Norwegian University of Science and Technology

Master's thesis

Intraday bidding optimization for a Nordic hydropower producer using fundamental drivers to forecast the intraday market

Cabrol, Julien

Master of Energy and Environmental Engineering
Submission date: 14.06.2021
Supervisor: Klæboe, Gro
Department of Electric Power Engineering
Norwegian University of Science and Technology
Contact: julien.cabrol7@gmail.com

Abstract

This thesis investigates the potential profits of a hydropower producer participating in the intraday market. It explores if forecasting the intraday market using fundamental drivers and machine learning can improve the intraday bidding. The study is motivated by the intraday market in NO3 recently seeing increased volumes. Recent studies have also found relations between the intraday market and fundamental drivers, while others have started building tools to predict the intraday prices based on this.

An optimization and simulation framework is proposed for the intraday bidding and hydropower scheduling problem. The framework uses a rolling-horizon approach to go through the intraday market from it opens until it closes. At each hourly time-step, a two-stage stochastic mixed-integer program will accept profitable limit orders from the real order book, considering the production plan, the water value and future trading opportunities. Scenarios of future trading opportunities are generated by forecasting the intraday premium, volume and occurrence of trades for each product. The forecasting is done with random forest regression or neural networks, and uses fundamental drivers as input variables.

For a case study with a hydropower producer in the bidding zone NO3, and 256 days in 2020, the benefit of participating in the intraday market converges to around 3 % for the bidding model without forecasting. The bidding model with forecasting does on average outperform the bidding model without forecasting. However, more data and testing is needed to reach a conclusion on the performance of this model. This uncertainty is mostly due to the different performance of the forecasting methods under different market conditions. Testing the bidding model outside of the abnormal year 2020, and development of the forecasting methods is therefore identified as the most important improvements to obtain more reliable results.

Sammendrag

Denne masteroppgaven undersøker den potensielle fortjenesten for en vannkraftprodusent fra NO3 som deltar i intradagsmarkedet. Oppgaven utforsker om forecasting av intradagsmarkedet ved hjelp av markedsdrivere og maskinlæring kan forbedre budgivningen. Studien er motivert av at intradagsmarkedet for NO3 nylig har sett økte volumer. Nyere studier har også funnet sammenhenger mellom intradagsmarkedet og markedsdrivere, mens andre har begynt å bygge verktøy for å forecaste intradagsprisene basert på markedsdriverne.

Et optimaliserings- og simuleringsrammeverk for intradagshandel og produksjonsplanlegging er utviklet. Rammeverket bruker en rolling-horizon approach til å gå gjennom intradagsmarkedet fra det åpner til det stenger time for time. Hver time vil et to-trinns stokastisk blandet heltallsproblem akseptere lønnsomme ordre fra den faktiske ordreboka, basert på produksjonsplanen, vannverdien og fremtidige handelsmuligheter. Scenarier for fremtidige handelsmuligheter genereres ved å forecaste intradagspremiumen, volumet og forekomsten av handler for hvert produkt. Forecastingen utføres med random forest regresjon eller nevrale nettverk, og bruker markedsdrivere som inputvariabler.

For en casestudie med en vannkraftprodusent i NO3, og 256 dager i 2020, konvergerer fortjenesten ved å delta i intradagsmarkedet til rundt 3 % for budmodellen uten forecasting. For budmodellen med forecasting er det behov for mer arbeid og data for å konkludere på ytelsen av modellen. Dette er til tross for at denne modellen i gjennomsnitt presterer bedre enn budgivningsmodellen uten forecasting for de testede dagene i 2020. Denne usikkerheten skyldes for det meste den forskjellige ytelsen til forecastingsmodellene under forskjellige markedsforhold. Testing av budmodellen utenfor det unormale året 2020, og utvikling av forecastingsmetodene blir derfor identifisert som de viktigste forbedringene for å gjøre resultatene mer pålitelige.

Contents

Chapter 1

Introduction

This thesis is a continuation of the work done in the specialization project [1]. Below is the introduction from the specialization project with changes that reflect the modelling improvements made in the master's thesis.

The last few years has seen increased volumes in the intraday market for NO3 (see 2.1). This is tied to increased penetration of wind power in the Nordics. The wind power producers rely on forecasts to bid in the energy markets, and because of the uncertainty of the forecasts, the post-spot markets are needed to make sure their commitments matches their actual production. Another reason for the increased volumes is that the intraday markets for the Nordics have recently been coupled with other European intraday markets through Single Intraday Coupling (SIDC) [2]. Some of these markets have high penetration of variable renewable energy, and therefore higher volumes in their intraday markets. For flexible hydropower producers in NO3, this means more opportunities to supply balancing services and take advantage of the different prices in the markets. The reservoirs and the flexible production units make them able to wait for the better prices and respond quickly when balancing services are needed. The ability to change production plans close to delivery or in real-time, is a scarce resource that is becoming more and more important in this market setting with a lot of variable production. Offering balancing services should in principle therefore be more lucrative than just selling power in the day-ahead market. This thesis focuses on modelling the intraday market, and the balancing market is left for future work.

Since the spot market has been, and still is the dominant power market, bidding strategies for the intraday market is not much developed. The literature on intraday bidding optimization for hydropower producers is lacking and the few models [3–6] that exist found little benefit of trading in the intraday market. However, there are different types of models [7, 8] which have found benefits and therefore shows a potential for the intraday market. The first is an optimization model with full foresight and the latter simulates through the real order book, but also uses optimization to develop bidding curves that can be used as decision support. Other recent simulation models in the literature that uses an order book is [9, 10]. The benefit of optimization is that it handles the temporal structure and the combination of the resource and bidding problem well, while the benefit of simulating through an order book is that it imitates the real structure of the continuous order book and that actual market prices can be used. Another recent development is studies on how fundamental drivers impacts the intraday price. Wind power, demand and spot prices, to mention some, has by numerous studies [11–17] been found to be fundamental drivers for the intraday price. Forecasting the intraday price based on these variables using machine learning methods [18–22] is also a recent development in the literature that could be adapted to improve the decision making in bidding problems.

The specialization project [1] preceding this thesis developed a simulation and optimization framework that combines hydropower scheduling, intraday bidding in the real order books and intraday forecasting with fundamental drivers. This combination was a gap in the literature, with the main challenges being combining optimization and simulation in a framework, and linking the intraday forecasting to the bidding in the order books. Also, since most of the literature on the intraday market covers price modelling, volume modelling was an important gap to fill, both for forecasting and use in an intraday bidding model. To predict and restrict the volume of orders in the future intraday market scenarios should give a more realistic representation of the future trading opportunities.

In the specialization project, this approach for a case study in NO3 found it profitable to participate in the intraday market, but the main forecasting model did not improve the intraday bidding. The goal of this thesis is therefore to continue this work by improving the bidding and forecasting frameworks, to be able to better evaluate the benefit of participating in the intraday market when using fundamental drivers to forecast it. The main improvements in this thesis are the spot market and order book modelling, the general forecasting framework, and the shift to use machine learning methods for the forecasting. The findings of this thesis suggest that more data and testing is needed to conclude on the performance of the bidding model with forecasting, despite the main model on average outperforms the bidding model without forecasting for the days tested in 2020. This is mostly due to the uncertain performance of the forecasting methods in different market conditions.

The rest of the paper is organized as follows. Section 2 gives the theoretical background for this study. Section 3 presents modelling of the bidding problem, while 4 covers the elements relevant to forecasting. Section 5 gives an overview of the case study and problems that will be solved. Section 6 contains the results and discussions, while section 7 gives the conclusion.

Chapter 2

Background

2.1 Hydropower scheduling

This thesis will cover reservoir hydropower. This is a unique technology since the variable production cost is almost zero. What differentiates reservoir hydropower from renewable energy sources like wind and solar is that one also can decide when to produce the power. The cost of production can therefore be represented by an opportunity cost, the expected marginal value of having an extra unit of water in the reservoir, which is called the water value. The limitations of the power plant that will impact the water value is the production capacity and efficiency of the production unit, and the size of the reservoir. The external factors are the uncertain market prices and reservoir inflow, and also environmental constraints. The hydropower producer wants to trade power at good prices, produce at an efficient production level, and without spilling water. A watercourse can also have several production units and reservoirs. Then these elements are connected and will impact each other. This thesis will cover a simple watercourse with only one production unit, one reservoir and no environmental constraints.

With the flexibiliy of the reservoir and production unit, and the generally low ramping/start-up costs, hydropower can store water for when the market needs it the most. This should coincide with selling power at the best possible prices. Long-term hydropower wants to adjust to supply and demand trends in the market - medium-term it wants to conserve water between yearly spring floods - and short term it wants to take advantage of for example demand spikes or the uncertain production of renewable energy sources. For the short-term, buying power is also an alternative, and the power producer can reduce its commitments by buying power if the prices are lower than the cost of the already sold power. The time resolution and level of detail in the modelling will increase towards shorter term problems. Long-term, seasonal or short-term models are usually also coupled in a hierarchy. The long term model will find watervalues, which can be used by the seasonal model to calculate new watervalues, which can again be used by the short term model. This is called price coupling, while other alternatives are volume or demand coupling [23]. At the end we have a watervalue for a given day that will depend on the level of the reservoir and can be given as cuts. For a hydropower plant with a large reservoir, using a constant daily watervalue is a valid assumption since production will not affect the water level that much. The final goal of short term hydropower scheduling, which this thesis focuses on, is to optimize the revenue of the power sold to and bought from the different markets, minus the value of the water resources used and the costs of production and trading.

2.2 Power markets

The objective of the Nordic power market is to be a competitive market where the market participants can decide their own dispatch by bidding in the markets. The wholesale power markets is where power producers trade their physical power. In the Nordic setting it consists of the day-ahead, intraday and balancing markets. These markets are connected to power delivery at the same hours, but the opening hours, market mechanisms and delivered services differs between the markets.

2.2.1 Day-ahead market

In the day-ahead market, which is a daily auction also referred to as the spot market, participants can buy or sell power deliveries for the next day by bidding on hourly products. Complex bid combinations of hourly products are possible. The Nordic system is part of Single Day-Ahead Coupling (SDAC), which by the time of writing consists of most European countries [24]. Market participants are divided into geographical bidding zones where the power flow between the zones are restricted by transmission capacities decided by the transmission system operators (TSOs). When the market is cleared, bids on exchanges in all zones are matched by taking into account the transmission capacities between zones. Prices, volumes and flows are decided for each hour and bidding zone. The goal of SDAC is to maximize the social welfare of the whole system, and the hourly price in each zone is therefore the marginal cost of power in the zone. At Nord Pool, participants can bid up to 12:00 CET/CEST the day before delivery, and the result of the market clearing is published at 12:42 or later [25].

2.2.2 Intraday market

As electricity on the power grid is not storable, there has to be equilibrium between production and consumption to keep the system stable. With uncertainty in both production and consumption, it is necessary with possibilities for market participants to adjust their commitments after the day-ahead market is cleared. The first opportunity for rebalancing comes in the intraday market. The background section on the intraday market from the specialization project [1] preceding this thesis, is presented below with some small changes and additions.

The intraday market a Norwegian power producer can participate in has historically been run by Nord Pool. It is called Elbas and uses a continuous double auction mechanism. For a given hourly product, market participants can make buy and sell orders at chosen prices and volumes, which are continuously either matched with previous orders or stored in the order book to be evaluated against future orders. Nord Pool today offers limit orders, block orders for consecutive production hours, iceberg orders, fill-or-kill orders and immediate-or-cancel orders [26]. Similar order types can be found at EPEX Spot, which launched for the Nordic region in May 2020 [27].

In 2018, Single Intraday Coupling (SIDC) was launched in 14 countries in the Nordics and Central Europe, and then expanded to 22 countries in 2019 [2]. This initiative gave cross-border intraday trading to these countries by implementing a shared order book. With this solution, the order book for a bidding zone will show all the orders from different power exchanges and zones as long as there is available transmission and ramping capacity for the order to be delivered. Normally the transmission capacity is first given when the intraday market opens, and then updated continuously as the market develops. For some transmission borders the capacity can be allocated explicitly or through capacity auctions as changes to the system happens.

The opening times of the intraday market varies between the bidding zones and products. For a Norwegian participant, one can start trading a product the day before delivery at 14:00 CET/CEST, until one hour before delivery. At first one can trade within the bidding zone - and from 15:00 with the Nordics, Baltics and Poland, from 18:00 with Germany, from 21:00 with the Netherlands and from 22:00 with the remaining countries [28].

2.2.3 Other markets

The other markets are less relevant for this thesis. Balancing markets are the last opportunity to achieve equilibrium in the supply and demand of power. Balancing is done in real-time by the TSO, which has the responsibility to pre-acquire enough of the different balancing services to ensure stability of the system. The last markets are the financial markets. In these markets participants can manage risk by trading long or short-term financial products without physical delivery.

2.2.4 NO3 market volumes

This section will give insight into the development of the intraday market and why this market could become a more important part of the Norwegian power market than it is now. The work from the specialization project [1] is presented below with some major changes and additions.

Table 2.1 shows the development of the day-ahead, intraday and tertiary balancing market volumes in GWh for NO3, which is the bidding zone relevant for this thesis. The increased intraday volumes in 2019 could be tied to the doubling of wind production in NO3 from 2018 to 2019 [29]. More wind power, which is variable and hard to forecast precisely should increase the need for flexibility in NO3. The volume increase in 2019 could also be linked to the launch of Single Intraday Coupling (SIDC) [2], which made it easier to trade with other bidding zones and power exchanges in the intraday market. NO3 which is hydropower dominated and therefore very flexible could offer its flexibility in the intraday market to less flexible bidding zones or to bidding zones with more developed intraday markets. The intraday marked volumes for NO3 are still only around 1% (ca. 0.29 TWh vs 27.4 TWh) of the total volumes if regulating power is not considered. In comparison, the number for the Nordic countries is 2.3% (8.2 vs 352) [30]. For the German market, which in addition has intraday auctions and quarter-hourly trading, the number is around 19% (53.7 vs 226.4) [31]. In these calculations the volumes are not double counted if both the buyer and seller is in the same area. Germany is an interesting market since they with marked design managed to reduce the volume in the balancing market even when the renewable energy share increased [32]. Koch and Hirth [32] found that 17% of the balancing energy decrease could be explained by quarter-hourly products. Quarter-hourly products is coming to the Nordic markets around 2023 for the intraday and tertiary balancing market [33]. This shows that future market design could also be an important driver for the intraday market. Another driver that could accelerate the use of the intraday market is higher transmission capacities from the Nordics to Europe, because of new sea interconnectors [34].

	2015	2016	2017	2018	2019	2020
Buy DAM	21751	25706	25166	26970	27387	26742
Sell DAM	14766	19814	21547	19086	22179	26080
Buy IDM	31.8	46.8	51.0	52.0	109.4	130.0
Sell IDM	88.5	67.9	70.4	81.9	182.1	143.5
Down BM	110.3	221.1	329.8	288.6	397.6	543.6
Up BM	62.3	91.4	108.9	87.8	119.9	113.6

Table 2.1: Day-ahead (DAM), intraday (IDM) and tertiary balancing market (BM) volumes in GWh for NO3. Source Nord Pool.

The analysis in Appendix A on the Nord Pool ticker data shows who NO3 trades with in the intraday market from 2018 to 06.2020. Most notably most of the trades are with other bidding zones and at much better prices than the spot price, both when selling and buying power. This reinforces the notion that there is a need for flexibility in the intraday market of other bidding zones and that the intraday market is a good opportunity for participants in NO3 to make more profits. The big contrast between the high profitability and the low volumes shows that either participants in NO3 lack the tools to trade in the intraday market in a cost effective way, or that bottlenecks in the transmission system limits the opportunities to trade. Also, Germany and Denmark being the most profitable areas to trade with shows a potential for when other areas increase their variable production, as these are areas with high amounts of renewable energy [35].

2.3 Related litterature

In this section the relevant literature on the intraday bidding problem will be presented. It will cover the literature that involves reservoir hydropower and optimization, but mention other types of models when relevant. It will also cover relevant price modelling and prediction methods used in the literature. A review of the state of art and related work was carried out in the specialization project [1]. This is amended with new papers on price forecasting that have become available since, and more details on price forecasting as this is more relevant for this thesis than in the specialization project.

2.3.1 Intraday bidding problem

The intraday bidding problem for a hydropower producer is not much discussed in the literature. The few papers and models that exist, models the full intraday market, but in different ways. The model of Faria and Fleten [3]) models the day-ahead and intraday markets as a two-stage stochastic mixed-integer program. The intraday stage in this paper has collapsed the whole intraday market into one trading opportunity per product. They optimize the day-ahead bidding by including scenarios of the trading possibilities in the intraday market. This bidding strategy is called coordinated bidding. The relevant challenges of coordinated bidding are covered in Aasgård *et al.* [36], which is price forecasting and scenario generation. They also emphasise that there is lost potential in not incorporating the post-spot markets in the day-ahead bidding, but that the existing models did not find significant benefits of coordinated bidding. Fodstad *et al.* [7], which compares trading in the day-ahead market with coordinated bidding with full foresight found that coordinated bidding gave 6.4-7.9% increase in profits for the German intraday market. Coordinated bidding is not a focus for this project, and will be left for future work.

The work done in [4–6] models bidding in the intraday market after day-ahead settlement. The model of Engmark and Sandven [4] is a multistage stochastic mixed-integer program. It has a rolling-horizon approach through the intraday market where the intraday scenarios are generated at the start of the day, while the balancing market scenarios are updated for each step of the rolling-horizon. The next model is Akersveen and Graabak [5], which is a multistage stochastic linear program that models the intraday market using a scenario tree. This paper also discusses and tests different modelling assumptions and choices for the problem. The last model is Bovim and Naess [6], which uses stochastic dynamic programming for the intraday problem. To reduce the problem size, they aggregate the problem into 6 stages and 4 products. Their model also have 5 price states and 4 production levels.

2.3.2 Price modelling

Price modelling forms the basis for the intraday market modelling and prediction. Weron [37] describes five different approaches for electricity price forecasting, which can be used alone or combined into hybrid methods. The paper gives a good overview over these methods, and gives examples on how these methods are used in the literature. Even though most of the examples are about predicting the spot market, they are also valid for the intraday market. The five approaches are as follows:

- Multi-agent methods: Focuses on the participants in the market, how they interact and how equilibrium between supply and demand is reached.
- Fundamental methods: Uses relations between fundamental drivers to find the price.
- Reduced-form methods: Replicates the statistical characteristics of the system, like for example price movement with spikes and volatility.
- Statistical methods: Finds relations between previous prices and system variables and uses it to predict the price.
- Computational intelligence: Uses machine-learning and -intelligence to solve problems the traditional methods struggle with.

Intraday modelling for bidding problem

In this section we will present the price, volume and bidding models used in the reviewed intraday bidding and production planning models. What most of the optimization models have in common is that they use stochastic price processes. The models of Faria and Fleten [3]), Aïd *et al.* [38], Bovim and Naess [6] and Akersveen and Graabak [5] models the price either as ARMAX processes, Brownian motions or Markov processes. The only optimization model that do not use a stochastic price process is the model of Engmark and Sandven [4]. They instead use the historical intraday premium to volume data to make price scenarios.

When it comes to the volume and bid acceptance, the models also differ. In some of the models, only the price is modelled and the volume a participant bids into the market will be accepted and will only be restrained by the resource and cost problems of the full models. This is the case for [6, 38]. [4, 5] on the other hand, have some sort of acceptance rate which gives the chance of the bid being accepted, while [3] restrict the volume as a percentage of the capacity of the plant. [4, 5] also models the price as an order book, where the participants can bid at different price levels.

The second type of models that have been reviewed have very different price models, and these papers all uses simulations instead of optimization. The first model, Dideriksen *et al.* [8], simulates through each new arrival of orders in the intraday order books. At each point in time it uses the marginal cost curve of the hydropower plant to decide what orders to accept. They found profits of 2 % compared to only considering the day-ahead market. The two last models uses fundamentals combined with statistics or machine learning to predict the price. Koch [9] uses linear regression to predict the balancing market from information in the power system. They use for example the time of day and wind forecast deviations, and if the predicted balancing market price is worse than some prices in the order book, these orders are accepted. The last model, Bertrand and Papavasiliou [10], uses machine learning to make decisions about what orders to accept in a point in time from the order books of the different production hours. Variables for the machine learning are the storage level, time to closure and flexibility of the market from the dayahead curve.

The use of fundamental drivers, statistics and intelligent price prediction methods, is also something that has been seen more in general intraday market analysis in recent years. With higher amounts of intermittent power in the system, the dynamics of these sources might start to dominate the market - a market that previously had low liquidity and was hard to predict due to the nature of unforeseen power events. The development presented below supports the inclusion of fundamentals into the price processes of optimization models.

Intraday modelling based on fundamental drivers

The goal of this section is to cover the development in the field of intraday modelling. The literature only covers price modelling, so modelling of the intraday volume is identified as a huge gap in the literature. The section will cover how fundamental drivers together with statistics and computational intelligence is used to describe or predict the intraday price. What parameters and exogenous variables that are used in the different methods, will also be mentioned. Only the intraday part of the papers presented are mentioned in this section. Shinde and Amelin [39], which gives an overview of the intraday market, points to many of the papers mentioned below.

The first fundamental models looks at the volume weighted intraday price or premium of each production hour and tries to describe it using some sort of regression models. The simplest models doing this are Hagemann [11], Kiesel and Paraschiv [12], Karanfil and Li [13], Soysal *et al.* [14], Ziel [15] and Hu *et al.* [16]. These models often uses present and previous/lagged market or exogenous variables to describe the intraday price. Using previous intraday price information to predict the future intraday price makes most of them autoregressive models. The exogenous variables they use are often the demand, and actual and forecasted production from different sources. The analysis is often evaluating the coefficients the regression model gives for each variable to find the impact the variable has on the intraday price. One example of such a model is Hu *et al.* [16]. They analyse how the wind, non-wind and load forecast errors, together with outages and intraday transmission capacity impacts the intraday price premium in the Swedish market areas. The key findings in this paper is that the wind forecast error impacts the premium, and that the unplanned power plant outages does not. Berger *et al.* [17] also uses a similar model to the previous ones, but instead of evaluating the coefficients of the regression variables, they evaluate the regression model performance on an out-of-sample dataset in terms of mean squared error. The authors highlights the changes in demand and onshore wind infeed as intraday price drivers in the German intraday market.

Another type of papers builds price prediction models with different types of methods and fundamental variables. These papers often uses more complex methods, with preprocessing, variable selection, out-of-sample testing, ensemble methods and comparing the results with benchmark forecasting methods. The first two papers, Monteiro *et al.* [40] and Andrade *et al.* [18], are on the Iberian intraday market which at the time had 6 intraday auctions throughout the day. They respectively use machine learning and regression, and for each auction they forecasts prices for all open products. Their key findings are that the best intraday price prediction models only used previous day-ahead and intraday prices as variables. This is interesting considering they included variables that other studies have found to be price explanatory, like actual generation for different sources, demand and weather, and their forecasts. Janke and Steinke [20], which forecasts the volume weighted price distribution for the German market using quantile regression and neural networks, also found the neighbouring prices to be the best variables. This is in contrast to the Lasso regression model of Marcjasz *et al.* [41] for the German intraday market and the deep learning methods of the next papers. They all use several exogenous variables in their best models. Kolberg and Waage [19] compares different forecasting methods on the Swedish intraday market. They use images of weather forecasts, day-ahead, intraday and regulating prices, transmission capacities, urgent market messages and time dummies. The deep learning methods in this paper outperforms the other methods and the benchmark model by 12-25%, but a breakdown of the most important variables was not given. Other papers where the deep learning models performs the best are Oksuz and Ugurlu [21] and Scholz *et al.* [22], which are for the Turkish and German intraday markets. The last model we will mention is the unique model of Kulakov and Ziel [42] as it shows the importance of renewable energy in the intraday market. This model shifts the day-ahead market curve based on wind and solar forecast errors to get the intraday price. This was done for the German market and it outperformed the regression benchmark models in the paper.

Other details that are relevant for the choices made in this thesis will be mentioned below. Firstly most of the papers on continuous intraday markets predicts one volume weighted price per product. Either for the whole product [21] or for the X last hours of each product [19, 20, 41]. The benefit of only predicting the last hours of a product is that one can use previous prices from the day in the prediction. This is done by [19, 20, 22, 41]. Differently from the other models, [22] models quarter-hourly intervals of the last 4 hours before production. It is also the only model that inputs times series of the previous development of the product before predicting it, that uses order book data, and that uses frequently updated wind production forecasts instead of the day-ahead forecast. A last detail is that [21] finds that predicting the difference between the spot price and intraday price gives much better results than directly forecasting the intraday price. That approach is therefore used in this thesis.

Generally in these papers, the order of worst to best performance in terms of mean squared error or mean average error, are regression methods and then different machine learning methods. The order of the machine learning methods are random forest and gradient boosting, simpler neural networks like multi-layer perceptrons, and at last recurrent neural networks like long short-term memory or gated recurrent unit. The reason why the recurrent neural networks performs the best is that they can directly model the temporal structure of the data, which is important for time series data. This indicates that the dynamics in the intraday market are very complex, and that non-linear and intelligent methods are needed to more accurately predict it. In this thesis machine learning will therefore be used for the intraday market forecasting, but modelling the temporal structure will be left for future work. Out of the variables discussed, the day-ahead price, wind power production and congestion variables will be used in this thesis to predict the intraday market.

Chapter 3

Bidding problem

This section will describe the framework and modelling choices made for the hydropower producer participating in the power markets.

3.1 Market modelling

This thesis will model a hydropower producer that controls one power plant and that participates in the day-ahead and intraday market. Each day will be modelled independently, which means that we ignore the fact that markets between the days are constantly overlapping. The model first finds the spot commitments day-ahead and uses this as a starting point for trading in the intraday market. The balancing market is ignored to reduce the modelling complexity, so the hydropower producer has to cover its potential imbalances before the intraday market closes (more on risk in section 3.3.2). One strength of the assumption of excluding the balancing market is that the hydropower producer usually does not need balancing services. Except for unexpected downtime, it does not have imbalances incur after the close of the intraday market, like a wind producer would. The weakness however, is that the hydropower producer would like to offer balancing services in the balancing markets. This is especially the case for the Norwegian market, where at the moment there is more liquidity in the balancing markets than in the intraday market (see table 2.1). Regardless, modelling of all markets would be needed to find out what the best trading strategy over different markets is, but that is out of the scope of this thesis.

3.2 Day-ahead model

To be able to model the intraday market, the hydropower producer needs to have a starting point of commitments decided in the day-ahead market. The specialization project [1] used the actual/historical production as a starting point for the intraday market. The problem with this is that it includes the scheduling decisions happening after the spot market trading. The modelled hydropower plant is also part of the portofolio of Trønderenergi, so other power plants could impact the dispatch decisions. The day-ahead model chosen for this thesis finds the spot commitments assuming the hydropower producer has perfect foresight and therefore bids perfectly in the spot market, not considering the post-spot markets. The prices are fixed to the actual spot prices that occurred in the market and the model only decide on the sold volume for each hourly product. The first benefit of this is that the spot commitments now only are decided based on the modelled power plant and the spot market. The second benefit is that one can use the same hydropower modelling for both the spot and intraday models, so the dispatch decisions are based on the same information in both markets. A last benefit is that the intraday profit is no longer affected by the accuracy of the spot market bidding of the power producer. In reality, a part of the intraday profit could be from adjusting non-optimal spot commitments, so that aspect is ignored in this thesis.

See the day-ahead model objective function 3.1. It optimizes the revenue from selling power at different hourly products $m \in M = \{0, 23\}$ of the next day. It also optimizes the water value of the reservoir at the end of the day and minimizes the startup costs of the production unit throughout the day. Trading costs at Nord Pool are not included in the modelling. The full spot market model can be found in appendix B.2, and the notation in appendix B.1. The general modelling choices and notation is the same as for the intraday market, so this will be covered in the section about the intraday modelling.

$$max\ z = \sum_{m \in M} Price_m^{DA} p_m + Watervalue * r_{m=23} - \sum_{m \in M^{startup}} C^{startup} a_m \tag{3.1}$$

3.3 Intraday model

The intraday bidding problem in this thesis integrates the use of the historical/real order book, forecasting, and production planning into an optimization and simulation framework. This combination is inspired by the models presented in the literature review that combines some of these elements, but as far as we know this work is the first attempt at combining all of them. Simulating through the real order book uses the actual prices in the market and gives a more realistic representation of the continuous intraday structure - forecasting should improve the bidding decisions - and an optimization framework should handle the temporal structure and the combination of the resource problem, intraday bidding and forecasting well.

With the continuously changing order books towards production and the uncertainty of the market development and inflow, the intraday bidding problem is a multi-stage stochastic programming (MSP) problem. The hydropower producer all throughout the day has to evaluate if the production schedule can be optimized by accepting orders that are in the order book now, or wait for better orders later. It also has to consider arbitrage opportunities. As the water in the reservoir and the production capacity are limited resources, accepted orders will impact the trading one can do later. The scheduling decisions also have to be made under the uncertainty of the market development and the inflow. In this thesis only the intraday market has uncertainty, while the inflow is modelled deterministically with the actual/historical inflows.

When wanting to bid in the real-time order book and using updated forecasts throughout the day, the full intraday bidding problem cannot be solved by running the MSP once. Therefore, this thesis uses a rolling horizon approach to solve a sequence of two-stage stochastic mixed integer programming (TSMIP) problems. The first stage of the TSMIP is to decide what orders to accept from the real-time order book, the trading opportunity at an exact moment, and the second stage is to decide what orders to accept from the forecasted order book scenarios, which should represent the trading opportunities for a product for the rest of the day until the product closes. The rolling horizon approach simulates through the intraday market in 32 hourly steps, from 14:00 day-ahead when the market opens, to 22:00 day-of, which is the last trading opportunity. At each hourly step of the rolling horizon approach, the model is fed the latest market data (real-time and forecasted order books), and the updated resource variables (previous commitments and initial reservoir for the first production hour) that depends on decisions taken earlier in the market. Then the TSMIP model is solved, before this procedure is repeated. A big difference between this modelling and the previous optimization models [4–6] is that we separate between the realized (real order book) and forecasted price(forecasted order book), whereas they used a price model to represent different market scenarios. Separating between realized and forecasted price was done before by the intraday simulation model of Koch [9].

Figure 3.1 shows a diagram of the different steps of the day-ahead and intraday bidding models, and the information flow between the models. Disregarding the general hydropower modeling, the only information the day-ahead model needs is the reservoir level from the day before, the spot market information, the deterministic inflow and the daily constant watervalue. After solving the problem it will send the decided commitments, and the reservoir level from the day before as inputs to the intraday model. Then the intraday model, the TSMIP, will be solved based on the initial spot commitments, initial reservoir and real-time and forecasted order books for the hour 14:00-15:00 day-ahead. The accepted orders from the real-time order book will be added to the spot commitments, and the rolling horizon will go to the next step which is 15:00-16:00 day-ahead to solve the new bidding problem. This procedure will continue until all intraday products are closed. The initial reservoir, which is for the first product/production hour, is only hourly updated as products start to close, which is after the hour 22:00-23:00 day-ahead. The bidding model only needs the initial reservoir since it calculates the reservoir development for the rest of the hours inside the TSMIP.

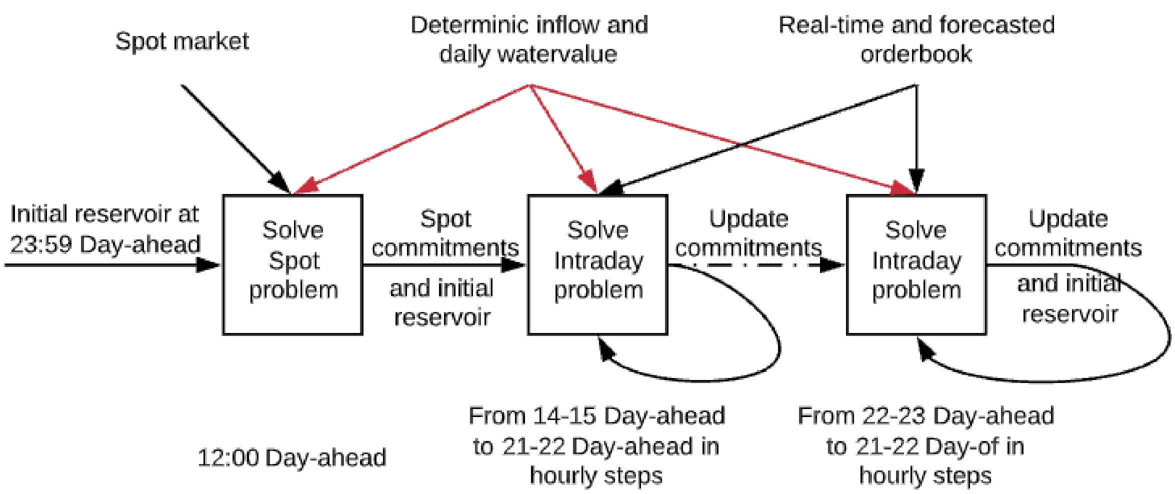

Figure 3.1: Market modelling of the day-ahead and intraday models with inputs and outputs.

Assume the model is at an arbitrary time between 14:00 day-ahead and 22:00 day-of production, which is the opening time of the intraday market. Depending on the time, the model can trade in all products that have not closed yet. Trading in the real-time order book closes one hour before production, so trading in the forecasted order book should close before this. If a product is about to close, there is no future trading opportunity. In this thesis trading in forecasted order books therefore closes two hours before production, that is one hour before the real-time order book closes. Figure 3.2 shows how the market changes as the rolling horizon approach goes through the 32 hourly steps of the intraday market. It shows what real-time and forecasted products are open to trade with at different times during the day. For example during the hour 14 day-ahead one can trade in all products in the real-time and forecasted order books. And during the last hour of the market, hour 21 day-of, one can only trade in the product 23 in the real-time order book as all other real-time products have closed and there is no future trading opportunities. The initial reservoir is from the hour before the first open real-time product of the time step.

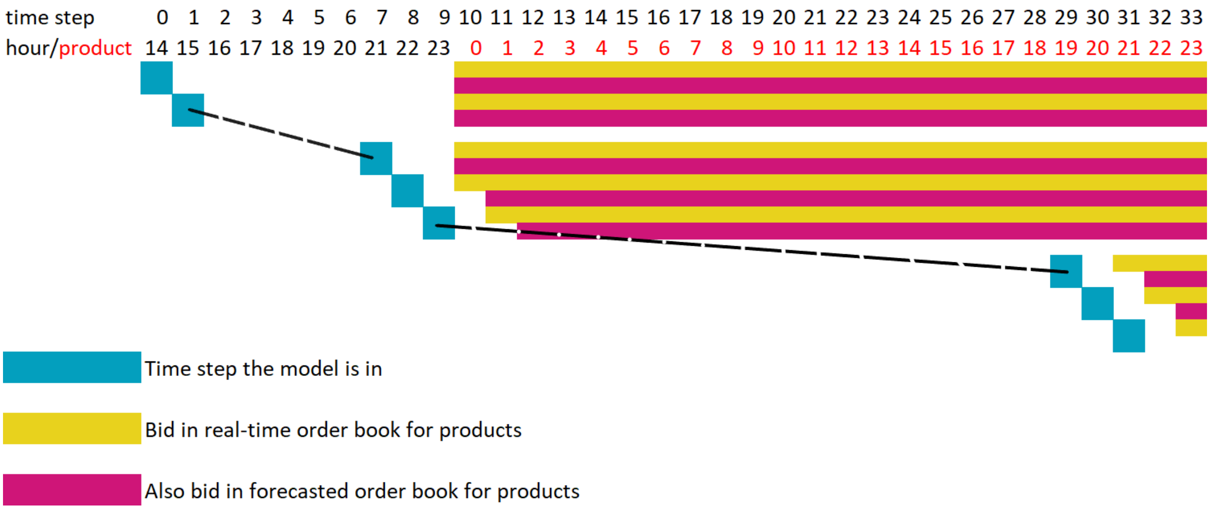

Figure 3.2: Example illustrative of the products one can bid in for some time steps.

The following sections will present the details of the general TSMIP modelling for an arbitrary time-step of the intraday market. The last sections are on the real-time order book modelling and on the different bidding models used in this thesis.

3.3.1 General intraday modelling

The hydropower producer is assumed to be a price taker in the bidding problem, so it can only accept orders that are already in the order book. Not being able to place orders for others to accept, loses some trading opportunities, but it simplifies the modelling a lot. It is also assumed that accepted orders will not impact how the market develops, they will only be removed from the order book. This assumption would be valid for small trades in a liquid market, but since the intraday market of NO3 lacks liquidity at certain times this is an important simplification. Together, these assumptions avoids realistic but complex market mechanics and makes it possible to use the historical/real-time order book without having to build a complicated order book/market model with it.

The goal of the TSMIP is to trade and optimize the production schedule so that the resource use of water is as good as possible at the end of the day. This includes the revenue from intraday trading in the real-time order books, the value of the reservoir at the end of the day and the costs of ramping up production from zero between production hours. Equations 3.2, 3.3 and 3.4 shows the objective function of the bidding problem at an arbitrary time step during the intraday market. The part 3.2 is the revenue from accepting orders in the real-time order book. The model can decide how much volume v it wants to accept at the price of the order $Price$, from the different buy or sell orders in the market. The summation is over all open hourly products M, and all orders O. This thesis assumes no intraday trading costs. The part 3.3 shows the revenue from accepting orders in the forecasted order book for open products M^{pred}, while 3.4 is the value of the reservoir at the end of the day minus the startup costs of the generator throughout the day. The parts 3.3 and 3.4 are summed up over all the forecasted intraday scenarios S with the probability ρ. A more thorough explanation of the notation can be found in appendix B.1. The scenario generation for the forecasted order books is explained in chapter 4.

$$max\ z = \sum_{m \in M} \Big(\sum_{o \in O_m^{buy,real}} Price_{mo}^{buy,real} v_{mo}^{buy,real} - \sum_{o \in O_m^{sell,real}} Price_{mo}^{sell,real} v_{mo}^{sell,real} \Big) \tag{3.2}$$

$$+ \sum_{s \in S} \rho \Big(\sum_{m \in M^{pred}} \Big(\sum_{o \in O_{sm}^{buy,pred}} Price_{smo}^{buy,pred} v_{smo}^{buy,pred} - \sum_{o \in O_{sm}^{sell,pred}} Price_{smo}^{sell,pred} v_{smo}^{sell,pred} \Big) \tag{3.3}$$

$$+ Watervalue * r_{s,m=23} - \sum_{m \in M^{startup}} C^{startup} a_{sm} \Big) \tag{3.4}$$

Each order in both the real-time and predicted order books is represented in the model with a volume parameter and a volume variable. The variable v is the amount our model accepts from the order, while the parameter V gives the actual volume of the order. This is handled in the restrictions 3.5, 3.6, 3.7 and 3.8.

$$0 \leq v_{mo}^{buy,real} \leq V_{mo}^{buy,real} \qquad \forall\ m \in M,\ o \in O_m^{buy,real} \tag{3.5}$$

$$0 \leq v_{mo}^{sell,real} \leq V_{mo}^{sell,real} \qquad \forall\ m \in M,\ o \in O_m^{sell,real} \tag{3.6}$$

$$0 \leq v_{smo}^{buy,pred} \leq V_{smo}^{buy,pred} \qquad \forall\ s \in S,\ m \in M,\ o \in O_{sm}^{buy,pred} \tag{3.7}$$

$$0 \leq v_{smo}^{sell,pred} \leq V_{smo}^{sell,pred} \qquad \forall\ s \in S,\ m \in M,\ o \in O_{sm}^{sell,pred} \tag{3.8}$$

3.3.2 General hydropower modelling

In this section the hydropower modelling will be presented. The problem consist of one reservoir and one production unit. Most of the hydropower restrictions are for all products M and for all scenarios S. See Appendix B.1 for the explained notation. The hydropower modelling is kept the same as in the specialization project [1]. Therefore, that work with some small changes is presented below.

The hydropower model is connected to the bid model via the restriction that consist of 3.9 and 3.10, which states that for each scenario and product/production hour, the model has to produce the same amount of power that it has committed to in the spot and intraday market. The commitments consist of the previous commitments $V^{initial}$, and the accepted orders from the real-time and predicted order books. Restriction 3.11 is a special case for the products that are still open, but that the model does not predict since the products are about to close. As shown in restriction 3.11 and figure 3.2, bidding in forecasted order books of products start to close at t = 8, which is the hour 22:00-23:00. Because of the possibility of negative prices in the markets, the production have to equal the commitments, instead of being greater than.

$$p_{sm} = V_m^{initial} + \sum_{o \in O_m^{buy,real}} v_{mo}^{buy,real} - \sum_{o \in O_m^{sell,real}} v_{mo}^{sell,real} \tag{3.9}$$

$$+ \sum_{o \in O_m^{buy,pred}} v_{smo}^{buy,pred} - \sum_{o \in O_m^{sell,pred}} v_{smo}^{sell,pred} \qquad \forall\, s \in S,\ m \in M^{pred} \tag{3.10}$$

$$if\ t > 7:\ p_{sm} = V_m^{initial} + \sum_{o \in O_m^{buy,real}} v_{mo}^{buy,real} - \sum_{o \in O_m^{sell,real}} v_{mo}^{sell,real} \qquad \forall\, s \in S,\ m = t - 8 \tag{3.11}$$

The restrictions represented by 3.9, 3.10 and 3.11 also lets the hydropower producer speculate in the intraday market. The production bounds presented later in equation 3.15 limits the production between 0 and the maximum production capacity. For equation 3.11, which represents the last hour of trading for a product, this means that the commitments have to be inside the production capacity. But for the restriction represented by 3.9 and 3.10 the model can commit to higher or lower production capacity as long as it expects to cover these positions later in the intraday market. This is represented by trading in the forecasted order books. Even though hydropower producers are generally risk-averse and would likely not take this risk in the real world, this is a wanted behaviour when testing a forecasting method. A bad forecasting method will make the model accept orders that it will have to cover later at worse prices, and we do not want the production capacity of the generator to reduce the impact and visibility of this behaviour.

Kong *et al.* [23] presents the following equation that shows that production is a product of gravity, generator and turbine efficiency, net head and discharge : $p = G * \eta^{gen}(p) * \eta^{turb}(h^{net}, q) * h^{net} * q$. Also the net head is a complex function of the gross head minus the main head losses, which are the penstock/main tunnel, canal intake and tailrace head losses. The non-linearitiess makes this modelling clearly too complex for an optimization model. This is why it is common to model this relation with a piece-wise linear approximation of the production to discharge (PQ) curve. This can be for different or for a fixed head level. This thesis will do the latter to keep the restrictions linear.

The PQ-curve goes from Q^{min} and P^{min} to Q^{max} and P^{max}. The area between 0 and Q^{min} is infeasible for production. In the modelling this is handled with a binary variable, u, called the production status that keeps the production and discharge either at 0 or inside the PQ-curve. Constraints 3.12 shows that the discharge is a sum of the minimum possible discharge times the production status, the sum of the discharge segments in the PQ curve and the spillage. Each discharge segment has a maximum value that constrains it, as can be seen in constraint 3.13. This thesis models the PQ-curve with 10 discharge segments.

$$u_m * Q^{min} + \sum_{j \in J} q_{smj}^{segment} + q_{sm}^{spillage} = q_{sm} \qquad \forall\, s \in S,\ m \in M \tag{3.12}$$

$$0 \leq q_{smj}^{segment} \leq Q_j^{segment,max} \qquad \forall\, s \in S,\ m \in M \tag{3.13}$$

Constraint 3.14 models the production as a sum of the minimum possible production times the production status and the sum of the discharge segments times their equivalent

PQ-efficiencies ϵ. Constraint 3.15 shows that the production is bounded by the minimum and maximum possible production times the production status. The production status in this thesis has been chosen to be the same regardless of the scenario.

$$u_m * P^{min} + \sum_{j \in J} \epsilon_j * q_{smj}^{segment} = p_{sm} \qquad \forall\, s \in S,\ m \in M \tag{3.14}$$

$$P^{min} * u_m \leq p_{sm} \leq P^{max} * u_m \qquad \forall\, s \in S,\ m \in M \tag{3.15}$$

The reservoir balance modelling is given by constraint 3.16. The reservoir volume is determined by the reservoir volume in the previous hour, the deterministic inflow and the discharge. The discharge and inflow, which are in m^3/s, are transformed to the hourly discharge format of the reservoir. Constraint 3.17 initialises the reservoir balance with the reservoir volume of the hour before the first open product. In reality, when bidding day-ahead, there are still intraday markets open for the previous day, which means that the initial reservoir is not final. This is a consequence of the overlapping of the intraday days that we model separately. This thesis assumes that the first initial reservoir is known and we initialise it from the historical reservoir. In addition the reservoir is restricted by a lower and upper bounds as seen in 3.18.

$$r_{sm} = r_{s,m-1} + F * L * (I_m - q_{sm}) \qquad \forall\, s \in S,\ m \in M^{res} \tag{3.16}$$

$$r_{sm} = R^{initial} \qquad \forall\, s \in S,\ m = max(t\text{-}8, 0) - 1 \tag{3.17}$$

$$R^{min} \leq r_{sm} \leq R^{max} \qquad \forall\, s \in S,\ m \in M \tag{3.18}$$

The reservoir volume of the final product is also represented in the objective function 3.4. Timed with the watervalue it gives the final value of the reservoir. An important simplification here is the constant water value of the reservoir throughout the day. This is only a valid simplification for medium and large reservoirs. The reservoir for our case study is small/medium so the use of water value cuts would improve this modelling, but also make the problem more complex and time consuming. Another detail is that water time delays are not represented as the model only consists of one reservoir and one production unit.

Hydropower plants have relatively low start up/shut-down cost compared with many other power generating technologies. However, it is still important to include these costs in the model, since the wear and tear of the generator should be accounted for to reduce unnecessary and costly production changes. The startup cost is assumed to be 150 € and the binary startup variable a is connected to the production status by constraint 3.19. Since we do not want the model to take advantage of the fact that the horizon is one day at a time, the startup cost is coupled to the previous day or product, and to the next day. For this to be possible, restriction 3.20 and 3.21 initialises the production statuses. The start-up costs are also included in the objective function 3.4.

$$a_m \geq u_m - u_{m-1} \qquad \forall\, m \in M^{startup} \tag{3.19}$$

$$u_m = U^{initial} \qquad \forall\, m = max(t\text{-}8, 0) - 1 \tag{3.20}$$

$$u_m = U^{last} \qquad \forall\, m = 23 + 1 \tag{3.21}$$

3.3.3 Real-time order book

As the real-time order book is continuous, it has to be discretized to fit the hourly modelling above. What was done in the specialization project [1] is to take a snapshot of the order book at minute 55 of every hour. This is pessimistic modelling of the intraday market as many good orders will not be available exactly when the snapshot was taken. Only parts of the intraday opportunities is therefore represented. This is a problem since the forecasting is done for the whole market and becomes much more optimistic than the trading opportunities in the real-time order book. This leads to the model trading based on forecasted prices that will not materialise in the pessimistic real-time order book, which can lead to non-optimal trades. This snapshot model will be used for the analysis in section 5.3.1 to find the realistic benefit of intraday trading, but is not the main model of this thesis. It has all the mathematical elements presented in the sections above.

In this thesis optimistic modelling of the real-time order book was implemented. The concept is to aggregate all orders during the hour and letting the bidding model choose from all of them. This gives too unrealistic intraday profits because of arbitrage, so we add a condition that forces the model to either just accept buy orders or just accept sell orders for each product. This reduces the arbitrage possibilities, but the benefit of intraday trading will still be exaggerated. The reason for this is that the bidding model can choose or arbitrage between orders that was not available at the same time. The benefit of this modelling is that the forecast now is pessimistic compared to the exaggerated real-time order book, so it will not overestimate the future trading opportunities. Besides the exaggerated benefits, another downside is that it adds the restrictions 3.22, 3.23 and 3.24 and more binary variables to the problem, $buying$ and $selling$. Restriction 3.22 sets $buying$ to 1 if the model is buying from a product m, while restriction 3.23 does the same for selling. Restriction 3.24 limits the possibility to only do one of them. The constant 1000 MW is chosen high enough over the production capacity of 37 MW, so that it does not affect the models ability to speculate as described in section 3.3.1.

$$\sum_{o \in O_m^{buy,real}} v_{mo}^{buy,real} \leq buying_m * 1000 \qquad \forall\, m \in M \tag{3.22}$$

$$\sum_{o \in O_m^{sell,real}} v_{mo}^{sell,real} \leq selling_m * 1000 \qquad \forall\, m \in M \tag{3.23}$$

$$buying_m + selling_m \leq 1 \qquad \forall\, m \in M \tag{3.24}$$

Another order book assumption is that all orders are collected as single orders even if they are part of more complex order types. Block orders are possible to differentiate from single orders, so it would be possible in future work to either remove them or model block orders properly. The latter could add unnecessary complexity to the problem. Other types of orders are not possible to differentiate from single orders in our raw order book data.

Another detail that could impact the results is the updating of orders. Orders can either be updated when they are partly accepted or when the bidder wants to change the price or volume of their order. When aggregating for a whole hour, this can lead to several versions of the same order. This thesis chooses to use the first order if this is the case and delete the others. This is possible to do since changed orders most often have the same order ID. Some market participants uses different order IDs when changing orders, thus the aggregation can not identify that some orders are just different versions of each other. Then all the versions of the same orders are kept. Changed orders is not an issue when using snapshots.

3.3.4 Different intraday bidding models

This section will give a short summary of the variations of the TSMIP intraday bidding models that will be used in this thesis.

The main model is the intraday bidding model with the hourly aggregated order book, and restrictions to only buy or sell described in section 3.3.3. It consist of all of the mathematical equation presented in the sections above. The next model is the bidding model that uses hourly snapshot. It consist of all of the mathematical equation presented in the sections above, except for the equations 3.22, 3.23 and 3.24. The main difference between the snapshot and aggregation model is how the raw order book data is processed.

These two models can also be run without forecasting. Then the amount of scenarios is set to one, and the elements regarding the forecasted order books are removed. The mathematical elements removed are 3.3, 3.7, 3.8 and 3.10. The last model variation is to run the two models without forecasting at a higher frequency than hourly. The TSMIP stays exactly the same, but the modelling framework slightly changes from what is presented in figure 3.1. If for example the bidding model is run frequently using new snapshots every minute, the accepted orders will be added to the initial commitments between every model run. However, the initial reservoir will still be updated as described in figure 3.1, which is hourly when markets are closing.

Chapter 4

Forecasting

This section will present the forecasting methodology used to generate the forecasted order books. First the general concept and motivation is presented before the modelling steps of the forecasting is covered.

4.1 Concept and motivation

In the forecasting, the trading opportunities in the order books will be represented by the actual trades that happened in the market(Elbas ticker data [43]). The continuous order book data will therefore not be used in the forecasting. Representing the trading opportunities in the order books with the actual trades in the market is a simplification for several reasons. One of them is that the hydropower producer is a price taker in this thesis, so all the trades in the ticker data will not be available when the hydropower producer is trading in the order book data. The intraday market forecasts will therefore be more optimistic than the trading opportunities in the full continuous order book data. However, the forecasted intraday market will still be less optimistic that the order book modelling chosen in section 3.3.3.

The general forecasting concept in this thesis is to predict the daily volume and volume weighted price for each open product, and then adjust the volume based on how much time has passed in the intraday market. This is done for the trades where NO3 is buying and selling to get one buy order and one sell order per product. This gives a model that can be used regardless of the time step of the intraday market the model is in. Predicting one order (only volume weighted price over period) per product per day is what is done in most of the forecasting literature in section 2.3.2. Choosing a similar concept for this thesis makes it possible to use the same methods as those used in the forecasting literature. The only differences are that the forecasting has to be done for buying and selling, and for the price and volume, which is new in this thesis.

The motivation for forecasting two-sided trading opportunities comes from the analysis of the ticker data for NO3 in appendix A. It shows that NO3 mostly trades with other bidding zones and that NO3 both buys and sells at better prices than the spot price of NO3. This means that different trading opportunities can appear in different bidding zones at the same time, which would not be possible to represent with one price and one volume. Also some market situations could lead to volatility and therefore arbitrage opportunities during the day. Another reason is that some variables, like transmission capacity from Germany to NO3 is only relevant for the buying opportunities of NO3, while the transmission capacity from NO3 to Germany is only relevant for NO3 selling. A more complex version of this was attempted in the specialization project [1], where one buy order and one sell order per product was forecasted for each bidding zone and put into the forecasted order book. This was unsuccessful probably because the bidding zones were forecasted independently of each other in the stochastic scenarios.

4.2 Modelling steps

Forecasting consists of many steps including variable selection, preprocessing of the data, choosing and training the forecasting method, and at last use of the final model in a forecasting setting. This section will present the steps used to develop the forecasting model in this thesis.

4.2.1 Data and feature engineering

A forecasting model is built by finding relations between inputs and outputs in historical data. The model is then used to forecast the output for a known input. In this thesis, the inputs are the spot price and the fundamental market drivers, while the outputs are the daily volumes and volume weighted prices of intraday products. Table 4.1 is a summary of the data sets used in the forecasting model.

Data set	**Source**	**Resolution**	**Assumed available**
Elbas ticker data	Nord Pool [43]	Continuous	Continuosly
Spot prices	ENTSOE [44]	Hourly	Before IDM opens
Initial intraday capacity	ENTSOE [45]	Hourly	14:00 Day-ahead
Day-ahead wind forecast	SKM Market Predictor	Hourly	12:00 Day-ahead
Day-ahead wind forecast	ENTSOE [46]	Hourly	18:00 Day-ahead
Intraday wind forecast	ENTSOE [46]	Hourly	08:00 Day of production
Wind production	ENTSOE [47]	Hourly	After production hour

Table 4.1: Market data used in the prediction model.

Outputs: Intraday variables NO3: Buy and sell volume [MWh], trade occurrence [binary], and premium [€ /MWh]

The Elbas ticker data from Nord Pool is filtered to keep the trades where NO3 is one of the participants. One data set is made for when NO3 is the buyer and one where NO3 is the seller in the trades. The trades in each data set are then aggregated for each production hour so that one is left with a total volume and a volume weighted price for each product. Then the buy and sell intraday price premiums, $premium = price^{ID} - price^{spot}_{NO3}$, are found using the volume weighted intraday prices and the NO3 day-ahead price for each production hour. The premiums are used to isolate the price deviations in the intraday market from the initial spot price in NO3. This approach was used successfully by Oksuz and Ugurlu [21], which found modelling the premium around the spot price gave better results than modelling the intraday price directly. Preliminary tests in this thesis came to the same conclusion. When the buy and sell premiums have been predicted, they will just be added to the spot price to represent the buy and sell intraday prices.

As the intraday market for NO3 lacks liquidity, there are many products with 0 total volume. A challenge is therefore to represent this structure in the forecast. The approach that works the best is to separate the volume into two parts. The first element is the occurrence of a trade, which is set to 0 if no trades happened for a product, and to 1 if one or more trades happened. When predicting the occurrence one will get a value between 0 and 1, and by using this as a probability one can generate a random number to say if the value should be 1 or 0. The second element is predicting the volume if there is an event, which is the same approach as for the premium. The input data to this model is therefore all products in the historical data with trading volume other than zero. On the other hand, the input to the occurrence model is all historical products in the dataset, since it here is also relevant with products without trades.

In total there is therefore 2*3=6 forecasting models to predict the intraday market. For the buy and sell order separately, for premiums, volumes and occurrence of trades/volumes. If volume*occurrence is other than 0, the order is kept and will be used in the forecast.

Inputs: Fundamental market drivers in 4 key bidding zones: Wind production, wind production forecast error, spot premium and intraday capacity

The analysis in appendix A found that NO3 mostly trades with other price areas. We therefore assume that variables from other bidding zones are the most important for describing the trading opportunities of a hydropower producer in NO3. Variables from the 4 most profitable areas to trade with are chosen, which are Germany, DK1, SE3 and Finland (see A). The reason why other studies have not taken this approach is that they generally have studied bigger and more liquid markets with less flexibility than NO3.

The first variables chosen are wind related. Relations between the intraday market and the wind production have by numerous studies been found to be price explanatory [11–13, 15–17]. We therefore use the hourly wind production and forecast errors as variables in the prediction model. The hourly wind production for a bidding zone can be used directly from the data source. The wind forecast error however, $error_{area} = F^{DA}_{area} - P_{area}$, is the day-ahead wind production forecast minus the actual wind production. These are the elements used when analysing the historical data to find the relations. Differently, when predicting future market prices and volumes, the actual hourly production is not known, so it has to be substituted with the latest wind production forecast. We have wind production forecasts from the day-ahead at 12:00 and 18:00, and intraday at 08:00. These are used to interpolate what the forecast would have been at a given time step of the intraday market. For hours after 08:00 intraday, the wind production forecast from 08:00 is used. The assumption is that the intraday market forecast should become more accurate as more accurate wind production forecasts closer to the production hours are being used. It is also convenient that most of the trades happens close to the production hour(see section 4.2.2), when the wind forecasts are more accurate. This will be tested by comparing intraday bidding results when using the first forecast of the day, updated forecasts every hourly time step, or the actual production (see section 5.3.4). Frequently updated wind forecasts have been used by [22] before, while other forecasting papers used the day-ahead wind forecasts [18, 20, 21, 40–42] in the prediction. In addition to yearly growth, the wind production data in table 4.1 shows strong seasonality. These are both factors that could impact the performance of the forecasting model.

The next variables are related to congestion between NO3 and the bidding zones chosen, and are assumed known before the intraday market opens. The spot premium, can be described by $spot\_premium_{area} = price^{spot}_{area} - price^{spot}_{NO3}$. The assumption is that since NO3 trades with different price areas, the intraday price premiums should be connected to the spot price of these bidding zones. Since the spot premium between bidding zones is also an indication of the level of congestion in the system, there could also be complex relations between the spot premiums and the expected trading volume. The spot premium has not been used before in the intraday literature because other studies have been less interested in variables from other bidding zones than the one they study.

The second congestion variable is the initial intraday capacity between NO3 and the bidding zone. This is the excess capacity not used in day-ahead market, and should also indicate a level of congestion that could be connected to the intraday premiums and trading volumes. To get the initial intraday capacity between two areas, routes are made connecting several border capacities together. The route with the highest capacity could be found using optimization, but in this thesis it is found manually. For each bidding zone, several likely routes are chosen, the maximum capacity of each route is calculated and the highest capacity from this is kept. One example: If the capacity from NO3 to SE2 is 500 MW, and the capacity from SE2 to SE3 is 300 MW, the capacity from NO3 to SE3 is 300 MW. If this is the route with the highest capacity from NO3 to SE3, we use the capacity of this route. The fundamental driver for NO3 buying is the capacity from the area to NO3, and the fundamental driver for NO3 selling is the capacity from NO3 to the area. The initial intraday capacities between borders have been used directly before by [16, 19].

In total this leads to 15 price and volume explanatory variables used in this thesis. 4 areas * 4 variables - 1. Data for the day-ahead wind forecast in Finland is not available, so the wind forecast error for Finland can not be calculated. While the outputs are forecasted separately, the 15 inputs covered in this section are the same for the premium, volume and occurrence models.

The priority in this thesis was not to test as many price explanatory variables as possible, so better variables than the ones chosen probably exist. Other variables that could have been included based on findings in the literature review, are previous intraday prices [18, 20, 40] and load information [14, 16, 17]. Intraday prices from previous days is somewhat already represented by the day-head price, intraday trades from the same product are hard to include in our general modelling framework, and load information is less interesting since we do not have updated load forecasts throughout the day. In addition [21] used the spot price to describe the premium, which is intuitive since there should be a link between their order of magnitude, and [17, 19, 22, 41] used dummy calendar variables that should identify daily, weekly or seasonal patterns. Future work could also include the analysis of variables from more bidding zones that are less profitable to trade with, and more variables from NO3 should definitely be tested.

4.2.2 Pre- and post-processing of data

Removing bad data

The first data processing step is to look through the data sets for inaccurate or missing data. In this thesis, if that is the case, the production hour is removed from the data set.

When running the bidding model, data from all hours/products of the day are needed.

Normalization

Regression and machine learning generally performs better when the input data is in the same order of magnitude. Options are to normalize the data by rescaling it between the values 0 and 1 or to standardize the data by rescaling it to have a mean of 0 and a standard deviation of 1. The downsides of each method is that normalization is sensitive to outliers, while standardization makes the assumption of a Gaussian distribution. Preliminary tests found normalization to work best for the problem in this thesis, and is therefore used.

Volume adjustments

To reduce the number of input variables in the forecasting model, the hourly relations are modelled outside of the forecasting model. The specialization project [1] found no clear relations between the premium and the hour of day, but found that the total trading volume of products coincided with the usual load patterns. This therefore has to be adjusted since we do not want the forecasting model to predict high volumes during the night. This is done by finding the average daily volume pattern in the historical data, and use it to make the volumes independent of hour of day. Then, after the forecasting is done, the volume forecasts are adjusted based on what hour of day they are for.

The second volume adjustment is based on how much time has passed in the intraday market. We do not want to predict the full volume of a product just before the product closes. Again the specialization project [1] found no important relations between the premiums and the time left of trading for a product. For the volume it was found that the trading volume increases towards delivery of the product. This pattern is found in the historical data separately for buying and selling, and for every hourly product. See figure 4.1 for the volume and premium relations to the trading time of the product 00:00-01:00. After the forecast, the volume is adjusted based on these patterns. One example: We make a prediction 4 hours before delivery for a product. The historical data says that, on average, 70 % of the volume is in the last 4 hours before delivery for this product. Therefore the forecasted volume for this product is reduced by 30 %.

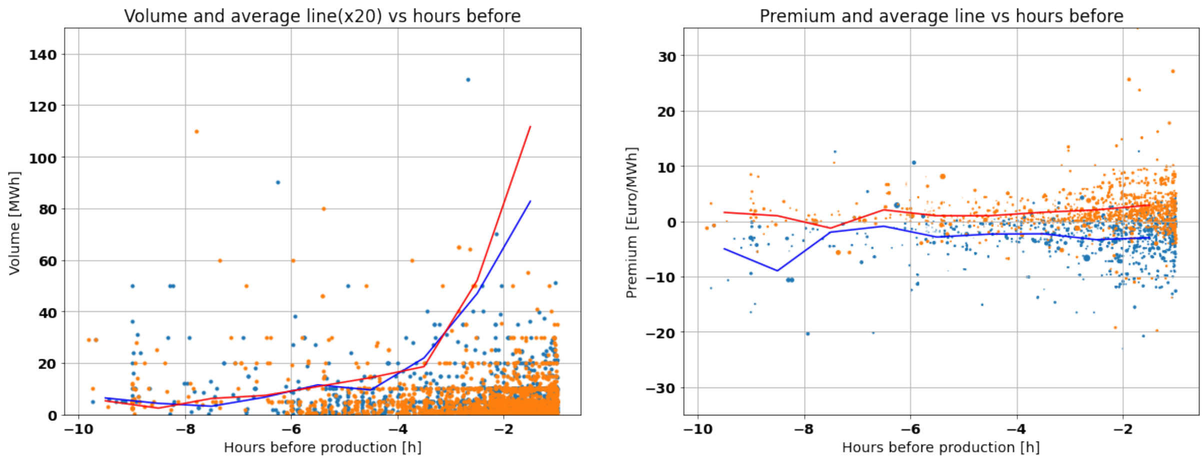

Figure 4.1: Empirical analysis of the relation between the intraday market and the trading time for daily product 0 (00:00-01:00). The market for this product opens at -10 (14:00) and closes at -1 (23:00). NO3 selling is orange, and buying is blue. Elbas ticker data from 2018-06.2020.

4.2.3 Train, validate and test

The standard way of testing a forecasting method is to split the available data into three independent data sets - training, validation and testing data. The training data is the historical data the forecasting method will find relations between to build a forecasting model. The concept is to avoid overfitting the prediction model to the training data by evaluating and changing the model based on results for the validation data. When the prediction model is final, the model can be evaluated on the testing data that is independent. This will also be done for this thesis, where the training and validation data will be used to develop both the forecasting and intraday bidding models and frameworks, while the testing data will only be used as a final evaluation.

4.2.4 Forecasting methods

The goal of this thesis is to choose a forecasting method that can well describe the dynamics in the intraday market, and that gives forecasts that works well in the intraday bidding problem. The forecasting methods evaluated in this thesis are well covered in the intraday forecasting literature, especially in [19, 21, 22].

Simple models are needed to compare against the main models chosen for this thesis. The benchmark forecasting method for the intraday price is to predict the spot price, which equates to predicting a zero premium, or to predict the average intraday premiums in the historical data. For the volume, randomly sampled volumes from the historical data is used. This is a very basic benchmark that the forecasting methods should be able to beat. Other simple methods tested are simple linear regression and linear regression with least absolute shrinkage and selection operator (LASSO). The latter uses regularization to avoid overfitting and to only include the most relevant input variables. The tool used for these methods are the scikit learn models with default parameters [48, 49].

The main model in this thesis is the machine learning method random forest regression (RF). This method generates several random decision trees and uses ensemble learning to choose the best "average" prediction from these trees. Also here the scikit learn tool with default parameters is used [50]. In addition, the feature selection tool [51] of scikit learn can be used to evaluate the importance of every variable in the random forest training.

The next model tested is the neural network method multi-layer perceptron (MLP). This method has a neural network with non-linear transformation functions between the inputs and the outputs of the model. The model chosen has two hidden layers with 15 (one per variable) and 4 (one per bidding zone) neurons in each layer with relu activation. A structure with 30 and 8 neurons in the two hidden layers performed better, but the amount of neurons is kept low to reduce the time use of the forecasting method. This model is built with the module tensorflow keras sequential model [52], with the Adam or SGD optimizer, the loss function mean squared error, batch sizes of 200, and a number of epochs/iterations between 10 and 30. More complicated neural networks like recurrent neural networks were evaluated in the preliminary testing, but they are not kept as they only performed a little better, gave forecasts with similar characteristics and were notably slower than the MLP. These are the recurrent neural networks long short-term memory (LSTM) and gated recurrent unit (GRU).

The forecasting in this thesis is modelled as a regression problem where no temporal structure of the data is taken advantage of. That is, even though we predict a time series from time series data, the outputs are separate point forecasts based on single data points from different variables. It could have been possible to indirectly include the temporal structure in all of the methods above by splitting up the time series and input them as independent variables. The recurrent neural networks are also able to use the temporal structure directly, which is why these methods performs well in the literature [19, 21, 22] compared to other methods. Temporal structures was not modelled in this thesis, either directly or indirectly as it increases the number of variables and makes the forecasting more time consuming.

4.2.5 Scenario generation

The final step is to use the data and forecasting method for generating different intraday market scenarios. Forecasting consists of first fitting the model to the previous historical data and then use the model to predict a future intraday market scenario based on the known input variables at the time. When a model has been fitted it will output the same forecast for the same input variables, so a unique fitted model is therefore needed for every scenario. In this thesis we randomly sample 20 % of the historical data that is before the forecasting day, and use this to fit the forecasting model. This random sampling makes the models and the resulting forecasts different from each other.

Ideally one would make new forecasting models for every hourly time step, so that the latest historical data is included in the model. With a high number of scenarios, many hourly time steps and fitting the data already being time consuming, simplifications have to be made so that testing in this thesis can be done in a reasonable amount of time. A simplification that does not seem to impact the results too much, is implemented, which is to fit the forecasting models of each scenario once a month. The downside is that on the last day of the month, one will use a forecasting model that is not fitted based on data from the 30 last days. Forecasting of days early in the month are less impacted by this simplification as they only lack data from the last couple of days.

Chapter 5

Case study

5.1 Research setting and data

The hydropower plant that will be used for the testing in this thesis is Søa [53], which is owned by TrønderEnergi. The watercourse first goes through two regulated reservoirs with a total of 67 mill m^3, and then to a production unit with a capacity of 37 MW. One of the simplifications made in this thesis is to merge the two reservoirs into one, which will only change the accuracy of the reservoir development when the reservoirs are spilling. The data needed for the hydropower plant is the production-discharge curve, the reservoir head level to volume relations, the daily watervalues, and historical/actual inflows, reservoir levels and production. The historical/actual data is used to initialize each day independently, and also to check that our hydropower modelling behaves similarly to the historical data. The hydropower plant is in the bidding zone NO3, so the market data used will be relevant for this area. In addition to the data presented in forecasting section 4.2.1, data that describes the full development of the intraday order books in NO3 is provided by TrønderEnergi.

As mentioned in the forecasting section 4.2.3, the data is split into training, validation and testing data. The training data is chosen to be 2018 and 2019, the validation data the first half of 2020 and the the independent testing data the second half of 2020. The bidding model will be run for the validation and testing data. The reason for only running the bidding problem in 2020 is because of data availability of order book data and wind production forecasting data, so backtesting the bidding model all the way back to 2018 is not possible. A consideration to make when evaluating the results is that 2020 is an abnormally wet year, which leads to average spot prices in NO3 of around 10 € /MWh, while it is around 40 € /MWh for the two previous years. Predicting the intraday market in 2020 based on data from 2018 and 2019, might therefore not give the most reliable results. However, as the model moves into 2020, more and more data from 2020 will be used in the training. Table 5.1 gives an overview of the average spot price, intraday premium and volume for the days in each period.

	NO3 spot [€]	**Buy prem [€]**	**Sell prem [€]**	**Buy vol [MW]**	**Sell vol [MW]**
Training	41.30	-2.60	1.08	9.1	15.1
Validation	12.20	-2.48	1.52	21.0	22.4
Testing	9.30	-1.44	2.00	11.0	13.3

Table 5.1: Average spot price, premium and volume for the training, validation and testing data.

Production hours with bad or missing data is also removed. When fitting a forecasting model for 02.01.2020, out of the 17520 hours in 2018-2019, 17295 hours can be used for the buy and sell occurrence models, 5502 hours can be used for the buy premium and volume models, and 7210 hours can used for the sell premium and volume models (see section 4.2.1). Most of the production hours removed are because of NO3 not buying or selling in the intraday market for those production hours. For the bidding model, whole days have to be removed if data is missing. For the first half of 2020, 131 days is modelled, and for the second half 125 days. Most of the days removed are days where the reservoir spills, which is not modelled realistically in our model with aggregated reservoirs and no spillage curve. Days where the reservoir development deviates too much from the historical data, is therefore removed. Other reasons for removing days are summer-time shifts or missing capacity, order book and wind forecast data.

The models are built with Python 3.8. The optimization problem is built with Pyomo 5.7 and solved with Gurobi 9.0.3. And the forecasting is done with scikit learn 0.23.2 or Keras Tensorflow 2.3. The full models are also run with multiprocessing to reduce the time use.

5.2 Forecasting problems

The following section will present the tests related to the forecasting methods outside of the bidding model. These tests are not independent from each other, and the result of one test can be used to build the general forecasting and intraday bidding models used in the other tests.

5.2.1 Forecasting methods

First we test how the different forecasting methods presented in section 4.2.4 perform in terms of mean squared error between the average forecast and the historical data. Mean squared error is chosen as it is used in most of the machine learning methods, which are optimized for this loss function. Mean squared error compared to the mean absolute error better handles big error values in the data sets. This forecast testing is only done for the buy and sell premiums. The characteristics of the volumes, with low amount of occurrences, makes the mean squared error less meaningful. Also, volume forecasting for the intraday market is not done in the literature, so we would have nothing to compare the results to. The volume forecasts will therefore be evaluated directly in the bidding model.

5.2.2 Variable selection

The second forecasting test is variable selection for the random forest tree regression method - for the premium, volume and occurrence models. Preliminary testing on the validation data finds that the forecasts of random forest regression performs the best for the bidding model, and that variable selection is important. The feature selection tool [51] of scikit learn with the estimator RandomForestRegressor, is therefore used to evaluate the importance of each variable for the training data.

5.3 Intraday bidding problems

The following section will present the tests that are run with the intraday bidding model. The intraday bidding model consist of many different elements, and testing is needed to see how the bidding model behaves for different modelling choices. These tests will be used to decide on important modelling decisions for this thesis, and also to evaluate the performance of the bidding model. The different bidding models used in the different tests are briefly presented in section 3.3.4.

The intraday bidding models will be evaluated as done in the specialization project [1]. They will be evaluated based on the profits from participating in the intraday market, which is a sum of the revenue from the intraday trades, the change in watervalue of the final reservoir and the change in startup-costs compared to only considering day-ahead commitments. The profits are divided by the initial spot revenues to get a percentage benefit.

$$Benefit\ [\%] = \frac{Profit}{Initial\ spot\ revenue} * 100$$

5.3.1 Order book

As discussed in section 3.3.3, the order book can be modelled in different ways. Part of the case study will be to test how our modelling with aggregation and the restriction to only buy or sell compares to the snapshot approach, and the approach that would give correct results, which would be to run the the optimization model after each new order arrival. The latter is of course not doable practically, but taking snapshots or aggregate every 15, 5 or 1 minute, and run the model at these time steps instead of every hour should give a good estimate of what the solution converges towards. To reduce the time-use of these models, this is only done for intraday bidding models without forecasting.

5.3.2 Forecasting methods

The next step is to test how the different forecasting methods perform in the intraday bidding model. Models will be run without forecasting, with the benchmark forecasting methods and the machine learning methods; random forest regression and multi-layer perceptrons. The two benchmark models both uses random sampled volumes from the historical data, and the spot price or the average spot premiums from the historical data. The machine learning models are also run with premium forecasts from the method and a volume forecast from the benchmark, and then with both premium and volume forecasts with the methods.

5.3.3 Scenarios

The most unstable model will also be run with different amount of scenarios to see how many scenarios are needed for a reliable results over the whole period. This is done for the bidding model with random forest regression models for the premium and volume, since this model gives the most variability in the forecasts.

5.3.4 Updated wind

At last we test to what degree the wind production relations are picked up by the forecasting method (see 4.2.1). The assumption is that the actual wind production for a product should have a more important relation to the intraday market than the wind forecast we substitute it with. The intraday market forecast should become more accurate the closer we get to the production hour, when the wind production forecast is at its most accurate. The first test is therefore to see if it is better to predict the intraday market using the first wind production forecasts of the day, or the actual production which would not be available at the time of forecasting. If the latter is the case, a strong relation is confirmed and we can also test if it improves the results to update the intraday market forecasts every hour based on the latest wind production forecasts.

Chapter 6

Results and discussion

6.1 Forecasting

6.1.1 Methods and data

Table 6.1 shows the mean squared error results when forecasting the buy and sell premiums with different methods and for different time periods. The tested forecasting methods are the spot price, the average historical premiums, linear and Lasso regression, random forest regression and multi-layer perceptrons. The random forest regression method has performed variable selection and uses 12 of the variables, while the other methods uses all the 15 variables presented in section 4.2.1. The three time periods are respectively in the training, validation and testing data, and the historical data used to train the forecasting models is from 01.01.2018 to the start of the time period for all the models. The results show that the neural network model generally performs the best, but that the Lasso regression and average spot premiums often performs similarly. Interestingly, the variable selection in the Lasso regression chooses no variables and predicts the average premium in the historical data instead, which indicates that the average buy or sell premium is a good benchmark model. This also raises questions of how good the neural network model is, as it only manages to beat this benchmark convincingly once(or twice*) - for the training data and the buying premium. Otherwise, the model performance compared to each other seem consistent with the literature.

	S2 2019 / train		**S1 2020 / val**		**S2 2020 / test**	
Method	**Buy prem**	**Sell prem**	**Buy prem**	**Sell prem**	**Buy prem**	**Sell prem**
Spot	12.24	10.80	12.74	13.12	9.36	22.00
Avg prem	12.16	9.98	7.54	12.06	10.18	20.43
Linear	13.71	12.77	10.98	16.30	14.77	27.86
Lasso	12.19	10.00	7.54	12.06	10.28	20.44
RF	11.60	10.70	8.20	13.10	9.52	27.77
MLP	10.18	10.03	7.60	11.98	10.11(8.22*)	20.42

Table 6.1: Mean squared error [$€^2$] between the average premium forecasts and the actual premiums for different forecasting methods and time periods (semesters). *10.11 is the solution when using the same model parameters as in the training and validation data, but the model found a better solution for a different amount of epochs.

The performance also varies between the different periods. Especially the results for the testing data and the selling premium stands out. Table 5.1 shows that the intraday market is quite different in this period, with a much higher sell premium than the historical data used to predict it. This indicates that the performance of the forecasting framework can be vulnerable to the market characteristics changing. Using a shorter horizon than back to 01.01.2018 could reduce this problem, but this would also give the forecasting model less data to train on.

A last observation from the different forecasting methods is that random forest regression gives much more variance between the different forecasts than the neural network forecasts. This is true both for the premium, volume and occurrence models. This is of course also in contrast with the benchmark and linear regression models, which have either none or little variance. Variance should be good when generating scenarios, so that the scenarios are different from each other and covers a bigger solution space.

6.1.2 Variable selection

Table 6.1 shows the importance factor of the exogenous variables when fitting them to buy and sell premiums, volumes and trade occurrences for the random forest method and the training data. The importance factor is the Gini importance, which indicates a higher importance with a higher number. Simply put, it indicates the average quality of the splits in all the decision trees when each variable/feature is used. The importance factor is then normalized, so that the total importance for all the variables is 1 for each of the 6 forecasting models. The results show that spot premium to Germany, DK1 and SE3 are important variables for the premiums and that the transmission capacities are less important. For the volume and trade occurrence training, the best variables are generally the wind variables and the worst are generally the spot premiums and transmission capacities. This analysis is used to remove the 3 worst input variables for each output variable, which improves the results in the bidding model. Since the variable selection tool uses random forest regression, the variables are only removed when using this method, and not when the neural network model is used.

Area	Variable	Buy prem	Sell prem	Buy vol	Sell vol	Buy prob	Sell prob
DE	Wind	0.073	0.057	0.083	0.077	0.095	0.085
	Wind error	0.072	0.079	0.095	0.071	0.085	0.080
	Spot prem	0.139	0.191	0.056	0.075	0.070	0.073
	Cap	0.030	0.023	0.048	0.051	0.045	0.036
DK1	Wind	0.071	0.055	0.079	0.066	0.091	0.082
	Wind error	0.058	0.048	0.080	0.086	0.080	0.075
	Spot prem	0.082	0.172	0.044	0.064	0.049	0.048
	Cap	0.029	0.037	0.055	0.044	0.052	0.048
SE3	Wind	0.075	0.056	0.095	0.089	0.094	0.087
	Wind error	0.070	0.054	0.086	0.078	0.088	0.080
	Spot prem	0.108	0.038	0.021	0.016	0.021	0.038
	Cap	0.033	0.039	0.049	0.085	0.049	0.092
FI	Wind	0.087	0.063	0.095	0.076	0.094	0.086
	Wind error	0.035	0.028	0.061	0.036	0.037	0.036
	Spot prem	0.039	0.059	0.052	0.086	0.051	0.054

Figure 6.1: Variable importance for the random forest method and the training data. The importance factors sums up to 1 for every output variable. The exogenous variables are the wind production, the wind production forecast error, the spot premium and the capacities between NO3 and the bidding zone.

6.2 Intraday bidding

6.2.1 Order book modelling

Table 6.2 shows that the calculated benefit of participating in the intraday market will depend on the modelling of the order book. Hourly snapshots will give benefits of around 1 and 3 percent for the first and second half of 2020, since it includes less of the trades in the order book modelling. The main modelling in this thesis, hourly aggregated order books with the restriction to only buy or sell, gives benefits of around 4 and 8 percent. This is because it gives the bidding model the ability to compare and arbitrage between trades that was not available at the same time. With snapshots and aggregation at higher frequency, the actual benefit of participating in the intraday market is found to converge towards around 2 and 5-6 percent for the two periods, or around 3 percent in total. It is important to remember that these models are without forecasting. Running the bidding model with forecasting once per minute to get more realistic results is very time consuming and not feasible for this thesis. Fitting several forecasting models and running the optimization frequently is very time consuming, and the same is true for testing different parameters and forecasting methods.

Order book modelling	S1 2020 / val	S2 2020 / test
Hourly snapshots	1.06	3.13
15 min snapshot	1.33	3.98
5 min snapshot	1.46	4.11
Minute snapshots	1.63	4.60
Snapshot at every new trade		
Minute agg buy or sell	2.63	6.39
5min agg buy or sell	3.19	6.90
15 min agg buy or sell	3.64	7.37
Hourly agg buy or sell	3.94	8.09

Table 6.2: Intraday benefit with different order book modelling for the validation and testing data. These models are without forecasting.

The goal of including the real order book in the modelling was for the benefits of participating in the intraday market to be more realistic. We have above shown that it is possible with our modelling framework to find a good approximation when the model is without forecasting. With forecasting however, the complexity and time use of the problem forces the order book modelling to be on an hourly time frame, and either pessimistic(snapshot) or optimistic(aggregated). The choice of optimistic order book modelling in this thesis gives exaggerated benefits of intraday trading, but this is less important since the main objective is to be able to compare a bidding model with forecasting, to a bidding model without forecasting. However, we would be more confident in the results of the forecasting if this comparison could be done for bidding in the actual order book, rather than the optimistic/aggregated version.

6.2.2 Forecasting methods

Table 6.3 show the percentage profit of participating in the intraday market when no or different forecasting methods are used to generate the future order book scenarios. The table shows the results for the first and second half of 2020, and total benefits for all 256 days tested in 2020. To give some context to these numbers, the initial spot revenue found by the day-ahead model in the first half of 2020 was 1.23 mill €, while it was 562 thousand € for the second half of 2020. For both periods this gives a benefit of around 50 thousand € for participating in the intraday market for the hydropower plant Søa. With the more realistic modelling found in section 6.2.1, the benefits are around 30 thousand € per period, which is still significant. With the same amount of days in both periods and similar spot prices, it also shows that the production is much higher in the first half of 2020, almost double that of the second half.

	No forecast	Spot	Avg prem	RF prem	MLP prem	RF both	MLP both
S1 2020	3.94	4.01	4.02	4.13	4.10	4.32	4.08
S2 2020	8.09	8.03	7.99	8.04	8.08	7.75	8.13
Total 2020	5.24	5.27	5.27	5.36	5.35	5.40	5.35

Table 6.3: Percentage benefits of participating in the intraday market compared to only participating in the day-ahead market, for different forecasting methods and for the first and second half of 2020. Spot, avg premium, RF prem and MLP prem uses the forecasting method for predicting the premiums, but the volume benchmark model.

The results show that all the forecasting models improve the intraday bidding for the first half of 2020, and that the full random forest model performs the best. This is interesting since this is not the best performing model in terms of mean squared error as seen in section 6.1.1. One reason for why this model performs the best in the bidding model could be the variance in the random forest forecasts of the premiums, volumes and occurrences. The neural network model has less variance in the forecasts, so the importance of variance could have been tested by adding random variance to the neural network forecasts, or even the average premium models.

In the second half of 2020, most of the models perform worse than not predicting the intraday market. The full random forest model now performs the worst, while the full neural network model barely improves the bidding. So similarly to the mean squared error results in section 6.1.1, the forecasts have become worse. The most likely explanation for this is again the difference between the periods highlighted in table 5.1, with just half the volumes in the second half of 2020 compared to the first half, and also vastly different buy and sell premiums. Again this could mean that the forecasting models have to be adapted more to the latest data, rather than fitting data back to 2018.

When it comes to the benefit of trading in the intraday market for both periods, the full random forest method still performs the best for 2020, even though the full MLP model gives a positive benefit in both periods. Another interesting metric is how many of the 256 days it would be better to just use the bidding model without forecasting in terms of benefit/profits. For the random forest regression, it improves the bidding 95 days, worsen it 87 days and bids the same 74 days compared to bidding without forecasting. For the neural network model it improves the bidding 97, worsen it 87 and bids the same 72 days. This result shows that bidding decisions influenced by the forecasting almost worsens the bidding as many times as it improves it. But in total the forecasting slightly improves the bidding.

More generally, these results suggest that the bidding models have to be tested on a longer time period to better understand how the forecasting methods perform in different market conditions. Especially testing the bidding model outside of the abnormal year 2020, is important. How this modelling performs in a more liquid market than NO3, also warrants further investigation. Other directions for future work are to improve the forecasting to take advantage of the temporal structure in the time series data, to see the real potential of using recurrent neural network models. Including other exogenous variables in the forecasting, and variables from more bidding zones, could also improve the performance of the forecasting models. Compared to the most complex intraday forecasting models in the literature, especially [19, 22], our forecasting frameworks are quite simple, and there should be big room for improvements. However, the testing in this thesis has shown that it already takes a lot of resources to first fit and forecast hundreds of scenarios with the forecasting methods, and then run the bidding optimization. So with more complex forecasting frameworks this could become an issue in future work. The same is true for more complex hydropower modelling. Future work could use water value cuts instead of a daily constant water values, or make the production to discharge curve dependent on the head level. A simpler addition to the model that would make the results more realistic without increasing the resource use, is intraday trading costs.

The gap between the resources needed for a bidding model with forecasting, compared to a bidding model without forecasting is also interesting. The bidding optimization without forecasting does not need to handle all the forecasting data, fit the forecasting models or run optimization with hundreds of scenarios. This means that the forecasting has to improve the bidding to a certain degree before it would even be considered used by hydropower producers. The slight benefits of forecasting found in this thesis, together with the uncertainty in the forecasting performance in other time periods, would not justify using the bidding model with forecasting over the model without forecasting. Another question is therefore if our modelling framework can easily be adapted to be used in real life. This is possible for the snapshot bidding model without forecasting. It could look at the order books at an arbitrary point in time, run the optimization model, and have a bidding solution in no more than a couple of seconds. This model is similar to the bidding framework of Dideriksen *et al.* [8]. The big difference is that [8] uses optimization to calculate the bidding curves of the hydropower plant after each change of the production plan, and would know almost instantly if it should accept new order arrivals. For our snapshot model without forecasting, the desired order could be gone by the time the optimization model is finished. This would be an even bigger problem for a slower snapshot model that contains forecasted scenarios. On the other hand, the aggregated bidding model with restrictions on buying and selling used in this thesis, is practical to test the performance of forecasting in this thesis, but could not be used in real life because of the aggregation of the order book.

6.2.3 Scenarios

The stability test for the bidding model with the full random forest model (RF both) and the validation data can be seen in figure 6.2. It shows that the average forecast quickly stabilizes, but that the variance is still quite high also as the number of scenarios increases more. This is a consequence of trading in the order book being very sensitive to the scenarios. There is also a lot of randomness both in the random forest regression method, and in how we choose the 20 % of data to fit each forecasting model. Running the bidding model twice with 300 forecasted scenarios and taking the average benefit is decided to be sufficient for our needs. Increasing the number of scenarios over 600 would still give a high variance, but most importantly, it would be too time consuming to test different forecasting methods and parameters. Also the neural network models in this thesis will be run with 300 scenarios. Remember that the full random forest model is our most variable model, so 300 scenarios should be sufficient also for the neural network model.

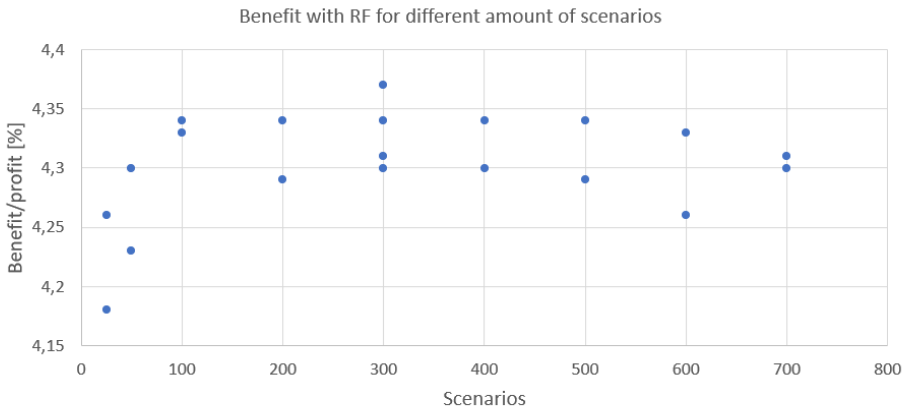

Figure 6.2: Benefit of intraday trading with full random forest regression and different scenarios for the validation data. The bidding model is run two times for each amount of scenarios.

6.2.4 Updated wind

The first wind test for the validation data finds no benefit in predicting the intraday market using the actual production (4.3 %) rather than using the first wind production forecast of the day (4.32 %). This indicates that the relations between the intraday market and the wind production is not fully picked up by the forecasting model. Updating the intraday market forecasts every hour is therefore not needed in this thesis, since the wind is our only variable that updates throughout the day. Instead, all the models in this thesis uses the same forecast for all time steps (but not for all products), the forecast generated when the intraday market opens. Updating the forecasts every hour actually performs even worse (4.17 %), probably because the bidding model now changes strategy every hour based on the new forecast. It would take a lot of scenarios for the general characteristics of the scenarios to be similar every time.

One thing that can improve the forecasting model training based on the wind is to account for the seasonality and the yearly growth of the wind production. This was unsuccessfully attempted in the preliminary testing by including one-hot encoded calendar dummy variables in the forecasting models. Another alternative is to just remove the seasonality and yearly growth from the wind production data before inputting it to the forecasting model.

Chapter 7

Conclusion and future work

The goal of this thesis was to incorporate bidding in the real order book, forecasting, and hydropower scheduling into an optimization and simulation framework in a best possible way. This was achieved with the modelling from the specialization project [1], which uses a rolling-horizon approach to go through the intraday market from it opens until it closes. At each hourly time-step, a two-stage stochastic mixed-integer program will accept profitable limit orders from the real order books, considering the production plan, the water value and future trading opportunities in the forecasted order books. This general framework was kept, but important changes where made to be able to better evaluate the benefit of participating in the intraday market when using fundamental drivers to forecast it. The first change was to use perfect spot trading. The other changes were related to forecasting, which was to change the order book modelling, improve the forecasting framework, and to improve the forecasting by using machine learning methods. Scenarios of future trading opportunities are now generated by forecasting the intraday premium, volume and occurrence of trades for each product. The forecasting is done with random forest regression or neural networks, and uses fundamental drivers as input variables.

For a case study with a hydropower producer in NO3, and 256 days in 2020, the benefit of participating in the intraday market converges to around 3 % for the bidding model without forecasting. For the bidding model with forecasting, more work and data is needed to conclude on the performance. This is despite this model on average outperforming the bidding model without forecasting for the days tested in 2020. This uncertainty is mostly due to the different performance of the forecasting methods under different market conditions. Testing the models outside of the abnormal year 2020 and development of the forecasting methods, is identified as important factors to make the results more reliable. This could be taking advantage of the temporal structures of the time series data with the neural network models LSTM or GRU, or testing more exogenous variables from different bidding zones as input variables. Another result is that the wind relations were not well picked up by the forecasting methods, so updating the forecasts every hour was not necessary. This also shows room for improvement in the forecasting methods.

One of the biggest challenges of combining intraday bidding, forecasting and hydropower scheduling into a framework, is the resources needed to fit the forecasting models and run the bidding model with many scenarios. Only fitting the forecasting methods once a month, is the implemented simplification that did not impact the results to much, but reduced the resource use a lot. Simpler forecasting methods was also used for similar reasons.

The order book modelling on the other hand, impacts the benefit of intraday trading a lot. We are forced towards hourly modelling of the order book, either optimistic(aggregation) or pessimistic(snapshot), and away from the actual benefit of trading in the intraday market. This thesis has shown that hourly aggregating of the order book still works well to compare a bidding model with forecasting, to a model without. For bidding models without forecasting, the resources needed decreases a lot. The bidding model can therefore be run at a higher frequency, and can be used to estimate the actual benefit of intraday trading. The snapshot bidding model without forecasting can also easily be adapted to be used in real life, which is not the case for the aggregation bidding model with forecasting.

Bibliography

[1] J. Bellika and J. Cabrol, 'Intraday bidding problem for a Nordic hydropower producer: Considering wind power and other fundamental drivers,' 2020.

[2] *Single Intraday Coupling (SIDC)*. [Online]. Available: https://www.entsoe.eu/network_codes/cacm/implementation/sidc/.

[3] E. Faria and S. E. Fleten, 'Day-ahead market bidding for a Nordic hydropower producer: Taking the Elbas market into account,' *Computational Management Science*, vol. 8, no. 1-2, pp. 75–101, Apr. 2011, ISSN: 1619697X. DOI: 10.1007/s10287-009-0108-5. [Online]. Available: https://link.springer.com/article/10.1007/s10287-009-0108-5.

[4] E. Engmark and H. Sandven, 'Edda Engmark Hanne Sandven Stochastic multistage bidding optimisation for a Nordic hydro power producer in the post-spot markets,' Tech. Rep., 2017. [Online]. Available: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2469856.

[5] S. Akersveen and J. Graabak, 'A Multistage Stochastic Optimization Model for Trading on the German Intraday Power Market,' Tech. Rep., 2018. [Online]. Available: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2577833.

[6] S. R. Bovim and H. R. Naess, 'A Stochastic Dynamic Programming Approach to the Bidding Problem in the Intraday Electricity Market,' Tech. Rep., 2018. [Online]. Available: https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2577207.

[7] M. Fodstad, M. M. Aarlott and K. T. Midthun, 'Value-Creation Potential from Multi-Market Trading for a Hydropower Producer,' *Energies*, vol. 11, no. 1, pp. 1–15, 2017. [Online]. Available: https://ideas.repec.org/a/gam/jeners/v11y2017i1p16-d123931.html.

[8] A. F. Dideriksen, S. Sekkesater, S. E. Fleten, E. K. Aasgard and H. I. Skjelbred, 'Operational use of marginal cost curves for a hydropower producer trading in the intraday market,' in *International Conference on the European Energy Market, EEM*, vol. 2019-September, IEEE Computer Society, Sep. 2019, ISBN: 9781728112572. DOI: 10.1109/EEM.2019.8916213.

[9] C. Koch, 'Intraday Imbalance Optimization: Incentives and Impact of strategic Intraday Bidding in Germany,' *SSRN Electronic Journal*, Dec. 2019, ISSN: 1556-5068. DOI: 10.2139/ssrn.3495047. [Online]. Available: https://papers.ssrn.com/abstract=3495047.

[10] G. Bertrand and A. Papavasiliou, 'Adaptive Trading in Continuous Intraday Electricity Markets for a Storage Unit,' *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 2339–2350, May 2020, ISSN: 15580679. DOI: 10.1109/TPWRS.2019.2957246.

[11] S. Hagemann, 'Price Determinants in the German Intraday Market for Electricity: An Empirical Analysis,' *SSRN Electronic Journal*, Nov. 2013, ISSN: 1556-5068. DOI: 10.2139/ssrn.2352854. [Online]. Available: https://papers.ssrn.com/abstract=2352854.

[12] R. Kiesel and F. Paraschiv, 'Econometric Analysis of 15-Minute Intraday Electricity Prices,' *SSRN Electronic Journal*, Oct. 2015. DOI: 10.2139/ssrn.2671379. [Online]. Available: https://papers.ssrn.com/abstract=2671379.

[13] F. Karanfil and Y. Li, 'The role of continuous intraday electricity markets: The integration of large-share wind power generation in Denmark,' in *Energy Journal*, vol. 38, International Association for Energy Economics, 2017, pp. 107–130. DOI: 10.5547/01956574.38.2.fkar.

[14] E. R. Soysal, O. J. Olsen, K. Skytte and J. K. Sekamane, 'Intraday market asymmetries - A Nordic example,' in *International Conference on the European Energy Market, EEM*, IEEE Computer Society, Jul. 2017, ISBN: 9781509054992. DOI: 10.1109/EEM.2017.7981920.

[15] F. Ziel, 'Modeling the impact of wind and solar power forecasting errors on intraday electricity prices,' in *International Conference on the European Energy Market, EEM*, IEEE Computer Society, Jul. 2017, ISBN: 9781509054992. DOI: 10.1109/EEM.2017.7981900.

[16] X. Hu, J. Jaraite and A. Kažukauskas, 'The Effects of Wind Power on Electricity Markets: An Evaluation Using the Swedish Electricity Market Data,' *SSRN Electronic Journal*, Jun. 2020, ISSN: 1556-5068. DOI: 10.2139/ssrn.3610800. [Online]. Available: https://papers.ssrn.com/abstract=3610800.

[17] J. Berger, T. Yalcinoz and K. Rudion, 'Investigating the Intraday Continuous Electricity Market using Auto Regression Integrated Moving Average Model with Exogenous Inputs,' in *Proceedings - 2020 IEEE International Conference on Environment and Electrical Engineering and 2020 IEEE Industrial and Commercial Power Systems Europe, EEEIC / I and CPS Europe 2020*, Institute of Electrical and Electronics Engineers Inc., Jun. 2020, ISBN: 9781728174532. DOI: 10.1109/EEEIC/ICPSEurope49358.2020.9160595.

[18] J. Andrade, J. Filipe, M. Reis and R. Bessa, 'Probabilistic Price Forecasting for Day-Ahead and Intraday Markets: Beyond the Statistical Model,' *Sustainability*, vol. 9, no. 11, p. 1990, Oct. 2017, ISSN: 2071-1050. DOI: 10.3390/su9111990. [Online]. Available: http://www.mdpi.com/2071-1050/9/11/1990.

[19] J. K. Kolberg and K. Waage, 'Artificial Intelligence and Nord Pool's Intraday Electricity Market Elbas: A Demonstration and Pragmatic Evaluation of Employing Deep Learning for Price Prediction Using Extensive Market Data and Spatio-Temporal Weather Forecasts NORWEGIAN SCHOOL OF ECONOMICS,' Tech. Rep., 2018. [Online]. Available: https://github.com/johannes-kk/Elbas-Deep-Learning..

[20] T. Janke and F. Steinke, 'Forecasting the Price Distribution of Continuous Intraday Electricity Trading,' *Energies*, vol. 12, no. 22, pp. 1–14, 2019. [Online]. Available: https://ideas.repec.org/a/gam/jeners/v12y2019i22p4262-d285033.html.

[21] I. Oksuz and U. Ugurlu, 'Neural network based model comparison for intraday electricity price forecasting,' *Energies*, vol. 12, no. 23, p. 4557, Nov. 2019, ISSN: 19961073. DOI: 10.3390/en12234557. [Online]. Available: www.mdpi.com/journal/energies.

[22] C. Scholz, M. Lehna, K. Brauns and A. Baier, 'Towards the Prediction of Electricity Prices at the Intraday Market Using Shallow and Deep-Learning Methods,' in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12591 LNAI, Springer Science and Business Media Deutschland GmbH, 2021, pp. 101–118, ISBN: 9783030669805. DOI: 10.1007/978-3-030-66981-2{_}9.

[23] J. Kong, H. I. Skjelbred and O. B. Fosso, *An overview on formulations and optimization methods for the unit-based short-term hydro scheduling problem*, Jan. 2020. DOI: 10.1016/j.epsr.2019.106027.

[24] *SDAC*. [Online]. Available: http://www.nemo-committee.eu/sdac.

[25] *Day-ahead market | Nord Pool*. [Online]. Available: https://www.nordpoolgroup.com/the-power-market/Day-ahead-market/.

[26] *Order types | Nord Pool*. [Online]. Available: https://www.nordpoolgroup.com/trading/intraday-trading/order-types/.

[27] *EPEX SPOT and ECC successfully launch Intraday markets in the Nordic region | EPEX SPOT*. [Online]. Available: https://www.epexspot.com/en/news/epex-spot-and-ecc-successfully-launch-intraday-markets-nordic-region.

[28] *Nord Pool SIDC gate opening and gate closing times*. [Online]. Available: https://www.nordpoolgroup.com/4a4d20/globalassets/download-center/xbid/nord-pool-sidc-gate-opening-times-gate-closing-times.pdf.

[29] *Vindkraftdata - NVE*. [Online]. Available: https://www.nve.no/energiforsyning/kraftproduksjon/vindkraft/vindkraftdata/.

[30] *2019 Annual Report*. [Online]. Available: https://www.nordpoolgroup.com/49ea17/globalassets/download-center/annual-report/annual-report_2019.pdf.

[31] *New trading record on EPEX SPOT in 2019 | EPEX SPOT*. [Online]. Available: https://www.epexspot.com/en/news/new-trading-record-epex-spot-2019.

[32] C. Koch and L. Hirth, 'Short-term electricity trading for system balancing: An empirical analysis of the role of intraday trading in balancing Germany's electricity system,' *Renewable and Sustainable Energy Reviews*, vol. 113, no. July, p. 109 275, 2019, ISSN: 18790690. DOI: 10.1016/j.rser.2019.109275. [Online]. Available: https://doi.org/10.1016/j.rser.2019.109275.

[33] *Roadmap and projects – nordicbalancingmodel*. [Online]. Available: https://nordicbalancingmodel.net/roadmap-and-projects/.

[34] *Nordic Grid Development Plan 2019*. [Online]. Available: https://www.statnett.no/contentassets/61e33bec85804310a0feef41387da2c0/nordic-grid-development-plan-2019-for-web.pdf.

[35] *Wind power in the European Union - Wikipedia*. [Online]. Available: https://en.wikipedia.org/wiki/Wind_power_in_the_European_Union.

[36] E. K. Aasgård, S. E. Fleten, M. Kaut, K. Midthun and G. A. Perez-Valdes, 'Hydropower bidding in a multi-market setting,' *Energy Systems*, vol. 10, no. 3, pp. 543–565, Aug. 2019, ISSN: 18683975. DOI: 10.1007/s12667-018-0291-y. [Online]. Available: https://doi.org/10.1007/s12667-018-0291-y.

[37] R. Weron, *Electricity price forecasting: A review of the state-of-the-art with a look into the future*, Oct. 2014. DOI: 10.1016/j.ijforecast.2014.08.008.

[38] R. Aïd, P. Gruet and H. Pham, 'An optimal trading problem in intraday electricity markets,' *Mathematics and Financial Economics*, vol. 10, no. 1, pp. 49–85, Jan. 2016, ISSN: 18629660. DOI: 10.1007/s11579-015-0150-8. [Online]. Available: http://www.fime-lab.org.

[39] P. Shinde and M. Amelin, 'A literature review of intraday electricity markets and prices,' in *2019 IEEE Milan PowerTech, PowerTech 2019*, Institute of Electrical and Electronics Engineers Inc., Jun. 2019, ISBN: 9781538647226. DOI: 10.1109/PTC.2019.8810752.

[40] C. Monteiro, I. Ramirez-Rosado, L. Fernandez-Jimenez and P. Conde, 'Short-Term Price Forecasting Models Based on Artificial Neural Networks for Intraday Sessions in the Iberian Electricity Market,' *Energies*, vol. 9, no. 9, p. 721, Sep. 2016, ISSN: 1996-1073. DOI: 10.3390/en9090721. [Online]. Available: http://www.mdpi.com/1996-1073/9/9/721.

[41] G. Marcjasz, B. Uniejewski and R. Weron, 'Beating the naïve-combining lasso with naïve intraday electricity price forecasts,' *Energies*, vol. 13, no. 7, p. 1667, Apr. 2020, ISSN: 19961073. DOI: 10.3390/en13071667. [Online]. Available: www.epexspot.com.

[42] S. Kulakov and F. Ziel, 'The Impact of Renewable Energy Forecasts on Intraday Electricity Prices,' *Economics of Energy & Environmental Policy*, vol. 10, no. 1, Mar. 2019. DOI: 10.5547/2160-5890.10.1.skul. [Online]. Available: http://arxiv.org/abs/1903.09641%20http://dx.doi.org/10.5547/2160-5890.10.1.skul.

[43] *Other products*. [Online]. Available: https://www.nordpoolgroup.com/services/power-market-data-services/Product-details/.

[44] *Day-ahead Prices*. [Online]. Available: https://transparency.entsoe.eu/transmission-domain/r2/dayAheadPrices/show.

[45] *Explicit Allocations - Intraday*. [Online]. Available: https://transparency.entsoe.eu/transmission/r2/explicitAllocationsIntraday/show.

[46] *Generation Forecasts for Wind and Solar*. [Online]. Available: https://transparency.entsoe.eu/generation/r2/dayAheadGenerationForecastWindAndSolar/show.

[47] *Actual Generation per Production Type*. [Online]. Available: https://transparency.entsoe.eu/generation/r2/actualGenerationPerProductionType/show.

[48] *sklearn.linear_model.LinearRegression — scikit-learn 0.24.2 documentation*. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html.

[49] *sklearn.linear_model.Lasso — scikit-learn 0.24.2 documentation*. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.Lasso.html.

[50] *sklearn.ensemble.RandomForestRegressor — scikit-learn 0.24.2 documentation*. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestRegressor.html.

[51] *sklearn.feature_selection.SelectFromModel — scikit-learn 0.24.2 documentation*. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.feature_selection.SelectFromModel.html.

[52] *The Sequential model | TensorFlow Core*. [Online]. Available: https://www.tensorflow.org/guide/keras/sequential_model.

[53] *Soa - TronderEnergi*. [Online]. Available: https://tronderenergi.no/produksjon/kraftverk/soa.

Appendix A

NO3 trading counterparts

This analysis was conducted in the specialization project [1].

The first thing we notice when looking at the Elbas ticker data for NO3, is that most of the trades happens with other price areas, most of them outside of Norway. From table A.1, we see that OPX, DK1, SE3 and FI are the most profitable areas to trade with. For each price area, we have calculated the total volume NO3 has bought and sold to that area and the volume weighted price of those trades. To find the area that was the most profitable to trade with, we define profitability as trades happening at better prices than the NO3 day-ahead price. Sellers in NO3 want positive premiums, while buyers wants the intraday price to be lower than the day-ahead price. For each trade we multiply the premium with the volume and aggregate the trades for each price area.

In the ticker data from Nord Pool, OPX are the market participants that uses other power exchanges than Nord Pool. In this thesis we consider OPX to be Germany, even though traders marked OPX can be in all market areas. This is especially true after EPEX SPOT launched their intraday market for the Nordics. Doing some analysis on the Elbas ticker data for OPX, we find that 60% of the volume is with Nord Pool participants in German areas. Another indication is that 28% of the volume of OPX is in quarter hour products which is only available in Germany, Austria and Slovenia. For EPEX SPOT this number is around 15% for the continuous intraday market, and around 25% if the intraday auctions also are included [31]. In this thesis we therefore group the OPX trades with the German areas AMP, TTG and 50Hz. This simplification is also motivated by the fact that Germany is the only central European country we have wind forecasts for.

	Profits [€]			Volumes [MWh]			Prices [€/MWh]	
Area	**Total**	**Sell profit**	**Buy profit**	**Total**	**Sell qty**	**Buy qty**	**Sell price**	**Buy price**
OPX	685897	388596	-297301	200773	126770	74004	34,91	25,93
DK1	187359	29997	-157362	53358	15666	37692	32,20	21,08
SE3	90486	54842	-35644	49955	35269	14685	35,24	32,65
FI	81174	45236	-35938	40620	26584	14036	35,53	30,45
AMP	59250	33016	-26234	17952	12182	5770	35,76	25,92
DK2	52525	28573	-23953	21488	15343	6145	34,59	29,32
SE1	42874	15263	-27611	24015	11836	12179	28,60	24,39
SE4	40721	29806	-10915	14371	10595	3776	34,33	31,98
SE2	21732	-8191	-29923	67577	38810	28767	28,24	26,25
NO5	21074	15121	-5953	20279	12945	7334	27,84	20,40
NO4	20472	7287	-13184	24430	9447	14984	28,75	31,02
TTG	16760	9449	-7310	4935	3472	1463	34,32	20,47
EE	15760	11707	-4053	6569	4773	1796	39,53	35,98
NL	13576	6854	-6722	2949	1706	1243	36,61	23,19
50HZ	10336	4381	-5954	2642	1553	1089	39,46	24,89
LV	7947	574	-7373	3404	979	2425	36,44	30,18
LT	6260	1915	-4345	3844	2590	1254	32,57	36,93
NO1	5891	3201	-2690	6119	4103	2016	31,04	27,09
NO2	3755	-3479	-7234	21232	12155	9077	30,15	31,77
BE	1995	1198	-796	494	342	153	25,78	15,29
FR	1460	1034	-426	487	316	171	34,35	26,08
AT	191	195	5	67	48	19	22,99	35,28
NO3	0	-2414	-2414	13193	6597	6597	28,35	28,35

Table A.1: Areas most profitable to trade with for NO3 if profitability is defined as intraday price premium times the volume. All trades from 2018 to 06.2020 in Elbas ticker data from Nord Pool included.

Appendix B

Mathematical modelling

B.1 Notation

Sets and indices

$t \in \{0, 31\}$ - time steps in the intraday market. 0 represents the first time step after the intraday market opens, the hour 14:00-15:00 day-ahead. 31 represents the hour 21:00-22:00 intraday, which is the last time one can bid in the intraday market (for product 23:00-00:00). 30 represents the hour 20:00-21:00 intraday, which is the last time one can bid in the forecasted order book.

m - Index for market/product/production hour

$M = \{max(t\text{-}8, 0), 23\}$ - all open products at time step t

$M^{pred} = \{max(t\text{-}7, 0), 23\}$ - products that should be predicted at time step t

$M^{res} = \{max(t\text{-}7, 0)\text{-}1, 23\}$ - products relevant for the reservoir coupling at time step t

$M^{startup} = \{max(t\text{-}7, 0), 23+1\}$ - products relevant for startup coupling at time step t

$M^{status} = \{max(t\text{-}7, 0)\text{-}1, 23+1\}$ - products relevant for production status at time step t

$s \in S$ - predicted orderbook scenarios

o - index for orders

$O_m^{buy,real}$ and $O_m^{sell,real}$ - buy and sell orders in the real-time order book for each product m

$O_{sm}^{buy,pred}$ and $O_{sm}^{sell,pred}$ - buy and sell orders in the predicted order book for each product m and scenario s

$j \in J$ - power-discharge curve segments

Parameters

$Watervalue$ - constant watervalue for the day

$Price_{m}^{DA}$ - price in the spot market for product m, used in the perfect spot model

$Price_{mo}^{buy,real}, Price_{mo}^{buy,real}$ - prices in real-time orderbook for product m and scenario s

$V_{mo}^{buy,real}, V_{mo}^{sell,real}$ - maximum volumes in real-time order book

$Price_{smo}^{buy,pred}, Price_{smo}^{sell,pred}$ - prices in forecasted order book

$V_{smo}^{buy,pred}, V_{smo}^{sell,pred}$ - maximum volumes in forecasted orderbook

$V_{m}^{initial}$ - initial committed volume for product, includes spot commitment and intraday commitments from previous time steps

ρ - probability of scenarios

ϵ_{j} - efficiency of production discharge segment

$C^{startup}$ - start-up cost for turbine [€]

$R^{initial}$ - initial reservoir volume

R^{min}, R^{max} - minimum and maximum reservoir volume

P^{min}, P^{max} - minimum and maximum production, when producing

$Q_{j}^{segment,max}$ - maximum discharge for each power-discharge segment j

Q^{min} - minimum discharge, when producing

$U^{initial}, U^{last}$ - initial and last production status for coupling to previous day or hour, and coupling to next day

F - $3600/10^{6}$, conversion factor between m^{3}/s and $Mm^{3}/hour$

L - length of time step in hours

I_{m} - deterministic inflow for product m

Variables

$v_{mo}^{buy,real}, v_{mo}^{sell,real}, v_{smo}^{buy,pred}, v_{smo}^{sell,pred}$ - volumes accepted from order books

r_{sm} - reservoir volume for scenario and product

p_{sm} - production commitment for scenario and product

q_{sm} - discharge for scenario and product

$q_{sm}^{spillage}$ - spillage for scenario and product

$q_{smj}^{segment}$ - discharge for discharge segment j, scenario s and product m

u_m - binary production status variable

a_m - binary startup variable, 1 if startup from product m-1 to m

$buying_m$, $selling_m$ - binary buying and selling variables, 1 if the bidding model is selling or buying in the order book for a product m

B.2 Perfect spot trading

$$max\ z = \sum_{m \in M} Price_m^{DA} p_m + Watervalue * r_{m=23} - \sum_{m \in M^{startup}} C^{startup} a_m$$

$$u_m * Q^{min} + \sum_{j \in J} q_{mj}^{segment} + q_m^{spillage} = q_m \qquad \forall\ m \in M$$

$$0 \leq q_{mj}^{segment} \leq Q_j^{segment,max} \qquad \forall\ m \in M$$

$$u_m * P^{min} + \sum_{j \in J} \epsilon_j * q_{mj}^{segment} = p_m \qquad \forall\ m \in M$$

$$P^{min} * u_m \leq p_m \leq P^{max} * u_m \qquad \forall\ m \in M$$

$$r_m = r_{m-1} + F * L * (I_m - q_m) \qquad \forall\ m \in M^{res}$$

$$r_{-1} = R^{initial}$$

$$R^{min} \leq r_m \leq R^{max} \qquad \forall\ m \in M$$

$$a_m \geq u_m - u_{m-1} \qquad \forall\ m \in M^{startup}$$

$$u_{-1} = U^{initial}$$

$$u_{24} = U^{last}$$