TIØ4317 Written report draft

Problem statement

The electricity market is characterized by significant price volatility, with prices fluctuating greatly across seasons and often even within the day. On the supply side, weather and climate conditions affect the production of renewable energy, which dominates Norway's electricity generation (SSB, 2025). Additionally, geopolitical events such as the Russian invasion of Ukraine, and changes in international trade regulations can disrupt cross-border electricity flows, further influencing prices. On the demand side, seasonal temperature variation affects electricity consumption, especially due to increased heating needs during the winter months.

In recent years, electricity prices in Norway have experienced extreme spikes, with electricity prices ranging from close to 0 EUR/MWh to more than 600 EUR/MWh over the past five years (Entso-e, 2025). This project aims to empirically investigate the key factors influencing electricity prices in Norway during the period from 2020 to 2025, addressing the research question *What factors determine the valuation of electricity prices in Norway?*

Data

To answer the research question, we use daily frequency data from the period 2020 to 2025. This time period is chosen as it includes a wide range of electricity price movements, including some of the most extreme spikes seen in recent years. This makes it an interesting period for analyzing the potential drivers behind electricity price fluctuations. Although electricity prices are recorded at a much higher frequency to account for the rapid movements, we use daily data as a practical compromise. Most of the independent variables in our model are only available at daily frequencies, and matching all variables to this level allows for a consistent dataset while still capturing much of the market volatility.

The dependent variable of the model is the electricity price for Norwegian price zone N01. This data is retrieved from the European electricity data platform Entso-e (2025) and is originally provided at a quarterly frequency. To align with the frequency of the independent variables, we calculate daily average prices.

The independent variables are selected based on their expected influence on electricity prices, such as weather conditions, energy commodity prices, exchange rates, and geopolitical events. Below is a brief overview of each variable, including the data source, unit, and motivation for inclusion in the model.

Since Norway is a net exporter of electricity, the exchange rate between the Norwegian krone (NOK) and the euro (EUR) may influence domestic electricity prices through its effect on cross-border trade. A weaker NOK can make Norwegian electricity cheaper for foreign buyers, potentially increasing exports and pressuring domestic prices. The exchange rate data, measured as NOK/EUR, is collected at daily frequency from Eikon.

As Norway exports electricity, prices in nearby countries are likely to influence Norwegian electricity prices. To account for this, we include daily spot prices from the SE3 price zone in Sweden, retrieved from Entso-e (2025). This variable serves as a proxy for regional market dynamics and external demand. The original day-ahead prices are reported on an hourly basis, but for consistency with the rest of the dataset, we compute daily averages. The unit of measurement is EUR/MWh.

Alternative energy sources such as oil and gas are included in the model due to their role as substitutes for electricity. Gas prices are collected from (Germany's Trading Hub Europe (THE) as it is the key destination for Norwegian pipeline gas exports/ the Dutch TTF as it serves at the benchmark for gas trading in continental Europe). Oil prices are represented by Brent crude, which is the most traded benchmark for oil and particularly relevant for the North Sea region. Daily spot prices for both gas and oil were retrieved from Eikon in the units EUR/MWh for gas and USD/barrel for oil.

To account for the effect of weather related demand fluctuations, particularly in winter when heating needs increase, we include average daily temperature data from five representative weather stations in southeastern Norway (geographic location of price zones N01 and N02). Lower temperatures are expected to drive demand for electric heating, and thus raise electricity prices. The temperature data is measured in °C and is obtained from Norsk Klimaservicesenter (2025).

Lastly, to capture major geopolitical shocks, we include a dummy variable for the Russian invasion of Ukraine, which had a great impact on European energy markets. This variable takes on the value 1 from 2022 onwards, reflecting the structural break associated with the war and energy crisis that followed.

Data preprocessing

In the process of data preprocessing, we chose to only focus on four sources of data out of the data collected in order to have a model that is easier to interpret and analyze. We chose to include electricity prices for NO1, exchange rates, Brent crude oil prices, and temperature readings. The goal of the preprocessing stage was to ensure that all datasets were consistent in format, aligned by date, and free of missing values.

The electricity price data for the NO1 region was extracted from a CSV file, where each observation contained a timestamp range and a corresponding EUR/MWh value. We first removed all missing entries, extracted the start time from the datetime range string, converted it to date format, and then aggregated the values to a daily resolution by computing the daily average price.

The exchange rate data consisted of daily NOK/EUR exchange rates. These values were parsed and converted to date format, and only the relevant columns were retained for merging.

For the Brent crude oil spot prices, we selected the daily closing price as the indicator of interest. After renaming and formatting the date column, the dataset was reduced to only include the date and the closing price, which we labeled Brent_oil.

The weather data required multiple cleaning steps. The timestamp column was cleaned to remove excess whitespace and then parsed using flexible datetime formats to ensure correct conversion. Temperature values were originally stored as strings with commas as decimal separators, which we replaced with periods before coercing the values to numeric. Finally, we aggregated the hourly temperature readings to a daily average for each date.

After independently cleaning each dataset, we performed inner joins on the "Date" column to create a complete dataframe that contained only the dates for which all variables were available. The final dataset includes daily values for electricity prices, exchange rates, crude oil spot prices, and average temperature. The cleaned dataset was exported as a CSV file "ARIMAX_Input_Data.csv".

Empirical analysis

Methodology

ARIMAX-Model

ARIMAX (AutoRegressive Integrated Moving Average with Explanatory variables) is an advanced time series model that uses both the ability to capture patterns as trends, seasonality and autocorrelation in data from the ARIMA-model, with extra external explanatory variables. The model makes it possible to analyze how several external factors directly affect a time series, in this case the electricity price in Norway. This provides a more comprehensive and precise understanding of how various factors interact and affect the price of electricity over time.

Electricity prices in Norway are complex and influenced by a range of external factors, such as weather conditions, price levels in the European market, and seasonal variations in energy usage. The ARIMAX model performs very well in this context because it can integrate both historical price data and external variables, providing a more precise and realistic picture of price development. The model not only allows us to understand past price fluctuations but can

also predict future prices based on changes in important explanatory variables (Torró, 2007). This is why we have chosen ARIMAX as one of the models to analyze which factors contribute the most to determining electricity prices.

Stationarity

In time series forecasting, stationarity is a key assumption for models such as ARIMA and ARIMAX. Stationarity implies that the statistical properties of the series, such as mean, variance, and autocorrelation, remain constant over time. Without this condition, model estimates may become unreliable, as patterns in the data may evolve and mislead the model's parameters.

There are two main types of stationarity, namely strict and weak. Most time series models assume weak stationarity, which holds if the mean, variance, and covariance between time steps remain stable. Non-stationary data can exhibit trends, seasonal components, or structural breaks that must be addressed before modeling (Svetunkov, 2023). In our case, we worked with daily electricity prices (EUR/MWh) for the Norwegian price area NO1. When visualizing the raw data, we observed apparent trend-like fluctuations and shifts in variance, suggesting the series was non-stationary. To ensure stationarity, we transformed the series by applying first-order differencing, defined as:

$$\Delta P_t = P_t - P_{t-1}$$

This transformation removes linear trends and stabilizes the mean of the time series. We did not apply a logarithmic transformation since our prices did not exhibit exponential growth patterns or multiplicative seasonality

In order to test for stationarity, we performed a ADF (Augmented Dickey-Fuller) test. The test checks the presence of a unit root, which is a sign of non-stationarity. The test uses the following regression model:

$$x_t = c + \beta x_{t-1} + \sum_{i=1}^{p-1} \phi_i \Delta x_{t-1} + e_t$$

Where the hypotheses are as follows:

4

$$H_0: \phi = 1 \text{ , The series has a unit root}$$

$$H_A: \phi < 1 \text{ , The series is stationary}$$

We used the adf.test() function from R's tseries package to run the test directly in our script. This built-in function computes the test statistic and corresponding p-value, allowing us to formally assess the stationarity of both the raw and differenced data.

The output from the adf.test() function P-value = 0.02029, which indicates stationarity at the 5% confidence level. We also perform the ADF test for the first differenced series where the P-value < 0.01, which indicates strong evidence of stationarity. The results confirm that the series is stationary at a level for use in our modeling and we can proceed with the forecasting.

Model parameters

The ARIMAX is characterized by three main parameters:

- p: the order of the autoregressive (AR) component

- d: the degree of differencing required to make the time series stationary

- q: the order of the moving average (MA) component

In addition to these, the ARIMAX model includes a set of exogenous regressors (X) that are assumed to influence the target variable. For our forecasting of electricity prices, we included Brent oil price, NOK/EUR exchange rate, and daily average temperature as relevant exogenous inputs.

To determine the optimal model structure, we used the auto.arima() function from the forecast package in R. This function automatically performs model selection by testing various combinations of (p, d, q) parameters and choosing the one that minimizes the Akaike Information Criterion (AIC). The differencing parameter d was chosen internally by the function based on tests for stationarity, while the best values for p and q were selected through an information-criteria-based search over different autoregressive and moving average terms.

By relying on the automated procedures in auto.arima(), we ensured an objective and statistically grounded approach to parameter selection. This not only reduced the need for manual tuning but also provided a reproducible and efficient pipeline for forecasting. The inclusion of exogenous variables allowed the model to account for key external drivers of

electricity price movements, complementing the internal temporal structure captured by the ARIMA components.
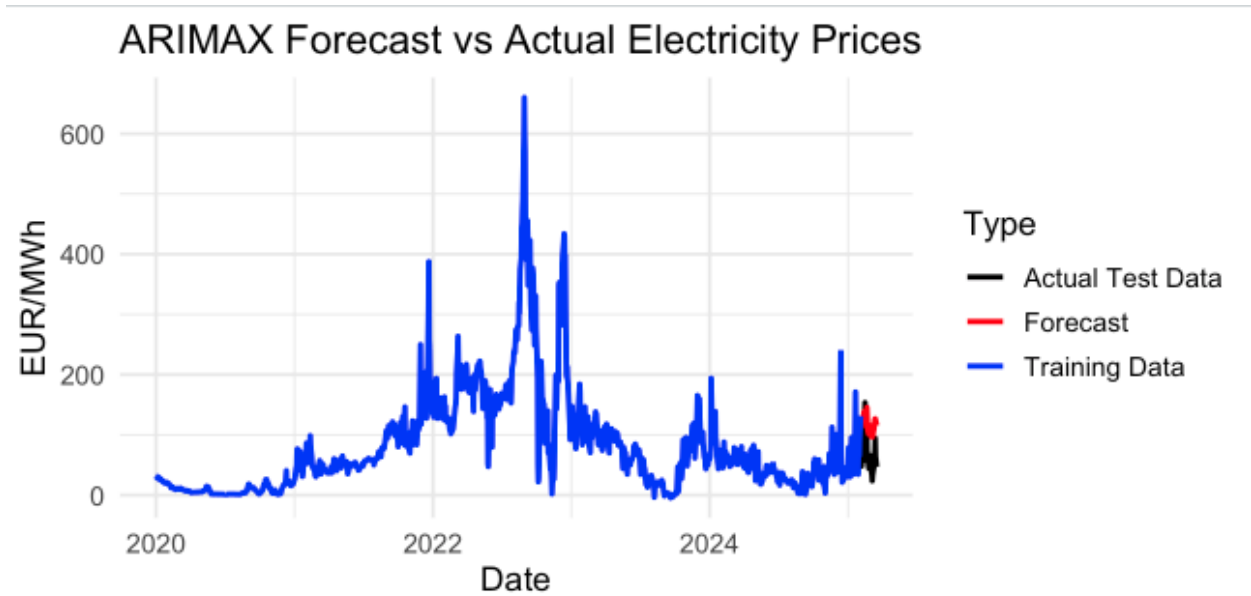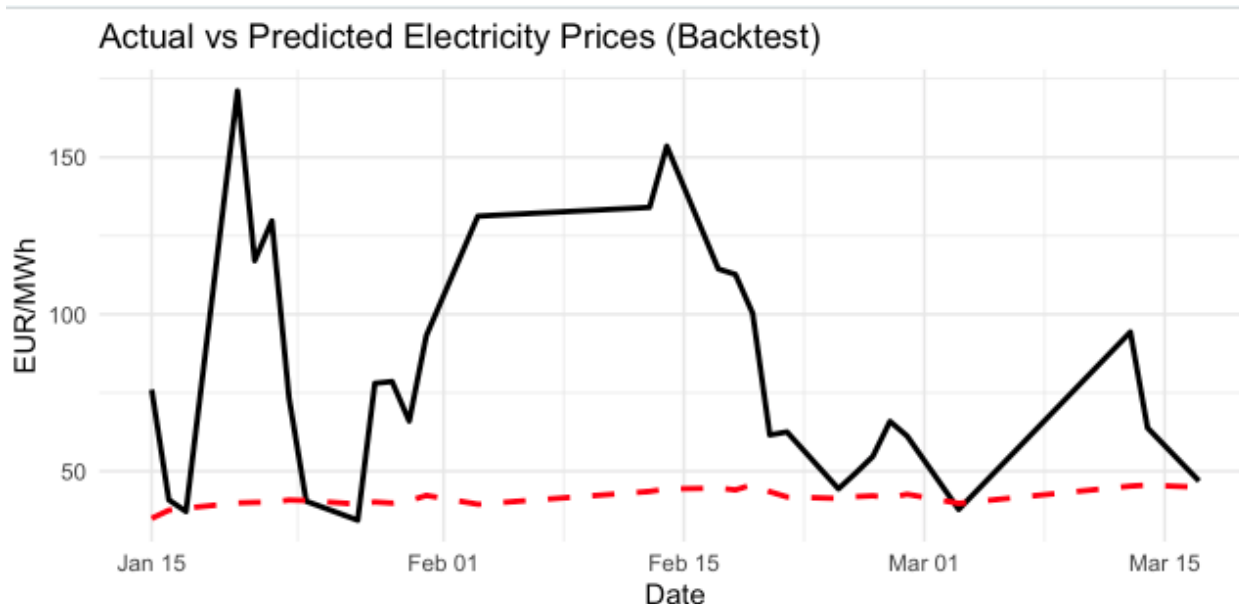
PLOTS

Results

Table 1 shows the results from running the ARIMAX model on the selected variables.

Variable	Coefficient	Stardard Error
Brent oil price	-0.3272	0.3176
Exchange rate	1929.79	1055.51
Temperature	-0.2829	0.0939
ar1, ar2, ar3, ar4		

Table 1: ARIMAX results

ARIMAX Forecast vs Actual Electricity Prices

Actual vs Predicted Electricity Prices (Backtest)

The coefficient for Brent oil price is negative, suggesting that an increase in oil price is associated with a decrease in electricity prices. This result is somewhat counterintuitive, as one might expect higher oil prices to increase electricity demand through substitution effects, thereby raising electricity prices. The link between electricity and alternative fuels may be weaker than initially anticipated. However, the standard error is nearly as large as the estimated coefficient, indicating low statistical significance. Therefore, we cannot confidently determine whether this observed relationship is meaningful or due to random variation.

The exchange rate shows a large positive coefficient, implying that a higher NOK/EUR rate leads to higher electricity prices. This may be explained by the increased competitiveness of Norwegian electricity in foreign markets when the NOK is weak, potentially increasing electricity exports and tightening domestic supply, driving domestic electricity prices up. The resulting t-value of 1.83 suggests that the coefficient is statistically significant at the 10% level.

The temperature variable has a negative coefficient, indicating that higher temperatures are associated with lower electricity prices. This supports the hypothesis that colder temperatures increase heating needs, raising electricity consumption and prices. The standard error leaves a t-value of approximately 3, indicating that this effect is highly statistically significant, even at the 1% level.

- R^2 - can we find the R squared value of the regression? How much of the variation in electricity prices can be explained by the variables we tested?
- MAE

- MSE

Conclusion

According to our empirical analysis, the main drivers of change in electricity price are temperature and exchange rate. As seen in Table 1, the coefficient of the exchange rate is much greater than the temperature coefficient, however the temperature can change in a larger range than the exchange rate typically does. The temperature is also estimated to have a more significant effect on electricity prices. Alternative fuel prices like oil seem to be less important when determining electricity prices.

Use of LLM

Throughout this project, the group used large language models, specifically ChatGPT, as a supportive tool. First and foremost, this AI tool was employed to efficiently identify relevant literature and facts, as well as to confirm our understanding of theories and concepts. In this way, ChatGPT served as a sparring partner, helping us clarify and refine our comprehension.

Regarding writing assistance for the report, ChatGPT was used to check for minor spelling and grammatical errors. This was beneficial, as it ensured consistency and professionalism throughout the entire report. ChatGPT helped to do this in an efficient manner, because it significantly reduced the time required for proofreading.

In terms of the programming work in this project, AI was very helpful when it came to the debugging phase. The code itself was written independently, with ChatGPT used only as a lookup for functions or R syntax. However, when the code failed to compile or did not work as intended, ChatGPT significantly reduced troubleshooting time by quickly identifying errors that we could then correct.

To summarize, large language models, specifically ChatGPT, were very helpful in reducing the time required to identify relevant theories, proofread the report and troubleshoot when the code was not working.

Copilot was also used to help with syntax, especially for installing correct packages and ensuring the usage of the correct kernel in the notebook.

Bibliography list

Adebiyi, A. A., Adewumi, A. O., & Ayo, C. K. (2014). Comparison of ARIMA and Artificial Neural Networks models for stock price prediction. *Journal of Applied Mathematics*, *2014*, Article ID 614342. https://doi.org/10.1155/2014/614342

Entso-e. (2025). *Day-ahead prices.* Retrieved March 18, 2025, from https://transparency.entsoe.eu/transmission-domain/r2/dayAheadPrices/show?name=&defaultValue=false&viewType=GRAPH&areaType=BZN&atch=false&dateTime.dateTime=17.03.2025+00:00%7CCET%7CDAY&biddingZone.values=CTY%7C10YNO-0--------C!BZN%7C10YNO-2--------T&resolution.values=PT15M&resolution.values=PT30M&resolution.values=PT60M&dateTime.timezone=CET_CEST&dateTime.timezone_input=CET+(UTC+1)+/+CEST+(UTC+2)

Norsk Klimaservicesenter. (2025). *Observasjoner og værstatistikk.* Retrieved March 18, 2025, from https://seklima.met.no/

SSB. (2025, March 14). *Electricity*. Retrieved March 18, 2025, from https://www.ssb.no/en/energi-og-industri/energi/statistikk/elektrisitet

Svetunkov, I. (2023). *Forecasting and Analytics with the Augmented Dynamic Adaptive Model (ADAM)* (1st ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781003452652

Torró, H. (2007). *Forecasting weekly electricity prices at Nord Pool*. University of Valencia. Retrieved from https://www.econstor.eu/bitstream/10419/74225/1/NDL2007-088.pdf