Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Economics and Management
NTNU Business School

Jan-Bendik Kristiansen & Marius Hølen Bakken

Green Hydrogen Production Optimization with Artificial Intelligence in Northern Norway

A Study of ARMAX-GARCH and Seq2Seq-LSTM Power Price Forecasting and Cost Minimization for Increased Predictability and Profitability in a Power Intensive Industry.

Master's thesis in Business Analytics
Supervisor: Johannes Mauritzen
May 2023

NTNU
Norwegian University of
Science and Technology

Jan-Bendik Kristiansen & Marius Hølen Bakken

Green Hydrogen Production Optimization with Artificial Intelligence in Northern Norway

A Study of ARMAX-GARCH and Seq2Seq-LSTM Power Price Forecasting and Cost Minimization for Increased Predictability and Profitability in a Power Intensive Industry.

Master's thesis in Business Analytics
Supervisor: Johannes Mauritzen
May 2023

Norwegian University of Science and Technology
Faculty of Economics and Management
NTNU Business School

NTNU
Norwegian University of Science and Technology

Preface

This thesis is written as part of our Master of Science degree in Economics and Business Administration, with a major in Business Analytics. We would like to thank our supervisor Johannes Mauritzen for his guidance and impeccable knowledge. We would also like to thank Nordkraft AS for valuable information regarding the NO4 power market, and the anonymous corporation providing essential information on hydrogen production.

We take full responsibility for the content of this thesis.

Trondheim, May 2023

Jan-Bendik Kristiansen

Marius Hølen Bakken

Abstract

Northern Norway's abundant hydro and wind-based power resources, in combination with its historically low and stable power prices, have made it an attractive location for power intensive industries. Meanwhile the price and volatility have increased the last two years, causing problems with predictability and concerns about profitability for corporations highly dependent on the power price. This thesis suggests a way to combat this unpredictability with the usage of power predictive models as model guided optimization for green hydrogen production.

This thesis compares two prominent methods for power price forecasting: traditional statistical approaches, represented by the ARMA-family of models, and machine learning techniques, specifically LSTM neural networks. The thesis' most effective statistical method, ARMAX-GARCH, achieved an RMSE and MAE of 14.56 and 6.46 øre/kWh, respectively, while the best machine learning method found, Seq2Seq-LSTM, demonstrated significantly improved performance with an RMSE of 6.43 and an MAE of 2.1 øre/kWh. This was tested on a period from 2016 to 2023 with five day predictions. Furthermore, the study revealed that applying the Seq2Seq architecture to an LSTM structure surpasses the performance of a conventional LSTM model, resulting in a 53% reduction in RMSE for five-day forecasting in the NO4 price area.

To translate these predictive models into practical financial implications for power intensive industries, a hydrogen production case was created using Mixed Integer Programming. Herein, the financial impacts of the forecasting models were compared against a baseline scenario that assumes a constant daily production of hydrogen. The model guided optimization aimed to shift hydrogen production, based on the predictive models identifying the most cost-effective production days within a work week. The results revealed that the Seq2Seq-LSTM guided optimization led to savings of 9.44%, while the ARMAX-GARCH model resulted in savings of 1.81% for a production period of eight weeks. In conclusion, we propose a hybrid AI solution that combines a Seq2Seq-LSTM neural network with a MIP-algorithm to increase predictability and enhance profitability for spot-price-based hydrogen production in the NO4 region.

Sammendrag

Nord-Norges rike ressurser på vann- og vindkraft, i kombinasjon med historisk sett lave og stabile kraftpriser har gjort det til et attraktivt område for kraftkrevende industri. Til tross for dette, har prisen og volatiliteten økt de siste par årene, noe som har skapt problemer for forutsigbarhet og lønnsomhet tilknyttet industrier sterkt avhengig av kraftprisen. Denne masteravhandlingen foreslår måter å imøtekomme denne uforutsigbarheten på, ved bruk av kraftpris-prediktive modeller og modellassistert optimering for produksjon av grønt hydrogen.

Oppgaven sammenligner to fremstående metoder for prediksjon av kraftpriser: tradisjonelle statistiske metoder, representert av ARMA-modeller, og maskinlæringsteknikker, spesifikt LSTM rekurrente nevrale nettverk. Oppgavens mest effektive statistiske metode, ARMAX-GARCH, oppnådde en RMSE og MAE på 14,56 og 6,46 øre/kWt. Av de undersøkte maskinlæringsteknikkene, oppnådde Seq2Seq-LSTM sterkt forbedret resultater med RMSE og MAE på 6,43 og 2,1 øre/kWt. Dette ble observert i løpet av en testperiode fra 2016 til 2023, ved bruk av femdagersprediksjoner. I tillegg til dette, viser studien en reduksjon i RMSE på 53%, ved å legge til Seq2Seq-arkitektur til en LSTM-struktur for femdagersprediksjoner i kraftprisområdet NO4.

For å overføre disse modellene inn i en praktisk økonomisk kontekst for kraftkrevende industrier, ble en produksjonscase for hydrogen skapt ved bruk av Mixed Integer Programming. Her ble modellenes økonomiske påvirkninger sammenlignet med et baseline-scenario som antar konstant daglig produksjon av hydrogen. Den modelassisterte optimeringen søker etter å flytte hydrogenproduksjon, basert på de prediktive modellenes evne til å identifisere de mest kostnadseffektive produksjonsdagene innenfor en arbeidsuke. Studien viser at en Seq2Seq-LSTM assistert optimering resulterte i besparelser på 9,44%, mens ARMAX-GARCH modellen resulterte i besparelser på 1,81% for en produksjonsperiode på åtte uker. Konkluderende foreslår vi en hybrid KI-løsning som kombinerer et Seq2Seq-LSTM nevralt nettverk med en MIP-algoritme, for å øke forutsigbarheten og lønnsomheten for spotprisbasert hydrogenproduksjon i NO4 området.

Table of Contents

Table of figures

1 INTRODUCTION

Northern Norway has emerged as a hotbed for investment in power intensive industries such as hydrogen, steel, ammonia and battery production. The region's rich hydro and wind-based power resources, in combination with historically low and stable power prices have made it an attractive location for corporations seeking to harness the power of renewable energy (Brembo & Olaisen, 2021). Despite this, geopolitical events have caused prices and volatility to rise. Unpredictability in power prices can lead to problems with optimal production planning, higher levels of uncertainty in profitability and poorer decision making. This is specifically the case for green hydrogen production, as power expenses can account for as much as 90% of total variable production costs (anonymous, personal communication, January 2023). Therefore, increasing predictability to make informed decisions is crucial for maintaining the profitability of green production in the region.

This chapter aims to provide a brief overview of the European and Norwegian power markets, focusing on its price development and the potential for heightened prices and volatility. Furthermore, the chapter clarifies the concept of green hydrogen production as it pertains to this thesis. Following this, the chapter introduces the thesis' research questions and outlines the structure of the work ahead.

1.1 Actualization

1.1.1 The European power market

Over the last few years, the Norwegian power market has been more integrated with the European market (Energifakta Norge, 2022). The main argument for this integration is to move power from areas with power surpluses to areas with power shortfalls throughout Europe.

The power market in Europe has been exposed to different impact factors, which have affected the Norwegian power market. In 2008, Norway had 14 power cables exchanging power with neighbouring countries. A power cable between Norway and Germany was opened in 2020, and about a year later, a cable between Norway and Great Britain opened. This contributed to integrating the southern parts of Norway closer to the European market (Kampevoll & Lorch-Falck, 2022). A couple of months after the opening of the latest power cable, the Russian war on Ukraine broke out. With the sanctions on Russia due to the war, Europe stopped importing Russian gas to the union. The shortage of gas contributed to driving the power prices up on the continent, which also influenced the power price in Norway as a consequence of the integrated market (Bugge, 2022). Another factor affecting the power price is the European Union’s decision to cut the union’s climate emissions (Øvrebø, 2023). This implies that the union needs to start producing power from renewable energy sources and phasing out non-renewable sources. According to Cevik and Ninomiya (2022), this is a source of significant uncertainty on the intermittent and volatile production of renewable assets that cause supply-demand imbalances, greater instability in the electricity grid, and more volatile pricing behaviour.

1.1.2 The Norwegian power market

The Norwegian power market is organized into five price areas (Statnett, 2022). This segmentation stems from the fact that most of the Norwegian power is generated from weather-dependent sources. In 2021, 88 % of Norwegian power came from hydroelectric power and 9 % from wind power (Norges vassdrags- og energidirektorat, 2022). Due to varying weather in different parts of the country and limitations in the transfer capacity between the different areas, the power price is not identical for the entire country. The transfer capacity from Northern Norway to Sweden and the Southern part of Norway is limited. This in combination with high power production capacity in the Northern part of Norway, are two of the main reasons for lower power prices in the region (Skagerak Energi, n.d.). The following table shows the different price areas in Norway.

Price area	Region
NO1	Eastern Norway
NO2	Southern Norway
NO3	Central Norway
NO4	Northern Norway
NO5	Western Norway

Figure 1 - Norwegian power price regions

The power price in Norway usually includes both value added tax (VAT) and electricity tax. However, in the Northern parts of the country, private consumers are exempted from these taxes (The Norwegian Ministry of Finance, 2019). Therefore, the following graph illustrates the power prices in the five different price areas, excluding VAT and electricity tax.

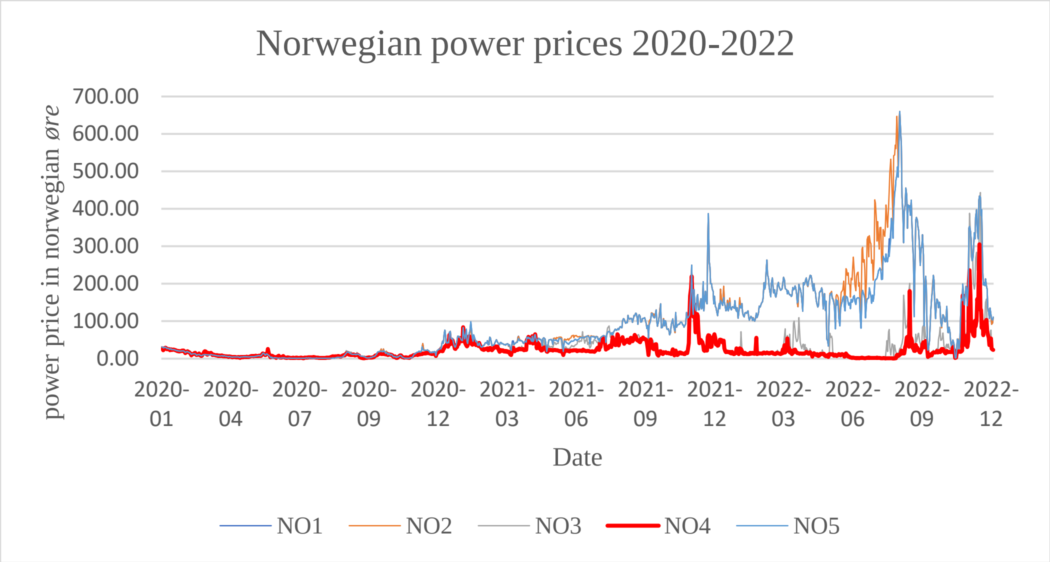


Figure 2 - Norwegian power prices 2020-2022. Source: Montel

1.1.3 Green hydrogen production

The increasing interest in green hydrogen production in recent years can be attributed to its potential to mitigate climate change by reducing greenhouse gas emissions. Green hydrogen is typically produced using renewable energy sources like wind and hydro power in a process that involves electrolysis – splitting water molecules into hydrogen and oxygen (Engie, 2021). However, this sustainable process is not without its challenges, mainly due to its high costs. Power expenses can constitute as much as 90% of total variable production costs (anonymous, personal communication, January 2023). With the profitability heavily relying on the power price, creating methods for optimizing decision-making with regards to the power price are essential for achieving sustainable and cost-effective production. Thus, we will focus specifically on ways to increase predictability and reduce costs in green hydrogen production.

1.2 Research questions

In this thesis, we delve into the NO4 price area of Northern Norway to explore the performances of two branches of predictive models for power price forecasting. Statistical models, which play a major part in price prediction and modelling in the power market, are compared with machine learning methods. ARMA-models, a family of statistical predictive models and artificial neural networks are trained and tested to predict the power price, using datasets for historical power prices, weather conditions and hydroelectric reservoir statuses. The models are fine-tuned specifically for the NO4 price area. However, the goal for this study is not just to compare the performance of these models. Rather, we aim to forge the forecasted outputs in a hydrogen production optimization case that accounts for production targets, constraints, and aims to minimize costs by shifting production based on the outputs from the predictive models using Mixed Integer Programming (MIP). Ultimately, the thesis seeks to address two fundamental questions:

> *How effective are ARMA models and LSTM neural networks in boosting predictability with power price forecasting in the NO4 region, and which model type exhibits superior performance?*
> *To what extent can increased predictability in the power market increase profitability in green hydrogen production, and which model type will yield the greatest financial benefit?*

By adjusting and planning hydrogen production based on the forecasted prices, the thesis' findings can help stakeholders navigate a future where volatile and high energy prices cause concern for spot-price-based production. We believe that the research could shed light on possible strategies for businesses that are significantly dependent on power prices. Additionally, the study investigates the comparative profitability of employing statistical methods versus artificial intelligence in optimizing spot-based production. The thesis also offers a comprehensive comparison of the predictive accuracy of state-of-the-art machine learning against traditional statistical models for power prices in the NO4 region.

1.3 Thesis structure

Chapter 2 of the thesis reviews previous literature on power price forecasting using both ARMA models and LSTM networks. Chapter 3 gives a thorough explanation of the theoretical foundations of statistical methods such as ARMA models and offers a comprehensive overview of neural networks such as LSTM and sequence to sequence structures. The chapter also gives insights into hyperparameters, training neural networks and mixed integer programming. In chapter 4, we detail the methodology used to address the research questions. This includes data processing, the development of the forecasting models, and the creation of a hydrogen production case. The case will serve insights into the profitability of relying on the predictive models, compared to a baseline scenario where the production is based on average daily production. Chapter 5 presents an overview of the performance, evaluation and results of the forecasting models developed independently, before presenting the results from the production optimization case. The chapter provides an assessment of the forecasting models' effectiveness in predicting power prices and their impact on hydrogen production planning. Chapter 6 offers in-depth discussions about the performance of the models, the models' implications on the MIP-optimization, real world application, limitations and critique. Finally, our conclusion and our proposal on how to increase predictability and enhance profitability for green hydrogen production in the NO4 area are presented in chapter 7.

2 LITERATURE REVIEW

Since the liberalization of the power markets in the 90s in both Norway and Europe, the importance of power price forecasting has motivated researchers for testing and creating new methods and models for this purpose (Cerjan et al., 2013). Through the next decades, different models for power price predictions were tested, such as Autoregressive Moving Average (ARMA) models and machine learning techniques like artificial neural networks. The ARIMA (autoregressive integrated moving average) model and its extensions, is a widely used group of statistical models for time series modelling, when basing the forecast on historical data from the target variable being forecasted. We have found several papers forecasting single step with ARMA models, but noticeably fewer on multistep forecasting. Power price forecasting with neural networks is still an ongoing field of research, however in recent years multiple articles surrounding the subject have been published, especially related to the usage of Long Short-Term Memory recurrent neural networks (LSTM). In addition to that, new techniques for time series forecasting with LSTMs has also emerged. The following section contains an overview of relevant research papers which position this thesis with existing literature.

2.1 Forecasting with ARMA/ARIMA models

Studies regarding power price forecasting from the Nord pool area using ARMA models have been conducted to some extent. Hipolit Torro (2007) wrote a paper on weekly power price prediction with daily observations from the Nordic Power Exchange area using ARIMAX (ARIMA with exogenous variables). He used temperature, precipitation, reservoir levels, and a basis (futures price less the spot-price) to reflect the seasonal patterns and found the ARIMAX model to have a significantly lower MSE (mean squared error) than the Myopic method. This method takes the present spot-price as the forecasted price, and the Futures method which takes present futures prices as forecasted price.

Swider and Weber (2007) used the ARMA model with and without GARCH (generalized autoregressive conditional heteroskedasticity) for power price forecasting developments in two price areas in Germany on the day-ahead market. They found that applying the GARCH approach on one of the areas slightly improved the MAPE (mean absolute percentage error) compared with an ARMA model without GARCH. Furthermore, the paper notes that there are indications that the GARCH approach improved the representation of the price distribution on all the considered markets.

Chang et al. (2010) used a rolling window approach on their VARMA-GARCH (vector ARMA-GARCH) and VARMA-AGARCH (vector ARMA-asymmetric GARCH) models to forecast the conditional correlation in the crude oil price one day ahead. The approach was added to the models to explore the time-varying nature of the target variable. The paper explains the rough features of the rolling window technique.

In 2009, a study of the main methodologies, including different ARMA models, used in power price forecasting was released by Aggarwal et al. (2009). The paper was not able to point out any systematic evidence that one of the models outperformed the other, but the writers believed this was because of short power price history, due to the liberalization of the market in the 90s, and large differences in price developments in the distinctive power markets.

2.2 Forecasting with neural networks

In a 2021 study, Memarzadeh and Keynia aimed to forecast the electricity market for load and price in Spain and Pennsylvania-New Jersey-Maryland, USA. Their power price forecast achieved a mean absolute percentage error (MAPE) of 0.93 for Spain and 2.2 for Pennsylvania, demonstrating the effectiveness of their approach. The study specifically utilized an LSTM recurrent neural network and concluded that it had strong capabilities for accurate power price forecasting.

Despite being relatively few studies on the use of LSTM for power price prediction in Norway, and especially northern parts of Norway, a group of researchers from NTNU did conduct a study on the topic in 2022. The researchers compared the use of a simple ANN and a LSTM model for day ahead power price forecasting in different price regions (Vamathevan et al., 2022). They found that their LSTM network performed better in the more volatile areas like NO1 and NO2 compared to NO4. However, they used the same model for each price area and did not tune them specifically to each region (Ü, Cali, personal communication, February 13, 2023).

Aranguren, Fragoso, and Aguilera’s 2022 study proposes a new approach for forecasting oil production in the Eagle Ford Shale (Texas, USA) using a Seq2Seq-LSTM (sequence to sequence) based learning framework. This methodology utilizes an encoder-decoder architecture to generate future declining production rates, incorporating recurrent neural networks and Seq2Seq-architectures that are commonly used for language processing in translation tasks. The model achieves this by utilizing historical data sequences to predict future oil rates. The study demonstrates that the Seq2Seq architecture outperforms a regular LSTM architecture and suggests that it can be a reliable and efficient technique for time series forecasting, in this specific example, within the oil sector.

In their 2019 study, Gong et al. developed a short-term load prediction model based on a Seq2Seq architecture. The model utilizes a LSTM neural network to address the time and non-linear characteristics of power system load data and analyse the periodic fluctuation characteristics of users' load data. The study compared the predictive performance of the model under different types of attention mechanisms and ultimately adopted the Seq2Seq-LSTM model. The results demonstrated the potential of Seq2Seq models in short-term prediction and its ability to improve prediction accuracy. The study also provided a thorough explanation of how the Seq2Seq and LSTM structure works. The proposed model has the potential to aid power-related departments in developing more effective power utilization plans.

3 THEORETICAL FRAMEWORK

This chapter aims to provide a comprehensive understanding of the key theories and models that constitute the core of this research. First, the chapter provides theory on forecasting models such as ARMA, LSTM recurrent neural networks, and sequence to sequence structures. The chapter then highlights hyperparameters, training, and evaluation metrics used in these models. Lastly, the chapter introduces the theory of Mixed Integer Programming (MIP), which is the algorithm used for the hydrogen production case in this thesis.

3.1 The ARMA family of models

There are several different types of models with the base of an autoregressive moving average model. The ARMA models are widely used to predict time series forecasting as they aim to describe the autocorrelation in the data (Athanasopoulos & Hyndman, 2018).

3.1.1 The ARMA model

A simple ARMA model is based on two components, the autoregressive model (AR) and the moving average model (MA). The autoregressive models are based on the idea that past values of a variable can explain the current and future values of the same variable (Shumway & Stoffer, 2017, p. 75). The model is often referred to as an AR(p) model, which is an autoregressive model of order p (Athanasopoulos & Hyndman, 2018). At which order of p the model will be optimal, depends on the autocorrelation between the lagged variables. The model can be written as followed (Athanasopoulos & Hyndman, 2018).

$$y_t = c + \emptyset_1 y_{t-1} + \cdots + \emptyset_p y_{t-p} + \varepsilon_t$$

The model can almost be interpreted as an ordinary regression, but contains p number of lagged variables of the same variable instead of different exogenous variables. The ε_t present uncertainty in the data that is not accounted for in the model, also called the white noise. The $\emptyset$ equals the coefficient of the relationship between the lagged variables (Turing, n.d.).

The moving average model referred to as the MA(q) model, is based on the same principle as an AR(p) model but uses past errors in the forecast to make future predictions (Athanasopoulos & Hyndman, 2018). When combining the AR(p) and the MA(q) model, the ARMA model shows.

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

For the ARMA model to be functional, there are parameters that need to be appropriately specified (Makridakis & Hibon, 1997). The observed data in the time series needs to be stationary, the number of AR(p) and MA(q) need to be in an appropriate order and the values of the parameters in the model must be estimated with an appropriate loss function. For the dataset to be stationary, the mean and the autocorrelation of the observations at any given time must be constant (Shumway & Stoffer, 2017, p. 20). To achieve stationarity in the time series, a common approach is to use differencing (d) which implies to calculate the difference between one observation to the next (Athanasopoulos & Rob, 2018). The combination of an ARMA model and a reversing of the differenced time series gives an ARIMA(p,d,q) model. A common approach to determine whether the data is stationary is to use an Augmented Dicky Fuller test (Prabhakaran, 2019). The selection of the parameters for the model can be challenging (Athanasopoulos & Rob, 2018). Languages such as Python or R may help select the parameters automatically. If the time series contains non-stationary data, the algorithm first determines the order of differencing necessary to get the data stationary, and then predicts the order of p and q. When selecting the orders of AR(p) and MA(q), the algorithm minimizes the AIC, Akaike's Information Corrected Criterion (Akaike, 1974).

The ARIMA(p,d,q) model is widely used to forecast time series, however, a simple ARIMA will not account for seasonal correlations (Athanasopoulos & Rob, 2018). When adding the seasonal compound to the ARIMA, the models are capable of calculating seasonal time series. To get the time series stationary, it might be necessary to do both a seasonal difference and an ordinary difference. While the idea behind the seasonal difference is the same as the ordinary, the seasonal difference computes the differences between an observation and the same observation in a previous period. These types of models are called SARIMA (seasonal ARIMA). These models also have a seasonal component which makes the model capable of modelling seasonal data (Athanasopoulos & Rob, 2018). For further information regarding this model, see Athanasopoulos and Rob (2018).

3.1.2 ARMA with exogenous variables

So far, the chapter has focused on the ARMA and the ARIMA models with purely one dependent variable predicted by the historical observations of the same variable. An ARMA model can be expanded to accommodate exogenous variables, which is often referred to as an ARMAX or ARIMAX model. These models predict future values of the target variable by using both past values from the target and exogenous variables (Cools et al., 2009). The ARMAX can use several different series of data input to predict the target variable (Changtzeng et al., 2019).

The model uses impulse response weights, where the values of the weights represent the response on the target variable, depending on the changes in the exogenous variable from one observation to the next. When predicting with an ARMAX, the model is not able to predict the values of the exogenous variables, only the target variable.

3.1.3 ARMAX with GARCH

A common issue when modelling ARMAX models is heteroskedasticity, which occurs when the residuals in the model are not constant. One way to deal with this issue is to combine the ARMAX model with an ARCH or GARCH model (Chen et al., 2011). The presence of heteroscedasticity in the model indicates that the GARCH model is appropriate to combine with the ARMA model (Chen & Yu, 2013). The GARCH model measures and predicts the variance of each error term (Engle, n.d.). A weighted average of squared residuals of the past is used with a declining weight to predict future values.

When applying the residuals of the fitted ARMAX into the GARCH model, the model predicts the future residuals from the model. When forecasting using both ARMAX and GARCH, the forecast has accounted for both the future estimate from the ARMAX and the predicted residuals from the GARCH model. The ARMAX(p,q)-GARCH(r,m) model can be presented mathematically as followed where the first model is the ARMAX and the second is the GARCH. The R_t term in the first model is the daily observation of the dependent variable and the $E(R_t|F_{t-1})$ term is the conditional mean of the daily observation from the ARMAX model. The σ_t presents the volatility and η_t represents an error term distributed according to a given distribution f with parameter set θ in the GARCH model (Porshnev et al., 2016).

$$y_t = R_t - E(R_t|F_{t-1})$$

$$y_t = \sigma_t * \eta_t \,, \eta_t \sim f(\theta)$$

3.1.4 Sliding window approach

To understand the model's performance when predicting five timesteps ahead, several different methods can be applied to the model. A standard ARMA forecast with a test set containing numerous predictions ahead will not be presentable for the model's performance on five-time steps ahead. The approach, also called a rolling window approach, involves calculating the model on a fixed contiguous block (window) of prior observations and then predicting a number of given timesteps ahead of the window (Brownlee, 2017b). The algorithm then repeats this process until the model has iterated through the dataset. The approach is commonly used to backtest a statistical model on historical data to evaluate the model (Zivot & Wang, 2006).

3.2 Neural networks

Neural Networks, also referred to as Artificial Neural Networks (ANNs) or Simulated Neural Networks (SNNs), is a branch of machine learning that form the backbone of deep learning algorithms (IBM, 2019). Inspired by the structure and functioning of the human brain, without being an accurate representation, these networks consist of a series of artificial neurons, often called nodes, arranged in layers: an input layer, one or more hidden layers, and an output layer (IBM, 2019). The nodes in the network are connected, each with its own weight and bias value. When a node receives an input, it applies a non-linear activation function to the weighted sum of the inputs and the biases and produces an output that is passed to the next layer. Neural networks are a diverse set of machine learning models, each with its strength and limitations, used for a wide range of tasks. Feedforward neural networks, convolutional networks (CNNs), recurrent neural networks (RNNs), and generative adversarial networks (GANs) are commonly used for language translation, image recognition, time series prediction and text generation. Neural networks offer powerful tools for machine learning with many applications across various fields (IBM, 2019).

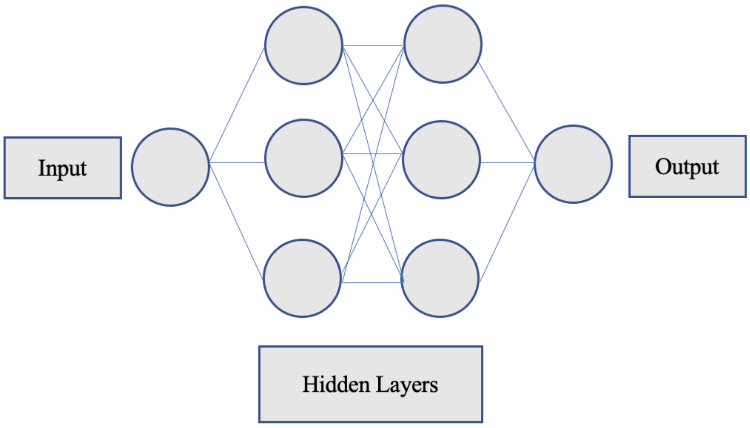


Figure 3 - Basic ANN architecture

One of the core building blocks of neural networks are the layers. Francois Chollet (2018, p. 28) argues that the layers in neural networks best can be described as data-processing filters. Some data goes in (input), and some data goes out (output), with the goal of filtering the data into a more useful form. Usually what defines deep learning models from typical machine learning is two or more hidden layers, however many state-of-the-art deep learning models tend to be much deeper, with tens or even hundreds of hidden layers within the neural network (Chollet, 2018, p.8).

3.2.1 Recurrent neural networks

One type of neural network that excels at time series forecasting is recurrent neural networks. The reason why RNNs is a type of neural networks that tends to perform well on time series forecasting is due its ability to remember (Chollet, 2018, p.196). Biological intelligence uses an incremental approach in processing information and continually updates its internal representation, constructed from previous experiences and newly acquired information (Chollet, 2018, p.196). This is essentially, although simplified, the same way RNNs process sequences of information, by iterating through the sequence of elements and preserving a state containing information that is relative to what the model has seen so far.

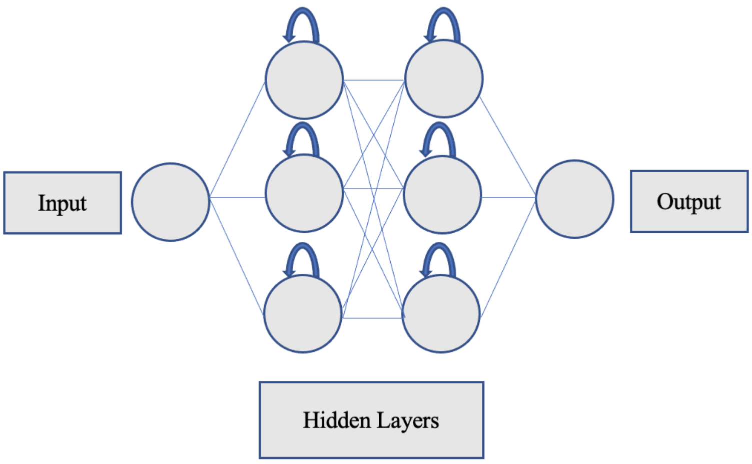


Figure 4 - Basic RNN architecture

In general, the aim of RNNs is to detect dependencies in sequential data by finding correlations between different points within a sequence. There are essentially two kinds of dependencies. Short-term, which describes a dependence in the recent past, and long-term, which shows dependencies between points that are further away from each other (Weller, 2018). Finding and understanding these dependencies are crucial for the neural network to work and predict accurate results and forecasting future trends. This is where standard RNNs have an issue. The problem with a standard RNN is that they are only able to detect short term dependencies, and the reason for this is called the vanishing gradient problem (Arbel, 2018). In order to understand the vanishing gradient problem, understanding how RNNs are trained is beneficial.

When training RNNs the inputs are fed into a recurrent unit that undergoes weighting, where they are multiplied by weight matrices. These weights reflect the significance assigned to each value by the model. The network improves its predictions by modifying the weight matrices to values that lead to better predictions (Weller, 2018). After processing a complete sequence, the network evaluates the accuracy of its predictions (compared to the actual values known as the labels) by calculating the error. The network then backpropagates through the entire sequence, adjusting the individual weight matrices to minimize the error. This method is known as backpropagation. However, because of the additional time dimension, RNNs need to use a backpropagation that not only goes back through the different hidden layers where an optimizer function adjusts the weight matrices, but also goes back in time adjusting all the weights of previous time steps. Training and backpropagation can become a problem if the sequence is long, and the network must go back after every prediction. Truncated backpropagation training (TBPTT) solves this issue by splitting up the sequence (Brownlee, 2017a). Backpropagation is only applied to the length of the truncated subsequence at each instance. However, this means that the network can only learn dependencies within those subsequences, so it is important to consider the length when defining and optimizing the RNN. This is when the so-called vanishing gradient problem can appear. The further the model goes back the sequence, the less importance the learned values can have on the current predictions, and it prevents the model from learning long term dependencies (Weller, 2018). When forecasting with long time series data, this problem can easily appear, therefore it is essential to find a way around it.

3.2.2 Long short-term memory

When doing time series forecasting one can expect the data to have some degree of long-term dependency. A plausible assumption is that the price is influenced by a wide range of factors such as previous prices, weather patterns, and the amount of water in the hydroelectric reservoirs. All these factors can have long-term dependencies, which can lead to the vanishing gradient problem. This means that the standard RNN can be a poor choice for the purpose of predicting power prices. A more complex form of RNNs is called Long Short-Term Memory (LSTM). LSTM is a type of RNN that solves the vanishing gradient problem, and it solves it by introducing memory cells (Chollet, 2018, p.202). Memory cells are like small units that store information for an extended period of time, as well as gates that control when information is allowed to enter or leave memory cells. All recurrent neural networks have the form of a chain of repeating modules. In a standard RNN, this repeating module will have a simple structure, such as a single hyperbolic tangent layer (tanh) (Olah, 2015), as seen in figure 5.

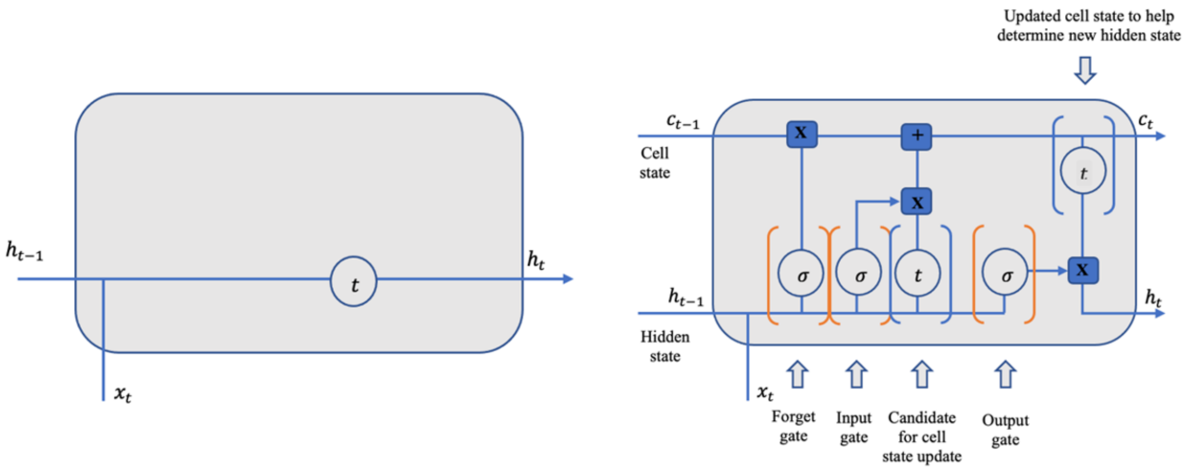


Figure 5 - RNN and LSTM cell structure

LSTMs also have a chain like structure, however instead of a single layer, there are four. The LSTM *forget gate* determines which information is relevant and which can be disregarded. It takes the *current input* x_t and the previous *hidden state* h_{t-1}, and passes them through a sigmoid function, producing values between 0 and 1. A value closer to 1 indicates that the old output is necessary, and this value is later used for point-by-point multiplication by the cell (Olah, 2015). The LSTM *input gate* updates the cell status by first passing the *current state* x_t and the previous *hidden state* h_{t-1}, through a sigmoid function, then passing the same information through a tanh function to create a vector of values between -1 and 1. The outputs from both activation functions are used for point-by-point multiplication (Olah, 2015).

Once the network has enough information from the *forget gate* and the *input gate*, it decides which information to store in the new state of the cell. The previous *cell state* C_{t-1}, is multiplied by the *forget vector*, and any values that results in 0 are dropped from the cell state. The output value of the input vector is then added point-by-point, resulting in a new cell state C_t. The *output gate* determines the value of the next *hidden state*, which contains information on previous inputs. It passes the *current state* and previous *hidden state* through a tanh function, and the outputs are multiplied point-by-point. Based on the final value, the network decides which information the *hidden state* should carry. This *hidden state* is used for prediction, and both the *new cell state* and the new *hidden state* are carried over to the next step (Olah, 2015).

3.2.3 Sequence to sequence models

Sequence to sequence (Seq2Seq) structured models, a type of neural network structure that has gained popularity in recent years, first introduced in the relatively recent 2014 paper, "Sequence to Sequence Learning with Neural Networks" by Sutskever et al. The paper proposed an encoder-decoder architecture based on recurrent neural networks (RNNs), specifically LSTM networks, to model input and output sequences of data. The encoder and the decoder are two independent neural networks jointly trained together.

In this architecture, the encoder LSTM processes the input sequence one element at a time and generates a fixed-length context vector that summarizes the input sequence. The decoder then uses this context vector as a starting point to generate the output sequence one element at a time. This approach allows the model to handle variable-length input and output sequences, making it suitable for tasks such as machine translation, text summarization and speech recognition, typically where the output can be of a different length than the input (Wu et al., 2016).

While Seq2Seq models were originally developed for natural language processing tasks, they have also been applied to time series forecasting. These structures have shown to be especially effective for multi-step forecasting (Gong et al., 2019). By taking a desired sequence of past data points as input, the model can generate a single predicted value for the next step in the sequence. This means that the network can be trained to predict an arbitrarily future sequence of data points, called timesteps, given an arbitrarily past sequence of data points often called window size (Gong et al., 2019; Aranguren et al., 2022). In simple terms the model uses a window size together with the learnt weights of the model to generate a desired sequence.

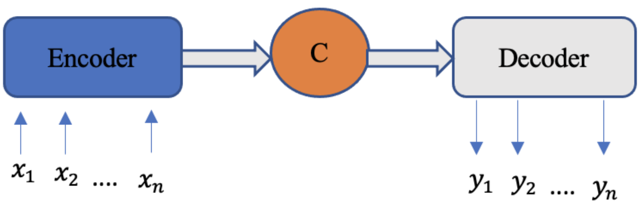


Figure 6 - Seq2Seq codec

The figure above (figure 6) shows a Seq2Seq codec, and the figure is heavily inspired by the one created by Gong et al. (2019). The codec is composed of an encoder, an intermediate vector C, and a decoder. Usually, codecs are made up of multi-layer RNN or LSTM structures, where the intermediate vector C incorporates the sequence of the x_1, x_2 ... x_m encoding data. At time t, the input to the decoder is the output of the previous moment y_{t-1}, the hidden layer state of the previous moments s_{t-1} and C. Subsequently, the decoder produces the hidden layer state s_t which predict the output value at time t (Gong et al., 2019). For further enlightenment on the general workings of the structure, see Sutskever et al. (2014).

3.2.4 Optimizer

The fundamental function of optimizations algorithms is to find the optimization values to the appropriate neural network weights to minimize the objective function (loss function) (Llugsi et al., 2021). Some common optimizers are stochastic gradient descent (SGD), Adam, Adagrad, RMSProp and Newton's method. However, this thesis will focus on the Adam optimizer. Adam is an algorithm for first-order gradient-based optimization of stochastic functions. The algorithm is easy to implement with computationally efficient features with little memory requirements. Like other gradient-based optimization algorithms, Adam is used to update the weights of a model based on the gradient of the loss function with respect to the weights. However, Adam differs from other optimization algorithms in that it uses an adaptive learning rate, which helps it converge more quickly and avoid getting stuck in local optima (Kingma & Ba, 2017).

3.2.5 Activation functions

Activation functions are essential components of neural networks that introduce non-linearity to the output of neurons. The primary purpose of activation functions is to transform the weighted sum of inputs and biases into an output that can be fed to the next layer of the network. Without the ability to model non-linear relationships, neural networks would not be much more than an unnecessary complex linear classifier or regression (Goled, 2021). Some of the most common activations functions are sigmoid, hyperbolic tangent (tanh) and rectified linear unit (ReLU) (Nielsen, 2018). The sigmoid function is a smooth differentiable approximation of a threshold unit and compresses the inputs into the range between 0 or 1, and it is therefore widely used as an activation function for classification tasks (Nielsen, 2018, p.8). However, for regression problems the sigmoid function is often not a suitable activation function because of the output values it produces (0,1).

The tanh function is a mathematical function similar to the sigmoid function (Nielsen, 2018, p.121). It is a non-linear function that maps input values to a range between -1 and 1, which makes it useful for certain types of problems, particularly in output layers where the output values need to be in the range of -1 to 1.

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

The ReLU function is a popular activation function used in neural networks. It is a non-linear function that has been found to perform well in many applications (Nielsen, 2018, p.123). The

ReLU function sets any negative input to 0 and any positive input to itself, resulting in an output that is always non-negative. The ReLU function is described as:

$$f(x) = \max(0, x)$$

Where the x is the input value of the function, and $\max(0, x)$ returns the maximum value between 0 and x. The function is computationally efficient, as it only involves a simple thresholding operation, which makes it particularly useful in large neural networks. One of the advantages of the ReLU function over other activation functions such as the sigmoid or tanh is that it does not suffer from the vanishing gradient problem (Brownlee, 2020). This can make it easier to train deep neural networks. The vanishing gradient problem occurs when the gradients of the activation function become very small, making it difficult for the network to learn and update weights properly. However as pointed out by Michael Nielsen (2018, p.124), we do not have a clear understanding of when, exactly, rectified linear units are preferable, nor why. Choosing the right activation function is not obvious, and in most cases trial and error is the best strategy.

3.2.6 Training, overfitting and regularization

Neural networks are in general known to be difficult to train, and recurrent neural networks is no exception. This is especially true when it comes to deep neural networks (Nielsen, 2018, p.152). Adding more layers to the network might improve how the network figures out complex pattern in the data, but it can also cause problems. Different layers in the deep neural network can learn at vastly different speeds. In particular, when later layers in the network are learning well, early layers often get stuck during training, and learns almost nothing at all (Nielsen, 2018, p.154). The opposite can also occur: the early layers may be learning well, but the later layers can become stuck. A lot of these issues boils down to the vanishing gradient problem, as already discussed. Another common problem is overfitting, which is related to the issue of training neural networks.

Overfitting occurs when the model is struggling to generalize on new data (Nielsen, 2018, p.74). For instance, the loss function indicates that the model is performing quite well on training data, but when evaluating the model on test data, the loss function skyrockets. This is happening because the model is overfitted to the training data, in such a matter that instead of learning, it starts remembering the patterns of the data, thus not being able to generalize. There are ways to combat overfitting. Reducing or increasing complexity of the architecture, adding more training data, hyperparameter tuning or regularization techniques.

Regularization is a way of compromising between finding small weights and minimizing the original cost function (Nielsen, 2018, p.79). The main aim of regularization is to reduce the over-complexity of the model and help the model learn a simpler function to promote better generalization on the test data. Two common types of regularization techniques are L1 and L2 regularization (Neumann, 2014). L1, also called Least Absolute Shrinkage and Selection Operator (LASSO) can be formulated:

$$Loss(L1) = \sum_{i=0}^{N}(y_i - \sum_{j=0}^{M} x_{ij}W_j)^2 + \lambda \sum_{j=0}^{M} \mid W_j \mid$$

L1 regularization adds a penalty term to the loss function that is proportional to the absolute value of the weights in the model. This encourages the model to use fewer features and select only the most important ones.

L2, also called Ridge regularization can be formulated:

$$Loss(L2) = \sum_{i=0}^{N}(y_i - \sum_{j=0}^{M} x_{ij}W_j)^2 + \lambda \sum_{j=0}^{M} W_j^2$$

L2 regularization adds a penalty term that is proportional to the square of the weights in the model. This encourages the model to have smaller weights and reduces the impact of outliers in the data.

Adding dropout layers is another way to combat overfitting. Dropout involves randomly dropping out, or "turning off", some of the neurons in a layer during training (Srivastava et al., 2014). This helps to prevent the network from relying too heavily on any one particular feature or neuron and encourages the network to learn more robust and generalizable representations of the input. During training, each neuron in the layer is assigned a probability *p* of being dropped out. At each iteration of training, each neuron is either kept with probability *p*, or dropped out with probability *1-p*. The neurons that are kept are then scaled by a factor of *1/(1-p)* to compensate for the fact that fewer neurons are active (Srivastava et al., 2014).

During inference, all neurons are used, and no scaling is applied. Typically, dropout layers are added to the network's architecture after the activation function of a fully connected layer. Moreover, the performance of deep neural networks is highly dependent on the type of hyperparameters chosen for the model (Nielsen, 2018, p.168). This includes activation functions, optimizer algorithms, loss function, general architecture of the network as well as batch size and epoch settings for training. Additionally, in machine learning, it is common to split the data into three subsets: training, validation and test. The training set is used to train the models into learning the weights and relationships of the data using a chosen algorithm. The validation set is used to fine-tune the models hyperparameters, and to perform model selection. Finally, the test set is used for a final unbiased evaluation of the models' abilities to generalize to new, unseen data. The test set is meant to work as a simulation on how the model would perform when applied in practice (Toisoul, 2020).

3.2.7 The Black Box Phenomenon

Black box models are machine learning systems that are characterized by their opacity: they do not reveal their internal mechanisms, and their parameters cannot be understood through inspection (Molnar, 2019). This lack of transparency is the heart of the black box phenomenon, a problem heavily discussed within the community of artificial intelligence.

Despite superior performance in a variety of domains, the lack of interpretability in deep neural networks is believed to be hurting their adaptation. Users may not trust systems whose decisions processes, they do not understand (Ahn et al., 2021). The lack of interpretability arises from several factors such as the complexity of the architecture. Deep neural networks can contain numerous layers and neurons, leading to the creation of thousands, or even millions of parameters. For example, the highly discussed ChatGPT, is a neural network that consists of 175 billion parameters (Hughes, 2023), and although this is a complexity far exceeding most deep neural networks, even thousands of parameters can cause challenges with interpretability.

As previously discussed, the activation functions in neural networks are non-linear. The outputs are a complex combination of the inputs, and it can be challenging to understand how the network arrives at its final predictions. The surge in performance has often been achieved through this non-linear complexity, turning such systems into black box approaches and causing uncertainty in the way they operate and, ultimately, the way they come to decisions (Linardatos et al., 2020). In general, while scientists know how individual artificial neurons and layers in a deep neural network function in a mathematical sense, it can be difficult to fully understand how they work together to produce a specific output. Each neuron takes in a set of inputs and performs a mathematical operation to produce an output, which is then passed on to the next layer of neurons (Nielsen, 2018, p.3). However, as the number of neurons and layers increases it becomes difficult to fully interpret them, as the relationship between the input and outputs might cause a high level of complexity.

3.3 Loss functions and evaluation metrices for the models

A loss function is a crucial component of training a neural network. The loss function measures the difference between the network's predicted output and the true target output. The goal of training and validating is to minimize the loss function. In other words, the loss function acts as a feedback mechanism for the network, allowing it to learn from its mistakes and make better predictions over time. The loss function gives crucial information regarding how well the neural network performs at its given task (Chollet, 2018, p.60). There are many different loss functions that can be used depending on the type of problem being solved and the type of network being used. For example, mean squared error is a commonly used loss function for regression problems, while cross-entropy is often used for classification problems (Chollet, 2018, p.60). The choice of loss function will affect the network's training process, as well as the final predictions that the network produces. Since the problem at hand is a regression problem, it makes sense to look at two commonly used loss functions for these types of tasks, MSE and RMSE.

$$MSE = \sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}$$

Mean Squared Error (MSE) evaluates the model performance based on how much the prediction ($\hat{y}$) deviates from the actual observation (y), divided by total observations (n). Squaring the difference penalizes larger errors more harshly than smaller ones, which is desirable in this case. While MSE is a useful loss function, it can be challenging to interpret as an evaluation metric since its value is not in the same units as the value being predicted. This is however easily solved by introducing Root Mean Square Error (RMSE), which is simply the square root of the MSE, as this produces a metric that is measured in the same unit as the unit we want to predict (Singh, 2022). Since the RMSE value is given in the predicted unit, there is no universal threshold for what constitutes an acceptable RMSE value, as the value's interpretation is highly context dependent.

$$RMSE = \sqrt{\sum_{i=1}^{n} \frac{(y_i - \hat{y}_i)^2}{n}}$$

The evaluation of the ARMA-model in this thesis will utilize RMSE as a metric to enable a comparison between statistical approaches and neural networks. However, RMSE will not be the sole determinant in selecting the most appropriate statistical model. Akaike's Information

Criterion (AIC) (Akaike, 1974) is a frequently used measure for model selection in statistical modelling, including ARMA-based models. AIC is a relative measure that compares the quality of fit of different statistical models to the same dataset. Another selection metric is the Bayesian Information Criterium (BIC) (Stone, 1979). BIC measures the trade-off between complexity and model fit and penalizes the model for increased complexity. In general, a lower AIC or BIC relative to the other compared values indicates a better fit.

$$AIC = -2 * \ln(L) + 2 * k$$

$$BIC = -2 * \ln(L) + 2 * \ln(n) * k$$

Where L is the value of the likelihood, n is the number of recorded measurements, and k is the number of estimated parameters. Additionally, as an added evaluation metric of the model performances, the Mean Absolute Error (MAE) is also used for both the statistical and machine learning approaches. As with the RMSE, the error value units match the predicted target value units. The score is linear, and not weighted like the RMSE, which can be particularly helpful if the error distribution is not Gaussian (Chai & Draxler, 2014).

$$MAE = \sum_{i=1}^{n} \frac{|y_i - \hat{y}|}{n}$$

3.4 Mixed-integer programming

In order to make a production plan and illustrate the financial impact by applying the models' predicted power prices, an optimization problem is created. The most common associations when dealing with optimization and linear programming (LP) are limited resources and competing activities (Hiller & Lieberman, 2021, p. 32). The motivation behind LP is to calculate the optimal solution, based on different types of restrictions. The optimal quantity (object function) may be minimizing cost or maximizing profit. When applying integer solution and binary variables to the problem, the problem becomes a mixed-integer optimization problem (MIP).

When calculating a mixed-integer problem in Microsoft Excel, the software uses the branch-and-bound technique (Frontline Systems, n.d.). The principle behind the branch-and-bound technique is to divide and conquer (Hiller & Lieberman, 2021, pp. 478-481). Since mixed-integer problems can be difficult to solve, the algorithm divides the main problem into different smaller sub-problems with different values for the binary restriction. This step of the algorithm is called branching. The algorithm then calculates the optimal solutions for each subset and discards the subset if the solution is not clearly optimal, called bounding. The algorithm uses an LP relaxation, which relaxes, or deletes a set of constraints that could make the problem difficult to solve. At this point, a common approach for the algorithm is to a apply a common LP method called simplex. The simplex method starts at one of the points, determining whether the point is optimal. If the calculated point is not optimal, the algorithm continues along one of the restrictions to the next point, and the iteration continues until the optimal point is identified. For all the created subsets, fathoming tests are run on the set to determent whether the subset is being dismissed or kept running the algorithm again on new subsets. For further read on the algorithm, see Hiller and Liberman's "Introduction to Operations Research" from 2021.

4 Methodology

This chapter outlines the components of our research methodology. Firstly, we present the data, target and exogenous variables, along with key assumptions about the data. The chapter then proceeds to showcase the specific methods employed in finding and creating the best suited ARMA model. This is followed by a comprehensive explanation of the procedures used in training, validating, and testing the Seq2Seq-LSTM model, along with its machine learning baseline, a conventional LSTM. Lastly, the chapter presents the hydrogen production case, along with its mathematical principles and production assumptions.

4.1 The dataset, target- and exogenous variables

The dataset used in this thesis contains daily observations of the power price in Northern Norway and several meteorological exogenous variables from the region. The data is observations from the time period between January 1st 1999 and January 1st 2023. The dataset has been created by collecting data from online sources like NVE, Klimaservicesenter, and Montel.

The target variable for forecasting is the price area NO4, measured in mean daily øre/kWh, collected from Montel. Furthermore, the exogenous variables in the time series are the percentage of water levels in hydroelectric reservoirs (NVE), temperature, wind and rainfall (Klimaservicessenter). Hydroelectric reservoirs and rainfall are included to give information on hydroelectric capacity. 90% of the Norwegian power in a normal year is produced by hydroelectric power (Energifakta Norge, 2022). The reason for including wind is due to the increasing wind power development in the Northern parts of the country. According to Nordkraft AS (Tore Schjelderup, personal communication, January 2023), the temperature influences the power usage, which influences the power price. Other variables also influence the power price in NO4, such as the power price in the northern part of Sweden and the transfer capacity between the two regions. However, incorporation of these variables into the models used in this thesis proves challenging, as these variables emerge concurrently with the target variable.

All the variables used in the time series were obtainable with daily observations, except for the amount of water in the hydroelectric reservoirs. Information regarding this variable was only available to contract with weekly observations, therefore, this thesis has used one week's observation for all the days in the same week. Moreover, there were no missing values in the observations for the price variable. Because the time series only contained a few missing observations in the *meteorological variables,* approximately 10 missing observations in each accounting for about 0.1% of the time series, the missing variables have been set to 0.

4.1.1 Key assumptions

This thesis has made some assumptions about the collected data. We assume that the reported power price in NO4 as well as the meteorological variables are correct. The power prices are not corrected for inflation, are without VAT and electricity tax. Other macroeconomic assumptions have not been made in the dataset. Because this thesis has collected data from different sources, we assume further that all the observations are in fact collected from the same day.

4.2 Determining the number of time steps to forecast

For the forecasting models to be practically useful in optimizing and planning hydrogen production, they must be capable of accurately predicting beyond just one time step. The accuracy of the models' predictions generally decreases the further into the future they are programmed to forecast, unless there is a significant issue with the models' interpretation of seasonal patterns. To strike a balance between accuracy, the number of forecasted time steps, and its practical use, a trade-off must be considered. In addition to this, forecasting too far into the future might be problematic, as the models use weather data as exogenous variables. Weather forecasts are typically 90% accurate when predicting five days ahead, but accuracies of weather forecasts beyond this point decreases substantially. For instance, weather forecasts for 10 days are typically only accurate half of the time (National Oceanic and Atmospheric Administration, n.d.).

An (S)ARMAX model requires a continuous supply of exogenous variables to generate power price forecasts for each time step. However, LSTM networks can be trained to predict both the exogenous variables and the target variable, thereby overcoming this limitation. To ensure a fair comparison between the two models, we will evaluate their performance in predicting five timesteps ahead. A forecast of five days ahead would be sufficient for planning hydrogen production for one work week while still being based on accurate weather forecasts. However, we will also view the LSTMs general ability to forecast beyond five timesteps to test its flexibility, as potentially accurate further ahead forecast would allow for long-term planning, which leaves more room for the MIP-algorithm to shift production to the least costly days.

4.3 Creating the ARMA model

The programming language Python has been used for the creation of the ARMA models. The *autoarima* function from the *pmdarima* packages was used in order to help navigate the correct order of the AR(p) and MA(q). The target of the function is to minimize the AIC value to achieve the order of a model with the best performance. The models were fitted on 70% of the time series and tested on the remaining 30%. The AIC values as shown in Appendix A03, are calculated on the first part of the time series.

4.3.1 Stationarity and heteroskedasticity

The ARMA model presupposes stationarity in the variables. A common test for stationarity in the variables is the Augmented Dickey-Fuller test. The test was conducted on both the target variable and all the potential exogenous variables, where all the variables turn out to be stationary (see Appendix A01). Hence, no differencing of the variables was necessary to run the model.

The models were tested for heteroskedasticity using a Breusch-Pagan test which showed significant heteroscedasticity in the models (see Appendix A02). Hence, the GARCH model as presented in Chapter 3.1.3 was applied to counteract this issue. To find the order of GARCH(r,m) which achieved the lowest AIC value, a *for loop* in Python was used where *r* and the *m* order of the model was restricted to take any value between 1 and 10.

4.3.2 ARMAX-GARCH with exogenous variables

A rule of thumb when creating models is to keep the model as simple as possible. Hence, an ARMA model without exogenous variables was the first model created. In order to decide which exogenous variables to include in the model, the ARMA was tested with different combinations of the variables available in the time series. To capture any potential seasonal trends in the time series, a SARMA model was also fitted the same way as the ARMA model.

An issue when forecasting with (S)ARMAX is that the models need the future values of the exogenous variables to make predictions on the price as described in Chapter 3.1.2. Hence, the time series needed to be updated every week to generate predictions for the upcoming week. Due to our assumption that a weather forecast five timesteps ahead is comparable with the actual weather, the model was given these variables as exogenous variables.

4.3.3 Applying sliding window

In order to evaluate the models' performance, a sliding window approach was conducted on the models as described in Chapter 3.1.4. The approach is illustrated in the following chart.

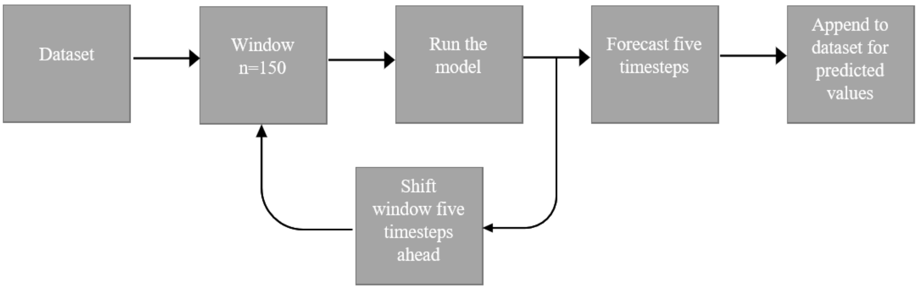


Figure 7 - The sliding window approach

The architecture achieving the lowest RMSE was with a window size of 150, meaning the algorithm uses the last 150 observations from the time series to predict the next five timesteps from the ARMAX-GARCH and SARMAX-GARCH models. This algorithm was applied on the last 30% of the time series, meaning the model was refitted every loop due to changes in the training time series, which means that the model adapted to the changes over time.

The ARMAX- and the SARMAX-GARCH models were ran in Python on a computer with an Intel(R) i5 CPU. The duration of running the model-script with the sliding window approach was about 10 minutes. Running the ARMAX, ARMAX-GARCH and SARMAX-GARCH models without the sliding window had a calculation time on approximately 1-3 minutes.

4.3.4 Interpretation of the ARMAX-GARCH model

The ARMAX-GARCH model with order (2,2) (3,7) can be interpreted as followed: The first number describes the order of the autoregressive model (AR). Because the order of this thesis' model is 2, the model uses the two previous observations from the time series to predict a new observation. The model weights the previous observation depending on the coefficients in the ARMAX model. The coefficients are the effect of the previous observations on the current observation. These weights determine the effects of previous observations on the current observation. Because the previous observations are calculated the same way as the current, all the observations in the time series will affect the current observation with decreasing weights.

The moving average (MA) has the same order as the AR model, order 2, which determent the number of previous observations the model uses. Unlike the AR model, the MA model observes the error between the estimated and actual observations and models a linear combination of the observed errors. Because the MA model is an order of 2, the errors from the two previous observations are multiplied by the coefficients, creating a linear combination which results in the prediction of the current variable. The exogenous variables impact the prediction of the target variable. The coefficients of the exogenous variables indicate the effect on the target variable. When combining the AR, MA and exogenous variables, the model predicts the power price based on the past values from the AR, the part errors from the MA, and the effects of the exogenous variables. The reason for combining the GARCH model with the ARMAX is due to the significant heteroskedasticity in the ARMAX model. The seasonal ARMAX-GARCH can be interpret the same way with one additional component. The SARMAX also take an observation *m*, or in this thesis' case seven timesteps back, as a lagged variable which affects the seasonal component in the model. The interpretation of the models is the same for all the windows in the sliding window approach.

4.4 Creating the LSTM models

For creating the LSTM models, *TensorFlow* and *Keras* packages were used. These packages are one of the most common deep learning packages for programming languages such as Python and R. The two LSTM networks in this thesis was created with the R programming language.

4.4.1 Hold out validation – training, validating and testing the models

Since hyperparameter tuning and model selection was needed to map out the best LSTM model, we chose to split all available data into a 50-20-30 split, with 50% used for training, 20% for validation and 30% for test. This is not a very common approach to splitting the dataset, but the choice is based around wanting a large test set. The choice to use a larger test set was made in order to evaluate the model's performance on different statistical distributions. Since the dataset had areas with varying levels of volatility. The chosen approach involved training on 50% of the dataset and performing hyperparameter tuning and model evaluation on the 20% validation set. It is crucial to evaluate the performance of models on a designated validation set during the parameter tuning process, instead of relying on the test set. Tuning models based on the feedback from the test set can lead to overfitting, as the models may be too specialized to the test set and lose their ability to generalize to new, unseen data (Toisoul, 2020).

Once the models with the lowest RMSE score on the validation set was identified, the models were retrained by adding the validation set to the training set, resulting in a 70/30 training/test split. Adding the validation set to the training data for retraining allowed for the use of more data for training but keeping the test unseen to the models, an approach like this is often desired as it allows for more training (Toisoul, 2020). When they were done training, the models were finally evaluated on the unbiased test set.

4.4.2 Hyperparameter tuning and model selection – Conventional LSTM

As discussed in chapter 3.2.1 due to the vanishing gradient problem, an LSTM was preferred over a standard RNN. The complexity of the LSTM model might improve forecasting ability, and it has proven to work quite well on numerous time series regression problems, including power price forecasting in previous literature. As previously discussed, the performance of neural networks greatly depends on the architecture of the model. Hyperparameters, tuning and architecture were inspired by a combination of multiple sources, including Chollet (2018), Sutskever et al. (2014), Nielsen (2018) and Olah (2015). In addition to that, previous literature such as Aranguren et al. (2022) and Gong et.al (2019) have also been a great inspiration.

To map which architecture performed the best, a shallow standard LSTM network was created. The performance of the shallow network was not satisfying, however, it served as a foundation for building a more complex model. Layers and the number of neurons were then increased and validated on a validation set, in order to choose a model with the lowest RMSE score. As the model increased in complexity, the network showed signs of overfitting to the training data, as the generalization to the validation set was poor. In order to reduce the chance of overfitting, different regularization techniques were implemented. As discussed in chapter 3.2.6, overfitting can easily occur. Experimenting with different regularization techniques found that adding LASSO regularization (L1) on the first hidden layer was beneficial, with a penalty value of 0.01. A penalty of 0.01 for the regularization parameter means that the penalty applied to the weights is relatively small, but in this case, still significant enough to encourage the model to learn sparse weights and prevent overfitting. Moreover, the model responded better to L1 regularization than L2. Furthermore, dropout layers were added between the second and third hidden layer with a value of 0.2. The regularization technique drops out (setting to zero) 20% of the neurons in a layer during training, and it forces the model to learn multiple independent representations of the same data and helps the model from relying too heavily on any one feature or set of features. This, in addition to the reduction of the model's overall complexity and the careful surveillance of the model's loss function to the number of epochs in training, helped remove any issues with overfitting. Experimentation showed that a batch size of two with number of epochs defined to 50 gave decent results without causing overfitting.

The activation function chosen for the LSTM layers is tanh, however, the dense layers in the model (fully connected layers) are activated using ReLU. During experimenting, it was found that the combination of these activation functions was optimal. The model seemed to especially benefit from the ReLU activation functions in the dense layers. The standard default activation settings for LSTM layers in Keras is the tanh function, and it was found no improvements in altering the default setting to ReLU. However, based on literature, a combination of activation functions is common. The optimizer used for the model was Adam. For further explanation, see chapter 3.2.4.

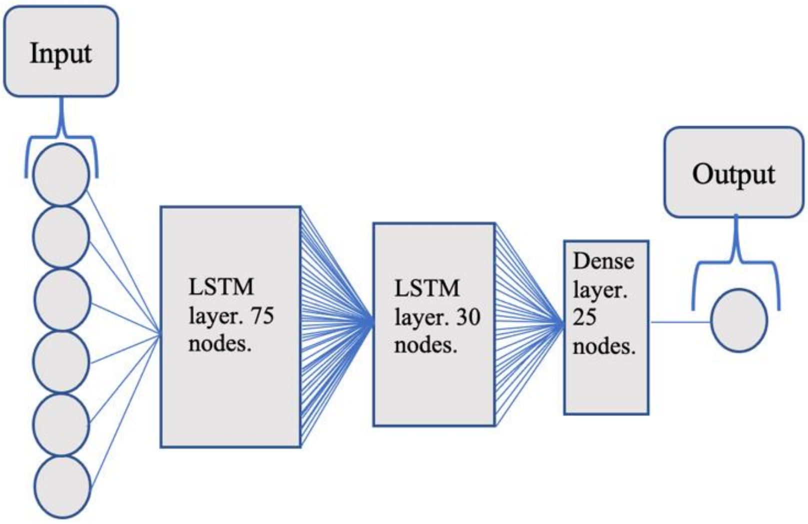


Figure 8 - Conventional LSTM topology

4.4.3 Applying sequence to sequence

So far, the model created is an LSTM model without a sequence-to-sequence architecture. The created model as shown in figure 8 served as a baseline for power price forecasting using neural networks in this thesis, and the structure is similar to what typically have been used for power price prediction using recurrent neural networks in previous literature. A sequence to sequence structure was used by Gong et al. (2019) when forecasting power load, and the method was found by them to be superior compared to LSTM models without the structure.

To create a sequence to sequence architecture, the structure and code needed to be heavily altered. Firstly, we created a structure that consists of two LSTMs: an encoder and a decoder LSTM, connected in a sequence. The encoder LSTM takes in the input sequence and generates a context vector that summarizes the input sequence. The second LSTM, the decoder, is then programmed to take the context vector and generate the output sequence. The output sequence is a sequence of desired timesteps. This is a methodically substantial difference from the conventional LSTM created in the first instance. The standard LSTM (figure 8) is used for single sequence prediction, where the model takes in a sequence of inputs and generates a single output until it has reached the desired timesteps. Despite the structural changes, it was found reasonable to keep the regularization techniques the same to prevent overfitting, as well as the activation functions. However, the overall topology of the network was changed.

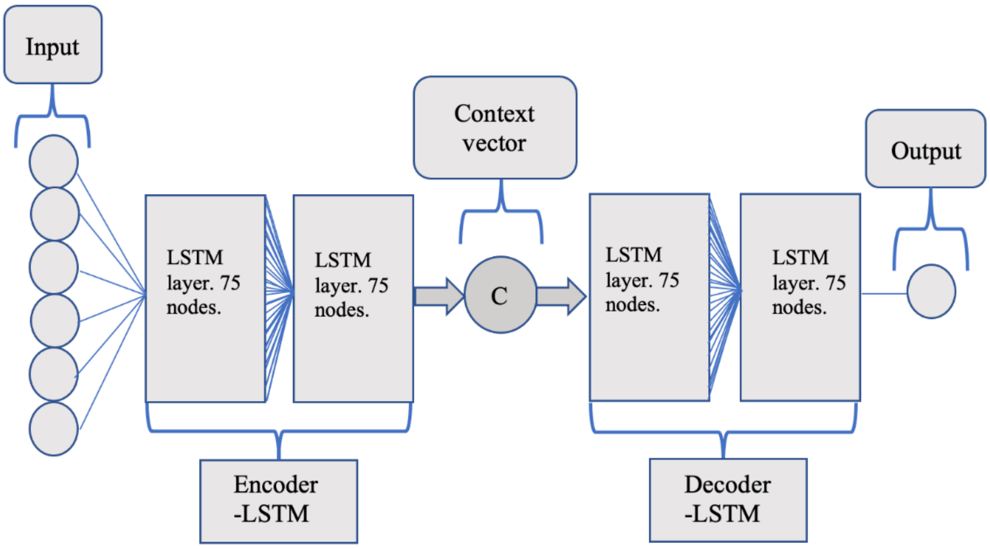


Figure 9 - Seq2Seq-LSTM topology

For the Seq2Seq-LSTM model, experimentation showed that the best results were achieved with a network structure that consists of two hidden layers in each network. This adds up to four hidden layers in total with the number of nodes consistently set to 75 for each hidden layer. It is also worth noting that in the architecture of the standard LSTM, it was discovered that a combination of LSTM layers and dense layers was appropriate, whereas with the Seq2Seq-LSTM, it was decided to only use LSTM layers.

In order to map out the best combination of window size and forecasted time step, different combinations were tried. Firstly, the number of time steps was set to 5, 14 and 30, and different window sizes between 30 and 150 were also tried. The complexity of the model was shown to be heavily influenced by the number of observations used for the window size, which did noticeably impact the duration of the training process. It was found no noticeable upside using a larger window size than 50, thus setting the window size to 50.

After discovering the optimal architecture and hyperparameters as described, the model was selected based on its performance on the validation set. Subsequently, the model was retrained with an increased training set using the selected architecture and hyperparameters, and finally evaluated on the test set. The same approach as with the conventional LSTM.

Training duration was only about 35 minutes for the Seq2Seq-LSTM and about one hour for the conventional LSTM. The models were trained on a computer with an intel core i5 CPU. The training time for these models is not necessarily very time-consuming in the world of deep learning, however, due to the amount of experimenting with different architectural structures and hyperparameter tuning, it can lead to a very time-consuming project. Therefore, due to the time constraints of the thesis, it is possible that there exist other model combinations and architectural structures that could potentially perform better than the ones discovered by us.

4.4.4 Interpretation of the LSTM models

During training, the Seq2Seq-LSTM was trained to forecast the desired number of timesteps ahead based on the specifications given in the encoder/decoder structure of the model. The model forecasts the next timesteps of the target and exogenous variables by taking in the historical data of all the variables, processes them through the encoder and decoder networks of the model, and outputs the predicted values for the next timesteps. This process is repeated for each new time step.

One of the key differences between the two LSTM models presented above, is how the two models are programmed to work based on their architecture. In the conventional LSTM, the model predicts one future value at a time, and the predicted value is then fed back into the model as input for the next prediction. In theory, this can lead to accumulated errors and a lower accuracy as the number of predicted time step increases to the desired number. In contrast, the Seq2Seq-LSTM is explicitly designed to handle multiple future predictions by generating the entire output sequence at once based on the input sequence and the learned weights. As discussed in Chapter 3.2.3 however, this have mostly been a success for language models, and the amount of Seq2Seq-modelling for time series forecasting is not as widespread as regular LSTMs, despite being quite successful for Aranguren et al. (2022) and Gong et.al (2019).

During the testing phase of the Seq2Seq-LSTM, the model was programmed to generate the desired number of predictions using the learned weights from the data, and a fixed window size with available information up to the point of prediction and compares them to the actual values in the test set. This process was repeated until all data in the test set had been used. This type of forecasting is quite similar to the sliding window method used in the (S)ARMAX-GARCH. The (S)ARMAX-GARCH forecasts uses a fixed size window to forecast the next timesteps, whereas the Seq2Seq-LSTM model uses the entire available learned data and a fixed window to forecast the next timesteps. The similarity, however, revolves around the model only forecasting a fixed sliding time step multiple times, instead of forecasting the entire test set in one go. Additionally, both LSTM models have also been programmed to forecast the exogenous variables for timesteps above five days.

4.5 Hydrogen production case

In this chapter, we have created a cost-minimization case to analyze the financial value of the forecasting models. The models have been programmed to forecast five timesteps (days) ahead, to provide input for cost minimization in hydrogen production for one work week. In this case, the models are used to forecast on five days interval, Monday to Friday, for eight weeks, starting from Monday January 2nd 2023. The data-inputs used for this case are data points beyond what the models were trained, validated and tested for. The outputs of the models were then used as inputs for the production period in the optimization problem explained below. The creation of the case is done in collaboration with a Norwegian hydrogen production corporation (anonymous, personal communication, 2023).

The corporation owns a factory with a power capacity of 100 MWh (Megawatt hour) which makes the total production capacity 2400 MWh throughout one day of production. This corporation operates an electrolyser for hydrogen production, which has a stack efficiency of 60%. This means the electrolyser can effectively convert 60% of supplied power into a measurable quantity of hydrogen (metric tons), quantified in MWh (units of power-equivalent hydrogen). To convert MWh of hydrogen into tons, the MWh is divided by the constant 33.3 which gives the produced quantity. The factory needs to produce 170 tons of hydrogen each week to reach the corporation's production target. Due to the calculation described above, the factory has a maximum capacity for production each day of 43.24 tons.

	Hour	Day
Megawatt hour	100.00	2400.00
Stack efficiency (%)	0.60	0.60
MWh available for production	60.00	1440.00
MWh hydrogen conversion factor	33.30	33.30
Hydrogen in tons	**1.80**	**43.24**

Figure 10 - Power to hydrogen conversion

The transportation of produced hydrogen from the plant incurs a cost. However, this expense only applies on the days when production occurs. This cost is added to the optimization problem to illustrate the impact of a hypothetical storage cost or a transportation cost. Since the hydrogen production case illustrates issues with variable production costs, and to create a better understanding of the power price influence on hydrogen production, fixed costs are deemed irrelevant. The optimization problem has the following mathematical expression.

$$MIN\ \Sigma(D_i * T_i) + \Sigma(PP_i * PQ_i) \qquad i \in 1{,}5$$

Subject to

$$PQ_i \leq 43.24$$

$$\Sigma PQ_i \geq 170$$

$$T_i = 150000$$

$$D_i = BIN$$

$$D_i = Dummy\ if\ transportation\ on\ day\ i$$

$$T_i = The\ transportation\ cost\ on\ day\ i$$

$$PP_i = The\ power\ price\ on\ day\ i$$

$$PQ_i = The\ production\ quantity\ on\ day\ i$$

The solver function in Excel which has a MIP calculation function, is used to minimize the cost for the corporation. Because this thesis is forecasting the power price with two different models, the optimization was done 16 times, every week in the production period for both models. A maximization and a minimization with respect to the actual power price was also conducted to analyse a cost-feasible area if the models missed on the weekly trends in the power price. This is conducted in chapter 5.3.3.

The total costs when optimizing with the predicted power price will only be accurate if the models manage to predict the actual power price with 100% accuracy. It is reasonable to think that this might not actually occur. Hence, two different scenarios were created to illustrate the different financial impacts the models may have.

Scenario 1 – The baseline

Given no knowledge of future power prices, a reasonable assumption revolves around the corporation producing the same quantity of hydrogen each day with no respect to the actual power price. Hence, this scenario accounts for an average production quantity each day and the actual power price for each week. The scenario can be seen as a baseline model of what the cost will be if the company does not actively guide the production with the use of power predictive models.

Scenario 2 – The actual cost

The second scenario presupposes that the company produces according to the ARMA model or the LSTM model. The scenario contains actual observations of the power price for each week and will illustrate the actual cost each week when producing according to this thesis' models. *Scenario 2* must outperform *Scenario 1* for the models to be useful cost saving tools.

The same optimization has been created for all eight weeks in January and February 2023, hereby called the production period. Both the predicted power prices from the ARMA model and the LSTM have been applied in the optimization. The technical calculations for the eight weeks for both models are visualized in Appendix A05.

5 Results

This chapter is structured to present our research results, divided into two main areas. The initial sections disclose the outcomes of the forecasting models on the test set. The chapter begins with the results from the ARMAX-GARCH and SARMAX-GARCH models, before moving on to reveal the findings from the baseline LSTM and Seq2Seq-LSTM. The final section delves into the models' forecasting performance during the production period and the optimization results, with particular emphasis on the financial impacts of the models. Lastly, a risk analysis of the optimization results was conducted.

5.1 ARMAX and SARMAX model performance

To determine the optimal model for the time series, the model with the lowest AIC was used in the forecast conducted in this thesis. The two models achieving the lowest AIC values were ARMAX-GARCH and SARMAX-GARCH models as shown in Appendix A03. Applying all the exogenous variables from the time series, *fyllingsgrad, temperatur, vind* and *nedbør,* both models decreased the AIC values and increase the models' fit to the time series. As shown in the appendix, the ARMAX-GARCH model managed to get an AIC of 31 751 and a BIC of 31 831. The SARMAX-GARCH model was tested with different seasonal intervals but achieved the best AIC and BIC values with seven as the seasonal component. The achieved AIC and BIC values were 30 301 and 30 375. The order of these models are for the ARMAX-GARCH(2,2)(3,7) and for the SARMAX-GARCH(2,2)(1,0,1,7)(4,6). Because all the variables in the time series were stationary, no differencing were needed, hence the models are ARMA, not ARIMA models.

Model	AIC	BIC	RMSE	MAE
ARMAX-GARCH(2,2)(3,7)	31 751	31 831	14.56	6.46
SARMAX-GARCH(2,2)(1,0,1,7)(4,6)	30 301	30 375	15.00	6.41

Figure 11 - AIC and BIC values of the ARMA models

Both models were evaluated using the sliding window approach as mentioned previously. The ARMAX-GARCH model, when forecasting five timesteps ahead, achieved an RMSE of 14.56 øre/kWh and an MAE of 6.46 øre/kWh. With the same evaluation criteria, the SARMAX-GARCH achieved an RMSE of 15.00 and an MAE of 6.41. The plot of the models' predictions and the actual observations show that the models also were capable of predicting some trend in the power price. The following graph shows a five-timestep prediction of the ARMAX-

GARCH and the SARMAX-GARCH. The actual power price is colored with blue, and the predicted power price is colored with red.

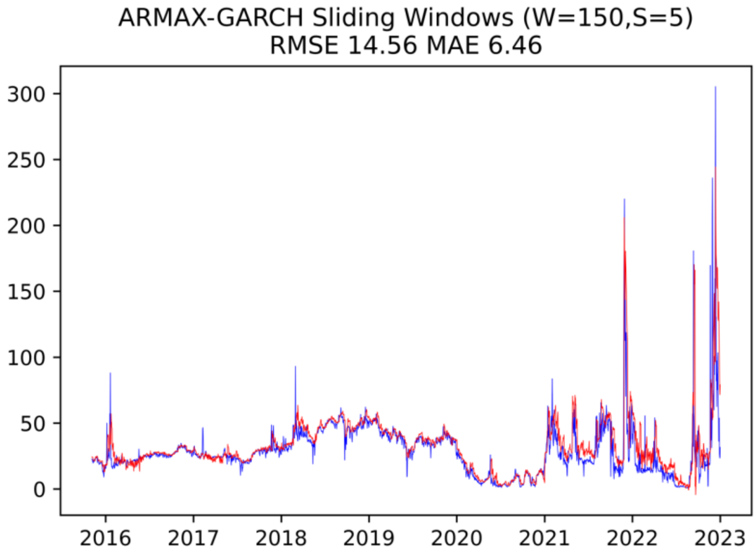


Figure 12 - ARMAX-GARCH with sliding windows, test results

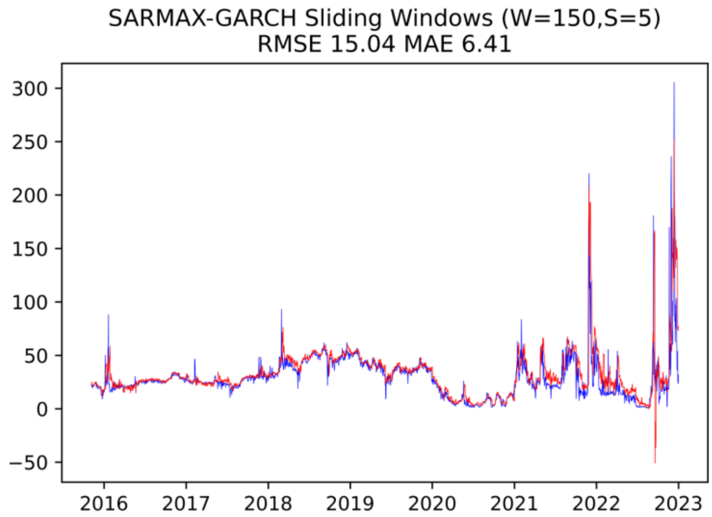


Figure 13 - SARMAX-GARCH with sliding windows, test results

The ARMAX-GARCH and SARMAX-GARCH models achieved almost the same results during the test period. The Appendix A03 also shows small differences between the ARMA models when comparing the AIC and BIC from the models. Because the ARMAX-GARCH model is a less complex model and achieved a slightly better RMSE value than the SARMAX-GARCH, the ARMAX was used in the optimization problem. Further experimenting with the seasonal component up to a year might increase the SARMAX-GARCH performance and lowering the RMSE/MAE. To gain a better understanding of how well the ARMAX-GARCH model will perform when predicting the power price during this thesis' production period, the

RMSE and MAE were calculated for the same predictions as over, but for the end of the time series (from 2022).

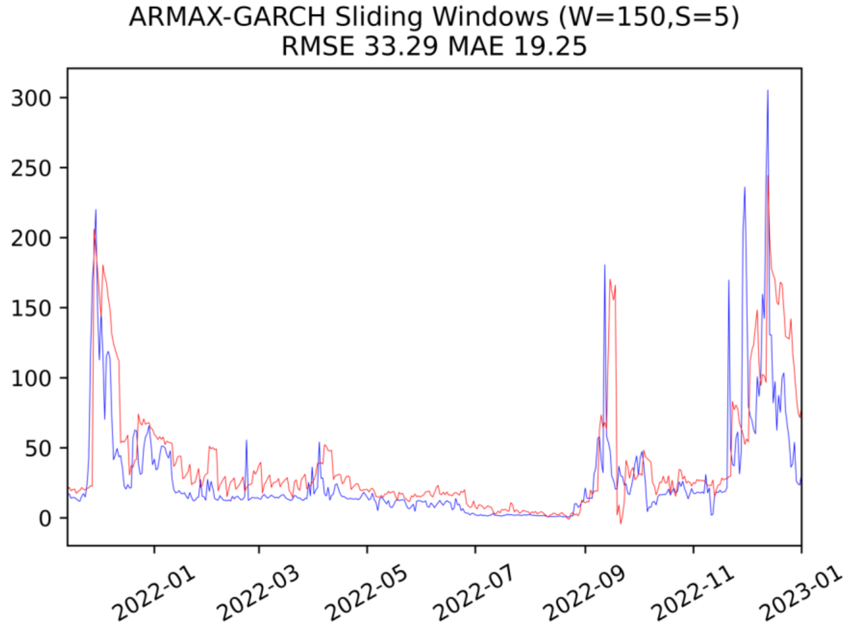


Figure 14 - ARMAX-GARCH with sliding windows from 2022, test results

Due to high volatility in the power price at the end of the time series, the model struggles to predict the power price in this period accurately. In the last two years of the time series, the model has an RMSE of 33.29 øre/kWh and an MAE of 19.25 øre/kWh, which is substantially higher than the RMSE for the entire period. The model seems to be disturbed by the sudden spikes, and as a result it leads to delays in the predicted prices.

5.2 The LSTM model performances

In line with the methodology described in the previous chapter, two LSTM models were trained, validated and tested. The results in this chapter give an overview of the models' performance on the test set. Three different timesteps were tested, and each of the two models were trained to forecast five, fourteen and thirty timesteps ahead.

5.2.1 Conventional LSTM model performance

As described in the methodology, a conventional LSTM model was created to establish a baseline for the machine learning models. We experimented with different timestep values, and it was discovered that the best practical result was achieved with a timestep of five. The results showed an RMSE of 13.75 øre/kWh and an MAE of 6.27 øre/kWh, indicating the degree of the model's forecasted values deviates from the actual values.

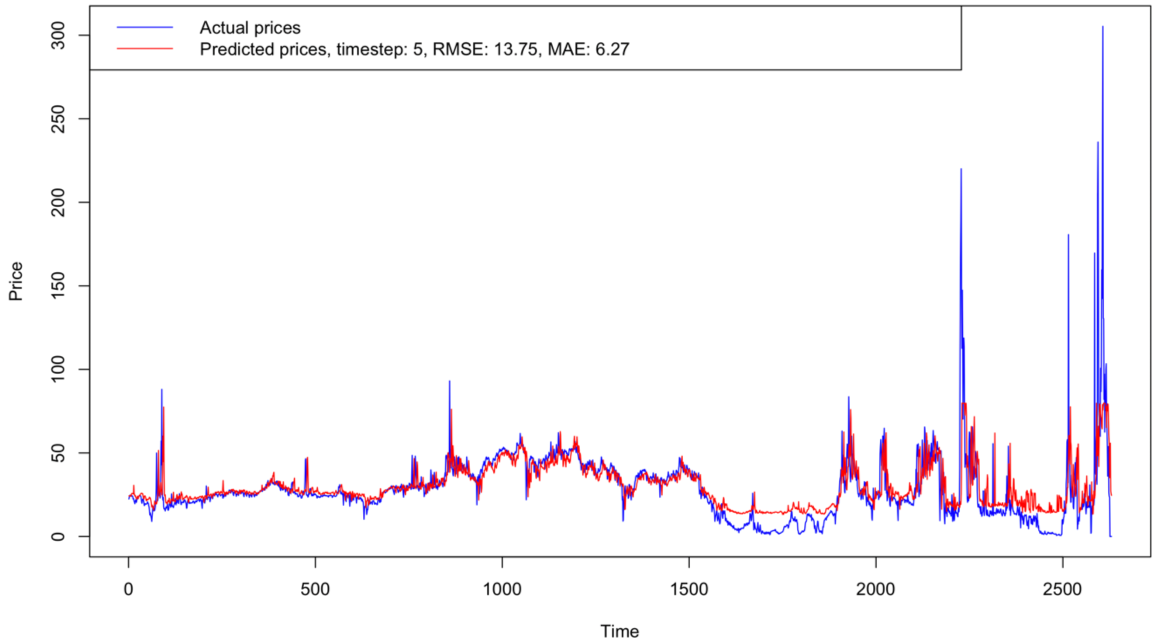


Figure 15 - Five-day forecast with conventional LSTM, test results

The conventional LSTM was able to capture the underlying trend of power prices, resulting in reasonably accurate predictions during the initial stages of the test set. However, as the test set progressed and the periods became more volatile, the model's accuracy decreased, particularly when predicting extreme values. The model struggled to capture the highest and lowest prices in the data, with the highest price recorded being 305.38 øre/kWh, while the highest price captured by the model was only 95 øre/kWh. Similarly, the lowest price recorded was 0.7 øre/kWh, while the lowest price predicted by the model was 10.10 øre/kWh. These difficulties are illustrated in figure 15, which depicts the challenges of predicting extreme values. Value within the range of 10.10 and 95 are however well depicted by the model, resulting in reasonably accurate forecasts.

Due to the LSTMs ability to forecast its own exogenous variables, the LSTM network was also tested on further ahead timesteps. The further into the future the model was programmed to predict, the lower the accuracy. In figure 16, the different timesteps and evaluation metrics are shown. This indicates that the model is not able to accurately predict power prices at longer timesteps with the architecture chosen for the conventional LSTM in this thesis, with especially high RMSE and MAE values.

Standard LSTM		
Timestep	RMSE	MAE
5	13.75	6.27
14	19.88	11.83
30	25.12	14.67

Figure 16 - Different timesteps conventional LSTM

5.2.2 Seq2Seq-LSTM model performance

When forecasting with five timesteps ahead as the baseline, the Seq2Seq-LSTM yielded an RMSE of 6.43 øre/kWh, almost half of the RMSE obtained by the conventional LSTM model. This indicated significantly reduced degree of deviation from the actual values. Additionally, the model displayed an MAE value of 2.1. Specifically, this suggests that, on average the model only missed each prediction by 2.1 øre/kWh, highlighting high accuracy in forecasting the power price five days ahead.

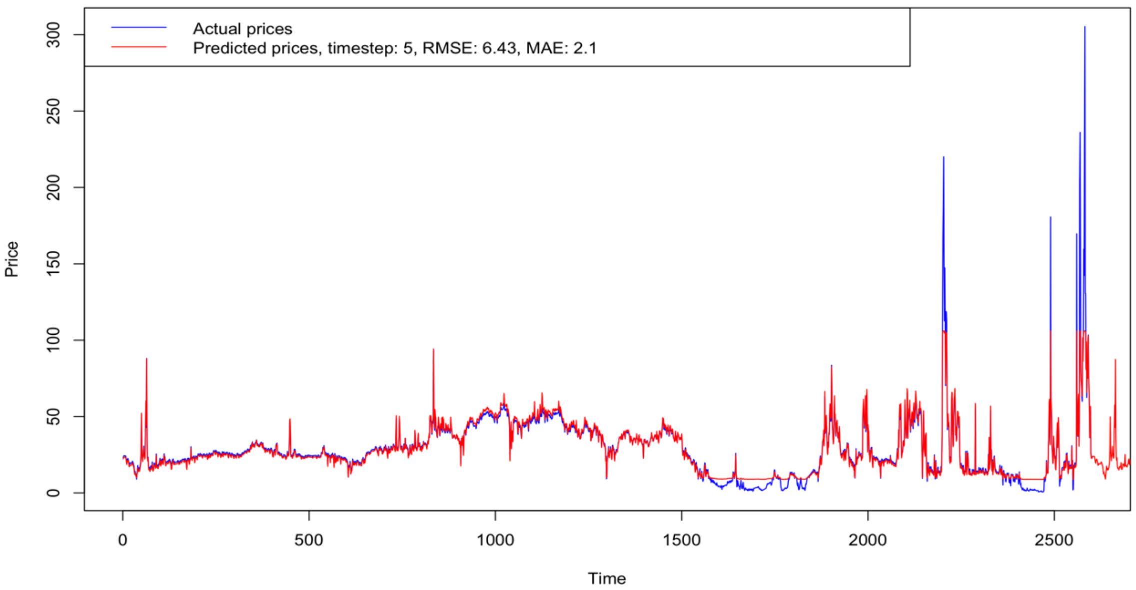


Figure 17 - Five-day forecast with Seq2Seq-LSTM, test results

The Seq2Seq-LSTM model demonstrated satisfactory performance in predicting the power prices, accurately capturing the underlying trend with high precision. In comparison to the conventional LSTM, the Seq2Seq-LSTM model demonstrated better performance in handling extreme values, being able to predict prices as low as 3.54 øre/kWh and as high as 121.97 øre/kWh. Despite its ability to handle extreme value better than the conventional LSTM, the model still struggled to accurately forecast the extremes. The model's limitations in predicting values outside of this range indicate that it may not fully account for the underlying factors contributing to extreme values. This is not too surprising, given that the model was trained and validated on a less volatile portion of the dataset that included fewer extreme values than the test portion.

Due to the relatively high accuracy of the Seq2Seq-LSTM model's performance when forecasting power prices five days ahead, we also tested the model's ability to predict prices for a timestep of 14. This could potentially lead to further ahead planning for the optimization problem.

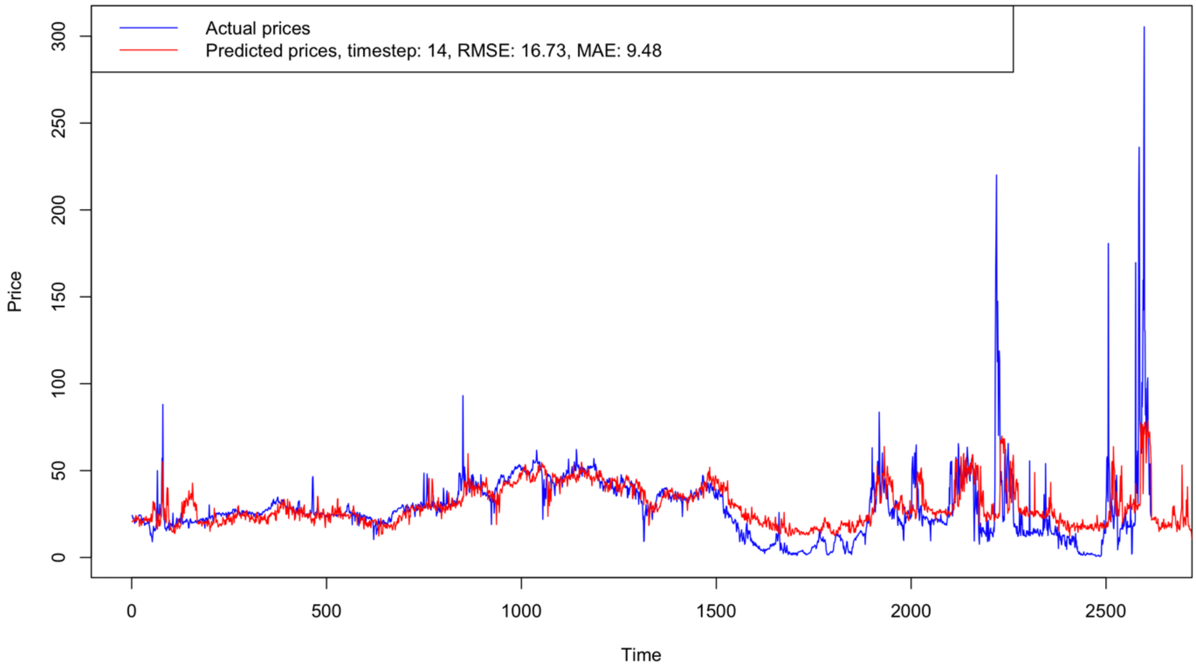


Figure 18 - Fourteen-day forecast with Seq2Seq-LSTM, test results

We found that the model had the ability to accurately capture a considerable portion of the underlying trend within the data, although the overall level of accuracy was not particularly high. While the RMSE of 16.73 and MAE of 9.48 were slightly better than the conventional LSTM with a timestep of 14, the results were not deemed satisfactory, as the model operates with too much uncertainty. However, it is worth noting that the model's relatively accurate depiction of the underlying trend still holds some promise for practical applications.

As described in the methodology, we also tested for the possibility of forecasting 30 days ahead. Given the inaccurate results of forecasting 14 days ahead, satisfying results were not expected, but the experimenting with different timesteps could potentially provide valuable information about the model's overall ability and limitations. As shown in the figure below, the results of the 30-day prediction were unsatisfactory. The model struggled to accurately capture the underlying trend, although some tendencies were apparent. The overall RMSE and MAE scores were high, indicating that the model's practical value was limited for this level of forecasting horizon. Despite these limitations, it is important to acknowledge that this specific Seq2Seq-LSTM model may not be the optimal architecture for long-term forecasting. Other model architectures or techniques may be more effective in accurately predicting the NO4 power price for extended periods.

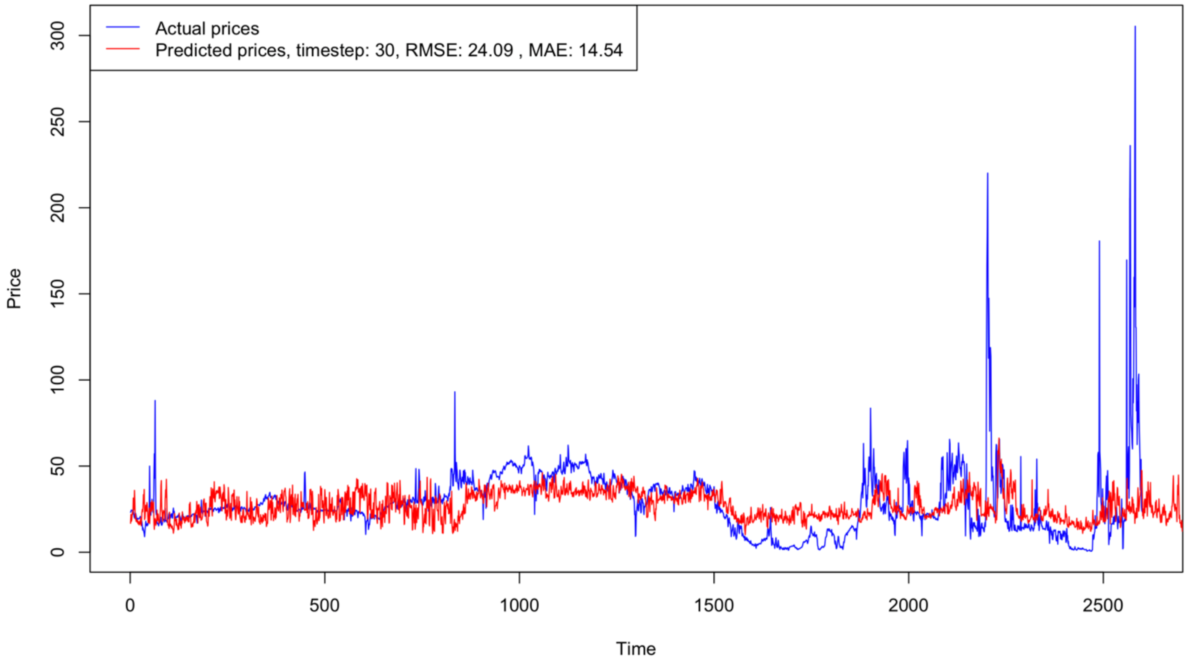


Figure 19 - Thirty-day forecast with Seq2Seq-LSTM, test results

Seq2Seq-LSTM		
Timestep	RMSE	MAE
5	6.43	2.1
14	16.73	9.48
30	24.09	14.54

Figure 20 - Different timesteps with Seq2Seq-LSTM

The results above demonstrate that the Seq2Seq-LSTM model exhibits superior forecasting performance when predicting power prices with a timestep of five, in comparison to the conventional LSTM model. While the Seq2Seq-LSTM model maintains a slightly better performance as the timestep increases, this improvement is not linear.

The results of the model performance evaluation emphasize the importance of selecting an appropriate number of timesteps ahead for the specific forecasting task. While the Seq2Seq-LSTM model performed quite well when forecasting five timesteps ahead, it struggled with further ahead forecasts. Although its accuracy was better than the conventional LSTM, this was especially true when forecasting five timesteps ahead. The visualizations and error metrics indicates that as the number of timesteps increases, the uncertainty also increases, and the model predictions tend to fluctuate more, and trends may be missed. This suggests that the model struggles to capture the nuances of the power price data when forecasting further ahead, and additional factors beyond the inputs used in the model may be needed to improve its performance. This is unsurprising, but it highlights the need to consider the specific purpose and constraints for the application when selecting an appropriate number of timesteps for the forecasting task.

For instance, when using the Seq2Seq-LSTM model for forecasting power prices for hydrogen production optimization, a timestep long enough to serve the practical purpose of planning the production was needed. Therefore, a trade-off between predictability and practicality was evaluated, thus choosing the model that can forecast five timesteps ahead with an RMSE of 6.43 and a MAE of 2.1 for further analysis. This allows the model to predict a workweek interval before being updated and run to forecast again every Sunday with updated data. The findings suggest that the Seq2Seq-LSTM model with a timestep of five is particularly effective at capturing the underlying trend within the data, resulting in higher accuracy. However, it is essential to note that the decision to choose this model may not be optimal for all applications, and it is crucial to consider the specific needs and constraints of each individual case. This also puts it in easy comparison to the ARMAX-GARCH model.

5.3 Model-assisted hydrogen production

5.3.1 Comparing the power price predictions during the production period

The potential savings for the factory is illustrated in the optimization example described in Chapter 4.5. Both the ARMAX-GARCH model and the Seq2Seq-LSTM network are capable of capturing tendencies in the weekly power prices. The meaning of trend in this chapter is how well the models can rank the days in a week from the lowest to the highest power prices. The following table illustrates the weekly trends in the predicted power prices from the models. The lowest prices each week are colored in green, while the highest prices are colored in red.

Lowest
Second-lowest
Middle
Second-highest
Highest

Week 1

	Actual	Seq2Seq-LSTM	ARMAX-GARCH
02.01.2023	80.72	89.79	61.54
03.01.2023	75.71	85.34	76.04
04.01.2023	51.79	52.95	91.47
05.01.2023	45.15	45.78	104.58
06.01.2023	47.51	48.18	112.98

Week 2

	Actual	Seq2Seq-LSTM	ARMAX-GARCH
09.01.2023	47.35	48.02	50.03
10.01.2023	39.07	39.88	69.70
11.01.2023	37.10	38.22	64.74
12.01.2023	36.65	37.86	61.87
13.01.2023	48.47	49.19	56.19

Week 3

	Actual	Seq2Seq-LSTM	ARMAX-GARCH
16.01.2023	63.04	69.79	52.54
17.01.2023	45.88	46.52	51.48
18.01.2023	45.38	46.02	60.99
19.01.2023	50.56	51.49	55.57
20.01.2023	52.26	53.54	53.16

Week 4

	Actual	Seq2Seq-LSTM	ARMAX-GARCH
23.01.2023	43.40	44.02	62.28
24.01.2023	34.50	36.21	69.03
25.01.2023	24.33	27.04	71.90
26.01.2023	39.71	40.45	74.79
27.01.2023	40.23	40.93	69.50

Week 5

	Actual	Seq2Seq-LSTM	ARMAX-GARCH
30.01.2023	26.65	29.55	78.21
31.01.2023	48.81	49.55	62.30
01.02.2023	44.44	45.07	69.33
02.02.2023	56.52	59.53	71.97
03.02.2023	54.07	55.93	81.62

Week 6

	Actual	Seq2Seq-LSTM	ARMAX-GARCH
06.02.2023	27.80	30.72	49.88
07.02.2023	24.56	27.31	55.94
08.02.2023	21.31	23.28	53.38
09.02.2023	19.14	20.78	63.77
10.02.2023	21.78	23.87	68.86

Week 7

	Actual	Seq2Seq-LSTM	ARMAX-GARCH
13.02.2023	23.91	26.53	60.18
14.02.2023	23.36	25.85	57.96
15.02.2023	22.59	24.89	47.05
16.02.2023	23.75	26.34	56.23
17.02.2023	26.73	29.67	53.96

Week 8

	Actual	Seq2Seq-LSTM	ARMAX-GARCH
20.02.2023	26.62	29.56	46.61
21.02.2023	55.50	57.99	53.63
22.02.2023	24.22	26.91	59.13
23.02.2023	24.02	26.67	51.36
24.02.2023	25.76	28.66	54.58

Figure 21 - Trend plot

Although both models are able to capture some trends in the weekly power prices, the Seq2Seq-LSTM network clearly outperforms the ARMAX-GARCH model. The network is able to capture a trend identical to the actual power price, which is not the case for the ARMAX-GARCH model. This is an important observation regarding the optimization problem, as the factory does not need to produce on the most expensive day of the week, because the production target is possible to achieve with only four production days.

Another key finding pertains to the differences between the predicted and the actual power prices in the production period. In this regard, the LSTM network yields an RMSE of 3.01, while the ARMAX model has an RMSE of 30.06 for the same timeframe. The graph presented below displays the deviations between the predicted and actual power prices for all days forecasted out of frame. Notably, the ARMAX-GARCH model exhibits significantly more volatile deviation compared to the network, indicating that it is associated with a higher degree of uncertainty in its predictions. It is important to highlight that the RMSE values for the production period and the test period cannot easily be compared due to the significant difference in their respective lengths. The test set was considerably longer and subject to more extreme observations and higher volatility, particularly towards the end of the test timeframe.

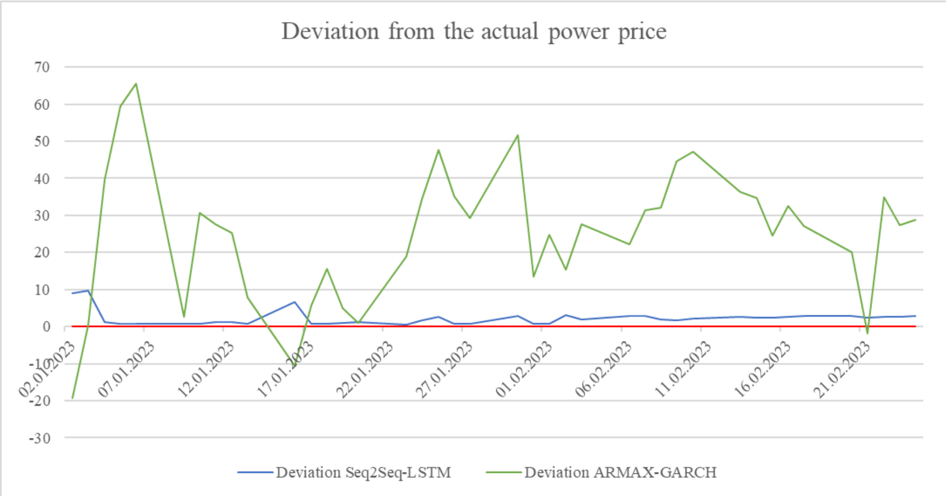


Figure 22 - Deviation from the actual power price during the production period

The graph above showcases deviations in the models, where results compared to the red line (zero) indicates the models' accuracy to actual power prices. These disparities between the models can influence the extent to which they contribute towards reducing costs. The differences between the two forecasted results will influence how the MIP-algorithm finds the optimal solution in each of the two cases, and thus, potentially provide different production plans. The graph below illustrates the predictions from both the ARMAX-GARCH model and the Se2Seq-LSTM network, as well as the actual power price during the production period.

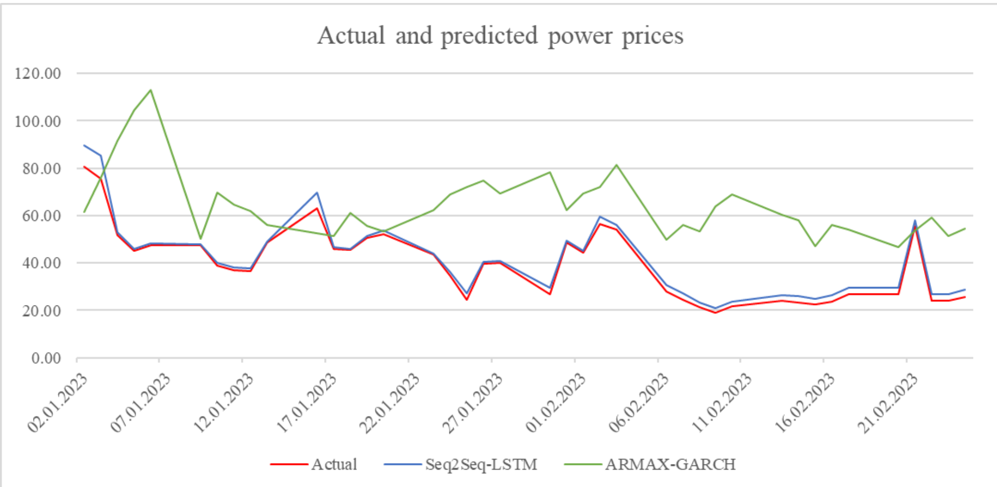


Figure 23 - Predicted power prices during the production period

5.3.2 Cost minimization with the predicted prices

Drawing upon the optimization problem presented in chapter 4.5 and utilizing forecasted prices as input for the optimization problem, scenarios can be generated that highlight the financial benefits and potential drawbacks of utilizing the models. The tables presented below demonstrate the extent to which the models' impact and differentiates cost minimization, as well as how they deviate from the actual cost of production. This is achieved by examining how each model influences the distribution of produced hydrogen. It is important to note that the quantity of production for each day is not consistent between the two models. They make distinct predictions on power prices, thus influencing the optimization problem to make different choices. For instance, the MIP-algorithm will either reduce or avoid production on the day with the highest predicted power price. This is an important observation, as it emphasizes the importance of not only accurate day-to-day power price predictions, but the importance of the models' comprehension of trends. The degree to which the models deviate from the actual power price may not be as crucial as its ability to accurately identify the trend, and correctly point out the most expensive day.

ARMAX-GARCH

	Week 1	*Week 2*	*Week 3*	*Week 4*	*Week 5*	*Week 6*	*Week 7*	*Week 8*	Total	Key figures
Model	8 434 715	6 081 146	5 614 327	7 026 432	7 234 449	5 846 117	5 669 196	5 458 285	51 364 667	
Scenario 2	6 606 252	4 608 427	5 598 319	3 978 814	4 789 982	2 795 830	2 875 771	3 723 044	34 976 438	
Deviation	**1 828 463**	**1 472 719**	**16 008**	**3 047 618**	**2 444 467**	**3 050 287**	**2 793 425**	**1 735 241**	**16 388 229**	**46.86 %**
Scenario 2	6 606 252	4 608 427	5 598 319	3 978 814	4 789 982	2 795 830	2 875 771	3 723 044	34 976 438	
Scenario 1	6 427 606	4 687 037	5 601 854	4 187 548	5 099 346	2 912 313	3 020 816	3 695 984	35 632 505	
Savings	**178 646**	**- 78 610**	**- 3 536**	**- 208 734**	**- 309 364**	**- 116 484**	**- 145 045**	**27 060**	**- 656 066**	**-1.84 %**

Seq2Seq-LSTM

	Week 1	*Week 2*	*Week 3*	*Week 4*	*Week 5*	*Week 6*	*Week 7*	*Week 8*	Total	Key figures
Model	6 033 385	4 456 338	5 253 369	4 003 620	4 830 194	2 840 724	3 042 870	3 234 438	33 694 938	
Scenario 2	5 759 057	4 366 005	5 171 718	3 864 140	4 686 141	2 642 457	2 807 192	2 970 968	32 267 678	
Deviation	**274 328**	**90 333**	**81 651**	**139 480**	**144 053**	**198 267**	**235 678**	**263 470**	**1 427 260**	**4.42 %**
Scenario 2	5 759 057	4 366 005	5 171 718	3 864 140	4 686 141	2 642 457	2 807 192	2 970 968	32 267 678	
Scenario 1	6 427 606	4 687 037	5 601 854	4 187 548	5 099 346	2 912 313	3 020 816	3 695 984	35 632 505	
Savings	**- 668 549**	**- 321 032**	**- 430 136**	**- 323 408**	**- 413 205**	**- 269 856**	**- 213 624**	**- 725 016**	**- 3 364 827**	**-9.44 %**

Figure 24 - Production results

The key figure *Deviation* illustrates how much the models deviate from the actual production expenses when producing the quantity according to the predicted power prices. This figure is the models' predicted expenses subtracted *Scenario 2* as described in chapter 4.4. The optimization based on the ARMAX-GARCH-predicted power prices overestimates the total production costs by 47% while the optimization based on predictions from the Seq2Seq-LSTM only overestimates by 4%. This discrepancy indicates a deviation from the actual production cost.

Despite the models' relatively high deviations, it is important to recognize the noteworthy figures on cost savings. The *Savings* figures are based on producing the optimal quantity of hydrogen, based on each model over eight weeks, but using the actual power prices. This is compared with an average production quantity through the week, illustrated with *Scenario 1*. These savings illustrate the potential benefits of utilizing these models for cost minimization for hydrogen production. The ARMAX-GARCH model produces cost savings for factory in six of the eight production weeks, while incurring additional expenses in the first and the last week of the production period. As a result, the total cost savings at the end of the period amount to NOK 656 066, which is equivalent to a saving of 1.84% through the production period with respect to the ARMAX-GARCH-influenced optimization. Furthermore, if the model was required to adhere to the same weekly transportation cost as *Scenario 1*, the actual cost reduction achieved through production according to the model would amount to NOK 506 066.

In contrast the Seq2Seq-LSTM network consistently reduces costs every week during the production period, resulting in total cost savings of NOK 3 364 827 for the production period. This translates to 9.44% savings when producing according to the neural network with actual power prices, as opposed to equally distributed production – *Scenario 1*. A noteworthy observation is that the models generate vastly different savings in the most expensive weeks (week 1 and 8). This is because the power prices in the other weeks are lower and relatively more stable, resulting in the optimization and decision based on the models having less impact compared to *Scenario 1*. The following charts illustrates the weekly and accumulated weekly savings when producing according to the quantity suggested, based on the predicted power prices from the predictive models.

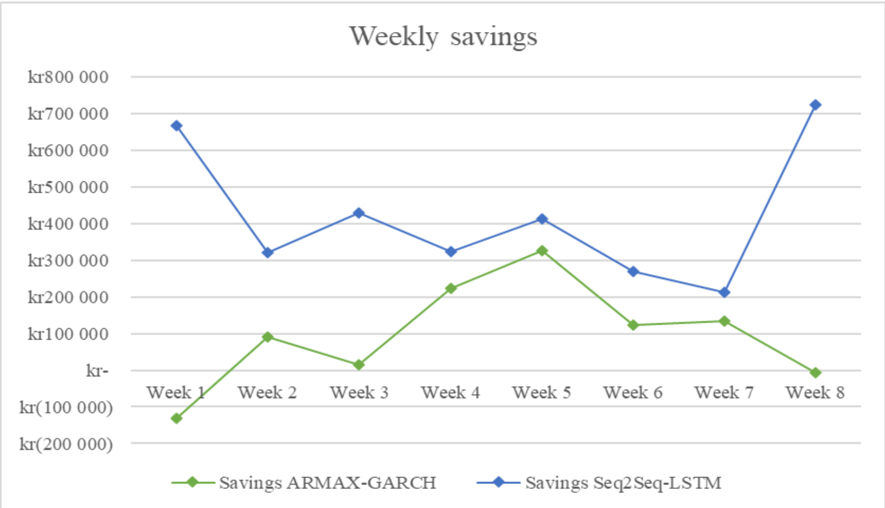


Figure 25 - Weekly savings during the production period

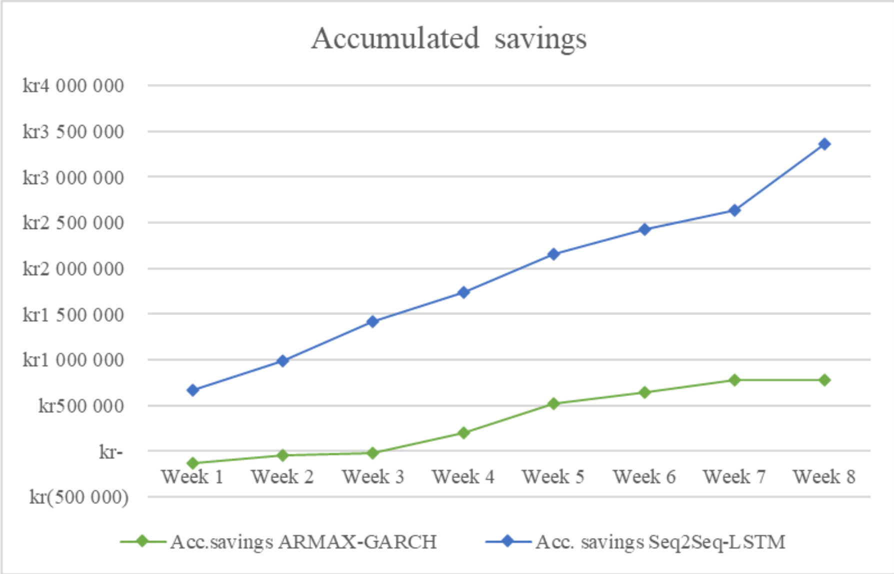


Figure 26 - Accumulated savings during the production period

5.3.3 Risk analysis

Given the relatively short duration of the production period, it presents a challenge to draw conclusions about the potential for long-term savings, and the general risks with model-guided production. Given that each week entails unique production circumstances, there lies a challenge with quantifying the precise risk associated with each model. However, to illustrate the potential risk during the production period, and to illustrate a potential general risk of combining these models with the MIP-algorithm, we have created following plots.

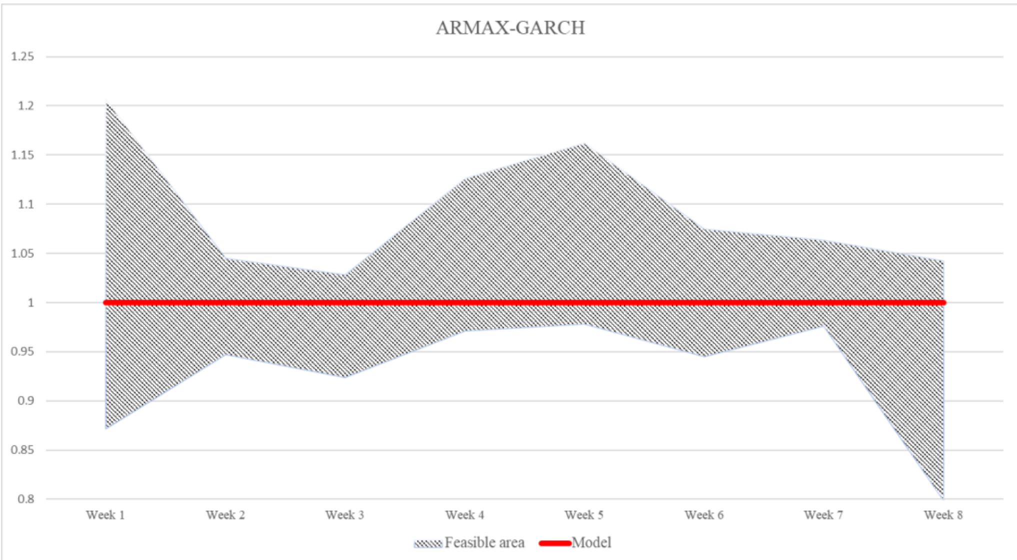


Figure 27 - ARMAX-GARCH risk analysis

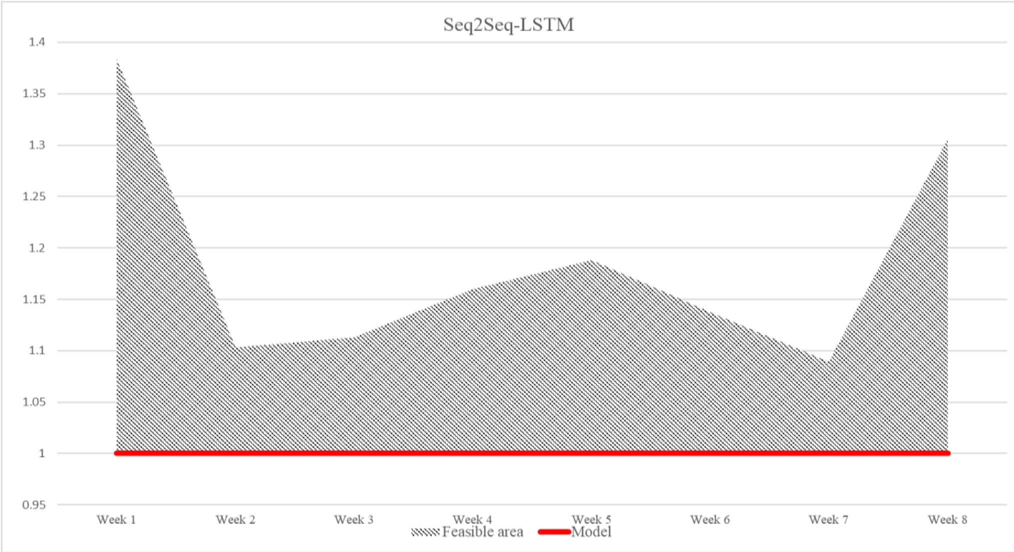


Figure 28 - Seq2Seq-LSTM risk analysis

The Y-axis illustrates the financial impacts of the models, based on how the models theoretically could perform. The red line marks the models' actual performance during the production period. The area below the red line but within the feasible area, indicates the potential areas where the model could have performed, assuming improved trend accuracy. Conversely, areas above the red line illustrate potential performance levels if the models had poorer trend depiction.

The feasible area is bounded by two lines. The upper boundary represents a pessimistic scenario where the model entirely misses the trend. This boundary is created by maximizing costs based on the actual power prices, reflecting what the models would have done if they completely missed the trend. The lower boundary illustrates an optimistic scenario where the models perfectly capture the trend. This boundary is formed by minimizing costs based on the actual power prices.

In a theoretical scenario where the ARMAX-GARCH model more accurately captured the trend, its performance could have improved by 12% in week 1 and 20% in week 8. However, the model could also have performed 20% and 4% worse in the same weeks, indicating that the models do not miss trends completely. In contrast, the Seq2Seq-LSTM model displayed a perfect trend alignment during the production week, leading to an optimal performance when combined with the MIP-algorithm. It is important to note that this does not imply a 100% accurate day-to-day price prediction, but rather a 100% accurate day-to-day depiction of the trend, which has optimal effect on the optimization.

6 Discussion

6.1 The superior performance of the Seq2Seq-LSTM

The choice to use a conventional LSTM network as the baseline machine learning model, is due to the model's extent use in recent time series research. Conventional LSTM models have been effectively utilized in power price forecasting in previous literature, especially regarding single time step forecasting. However, this thesis found a Seq2Seq-LSTM model to be superior to a conventional LSTM model. This is in line with previous research done by Aranguren et. al (2022) and Gong et. al (2019). The proposed structure showed a reduction in RMSE of 53% compared to the thesis' baseline model. The structure seems to be especially well suited for multi-step forecasting. Due to the black box nature of the model, it can be difficult to fully comprehend the reasons behind the high accuracy, but we have some hypothesis as to why.

In a conventional LSTM model, the input sequence is mapped to a sequence of hidden states, and the final hidden state is then used to generate the output sequence (see figure 5). This has the potential to result in fewer dependencies captured. The encoder-decoder architecture of the Seq2Seq-LSTM may allow the model to learn the input-output mapping in a more structured way. The encoder maps the input sequence to a fixed-length internal representation, which tries to capture the most important information about the input sequence. The decoder then uses this internal representation to generate the output. For instance, the encoder-LSTM might focus on capturing trends and dependencies in the data while the decoder-LSTM focuses on generating the forecast based on these trends and dependencies. Additionally, the Seq2Seq-LSTM model's ability to handle variable length input and output sequences might make it specifically suited for the forecasting task in this thesis. The encoder-decoder architecture allows the model to handle a variable length input, in this case a window of the last fifty days, to predict a sequence of five days. Due to the autoregressive nature of time series forecasting, the overall flexibility of the Seq2Seq-LSTM might be the main reason for the increased accuracy. It is also important to note that the Seq2Seq-LSTM model converged faster during training compared to the conventional LSTM. Pinpointing the exact reason is challenging, but one possible explanation could be the increased complexity of the model, which may have led to a better fit with the data.

In comparison to the ARMAX-GARCH model, deep learning techniques such as Seq2Seq-LSTM benefits from effective capturing of non-linear relationships in data. Weather conditions are typically non-linear (see Appendix A04). Demand for electricity might increase or decrease with temperature, but not at a constant rate. For instance, on cold days, demand might rise sharply due to increased usage of heating. Due to the Seq2Seq-LSTM-networks gating and memory mechanisms (see figure 5), the model can handle long-term dependencies in data far more effectively than the ARMAX-GARCH model. The gates can learn which data in a sequence is important to keep or discard, thereby reducing the impact of irrelevant inputs. The ARMAX-GARCH model assumes that the current value of the time series depends on the fixed values of past values. Additionally, the weight given to past observations decreases as the model go further back in time. This is because the model assumes that more recent data points are more relevant for predicting the current value, which restricts the model's ability to capture long-term dependencies. Furthermore, the ARMAX-GARCH model assume stationarity, and even though the ADF-test (see Appendix A01) showed the data was stationary, the time series heavily changes its underlying characteristics towards the end of the time series. Therefore, because statistical models such as ARMAX-GARCH assumes that incidents that have occurred in the past will continue, the model were not able to capture the volatility at the end of the time series as well as the Seq2Seq-LSTM. This is also one of the reasons why the ARMAX-GARCH model performs poorly during the production period in the hydrogen case. The high volatility and price spikes that happened towards the end of the test set is added to the window size, and the ARMAX-GARCH assumes such high price spikes and volatility to happen during the production period.

Even though the daily forecasts during the production period were not perfectly depicted by the Seq2Seq-LSTM, with two deviations as high as 10 øre/kWh, the model is perfectly able to rank the days in terms of lowest to highest price. Despite this, the model is not able to generalize on extreme values as seen in figure 17. Some of this can probably be related to the lack of such extreme values in the training set. Due to the time series structures of the model, the training, validation and test needs to be split in a time series manner, meaning techniques such as k-fold cross validation in a random split is not suited. The model can be, and should be, retrained to cope for this; however, evaluating the model's ability to generalize to unseen instances of high prices can prove challenging. This is primarily because real-time instances of extreme values are required to truly assess the model's capabilities. Another possible assumption is that there are other architectural structures and hyperparameters which would increase the model's performance on handling extreme values.

Despite the high performance of the Seq2Seq-LSTM, the ARMAX-GARCH model are much more interpretable. This can be a major advantage for insights, scenario simulations and a broader understanding of the underlying factors' implications of the power price. Generally, they are also easier to implement than more complex models. Therefore, even though complex machine learning methods may be better at accurately depicting the power price and its usage in this thesis, statistical methods can be much better suited for other tasks.

6.2 Critique of the optimization case

We needed to create a realistic case of hydrogen production to illustrate the financial impact of the models. The production target and the daily production of the case factory mirrored that of a real-world hydrogen production facility, according to the anonymous hydrogen corporation. The inclusion of a transportation cost, which illustrates a transportation, storage- or start-up cost is also what to expect in a real-world factory. Despite including the transportation cost, the optimization case neglects to some extent the storage cost of the produced hydrogen. This is because the cost varies depending on the locations and the production quantity of the factories, which makes it difficult to quantify. It is reasonable to believe that the cost of storage, transportation and start-up could exceed the amount of NOK 150 000 used in this thesis' production case.

A factor affecting the *savings* figures presented in this thesis is when producing according to an average daily quantity (*Scenario 1*), the transportation cost will apply for all days. Unlike producing according to the models, because the model can achieve the production target with only four production days each week. Calculating for this issue, the ARMAX-GARCH and the Seq2Seq-LSTM network achieve savings of -1.5% and 6.3% which implies that producing according to the statistical model will gain higher costs during the production period than weekly average production.

The case does not include hourly changes in the power price, which will affect the production cost throughout the day. The case also assumes the hydrogen is produced at 100% spot-price to highlight the importance of future power price predictions. According to the corporation who provided details for the optimization case, a split of 70% fixed and 30% spot-price are realistic. This is not accounted for in the case as we highlight the spot-price's impact on the production costs. With these factors in mind, the actual savings of applying this thesis' methods might not be as high as illustrated.

The models' usage for cost minimization during low-volatility periods with low electricity prices, such as the summer, is likely to result in less significant savings compared to using the model during high-price and high volatile periods. This is because stable or constant prices throughout the week can lead to greater indifference in terms of which day production is planned, thereby reducing the benefits of the models' predictive capabilities. Conversely, even higher prices and volatility than seen during the production period could result in higher percentage savings than presented in this thesis.

6.3 Model implication on the optimization

When applying the models to the hydrogen production case, the complexity of model evaluation increases. Predicting power prices with as close to perfect accuracy as possible is not as important as the models' ability to depict the trend accurately and rank the days in terms of costliness. Theoretically, the models could miss the actual power price by a large margin, but if it correctly identifies the costliest days, this deviation will not affect the MIP-algorithm's production decision. Thus, contribute to cost reduction despite high deviation from the actual price. Because of this, merely evaluating the models' performance based on loss metrics like RMSE and MAE may not be a sufficient indicator of how the models will perform in practice. This is especially true when combining the models with a MIP-algorithm, which priorities shifting production to the least costly days and places larger importance on trend than simply accurate price predictions. While RMSE and MAE (see chapter 3.3) measures the deviation of predicted price from the actual price, they do not account for trend. It can be assumed that there is coherence between low RMSE/MAE and the models' ability to accurately depict the trend, but it is not the exact same thing. Therefore, models that may have a higher RMSE/MAE compared to other models might work better when used in combination with optimization algorithms such as MIP.

Due to the Seq2Seq-LSTMs flexibility, the model can forecast its own exogenous variables, and this makes it suitable for further ahead forecasts. In chapter 5 (figure 18), it can be viewed that the Seq2Seq-LSTM model is capable of forecasting 14 days ahead with an RMSE/MAE of 16.73 and 9.48, respectively. Moreover, the model seems to be able to capture some trend based on the figure. Planning with 14 days could potentially leave more room for the MIP-algorithm to make better decisions, and greater options for the algorithm to shift production to the least costly days. Therefore, it is not obvious that the five-step model chosen is the best suited model for this practical purpose.

6.4 Real-world application

For a real-world application of these models, it would be crucial to keep the models updated with the most recent data. To achieve this, we propose a weekly update cycle where in every Sunday, the models are provided with the actual values from the preceding week. This approach ensures that the models are always informed by the latest data before they are used to forecast the sequence for the upcoming five production days. Integrating the forecasting with the production optimization can be seamlessly achieved in the same environment, where the forecasted outputs are automatically used as input in the optimization. The Seq2Seq-LSTM model offers quick and computationally efficient retraining. Its relatively short training time, which does not necessitate substantial CPU power, makes it possible to retrain the model frequently. This regular retraining can help the model adapt to any ongoing changes in the power market.

As discussed in chapter 3.2.7, the Seq2Seq-LSTM is a black box algorithm. It is difficult to interpret its result and back trace its reasoning from input to output. In the case of this model, this is particularly relevant, because the ultimate intention is to use the forecasted values as input in an optimization for cost minimization. The forecasted prices serve as the main input of the optimization algorithm, and if the forecasted prices miss the trend, and wrongly indicate the most expensive production day as the cheapest, the production could lose a significant amount of money. Because of this, it would have been a major advantage to thoroughly understand how the model works. This is important to better understand why the model makes accurate predictions, and perhaps even more importantly, understand why it sometimes struggles with inaccurate predictions. This is especially true with regards to extreme values. The lack of interpretability might be hurting the adaptation and real-world applications, as users may not trust the system if they do not thoroughly understand it (Ahn et al., 2021). To address this challenge, researchers have developed various methods to interpreting the inner workings of deep neural networks, such as visualizing the activations of the individual neurons (Zeiler & Fergus, 2013) or identifying important features for a particular prediction (Ribeiro et al., 2016). In addition, the widely used SHAP-framework (Lundberg & Lee, 2017) could also help increase the interpretability. While these methods can provide some insights in how deep neural networks make their decisions, they may not fully capture the complexity of the network’s internal operations.

The risk analysis carried out in this thesis helps to depict the potential outcomes in both optimistic and pessimistic scenarios. Assuming the predictive models deliver perfect accuracy or complete inaccuracy in forecasting the weekly power price trends during the production period. Because the production period in the case only extends over eight weeks, quantifying the risk associated with applying these models to the real world cannot be easily generalized. However, the risk analysis gives an idea on how the models theoretically could impact the production cost. The risk analysis implies that the Seq2Seq-LSTM is overall associated with less risk than the ARMAX-GARCH model, due to its ability to accurately depict the trend during the production period.

While the application of these models has been demonstrated in the context of hydrogen production, they can be applied to other specific industries. Other sectors where electricity constitutes a significant input in the production process may also yield similar benefits from the application of these models. Thus, these forecasting tools may hold potential for wide ranging impact across various industries, wherever managing and forecasting power costs are critical to operations.

6.5 Regarding previous literature

Finding literature on power price forecasting in NO4 and cost optimization was not an easy task. In general, previous studies mainly conduct forecasting with one timestep ahead, which makes it difficult to evaluate the models’ performance compared to other findings. There have been studies regarding power price forecasting in the Nordic countries and Norway as previously mentioned, but we have not succeeded in finding studies with models specifically trained to forecast in the price area NO4. Nor have we been able to find studies containing power price forecasting combined with an optimization case. This is although a minor issue, as the project-specificness of the optimization case makes it hard to compare this part of the thesis with previous literature. To the best of our knowledge, studies combining power price forecasting with production optimization has not been conducted before.

7 Conclusion

The main objective of this thesis has been to investigate ways to increase predictability and profitability for green hydrogen production in the NO4 price region. The thesis has divided the task into two primary research questions. The first question focuses on examining the forecasting accuracy and performance of statistical and machine learning techniques to the NO4 power market. The Seq2Seq-LSTM model significantly outperformed the best-suited statistical approach, ARMAX-GARCH, and further reduced the RMSE by 53% compared to the baseline LSTM-model. As a result, we conclude that the Seq2Seq-LSTM model holds considerable potential in enhancing predictability in the power market over a five-day horizon, and exhibits promising capabilities for longer-term forecasting.

The primary motivation and sole purpose for developing these models is to explore their practical applicability and value in a power intensive industry. The second research question investigates how this increased predictability can contribute to improving the profitability of green hydrogen production. The Seq2Seq-LSTM guided optimization led to savings of 9.44% and reduced costs by NOK 3 364 827 over an eight-week production period compared to the average production scenario (*Scenario 1*). Additionally, this outperformed the ARMAX-GARCH influenced optimization that led to cost savings of NOK 656 066.

The Seq2Seq-LSTM guided optimization, with its proficiency in capturing trends and accurately identifying the most cost-effective days for production, directed production planning with far greater accuracy than the statistical approach. This ultimately demonstrated its effect in increasing both predictability and profitability for green hydrogen production compared to the ARMAX-GARCH model. A cost reduction of 9.44% throughout the production period, illustrates that there are major possibilities for increased financial sustainability for spot-price-based production. Furthermore, this shows the promising ability of including machine learning techniques as a major part of production planning, which can achieve important competitive advantages. Thus, in conclusion, we propose a hybrid AI solution that combines a Seq2Seq-LSTM recurrent neural network with a MIP-algorithm, for increased predictability and profitability for green hydrogen production in the NO4 price area.

7.1 Future work

As previously mentioned, gathering literature on power price forecasting in Norway and the NO4 price area proved quite difficult. More research on future power prices in Norway may contribute to a competitive advantage for power intensive industries and other power demanding services such as electric ferries and shipping. Due to the prerequisites in the exogenous variables in the ARMAX-GARCH model, this thesis is primarily focused on forecasting five timesteps ahead. In order to gain more predictability for the power consumers, expanding the time horizon further in combination with short time forecast might help increasing the predictability.

Examining the models' performances with more and less volatile power prices is important to gain a better understanding of how robust the models are, and how well they perform during different conditions. Extending the production period will also be helpful to gain more significant results on whether the models actually are able to gain any savings for the power consumers. Further research on both the statistical model and the neural networks used in this thesis, might increase the accuracy of the power price predictions. Different techniques, such as combining the linearity of the ARMAX-GARCH model with the Seq2Seq-LSTM, or trying different machine learning techniques, such as transformer neural networks can further help increase the predictive accuracy. Furthermore, applying similar Seq2Seq-LSTM structures into modelling hydrogen demand, and including this into similar optimization problems connected to production, can also aid the overall decision making of green hydrogen production.

Bibliography

Aggarwal, S. K., Saini, L. M., & Kumar, A. (2009, January). *Electricity price forecasting in deregulated markets: a review and evaluation.* Retrieved from Emerald insight: https://www.emerald.com/insight/content/doi/10.1108/17506220911005731/full/pdf

Ahn, D., Almaatouq, A., Gulabani, M., & Hosanagar, K. (2021, November). *Research Gate.* Retrieved from Will We Trust What We Don't Understand? Impact of Model Interpretability and Outcome Feedback on Trust in AI: https://www.researchgate.net/publication/356283322_Will_We_Trust_What_We_Don't_Understand_Impact_of_Model_Interpretability_and_Outcome_Feedback_on_Trust_in_AI/link/62b937186ec05339cca6b656/download

Akaike, H. (1974, December 6). *A New Look at the Statistical Model Identification.* Retrieved from IEEE Xplore: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1100705

Aranguren, C., Fragoso, A., & Aguilera, R. (2022). *Sequence to Sequence (Seq2Seq) Long Short-Term Memory (LSTM) for oil production forecast of shale reservoirs.*

Arbel, N. (2018, December 21). *DataDrivenInvestor*. Retrieved from How LSTM networks solve the problem of vanishing gradients: https://medium.datadriveninvestor.com/how-do-lstm-networks-solve-the-problem-of-vanishing-gradients-a6784971a577

Athanasopoulos, G., & Rob, H. J. (2018). *Forecasting: Principles and Practice.* https://otexts.com/fpp2/: OTexts.

Brembo, F., & Olaisen, S. R. (2021, October 26). *NRK*. Retrieved from Røkke-selskap lanserer stor industrisatsing i Narvik - kjøper omstridt industritomt: https://www.nrk.no/nordland/aker-satser-200-millioner-pa-industriarealer-i-narvik.-i-tillegg-satser-de-pa-batteriproduksjon.-1.15704235

Brownlee, J. (2017a, June 23). *Machine Learning Mastery*. Retrieved from A Gentle Introduction to Backpropagation Through Time - MachineLearningMastery.com: https://machinelearningmastery.com/gentle-introduction-backpropagation-time/

Brownlee, J. (2017b, March 31). *Simple Time Series Forecasting Models to Test So That You Don't Fool Yourself*. Retrieved from Machine Learning Mastery: https://machinelearningmastery.com/simple-time-series-forecasting-models/

Brownlee, J. (2020, August 20). *Machine Learning Mastery.* Retrieved from A Gentle Introduction to the Rectified Linear Unit (ReLU): https://machinelearningmastery.com/rectified-linear-activation-function-for-deep-learning-neural-networks/

Bugge, S. (2022, April 27). *E24*. Retrieved from Norske strømpriser stiger etter russisk gasskutt: https://e24.no/norsk-oekonomi/i/Xq161x/norske-stroempriser-stiger-etter-russisk-gasskutt

Cerjan, M., Krželj, I., Vidak, M., & Delimar, M. (2013, July 4). *Ieexplore*. Retrieved from A Literature Review with Statistical Analysis of Electricity Price Forecasting Methods: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6625068&tag=1

Cevik, S., & Ninomiya, K. (2022, November). *Chasing the Sun and Catching the Wind: Energy Transition and Electricity Prices in Europe.* Retrieved from IMF.org: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwj24t2-7cT-AhWUQ_EDHY8KC04QFnoECA4QAQ&url=https%3A%2F%2Fwww.imf.org%2F-%2Fmedia%2FFiles%2FPublications%2FWP%2F2022%2FEnglish%2Fwpiea2022220-print-pdf.ashx&usg=AOvVaw1qbW_OSWJxoPb7HtV

Chai, T., & Draxler, R. (2014, June 30). *Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature.* Retrieved from ResearchGate: https://www.researchgate.net/publication/272024186_Root_mean_square_error_RMSE_or_mean_absolute_error_MAE-_Arguments_against_avoiding_RMSE_in_the_literature

Chang, C.-L., McAleer, M., & Tansuchat, R. (2010, April 30). *Analyzing and forecasting volatility spillovers, asymmetries and hedging in major oil markets.* Retrieved from ScienceDirect: https://reader.elsevier.com/reader/sd/pii/S0140988310000708?token=BEC3F6A5E709689A9948D77E82718CDA82B891D49626898F82CC9924BC232A926E1BF88A40B26DE8A7226956BEF0AA7C&originRegion=eu-west-1&originCreation=20230420063514

Changtzeng, H.-C., Chang, D.-F., & Lo, Y.-H. (2019). *Detecting Concurrent Relationships of Selected Time Series Data for ARIMAX Model.* Retrieved from ICIC Express Letters: http://www.icicelb.org/ellb/contents/2019/10/elb-10-10-10.pdf

Chen, C., Hu, J., Meng, Q., & Zhang, Y. (2011, june). *Short-time Traffic Flow Prediction with ARIMA-GARCH model.* Retrieved from IEEE: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5940418

Chen, L., & Yu, L. (2013). *Structural Nonlinear Damage Identification Algorithm Based on Time Series ARMA/GARCH Model*. Retrieved from Sage journals: https://journals.sagepub.com/doi/pdf/10.1260/1369-4332.16.9.1597

Chollet, F. (2018). *Deep Learning with Python.* Manning.

Cools, M., Moons, E., & Wets, G. (2009). *Investigating the Variability in Daily Traffic Counts Through Use of ARIMAX and SARIMAX Models.* Retrieved from journals sagepub: https://journals.sagepub.com/doi/pdf/10.3141/2136-07

Energifakta Norge. (2022, May 13). *Kraftmarkedet.* Retrieved from Energifakta Norge: https://energifaktanorge.no/norsk-energiforsyning/kraftmarkedet/

Engie. (2021, October 27). *Green hydrogen production: hos does it work?* Retrieved from Engie: https://www.engie.com/en/renewables/hydrogen/green-hydrogen-production

Engle, R. (n.d.). *GARCH 101: An Introduction to the Use of ARCH/GARCH models in Applied.* Retrieved from web-static: https://web-static.stern.nyu.edu/rengle/GARCH101.PDF

Frontline Systems. (n.d.). *Excel solver - algorithms and methods used*. Retrieved from Frontline Solvers: https://www.solver.com/excel-solver-algorithms-and-methods-used

Goled, S. (2021, November 30). *analyticsindamag*. Retrieved from How Do Activation Functions Introduce Non-Linearity In Neural Networks?: https://analyticsindiamag.com/how-do-activation-functions-introduce-non-linearity-in-neural-networks/

Gong, G., An, X., Mahato, N. K., Sun, S., & Chen, S. (2019, August). *Mdpi.* Retrieved from Research on Short-Term Load Prediction Based on Seq2Seq Model: https://www.mdpi.com/1996-1073/12/16/3199

Hiller, F. S., & Lieberman, G. J. (2021). *Introduction to Operations Reshearch.* New York: McGraw-Hill Education - Forlag.

Hughes, A. (2023, February 2). *Science Focus.* Retrieved from ChatGPT: Everything you need to know about OpenAI's GPT-3 tool: https://www.sciencefocus.com/future-technology/gpt-3/

IBM. (2019). *IBM*. Retrieved from What are Neural Networks?: https://www.ibm.com/topics/neural-networks

Kampevoll, F., & Lorch-Falck, S. (2022, August 18). *Fortellingen om kraftkablene*. Retrieved from NRK: https://www.nrk.no/norge/xl/fortellingen-om-kraftkablene-1.16060842

Kingma, D. P., & Ba, J. (2017, January 30). *Arxiv.* Retrieved from Adam: A Method for Stochastic Optimization: https://arxiv.org/pdf/1412.6980.pdf

Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020, December 25). *mdpi.* Retrieved from Explainable AI: A Review of Machine Learning Interpretability Methods: https://www.mdpi.com/1099-4300/23/1/18

Llugsi, R., El Yacoubi, S., Fontaine, A., & Lupera, P. (2021, November 10). Retrieved from Comparison between Adam, AdaMax and Adam W optimizers to implement a Weather Forecast based on Neural Networks for the Andrean city of Quito: https://ieeexplore.ieee.org/abstract/document/9590681

Lundberg, S., & Lee, S.-I. (2017, November 25). *A unified Approach to Interpreting Model Predictions*. Retrieved from Cornell University arxiv: https://arxiv.org/abs/1705.07874

Makridakis, S., & Hibon, M. (1997). ARMA Models and the Box_jenkins Methodology. *Jurnal of Forecasting*, pp. 147-163.

Memarzadeh, G., & Keynia, F. (2021). *Elsevier*. Retrieved from Short-term electricity load and price forecasting by a new optimal LSTM-NN based prediction algorithm: https://reader.elsevier.com/reader/sd/pii/S0378779620307938?token=9746A781E01867B7B8487AFCBA8AA2DCBF8A021799A2F4C816E3B4EA899DA5DA450E42B690849D32ADA11C9020E803AA&originRegion=eu-west-1&originCreation=20230209100727

Molnar, C. (2019). *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable.* https://christophm.github.io/interpretable-ml-book/preface-by-the-author.html: Ebook.

National Oceanic and Atmospheric Administration. (n.d.). *How Reliable Are Weather Forecasts?* Retrieved from SciJinks: https://scijinks.gov/forecast-reliability/

Neumann, D. (2014). *Ridge Regression and Lasse.* Retrieved from Information Systems Research: https://www.is.uni-freiburg.de/resources/seminar-papers/Ridge_Regression_LASSO.pdf

Nielsen, M. (2018). *Neural Networks and Deep Learning.* Ebook, NeuralNetworksAndDeepLearning.com.

Norges vassdrags- og energidirektorat. (2022, January 4). *NVE*. Retrieved from Hvor kommer strømmen fra?: Norges vassdrags- og energidirektorat

Olah, C. (2015, August 27). *Colah's blog.* Retrieved from Understanding LSTM Networks: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Porshnev, A. V., Lakshina, V. V., & Red'kin, I. E. (2016, April 12). *Colud Emotional Markers in Twitter Posts Add Information to the Stock Market ARMAX-GARCH Model.* Retrieved from SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2763583

Prabhakaran, S. (2019, November 19). *Augmented Dicky Fuller Test (ADF Test) - Must Read Huide*. Retrieved from Machine learning +: https://www.machinelearningplus.com/time-series/augmented-dickey-fuller-test/

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). *kdd.org.* Retrieved from "Why should I Trust You?: Explaining the Predictions of Any Classifier: https://www.kdd.org/kdd2016/papers/files/rfp0573-ribeiroA.pdf

Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and Its Applications: With R Examples.* Springer International Publishing.

Singh, Y. (2022, October 14). *Proclus Acaemy*. Retrieved from 3 Regression Metrics You Must Know: MAE, MSE, and RMSE: https://proclusacademy.com/blog/explainer/regression-metrics-you-must-know/

Skagerak Energi. (n.d.). *Hvorfor er strømprisen i Nord-Norge tidvis mye lavere enn i sør?* Retrieved from Skagerak Energi: https://www.skagerakenergi.no/strompriser/hvorfor-er-stromprisen-i-nord-norge-tidvis-mye-lavere-enn-i-sor-article3751-3077.html

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov. (2014, June). Retrieved from Dropout: A Simple Way to Prevent Neural Networks from Overfitting: https://dl.acm.org/doi/pdf/10.5555/2627435.2670313

Statnett. (2022, October 3). *Statnett*. Retrieved from Derfor har vi prisområder: https://www.statnett.no/om-statnett/bli-bedre-kjent-med-statnett/om-strompriser/fakta-om-prisomrader/

Stone, M. (1979, April). *Comments on Model Selection Criteria of Akaike and Schwarz.* Retrieved from Jstor: https://www.jstor.org/stable/2985044?seq=1

Sutskever, I., Vinyals, O., & Le, Q. V. (2014, December 14). *arxiv.org.* Retrieved from Sequence to Sequence Learning with Neural Networks: https://arxiv.org/abs/1409.3215

Swider, D. J., & Weber, C. (2007, February 7). *Extended ARMA models for estimating price developments on day-ahead electricity markets.* Retrieved from ScienceDirect: https://reader.elsevier.com/reader/sd/pii/S0378779606001313?token=242AD522C1639BFA5388B7E737D54B60A7E724332A12B5A2BAF1B9987770E4DEF485ADDB70C055D4C58D879770D15E14&originRegion=eu-west-1&originCreation=20230420065535

The Norwegian Ministry of Finance. (2019). NOU 2019:11. *Enklere merverdiavgift med en sats*. The Norwegian Government.

Toisoul, A. (2020, November 21). *[Machine learning fundamentals] Training, validation and test sets.* Retrieved from Antoinetlc: https://www.antoinetlc.com/blog-summary/machine-learning-how-to-use-the-training-validation-set-and-test-sets

Torro, H. (2007, June 7). *Forecasting Weekly Electricity Prices at Nord Pool.* Retrieved from SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=991532

Turing. (n.d.). *Guide to Autoregressive Models*. Retrieved from Turing: https://www.turing.com/kb/guide-to-autoregressive-models

Vamathevan, G., Dynge, M. F., & Cali, U. (2022, August 10). *Ieexplore*. Retrieved from Electricity Price Forecasting for Norwegian Day-Ahead Market using Hybrid AI Models: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9921003

Weller, M. (2018, October 16). *Novatec Consulting GmbH*. Retrieved from Recurrent Neural Networks for time series forecasting | Novatec - Novatec: https://www.novatec-gmbh.de/en/blog/recurrent-neural-networks-for-time-series-forecasting/

Wu, Y., Schuster, M., Chen, Z., Le, Q., Norouzi, M., Macherey, W., . . . Dean, J. (2016, September 26). *Arxiv.* Retrieved from Google's Neural Machine Translation System: Briding the Gap between Human and Machine Translation: https://arxiv.org/abs/1609.08144

Zeiler, M. D., & Fergus, R. (2013, November 28). *arxiv.org.* Retrieved from Visualizing and Understanding Convolutional Networks: https://arxiv.org/abs/1311.2901

Zivot, E., & Wang, J. (2006). *Modeling Financial Time Series with s-plus.* Springer.

Appendix

A01 Augmented Dickey-Fuller test

```
ADF Statistic: -6.143216
p-value: 0.000000
PRIS: Stationary data

ADF Statistic: -10.083756
p-value: 0.000000
FYLLINGSGRAD: Stationary data

ADF Statistic: -5.953611
p-value: 0.000000
TEMPERATUR: Stationary data

ADF Statistic: -33.330627
p-value: 0.000000
NEDBØR: Stationary data

ADF Statistic: -12.735885
p-value: 0.000000
VIND: Stationary data
```

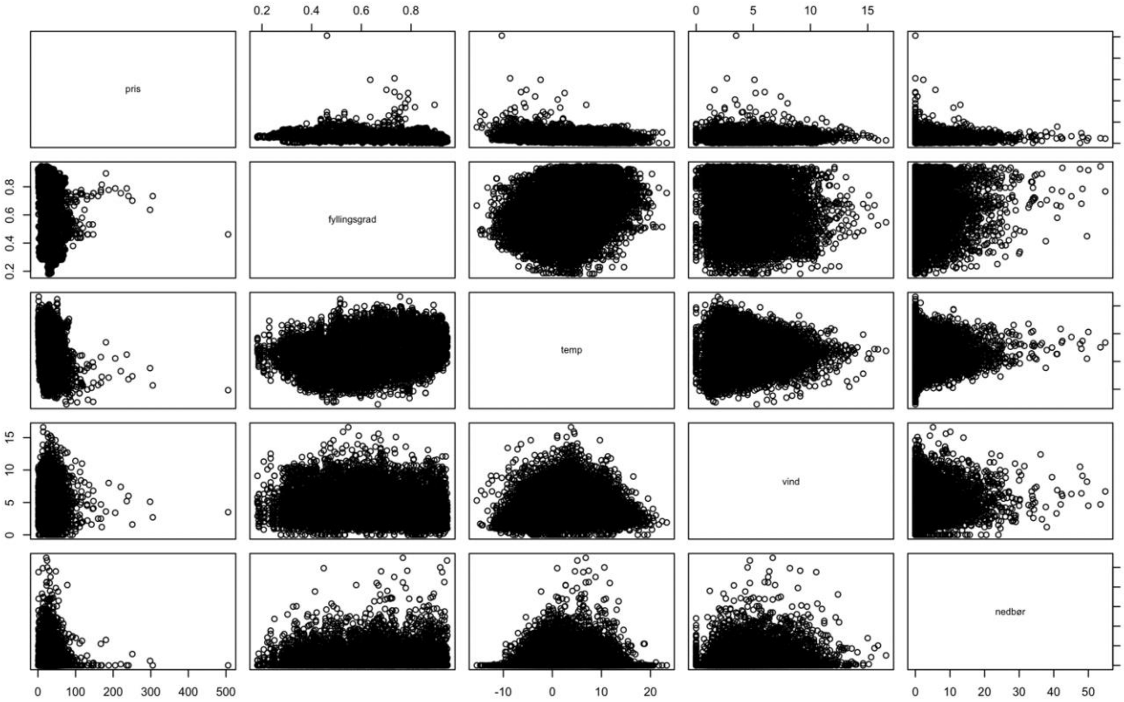
A02 Breusch-Pagan test for heteroscedasticity

```
Ho: Homoscedasticity    Ha: Heteroscedasticity
Breusch-Pagen-test, p-value:  0.0024242241971499665
H0 hypothesis is redjected due to: p-value < 0.05
```

A03 AIC and BIC values ARMAX-GARCH/SARMAX-GARCH

	AIC	BIC
ARMA	43 112	43 145
ARMAX fyllingsgrad	43 110	43 150
ARMAX temp	43 111	43 152
ARMAX nedbør	43 114	43 154
ARMAX vind	43 114	43 154
ARMAX fyllingsgrad & temp	43 110	43 157
ARMAX fyllingsgrad & nedbør	43 110	43 157
ARMAX fyllingsgrad & vind	43 111	43 158
ARMAX temp & nedbør	43 113	43 160
ARMAX temp & vind	43 114	43 161
ARMAX nedbør & vind	43 116	43 163
ARMAX fyllingsgrad, temp & nedbør	43 121	43 175
ARMAX fyllingsgrad, temp & vind	43 112	43 166
ARMAX fyllingsgrad, nedbør & vind	43 130	43 184
ARMAX temp, nedbør & fyllingsgrad	43 115	43 169
ARMAX Fyllingsgrad, temp, nedbør, vind	43 134	43 195
ARMAX temp, fyllingsgrad, nedbør & vind	43 147	43 207
ARMA-GARCH(4,6)	31 905	31 987
ARMAX-GARCH(4,6) fyllingsgrad	32 033	32 113
ARMAX-GARCH(4,6) temp	31 922	32 003
ARMAX-GARCH(4,6) nedbør	31 912	31 992
ARMAX-GARCH(4,6) vind	31 906	31 986
ARMAX-GARCH(4,6) fyllingsgrad & temp	32 031	32 112
ARMAX-GARCH(4,6) fyllingsgrad & nedbør	32 001	32 082
ARMAX-GARCH(4,6) fyllingsgrad & vind	32 036	32 117
ARMAX-GARCH(4,6) temp & nedbør	31 926	32 006
ARMAX-GARCH(4,6) temp & vind	31 917	31 997
ARMAX-GARCH(4,6) nedbør & vind	31 900	31 980
ARMAX-GARCH(4,6) fyllingsgrad, temp & nedbør	31 872	31 952
ARMAX-GARCH(4,6) fyllingsgrad, temp & vind	32 047	32 128
ARMAX-GARCH(4,6) fyllingsgrad, nedbør & vind	31 995	32 036
ARMAX-GARCH(4,6) temp, nedbør & vind	31 935	32 016
ARMAX-GARCH(4,6) temp, fyllingsgrad, nedbør & vind	**31 751**	**31 831**
SARMAX(1,0,1,7)	42 943	43 017
SARMAX-GARCH(2,2)(1,0,1,7)(3,6) temp, fyllingsgrad, nedbør & vind	**30 301**	**30 375**

A04 Scatter plot of the variables



A05 Optimization using ARMAX-GARCH and Seq2Seq-LSTM

Optimization week 1 ARMAX-GARCH and Seq2Seq-LSTM

Week 1	**Capacity restrictions**				**Hour**		**Day**
ARMAX-GARCH	43.24	<=	43.24	TON	100	MWH	2400
	43.24	<=	43.24	TON	40 %	Power reduction	40 %
	43.24	<=	43.24	TON	60	MWH of hydroger	1440
	40.28	<=	43.24	TON	33.3	MWH to Ton	33.3
	0	<=	0	TON	1.8	Ton pr unit	43.24
	Production target						
	170	>=	170	TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
02.jan Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.62	kr 1 476 849	kr 150 000	1	kr 150 000
03.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.76	kr 1 824 823	kr 150 000	1	kr 150 000
04.jan Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.91	kr 2 195 115	kr 150 000	1	kr 150 000
05.jan Thursday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	1.05	kr 2 337 928	kr 150 000	1	kr 150 000
06.jan Friday	43.24	0	33.3	-	60 %	-	-	0	1.13	kr -	kr 150 000	0	kr -

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
02.jan Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.81	kr 1 523 186	kr 150 000	1	kr 150 000
03.jan Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.76	kr 1 428 648	kr 150 000	1	kr 150 000
04.jan Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.52	kr 977 277	kr 150 000	1	kr 150 000
05.jan Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.45	kr 851 981	kr 150 000	1	kr 150 000
06.jan Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.48	kr 896 514	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
02.jan Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.81	kr 1 937 135	kr 150 000	1	kr 150 000
03.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.76	kr 1 816 904	kr 150 000	1	kr 150 000
04.jan Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.52	kr 1 242 867	kr 150 000	1	kr 150 000
05.jan Thursday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.45	kr 1 009 346	kr 150 000	1	kr 150 000
06.jan Friday	43.24	0	33.3	-	60 %	-	-	0	0.48	kr -	kr 150 000	0	kr -

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	7 834 715.44	Power expenses	kr	5 677 605.60	Power expenses	kr	6 006 251.51
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**8 434 715.44**	**Production cost**	**kr**	**6 427 605.60**	**Production cost**	**kr**	**6 606 251.51**

Week 1	**Capacity restrictions**				**Hour**		**Day**
Seq2Seq	0	<=	0	TON	100	MWH	2400
	40.28	<=	43.24	TON	40 %	Power reduction	40 %
	43.24	<=	43.24	TON	60	MWH of hydroger	1440
	43.24	<=	43.24	TON	33.3	MWH to Ton	33.3
	43.24	<=	43.24	TON	1.8	Ton pr unit	43.24
	Production target						
	170	>=	170	TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
02.jan Monday	43.24	0	33.3	-	60 %	-	-	0	0.90	kr -	kr 150 000	0	kr -
03.jan Tuesday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.85	kr 1 907 810	kr 150 000	1	kr 150 000
04.jan Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.53	kr 1 270 705	kr 150 000	1	kr 150 000
05.jan Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.46	kr 1 098 638	kr 150 000	1	kr 150 000
06.jan Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.48	kr 1 156 233	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
02.jan Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.81	kr 1 523 186	kr 150 000	1	kr 150 000
03.jan Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.76	kr 1 428 648	kr 150 000	1	kr 150 000
04.jan Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.52	kr 977 277	kr 150 000	1	kr 150 000
05.jan Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.45	kr 851 981	kr 150 000	1	kr 150 000
06.jan Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.48	kr 896 514	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
02.jan Monday	43.24	0	33.3	-	60 %	-	-	0	0.81	kr -	kr 150 000	0	kr -
03.jan Tuesday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.76	kr 1 692 527	kr 150 000	1	kr 150 000
04.jan Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.52	kr 1 242 867	kr 150 000	1	kr 150 000
05.jan Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.45	kr 1 083 519	kr 150 000	1	kr 150 000
06.jan Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.48	kr 1 140 154	kr 150 000	1	kr 150 000

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	5 433 385.40	Power expenses	kr	5 677 605.60	Power expenses	kr	5 159 067.32
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**6 033 385.40**	**Production cost**	**kr**	**6 427 605.60**	**Production cost**	**kr**	**5 759 067.32**

Optimization week 2 ARMAX-GARCH and Seq2Seq-LSTM

Week 2
ARMAX-GARCH

	Capacity restrictions				**Hour**		**Day**
	43.24	<=	43.24	TON	100	MWH	2400
	0	<=	0	TON	40 %	Power reduction	40 %
	40.28	<=	43.24	TON	60	MWH of hydroger	1440
	43.24	<=	43.24	TON	33.3	MWH to Ton	33.3
	43.24	<=	43.24	TON	1.8	Ton pr unit	43.24
	Production target						
	170	>=	170	TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
09.jan Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.50	kr 1 200 630	kr 150 000	1	kr 150 000
10.jan Tuesday	43.24	0	33.3	-	60 %	-	-	0	0.70	kr -	kr 150 000	0	kr -
11.jan Wednesday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.65	kr 1 447 289	kr 150 000	1	kr 150 000
12.jan Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.62	kr 1 484 769	kr 150 000	1	kr 150 000
13.jan Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.56	kr 1 348 459	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
09.jan Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.47	kr 893 495	kr 150 000	1	kr 150 000
10.jan Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.39	kr 737 251	kr 150 000	1	kr 150 000
11.jan Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.37	kr 700 077	kr 150 000	1	kr 150 000
12.jan Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.37	kr 691 586	kr 150 000	1	kr 150 000
13.jan Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.48	kr 914 629	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
09.jan Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.47	kr 1 136 315	kr 150 000	1	kr 150 000
10.jan Tuesday	43.24	0	33.3	-	60 %	-	-	0	0.39	kr -	kr 150 000	0	kr -
11.jan Wednesday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.37	kr 829 385	kr 150 000	1	kr 150 000
12.jan Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.37	kr 879 534	kr 150 000	1	kr 150 000
13.jan Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.48	kr 1 163 193	kr 150 000	1	kr 150 000

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	5 481 146.03	Power expenses	kr	3 937 036.80	Power expenses	kr	4 008 426.89
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**6 081 146.03**	**Production cost**	**kr**	**4 687 036.80**	**Production cost**	**kr**	**4 608 426.89**

Week 2
Seq2Seq

	Capacity restrictions				**Hour**		**Day**
	40.28	<=	43.24	TON	100	MWH	2400
	43.24	<=	43.24	TON	40 %	Power reduction	40 %
	43.24	<=	43.24	TON	60	MWH of hydroger	1440
	43.24	<=	43.24	TON	33.3	MWH to Ton	33.3
	0	<=	0	TON	1.8	Ton pr unit	43.24
	Production target						
	170	>=	170	TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
09.jan Monday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.48	kr 1 073 506	kr 150 000	1	kr 150 000
10.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.40	kr 957 048	kr 150 000	1	kr 150 000
11.jan Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.38	kr 917 211	kr 150 000	1	kr 150 000
12.jan Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.38	kr 908 572	kr 150 000	1	kr 150 000
13.jan Friday	43.24	0	33.3	-	60 %	-	-	0	0.49	kr -	kr 150 000	0	kr -

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
09.jan Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.47	kr 893 495	kr 150 000	1	kr 150 000
10.jan Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.39	kr 737 251	kr 150 000	1	kr 150 000
11.jan Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.37	kr 700 077	kr 150 000	1	kr 150 000
12.jan Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.37	kr 691 586	kr 150 000	1	kr 150 000
13.jan Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.48	kr 914 629	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
09.jan Monday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.47	kr 1 058 528	kr 150 000	1	kr 150 000
10.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.39	kr 937 610	kr 150 000	1	kr 150 000
11.jan Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.37	kr 890 333	kr 150 000	1	kr 150 000
12.jan Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.37	kr 879 534	kr 150 000	1	kr 150 000
13.jan Friday	43.24	0	33.3	-	60 %	-	-	0	0.48	kr -	kr 150 000	0	kr -

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	3 856 337.58	Power expenses	kr	3 937 036.80	Power expenses	kr	3 766 005.11
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**4 456 337.58**	**Production cost**	**kr**	**4 687 036.80**	**Production cost**	**kr**	**4 366 005.11**

Optimization week 3 ARMAX-GARCH and Seq2Seq-LSTM

Week 3

ARMAX-GARCH

Capacity restrictions				Hour		Day
43.24	<=	43.24	TON	100	MWH	2400
43.24	<=	43.24	TON	40 %	Power reduction	40 %
0	<=	0	TON	60	MWH of hydroger	1440
40.28	<=	43.24	TON	33.3	MWH to Ton	33.3
43.24	<=	43.24	TON	1.8	Ton pr unit	43.24
Production target						
170	>=	170	TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
16.jan Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.53	kr 1 260 865	kr 150 000	1	kr 150 000
17.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.51	kr 1 235 427	kr 150 000	1	kr 150 000
18.jan Wednesday	43.24	0	33.3	-	60 %	-	-	0	0.61	kr -	kr 150 000	0	kr -
19.jan Thursday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.56	kr 1 242 290	kr 150 000	1	kr 150 000
20.jan Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.53	kr 1 275 744	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
16.jan Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.63	kr 1 189 565	kr 150 000	1	kr 150 000
17.jan Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.46	kr 865 756	kr 150 000	1	kr 150 000
18.jan Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.45	kr 856 321	kr 150 000	1	kr 150 000
19.jan Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.51	kr 954 067	kr 150 000	1	kr 150 000
20.jan Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.52	kr 986 146	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
16.jan Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.63	kr 1 512 847	kr 150 000	1	kr 150 000
17.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.46	kr 1 101 037	kr 150 000	1	kr 150 000
18.jan Wednesday	43.24	0	33.3	-	60 %	-	-	0	0.45	kr -	kr 150 000	0	kr -
19.jan Thursday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.51	kr 1 130 289	kr 150 000	1	kr 150 000
20.jan Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.52	kr 1 254 146	kr 150 000	1	kr 150 000

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	5 014 326.65	Power expenses	kr	4 851 854.40	Power expenses	kr	4 998 318.90
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**5 614 326.65**	**Production cost**	**kr**	**5 601 854.40**	**Production cost**	**kr**	**5 598 318.90**

Week 3

Seq2Seq

Capacity restrictions				Hour		Day
0	<=	0	TON	100	MWH	2400
43.24	<=	43.24	TON	40 %	Power reduction	40 %
43.24	<=	43.24	TON	60	MWH of hydroger	1440
43.24	<=	43.24	TON	33.3	MWH to Ton	33.3
40.28	<=	43.24	TON	1.8	Ton pr unit	43.24
Production target						
170	>=	170	TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
16.jan Monday	43.24	0	33.3	-	60 %	-	-	0	0.70	kr -	kr 150 000	0	kr -
17.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.47	kr 1 116 396	kr 150 000	1	kr 150 000
18.jan Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.46	kr 1 104 397	kr 150 000	1	kr 150 000
19.jan Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.51	kr 1 235 667	kr 150 000	1	kr 150 000
20.jan Friday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.54	kr 1 196 908	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
16.jan Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.63	kr 1 189 565	kr 150 000	1	kr 150 000
17.jan Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.46	kr 865 756	kr 150 000	1	kr 150 000
18.jan Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.45	kr 856 321	kr 150 000	1	kr 150 000
19.jan Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.51	kr 954 067	kr 150 000	1	kr 150 000
20.jan Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.52	kr 986 146	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
16.jan Monday	43.24	0	33.3	-	60 %	-	-	0	0.63	kr -	kr 150 000	0	kr -
17.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.46	kr 1 101 037	kr 150 000	1	kr 150 000
18.jan Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.45	kr 1 089 038	kr 150 000	1	kr 150 000
19.jan Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.51	kr 1 213 349	kr 150 000	1	kr 150 000
20.jan Friday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.52	kr 1 168 293	kr 150 000	1	kr 150 000

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	4 653 368.86	Power expenses	kr	4 851 854.40	Power expenses	kr	4 571 717.93
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**5 253 368.86**	**Production cost**	**kr**	**5 601 854.40**	**Production cost**	**kr**	**5 171 717.93**

Optimization week 4 ARMAX-GARCH and Seq2Seq-LSTM

Week 4
ARMAX-GARCH

Capacity restrictions				Hour		Day
43.24	<=	43.24	TON	100	MWH	2400
43.24	<=	43.24	TON	40 %	Power reduction	40 %
40.28	<=	43.24	TON	60	MWH of hydroger	1440
0	<=	0	TON	33.3	MWH to Ton	33.3
43.24	<=	43.24	TON	1.8	Ton pr unit	43.24
Production target						
170	>=	170	TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
23.jan Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.62	kr 1 494 608	kr 150 000	1	kr 150 000
24.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.69	kr 1 656 596	kr 150 000	1	kr 150 000
25.jan Wednesday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.72	kr 1 607 353	kr 150 000	1	kr 150 000
26.jan Thursday	43.24	0	33.3	-	60 %	-	-	0	0.75	kr -	kr 150 000	0	kr -
27.jan Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.70	kr 1 667 875	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
23.jan Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.43	kr 818 958	kr 150 000	1	kr 150 000
24.jan Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.35	kr 651 015	kr 150 000	1	kr 150 000
25.jan Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.24	kr 459 107	kr 150 000	1	kr 150 000
26.jan Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.40	kr 749 328	kr 150 000	1	kr 150 000
27.jan Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.40	kr 759 140	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
23.jan Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.43	kr 1 041 522	kr 150 000	1	kr 150 000
24.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.35	kr 827 938	kr 150 000	1	kr 150 000
25.jan Wednesday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.24	kr 543 907	kr 150 000	1	kr 150 000
26.jan Thursday	43.24	0	33.3	-	60 %	-	-	0	0.40	kr -	kr 150 000	0	kr -
27.jan Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.40	kr 965 448	kr 150 000	1	kr 150 000

THE MODEL			SCENARIO 1			SCENARIO 2		
Power expenses	kr	6 426 431.80	Power expenses	kr	3 437 547.90	Power expenses	kr	3 378 814.25
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**7 026 431.80**	**Production cost**	**kr**	**4 187 547.90**	**Production cost**	**kr**	**3 978 814.25**

Week 4
Seq2Seq

Capacity restrictions				Hour		Day
0	<=	0	TON	100	MWH	2400
43.24	<=	43.24	TON	40 %	Power reduction	40 %
43.24	<=	43.24	TON	60	MWH of hydroger	1440
43.24	<=	43.24	TON	33.3	MWH to Ton	33.3
40.28	<=	43.24	TON	1.8	Ton pr unit	43.24
Production target						
170	>=	170	TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
23.jan Monday	43.24	0	33.3	-	60 %	-	-	0	0.44	kr -	kr 150 000	0	kr -
24.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.36	kr 868 975	kr 150 000	1	kr 150 000
25.jan Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.27	kr 648 911	kr 150 000	1	kr 150 000
26.jan Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.40	kr 970 727	kr 150 000	1	kr 150 000
27.jan Friday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.41	kr 915 007	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
23.jan Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.43	kr 818 958	kr 150 000	1	kr 150 000
24.jan Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.35	kr 651 015	kr 150 000	1	kr 150 000
25.jan Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.24	kr 459 107	kr 150 000	1	kr 150 000
26.jan Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.40	kr 749 328	kr 150 000	1	kr 150 000
27.jan Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.40	kr 759 140	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
23.jan Monday	43.24	0	33.3	-	60 %	-	-	0	0.43	kr -	kr 150 000	0	kr -
24.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.35	kr 827 938	kr 150 000	1	kr 150 000
25.jan Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.24	kr 583 876	kr 150 000	1	kr 150 000
26.jan Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.40	kr 952 969	kr 150 000	1	kr 150 000
27.jan Friday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.40	kr 899 358	kr 150 000	1	kr 150 000

THE MODEL			SCENARIO 1			SCENARIO 2		
Power expenses	kr	3 403 619.86	Power expenses	kr	3 437 547.90	Power expenses	kr	3 264 140.37
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**4 003 619.86**	**Production cost**	**kr**	**4 187 547.90**	**Production cost**	**kr**	**3 864 140.37**

Optimization week 5 ARMAX-GARCH and Seq2Seq-LSTM

Week 5	Capacity restrictions			Hour		Day
ARMAX-GARCH	40.28	<=	43.24 TON	100	MWH	2400
	43.24	<=	43.24 TON	40 %	Power reduction	40 %
	43.24	<=	43.24 TON	60	MWH of hydroger	1440
	43.24	<=	43.24 TON	33.3	MWH to Ton	33.3
	0	<=	0 TON	1.8	Ton pr unit	43.24
	Production target					
	170	>=	170 TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
30.jan Monday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.78	kr 1 748 416	kr 150 000	1	kr 150 000
31.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.62	kr 1 495 088	kr 150 000	1	kr 150 000
01.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.69	kr 1 663 795	kr 150 000	1	kr 150 000
02.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.72	kr 1 727 150	kr 150 000	1	kr 150 000
03.feb Friday	43.24	0	33.3	-	60 %	-	-	0	0.82	kr -	kr 150 000	0	kr -

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
30.jan Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.27	kr 502 886	kr 150 000	1	kr 150 000
31.jan Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.49	kr 921 045	kr 150 000	1	kr 150 000
01.feb Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.44	kr 838 583	kr 150 000	1	kr 150 000
02.feb Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.57	kr 1 066 532	kr 150 000	1	kr 150 000
03.feb Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.54	kr 1 020 301	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
30.jan Monday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.27	kr 595 771	kr 150 000	1	kr 150 000
31.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.49	kr 1 171 352	kr 150 000	1	kr 150 000
01.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.44	kr 1 066 480	kr 150 000	1	kr 150 000
02.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.57	kr 1 356 378	kr 150 000	1	kr 150 000
03.feb Friday	43.24	0	33.3	-	60 %	-	-	0	0.54	kr -	kr 150 000	0	kr -

THE MODEL			SCENARIO 1			SCENARIO 2		
Power expenses	kr	6 634 449.35	Power expenses	kr	4 349 346.30	Power expenses	kr	4 189 981.82
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**7 234 449.35**	**Production cost**	**kr**	**5 099 346.30**	**Production cost**	**kr**	**4 789 981.82**

Week 5	Capacity restrictions			Hour		Day
Seq2Seq	43.24	<=	43.24 TON	100	MWH	2400
	43.24	<=	43.24 TON	40 %	Power reduction	40 %
	43.24	<=	43.24 TON	60	MWH of hydroger	1440
	0	<=	0 TON	33.3	MWH to Ton	33.3
	40.28	<=	43.24 TON	1.8	Ton pr unit	43.24
	Production target					
	170	>=	170 TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
30.jan Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.30	kr 709 147	kr 150 000	1	kr 150 000
31.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.50	kr 1 189 111	kr 150 000	1	kr 150 000
01.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.45	kr 1 081 599	kr 150 000	1	kr 150 000
02.feb Thursday	43.24	0	33.3	-	60 %	-	-	0	0.60	kr -	kr 150 000	0	kr -
03.feb Friday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.56	kr 1 250 338	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
30.jan Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.27	kr 502 886	kr 150 000	1	kr 150 000
31.jan Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.49	kr 921 045	kr 150 000	1	kr 150 000
01.feb Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.44	kr 838 583	kr 150 000	1	kr 150 000
02.feb Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.57	kr 1 066 532	kr 150 000	1	kr 150 000
03.feb Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.54	kr 1 020 301	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
30.jan Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.27	kr 639 552	kr 150 000	1	kr 150 000
31.jan Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.49	kr 1 171 352	kr 150 000	1	kr 150 000
01.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.44	kr 1 066 480	kr 150 000	1	kr 150 000
02.feb Thursday	43.24	0	33.3	-	60 %	-	-	0	0.57	kr -	kr 150 000	0	kr -
03.feb Friday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.54	kr 1 208 756	kr 150 000	1	kr 150 000

THE MODEL			SCENARIO 1			SCENARIO 2		
Power expenses	kr	4 230 194.02	Power expenses	kr	4 349 346.30	Power expenses	kr	4 086 140.66
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**4 830 194.02**	**Production cost**	**kr**	**5 099 346.30**	**Production cost**	**kr**	**4 686 140.66**

Optimization week 6 ARMAX-GARCH and Seq2Seq-LSTM

Week 6	**Capacity restrictions**				**Hour**		**Day**
ARMAX-GARCH	43.24	<=	43.24	TON	100	MWH	2400
	43.24	<=	43.24	TON	40 %	Power reduction	40 %
	43.24	<=	43.24	TON	60	MWH of hydroger	1440
	40.28	<=	43.24	TON	33.3	MWH to Ton	33.3
	0	<=	0	TON	1.8	Ton pr unit	43.24
	Production target						
	170	>=	170	TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
06.feb Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.50	kr 1 197 030	kr 150 000	1	kr 150 000
07.feb Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.56	kr 1 342 459	kr 150 000	1	kr 150 000
08.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.53	kr 1 281 024	kr 150 000	1	kr 150 000
09.feb Thursday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.64	kr 1 425 604	kr 150 000	1	kr 150 000
10.feb Friday	43.24	0	33.3	-	60 %	-	-	0	0.69	kr -	kr 150 000	0	kr -

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
06.feb Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.28	kr 524 586	kr 150 000	1	kr 150 000
07.feb Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.25	kr 463 447	kr 150 000	1	kr 150 000
08.feb Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.21	kr 402 120	kr 150 000	1	kr 150 000
09.feb Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.19	kr 361 172	kr 150 000	1	kr 150 000
10.feb Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.22	kr 410 989	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
06.feb Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.28	kr 667 150	kr 150 000	1	kr 150 000
07.feb Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.25	kr 589 396	kr 150 000	1	kr 150 000
08.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.21	kr 511 402	kr 150 000	1	kr 150 000
09.feb Thursday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.19	kr 427 882	kr 150 000	1	kr 150 000
10.feb Friday	43.24	0	33.3	-	60 %	-	-	0	0.22	kr -	kr 150 000	0	kr -

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	5 246 117.30	Power expenses	kr	2 162 313.30	Power expenses	kr	2 195 829.75
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**5 846 117.30**	**Production cost**	**kr**	**2 912 313.30**	**Production cost**	**kr**	**2 795 829.75**

Week 6	**Capacity restrictions**				**Hour**		**Day**
Seq2Seq	0	<=	0	TON	100	MWH	2400
	40.28	<=	43.24	TON	40 %	Power reduction	40 %
	43.24	<=	43.24	TON	60	MWH of hydroger	1440
	43.24	<=	43.24	TON	33.3	MWH to Ton	33.3
	43.24	<=	43.24	TON	1.8	Ton pr unit	43.24
	Production target						
	170	>=	170	TON			

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
06.feb Monday	43.24	0	33.3	-	60 %	-	-	0	0.31	kr -	kr 150 000	0	kr -
07.feb Tuesday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.27	kr 610 526	kr 150 000	1	kr 150 000
08.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.23	kr 558 678	kr 150 000	1	kr 150 000
09.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.21	kr 498 683	kr 150 000	1	kr 150 000
10.feb Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.24	kr 572 837	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
06.feb Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.28	kr 524 586	kr 150 000	1	kr 150 000
07.feb Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.25	kr 463 447	kr 150 000	1	kr 150 000
08.feb Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.21	kr 402 120	kr 150 000	1	kr 150 000
09.feb Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.19	kr 361 172	kr 150 000	1	kr 150 000
10.feb Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.22	kr 410 989	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
06.feb Monday	43.24	0	33.3	-	60 %	-	-	0	0.28	kr -	kr 150 000	0	kr -
07.feb Tuesday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.25	kr 549 049	kr 150 000	1	kr 150 000
08.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.21	kr 511 402	kr 150 000	1	kr 150 000
09.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.19	kr 459 326	kr 150 000	1	kr 150 000
10.feb Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.22	kr 522 681	kr 150 000	1	kr 150 000

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	2 240 723.70	Power expenses	kr	2 162 313.30	Power expenses	kr	2 042 456.61
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**2 840 723.70**	**Production cost**	**kr**	**2 912 313.30**	**Production cost**	**kr**	**2 642 456.61**

Optimization week 7 ARMAX-GARCH and Seq2Seq-LSTM

Week 7
ARMAX-GARCH

Capacity restrictions				**Hour**		**Day**
0	<=	0 TON		100	MWH	2400
40.28	<=	43.24 TON		40 %	Power reduction	40 %
43.24	<=	43.24 TON		60	MWH of hydroger	1440
43.24	<=	43.24 TON		33.3	MWH to Ton	33.3
43.24	<=	43.24 TON		1.8	Ton pr unit	43.24
Production target						
170	>=	170 TON				

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
13.feb Monday	43.24	0	33.3	-	60 %	-	-	0	0.60	kr -	kr 150 000	0	kr -
14.feb Tuesday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.58	kr 1 295 719	kr 150 000	1	kr 150 000
15.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.47	kr 1 129 115	kr 150 000	1	kr 150 000
16.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.56	kr 1 349 419	kr 150 000	1	kr 150 000
17.feb Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.54	kr 1 294 943	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
13.feb Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.24	kr 451 182	kr 150 000	1	kr 150 000
14.feb Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.23	kr 440 803	kr 150 000	1	kr 150 000
15.feb Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.23	kr 426 273	kr 150 000	1	kr 150 000
16.feb Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.24	kr 448 163	kr 150 000	1	kr 150 000
17.feb Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.27	kr 504 395	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
13.feb Monday	43.24	0	33.3	-	60 %	-	-	0	0.24	kr -	kr 150 000	0	kr -
14.feb Tuesday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.23	kr 522 222	kr 150 000	1	kr 150 000
15.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.23	kr 542 119	kr 150 000	1	kr 150 000
16.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.24	kr 569 957	kr 150 000	1	kr 150 000
17.feb Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.27	kr 641 472	kr 150 000	1	kr 150 000

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	5 069 195.95	Power expenses	kr	2 270 815.80	Power expenses	kr	2 275 770.62
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**5 669 195.95**	**Production cost**	**kr**	**3 020 815.80**	**Production cost**	**kr**	**2 875 770.62**

Week 7
Seq2Seq

Capacity restrictions				**Hour**		**Day**
40.28	<=	43.24 TON		100	MWH	2400
43.24	<=	43.24 TON		40 %	Power reduction	40 %
43.24	<=	43.24 TON		60	MWH of hydroger	1440
43.24	<=	43.24 TON		33.3	MWH to Ton	33.3
0	<=	0 TON		1.8	Ton pr unit	43.24
Production target						
170	>=	170 TON				

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
13.feb Monday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.27	kr 593 089	kr 150 000	1	kr 150 000
14.feb Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.26	kr 620 353	kr 150 000	1	kr 150 000
15.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.25	kr 597 315	kr 150 000	1	kr 150 000
16.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.26	kr 632 113	kr 150 000	1	kr 150 000
17.feb Friday	43.24	0	33.3	-	60 %	-	-	0	0.30	kr -	kr 150 000	0	kr -

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
13.feb Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.24	kr 451 182	kr 150 000	1	kr 150 000
14.feb Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.23	kr 440 803	kr 150 000	1	kr 150 000
15.feb Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.23	kr 426 273	kr 150 000	1	kr 150 000
16.feb Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.24	kr 448 163	kr 150 000	1	kr 150 000
17.feb Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.27	kr 504 395	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
13.feb Monday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.24	kr 534 518	kr 150 000	1	kr 150 000
14.feb Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.23	kr 560 598	kr 150 000	1	kr 150 000
15.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.23	kr 542 119	kr 150 000	1	kr 150 000
16.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.24	kr 569 957	kr 150 000	1	kr 150 000
17.feb Friday	43.24	0	33.3	-	60 %	-	-	0	0.27	kr -	kr 150 000	0	kr -

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	2 442 870.02	Power expenses	kr	2 270 815.80	Power expenses	kr	2 207 192.15
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**3 042 870.02**	**Production cost**	**kr**	**3 020 815.80**	**Production cost**	**kr**	**2 807 192.15**

Optimization week 8 ARMAX-GARCH and Seq2Seq-LSTM

Week 8
ARMAX-GARCH

Capacity restrictions

43.24	<=	43.24 TON
43.24	<=	43.24 TON
0	<=	0 TON
43.24	<=	43.24 TON
40.28	<=	43.24 TON
Production target		
170	>=	170 TON

Hour		Day
100	MWH	2400
40 %	Power reduction	40 %
60	MWH of hydroger	1440
33.3	MWH to Ton	33.3
1.8	Ton pr unit	43.24

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
20.feb Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.47	kr 1 118 556	kr 150 000	1	kr 150 000
21.feb Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.54	kr 1 287 023	kr 150 000	1	kr 150 000
22.feb Wednesday	43.24	0	33.3	-	60 %	-	-	0	0.59	kr -	kr 150 000	0	kr -
23.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.51	kr 1 232 548	kr 150 000	1	kr 150 000
24.feb Friday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.55	kr 1 220 158	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
20.feb Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.27	kr 502 319	kr 150 000	1	kr 150 000
21.feb Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.56	kr 1 047 285	kr 150 000	1	kr 150 000
22.feb Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.24	kr 457 031	kr 150 000	1	kr 150 000
23.feb Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.24	kr 453 257	kr 150 000	1	kr 150 000
24.feb Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.26	kr 486 091	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
20.feb Monday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.27	kr 638 832	kr 150 000	1	kr 150 000
21.feb Tuesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.56	kr 1 331 900	kr 150 000	1	kr 150 000
22.feb Wednesday	43.24	0	33.3	-	60 %	-	-	0	0.24	kr -	kr 150 000	0	kr -
23.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.24	kr 576 437	kr 150 000	1	kr 150 000
24.feb Friday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.26	kr 575 875	kr 150 000	1	kr 150 000

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	4 858 284.85	Power expenses	kr	2 945 984.40	Power expenses	kr	3 123 044.05
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**5 458 284.85**	**Production cost**	**kr**	**3 695 984.40**	**Production cost**	**kr**	**3 723 044.05**

Week 8
Seq2Seq

Capacity restrictions

40.28	<=	43.24 TON
0	<=	0 TON
43.24	<=	43.24 TON
43.24	<=	43.24 TON
43.24	<=	43.24 TON
Production target		
170	>=	170 TON

Hour		Day
100	MWH	2400
40 %	Power reduction	40 %
60	MWH of hydroger	1440
33.3	MWH to Ton	33.3
1.8	Ton pr unit	43.24

MODEL	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Predicted powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
20.feb Monday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.30	kr 660 826	kr 150 000	1	kr 150 000
21.feb Tuesday	43.24	0	33.3	-	60 %	-	-	0	0.58	kr -	kr 150 000	0	kr -
22.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.27	kr 645 792	kr 150 000	1	kr 150 000
23.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.27	kr 640 032	kr 150 000	1	kr 150 000
24.feb Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.29	kr 687 788	kr 150 000	1	kr 150 000

SCENARIO 1	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
20.feb Monday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.27	kr 502 319	kr 150 000	1	kr 150 000
21.feb Tuesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.56	kr 1 047 285	kr 150 000	1	kr 150 000
22.feb Wednesday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.24	kr 457 031	kr 150 000	1	kr 150 000
23.feb Thursday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.24	kr 453 257	kr 150 000	1	kr 150 000
24.feb Friday	43.24	34	33.3	1 132	60 %	1 887	1 887 000	1	0.26	kr 486 091	kr 150 000	1	kr 150 000

SCENARIO 2	Max. cap. TON	Prod. in TON	TON to MWH	H2 in MWH	Stack efficiency	Required MWH	Required KWH	Production-day	Acutal powerp.	Power ex.	Transport ex	Dummy if transp.	Varible trans ex
20.feb Monday	43.24	40.28	33.3	1 341	60 %	2 236	2 235 540	1	0.27	kr 595 101	kr 150 000	1	kr 150 000
21.feb Tuesday	43.24	0	33.3	-	60 %	-	-	0	0.56	kr -	kr 150 000	0	kr -
22.feb Wednesday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.24	kr 581 236	kr 150 000	1	kr 150 000
23.feb Thursday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.24	kr 576 437	kr 150 000	1	kr 150 000
24.feb Friday	43.24	43.24	33.3	1 440	60 %	2 400	2 399 820	1	0.26	kr 618 194	kr 150 000	1	kr 150 000

THE MODEL			**SCENARIO 1**			**SCENARIO 2**		
Power expenses	kr	2 634 437.59	Power expenses	kr	2 945 984.40	Power expenses	kr	2 370 967.55
Transportation expenses	kr	600 000.00	Transportation ex	kr	750 000.00	Transportation ex	kr	600 000.00
Production target	**kr**	**3 234 437.59**	**Production cost**	**kr**	**3 695 984.40**	**Production cost**	**kr**	**2 970 967.55**

Norwegian University of
Science and Technology