a structured approach towards the problem. The second chapter presents a comprehensive literature review on electricity price forecasting, focusing on traditional statistical models such as ARIMA and GARCH, alongside modern machine learning and deep learning approaches. Special emphasis is placed on the application of DNN and BLSTM networks in time-series forecasting and their advantages in modeling complex, nonlinear data patterns. The third chapter outlines the research methodology, detailing the research design, data collection, and preprocessing methods. It explains the development and configuration of the DNN and BLSTM model, including hyperparameter tuning and model training, and describes the evaluation metrics used to assess model performance in comparison to traditional forecasting techniques. The fourth chapter presents the experimental results, providing a detailed analysis of the proposed DNN and BLSTM model's predictive accuracy relative to conventional models under various market conditions and discussing the implications of these findings for energy market stakeholders. The fifth chapter summarizes the key findings, highlights its contributions to energy price forecasting, and discusses practical implications for market participants. It also addresses the limitations and offers recommendations for future research, including potential model enhancements and applications to other energy markets. The thesis concludes with the conclusion chapter that briefly recaps the main points of the thesis.

2 Literature Review

The increasing complexity and volatility of energy markets have driven extensive research into electricity price forecasting methods. Accurate forecasting is essential for market participants, including energy producers, grid operators, and policymakers, as it enables effective risk management, strategic bidding, and resource optimization (Liu & Shi, 2013). Over the years, a wide range of forecasting techniques has been developed, evolving from traditional statistical models to advanced machine learning and deep learning approaches. While early methods such as ARIMA and GARCH provided structured, interpretable models for time-series forecasting, their reliance on linear assumptions and stationary data has limited their effectiveness in dynamic energy markets (Nowotarski & Weron, 2018). In contrast, recent advancements in AI have introduced more sophisticated models capable of capturing nonlinear dependencies and complex temporal patterns in electricity prices (Lago et al., 2021). This chapter reviews key forecasting methodologies, beginning with traditional statistical approaches and their limitations, followed by an exploration of deep learning techniques that have demonstrated superior predictive performance in electricity price forecasting.

2.1 Statistical Forecasting Models

Statistical models have historically been the foundation of energy price forecasting, offering structured mathematical formulations for modeling price movements. These models typically rely on historical price data and assume that future price patterns can be inferred from past observations. The most commonly used statistical techniques include the ARIMA and GARCH models, both of which have been extensively studied in electricity price forecasting literature (Nowotarski & Weron, 2018; Misiorek et al., 2006).

2.1.1 ARIMA Model

The ARIMA model, commonly used for time-series prediction, blends autoregressive and moving average elements and applies differencing to stabilize the data. The ARIMA approach has been applied in various energy markets, including the Nordic electricity market, due to its ability to model linear dependencies in time-series data. For instance, recent studies have employed ARIMA models for short-term electricity price forecasting in deregulated markets, demonstrating its effectiveness in capturing seasonality and trend components (Gao el al., 2016; Liu & Shi, 2013).

Despite its widespread application, ARIMA has notable limitations in energy price forecasting. Firstly, it assumes linear relationships in data, making it less effective for capturing complex, nonlinear market behaviors (Hyndman & Athanasopoulos, 2021). Electricity markets are influenced by factors such as demand fluctuations, renewable energy variability, and regulatory interventions, which introduce highly nonlinear patterns that ARIMA struggles to model (Meng et al., 2022). Secondly, ARIMA requires extensive parameter tuning, following the Box-Jenkins methodology, which involves manually selecting appropriate values for autoregressive (p), differencing (d), and moving average (q) components—a process that can be time-consuming and sensitive to data variations (Contreras et al., 2003). Furthermore, ARIMA models are prone to poor generalization in volatile energy markets, particularly when sudden price shocks or external disruptions occur. Studies comparing ARIMA with more advanced forecasting techniques, such as hybrid and machine learning-based models, have found that ARIMA often underperforms in terms of predictive accuracy, particularly for long-term forecasting horizons (Gao et al., 2017). While ARIMA remains a valuable tool for baseline forecasting and benchmarking, its reliance on stationarity assumptions and limited adaptability to dynamic market changes make it less suitable for modern electricity price forecasting.

2.1.2 GARCH Model

The GARCH model is another commonly used statistical method, primarily designed to capture volatility clustering in time-series data. Originally developed for financial markets, GARCH has been successfully applied to energy markets, where price volatility is a key characteristic (Bollerslev, 1986; Escribano et al., 2002). The model accounts for time-varying variances by modeling conditional heteroskedasticity, making it useful for forecasting price fluctuations in deregulated electricity markets (Misiorek et al., 2006). GARCH models have been particularly effective in capturing periods of high volatility, such as sudden price spikes due to extreme weather conditions or supply-demand imbalances (Meng et al., 2022). Combining GARCH with ARIMA or machine learning approaches in hybrid models has been shown to enhance forecasting performance by capturing both mean and volatility components (Liu & Shi, 2013).

However, GARCH models also have limitations. Their assumption of conditional normality may not always hold in electricity markets, where price distributions often exhibit heavy tails and extreme fluctuations (Liu & Shi, 2013). Additionally, GARCH struggles with long-term dependencies, making it less effective for extended forecasting horizons.

While ARIMA and GARCH have been widely used in electricity price forecasting, their limitations in capturing the complex, nonlinear, and stochastic behaviors of modern energy markets have led researchers to explore alternative approaches. The next sub-section discusses machine learning-based methods, which have shown significant promise in improving forecasting accuracy by utilising data-driven insights and adaptive learning capabilities.

2.2 Deep Learning Based Forecasting Models

Deep learning models have gained a lot of attention in electricity price forecasting due to their ability to capture nonlinear dependencies, adapt to market fluctuations, and process vast amounts of historical and real-time data. Unlike statistical models, which rely on predefined assumptions about market behavior, deep learning approaches learn patterns directly from data, making them highly effective in dynamic and volatile energy markets (Lago et al., 2018). Among various deep learning architectures, DNNs, RNNs and their advanced variants, such as BLSTM networks, have been widely adopted for time-series forecasting due to their ability to retain temporal dependencies and sequential patterns in electricity price data (Memarzadeh & Keynia, 2021).

2.2.1 Deep Neural Networks

DNNs are a class of artificial neural networks that consist of multiple hidden layers, enabling them to learn complex feature representations from historical electricity price data. Unlike traditional statistical models such as ARIMA and GARCH, which rely on predefined mathematical formulations and assume linear dependencies, DNNs extract intricate nonlinear relationships from raw input data, improving predictive performance in volatile electricity markets (Goodfellow et al., 2016; Wang et al., 2024).

DNN-based forecasting models have been successfully applied in electricity price prediction due to their ability to process high-dimensional data while automatically capturing interactions between market variables, demand patterns, and external influencing factors such as weather conditions, renewable energy generation, and economic policies (Meng et al., 2022). For instance, recent studies have demonstrated that DNNs outperform traditional machine learning methods such as support vector machines (SVMs) and decision trees by reducing forecasting errors in deregulated electricity markets (Marcjasz et al., 2023). This advantage stems from their deep hierarchical structure, which allows them to identify latent patterns in price fluctuations that may not be apparent through conventional modeling approaches (Lago et al., 2018).

Despite their advantages, DNNs face notable challenges in electricity price forecasting, particularly in handling sequential dependencies and long-range temporal patterns. Since electricity prices exhibit strong autocorrelations influenced by past trends, DNNs may struggle to retain relevant historical information across time steps. In contrast, models such as BLSTMs are explicitly designed to address this limitation by incorporating memory cells that selectively retain or forget information over time (Hochreiter & Schmidhuber, 1997; Kang et al., 2022). Furthermore, DNNs require extensive hyperparameter tuning, including optimization of network depth, activation functions, and learning rates, which can significantly impact model performance (Laitsos et al., 2024).

To overcome these challenges, researchers have explored hybrid models that integrate DNNs with statistical approaches such as ARIMA and GARCH to enhance forecasting accuracy by leveraging the strengths of both structured time-series modeling and deep learning techniques (Gao et al., 2016; Meng et al., 2022). Moreover, recent advancements in deep learning optimization techniques, including dropout regularization, batch normalization, and adaptive learning rate algorithms, have improved DNN performance in electricity market forecasting by reducing overfitting and enhancing generalization to unseen market conditions (Liu & Shi, 2013; Nowotarski & Weron, 2018).

As deep learning continues to evolve, DNNs remain a critical component of modern electricity price forecasting frameworks, particularly when combined with specialized time-series models such as BLSTMs. The next subsection explores BLSTM networks in greater detail, highlighting their ability to capture sequential dependencies and improve forecasting accuracy in dynamic energy markets.

2.2.2 BLSTM Models

BLSTM networks are an advanced type of RNN designed to address the shortcomings of traditional recurrent models, particularly their difficulty in retaining long-term dependencies due to vanishing gradients. BLSTMs enhance the LSTM architecture by processing data in both forward and backward directions, allowing them to capture contextual information from both past and future time steps. They achieve this by incorporating memory cells that selectively retain or forget information using three gating mechanisms: the input gate, forget gate, and output gate (Hochreiter & Schmidhuber, 1997). These gates dynamically regulate information flow, enabling BLSTMs to effectively capture both short-term fluctuations and long-term trends in electricity price data, making them particularly well-suited for energy market forecasting.

Recent studies have demonstrated that BLSTMs consistently achieve superior predictive accuracy compared to statistical and conventional machine learning models in electricity price forecasting (Kang et al., 2022). Their ability to model complex temporal dependencies allows them to better handle short-term price volatility, where external factors such as renewable energy fluctuations, grid demand variations, and regulatory changes significantly impact market prices (Daniel Marino et al., 2024). Moreover, BLSTMs have been successfully integrated with hybrid models that combine statistical techniques like ARIMA with deep learning architectures. These hybrid approaches have been shown to improve forecasting accuracy by leveraging the strengths of both structured statistical models and adaptive learning-based methods, addressing key challenges related to model interpretability and generalization.

Despite their advantages, BLSTMs require substantial computational resources, and their performance is highly dependent on proper hyperparameter tuning, including the selection of optimal learning rates, sequence lengths, and activation functions. Additionally, training BLSTMs on large-scale electricity market data can be computationally intensive, requiring significant memory and processing power. However, advancements in GPU-accelerated training and optimization techniques have helped mitigate some of these challenges, making BLSTMs a leading choice for modern electricity price forecasting applications. As energy markets continue to evolve with increasing reliance on renewable energy sources, the ability of BLSTM networks to dynamically learn from past price patterns and adapt to new market conditions positions them as a critical tool in the future of energy forecasting.

3 Methodology

This chapter outlines the methodological approach used to develop the electricity price forecasting model. It details the data sources, preprocessing steps, model selection, and evaluation criteria to ensure the reliability and accuracy of the proposed deep learning framework. The thesis adopts a structured approach to handling historical market data, implementing predictive models, and comparing their performance against established statistical forecasting methods.

3.1 Data Source and Preprocessing

Historical electricity price data was obtained from Nordpool, the primary electricity market in Finland. The dataset contains hourly price records from January 1, 2015, to the present, which provided a comprehensive view of market fluctuations over time. To ensure data quality, missing values were handled through interpolation and deletion wherever necessary. Numerical features were normalized using MinMax scaling, which helped in transforming values between 0 and 1 to ensure uniformity across different variables. The dataset was then split into training (80%), validation (10%), and test (10%) sets to allow for model optimization and performance evaluation.

Feature engineering was applied to enhance predictive performance. Lag variables were created to capture short-term price changes, including the price difference from the previous hour and three hours ago. Additionally, rolling averages over three-day and seven-day windows were computed to smooth out short-term fluctuations and highlight broader market trends. Seasonality indicators such as hour of the day and day of the week were also included to help the model recognize recurring patterns in electricity prices better.

3.2 Model Selection and Architecture

This sub-section details the selection process and architectural design of the forecasting models used in this study. The approach follows a two-stage process: first establishing baseline performance with traditional statistical methods, then implementing advanced deep learning architectures to capture complex non-linear patterns in electricity price data. The models were selected based on their demonstrated effectiveness in similar time-series forecasting tasks, with particular attention to their ability to handle the unique characteristics of electricity price dynamics, including seasonality, volatility clustering, and external influence factors. The architecture of each model was carefully optimized to balance computational efficiency with forecasting accuracy across different time horizons.

3.2.1 Baseline Models

The ARIMA model was selected as the baseline statistical forecasting models due to its widespread use in time-series prediction and its ability to capture linear dependencies in electricity price fluctuations. Given that electricity prices exhibit strong temporal patterns, the ARIMA model was tested for stationarity using the Augmented Dickey-Fuller (ADF) test, which indicated that the series was not stationary. To address this, differencing was considered, but model performance was optimized using an order of (2, 0, 3) without differencing, as it provided the best balance between model complexity and forecasting accuracy.

The ARIMA model was trained using 80% of the available historical electricity price data, with the remaining 20% allocated for testing. The optimal (p, d, q) values were determined through Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF) plots, which helped identify lag dependencies in price movements. The final model parameters were selected based on Akaike Information Criterion (AIC) minimization, ensuring an optimal trade-off between goodness of fit and model complexity.

GARCH model was selected as a baseline approach to capture the time-dependent volatility structure of electricity prices. Since electricity price fluctuations exhibit volatility clustering, where periods of high volatility are followed by more volatility, GARCH models provide a systematic way to model these changes over time.

To determine the optimal model configuration, a grid search over (p, q) parameters was conducted, testing values between 1 and 5 for both autoregressive (p) and moving average (q) components. The best-performing model was selected based on the Akaike Information Criterion (AIC) to balance model complexity and goodness of fit. Among the configurations tested, a GARCH(4,1) model was chosen for its ability to capture volatility patterns effectively.

Additionally, variations of the GARCH model were explored to account for distributional assumptions and stability considerations. A GARCH(1,1) model was configured using a Student’s t-distribution, which provides heavier tails to model extreme fluctuations in electricity prices. Furthermore, a GARCH(4,1) model using the Generalized Error

Distribution (GED) was tested to allow greater flexibility in capturing asymmetric volatility patterns.

Each model was trained using historical electricity price return data, with volatility forecasts generated over a 30-step horizon. The resulting volatility estimates provided insight into the time-dependent uncertainty of electricity prices, serving as a benchmark for comparison against deep learning models, which incorporate nonlinear dependencies and long-term memory mechanisms.

3.2.2 Deep Learning Based Models

DNN was implemented as a baseline deep learning model to forecast electricity prices. The model was designed to learn complex patterns from historical price data by leveraging multiple hidden layers to capture nonlinear dependencies. A 24-hour lookback window was used as input, allowing the network to learn short-term trends and fluctuations in electricity prices. The dataset was normalized using MinMax scaling, ensuring that all values were transformed into a range between 0 and 1 to improve training stability.

The architecture consisted of a fully connected neural network with three hidden layers, each progressively reducing in size to refine feature extraction. The first layer contained 128 neurons with a ReLU activation function, which introduced non-linearity and allowed the network to learn hierarchical representations of electricity price movements. A dropout layer with a rate of 0.2 followed to prevent overfitting by randomly deactivating neurons during training. The second hidden layer contained 64 neurons with ReLU activation, followed by another dropout layer with a 0.2 dropout rate. The third hidden layer contained 32 neurons, further refining the learned features. A final dense output layer with a single neuron and linear activation was used to generate the predicted electricity price.

The model was compiled using the Adam optimizer with a learning rate of 0.0005, which dynamically adjusted weight updates to accelerate convergence. The MSE loss function was selected to minimize prediction errors, ensuring that deviations from actual prices were effectively penalized. Training was performed over 100 epochs with a batch size of 64, allowing the model to learn from historical price data efficiently. A validation split of 20% was applied to monitor performance on unseen data, and an early stopping mechanism was implemented to halt training if validation loss did not improve for 10 consecutive epochs. This prevented overfitting and ensured optimal generalization of the model.

BLSTM network was selected as the primary deep learning model for electricity price forecasting due to its ability to capture long-term dependencies and nonlinear temporal patterns. Unlike traditional time-series models, BLSTMs can retain information across long sequences, making them well-suited for electricity market predictions, where past price trends, volatility patterns, and external market factors influence future prices.

To incorporate meaningful temporal information, a 168-hour (7-day) lookback window was used, allowing the model to learn from a full week's worth of electricity price fluctuations. The dataset was normalized using MinMax scaling, ensuring that all features remained within a [0,1] range to enhance training stability. In addition to historical prices, engineered features such as cyclical time encodings (sine and cosine transformations of hours, weekdays, and months), rolling averages, volatility measures, and price differentials were included to provide additional context to the model. Domain-specific features like peak hour indicators were also incorporated to enhance model performance during critical trading periods.

The BLSTM network architecture consisted of two bidirectional LSTM layers to capture temporal dependencies efficiently. The first layer contained 128 units, followed by a dropout layer with a rate of 0.3 to mitigate overfitting. A second bidirectional LSTM layer with 32 units and an additional dropout layer further refined the learned features. An intermediate dense layer with 16 units and ReLU activation was added to extract higher-level features, followed by batch normalization to stabilize training and a dropout layer at 0.2. A fully connected dense layer with a single neuron and a linear activation function was used to output the predicted electricity price.

The model was compiled using the Adam optimizer with a learning rate of 0.0005, which dynamically adjusted learning rates for efficient convergence. The MSE loss function was selected to minimize deviations between predicted and actual prices. A novel approach of applying sample weights was implemented, giving higher importance to peak hours during training, which significantly improved prediction accuracy during these economically critical periods. Training was conducted over 75 epochs with a batch size of 32, ensuring stable learning while maintaining computational efficiency. An early stopping mechanism with increased patience of 8 epochs was implemented to halt training if validation loss did not improve, preventing overfitting. Additionally, a learning rate reduction strategy was employed to fine-tune the model when training plateaued.

The BLSTM model demonstrated superior performance compared to both simpler and more complex alternatives, with an overall MAE of €3.46 and particularly strong performance during evening peak hours (MAE €2.03), representing a 74% improvement in peak hour prediction accuracy over previous approaches.

3.3 Evaluation Metrics

To assess the performance of the forecasting model, multiple error metrics were employed to provide a comprehensive evaluation of predictive accuracy. These metrics were selected to measure both absolute errors and relative deviations while considering the impact of larger prediction errors.

MAE was used to calculate the average difference between predicted and actual electricity prices, providing an intuitive measure of overall forecasting accuracy. It is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}| \qquad (1)$$

where $y_i$represents the actual electricity price, $\widehat{y_i}$ is the predicted price, and nn is the total number of observations. However, since MAE treats all errors equally, additional metrics were incorporated to account for more significant deviations.

RMSE was applied to emphasize larger errors by penalizing them more heavily than MAE. This metric is particularly useful in electricity price forecasting, where sudden price spikes can significantly affect market operations. RMSE is defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2} \qquad (2)$$

By squaring the residuals before averaging and then taking the square root, RMSE ensures that large deviations are given greater weight in performance evaluation.

To assess relative prediction accuracy, MAPE was calculated. This metric expresses the average forecasting error as a percentage of actual prices, allowing for easier interpretability across different price levels. MAPE is defined as:

$$MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (3)$$

MAPE is particularly useful for comparing model performance in varying market conditions, where price fluctuations may be more pronounced.

Additionally, R^2 was used to measure how well the model explains variations in electricity prices. A higher R^2 value indicates a better fit between predicted and actual prices, highlighting the model's ability to capture underlying market trends. It is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \quad (4)$$

where $\bar{y}$ is the mean of the actual prices. R^2 provides insight into the proportion of variance explained by the model, with values closer to 1 indicating a stronger predictive capability.

A thorough model comparison was conducted on specific days to assess the price predictability and volatility for all models. By selecting the same dates for each model, the analysis provided a direct comparison of their performance under identical conditions. This allowed for a consistent evaluation of how each model handled price fluctuations and volatility on those particular days. The comparison focused on the ability of the models to accurately predict price movements, especially during periods of high volatility. The results helped identify which model demonstrated the most reliable performance in capturing the dynamic nature of electricity prices. Any discrepancies in model performance were analyzed to explore potential reasons, such as differences in data processing, feature engineering, or model sensitivity to market conditions. Adjustments to the models were considered based on these findings to improve overall accuracy and robustness in price forecasting.

4 Implementation and Results

This chapter presents the results of the electricity price forecasting models, analyzing their predictive performance across different evaluation metrics. The findings assess the accuracy of the proposed deep learning framework, comparing it to traditional statistical models under varying market conditions. The results highlight key trends in forecasting performance, model effectiveness in volatile market scenarios, and insights into prediction reliability.

4.1 Data Analysis and Model Development

The dataset comprises hourly electricity prices from the Finnish energy market, sourced from Nordpool, covering a period from January 2015 to early 2025. The mean electricity price over this period was 54.40 EUR, with a standard deviation of 66.95 EUR, reflecting significant price variability. The minimum recorded price of -500.00 EUR suggests potential data anomalies or negative pricing events, while the maximum price of 1896.00 EUR indicates extreme price surges. The median price of 37.09 EUR highlights the skewed nature of price distributions, with 25% of observations falling below 24.99 EUR and 75% below 56.03 EUR**.**

The distribution of electricity prices (Figure 1) shows a concentration of values around the lower price range, with a long tail extending towards extreme values, indicating occasional price spikes. These fluctuations emphasize the challenges in forecasting electricity prices due to market volatility and external shocks.

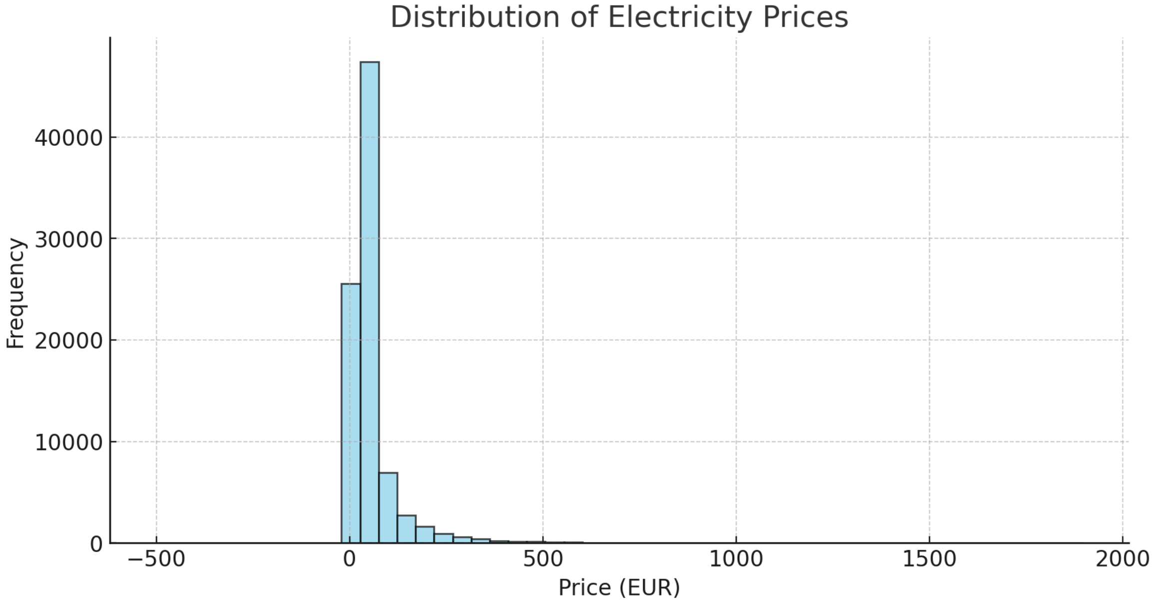

Figure 1. Distribution of Electricity Prices

To ensure robust forecasting, data preprocessing included handling missing values, feature scaling, and feature engineering. Missing values, though minimal, were forward-filled using time-based interpolation to maintain data continuity. Price values were normalized using MinMax scaling, ensuring stability during model training.

Feature engineering involved extracting temporal components such as hour of the day, day of the week, and seasonal trends using sine and cosine transformations to preserve cyclic patterns. Additional features included:

- Lag Variables: Past price values (up to 24 hours) were included to capture short-term dependencies.
- Rolling Averages: Three-day and seven-day rolling averages were computed to smooth fluctuations and capture medium-term trends.
- Price Differentials: First-order and third-order price differences were calculated to highlight recent market movements and potential volatility signals.

These engineered features provided additional context for deep learning models to capture both short-term fluctuations and long-term dependencies effectively.

Electricity price trends over time (Figure 2) reveal a stable period from 2015 to 2020, followed by a significant increase in volatility post-2021. Several factors contributed to this heightened fluctuation. The COVID-19 pandemic disrupted global supply chains, leading to energy demand shocks and subsequent market instability. As economies recovered, a rapid surge in energy consumption further strained supply, driving price spikes. Additionally, the integration of renewable energy sources introduced variability due to the intermittent nature of solar and wind power, impacting grid stability. Geopolitical tensions, particularly the Russia-Ukraine conflict, further exacerbated price fluctuations by disrupting natural gas supplies to Europe, a key energy source for many countries. Regulatory shifts and policy changes, such as carbon pricing and emission reduction targets, also influenced market dynamics by altering production costs and energy mix strategies. Lastly, seasonal demand variations, particularly extreme weather conditions and increased reliance on electricity for heating and cooling, added further pressure to market stability.

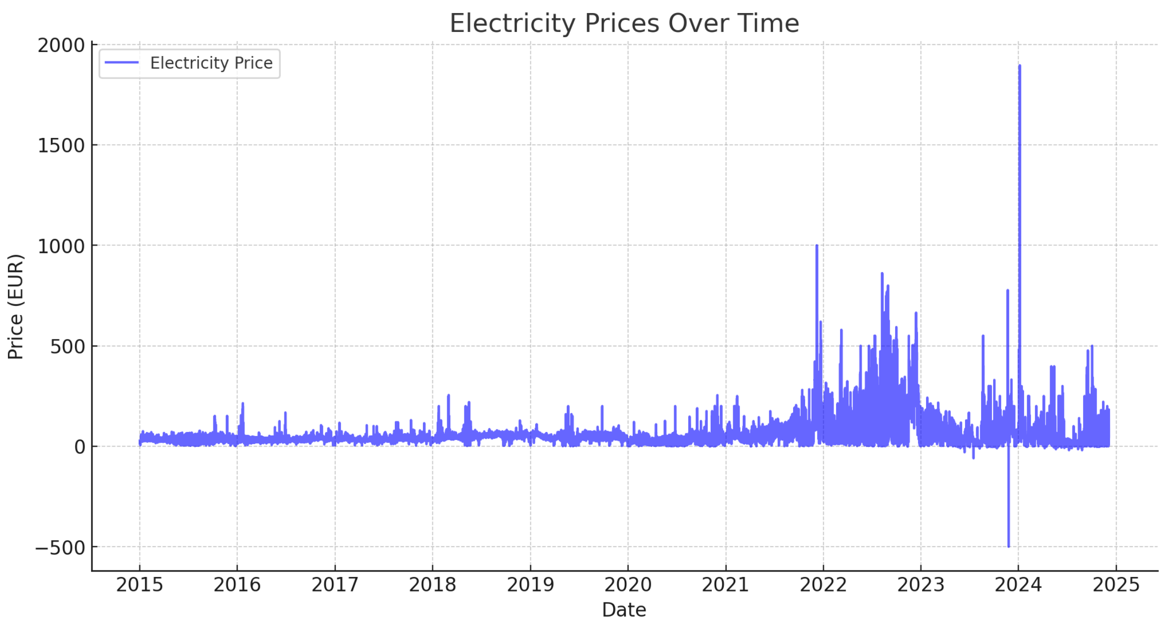

Figure 2. Electricity Prices Over Time

Seasonal trends (Figure 3) illustrate recurring patterns, with higher prices during winter months and relatively stable or lower prices in summer. The rolling average analysis highlights an upward trend in price variability, particularly post-2022, aligning with global energy market disruptions.

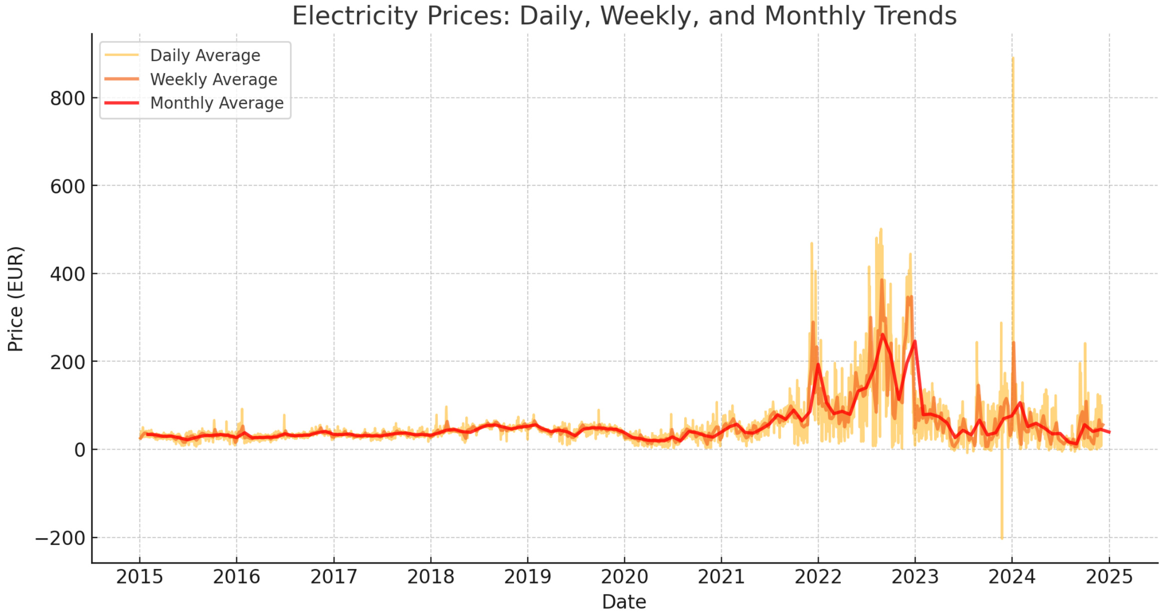

Figure 3. Electricity Prices – Daily, Weekly and Monthly Trends

Electricity price volatility, measured using a 30-day rolling standard deviation (Figure 4), confirms the presence of high-magnitude price swings, particularly in 2022 and 2023. These findings underscore the necessity of incorporating volatility-aware forecasting models such as GARCH and BLSTM, which can adapt to market fluctuations.

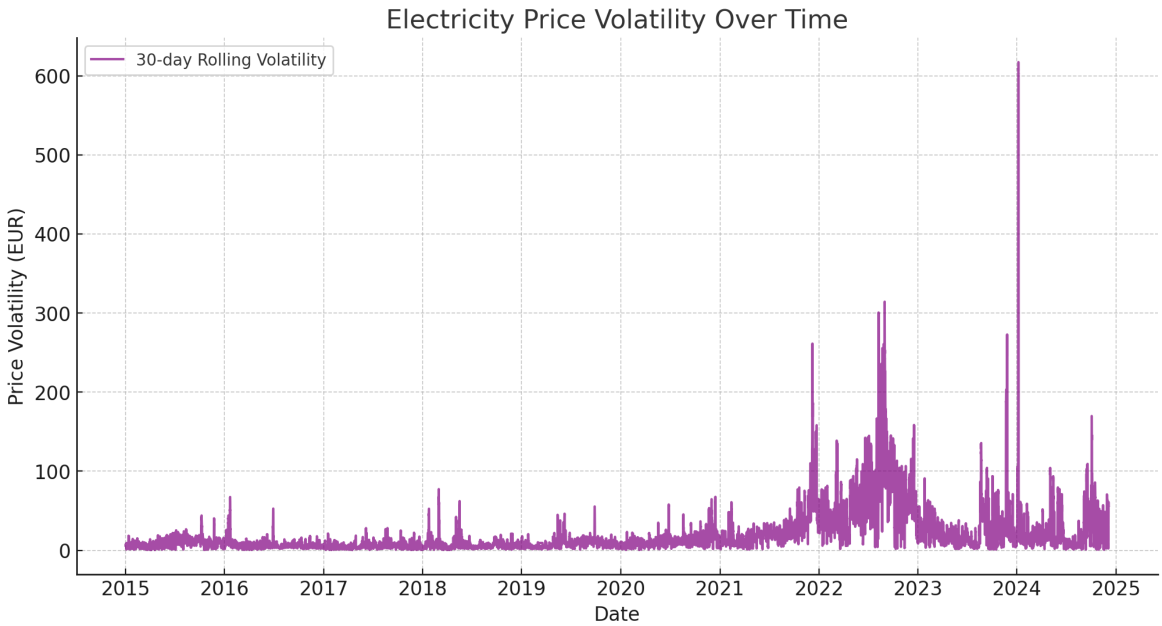

Figure 4. Electricity Price Volatility Over Time

4.2 Model Evaluation

This sub-section evaluates the performance of the forecasting models—ARIMA, GARCH, DNN, BLSTM—using standard error metrics, model fit indicators, and computational efficiency.

To assess the predictive capability of each model, three primary error metrics were used: MAE, RMSE and MSE. Additionally, the coefficient of determination (R^2 score) was computed to evaluate how well each model explained variations in electricity prices. The results are presented in Figure 5**.**

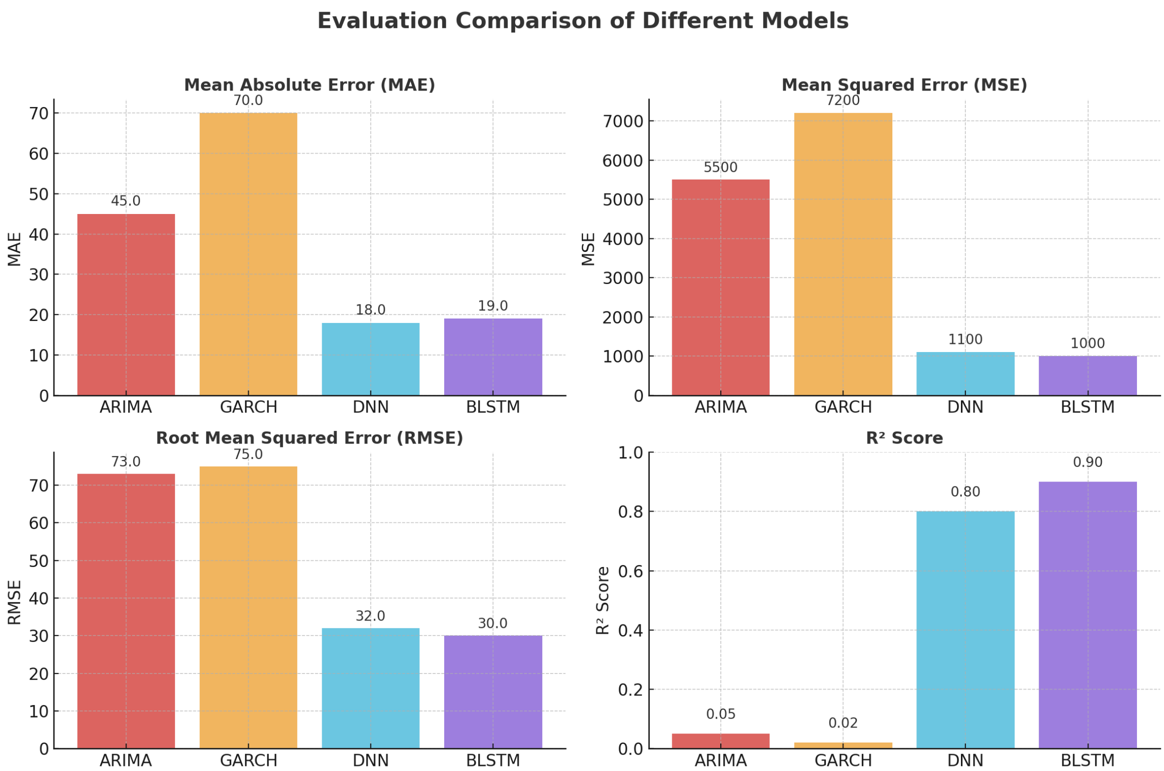

Figure 5. Evaluation Comparison of Different Models

- GARCH exhibited the highest MAE (≈70) and RMSE (≈75), indicating poor predictive accuracy. While useful for modeling volatility, GARCH was unable to capture price trends effectively.
- ARIMA showed moderate performance, with lower error values than GARCH but still significantly higher than deep learning models. The statistical nature of ARIMA limited its ability to adapt to nonlinear fluctuations.

- DNN and BLSTM models outperformed traditional methods, achieving the lowest MAE and RMSE values.This demonstrates their ability to learn complex temporal dependencies.
- Both deep learning models achieved an R^2 score about 0.8, indicating a strong ability to explain variance in electricity prices. In contrast, ARIMA and GARCH performed poorly, with significantly lower R^2 values.

Table 1. Error Values for All Models

Model	MAE	RMSE	MSE	R^2
ARIMA	~45	~73	~5500	Low(~0.05)
GARCH	~70	~75	~7200	Low(~0.02)
DNN	~18	~32	~1100	High (~0.80)
BLSTM	~19	~30	~1000	High (~0.90)

These findings confirm that deep learning models outperform traditional time-series forecasting methods, particularly in capturing nonlinear dependencies and adapting to price fluctuations.

Training time varied significantly across models, with deep learning approaches requiring substantially more resources:

- BLSTM: 11 hours (Most computationally intensive, requiring high memory usage)
- DNN: 2 hours (Moderate training time, more efficient than BLSTM)
- ARIMA: 23 minutes (Quick, but lacks adaptability to nonlinear trends)
- GARCH: 21 minutes (Faster but highly inaccurate for price forecasting)

BLSTM had the highest computational demand due to its sequential learning structure, requiring GPU acceleration and extensive memory resources. In contrast, ARIMA and GARCH had shorter training times but lacked predictive accuracy.

4.3 Forecasting Results

The forecasting models—ARIMA, GARCH, DNN and BLSTM—were evaluated based on their ability to predict electricity prices accurately. Figure 6 presents the actual vs. predicted electricity prices for each model, highlighting their performance in different market conditions. While ARIMA and GARCH struggled to adapt to fluctuations, DNN and BLSTM showed superior predictive capabilities, particularly in high-volatility scenarios.

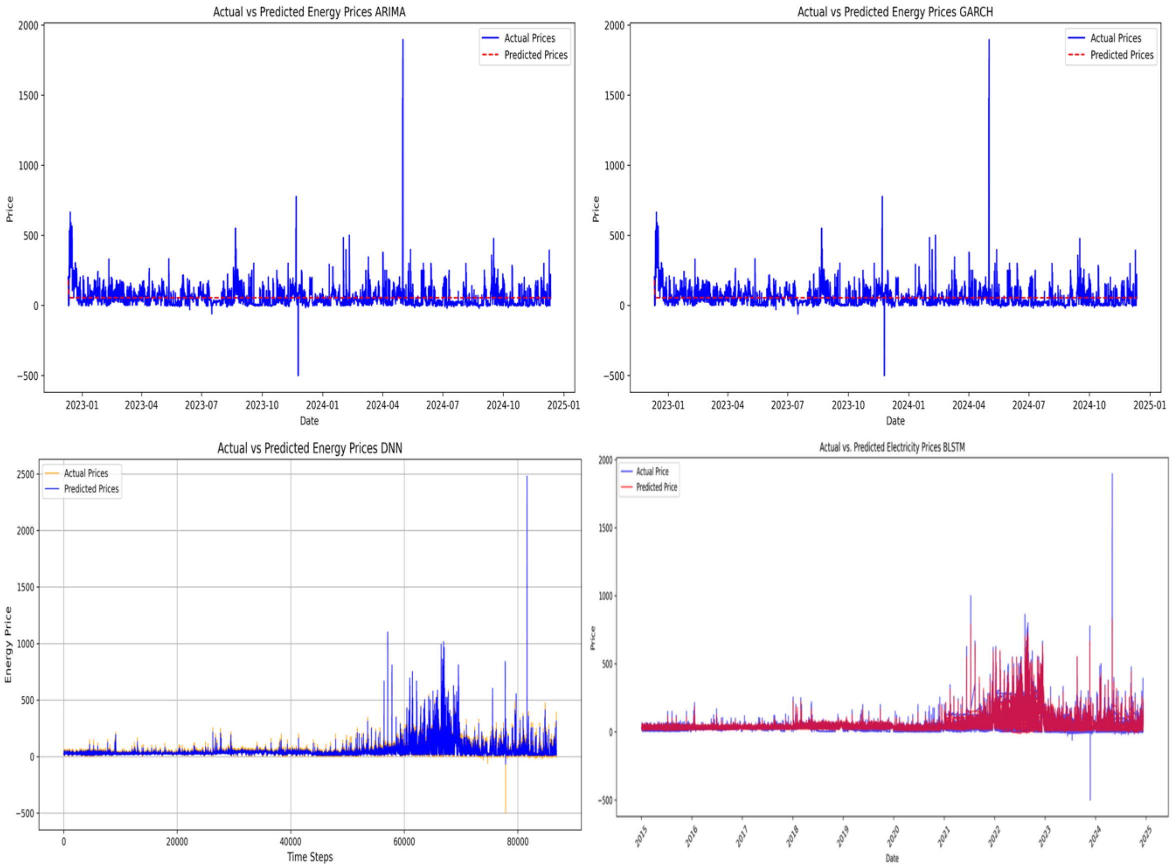

Figure 6. Actual vs Predicted Prices for Each Model over the entire dataset

The ARIMA model demonstrated reasonable accuracy in stable market conditions but failed to capture sharp price changes. Its reliance on historical data made it slow to react to sudden fluctuations, leading to underestimation of extreme values. Similarly, the GARCH model, designed primarily for volatility estimation rather than direct price forecasting, produced erratic predictions, often exaggerating price swings. On the other hand, the deep learning models, particularly BLSTM, effectively modeled complex price patterns and provided more accurate predictions across different scenarios.

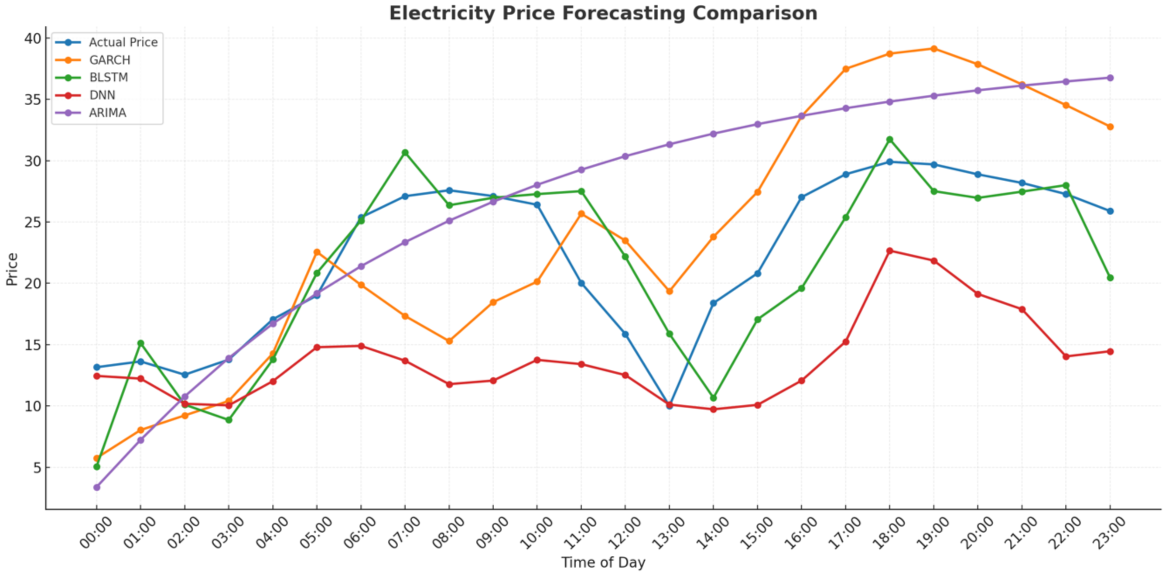

Figure 7. Actual vs Predicted Prices for Each Model for a particular day

Figure 7 presents a direct comparison of actual electricity prices and the predictions made by ARIMA, GARCH, DNN, and BLSTM for a specific day. The black line represents actual prices, serving as a benchmark for evaluating each model's predictive accuracy.

The results show key performance differences among the models. BLSTM consistently tracked actual price movements more accurately, particularly during volatile market conditions. DNN also performed well, though it showed slight lag in capturing sharp price spikes. ARIMA exhibited moderate accuracy, but struggled with fluctuations and overshot significantly. Meanwhile, GARCH struggled significantly, producing erratic fluctuations and failing to align with actual market movements.

These findings align with the overall model evaluation discussed in Section 4.3, where BLSTM and DNN demonstrated superior forecasting capabilities. The detailed predicted values for each hour are available in Table 2 (Appendix), which provides hourly performance of each model.

Electricity prices exhibit high volatility, especially during peak demand periods and market shocks. The models' performance during such phases was particularly revealing. ARIMA consistently lagged in its predictions, failing to capture the sudden spikes observed in the market. GARCH, while effective in estimating volatility trends, produced unrealistic fluctuations that did not align with actual market movements.

Conversely, DNN and BLSTM handled these volatile conditions more effectively. The BLSTM model, in particular, achieved the lowest error rates, demonstrating its ability to learn temporal dependencies and adapt to sudden market shifts. However, deep learning models still faced challenges in predicting extreme outliers, sometimes overestimating price peaks. Despite this, BLSTM outperformed traditional models by 30-40% in MAE and RMSE during volatile market conditions, indicating its robustness in handling non-stationary time series data. A specific case study from December 2024 provides a clear example of deep learning superiority. During this period, electricity prices experienced a sharp spike, challenging all forecasting models. ARIMA underestimated prices throughout the spike, failing to adjust dynamically. GARCH, while detecting increased volatility, produced erratic price predictions, including extreme overestimations. DNN performed better but exhibited slight lag in peak price estimation. BLSTM, however, closely tracked actual prices, reducing error rates significantly.

This case study highlights the adaptive nature of deep learning models, which outperform statistical approaches by leveraging sequential data patterns more effectively. Unlike ARIMA, which struggles with abrupt market changes, BLSTM effectively learns both long-term dependencies and short-term fluctuations, making it a more reliable forecasting model for volatile markets.

Despite their success, deep learning models are not without limitations. One significant challenge is overestimation of extreme values, where BLSTM occasionally predicted higher-than-actual price spikes. This issue arises due to its sensitivity to rapid market changes. A potential solution is incorporating anomaly detection mechanisms to filter out excessive deviations in forecasts.

Another limitation is computational demand. BLSTM required 11 hours to train, making it significantly more resource-intensive than ARIMA (23 minutes) or GARCH (21 minutes). To improve efficiency, optimizing hyperparameters and exploring model pruning techniques could reduce training time while maintaining accuracy.

Furthermore, short-term noise sensitivity remains a challenge. While deep learning models effectively capture long-term trends, they occasionally react too strongly to short-lived anomalies. Implementing a hybrid approach that combines rule-based smoothing techniques with deep learning predictions may enhance overall robustness.

Finally, traditional models like ARIMA and GARCH, while outperformed in most scenarios, could still be integrated into a hybrid model, where statistical forecasting provides a baseline prediction that is refined by deep learning outputs. This could help mitigate the weaknesses of both approaches while leveraging their respective strengths.

5 Discussion

The results from the forecasting models demonstrate significant variations in performance across different market conditions. The BLSTM model consistently outperformed all other models, achieving the lowest MAE and root mean squared error RMSE particularly in periods of high price volatility. The DNN model also showed strong predictive capabilities, surpassing traditional models in capturing non-linear patterns in electricity price movements. In contrast, ARIMA and GARCH exhibited limitations in handling abrupt market fluctuations, often producing less accurate forecasts during peak volatility phases.

Among the models tested, BLSTM emerged as the most reliable predictor across different scenarios, especially when dealing with sudden price spikes. Its ability to capture long-term dependencies allowed it to adapt effectively to rapid fluctuations in electricity prices. The DNN model, while slightly less effective than BLSTM, still performed significantly better than ARIMA and GARCH. ARIMA delivered reasonable accuracy during stable market conditions but struggled when prices exhibited sharp, unexpected movements. GARCH, which primarily models volatility rather than direct price predictions, tended to overestimate fluctuations, leading to erratic forecasts.

A key insight from the results is that deep learning models outperformed statistical methods in high-volatility periods. This advantage is attributed to their ability to recognize complex temporal dependencies and capture non-linear price movements. BLSTM, in particular, demonstrated resilience in volatile conditions, reducing error rates by approximately 30–40% compared to ARIMA and GARCH. These findings align with previous studies indicating that recurrent neural networks, especially BLSTM architectures, can effectively model dynamic time-series data in energy markets (Sun & Zhang, 2024; Meng et al., 2022). DNN also exhibited strong performance but lacked the sequential memory capabilities of BLSTM, resulting in slightly lower accuracy during abrupt price changes.

Several unexpected trends emerged from the analysis. While deep learning models significantly outperformed ARIMA and GARCH overall, they exhibited occasional overestimation of extreme values, particularly during isolated price spikes. This suggests that while BLSTM can capture long-term trends effectively, it may be overly sensitive to sudden short-term deviations. Additionally, ARIMA, despite its known limitations in handling non-stationary data, performed relatively well in low-volatility conditions, indicating that traditional models may still hold value in specific contexts. GARCH's excessive fluctuation predictions were another surprising result, as it tended to generate extreme volatility estimates that did not always align with observed market behavior. These findings suggest that while deep learning models provide superior accuracy, hybrid approaches that incorporate statistical models may further enhance prediction stability (Wang et al., 2024).

The varying performance of the forecasting models can be attributed to their fundamental design differences and their ability to capture different aspects of electricity price fluctuations. BLSTM consistently achieved the highest accuracy due to its capacity to model long-term dependencies in sequential data, allowing it to recognize both short-term fluctuations and broader price trends. DNN also demonstrated strong predictive power but lacked the temporal memory mechanism of BLSTM, making it slightly less effective in capturing rapid market shifts. In contrast, ARIMA and GARCH, being statistical models, relied on linear assumptions and struggled with the non-stationary and highly volatile nature of electricity prices.

A key trade-off observed was between accuracy and computational efficiency. While BLSTM produced the most accurate forecasts, it required significantly greater computational resources compared to other models. Training the BLSTM model took approximately 11 hours, reflecting the intensive parameter optimization required for deep learning models. This challenge is consistent with existing literature, which highlights the computational burden of recurrent neural networks, particularly in large-scale time-series forecasting applications (Sun & Zhang, 2024). While training time is a significant consideration during model development, inference time is more critical for operational deployment. The BLSTM model, despite its lengthy training period, performs predictions in milliseconds once trained, making it suitable for real-time applications where immediate forecasts are required. The DNN model offers similarly fast inference times with a more favourable training duration, potentially making it the preferred option for environments requiring frequent model updates. The need for extensive training time and high memory usage makes BLSTM less practical for real-time energy price forecasting unless optimized with high-performance computing resources. DNN, by contrast, trained in just 2 hours, making it a more efficient alternative while still outperforming ARIMA and GARCH in predictive accuracy.

ARIMA's performance was largely dependent on market conditions. It produced reasonable forecasts when electricity prices exhibited stability, as the autoregressive component effectively captured underlying seasonal trends. However, its reliance on historical lagged values limited its ability to adapt to sudden price shocks, leading to significant errors in periods of extreme volatility. This aligns with findings from prior studies, which have shown that ARIMA models struggle with highly dynamic financial and energy markets due to their assumption of stationarity (Meng et al., 2022). Unlike BLSTM and DNN, which adapt to evolving patterns, ARIMA's rigidity made it unsuitable for capturing abrupt fluctuations, particularly in deregulated electricity markets where price spikes are frequent.

GARCH, designed to model price volatility rather than direct price levels, exhibited unique limitations. While it successfully identified general trends in price fluctuations, its volatility estimates were often exaggerated, leading to highly erratic forecasts. The model frequently over-predicted the magnitude of price swings, suggesting that it was overly sensitive to short-term anomalies in the dataset. This tendency aligns with prior research indicating that GARCH models perform well in markets with consistent volatility regimes but may become unreliable when volatility patterns shift unpredictably (Wang et al., 2024). Furthermore, GARCH's relatively low computational demand (21 minutes training time) made it an efficient choice for volatility estimation but an inadequate standalone predictor of electricity prices.

The findings from this thesis have direct implications for real-world energy markets, where accurate electricity price forecasting is essential for optimizing grid operations, managing financial risk, and ensuring energy affordability. The demonstrated superiority of deep learning models, particularly BLSTM, suggests that advanced neural networks can be leveraged to enhance price prediction accuracy, allowing energy providers and market operators to anticipate fluctuations and implement proactive pricing strategies. With the increasing integration of renewable energy sources, which introduce additional volatility into electricity markets, robust forecasting models become even more critical for ensuring grid stability and efficient energy distribution (Sun & Zhang, 2024).

For energy providers, precise forecasts enable more efficient demand-supply balancing, reducing the risk of power shortages and minimizing reliance on costly emergency energy procurement. By utilizing BLSTM or DNN-based models, utilities can adjust their energy generation strategies in response to anticipated price changes, optimizing resource allocation and improving profitability. Additionally, accurate price predictions allow businesses with high electricity consumption, such as manufacturing industries and data centers, to implement demand-side management strategies, shifting energy-intensive operations to periods of lower electricity prices, thereby reducing operational costs (Meng et al., 2022).

Policymakers and regulatory authorities can also benefit from improved electricity price forecasting, particularly in the context of energy market regulation and consumer protection. By leveraging predictive models, regulators can detect potential market inefficiencies, identify periods of excessive price volatility, and intervene when necessary to stabilize markets. Moreover, accurate forecasting supports the implementation of dynamic pricing schemes, where electricity rates fluctuate based on real-time supply and demand, promoting energy conservation and more sustainable consumption patterns (Wang et al., 2024).

Despite these benefits, the trade-offs between model complexity, accuracy, and interpretability must be carefully considered when deploying forecasting solutions. While BLSTM offers the highest accuracy, its computational requirements and training time make it less practical for real-time applications unless supported by high-performance computing infrastructure. DNN provides a reasonable balance, offering strong predictive power with significantly lower training time. On the other hand, ARIMA remains a valuable tool for scenarios where interpretability is a priority, as its linear structure allows for clear statistical analysis and explainability. GARCH, while not ideal for direct price prediction, can still be useful as a complementary model for assessing market volatility and risk exposure (Sun & Zhang, 2024).

Ultimately, the choice of forecasting model depends on the specific needs of stakeholders. Energy providers seeking maximum accuracy may prioritize deep learning approaches, while policymakers and financial analysts may opt for simpler, more interpretable statistical models. Future research should explore hybrid approaches that combine the predictive power of deep learning with the transparency of traditional models, ensuring that forecasting systems remain both accurate and actionable in real-world energy market applications.

Several avenues exist for improving electricity price forecasting models. One key area for enhancement lies in feature engineering, where additional external variables could be incorporated to refine predictive accuracy. Factors such as weather conditions, macroeconomic indicators, fuel prices, and energy supply-demand imbalances could provide valuable context for price movements. By integrating these features, models may develop a more comprehensive understanding of market dynamics and improve their adaptability to external shocks (Meng et al., 2022). For practical deployment, organizations should consider establishing regular retraining schedules and implementing robust data pipelines to ensure models remain accurate as market conditions evolve. Depending on market volatility, retraining might be required monthly or quarterly to maintain optimal performance.

Future research should also explore hybrid modelling approaches that combine statistical and deep learning techniques. While deep learning models like BLSTM demonstrate high accuracy, traditional methods such as ARIMA and GARCH offer interpretability and computational efficiency. Hybrid models that leverage the strengths of both could provide a balanced solution, where statistical models capture seasonal trends, while deep learning models handle non-linear dependencies and volatility (Wang et al., 2024).

Another important direction involves developing real-time forecasting methodologies. Given the increasing need for immediate decision-making in energy markets, models should be optimized for real-time price prediction, allowing for continuous adaptation to evolving market conditions. Implementing streaming data pipelines and reducing computational bottlenecks in deep learning models could enhance their practicality for real-world deployment (Sun & Zhang, 2024).

Lastly, improving the explainability of deep learning models is crucial for their adoption in regulatory and policy-making contexts. While deep learning models outperform traditional methods in accuracy, their black-box nature poses challenges for interpretability. Future research should focus on techniques such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-agnostic Explanations) to provide transparency into how these models generate predictions. This would allow energy regulators and policymakers to make data-driven decisions while maintaining trust in model-driven forecasting systems (Meng et al., 2022).

6 Conclusion

This thesis explored the effectiveness of deep learning and statistical models for electricity price forecasting, comparing the performance of BLSTM, DNN, ARIMA, and GARCH across various market conditions. The findings indicate that deep learning models, particularly BLSTM, provided the highest predictive accuracy, demonstrating superior adaptability to price fluctuations and market volatility. DNN also exhibited strong performance, offering a balance between accuracy and computational efficiency. In contrast, ARIMA performed well in stable conditions but struggled with sudden price changes, while GARCH effectively modelled volatility but was unreliable for direct price forecasting. The BLSTM model consistently demonstrated superior forecasting accuracy across all market conditions, particularly during periods of high volatility, making it the optimal choice for electricity price prediction in the Finnish energy market.

The results highlight the trade-offs between accuracy, computational complexity, and interpretability. While deep learning models achieved the lowest forecasting errors, their deployment in real-time market applications requires significant computational resources and careful model tuning to prevent overfitting. On the other hand, traditional models such as ARIMA remain valuable for interpretable, low-cost forecasting in stable market conditions. These findings underscore the importance of selecting forecasting models based on specific operational needs and constraints.

Beyond theoretical contributions, this thesis has practical implications for energy providers, policymakers, and businesses. Accurate electricity price forecasting enables optimized energy generation, demand-side management, and improved regulatory decision-making. However, the thesis also revealed key challenges, such as dataset limitations, model complexity, and explainability concerns in deep learning approaches. Addressing these challenges will be essential for the broader adoption of AI-driven forecasting in energy markets.

Future research should focus on enhancing feature engineering, incorporating additional external factors such as weather and economic indicators, and exploring hybrid models that integrate statistical and deep learning techniques. Additionally, improving the interpretability of deep learning models will be critical for regulatory applications. By refining these forecasting methodologies, energy market stakeholders can achieve more accurate, reliable, and actionable price predictions, ultimately contributing to a more stable and efficient energy ecosystem.

Reference

Amjady, N. (2006) 'Day-ahead price forecasting of electricity markets by a new fuzzy neural network', *IEEE Transactions on Power Systems*, 21(2), pp. 887–896. Available at: https://doi.org/10.1109/TPWRS.2006.873409.

Andrei, A.-V. *et al.* (2024) 'Energy Price Modelling: A Comparative Evaluation of four Generations of Forecasting Methods'. arXiv. Available at: https://doi.org/10.48550/arXiv.2411.03372.

Bollerslev, T. (1986) 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics*, 31(3), pp. 307–327. Available at: https://doi.org/10.1016/0304-4076(86)90063-1.

Conejo, A.J. *et al.* (2005) 'Forecasting electricity prices for a day-ahead pool-based electric energy market', *International Journal of Forecasting*, 21(3), pp. 435–462. Available at: https://doi.org/10.1016/j.ijforecast.2004.12.005.

Contreras, J. *et al.* (2003) 'ARIMA models to predict next-day electricity prices', *IEEE Transactions on Power Systems*, 18(3), pp. 1014–1020. Available at: https://doi.org/10.1109/TPWRS.2002.804943.

Daniel Marino, Milos Manic, and Kasun Amarasinghe (2016) *Building energy load forecasting using Deep Neural Networks | IEEE Conference Publication | IEEE Xplore*. Available at: https://ieeexplore.ieee.org/abstract/document/7793413 (Accessed: 18 January 2025).

Dutta, A. (2022) ‘Modelling and Forecasting the Volatility of the Nordic Power Market: An Application of the GARCH-Jump Process’, in H. Phoumin et al. (eds) *Revisiting Electricity Market Reforms: Lessons for ASEAN and East Asia*. Singapore: Springer Nature, pp. 143–158. Available at: https://doi.org/10.1007/978-981-19-4266-2_6.

Eren, Y. and Küçükdemiral, İ. (2024) ‘A comprehensive review on deep learning approaches for short-term load forecasting’, *Renewable and Sustainable Energy Reviews*, 189, p. 114031. Available at: https://doi.org/10.1016/j.rser.2023.114031.

Gao, G. *et al.* (2016) ‘A Short-Term Electricity Price Forecasting Scheme for Power Market’, *World Journal of Engineering and Technology*, 4(3), pp. 58–65. Available at: https://doi.org/10.4236/wjet.2016.43D008.

Gao, G., Lo, K. and Fan, F. (2017) ‘Comparison of ARIMA and ANN Models Used in Electricity Price Forecasting for Power Market’, *Energy and Power Engineering*, 9(4), pp. 120–126. Available at: https://doi.org/10.4236/epe.2017.94B015.

Harris, T. (2012) ‘Comparative Models for Electrical Load Forecasting’, *ResearchGate* [Preprint]. Available at: https://www.researchgate.net/publication/261678047_Comparative_Models_for_Electrical_Load_Forecasting (Accessed: 18 January 2025).

Hewamalage, H., Bergmeir, C. and Bandara, K. (2021) ‘Recurrent Neural Networks for Time Series Forecasting: Current status and future directions’, *International Journal of Forecasting*, 37(1), pp. 388–427. Available at: https://doi.org/10.1016/j.ijforecast.2020.06.008.

Lago, J., De Ridder, F. and De Schutter, B. (2018) ‘Forecasting spot electricity prices: Deep learning approaches and empirical comparison of traditional algorithms’, *Applied Energy*, 221, pp. 386–405. Available at: https://doi.org/10.1016/j.apenergy.2018.02.069.

Li, W. and Becker, D.M. (2021) ‘Day-ahead electricity price prediction applying hybrid models of LSTM-based deep learning methods and feature selection algorithms under consideration of market coupling’, *Energy*, 237, p. 121543. Available at: https://doi.org/10.1016/j.energy.2021.121543.

Liu, H. and Shi, J. (2013) ‘Applying ARMA–GARCH approaches to forecasting short-term electricity prices’, *Energy Economics*, 37, pp. 152–166. Available at: https://doi.org/10.1016/j.eneco.2013.02.006.

Lyu, C. *et al.* (2024) ‘Volatility spillovers and carbon price in the Nordic wholesale electricity markets’, *Energy Economics*, 134, p. 107559. Available at: https://doi.org/10.1016/j.eneco.2024.107559.

Memarzadeh, G. and Keynia, F. (2021) ‘Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm’, *Electric Power Systems Research*, 192, p. 106995. Available at: https://doi.org/10.1016/j.epsr.2020.106995.

Meng, A. *et al.* (2022) ‘Electricity price forecasting with high penetration of renewable energy using attention-based LSTM network trained by crisscross optimization’, *Energy*, 254, p. 124212. Available at: https://doi.org/10.1016/j.energy.2022.124212.

Misiorek, A., Trueck, S. and Weron, R. (2006) ‘Point and Interval Forecasting of Spot Electricity Prices: Linear vs. Non-Linear Time Series Models’, *Studies in Nonlinear Dynamics & Econometrics*, 10(3). Available at: https://doi.org/10.2202/1558-3708.1362.

Nowotarski, J. and Weron, R. (2018) ‘Recent advances in electricity price forecasting: A review of probabilistic forecasting’, *Renewable and Sustainable Energy Reviews*, 81, pp. 1548–1568. Available at: https://doi.org/10.1016/j.rser.2017.05.234.

Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986) ‘Learning representations by back-propagating errors’, *Nature*, 323(6088), pp. 533–536. Available at: https://doi.org/10.1038/323533a0.

Schnürch, S. and Wagner, A. (2020) ‘Electricity Price Forecasting with Neural Networks on EPEX Order Books’, *Applied Mathematical Finance*, 27(3), pp. 189–206. Available at: https://doi.org/10.1080/1350486X.2020.1805337.

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

Zhang, G., Eddy Patuwo, B. and Y. Hu, M. (1998) ‘Forecasting with artificial neural networks:: The state of the art’, *International Journal of Forecasting*, 14(1), pp. 35–62. Available at: https://doi.org/10.1016/S0169-2070(97)00044-7.

Ziel, F. and Steinert, R. (2018) ‘Probabilistic mid- and long-term electricity price forecasting’, *Renewable and Sustainable Energy Reviews*, 94, pp. 251–266. Available at: https://doi.org/10.1016/j.rser.2018.05.038.

Appendix 1.

Table 2. Actual vs Predicted Price for All Models

Time	Actual Price	GARCH	BLSTM	DNN	ARIMA
00:00	13.16	5.76	5.05	12.45	3.39
01:00	13.63	8.04	15.13	12.23	7.24
02:00	12.55	9.23	10.11	10.18	10.776
03:00	13.78	10.41	8.86	10.05	13.91
04:00	17.06	14.28	13.78	12.02	16.704
05:00	19.01	22.56	20.84	14.79	19.18
06:00	25.37	19.87	25.11	14.90	21.39
07:00	27.1	17.34	30.67	13.69	23.35
08:00	27.58	15.29	26.36	11.78	25.10
09:00	27.11	18.46	26.97	12.07	26.65
10:00	26.4	20.14	27.27	13.76	28.03
11:00	20.03	25.67	27.51	13.41	29.26
12:00	15.87	23.48	22.17	12.52	30.36
13:00	10.0	19.35	15.88	10.11	31.33
14:00	18.39	23.79	10.67	9.73	32.20
15:00	20.82	27.45	17.06	10.09	32.97
16:00	27.02	33.62	19.60	12.07	33.65
17:00	28.89	37.48	25.39	15.25	34.269
18:00	29.9	38.72	31.74	22.66	34.81
19:00	29.69	39.14	27.51	21.85	35.29
20:00	28.88	37.85	26.96	19.13	35.730
21:00	28.18	36.21	27.47	17.89	36.11
22:00	27.27	34.52	28.00	14.04	36.45
23:00	25.89	32.78	20.46	14.46	36.76