Norwegian School of Economics
Bergen, Spring, 2023

Spot Price Forecasting

Evaluating the Impact of Weather Based Demand Forecasting on Electricity Market Predictions

Vebjørn Sveva & Torolv Sundt

Supervisor: Geir Drage Berentsen

Master Thesis, MSc in Economics and Business Administration, Business Analytics

NORWEGIAN SCHOOL OF ECONOMICS

Acknowledgements

This thesis marks the culmination of our Master´s in Economics and Business Administration, symbolising the completion of our academic journey at the Norwegian School of Economics (NHH). Both of us share a common major in Business Analytics, a field that has exposed us to fascinating and demanding subjects. This specialisation forms the basis for our decision to focus our thesis on a related topic. Undoubtedly, this undertaking has presented us with significant challenges. However, it has also provided us with tons of knowledge and experience in our chosen topic, for which we are immensely grateful.

First and foremost, we would like to express our gratitude to our supervisor, Geir Drage Berentsen, and PhD candidate, Xuan Li, for their invaluable guidance and support throughout the process of our thesis. Their expertise and direction have been of most importance in ensuring the correctness and effectiveness of our work. Secondly, we would like to thank Nord Pool for giving us access to excellent data, used in this thesis. Lastly, we extend our heartfelt appreciation to our friends and family for their support and encouragement throughout our academic journey. Their motivation has played a crucial role in keeping us inspired and focused every step of the way.

Abstract

This thesis uses electricity data sourced from Nord Pool and weather data obtained from Norsk Klimaservicesenter, seeking to forecast day-ahead spot prices by leveraging temperature-based demand forecasts. Through this analysis, we aim to examine the feasibility of developing a model that can be utilised by participants in the electricity market bidding process. A significant portion of our research efforts has been dedicated to exploring a SARIMAX model, which is widely employed in this field of research. However, we have also thoroughly examined and tested various alternative models to assess their viability by considering them as potential benchmarks.

The thesis is structured into several chapters, beginning with an initial introduction that provides an overview of the electricity market in Norway. This section serves to establish the context and background for our research. Following the introduction, we delve into the presentation of the data and methods employed to address our research question. This chapter outlines the specific datasets utilised and the methodologies implemented in our analysis. Finally, we conclude the thesis by presenting our results and the implications our study might have for the participants in the Nord Pool day-ahead market.

Our findings reveal a notable spurious correlation between temperature and spot price. However, we acknowledge that relying solely on weather variables is insufficient due to the influence of external factors on pricing decisions. Nevertheless, our research has yielded satisfactory results, with the best models achieving an overall error ranging between 5-10%. Our main model consistently performed well, although there were instances where alternative models outperformed it on specific days or weeks.

Contents

1. Introduction **8**

2. Literature review **9**

2.1 Studies on weather effects on demand and price **9**

2.2 Studies on forecasting electricity prices **10**

3. Theory **13**

3.1 Organisation of the power market **13**

3.1.1 Price zones *15*

3.2 Norwegian electricity production **16**

4. Data **17**

4.1 Variable selection **17**

4.1.1 Norsk klimaservice data *17*

4.1.2 Nord Pool data *18*

4.1.3 Created variable *19*

4.2 Pre- processing **19**

4.2.1 Data cleaning *20*

4.2.2 Aggregation of weather variables *20*

4.3 Initial analysis of data **22**

4.3.1 Time period for analysis *23*

4.3.2 Seasonality *24*

4.3.3 Correlation *27*

5. Methodology **30**

5.1 Time series modelling **30**

5.1.1 Forecasting horizon *30*

5.1.2 Cross validation *32*

5.1.3 Transformation *34*

5.1.4 Model selection *35*

5.2 Forecasting models **36**

5.2.1 SNAÏVE *36*

5.2.2 ETS 37
5.2.3 SARIMAX 39
5.2.4 Artificial neural network 46
5.2.5 Combination model 47

5.3 Development of models **48**
5.3.1 Electricity demand 48
5.3.2 Spot price 54

6. Results **61**

6.1 Error evaluation tools **61**
6.1.1 Mean Absolute Error 61
6.1.2 RMSE 62
6.1.3 MAPE 62

6.2 Consumption forecast results **64**
6.2.1 Summer results 64
6.2.2 Winter results 65

6.3 Spot price forecast results **67**
6.3.1 Benchmark for MAE 67
6.3.2 Summer results 68
6.3.3 Winter results 71

7. Discussion **74**

7.1 Consumption forecasting results **74**

7.2 Spot price forecasting results **75**
7.2.1 Benchmark models 75
7.2.2 SARIMAX model 76

7.3 Critique **77**
7.3.1 Limitations of model variables 77
7.3.2 Limitations of methodology 78

7.4 Implications of study **79**

8. Conclusion **81**

References **83**

Appendix **87**

List of figures

Figure 3.1 *Supply and demand curve on the spot market* 14
Figure 3.2 *Norway´s five price zones* 15
Figure 3.3 *Distribution of energy production in Norway in 2021* 16
Figure 4.1 *Spot price yearly aggregated* 23
Figure 4.2 *Consumption hourly aggregated* 24
Figure 4.3 *Spot price hourly aggregated* 25
Figure 4.4 *Relationship between consumption and spot price* 28
Figure 4.5 *Relationship between temperature and consumption* 29
Figure 5.1 *Illustration of the time series cross validation process* 32
Figure 5.2 *Illustration of a multilayer feed-forward neural network* 46
Figure 5.3 *The construction-process of SARIMAX model* 50
Figure 5.4 *ACF/ PACF plot of seasonally differenced consumption data* 51
Figure 5.5 *ACF/ PACF plot of the optimal model´s residuals* 53
Figure 5.6 *ACF/ PACF plot of differenced spot price data* 56
Figure 5.7 *The neural network architecture with optimal parameters* 59
Figure 6.1 *Actual vs. forecasted consumption values for zone west - summer (09.06.- 16.06.2019)* 65
Figure 6.2 *Actual vs. forecasted consumption values for zone west - winter (10.02.- 17.02.2019)* 66
Figure 6.3 *Actual vs. forecasted spot price for zone west - summer (09.06.- 16.06.2019)* 69
Figure 6.4 *Actual vs. Forecasted spot price with the best model for zone west - summer (09.06.- 16.06.2019)* 70
Figure 6.5 *Actual vs. forecasted spot price for zone west - winter (10.02.- 17.02.2019)* 72
Figure 6.6 *Actual vs. Forecasted spot price with the best model for zone west - winter (10.02.- 17.02.2019)* 73

List of tables

Table 4.1 *Hourly average consumption across all zones for each weekday* 27

Table 4.2 *Hourly average spot price across all zones, given in NOK per Mwh* 27

Table 6.1 *Electricity demand forecast - MAPE - summer (09.06.- 16.06.2019)* 64

Table 6.2 *Electricity demand forecast - MAPE - winter (10.02.- 17.02.2019)* 65

Table 6.3 *Spot price forecast - MAE - summer (09.06.- 16.06.2019)* 68

Table 6.4 *Spot price forecast - MAE - winter (10.02.- 17.02.2019)* 71

1. Introduction

The electricity market plays an important role in the socio-economic development of a country, serving as a vital foundation for various sectors such as industry, transportation, and households. Recent years have witnessed an increasing emphasis on understanding pricing patterns, driven by several factors.

Norway's electricity market has experienced significant transformations over the past decade, fuelled by efforts towards liberalisation and integration with international electricity networks. These changes have made comprehending price dynamics more challenging, as they are influenced by various factors such as weather patterns, market demand, fuel prices and regulatory policies (Statnett, n.d.). Notably, regulatory policies focused on promoting and integrating green energy sources have further amplified the impact of weather conditions on the electricity market. This heightened dependence on weather patterns arises from the green energy boom and the existing reliance on weather for renewable energy generation.

Consequently, understanding specific weather patterns and their impact on the electricity market has become increasingly important. By gaining insights into these relationships, stakeholders can navigate the complexities presented by the interplay of weather, demand, and regulatory factors.

This thesis aims to contribute to a deeper understanding of the influence of weather patterns on the electricity demand and pricing. Through comprehensive data analysis and the application of forecasting techniques, this research will develop models capable of predicting electricity prices while considering the interdependence of weather conditions. We aim to develop models, based on the available data, that empower electricity buyers with the ability to make informed bids in the electricity auction, thus enhancing their decision-making process.

2. Literature review

The aspects brought to light in the introduction have led to increased amounts of research on weather effects on the electricity market as well as research on how to forecast electricity demand and spot prices. This is because electricity, by nature, cannot be stored and is therefore more vulnerable to uncertainty regarding climatic situations.

2.1 Studies on weather effects on demand and price

Kang and Reiner (2022) conducted a comprehensive investigation on the influence of weather conditions on household electricity consumption in Ireland. Their findings reveal a noteworthy relationship between temperature and electricity demand, indicating a consistent negative effect. A decrease in temperature leads to an increase in consumption, suggesting that colder weather prompts higher energy usage among households.
Additionally, the study explores the correlation between rainfall and electricity demand. It is observed that residents tend to stay indoors during rainy periods, leading to a subsequent rise in electricity consumption during those hours. This connection between rainfall and increased energy usage varies depending on the time of day.

Furthermore, Kang and Reiner (2022) examine the disparities in weather's impact on electricity demand between workdays and weekends. While both workdays and weekends demonstrate a negative relationship between temperature and demand, the study identifies certain variations in their respective influences. Specifically, weekends exhibit a higher sensitivity to temperature changes compared to workdays.
However, it's important to note that as Kang and Reiner (2022) focus exclusively on the household sector, the observed differences in weather impact may not be as significant when considering total consumption across all sectors as we do in this thesis.

In their study, Tanaka et al. (2022) examine the relationship between temperature and electricity spot prices, highlighting an indirect connection. They find that temperature directly influences the total electricity consumption, which in turn, has a direct impact on market prices. The research focuses specifically on various spot price zones in Germany, enabling the authors to analyse the weather effects on each region separately.

Through their analysis, Tanaka et al. (2022) generate forecasts that reveal distinct weather change effects on consumption for different regions. This indicates that the response of electricity consumption to weather fluctuations varies across spot price zones in Germany. Moreover, the study emphasises that while the effect of weather changes on the spot market may be smaller in magnitude, it remains significant and noteworthy.

By providing insights into the complex relationship between temperature, electricity consumption and spot prices, Tanaka et al. (2022) contribute valuable information on the impact of weather on the electricity market. Their study underscores regional disparities and the significance of weather-related factors in price dynamics.

2.2 Studies on forecasting electricity prices

In terms of forecasting electricity prices, there are several research articles dedicated to the topic. Weron (2014) reviews the available solutions, revealing several different modelling approaches. The article highlights the attractiveness of a statistical approach because it allows for a physical interpretation of its components, making them easier to understand and study. It is underlined that the statistical method stands a good chance in the power markets due to the constant seasonality prevailing in all periods.

Weron (2014) provides an overview of the commonly used statistical models in electricity price forecasting. One widely applied benchmark model is the *similar-day approach*, which often employs the *naïve* method as a simple implementation. Another crucial class of models for electricity price time series is the autoregressive models. Among them, the *AutoRegressive Moving Average* (ARMA) model stands as the standard model that considers both the random nature and time correlations of the data.

Extensions of the ARMA model are also discussed, including the *AutoRegressive Integrated Moving Average* (ARIMA) model. The ARIMA model incorporates differencing as a tool to handle non-stationary series. Additionally, the Seasonal AutoRegressive Integrated Moving Average (SARIMA) model is introduced. This model incorporates seasonal components into the models to capture recurring patterns in the data.

Moreover, Weron (2014) highlights that many research papers on electricity price forecasting propose time series models that incorporate exogenous variables such as consumption and temperature. To accommodate the inclusion of these exogenous factors, generalised versions of the basic models are employed. Specifically, the *ARMAX*, *ARIMAX*, and *SARIMAX* models are mentioned as the generalised counterparts of ARMA, ARIMA, and SARIMA, respectively. These models enable the integration of external variables, allowing for a more comprehensive analysis of the factors influencing electricity prices.

In Weron's study (2014), *Computational Intelligence* tools are also acknowledged for their valuable characteristics in electricity price forecasting. These tools possess the ability to handle complexity and non-linearity, which often makes them more suitable for modelling specific scenarios compared to traditional statistical methods. It is important to note, however, that this advantage does not necessarily guarantee better point forecasts in general. Instead, their strength lies in their effectiveness for short-term forecasting, as evidenced by the excellent performance reported by several authors in the field.

One of the primary classes of Computational Intelligence techniques highlighted in the study is *Artificial Neural Networks* (ANNs). ANNs are well-suited for both single period forecasting and multi-period forecasting, such as predicting electricity prices for one hour or for a day consisting of 24 hours. The flexibility and adaptability of ANNs allow them to capture intricate patterns and relationships within the data, enabling more accurate predictions in short-term forecasting scenarios (Weron, 2014).

We find that the most similar approach to our case is presented in Kristiansen's study (2012), where an autoregressive model is introduced for forecasting Nord Pool day-ahead prices. The model incorporates forecasted demand as an explanatory variable, alongside lagged price values and dummy variables to account for seasonality. This approach demonstrated strong performance, returning weekly and hourly percentage errors averaging around 5%.

In Tanrisevers' et al. 's research (2021), the authors discuss the increasing complexity of bidding strategies in Day-Ahead markets, leading to intricate combinatorial auctions. As a result, more advanced optimization techniques are needed to efficiently clear these markets. The article specifically highlights the prevalence of order types based on linear functions, which determine the quantity a bidder is willing to trade at a given price. The objective of the thesis is to provide forecasts ahead of the price setting, enabling market participants to better plan their bidding strategies in this complex environment.

In our opinion, while there have been numerous studies focusing on weather effects on demand and forecasting spot prices using demand as a factor, there are fewer studies that consider the indirect influence of weather on prices. We will study this influence during this thesis, with the aim of providing a comprehensive understanding of relationships between weather, demand, and prices in electricity markets.

3. Theory

In this section, the purpose is to give the reader a general understanding of spot price as a concept and how the power market in Norway works in general. We explain how the power market is organised, and how the pricing process is conducted by Nord Pool. In addition, a short section on the grid zones in Norway and Norwegian energy production will follow.

3.1 Organisation of the power market

The power market in Norway comprises both the *wholesale* market and the *end-user* market. In the wholesale market, large quantities of electricity are bought and sold by various participants such as producers, brokers, suppliers, energy companies and large industrial customers. Meanwhile in the end-user market, individual customers obtain electricity from power suppliers. The end-user market in Norway is divided into three equally sized segments: households, larger customers, such as stores and businesses, and industrial customers (Energifakta Norge, 2022).

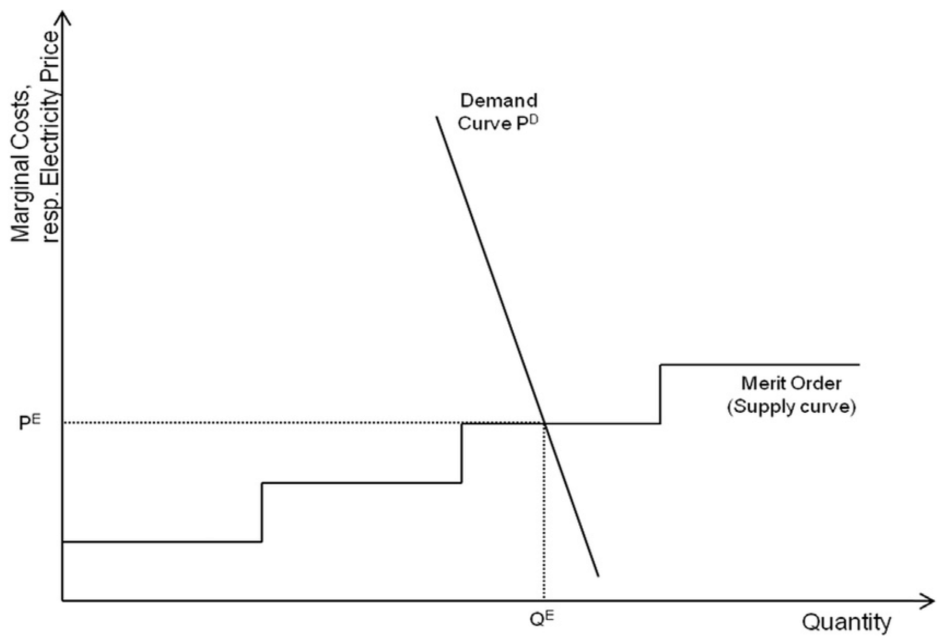
The *day-ahead* market, *continuous intraday* market, and *balancing* markets are the key components of the wholesale market. Among these, the day-ahead market is the primary platform for power trading and involves the largest volumes traded on Nord Pool. Each day at 10:00 CET, Nord Pool publishes the available capacities on interconnectors and in the grid, providing buyers and sellers with two hours to submit their final bids for the auction. The bidding process follows the *Single Day-ahead Coupling* (SDAC) initiative, which aims to enhance the overall efficiency of trading by establishing a pan-European cross-zonal day-ahead market (NEMO, n.d.).

During the auction buyers and sellers submit their bids, which are then matched with other orders. The price for each hour and bidding zone is determined at the meeting point of the sell price and buy price while considering network restrictions. The individual results are reported to buyers and sellers after publication, typically around 12:45 CET. Once the energy is purchased or sold, the customer is obligated to deliver or consume the corresponding amount of energy (Nord Pool, n.d.-a). This process ensures transparency and enables efficient trading in the day-ahead market.

The day-ahead market serves the essential purpose of establishing an equilibrium between the supply and demand of electricity. This equilibrium is crucial due to the inability to store electricity efficiently and the significant costs associated with supply failures. The day-ahead market achieves equilibrium by matching the bids made by producers indicating the quantity they are willing to supply at specified prices, with the bids made by consumers specifying the amount they plan to consume at different prices. This process ensures that the demand for electricity is met while minimising social costs (Energifakta Norge, 2022).

The hydropower's high storage capacity plays a vital role in the functionality of the Norwegian electricity system. The flexibility provided by hydropower enables easier matching of supply and demand in the Norwegian market. This flexibility allows hydropower producers to adjust their production levels based on the changing demand, contributing to the stability of the system (Energifakta Norge, 2021). *Figure 3.1* (Coester et al., 2018) illustrates the supply and demand curves in the spot market, visually depicting the relationship between price and quantity in the electricity market.

Figure 3.1
Supply and demand curve on the spot market



Overall, the goal of Nord Pool’s Day-ahead market is to maximise social welfare while taking into consideration network constraints provided by the transmission system operators (Nord Pool, n.d.).

3.1.1 Price zones

In addition to system prices, Nord Pool also sets area prices by considering congestion on the grid while balancing between purchase and sales bids. Norway's power grid system is divided into five zones, as seen in *figure 3.2* below (Statnett, 2023).

Figure 3.2
Norway´s five price zones

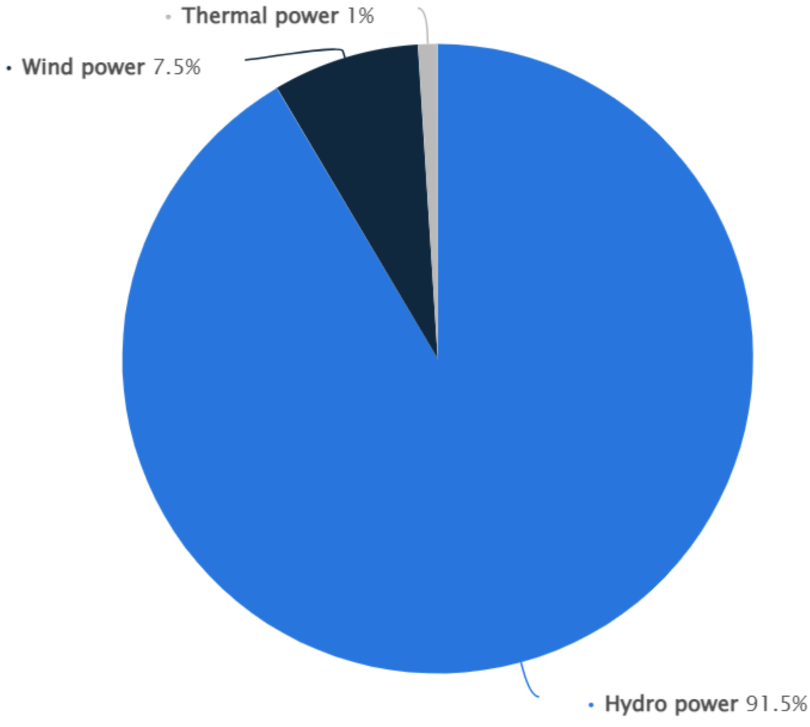
During this paper we will refer to NO1 as *east*, NO2 as *south*, NO3 as *mid*, NO4 as *north* and NO5 as *west*. These zones differ both in capacity on the supply side as well as on the demand side due to population differences.

Norway's weather-based power system creates varying situations throughout the country, hence the separation into zones. There is limited transmission capacity in the power grid system, with these bottlenecks proving a decisive factor for the different pricing of the areas (Statnett, 2023).

3.2 Norwegian electricity production

Norway is one of the leading countries in the world on renewable energy production and consumption. Over 90% of Norwegian energy production in 2021 came from hydropower, while the majority of the remaining 10% came from wind power, according to Statista (2023).

Figure 3.3
Distribution of energy production in Norway in 2021



A special feature of Norway's hydropower generation is its high storage capacity. More than 75% of the production capacity is flexible, meaning production can be rapidly increased and decreased at low cost. On the other hand, wind power is intermittent, which means electricity can only be generated when energy is available (Energifakta Norge, 2021).

4. Data

In this chapter, we present the data used throughout the analysis of the project, which is a combination of electricity market data supplied by Nord Pool and weather data gathered from Norsk Klimaservicesenter (Seklima, n.d.). All data gathered are hourly and spans over ten years, from 01.01.2013 until 31.12.2022. Additionally, we will present some initial analysis that has had an impact on our chosen methodological approach.

4.1 Variable selection

4.1.1 Norsk klimaservice data

In our thesis, we aim to examine the link between weather and electricity prices in Norway and develop a forecasting model that utilises weather data as a factor. While various weather phenomena such as wind, precipitation and temperature can influence electricity prices, we will focus primarily on temperature due to several justifications specific to the Norwegian context.

Firstly, when considering the importance of different weather types on spot prices, wind energy plays a relatively minor role in Norway's total electricity production mix. As a result, the contribution of wind energy to the overall supply is not significant, making it less relevant as a predictive factor for electricity prices. Similarly, variations in precipitation have minimal short-term impact on the availability of water resources for hydropower due to Norway's extensive network of water reservoirs, which very rarely reach critically low levels (NVE, n.d.). Consequently, precipitation is also excluded as a predictive factor.

Secondly, considering the appliance of our model on real-time weather forecasts, it is essential to consider the general accuracy of these forecasts for each weather metric. Temperature forecasts are relatively more accurate and less likely to exhibit significant fluctuations compared to wind and precipitation forecasts (Drigo, n.d.). Temperature is also less prone to displaying very local variations, which can be important when using multiple weather stations to cover an entire zone. Hence, including variables that have higher forecast uncertainty could lead to a less accurate electricity price forecasting model.

Taking these factors into consideration, we conclude that temperature has the most significant impact on electricity prices in Norway due to its direct link to consumption, as we will demonstrate later in the thesis.

Temperature

This variable contains hourly lowest observed temperature in each meteorological station considered, given in degrees Celsius. As Kang and Reiner (2022) depicts, there is reason to believe there is a direct relationship between temperature and consumption.

There are numerous alternatives for representing temperature in electricity consumption analysis, one common measure being heating degree days and cooling degree days. These measures might give a better picture in modelling consumption as they are created to directly assess the heating and cooling needs. However, the climate in Norway is colder than most countries, making the relationship between temperature and consumption for the most part flat negative. Hence, the use of actual observed temperature values is in our view more suitable for this project.

4.1.2 Nord Pool data

Spot price

Our main dependent variable in this project is the hourly day-ahead prices, expressed in Norwegian Kroner per megawatt-hour (NOK/MWh) for all five electricity zones. These prices are updated daily after the day-ahead price setting process conducted by Nord Pool. We will analyse and utilise historical data on these prices to develop our forecasting model and examine the relationship between prices and the weather variable we have selected.

Consumption

This variable represents hourly total consumed electricity in megawatt-hours (MWh/h) for all of Norway's five electricity zones, serving as an indicator of the demand side in the system price equilibrium model. As electricity consumption occurs in response to demand, this variable captures the total amount of electricity consumed in each hour for all zones in Norway. It plays a crucial role in understanding the relationship between electricity demand and the corresponding pricing dynamics. By incorporating this variable into our analysis, we can better assess the interplay between electricity consumption and pricing in the Norwegian electricity market.

4.1.3 Created variable

Weekday

This variable represents the day of the week in our time series dataset, providing information about the specific weekday for each observation. By including this factor variable in our model, we can capture and analyse patterns that are unique to specific weekdays. This variable essentially serves the same purpose as the dummy variable incorporated in the study developed by Kristiansen (2012). It is particularly valuable when dealing with time series data that exhibits weekly seasonal patterns, as the variable provides a more comprehensive understanding of the dynamics within the time series.

4.2 Pre- processing

In this project, the programming language *R* and the *integrated development environment* (IDE) *RStudio* are utilised for data handling and analysis. R is chosen due to its suitability for dataset analysis, along with its extensive collection of packages for data gathering, visualisation and time series modelling. The project can technically be divided into two main parts: pre-processing and analysis. Separate R scripts are employed for each part, with one script dedicated to pre-processing tasks that prepare the data for analysis. Another script contains the analytical steps, including model training, validation, and testing.

The following R packages are used in this project:
Tidyverse: A comprehensive collection of open-source packages (such as *ggplot2*, *dplyr*, etc.) that simplifies coding tasks by following a consistent design philosophy.
Readxl: Enables the loading and reading of Excel documents into R data frames.
Lubridate: Designed to facilitate working with dates in R, making date manipulation tasks more straightforward.

Stringr: Provides a set of functions specifically designed to simplify working with strings (text elements).
Strex: Contains additional string manipulation functions that complement the functionality provided by the stringr package.
Fpp3: A collection of packages created by the authors of the book "*Forecasting: principles and practice 3rd ed.*", Rob Hyndman and George Athanasopoulos. These packages cover a wide range of forecasting tasks, from time series data manipulation to the development of forecasting models.

4.2.1 Data cleaning

The cleaning process of this project is carried out with the goal of creating a single analysis-ready dataset that includes all utilised variables, matched on time. The steps involved in this process include merging the collected weather and electricity data for each zone, as well as replacing any missing values in the merged set.

4.2.2 Aggregation of weather variables

The weather data are aggregated based on the population distribution within each power zone. Here is a summary of the aggregation process for each zone:

East

This zone consists of two meteorological stations, Blindern in Oslo and Stavsberg in Hamar. The population in and around Oslo is the most densely populated area in Norway, so the weather station in Oslo has a greater impact on electricity demand. The impact of the weather stations is weighted with 85% for Oslo and 15% for Hamar, considering the potential different climate in the northern regions of eastern Norway. Mathematically, the aggregating can be written in the following manner:

$$temp_{East} = 0.85temp_{Oslo} + 0.15temp_{Hamar}$$

South

The *south* zone is complex to aggregate due to its large area, which includes the southern part of western Norway, the southern coast and the western parts of eastern Norway. To capture the climatic situation in this diverse zone, three meteorological stations are included. The most populated area in the southern part of western Norway is represented by a station in Sola, Stavanger, which is given a weight of 40%. The eastern area of the zone, including towns in Telemark and Vestfold, is represented by a station in Porsgrunn, weighted at 35%. The remaining 25% weight is assigned to the southern coast, represented by a station in Kjevik, Kristiansand.

$$temp_{South} \;=\; 0.4temp_{Stavanger} + 0.35temp_{Porsgrunn} + 0.25temp_{Kristiansand}$$

Mid

The *mid* zone includes the middle part of Norway, including Trøndelag and the northern area of the western coast. The population is evenly distributed between Trøndelag and the towns on the west coast such as Ålesund and Molde. Therefore, two meteorological stations are included with equal weights, one in Trondheim and the other in Vigra, Ålesund.

$$temp_{Mid} \;=\; 0.5temp_{Trondheim} + 0.5temp_{Ålesund}$$

North

The *north* zone covers all of Norway north of Trøndelag, which is a sparsely populated area. The population is concentrated in and around Tromsø, with smaller towns located further south. To represent these areas, one meteorological station is included in Tromsø and another in Skamdal, close to Mo i Rana, with equal weights.

$$temp_{North} \;=\; 0.5temp_{Tromsø} + 0.5temp_{Mo\ i\ Rana}$$

West

This zone includes the populated area in and around Bergen and extends east inland towards central parts of southern Norway, which have a different climate. The inland areas are sparsely populated, so most significance is given to the meteorological station in Bergen, with an assigned weight of 75%. The remaining 25% is assigned to a station in Sogndal.

$$temp_{West} = 0.75 temp_{Bergen} + 0.25 temp_{Sogndal}$$

Missing values occasionally occur in the weather data. In zones with two meteorological stations, missing values are replaced with the corresponding value from the other station at the same hour. In the *south* zone, which has three meteorological stations, missing values are replaced with the mean of the observations from the other two stations to simplify the process.

By aggregating the weather variables based on population distribution and handling missing values, the project ensures a representative view of the climate situation in each power zone for residents and industries in Norway.

4.3 Initial analysis of data

Performing initial analysis and detecting patterns in the data is crucial for model selection and development. Seasonalities and correlations between variables are important insights to consider. We create trend lines using the *smoothed conditional means function* (geom_smooth) from the ggplot2 package to help visualise these patterns and aid in understanding the data.

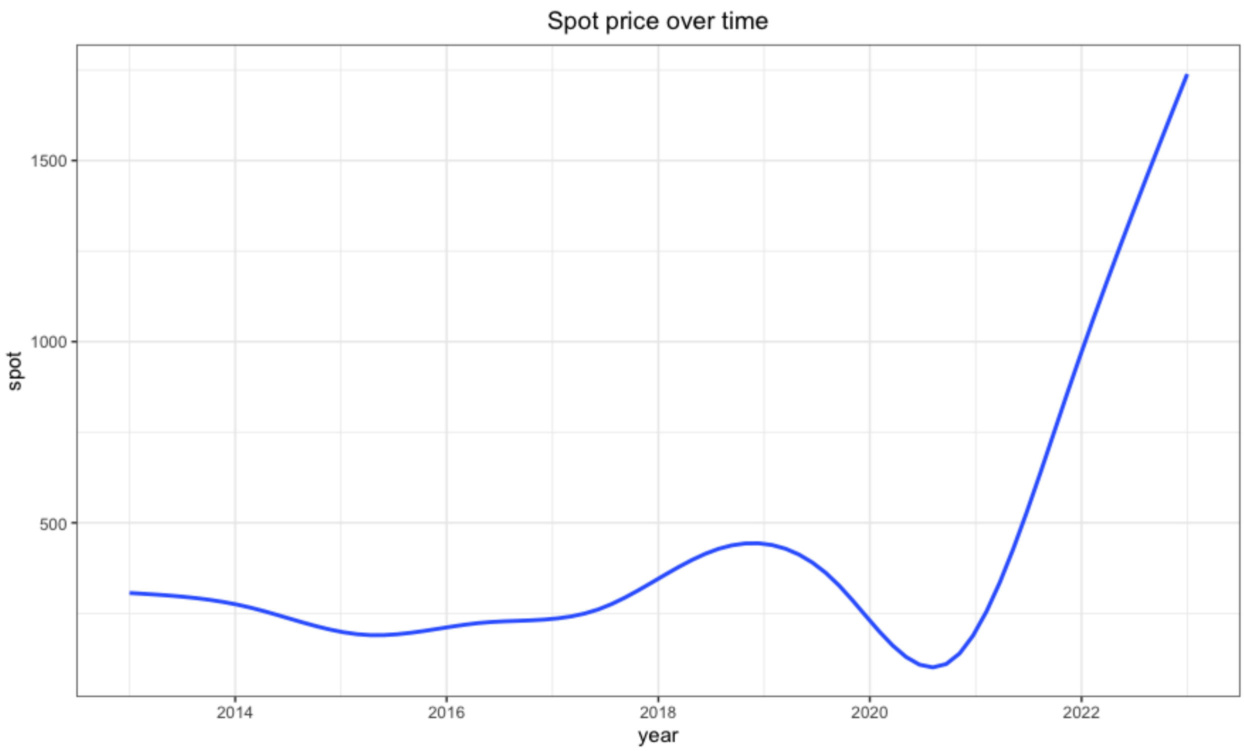
The smoothed conditional means function uses *locally estimated scatterplot smoothing* (LOESS) to create trend lines. LOESS is a technique that helps identify patterns in data by fitting a smooth curve to the scatterplot while considering the local neighbourhood of each point. This approach is particularly useful when dealing with datasets that have a large number of observations, making it difficult to plot each point individually (Wickham et. al., n.d).

By using the trend line, we can observe the general trend in the data and identify any notable patterns or trends. Conducting this initial analysis and visualising the data gives valuable insights into the data's characteristics, which can guide the model selection and development process.

4.3.1 Time period for analysis

In recent times, a confluence of events on the global stage has led to a significant surge in electricity prices throughout Europe. The primary factor behind this abnormal increase can be traced to a shortage of access to natural gas, compounded by a lack of Russian exportation. Several factors have contributed to this situation, including political tensions, disruptions in gas supply routes, and a shift towards renewable energy sources. Politically motivated conflicts and strained relations between certain nations have hindered the smooth flow of natural gas, leading to reduced availability and higher prices. Furthermore, the transition towards renewable energy, while commendable in the long run, has caused an interim dependency on natural gas as a backup source, intensifying the impact of its scarcity. As a result, Europe finds itself grappling with abnormally high electricity prices, Norway included (Statnett, n.d.).

Figure 4.1
Spot price yearly aggregated



As *figure 4.1* shows, the development in price from late 2020 have been extreme and has had little to do with abnormal weather or consumption. Hence, we will in this study focus on years pre-dating this development.

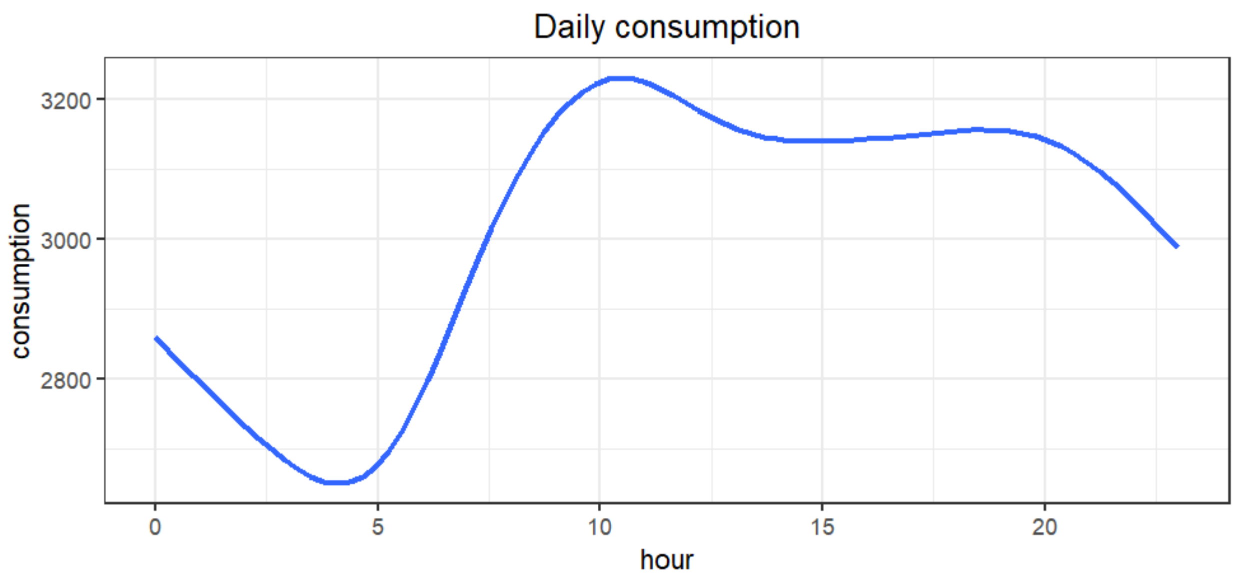
4.3.2 Seasonality

Our initial analysis uncovered multiple seasonalities in our data, especially concerning consumption, which is not surprising given its commonly observed relationship with temperature. We will present discoveries made below.

Daily seasonalities

The plots shown below indicate that there are daily seasonal patterns related to both consumption and price. The values are averaged across all zones, and hours.

Figure 4.2
Consumption hourly aggregated

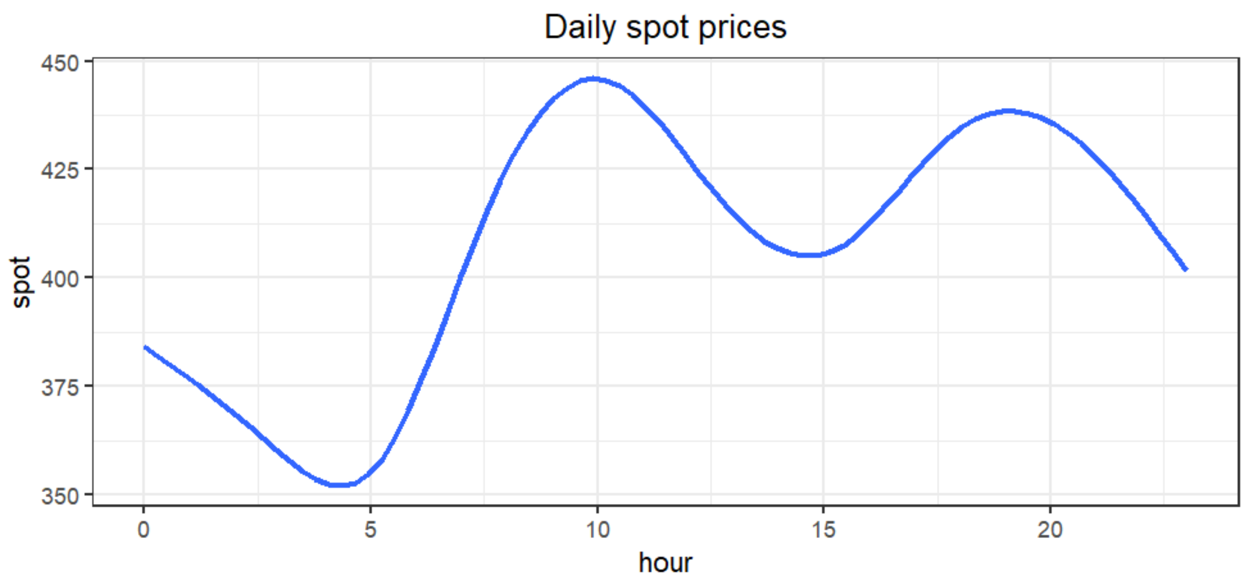


The pattern observed in *figure 4.2* provides valuable insights into the consumption behaviour in the electricity market. Commonly, a consumption peak appears daily around 10:00 - 11:00 and remains at high levels before declining rapidly from around 20:00 in the evening. The daily consumption peak during working hours aligns with the fact that approximately two-thirds of the total consumption is attributed to industry and businesses. As businesses and industrial factories typically operate at full capacity during the morning hours, the electricity demand increases, leading to the consumption peak around 10:00 - 11:00. This can be attributed to the increased energy usage in manufacturing processes, office buildings, and other commercial activities.

Furthermore, the sustained high levels of consumption throughout the day indicate that the demand remains relatively constant during working hours, reflecting the continued energy needs of industries and businesses. As the evening approaches, the electricity demand starts to decline rapidly from around 20:00, which can be attributed to the end of the workday and the reduced energy requirements during night-time hours.

It is also worth noting that residential homes typically experience higher heating demands during the morning hours, especially during colder periods, which can contribute to the consumption peak observed during that time. This can be attributed to the need for heating in residential buildings as people wake up and prepare for their day.

Figure 4.3
Spot price hourly aggregated



The graph of daily spot prices exhibits a similar pattern to the consumption plot, albeit with some differences. This observation suggests a potential correlation between the variables and provides valuable insights into the dynamics of the electricity market.

The observed similarity in patterns between the spot prices and consumption suggests a relationship between the two variables. This correlation can be attributed to the fundamental principles of supply and demand in the electricity market. As the consumption increases during

peak hours, the demand for electricity rises, which can put upward pressure on prices due to the limited availability of supply. Similarly, as the consumption decreases during non-peak hours, the demand decreases, which can lead to a decrease in prices.

The dip in prices appearing around 14:00 -15:00, distinct from the consumption pattern, indicates a unique behaviour in the electricity market. This dip can be influenced by various factors, such as changes in electricity supply or market dynamics. One possible explanation is the availability of additional electricity generation resources during that time, which can lead to increased supply and consequently lower prices. It could also be related to the behaviour of market participants, such as the scheduling of electricity production or the presence of specific contracts or pricing mechanisms during that time period.

The return of prices to levels similar to consumption towards the evening suggests a convergence of supply and demand conditions. As evening approaches and consumption decreases, prices have the potential to revert back to levels seen earlier in the day, assuming the supply conditions remain relatively stable.

Weekly seasonality

There are clear weekly seasonalities observed in the data. During weekends, electricity consumption is, on average, 150-200 MWh lower per hour compared to workdays. The lowest consumption is typically recorded on Sundays, as several businesses and industries are shut down during weekends. The table below shows average hourly electricity consumption in Mwh for each weekday, aggregated on all zones and rounded to the closest whole digit:

Table 4.1
Hourly average consumption across all zones for each weekday

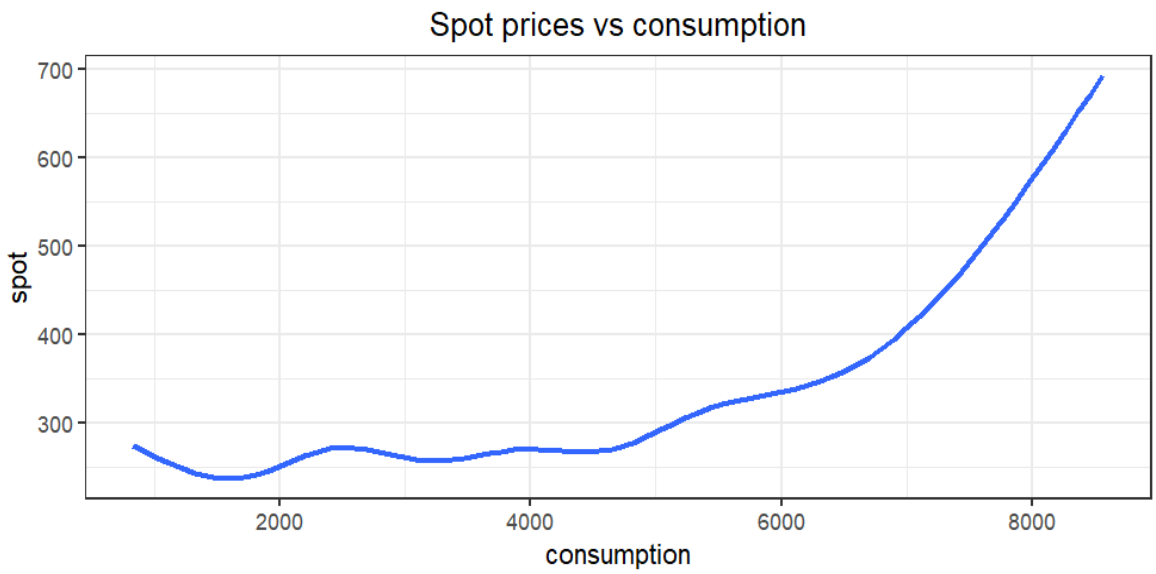
A similar pattern is present when aggregating hourly spot prices for each weekday. The similarities implies that there is a correlation between prices and consumption in our time series. Average price for each weekday in NOK is shown below, rounded to the closest whole digit:

Table 4.2
Hourly average spot price across all zones, given in NOK per Mwh

4.3.3 Correlation

The analysis shown above indicates that the price follows the consumption to a degree which implies that there might be a direct correlation between prices and consumption in our data, matching the discoveries made by Kristiansen (2012). Kang and Reiner (2022) define a direct negative relationship between consumption and temperature. These observations in addition to the article from Tanaka (2022) sets the foundation for our expectations when studying correlations between variables closer.

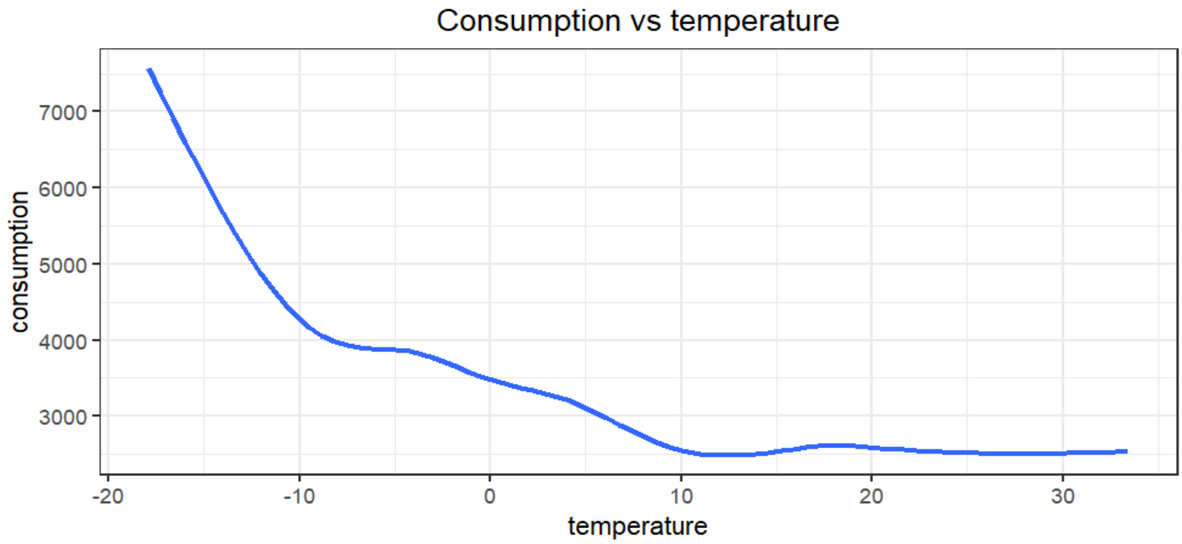
Figure 4.4
Relationship between consumption and spot price



The plot highlighting the relationship between energy consumption and spot prices provides a visual representation of the potential impact of consumption on prices. By filtering the data on pre-dating 2021, we aim to exclude the influence of external factors that have distorted the relationship between consumption and prices in recent times.

The clear pattern observed in the plot suggests that consumption indeed plays a significant role in influencing spot prices. As consumption levels increase, spot prices tend to follow a similar trend. This relationship aligns with the basic principles of supply and demand, where higher demand typically leads to higher prices.

Figure 4.5
Relationship between temperature and consumption



The above plot clearly demonstrates a noticeable pattern showcasing an inverse relationship between temperature and consumption. This indicates that temperature plays a crucial role in influencing consumption levels. This trend holds true for temperatures below approximately 10-12 degrees Celsius. The consumption stays generally close to a constant level above this point, showing that electricity consumption in Norway is almost exclusively used for heating rather than cooling. These findings have led us to believe that temperature might have a significant effect on the price, indirectly.

However, to include both temperature and consumption in the spot price model can complicate the interpretation of the price forecast, as already heavily correlated variables such as consumption forecasted via temperature and actual temperature might lead to multicollinearity. The results would be less reliable as well since the deployment of the model with forecasted temperature instead of actual observations would magnify the negative effect of weather forecast errors.

5. Methodology

This chapter will provide a comprehensive overview of the methodological approach employed in this project, focusing on the creation, deployment, and evaluation of our forecasting models. We delve into the theoretical underpinnings that are fundamental, covering essential concepts and principles. Furthermore, we will discuss the specific models we have utilised, highlighting their relevance and applicability. By presenting the theoretical foundation alongside the practical implementation of the models, we aim to provide an understanding of our methodology and its effectiveness in generating accurate forecasts.

5.1 Time series modelling

A time series refers to a sequential collection of observations recorded over time, typically at regular intervals. The specific time intervals considered may vary depending on the scope and nature of the project. Time series forecasting aims to predict the future continuation of such observations by analysing patterns inherent in the data. Time series data often exhibit various patterns, such as trends, seasonalities, or cyclic behaviours, making them valuable resources in forecasting projects. These time series are composed of different components, each representing an underlying pattern that contributes to the overall behaviour of the data (Hyndman & Athanasopoulos, 2021).

5.1.1 Forecasting horizon

In electricity price forecasting models, the forecasting horizons are typically categorised into short-term, medium-term, and long-term forecasts. Each category serves different purposes and timeframes in the energy industry.

Short-term forecasts are focused on hourly or daily predictions and are primarily utilised for day-to-day operational planning. They provide valuable insights into near-term electricity price fluctuations and are commonly used by market participants for activities such as scheduling energy generation and consumption, optimising trading strategies, and managing supply-demand imbalances.

Medium-term forecasts cover a time span ranging from a few days to a few months. They are employed for activities such as balance sheet calculations, risk management, and evaluating potential hedging strategies. These forecasts help market participants make informed decisions related to resource allocation, budgeting, and portfolio management.

Long-term electricity price forecasts extend over quarters or even years. They are typically utilised for strategic planning and investment decision-making. Market participants, including energy companies, investors, and policymakers, rely on long-term forecasts to assess the economic viability of new projects, evaluate profitability, and plan long-term energy procurement or infrastructure development (Weron, 2014).

In this project, the focus is on developing a short-term forecasting model for day-ahead electricity prices, aiming to assist market participants in making informed decisions regarding bidding prices and volumes for the upcoming day. The model utilises meteorological data to generate daily forecasts of hourly spot prices. By focusing on short-term forecasting, the model can leverage the more accurate and precise weather forecasts available for the near future.

Furthermore, studying patterns in different climatic situations is important as the impact of temperature on electricity demand varies throughout the year. By deploying the forecasting models on both winter and summer dates, the project aims to capture and analyse the distinct patterns and dynamics associated with different seasons. The combination of short-term forecasting and the consideration of different climatic situations enables the development of a more robust and accurate model for day-ahead spot price forecasts. This information can empower market participants to make well-informed bidding decisions, optimise their trading strategies, and effectively manage their operations in the electricity market.

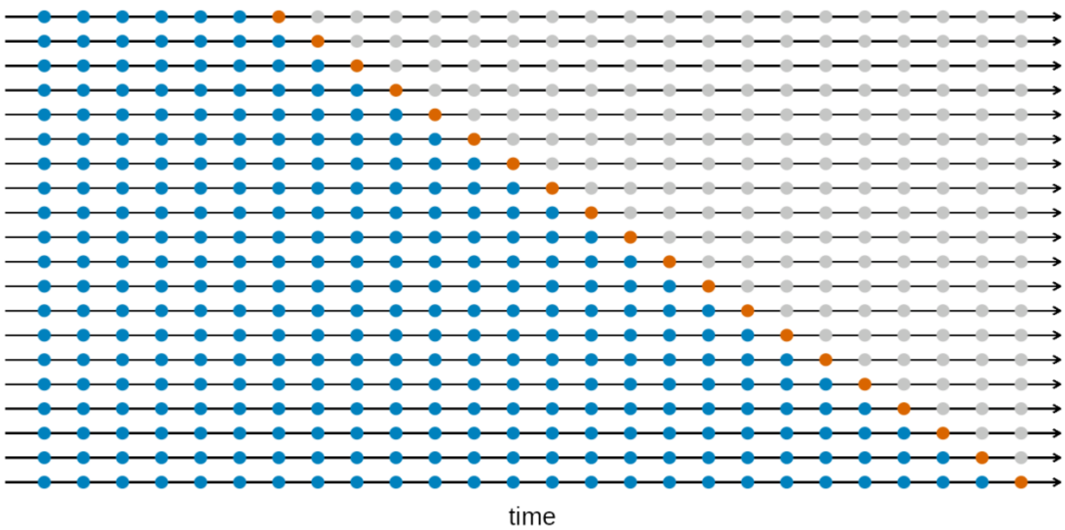
5.1.2 Cross validation

To ensure a reliable assessment of how well a forecast model generalises to unseen data, it is crucial to evaluate its performance on data that was not used during the model fitting process, which helps provide a reliable assessment of how well the model generalises to unseen data. To is commonly achieved by dividing into training and test sets, where the training data is used to estimate model parameters, while the test data used to evaluate the model's performance.

The division of data into training and test sets is commonly done by allocating a certain proportion of the full dataset to each set. It is typical to allocate around 75-80% of the data for training and the remaining portion for testing. This approach allows the model to learn from a substantial portion of the data while still providing an independent dataset for evaluation.

Another variation of the validation process is time series cross-validation. In this approach, a series of training sets and corresponding test sets are created. The test set typically consists of a single observation that immediately follows each training set. This method captures the temporal dependence in time series data and provides a more realistic evaluation of the model's performance (Hyndman & Athanasopoulos, 2021).
Chapter 5.10 in Hyndman & Athanasopoulos (2021) provides an illustration of the procedure:

Figure 5.1
Illustration of the time series cross validation process



The blue dots represent the training set while the orange ones represent the test set. The procedure starts with using a training set which contains n observations, forecasting n+1. The next step will include the original $n + 1$ value and forecast $n + 2$. To calculate the accuracy, one simply averages the error of the forecasted values from each "roll". Because of the nature of this procedure, the method is often referred to as an "evaluation on a rolling forecasting origin" (Hyndman & Athanasopoulos, 2021).

In our analysis, we are interested in forecasting day-to-day, which in our case with hourly data would mean forecasting 24 steps ahead over a period of one week. The initial training set will contain *one year* of data and will, for each roll, include one additional day - forecasting the next day. This will result in 7 folds of data containing forecasts from Monday till Sunday.

The decision to limit the training period to one year, specifically from summer/winter 2018 to the same period in 2019, is based on several factors. Firstly, we find that the accuracy of the forecasts does not significantly improve with training periods longer than a year. By focusing on a one-year training period, computational time can be optimised without sacrificing forecast accuracy. Secondly, when selecting the time period for analysis, the project aims to use data from "normal" years that are close to present time. This choice ensures that the training data reflects recent patterns and trends in electricity consumption and spot prices. By considering data from recent years, the models can capture the most relevant and up-to-date information for forecasting.

Additionally, the decision to start the training set one year before the chosen forecasting period (summer and winter) in 2019 avoids including a seasonal period twice. For example, if the training data were from January 2018 to July 2019, the months from January to July would be repeated in both the training and forecasting periods. This repetition could potentially introduce bias and unevenly fit the model, giving more weight to the seasonal impact of those specific months.

The time series cross-validation procedure has both been used to determine the accuracy of our models and to produce our final forecasts, as it is a good way of forecasting multiple steps ahead, with a rolling training set, resembling a real-life scenario with the latest data. As mentioned, we split the data used in a summer- and winter to capture the seasonal climate change. This is done both when addressing the optimal model to proceed with and when our final forecasts are computed.

5.1.3 Transformation

In time series forecasting, the objective is to accurately predict future values based on historical data and patterns. Simplifying these patterns can lead to improved model performance and more accurate forecasts. One approach to achieve this is through mathematical transformations. Two common methods are: taking the natural logarithm of the observations or using a power transformation (Hyndman & Athanasopoulos, 2021).

Power transformations involve raising the observations to a certain power, such as squaring or cubing them, to achieve the desired transformation. These transformations can help stabilise variances or linearise relationships between variables, making them more amenable to modelling and analysis.

Box-Cox transformation is a flexible technique that combines both logarithmic and power transformations. In this analysis, we will use a modified version of the Box-Cox transformation (Box, 1964) developed by Bickel and Doksum (1981), which allows for negative values (Hyndman & Athanasopoulos, 2021). Although consumption values are never negative, we choose this modified version to maintain consistency in our analysis, as we will apply the Box-Cox transformation to the spot price, which can be negative, to investigate if it improves the accuracy of forecasts or not.
The modified Box-Cox transformation can be represented as follows:

$$w_t = \begin{cases} log(y_t) & : \lambda = 0 \\ \frac{sign(y_t)|y_t^\lambda - 1}{\lambda} & : \lambda \neq 0 \end{cases}$$

Where:

w_t is the transformed observations,

y_t is the original observations and

λ is the transformation parameter.

The transformation will with $\lambda \neq 0$ use a power transformation. If $\lambda = 0$ a natural logarithm will be used. The estimation of λ can be obtained by utilising the R package fpp3, which uses the method presented in Guerrero (1993) to calculate the optimal value. The optimal value of λ will minimise the variation across the time series, often resulting in more accurate forecasts.

5.1.4 Model selection

Given the frequent instances where one must choose between models with varying parameters, the *Corrected Akaike Information Criterion* (AICc) plays a vital role in guiding the selection process towards identifying the optimal model. AICc is a bias-corrected version of *Akaike Information Criterion* (AIC) utilised to avoid small sample bias (Hyndman & Athanasopoulos, 2021). Even though one does not encounter this bias, the use of AICc is still generally used as the tradeoff in using AICc instead of AIC is minimal compared to the potential benefit in avoiding bias. As the extra calculations are no problem when working on a computer, the AICc will be used in this thesis. The AICc will take into account how well-fit the model is, along with penalising models which have an excessive number of parameters. The lower the AICc value is, the better fit the model has. The general AICc is typically modelled as following:

$$AIC = -2log(L) + 2k$$

$$AIC_c = AIC + \frac{2k(k+1)}{n-k-1}$$

Where:

L is the likelihood for the model,

k is the number of parameters estimated and

n is the number of observations in the data.

The model which returns the lowest AICc will be the one that fits the data the best, hence often producing good forecasts. When choosing between the same models with different parameters, we will, in most cases, utilise the AICc.

5.2 Forecasting models

Several forecasting models have been employed throughout this project. The main model used to incorporate exogenous effects into our model is the SARIMAX model. Additionally, we have included a variety of benchmark models to gain a better understanding of the data's behaviour from a forecasting perspective and to assess the actual impact of the exogenous variables. Statistical models deployed include SARIMAX, SNAÏVE and ETS. Furthermore, we have included a Neural Network as a computational intelligence model. Lastly, we also include a combinational model to study the effect of aggregating the different forecasting models.

5.2.1 SNAÏVE

The Seasonal Näive (SNAÏVE) model is based on the naïve method, where all forecasts are set to the value of the last observation. The difference is that instead of simply copying the last observation, the SNAÏVE sets each forecast to be equal to the last observation from the same season. This means that forecasts for a particular hour are set to the same hour of the previous day/week, or forecasts of a particular day are set to the same day of the previous month (Hyndman & Athanasopoulos, 2021). The SNAÏVE can be expressed as following:

$$\hat{y}_{t+h|t} = y_{t+h-m(k+1)}$$

Where:

h is the forecast horizon,

$\hat{y}_{t+h|t}$ is the forecasted value h timesteps ahead,

m is the seasonal period and

k is the integer part of $(\boldsymbol{h-1})/\boldsymbol{m}$, which represent the number of seasonal periods that have occurred during h.

The overall simplicity of the SNAÏVE model makes it suitable as a benchmark for comparing against other more computationally intensive models. The seasonal component in the model even makes it a suitable candidate as an accurate model for scenarios where values between periods do not differ heavily from each other (Hyndman & Athanasopoulos, 2021).

5.2.2 ETS

Exponential smoothing is a forecasting method that can be considered "one step" more complex than typical naïve methods. The easiest way in looking at the exponential smoothing model, is to think of it as a combination of the naïve and the simple average models. While the naïve method is "only" looking at the most recent observations and considers only those as important for future forecasts, the average method works as the polar opposite; every historical observation is taken into account, treating them as equals for future forecasts. Both methods can be seen as quite extreme, and a combination of these is often a preferred option. This is where the exponential smoothing model comes into play.

The idea behind exponential smoothing is toincorporate not only the most recent observation, but also older observations. However, the weight assigned to each observation diminishes as the observation gets older, meaning recent observations carry more weight in the forecast compared to older observations. This weighting scheme allows for a flexible and adaptive approach to capturing patterns and trends in the data (Hyndman & Athanasopoulos, 2021). A simple exponential smoothing (SES) model´s forecast equation can be represented as follows:

$$\hat{y}_{t+1|t} = \alpha y_t + \alpha(1-\alpha)y_{t-1} + \alpha(1-\alpha)^2 y_{t-2} + \dots$$

where:

$\hat{y}_{t+1|t}$ is the value of a one-step ahead forecast,

α is the smoothing parameter, taking a value between 0 and 1 and

y_t is the value of the observation at time t.

In 1957, Holt introduced an enhancement to exponential smoothing by incorporating a trend component into the model (Holy, 1957). This advancement allowed for the capture of trend patterns in time series data. Three years later, Winters further extended the model by including the ability to capture seasonal patterns as well (Winters, 1960). With these additions, the exponential smoothing model evolved into what is commonly known as ETS (Error, Trend, Seasonality), which encompasses the error term as well as trend and seasonal components in the forecasting equation.

However, the Holt-Winters (ETS) equation varies depending on the nature of the seasonal component. The seasonal component can be split into two varieties: additive and multiplicative methods. The preferred method depends on the characteristics of the data being analysed. If the seasonal variation remains relatively constant throughout the time series, an additive method is preferred, whereas if the seasonal variation varies in proportion to the time series, the latter is favourable (Hyndman & Athanasopoulos, 2021).
The following forecast equations can be modelled as:

Additive

$$\hat{y}_{t+h|t} = \ell_t + hb_t + s_{t+h-m(k+1)}$$ *(forecast equation)*

$$\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$$ *(level component)*

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$$ *(trend component)*

$$s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}$$ *(seasonal component)*

Where:

α, β, γ are smoothing parameters taking values between 0 and 1, respectively representing weights for the level, trend and seasonality,

m is the seasonal period,

h is the forecasting horizon,

$\hat{y}_{t+h|t}$ is the forecasted value h timesteps ahead and

k is the integer part of $(h-1)/m$, which represent the number of seasonal periods that have occurred during h.

Multiplicative

$$\hat{y}_{t+h|t} = (\ell_t + hb_t)s_{t+h-m(k+1)}$$ *(forecast equation)*

$$\ell_t = \alpha \frac{y_t}{s_{t-m}} + (1-\alpha)(\ell_{t-1} + b_{t-1})$$ (*level component)*

$$b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$$ (*trend component)*

$$s_t = \gamma \frac{y_t}{(\ell_{t-1} + b_{t-1})} + (1-\gamma)s_{t-m}$$ *(seasonal component)*

Where:

All parameters represent the same as in the additive model.

5.2.3 SARIMAX

The Autoregressive Integrated Moving Average (ARIMA) makes up the foundation for the most significant model applied in this project, the SARIMAX. It is essential to understand several key concepts before delving into variations of ARIMA forecasting models. This subsection will present the theoretical framework for the concepts utilised in the development of ARIMA models.

Stationarity

A time series process is considered stationary if its properties remain unchanged by a change of time origin, meaning the joint probability distribution must be unaffected by shifting times of observations backwards and forwards. In practice this means that for stationary models, mean and variance remain constant over time (Jain & Singh, 2003). In many cases the data is not stationary, especially in the case of seasonality dominated series such as the electricity consumption and the electricity price data studied in this project. The ARIMA process incorporates steps for making a non-stationary time series stationary, and there exists commonly applied methods to determine whether a time series is stationary or not.

Differencing

Time series differencing is a technique used to stabilise the mean and variance of a time series by eliminating any trend or seasonality present. The mean can be stabilised by removing any trend, for example by fitting a trend line and subtracting it prior to fitting the model. Variance can be stabilised by the use of transformations, which means converting the data to contain period-to-period differences rather than actual values (Nau, 2020). Consequently, the differenced series contains changes between consecutive observations in the original series and can be written as follows:

$$Y'_t = Y_t - Y_{t-1}.$$

Seasonal differencing

Seasonal differencing is distinct from standard differencing in that it captures seasonal changes instead of periodic changes, focusing on changes from one season to the next. For example, in an hourly dataset with daily seasonality one season will contain 24 hours, meaning the seasonal difference is:

$$Y_t - Y_{t-24}.$$

Stationarity testing

A unit root test is used in time series analysis to test whether the data is stationary by looking for the presence of a unit root. In unit root testing, it is assumed that the time series being tested can be presented in the following manner:

$$Y_t = D_t + z_t + \varepsilon_t.$$

Where:

D_t represents a deterministic component,

z_t represents a stochastic component and

ε_t is a stationary error.

The test aims to confirm whether z_t contains a unit root.

A commonly used tool to test for stationarity in time series data is the *Kwiatkowski-Phillips-Schmidt-Shin* (KPSS)-test. Like other hypothesis tests, the KPSS-test contains a null hypothesis that is evaluated on the provided data to determine the plausibility of the hypothesis. The p-value returned from the test serves as the evaluator and in the KPSS-test, the null hypothesis assumes that the data is stationary. Small p-values means we reject the null hypothesis - the data is not stationary. Large p-values however means we confirm the null hypothesis - the data is stationary.

The KPSS-test is based on a linear regression, assuming the following equation:

$$Y_t = \xi t + r_t + \varepsilon_t.$$

The equation consists of a deterministic trend ξt , a random walk r_t and a stationary error ε_t . r_t can be defined as:

$$r_t = r_{t-1} + u_t,$$

Where u_t are independent and identically distributed random values $(0, \sigma^2{}_u)$. The stationarity hypothesis is $\sigma^2{}_u = 0$, resulting in a constant random walk for r_t. Since ε_t is assumed to be stationary, if there is no trend in the data the model is stationary around a level r_0 (Kwiatkowski et al., 1992).

To determine whether seasonal differencing is appropriate, we can measure the seasonal strength of the model. The strength of seasonality can be defined as follows:

$$F_s = max(0, 1 - \frac{Var(r_t)}{Var(s_t + r_t)})$$

Where:

s_t is the seasonal component,

r_t is the remainder component, and

Var indicates the variance.

First part of the equation specifies the minimum value as 0, meaning that the strength is measured on a scale from 0 to 1, with a higher value indicating stronger seasonality. The relationship between the variance of the remainder and the seasonal component is the key to measure the seasonal strength. A significantly higher variation in seasonal variance means the second part of the equation returns close to zero, which in turn means the equation returns a value close to 1. Vice versa, a higher remainder variance will return F_s close to 0. When applying the test, a seasonal differencing is suggested if the seasonal strength $F_s >= 0.64$ (Hyndman & Athanasopoulos, 2021).

Autocorrelation

Autocorrelation refers to the measure of the linear relationship between lagged values of a time series. In the presence of autocorrelation, random errors are often positively correlated over time, indicating that each random error is more likely to be similar to the previous random error than if the errors were independent of each other. The correlation coefficient between two values in a time series is called the *autocorrelation function* (ACF). For a time series Y_t, the ACF can be expressed as:

$$Corr(Y_t, Y_{t-k}) \qquad : k = 1, 2, \dots$$

Where the value of k is the considered number of lags.

Another method of looking at autocorrelation is to focus on the direct association between Y_t and Y_{t-k}, filtering out the impact of the values in between. By doing a transformation on the time series and then calculating the correlation of the transformed series we obtain the *partial autocorrelation function* (PACF).

A common practice when studying autocorrelation is to plot these functions as ACF and PACF plots. These plots provide a graphical representation of the functions that are easier to interpret, which makes them useful as tools to identify orders in the model (The Pennsylvania State University, n.d.).

Autoregressive component

In *autoregressive* (AR) models, the variable of interest is forecasted using a linear combination of its own lagged values:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots \phi_p Y_{t-p} + \varepsilon_t$$

Where ϕ_p is the autoregressive coefficient for the belonging lag, and ε_t are the error term at time t. An advantage of autoregressive models is their flexibility in terms of dealing with a wide variety of different patterns in time series data (Hyndman & Athanasopoulos, 2021).

Moving Average component

As with autoregressive models, *moving average* (MA) models study past observations to predict the future. However, rather than looking at actual observations in the past, these models use past forecast errors to predict, using the following model:

$$Y_t = c + \varepsilon_t + \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \ldots + \theta_q\varepsilon_{t-q}$$

Where θ_q is the coefficient of the moving average term and ε_t denotes white noise.

ARIMA

The concepts of autoregression, differencing, and moving average are combined into ARIMA, constituting the components, and being denoted as parameters in the model. The model is classified as “ARIMA (p,d,q)”, where:

- p is the number of autoregressive terms
- d is the number of differences needed for stationarity
- q is the number of lagged forecast errors

The general forecasting equation for a differenced series can be written as

$$Y_t{}' = c + \phi_t Y'_{t-p} + \theta_q\varepsilon_{t-q} + \varepsilon_t$$

where $Y_t{}'$ is the differenced series, while the right side of the equation consists of lagged values of the time series along with lagged errors. The model also contains a constant *c*, which mainly has an impact on long-term forecasts (Hyndman & Athanasopoulos, 2021). By letting y denote the d^{th} difference of Y, we can describe Y_t as follows:

- $d = 0 : Y_t = y_t$
- $d = 1 : Y_t{}' = y_t - y_{t-1}$
- $d = 2 : Y_t{}'' = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) = y_t - 2y_{t-1} + y_{t-2}$.

For clarification, we would like to specify that for $d = 2$, the model does not display the difference from two periods ago, but rather signifies a difference of an already differenced Y (Nau, 2020).

Identifying order of ARIMA models

The first and most important step when fitting an ARIMA model is to identify the order of differencing needed to make a time series stationary, by applying tests such as the KPSS test.

After achieving a stationary time series, the next step is to identify numbers of AR and MA terms needed by looking at the ACF and the PACF plots (Nau, 2020). These are the general rules when identifying terms with ACF and PACF plots:

- **AR process (p, d, 0):**

A gradually declining ACF plot, combined with a sharp cut-off in PACF, indicates an AR process. The number of AR terms is decided by the number of significant PACF lags before the sharp cut-off.

- **MA process (0, d, q):**

A gradually declining PACF plot, combined with a sharp cut-off in ACF, indicates a MA process. The number of MA terms is decided by the number of significant ACF lags before the sharp cut-off.

- **ARMA process (p, d, q):**

If both plots are gradually declining, it is an indication of an ARMA process with both autoregressive and moving average terms. The number of terms to include for each component is not obvious, usually being decided by testing for different values and studying the estimation results with help from information criteria such as the AICc (Virenrehal, 2022).

SARIMA and SARIMAX

ARIMA models, in themselves, lack the ability to capture seasonal effects in time series. The Seasonal Autoregressive Integrated Moving Average (SARIMA) model is a version of the ARIMA model which adds a seasonal part to represent the seasonal effect. The components of the seasonal part have the same structure as the non-seasonal part but differ in that they operate over multiple lags equal to the number of periods in a season.
The model can be classified as ARIMA(p,d,q) x (P,D,Q), where:

- P denotes the number of seasonal AR-terms
- D denotes the number of seasonal differences
- Q denotes the number of seasonal MA-terms (Nau, 2020).

The SARIMAX further extends on SARIMA by adding exogenous predictor variables. This model is especially useful in cases with high outside influence on the predicted variable coming from other exogenous factors, as it is able to handle these external effects.

Identifying orders of SARIMAX model

Identifying orders of SARIMAX models follows a similar procedure as ARIMA models, starting with identifying whether differencing is needed. However, one should first determine whether a seasonal differencing is needed and then decide on whether a non-seasonal differencing is required. The main rules for identifying orders of SARIMAX models are as follows:

- If the seasonal pattern is strong and consistent, it is recommended to use one order of seasonal differencing. However, it is advised to not use more than one order of seasonal differencing or more than two orders of differencing in total.

- **Seasonal AR process (P, D, 0):**

A gradually declining ACF plot, combined with a sharp cutoff in the PACF plot, indicates a seasonal autoregressive model. We expect significant spikes once every seasonal period, for example every 7th lag for a daily series with weekly seasonality.

- **Seasonal MA process (0, D, P):**

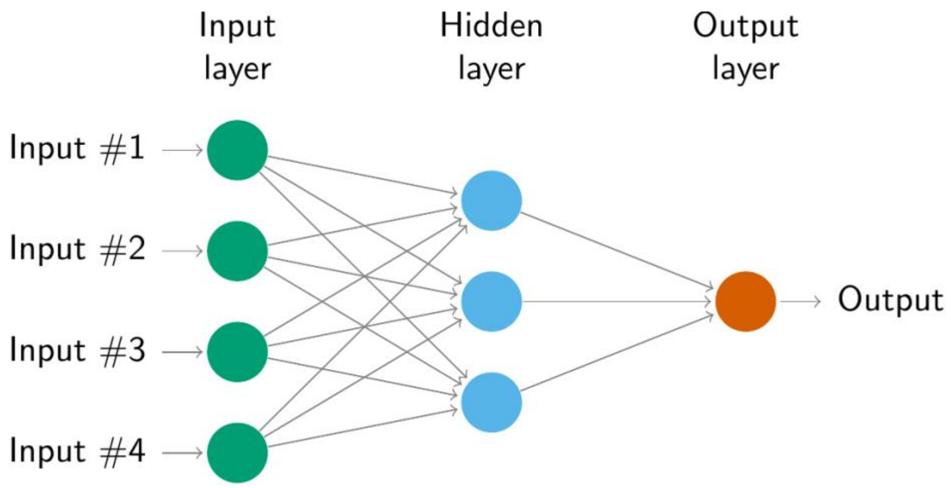
A gradually declining PACF plot, combined with a sharp cutoff in the ACF plot, indicates a seasonal moving average model. As with AR processes, we expect significant spikes once every seasonal period.

In general, a rule of thumb is to try avoiding using both terms, as this is likely to lead to an overfitted model (Virenrehal, 2022).

5.2.4 Artificial neural network

Computational Intelligence models create approaches that are capable of adapting to complex dynamic systems, by combining elements like learning and evolution. Neural network models are designed to mimic how the human brain works - of course in a very simplified manner. Their design can be seen as a network of interconnected *neurons* organised in *layers*. Typically in forecasting, the network consists of an *input layer* (predictors), one or multiple *hidden layers* (which makes the model non-linear with the use of activation functions) and an *output layer* (output/forecasts). This type of network is more commonly known as a *multilayer feed-forward network* (Hyndman & Athanasopoulos, 2021).
Chapter 12.4 in the book by Hyndman and Athanasopoulos (2021) provides an illustrative depiction of this type of network.

Figure 5.2
Illustration of a multilayer feed-forward neural network



Each layer receives inputs from the previous and produces an output, which is then fed forward to the next layer. The connections between every neuron have weights which differ. These weights initially adopt random values and are subsequently refined through a learning algorithm that aims to minimise the discrepancy between the predicted output and the actual output. To prevent the weights from growing excessively, a decay parameter is introduced and set to a value of 0.1 (Hyndman & Athanasopoulos, 2021).

The general linear function of weights assigned to the input layer can be defined as follows:

$$z = W \cdot x_t + b_i$$

This weighing process is then modified in the hidden layer, making it non-linear. For example, with the use of a *sigmoid* function:

$$f(z) = \frac{1}{1 + e^{-z}}$$

Where:

W is the weights assigned to the inputs,

b_i is the bias to each input,

x_t is the inputs and

f represents the sigmoid activation function.

When working with specific time series (as presented), it is typically advantageous to use lagged values of the response variable as inputs in the network. This approach makes the use of a *neural network autoregression* (NNAR) convenient.

When there are seasonal data, it is common to use both a number of lagged values (p) and a number of last observed values from the same season (P) as inputs (Hyndman & Athanasopoulos, 2021). If we combine these parameters with the number of nodes in the hidden layer (k) we end up with a seasonal autoregressive neural network with the notation of $SAR - NN(p, P, k)$.

5.2.5 Combination model

Combining the strengths of multiple forecasting approaches has been widely acknowledged as an effective strategy to enhance forecast accuracy. In a seminal article by Clemen (1989) titled "*Combining Forecasts: A Review and Annotated Bibliography*," the author concludes that the consensus among studies is clear: combining multiple forecasts results in improved forecast accuracy (Clemen, 1989).

Building upon this research, we propose a hybrid forecasting model that leverages the predictive capabilities of the several included models presented in this chapter. Each of these models brings unique strengths to the forecasting process.

The SARIMAX and ETS models are well-suited for capturing trend and seasonality patterns in time series data. They consider the historical patterns and dynamics to make accurate predictions. On the other hand, the neural network model excels at capturing complex nonlinear relationships within the data due to its ability to learn intricate patterns and dependencies. To provide a baseline reference, we also include the seasonal naïve model, which relies solely on historical observations and captures the seasonal patterns present in the data.

There are several ways to combine forecasting models. However, as stated Hyndman and Athanasopoulos in the book "*Forecasting: Principles and Practice 3rd ed.*": "using a simple average has proven hard to beat." (Hyndman & Athanasopoulos, 2021). Therefore, we propose the use of a model averaging algorithm.

5.3 Development of models

All models studied in this project have been developed through R with corresponding packages. We will present our approach in forecasting the electricity demand before we describe the implementation of the different types of models applied in forecasting the price.

5.3.1 Electricity demand

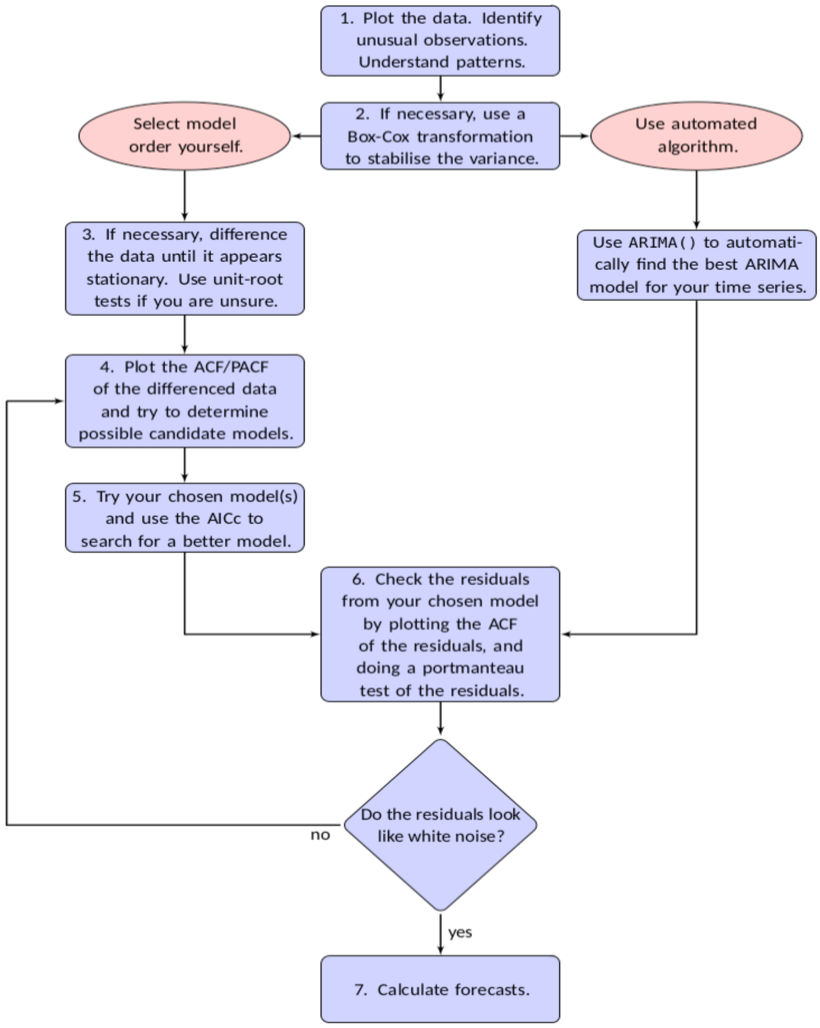
From the basis of observations made in the initial analysis, we have included temperature and a weekday factor variable as exogenous variables to capture the impact of temperature changes and weekly seasonality. We account for the exogenous variables and seasonalities by deploying a SARIMAX-model.

Implementation of SARIMAX model

Because we have different zones, which can differ between our forecasting periods in summer and winter, we start with a general training set consisting of one year of data (1. January 2018 - 1. January 2019) and analyse this period across each zone to look for similarities. Since the patterns are similar, it opens up the possibility to use the same model in every zone, both reducing computational time while also making our analysis easier to interpret and to deploy for an eventual end user.

In order to create a general forecasting model for all zones, we utilise the specified training set and focus on the *mid* zone. We consider the *mid* zone as representative due to its relative proximity to all other zones, making it an appropriate basis zone for constructing the model. The further construction of our model will be inspired by the process illustrated at chapter 9.7 in Hyndman and Athanasopoulos (2021):

Figure 5.3
The construction-process of SARIMAX model

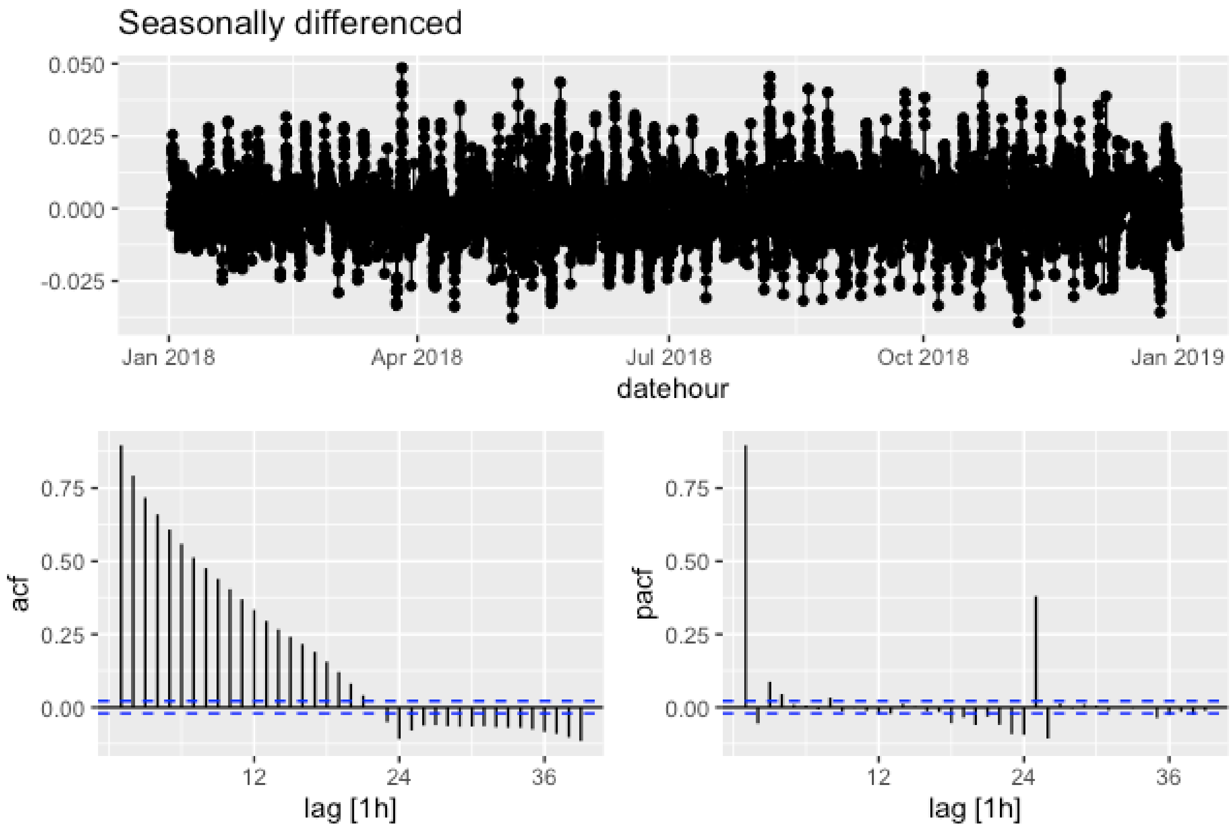


To begin the process, since there are no abnormal values or outliers in the consumption data, our first step is to stabilise the variance in the time series. We accomplish this by performing a modified Box-Cox transformation. Then, to further develop the model, we run multiple tests to make sure that the data we are analysing are stabilised.

First, we will test the data for the need of seasonal differencing by testing the seasonal strength (F_S) of the data. As imagined, the seasonal strength is way above the threshold of 0,64, meaning the need for seasonal differencing is present. Next, we test for first differencing using a KPSS-test, which returns a H_0 that is not rejected. Accordingly, the need for first differencing is *not* present.

Moving on, we determine the orders of our SARIMAX models by plotting the ACF/PACF on the differenced data:

Figure 5.4
ACF/ PACF plot of seasonally differenced consumption data



The ACF plot gradually declines, while the PACF plot displays a sharp cut-off. With the rules of non-seasonal orders in mind, this describes an autoregressive process. Since the PACF cuts off after one significant spike, it indicates an AR (1) process. Considering the seasonal orders, we see significant spikes in the PACF occurring in the first and 24th lag, indicating a seasonal AR (2) process. Ultimately, this results in a SARIMAX $(1,0,0)(2,1,0)_{24}$ process.

To make sure we use the best model available, we will include multiple manually created variations of this model, as well as an automatically detected model. The latter is based on the *Hyndman- Khandakar algorithm* (Hyndman & Khandakar, 2008) and will iterate through multiple steps:

1. Number of differencing needed - determined using repeated KPSS-tests and testing the seasonal strength (F_s)
2. Order of the model (p, q) (P, Q) are determined.
a. Four initial models are fitted.
b. The model with the smallest AICc will be set as the "current model".
c. Variations of the current model are considered, where it will vary p/q and P/Q and include/exclude c (constant) from the current model. The best model considered so far becomes the new current model.
d. Repeat step 2(c) until no lower AICc can be found.

The automatic models we have chosen are programmed such that no "default" shortcuts are implemented. This makes the model very computationally heavy as it will accordingly test every possible combination of the model's order. Additionally, it won't approximate the value of the likelihood for the models but rather use the exact computation, which often results in a well-performing model. In our case, the computation ended with a $(2, 0, 2)(2, 1, 0)_{24}$ model, which we will evaluate along with the seven manually made ones.

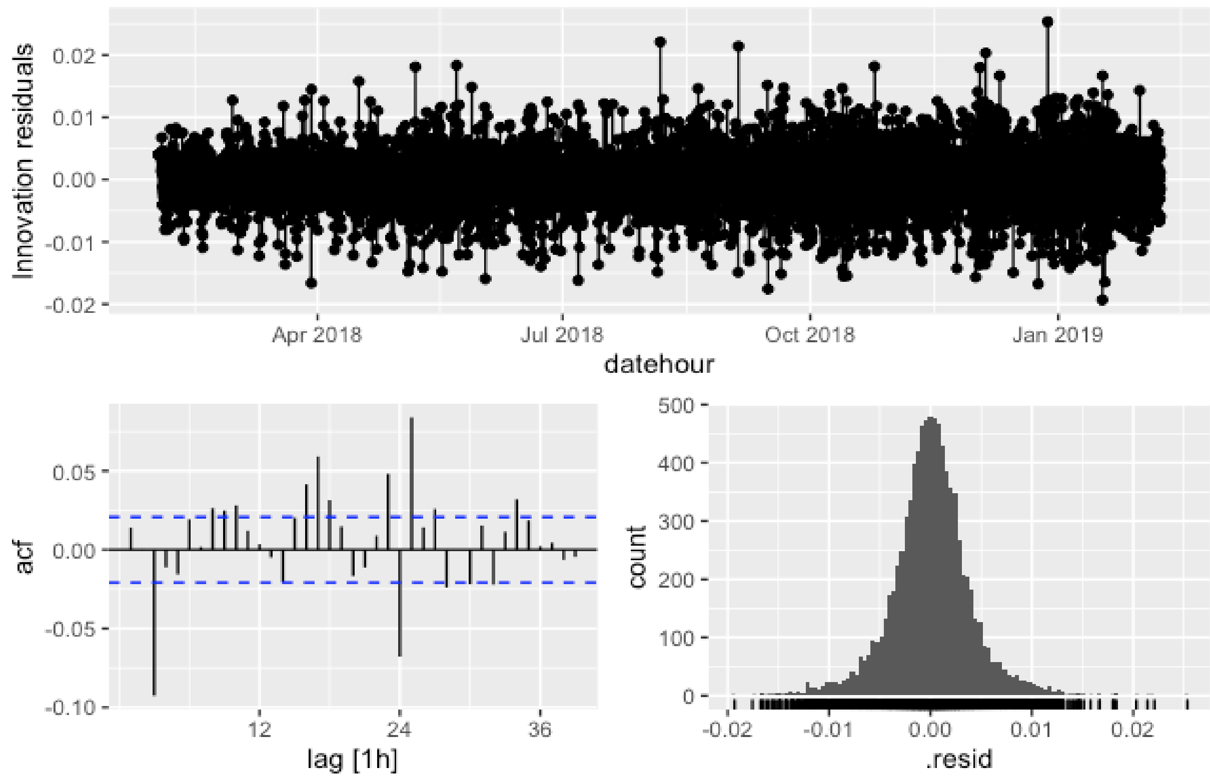
We decide upon which model to proceed with by comparing the AICc scores for the eight models created on the initial training set, continuing with the three models that return the lowest AICc scores. To ultimately decide for which model to use, we do a time series cross validation (tsCV) on the final three models. This is done because even though a model may have a good AICc score, it won't always be the model that produces the best forecasts. Nevertheless, a good AICc score indicates a good fit, which is why we will limit the cross validation to the three models with the best fit.

The three combinations of orders that have the best fit and will continue to the tsCV process are $(2, 0, 2)(2, 1, 0)_{24}$, $(1, 0, 1)(2, 1, 0)_{24}$, and $(1, 0, 2)(2, 1, 0)_{24}$. Because our final forecasting model will produce results for both summer and winter, we make two new training sets. We will run our tsCV on a summer and a winter set, each containing a year of data, producing forecasts for the week prior to our final forecasting horizon.

After running tsCV on both periods in each zone, we end up with a $\text{SARIMAX}(1,0,2)(2,1,0)_{24}$ model as the generally best performing model for forecasting demand. This is supported by conducting an accuracy test on each forecast and comparing the errors of the three models, giving each model a score from one to three respective to their error value, selecting the lowest error value as the best.

Having found the generally best performing model, we test the residuals for white noise. This is done to determine if all information in the data is captured, by testing whether any patterns remain in the residuals (Hyndman & Athanasopoulos, 2021). Testing can be done by examining the ACF plot of the residuals and/or by conducting a *portmanteau* test, such as a *ljung- box* test. As an example, we will present a plot of the residuals for the winter period in the *mid* zone (although all the fits are very similar in how the residuals behaves):

Figure 5.5
ACF/ PACF plot of the optimal model's residuals



By examining the ACF plot, multiple spikes can be observed above the blue dotted significance line, indicating that the residuals *do not* resemble white noise. Furthermore, a ljung- box test confirms this observation by rejecting H_0. However, passing all tests is not always feasible with the data provided (Hyndman & Athanasopoulos, 2021). In our case, considering the promising results obtained from the accuracy measures, we will proceed with the selected $(1,0,2)(2,1,0)_{24}$ model.

As the forecasting equation will vary depending on the order of the terms, an easy model definition (Aric LaBarr, 2021) will be provided to give a better idea of the specific model:

$$Y_t = y_t - y_{t-24}$$
$$\hat{Y}_t = c + \phi_1 Y_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \gamma_1 Y_{t-24} + \gamma_2 Y_{t-48} + \beta_1 wday_t + \beta_2 \text{min\_temp}_t + \varepsilon_t$$

Where:

$\hat{Y}_t$ is the forecasted value at time t,

c is the constant term or intercept,

ϕ_1 is the coefficient for lag- 1 of the autoregressive term,

θ_1 and θ_2 are the coefficients for lag- 1 and -2 of the moving average terms,

γ_1 and γ_2 are the coefficients for lag- 24 and -48 of the seasonal autoregressive term,

β_1 and β_2 are the coefficients for the exogenous variables $wday_t$and min_temp_t and

ε_t is the error term at time t.

5.3.2 Spot price

While we decided to solely focus on SARIMAX models to forecast consumption, we deploy a range of alternative models presented earlier in this chapter.

Implementation of SARIMAX model

When choosing the optimal SARIMAX model for the spot price, we will follow the exact same procedure as with the consumption model.

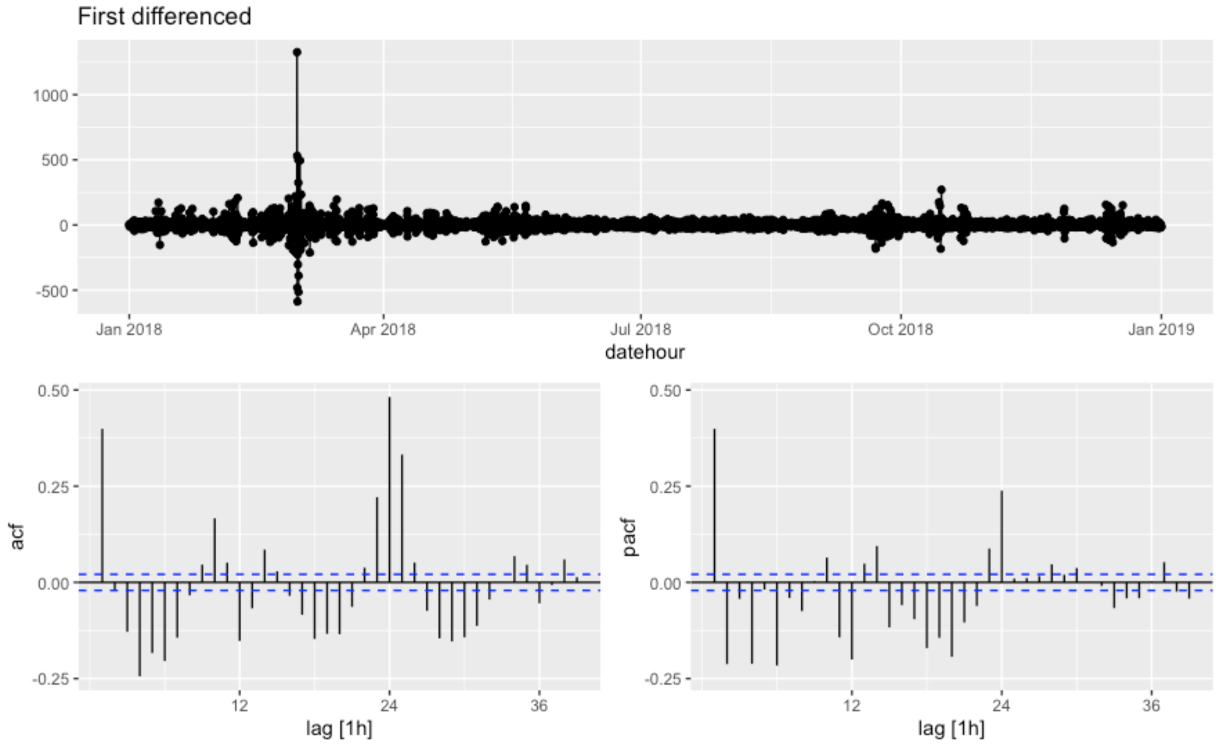
Given the nature of the electricity data, it is important to consider the presence of high values that may be perceived as outliers. Although they deviate somewhat from the majority of observations, these values are not erroneous but rather represent occasional peaks in the data. Excluding such values could lead to a loss of important information, as they are inherent to the underlying dynamics of the electricity data.

To assess the need for data transformation, we first examine whether the data requires any logarithmic or power transformations. However, after carefully evaluating the characteristics of the data and the time series, it has been concluded that such transformations are unnecessary. Consequently, we will test the need for first- and/or seasonal-differencing.

Following the earlier procedure, we start off by testing the seasonal strength of the data. This results in a F_S of 0.626, which is marginally lower than the significance level of 0.64. Based on the significance level criteria, it is determined that there is no need for seasonal differencing in the data. However, the results of the KPSS- test suggest there is a need for first differencing to achieve stationarity.

Having data that have passed both stationarity tests, the order of terms in the SARIMAX model can be detected by the ACF/PACF:

Figure 5.6
ACF/ PACF plot of differenced spot price data



Studying the plots, we detect a pattern where both the ACF and PACF gradually decline, which, according to the rules defined in the theoretical section, indicates an ARMA process with both AR and MA terms present. In deciding upon seasonal terms, we have to take into consideration the rule of thumb to not include both seasonal ACF and PACF terms. We decided upon a seasonal AR-process, as the ACF plot seems to decline at a slower rate than the PACF plot. As the first spike is significantly stronger than any other in the PACF-plot, we ultimately decide upon an AR (1) process, which leaves us with a $(1,1,1)(1,0,0)_{24}$ process.

As done earlier, we test the manually made one with 7 other variations, whereas one of the models is automatically detected with the Hyndman- Khandakar algorithm with no shortcuts. The automatic procedure ended with a model with the order of $(1,1,3)(2,0,0)_{24}$.

The three models returning the lowest AICc, thus advancing further into the tsCV step, are $(1,1,3)(2,0,0)_{24}$, $(1,1,2)(2,0,0)_{24}$, and $(1,1,2)(1,0,0)_{24}$. After computing rolling forecasts for each of the models, the model which generally scores the best across each of the zones and the two periods are the SARIMAX $(1, 1, 2)(2, 0, 0)_{24}$.

The Ljung- Box hypothesis test has also been applied to the optimal model to test for white noise in the residuals. However, the p-value is below the threshold of retaining H_0, consequently failing the test. Nevertheless, despite the presence of some patterns in the residuals that the model fails to detect, the accuracy of the temporary forecasts is promising.

Given that the spot price model and the consumption model exhibit a different order, we will provide the specific mathematical notation (Aric LaBarr, 2021) for the spot price SARIMAX model:

$$\hat{Y}_t = c + \phi_1 \Delta Y_{t-1} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \gamma_1 \Delta Y_{t-24} + \gamma_2 \Delta Y_{t-48} + \beta_1 wday_t + \beta_2 \text{min\_temp}_t + \varepsilon_t$$

Where:

$\hat{Y}_t$ is the forecasted value at time t,

c is the constant term or intercept,

ϕ_1 is the coefficient for lag- 1 of the autoregressive term,

ΔY_t represent the differenced data,

θ_1 and θ_2 are the coefficients for lag- 1 and -2 of the moving average terms,

γ_1 and γ_2 are the coefficients for lag- 24 and -48 of the seasonal autoregressive terms,

β_1 and β_2 are the coefficients for the exogenous variables $wday_t$and min_temp_t and

ε_t is the error term at time t

Implementation of ETS model

Because there are several possible combinations of the ETS model for each zone and forecasting period, we let R find the best model for us. This calculation tests an amalgam of combinations for smoothing parameters and determines whether the error, trend and seasonality are additive or multiplicative (or non-existent), choosing the model that minimises the AICc.

The chosen model for all zones and forecasting periods is $ETS(A, N, A)$, which indicates an additive error and seasonality with no trend. A common pattern observed in the smoothing parameters, α and γ (respectively belonging to the error and seasonality term), across the different fits is a relatively high α and a very low γ . This means that the model will react quickly to changes in price but won't overly rely on the seasonal patterns, as they may appear erratic.

Implementation of Neural Network model

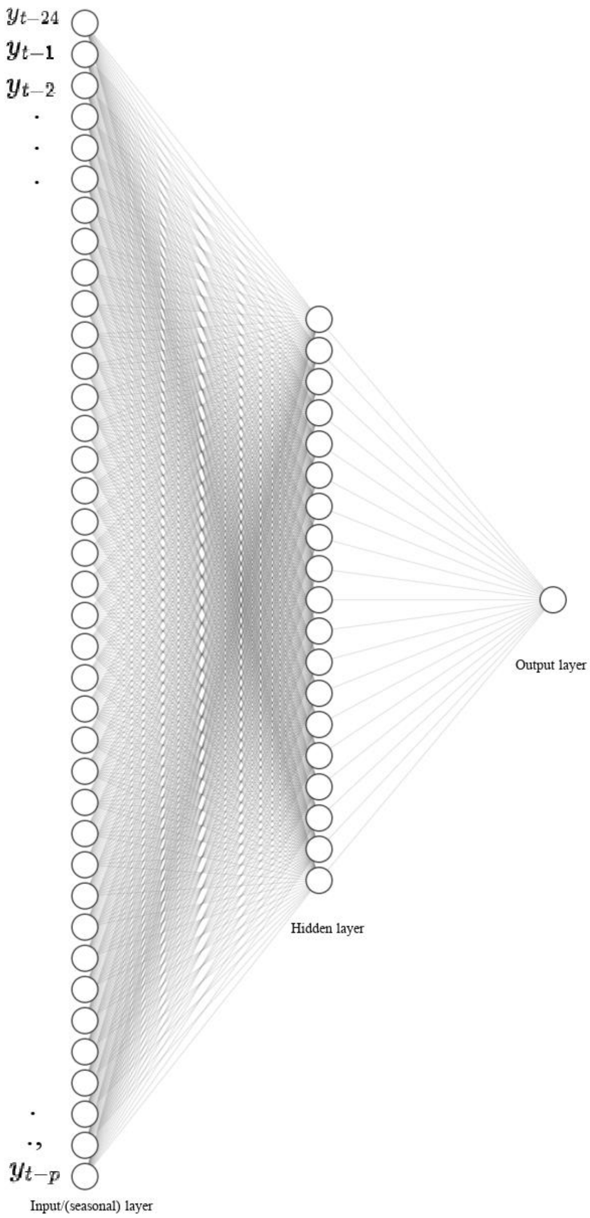
As with the ETS model, there are multiple variations which can be used in the parameters of a SAR- NN. To determine the best fit for our data, we will automatically select the parameters that minimise the overall error. Since we have seasonal data, the P is set to 1 by default, which means that the model will use one observation from the latest similar season (24 hours behind the point forecast). The optimal value of p will be the order in which the autoregressive term minimises the AICc (Hyndman & Athanasopoulos, 2021). At default, k will be set to

$$= (p + P + 1)/2.$$

When fitting the models to the different periods and zones, we observe some differences. However, the differences are minimal and consists mainly of ± 2 in the optimal value of the p-term.

The most frequently selected model is the $SAR - NN(37, 1, 19)$, meaning the network that minimises the error has 37 lagged observations of the dependent variable, 1 seasonal lagged variable and 19 nodes in the hidden layer. This network could be illustrated as follows (Lenail, n.d.):

Figure 5.7
The neural network architecture with optimal parameters



Implementation of SNAÏVE model

Being the simplest of our models, it is also the easiest to implement. The observations for each day will simply reflect the same hour the previous day.

Implementation of Combination model

Using a simple linear function, we calculate the average of the forecast values provided by each individual model. The linear function treats all models equally, without assigning any specific weights or preferences. Averaging the forecasts with equal weight ensures a model that appropriately reflects general errors and uncertainties associated with the models (Hyndman & Athanasopoulos, 2021).

6. Results

In this chapter, we will present our forecasting results for the different models presented in the methodology chapter. The specific weeks chosen for evaluation are from 10th February to 17th February 2019 and from 9th June to 16th June 2019. We will start by defining which evaluation tools are selected for different parts of the process before presenting results with visual plots and tables, followed by short briefs on the most interesting findings and patterns. To ensure a balanced presentation, the plots in the result chapter will exclusively depict zone *west* (please refer to the appendix for graphs related to the remaining zones). This decision is based on the geographical location of NHH within this zone, as we aim to avoid an excessive number of plots that might overshadow the main content of the thesis.

6.1 Error evaluation tools

Error in forecasting models refers to the difference between an observed value and the forecasted value, meaning the unpredictable part of an observation. Forecasts differ from residuals in that they are calculated on the test set and can involve multi-step forecasts, whereas residuals are calculated based on one-step. We can divide error evaluation tools into scale-dependent errors, percentage errors, and scaled errors. Percentage and scaled error evaluation tools are utilised to compare forecast accuracy across series with different units, while scale-dependent methods calculate errors on the same scale as the data (Hyndman & Athanasopoulos, 2021).

6.1.1 Mean Absolute Error

The most widely used measures of accuracy in electricity price forecasting are based on absolute errors. *Mean Absolute Error* (MAE) is commonly used in series with hourly observations by taking the mean over 24 observations for daily mean absolute error or by taking the mean over 168 (number of hours in a week) observations for weekly mean absolute error (Weron, 2014). MAE gets rid of offsetting issues deriving from a mix of negative and positive errors. The formula for MAE can be written as follows:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\text{Actual}_i - \text{Forecast}_i|$$

6.1.2 RMSE

Another commonly adopted method in electricity price forecasting is the *Root Mean Squared Error* (RMSE). As with MAE, RMSE avoids offsetting from a mix of negative and positive values by taking a square root of the deviation (Hyndman & Athanasopoulos, 2021). Mathematically, RMSE can be written as:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\text{Actual}_i - \text{Forecast}_i)^2}$$

MAE and RMSE are very similar in theory, so determining which one to use in practice can be difficult to interpret. The difference between the models is that RMSE assigns higher weights to more extreme errors. In the case of MAE, an error of 200 NOK in spot price returns twice as high error value compared to an error of 100 NOK. For RMSE however, the first case returns more than twice as high error value, meaning RMSE assigns higher weight to values further away from actual observations (Zach, 2021). Which metric to prefer ultimately depends on the nature of the project and the relative importance of avoiding more extreme errors.

6.1.3 MAPE

Another very popular evaluation metric is the *Mean Absolute Percentage Error* (MAPE), which differs from MAE only in that it returns percentages from the actual observation instead of raw value. This metric has the advantage of being unit-free, hence it is frequently used as an evaluation method in time series forecasting in general (Hyndman & Athanasopoulos, 2021). MAPE can be presented as follows:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\text{Actual}_i - \text{Forecast}_i}{\text{Actual}_i} \right| \times 100$$

As a rule of thumb, it is generally stated that a well performing model returns a MAPE not exceeding 5% (Swanson, 2015).

Comparing the different metrics, MAPE has an advantage in terms of interpretation, as percentage error gives a direct indication of the relative distance between forecasts and actual observations. However, a disadvantage occurs in cases where values drop close to zero. In these cases, smaller errors may display very large percentage errors, which can be problematic, especially if the lower values are not easier to precisely forecast. The strengths of MAPE outweigh the weaknesses in terms of our consumption forecasts since consumption values never drop anywhere near towards zero. However, for spot price, the case is the opposite, as price irregularly might drop to around zero for certain hours, meaning that not only will small absolute errors return dramatically large MAPE values, but many of the larger forecast errors will appear in these instances as well, since the rapid irregular drops are hard to predict (Weron, 2014). Since all forecasts are made on the same scale as the data, there is no necessity for considering using a scaled-independent evaluation method.

Our forecasts are divided into two parts: forecasts of electricity consumption and forecasts of spot price. The different characteristics of these parts leads us to the decision of using different metrics in each part, as we have no intention of comparing consumption forecasts with price forecasts. We will use MAPE for the consumption forecasts and MAE for the spot price forecasts. We use MAE instead of RMSE since many of the more extreme errors, in our view, occur because of abnormal situations in the data that cannot be traced well to changes in consumption and weather data, but rather to external factors not included in this project. Hence, we do not see it as beneficial to assign higher weights to more extreme errors.

6.2 Consumption forecast results

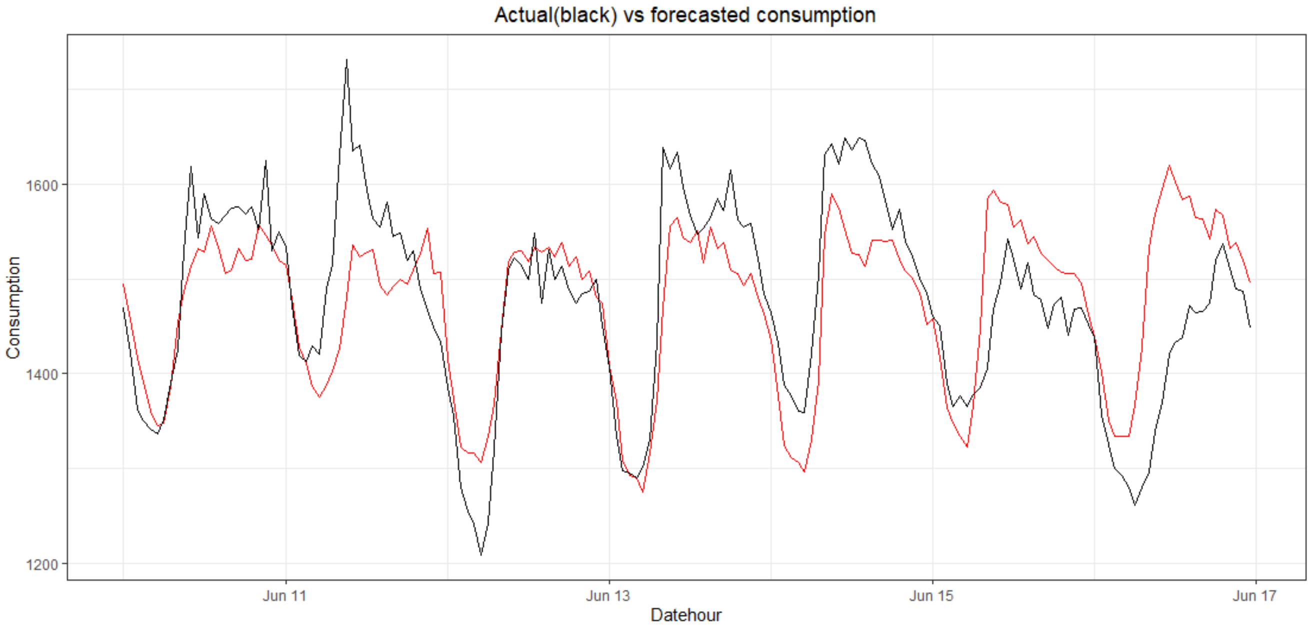
Presented below are the results of our demand forecasts for summer and winter obtained from our chosen SARIMAX model. The forecasted values will be further applied as exogenous predictors in our spot price forecasts.

6.2.1 Summer results

Table 6.1
Electricity demand forecast - MAPE - summer (09.06.- 16.06.2019)

As shown in *table 6.1*, the overall performance is adequate, with an overall MAPE average of 4.852%. The majority of forecasts satisfy the rule of thumb of a MAPE below 5%. Despite that, the model struggles to capture all patterns, specifically on Tuesdays and Sundays. Additionally, the model performs relatively poorly on Tuesdays and Saturdays in the *east* zone.

Figure 6.1
Actual vs. forecasted consumption values for zone west - summer (09.06.-16.06.2019)



The graph shows a pattern where the forecasts generally underestimate during the weekdays and overestimate during the weekends, though not to any dramatic degree. The biggest gap appearing on Tuesday seems to result more from abnormally high consumption rather than an undervalued forecast.

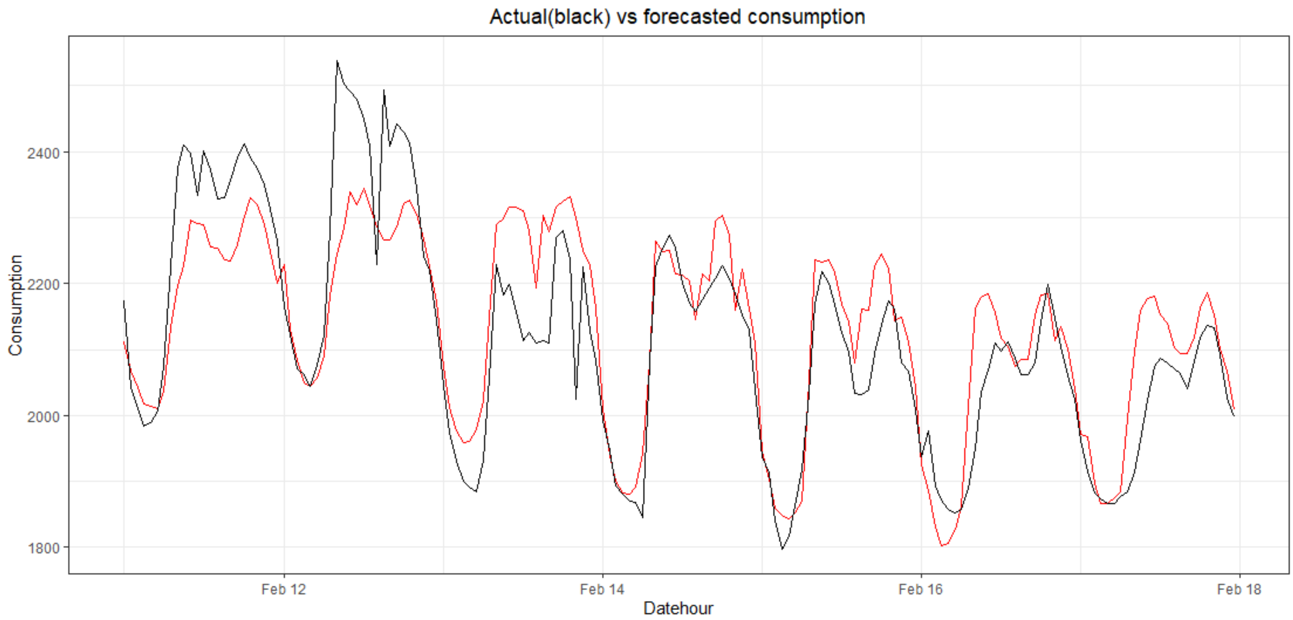
6.2.2 Winter results

Table 6.2
Electricity demand forecast - MAPE - winter (10.02.- 17.02.2019)

The winter forecasts improve upon the summer forecasts, returning an average MAPE of 3.65%. Additionally, there doesn't seem to be any general patterns for this week that the model fails to capture. Although a recurring issue seems to be present in the *east* zone, as it doesn't quite accomplish the same accuracy as the other zones. However, the general results are up to the mark.

Figure 6.2

Actual vs. forecasted consumption values for zone west - winter (10.02.- 17.02.2019)



As with the summer forecast, the model seems to underestimate early in the weekdays. However, the forecast improves further into the week and manages to capture the slight drop during the weekend, only slightly overestimating in certain periods.

6.3 Spot price forecast results

This part presents the price forecast results. As with consumption, we will present one table for each forecasted time period. Two plots are included for each period: one comparing the patterns of performance of all models, while the other displays the best performing model.

6.3.1 Benchmark for MAE

As we use MAE to evaluate our spot price models, it is advantageous to declare a benchmark value for a good score, as the numbers being on the same scale as the data do not provide the same information relative to the actual observed values as percentage errors do. Actual values during forecasted periods generally range from around 200 - 350 NOK during the summer, while generally ranging between 400 - 500 NOK during the winter. Since spot price, in comparison to consumption, is more sensitive to impact by external factors and therefore harder to predict by nature, and the lower values means small residuals return higher relative errors, we will adjust the rule of thumb benchmark for good results to around 10 percent for the lower values. This means that a MAE of 20-30 will be viewed as adequately precise in this part.

6.3.2 Summer results

Table 6.3
Spot price forecast - MAE - summer (09.06.- 16.06.2019)

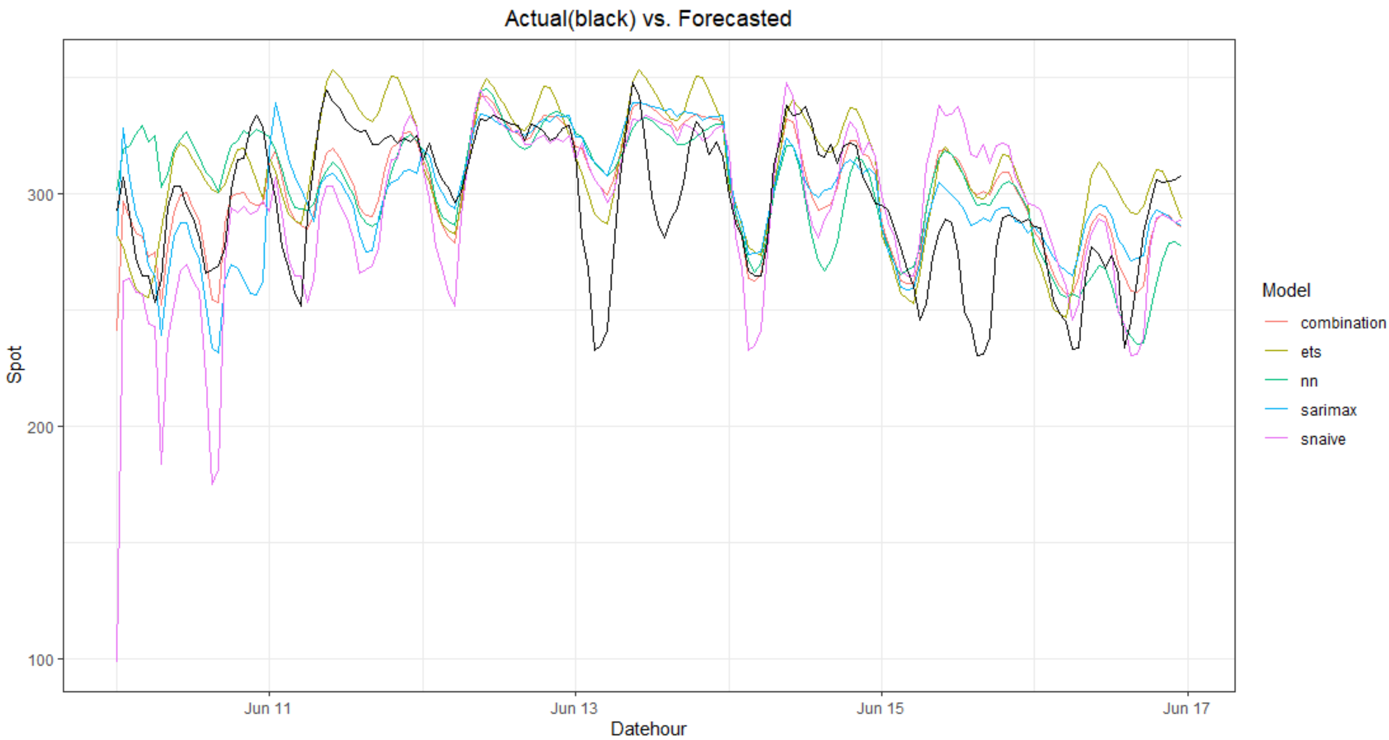
Table 6.3 displays varying results across days and zones. In general, the model that seems to reliably produce the best forecast is the combination model. The combination model produces the best forecast for *south*, *west*, and *east*, whilst being the second best in *mid* and *north*. Interestingly, SNAÏVE performs the best in these zones but struggles in the rest.

A pattern worth noting is the general underperformance in zone *north* and *mid*, especially on Monday. A closer look at the data uncovered an abnormal situation in these zones. Prices fluctuate from drops to almost zero at night-time to around 300 with steep hourly changes from

Monday to Wednesday, resulting in massively undervalued forecasts unable to adapt to the steep increases. We do not believe these high errors are an indication of a bad model performance as these situations are hard to predict, and we would rather judge the model based on performance on the other days for these zones. Nevertheless, the models in both zone *mid* and *north* still underperform quite a bit, with the exception being Friday.

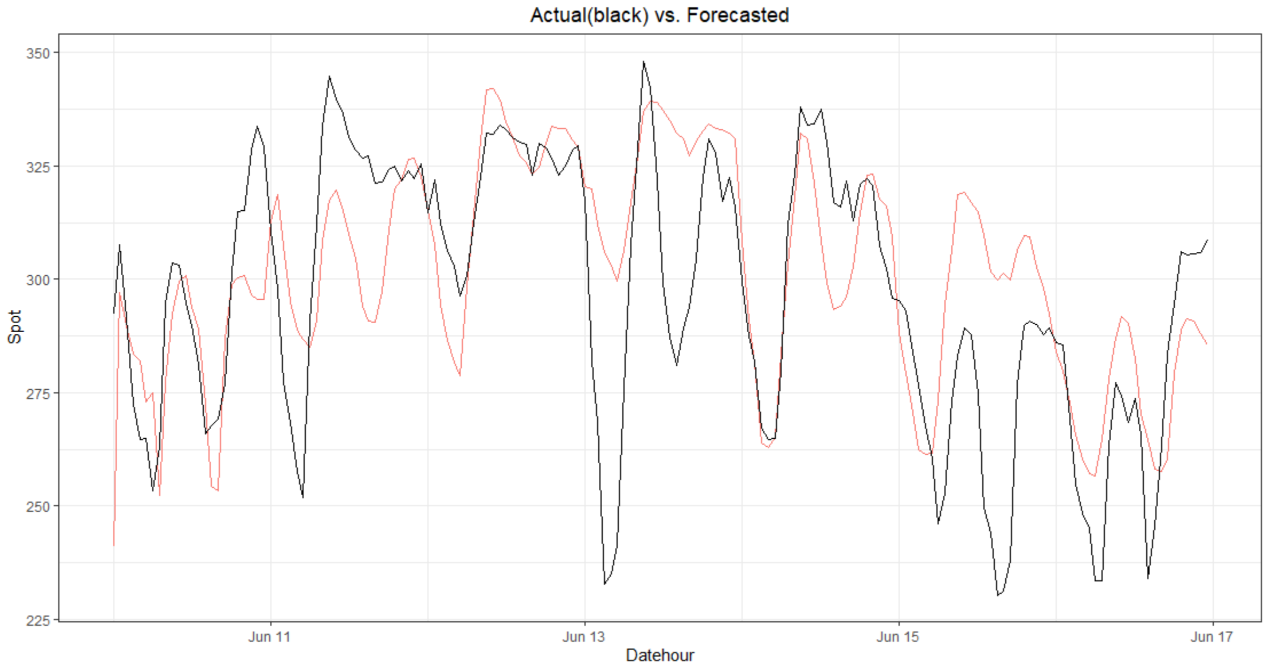
Figure 6.3

Actual vs. forecasted spot price for zone west - summer (09.06.- 16.06.2019)



As *figure 6.3* shows, the amount of negative and positive errors is approximately evenly distributed for most days. Certain models tend to mainly either overvalue or undervalue; the SNAÏVE model tends to undervalue forecasts for large parts of the week, while ETS, in particular, tends to overvalue. Otherwise, most models tend to shift between undervaluing and overvaluing, which explains why the aggregating combination model outperforms the others. Most underpredictions are made early in the week, while most overpredictions happen during the weekend, similar to the demand forecasts.

Figure 6.4
Actual vs. forecasted spot price with the best model for zone west - summer (09.06.-16.06.2019)



This plot shows the best performing model, the combination model. After starting out massively underpredicting on Monday, it manages to capture the rest of the pattern on Monday remarkably well. Otherwise, the performance is quite consistent, with some exceptions of high overpredictions in hours with especially low price observations. Note that the plot is scaled for the lines to utilise the whole area, resulting in the Y-axis ranging from 225 to 350, a relatively small interval. This may make the predictions appear less precise than they actually are. This good performance of the combination model confirms the evenly distributed pattern observed in the other models.

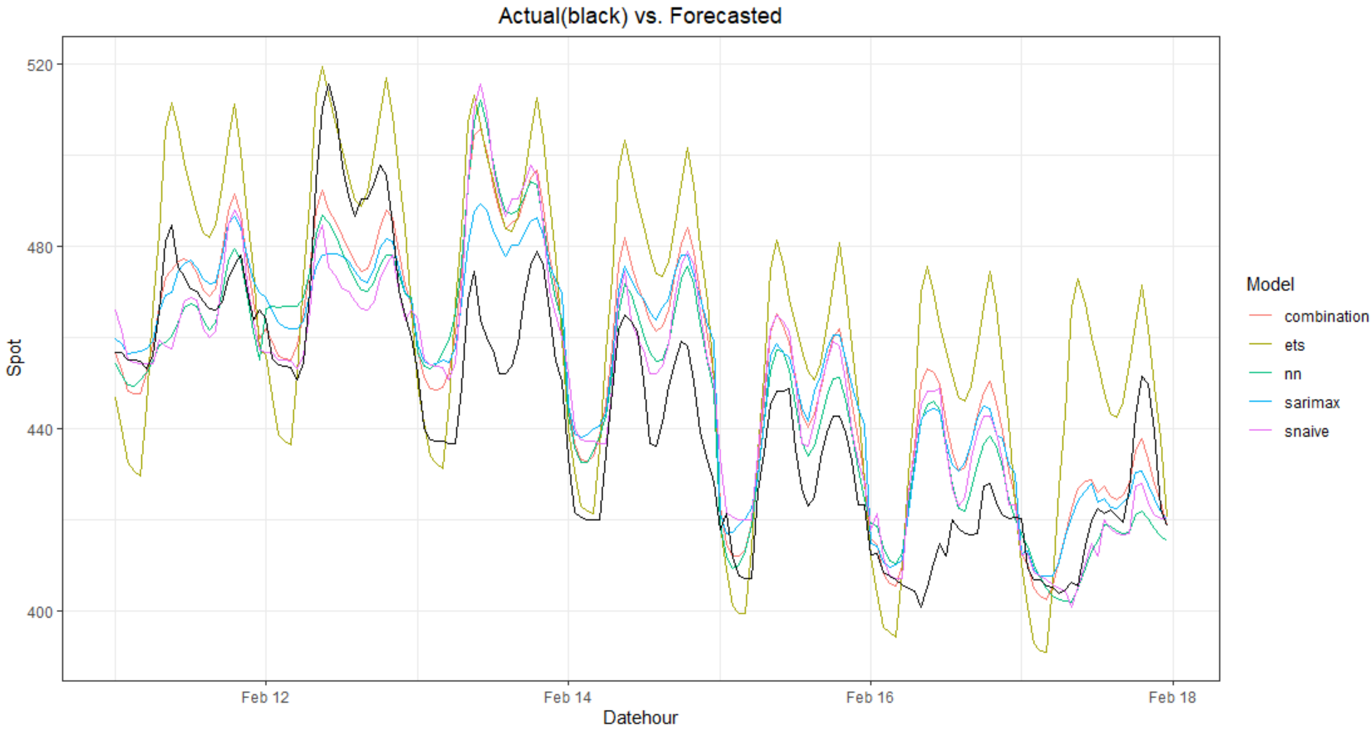
6.3.3 Winter results

Table 6.4
Spot price forecast - MAE - winter (10.02.- 17.02.2019)

In general, the results from winter forecasts are more precise, and more importantly, display less variation. The only instance where the weekly mean error surpasses 30 is for the ETS model, which, in general, is the only model somewhat lacking in performance. Interestingly, there is no model that really outshines the others. Contradicting the summer forecast, the combination model does not prove to be the best overall model in any zone for this period.

The best forecasts generally happen early in the week, with Monday and Tuesday boasting the best average results. All days seem to generally produce precise forecasts, with a slight exception on Saturday, where all models are struggling in the *mid* and *north* zone. The models seem to be struggling in approximately the same degree, which we can track back to a price dip appearing early Saturday that the models struggle to capture.

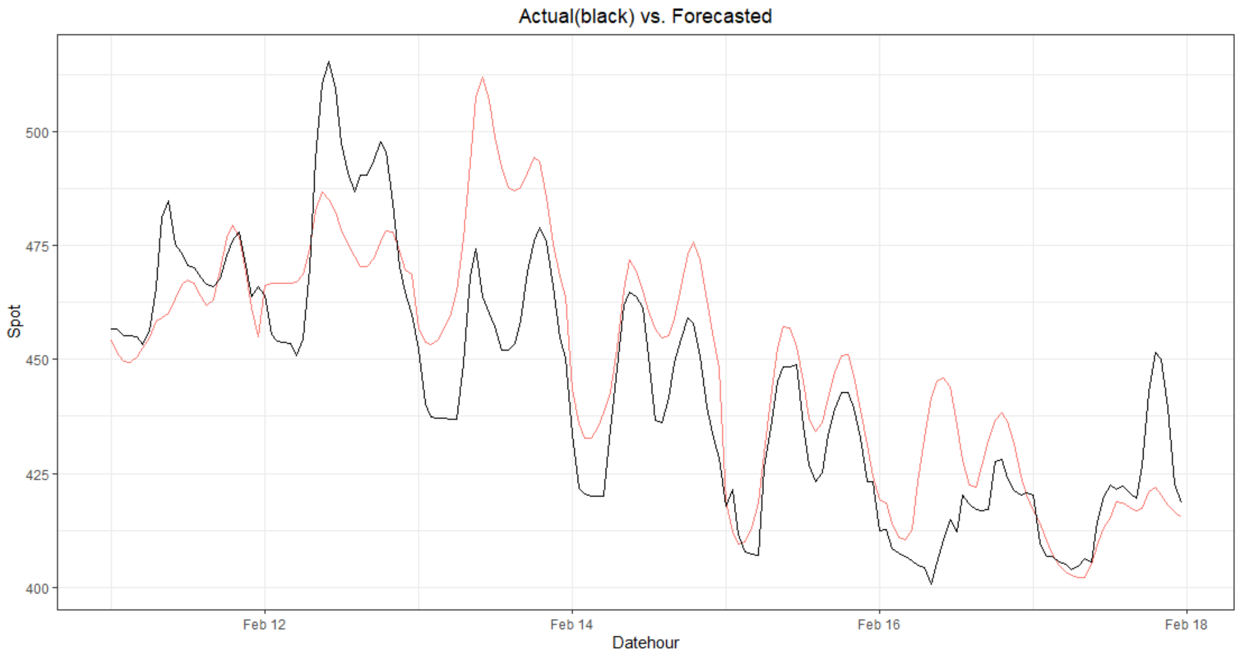
Figure 6.5
Actual vs. forecasted spot price for zone west - winter (10.02.- 17.02.2019)



The general performance of the model is displayed here, showing slight overpredictions for most of the week. The underperformance of ETS seems to be mostly consistent throughout, displaying overly sensitive reactions to most dips and peaks. This sensitivity damages the performance of the combination model, as it only serves to strengthen the error pattern shown in the other models.

Figure 6.6

Actual vs. Forecasted spot price with the best model for zone west - winter (10.02.-17.02.2019)



The best performing model for zone *west* during the winter is the neural network model. This model captures the beginning and ending of the week remarkably well, as well as the general slowly descending pattern towards the end of the week. The weakest point of the model is the overprediction on Wednesday, where it fails to detect the price dip from Tuesday. Generally, it can be seen in the plot that the predictions for each day strongly resemble the observations from the preceding day, indicating that the neural network model weighs the seasonal nature of the data considerably. This approach works better in this period than for the summer forecast.

7. Discussion

In this chapter, we interpret the results obtained from our forecasting models. Our goal is to draw meaningful conclusions based on our findings and discuss the implications that our findings might hold for the participants on the Nord Pool day-ahead market. Additionally, we will discuss some of the limitations of the study that future research could address to gain additional knowledge on the topic.

7.1 Consumption forecasting results

Generally, the patterns show similarities between summer and winter. Early in the week, the model tends to underestimate, while during the weekend, it tends to overestimate. This might imply that the weekly seasonal effect captured by the weekday variable in the model is not sufficient to efficiently capture the weekly pattern of increasing consumption on Mondays after the weekend and decreasing consumption after Fridays for weekends. The weekday variable appears to have minimal impact on the summer forecasts. However, for the winter forecasts, forecasted consumption decreases during the weekend, suggesting that the variable might have a significant impact. Consequently, it seems that the temperature variable may hold too much weight in the summer period, where changes in temperature have a lesser effect on consumption compared to colder periods.

Notably, the *eastern* zone deviated significantly in performance compared to the other zones. The *eastern* zone is the most populated zone in Norway, accounting for 40 % of the total population. It also contains the by far most densely populated region in the country in the urban areas in and around Oslo. As a result, the effect on consumption from weather in this zone might differ and moreover might be more sensitive to external effects not explained by the model. It is likely that the high amount of urban population, combined with the numerous businesses and industries in and around Oslo, leads to different sensitivities to the exogenous variables.

In terms of residential areas, a larger number of people live in apartment blocks with different heating requirements compared to residents of detached houses, which are more prevalent in other parts of the *eastern* zone. Additionally, the numerous businesses in central Oslo might affect the weight on the weekday variable for parts of the zone in terms of opening and closing hours. Ultimately, the differential impact of exogenous factors across urban and rural areas within the zone may have a negative effect on the overall performance of a generalised model.

Another noteworthy aspect is the enhanced forecast performance observed during winter compared to summer, which potentially supports the initial analysis findings suggesting a stronger correlation between temperature and consumption for lower temperatures. However, it is also possible that the improvements are simply a result of more stable data during winter, since the price forecasting models without the inclusion of weather data also show the same improvement for winter compared to summer. It is likely that the difference in performance can be explained by a combination of these factors.

In conclusion, the consumption forecasts perform well, which is not surprising considering the robust correlations with temperature and consistent seasonal pattern related to industries and businesses working hours. There are also far fewer external effects connected to consumption compared to price.

7.2 Spot price forecasting results

7.2.1 Benchmark models

As the results show, there is no clear indication of any model outperforming the others. However, we can interpret from our results that the ETS model underperforms on average, mostly due to being overly sensitive, resulting in overestimation of the peaks and lows of the forecasting period.

The SNAÏVE model displays some of the biggest variations in performance, notably outperforming other models on abnormal data such as the summer period in the *north* and *mid* zones, whilst relatively struggling to capture patterns during the same period for the other zones. The simplicity of the seasonally naïve method proves to be surprisingly effective for data with strong seasonal patterns, as long as the levels of highs and lows are somewhat stable throughout the analysis period. The *north* and *mid* zones exhibit very high volatility during the summer week, which the other models were unable to capture. However, the SNAÏVE model struggles on Monday but quickly adjusts and returns generally good results for the rest of the week, as the volatile pattern remains consistent. In general, the SNAÏVE model seems to be highly effective for stable seasonal patterns.

The neural network model displays some promising results, returning the most accurate forecasts for zone *west* during the winter. However, its performance is mediocre in other periods, potentially indicating an overreliance on the daily seasonal patterns, as shown in *figure 6.6*.

In cases where the models display differing patterns with both positive and negative errors, the combination models deliver great performance. This model effectively balances the strengths and weaknesses of the individual models, resulting in high stability. However, in cases where the patterns are more similar and one model underperforms, the combination model weakens, as for example in the winter forecast, where the ETS model follows a similar pattern as the others but with higher errors. Nonetheless, with the appropriate selection of models, we believe that a combination model can be the most appropriate and, as mentioned earlier, difficult to surpass.

7.2.2 SARIMAX model

The SARIMAX model produces generally adequately good forecasts. Out of the 10 weekly average results, only two periods/zones return unsatisfactory high MAE values. Furthermore, these weeks exhibit abnormal price fluctuations that no other model effectively captures either. However, the inclusion of exogenous variables does not seem to significantly enhance the model's performance, as the SARIMAX model does not outperform the other benchmark models. This suggests that the impact of the exogenous factors may not be as important as initially expected, compared to other external effects not considered in this project. Alternatively, the relationships between the variables may exhibit nonlinear dynamics instead of the linear relationship assumed by the SARIMAX model. Nonetheless, the SARIMAX model consistently produces stable and adequate results, indicating that it can rely on the incorporated exogenous effects in situations where the reliability of the autoregressive effect is weak.

In conclusion, the overall performance suggests that the SARIMAX model is capable of handling most situations reasonably well, and in cases of performance dips, the model tends to quickly improve, as shown by the lack of consecutive instances of bad MAE values.

7.3 Critique

7.3.1 Limitations of model variables

Throughout this thesis we have assumed spot prices to depend on lagged prices, demand and indirectly on temperature and weekday. In reality however, there are numerous other factors that also have an effect on the price, which we have chosen not to consider in this project. Below, we will present some of the most important factors that were not taken into account in our study.

Reservoir levels

Since almost all Norwegian electricity is produced from hydropower, the availability of water in reservoirs does have an impact on price setting. However, we argue that this relationship works in a more complex manner than a simple "less availability equals higher price" and vice versa. In the short term, reservoir levels may not have a strong influence on price, but they are important factors to be considered in long-term forecasting models.

Import and export

Norway partakes in the global power market, with constantly evolving capacity for import and export of electricity. Power flows in and out of the country, contributing to a reduced risk of power shortages. However, this also means that Norwegian prices are heavily influenced by prices abroad, both in Europe and the rest of the world. The interconnectedness of the global power market plays a significant role in shaping the electricity prices in Norway.

Coal prices

Unlike Norway's predominantly renewable production, several countries in Europe still rely on non-renewable sources such as coal and gas power plants. These countries need to purchase the raw materials required for power production. Therefore, high prices of coal and gas can have an impact on the amount of electricity produced in these countries, which, in turn, can influence electricity prices in Norway.

Exchange rates

The exchange rates of the Euro and the US Dollar are two additional factors that impact Norwegian electricity prices. In the day-ahead market, hourly prices are determined in Euros. An increase in the Euro exchange rate leads to a weakened Norwegian krone, resulting in higher recalculated prices in Norway. Furthermore, the Dollar exchange rate is relevant for acquiring coal in European coal power plants. Therefore, fluctuations in the Dollar rate can influence the amount of coal purchased and produced, which in turn can affect import prices for Norway (NTE, n.d.).

7.3.2 Limitations of methodology

Model selection

In this thesis, only one model incorporates the exogenous effect of demand and weather. The flexible nature of the SARIMAX model, in addition to the aim of comparing the autoregressive capability of this model with other statistical approaches that effectively model based on past observations, prove the basis for this decision. Nevertheless, it is clear that the inclusion of additional models that take exogenous factors into account could enhance the analysis and provide further benefits.

Aggregation of weather variable

The aggregation of weather variables in this thesis is based on an intuitive approach. Rough estimates of population distributions for each zone are computed, and meteorological data from stations located in different areas of the zones are retrieved to account for climatic differences. The zone division does not follow the same borders as the counties in Norway, which led to the preference for this estimation approach. The selection of meteorological stations was limited due to the lack of data for certain weather metrics or large amounts of missing data in most stations. Therefore, the selection was made in a best-case scenario manner, considering multiple imperfect estimations, particularly in the zones *south*, *west*, and *east*. However, the decision to use only temperature as a weather variable helps mitigate these imperfections, as temperature is usually a more stable metric within a specific area compared to rainfall and wind.

Actual / Forecasted weather

In this project, we utilise actual observed weather data obtained from Norsk Klimaservicesenter's database to generate the demand forecasts. However, the main objective of this forecasting project is to develop a model capable of predicting future spot prices by utilising weather forecasts rather than relying on actual observations. It is important to note that using weather forecasts instead of actual data might have a negative impact on the model's performance. However, it is worth mentioning that short-term temperature forecasts tend to be accurate, meaning the potential decrease in performance should be of small effect.

7.4 Implications of study

The results of our study indicate that the SARIMAX model, along with the other benchmark models, can be employed as valuable tools in electricity price forecasting. The characteristics of the price data make it a suitable and viable object for forecasting purposes.

The findings in our study indicate temperature as a useful predictor of electricity demand, and that there in addition are correlations between temperature and price as well. However, it is important to note that temperature is just one of many variables that influence electricity prices, meaning the impact on the model's performance is not easily interpretable in isolation. We also observed that variables such as rainfall and wind speed do not exhibit significant relationships with demand or prices, likely due to the ability to store hydropower and the relatively low percentage of wind power used in energy production. Concludingly, we find that weather effects on electricity price are present, but in a more nonlinear and complex manner than our forecasting models are generally able to capture, suggesting that further research and exploration is needed to better understand and incorporate the dynamics of the weather effects.

In terms of practical implications, our findings indicate that incorporating autoregression and seasonal effects into the price forecasting methodology proves to be an effective approach. The general results suggest that a forecasting model, combined with market participants' knowledge of potential external influences, should prove an effective strategy, aiding said participants in making informed decisions on the day-ahead market regarding bidding and pricing values.

8. Conclusion

This master thesis studies the concept of electricity price forecasting on the Nord Pool day-ahead market in Norway. The electricity market in Norway is today a part of an internationally integrated network, which, in turn, has led to increasingly complex price dynamics influenced by factors such as weather, policies, demand, and fuel prices. Weather patterns are becoming increasingly important as well due to the global focus on utilising renewable energy production. By conducting a study on electricity price forecasting and looking into the potential impact of weather variables, we aim to aid the future market participants in deciding upon bidding values and volumes on the Nord Pool day-ahead market.

Our analysis is done on a combination of lagged price values and market demand, represented by consumption in our analysis, both obtained from Nord Pool. Weather variables were gathered from Norsk Klimaservicesenter and included temperature, precipitation, and wind speed values. However, after some initial analysis, we decided to solely focus on temperature as this is the only weather variable that seems to significantly impact demand and price. We have incorporated the temperature data into our models as a predictor on demand, which in turn is used as a predictor on price due to the correlated nature of these variables. The temperature has a direct impact on consumption both in households and larger industries, while consumption has a direct impact on price as the day ahead market always aims to establish equilibrium between market demand and supply, meaning more resources are required to account for higher consumption, affecting the price.

The model used to incorporate temperature data is the SARIMAX model, a model which is able to flexibly capture the autoregressive nature of price, combined with impact from exogenous factors. Exogenous factors included are the aforementioned demand, as well as a weekday factor variable which is included to effectively capture the weekly seasonal variations in demand and price patterns. To evaluate the SARIMAX model performance, we have additionally included a range of benchmark models, ranging from a simple seasonal naïve model to more complex models such as a neural network model. The benchmark models produce forecasts solely based on past observations of price, making the direct influence of weather and other exogenous factors in the SARIMAX model easier to interpret.

Our models return generally promising results, suggesting that the autoregressive nature of electricity price on its own can be used to produce accurate forecasts, explained by weekly MAE values for all models rarely exceeding 30, which in most cases means less than 10% MAPE. We consider this as generally acceptable accuracy, as the nature of the interconnected market means a wide range of external factors not easily incorporated in forecasting models may affect price fluctuations by uneven amounts.

Our SARIMAX model performs adequately on average in comparison to the benchmark models, but the results suggest a relatively high stability in performance. The inclusion of the weekday factor variable, temperature, and demand helps explain the stability observed, as these external factors serve to mitigate the risk of substantial undervaluation or overvaluation of forecasts during periods of intense fluctuations.

Generally speaking, the findings of this study suggest that the SARIMAX model is a helpful tool for forecasting electricity prices, with weather and demand positively impacting the overall model performance. However, the relationship seems to vary depending on the season and the impact of external effects not included in this study, making the direct influence of weather complex and nonlinear. We find that the general results, combined with market participants' knowledge of potential external influences, may aid these participants in making accurate forecasts in real-life situations.

References

Aric LaBarr. (2021). *What are Seasonal ARIMA Models*. YouTube. https://www.youtube.com/watch?v=IK67f3IItfw

Bickel, P. J., & Doksum, K. A. (1981). An Analysis of Transformations Revisited. *Journal of the American Statistical Association*, *76*(374), 296–311. https://www.tandfonline.com/doi/abs/10.1080/01621459.1981.10477649

Box, G. E. P., & Cox, D. J. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society Series B-methodological*, *26*(2), 211–243. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Clemen, R. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, *5*(4), 559–583

Coester, A., Hofkes, M. W., & Papyrakis, E. (2018). An optimal mix of conventional power systems in the presence of renewable energy: A new design for the German electricity market. *Energy Policy*, *116*, 312–322. https://doi.org/10.1016/j.enpol.2018.02.020

Drigo, I. (n.d). *How to get the most accurate weather forecast. Is it really possible? –* Windy.app. https://windy.app/blog/most-accurate-weather-forecast.html

Energifakta Norge (2021). *Electricity production.* Energifakta Norge. https://energifaktanorge.no/en/norsk-energiforsyning/kraftproduksjon/?fbclid=IwAR3e_pAfS_sRNo5bX3ldKpzGbv5f5Wj3FIKowcAUea4ZNBmPKOEKobM_pBQ

Energifakta Norge (2022). *The power market.* Energifakta Norge. https://energifaktanorge.no/en/norsk-energiforsyning/kraftmarkedet/

Guerrero, V. H. (1993). Time-series analysis supported by power transformations. *Journal of Forecasting*, *12*(1), 37–48. https://doi.org/10.1002/for.3980120104

Holt, C. C. (1957). *Forecasting seasonals and trends by exponentially weighted averages* (ONR Memorandum No. 52). Carnegie Institute of Technology, Pittsburgh USA. Reprinted in the *International Journal of Forecasting*, 2004. https://www.sciencedirect.com/science/article/pii/S0169207003001134?via%3Dihub

Hyndman, R.J., & Athanasopoulos, G. (2021) *Forecasting: principles and practice*, 3rd edition, OTexts: Melbourne, Australia. https://otexts.com/fpp3/

Hyndman, R. J., & Khandakar, Y. (2008). Automatic Time Series Forecasting: The forecast Package for *R*. *Journal of Statistical Software*, *27*(3). https://doi.org/10.18637/jss.v027.i03

Jain, S., & Singh, V. (2003). Statistical Techniques for Data Analysis. *Developments in Water Science*, 207–276. https://doi.org/10.1016/s0167-5648(03)80058-8

Kang, J., & Reiner, D. (2022). What is the effect of weather on household electricity consumption? Empirical evidence from Ireland. *Energy Economics*, *111*, 106023. https://doi.org/10.1016/j.eneco.2022.106023

Kristiansen, T. (2012). Forecasting Nord Pool day-ahead prices with an autoregressive model. *Energy Policy*, *49*, 328–332. https://doi.org/10.1016/j.enpol.2012.06.028

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, *54*(1–3), 159–178. https://doi.org/10.1016/0304-4076(92)90104-y

Lenail, A. (n.d.). Neural Network SVG Creator. https://alexlenail.me/NN-SVG/index.html

Nau, R. (2020). *Statistical forecasting: notes on regression and time series analysis.* https://people.duke.edu/~rnau/411home.htm

NEMO (n.d.). *Single Day-ahead Coupling (SDAC)*. https://www.nemo-committee.eu/sdac

Nord Pool (n.d.-a). *Day-ahead market.* https://www.nordpoolgroup.com/en/the-power-market/Day-ahead-market

Nord Pool (n.d.-b). *Price formation.* https://www.nordpoolgroup.com/en/the-power-market/Day-ahead-market/Price-formation/

NTE. (n.d.). *Hva påvirker strømprisen*? https://nte.no/blogg/hva-pavirker-stromprisen/

NVE. (n.d.). *Magasinstatistikk.* https://www.nve.no/energi/analyser-og-statistikk/magasinstatistikk/

Seklima. (n.d.).*Observasjoner og værstatistikk.* https://seklima.met.no/

Statista (2023). *Distribution of electricity production in Norway 2021, by source*. https://www.statista.com/statistics/1025497/distribution-of-electricity-production-in-norway-by-source/

Statnett (n.d.). *Om strømpriser*. https://www.statnett.no/om-statnett/bli-bedre-kjent-med-statnett/om-strompriser/

Statnett (2023). *Fakta om prisområder*. https://www.statnett.no/om-statnett/bli-bedre-kjent-med-statnett/om-strompriser/fakta-om-prisomrader/

Swanson, D. A. (2015). *On the Relationship among Values of the same Summary Measure of Error when used across Multiple Characteristics at the same point in time: An Examination of MALPE and MAPE*. https://escholarship.org/uc/item/1f71t3x9

Tanaka, K., Matsumoto, K., Keeley, A. R., & Managi, S. (2022). The impact of weather changes on the supply and demand of electric power and wholesale prices of electricity in Germany. *Sustainability Science*, *17*(5), 1813–1825. https://doi.org/10.1007/s11625-022-01219-7

Tanrisever, F., Shahmanzari, M. & Büke, B. (2021). *European Electricity Day-Ahead Markets: A Review of Models and Solution Methods.* https://www.maths.ed.ac.uk/~bbuke/assets/European_DAM_Review_05_January_2021.pdf

The Pennsylvania State University. (n.d.). 10.2 - *Autocorrelation and Time Series Methods.* https://online.stat.psu.edu/stat462/node/188/

Virenrehal. (2022). Interpreting ACF and PACF plots - SPUR ECONOMICS. *SPUR ECONOMICS - Learn and Excel*. https://spureconomics.com/interpreting-acf-and-pacf-plots/

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), 1030–1081. https://doi.org/10.1016/j.ijforecast.2014.08.008

Wickham, H., Chang, W., Henry, L., Lin Pedersen, T., Takahashi, K., Wilke, C., Woo, K., Yutani, H. & Dunnington, D. (n.d). *Smoothed conditional means — geom_smooth*. https://ggplot2.tidyverse.org/reference/geom_smooth.html

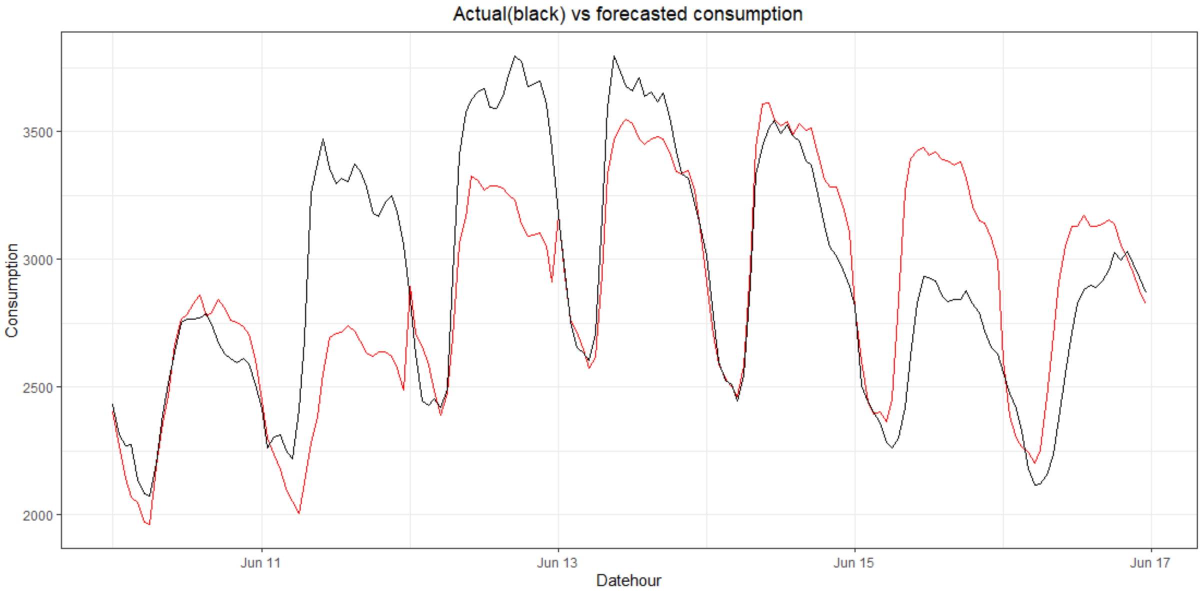
Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, *6*(3), 324–342. https://pubsonline.informs.org/doi/10.1287/mnsc.6.3.324

Zach. (2021). MAE vs. RMSE: Which Metric Should You Use? *Statology* https://www.statology.org/mae-vs-rmse/

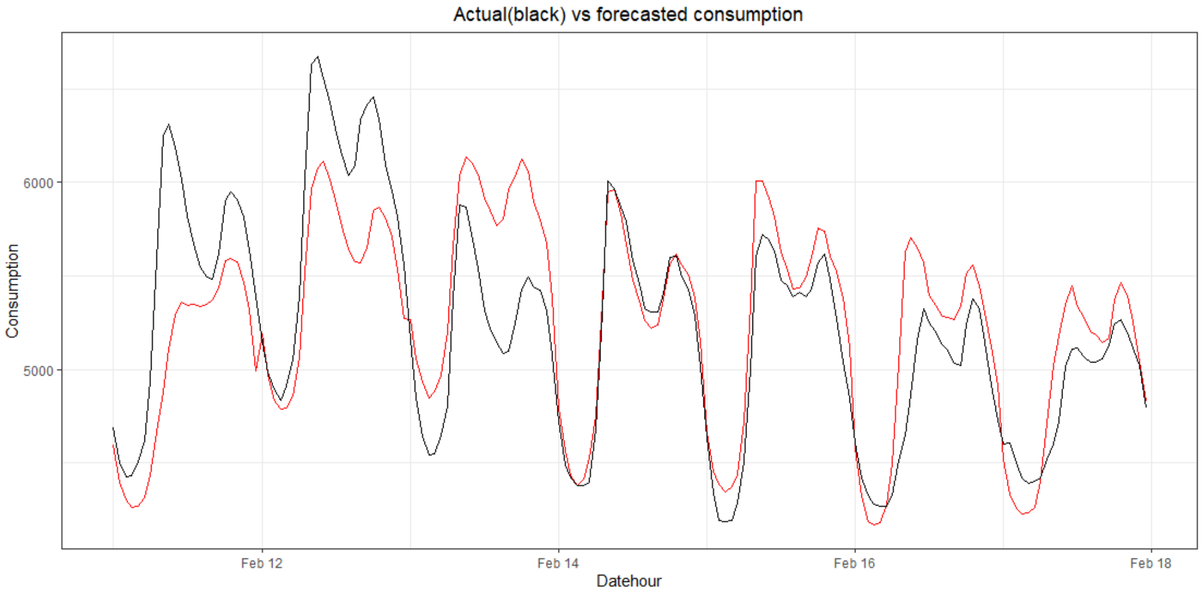
Appendix

Figures for zone *east*:

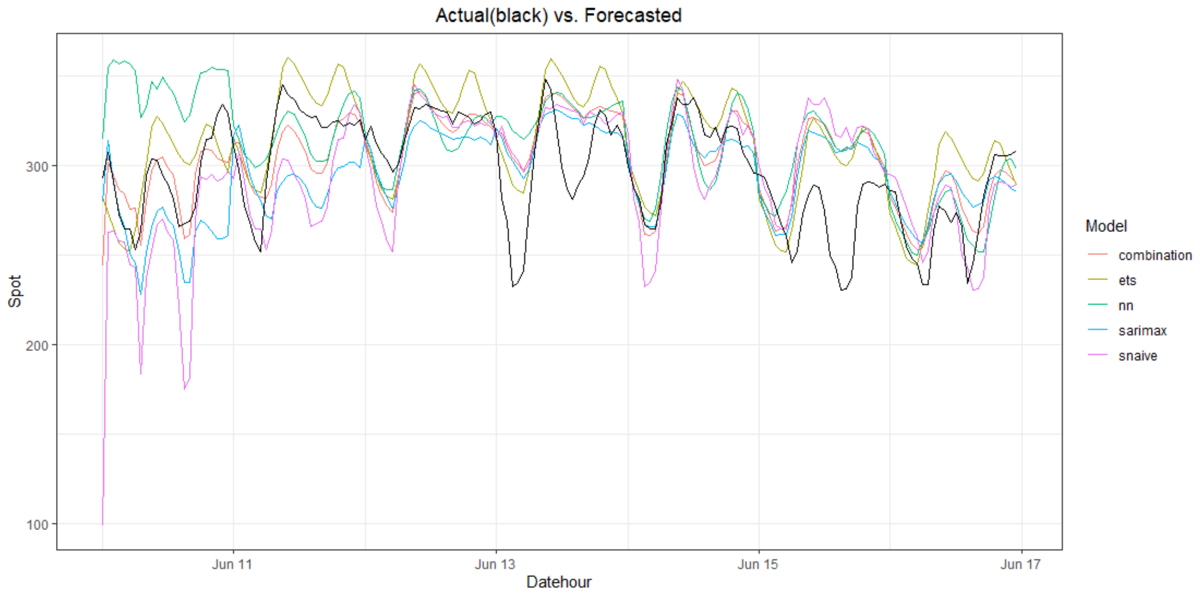
Summer consumption:



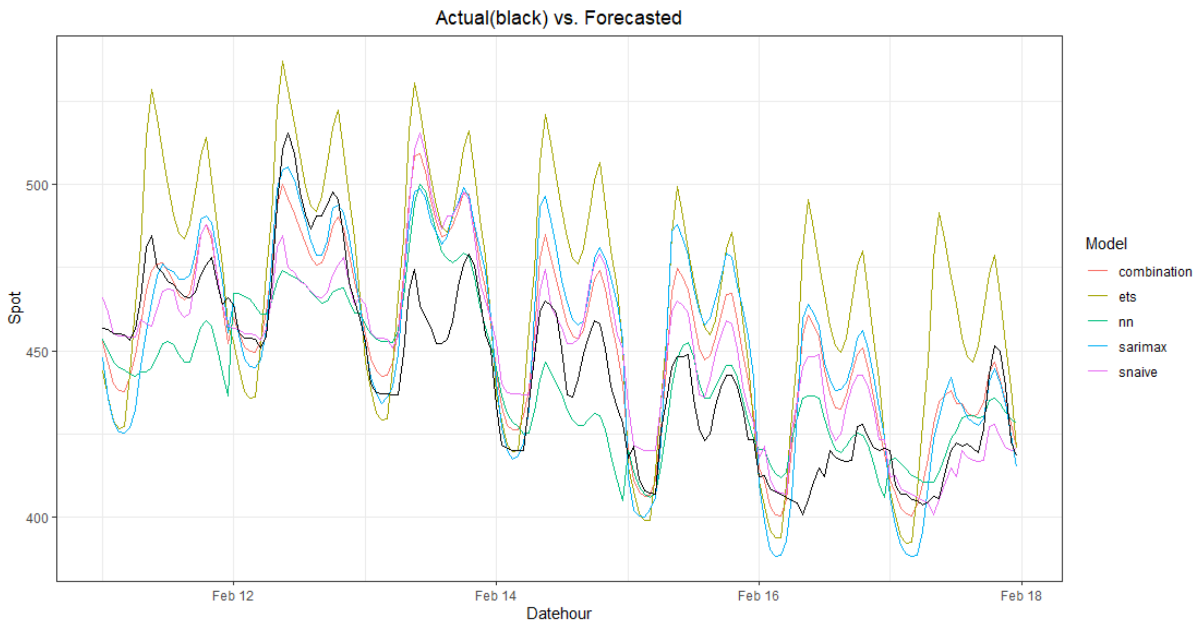
Winter consumption:



Summer spot:

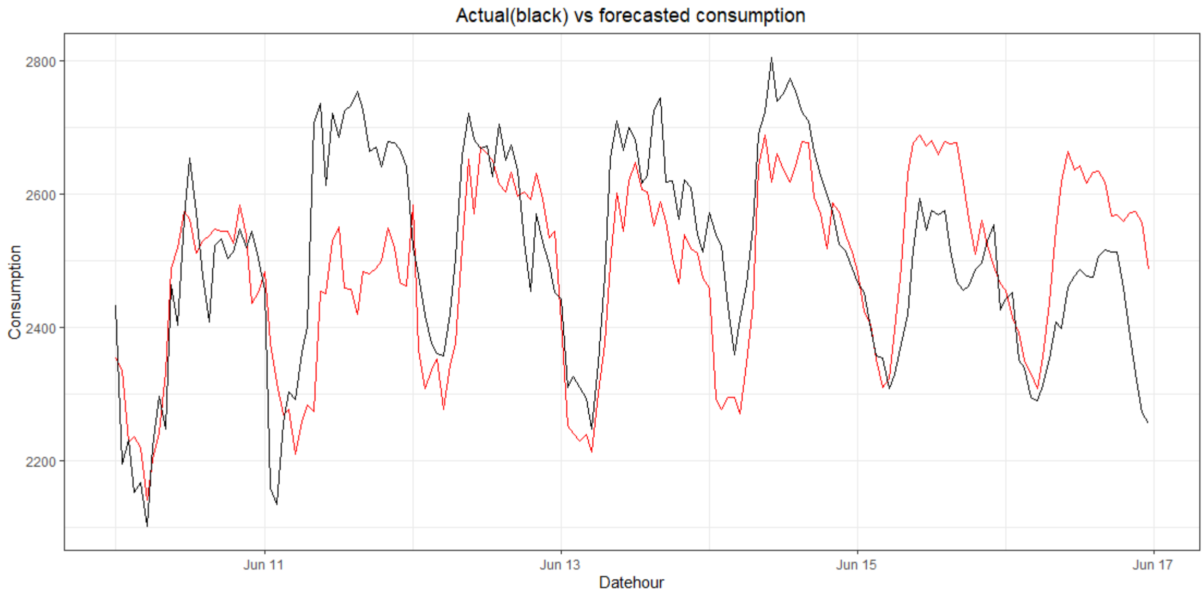


Winter spot:

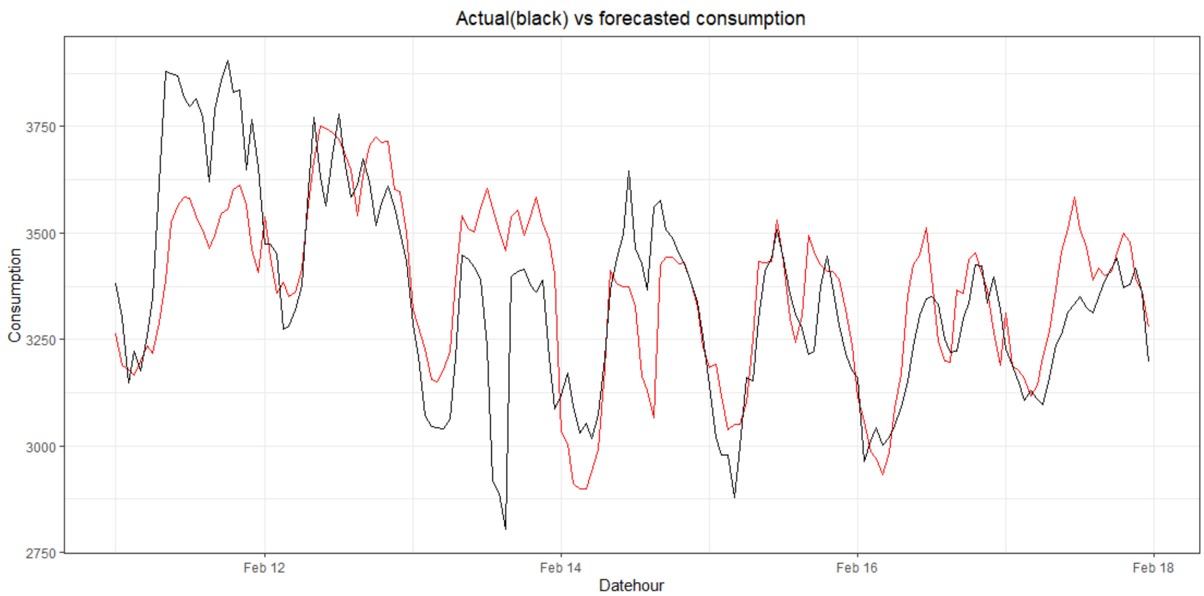


Figures for zone *mid*:

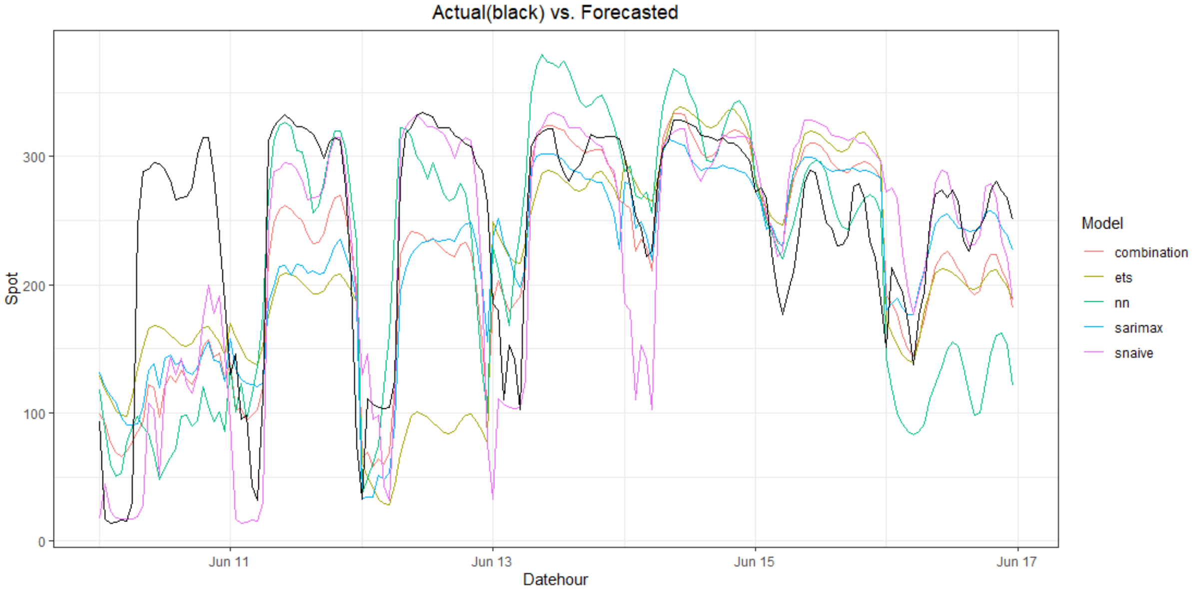
Summer consumption:



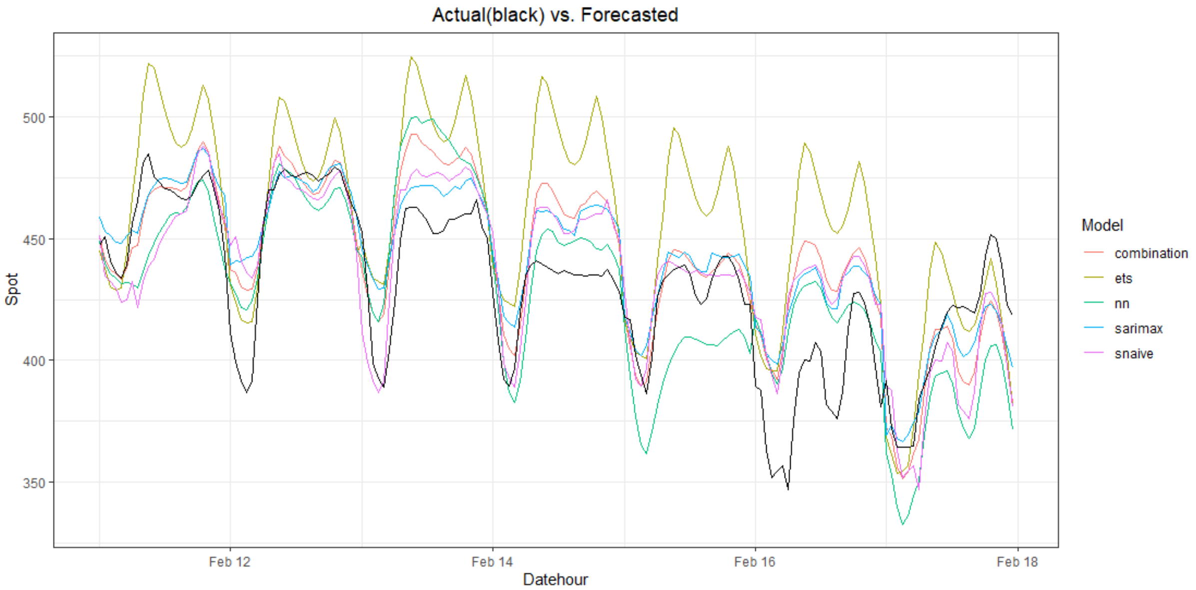
Winter consumption:



Summer spot:

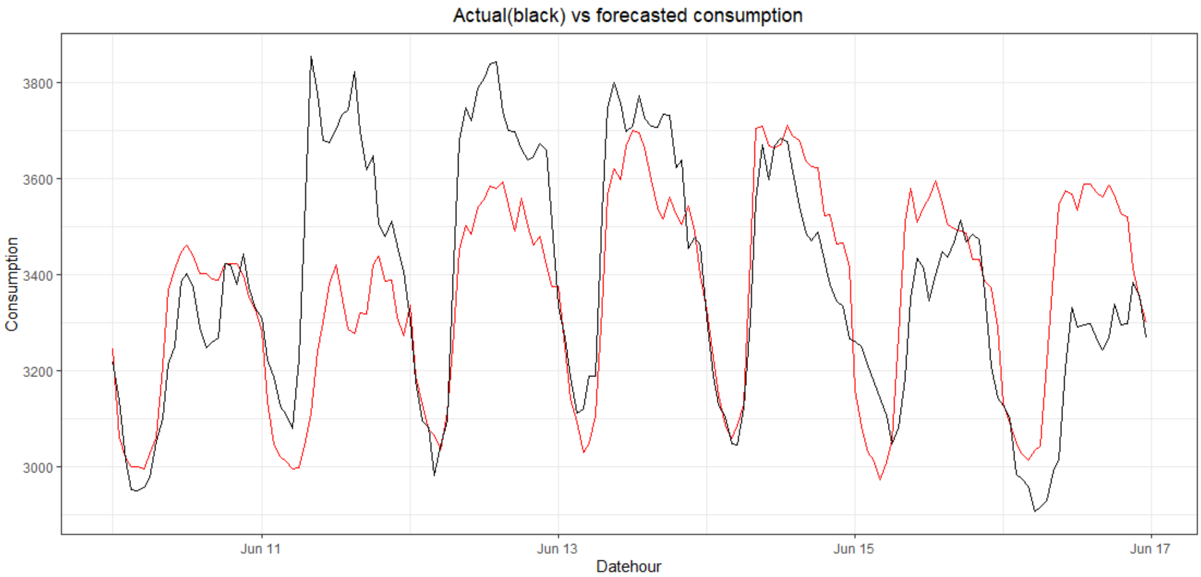


Winter spot:

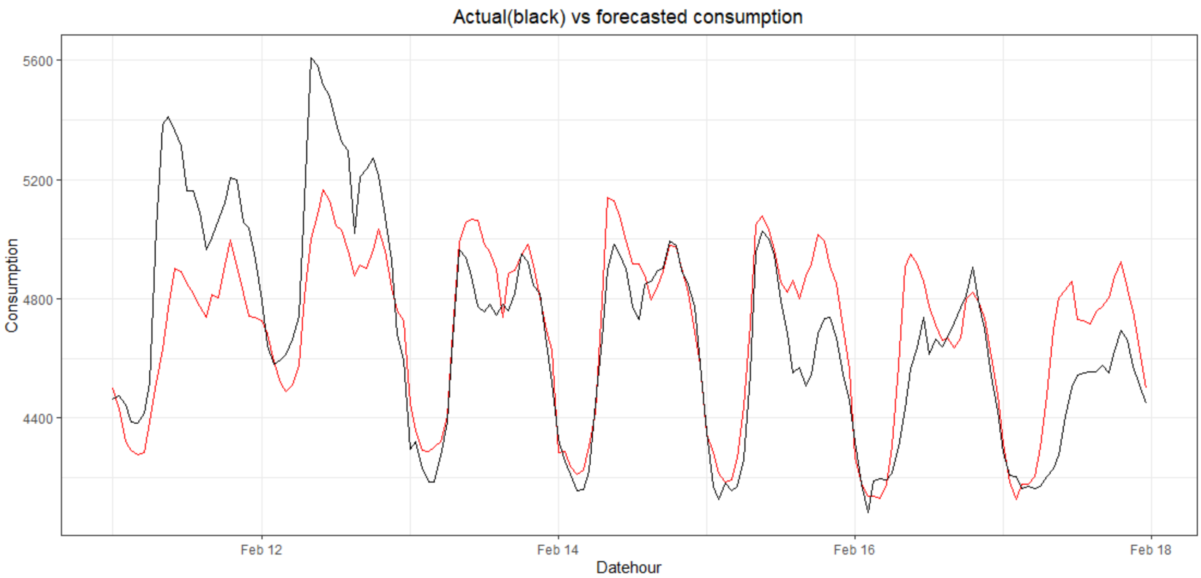


Figures for zone *south*:

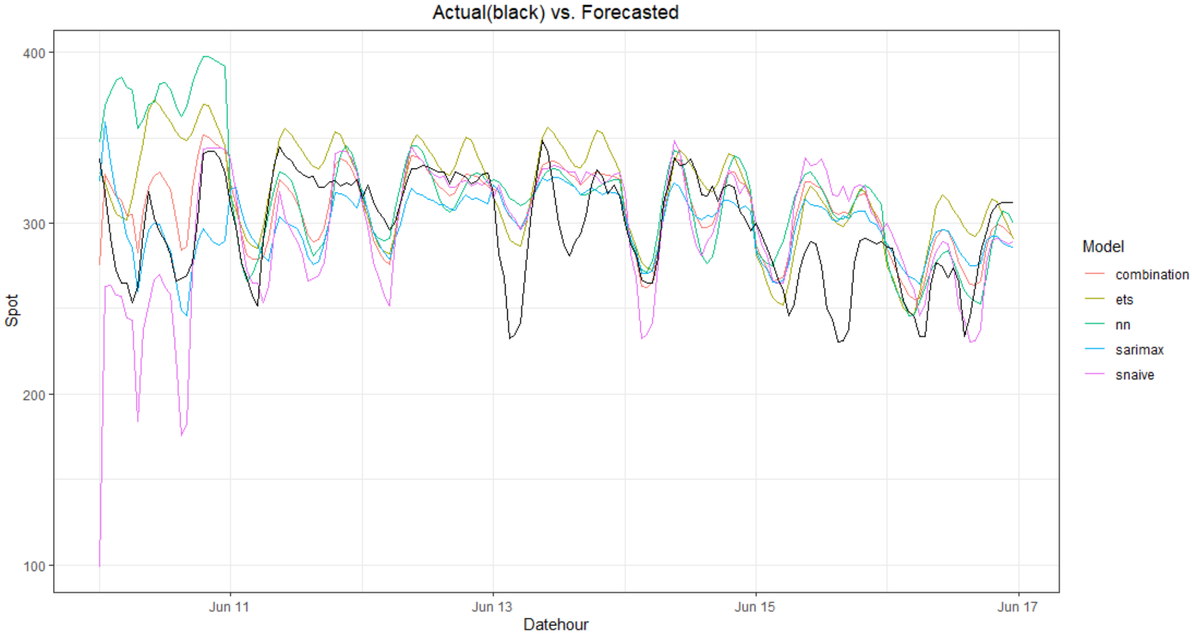
Summer consumption:



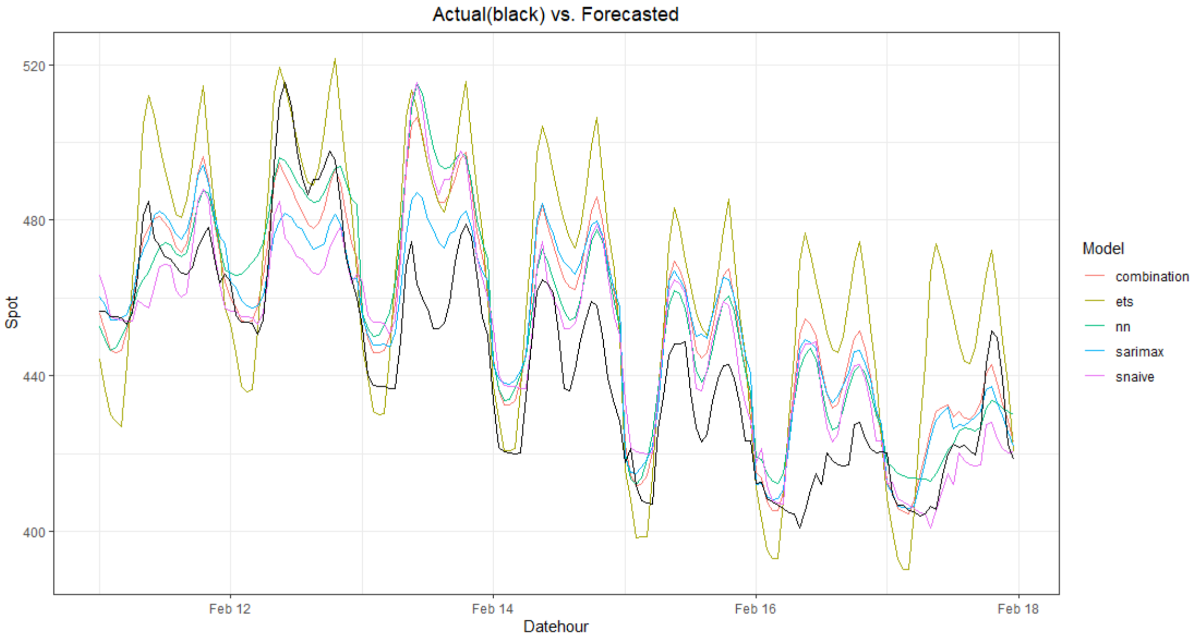
Winter consumption:



Summer spot:

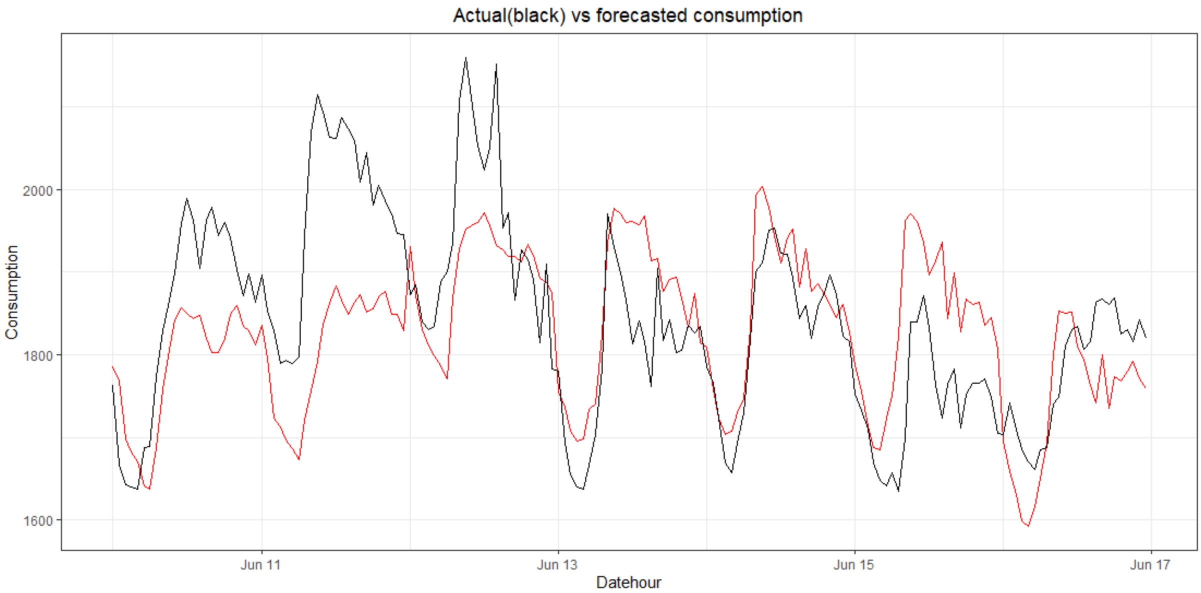


Winter spot:

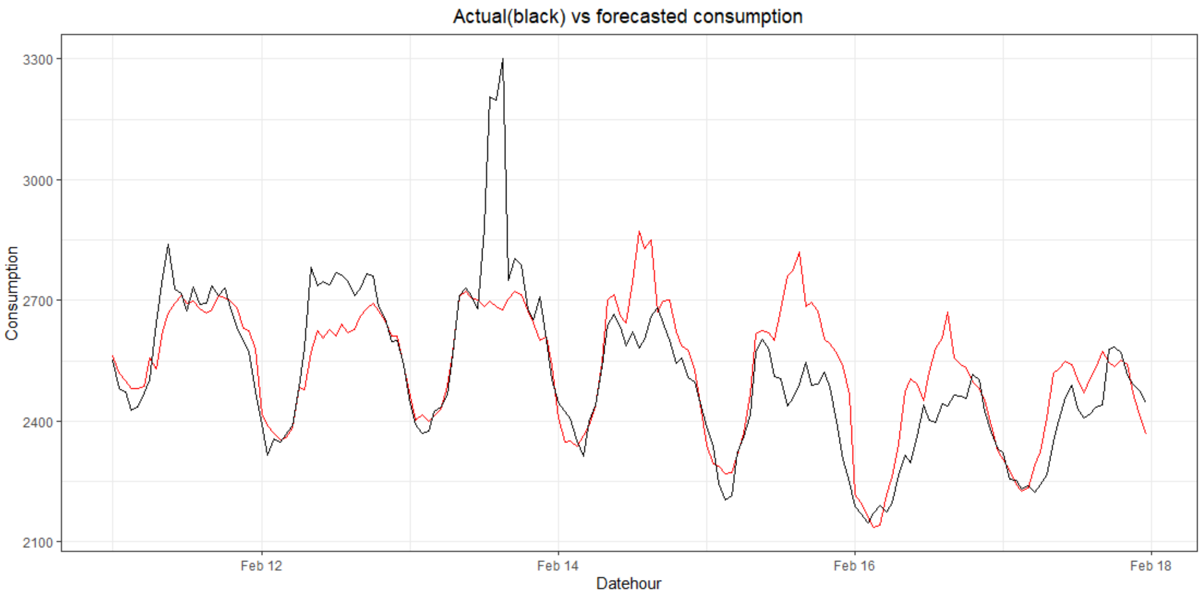


Figures for zone *north*:

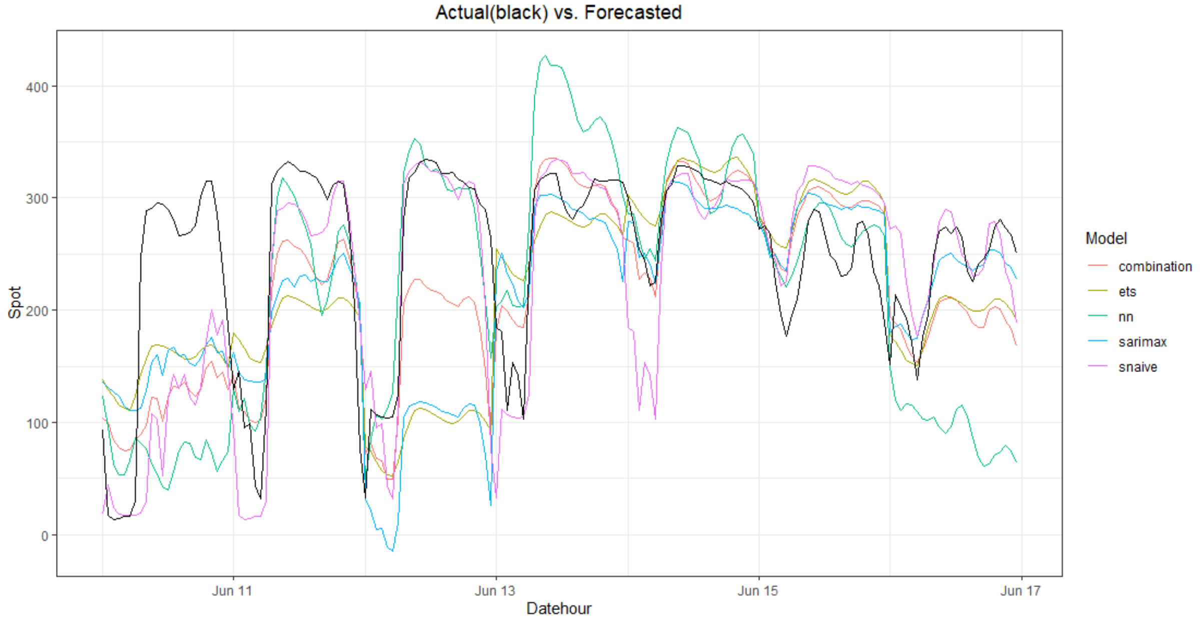
Summer consumption:



Winter consumption:



Summer spot:



Winter spot: