UNIVERSITY OF TWENTE.

DVEP ENERGIE

An Assessment of Wind Power Forecasting Models and its Financial Implications for the Traders

Master's thesis in Industrial Engineering and Management

JONAS AMTSFELD

Faculty of Behavioural, Management and Social sciences
UNIVERSITY OF TWENTE
December 12, 2019
Enschede, the Netherlands

Master's thesis 2019

An Assessment of Wind Power Forecasting Models and its Financial Implications for the Traders

JONAS AMTSFELD

UNIVERSITY OF TWENTE.

Faculty of Behavioural, Management and Social sciences
IEBIS
University of Twente
Enschede, the Netherlands 2019

An Assessment of Wind Power Forecasting Models and its Financial Implications for the Traders
JONAS AMTSFELD

Supervisory Commitee:

R.A.M.G Joosten
A. Abhishta

Faculty of Behavioural, Management and Social sciences
University of Twente
7522 NB Enschede

External supervisors:

Maarten Hofhuis
Maarten Vinke
Bart Hollema

De Vrije Energie Producent
Jan Tinbergenstraat 110
7559 SP Hengelo

Master's Thesis 2019

Cover: Exemplary windpark Getty Images (Tomasz Wyszoamirski).

Abstract

De Vrije Energie Producent (DVEP) is an energy supplier and Balance Responsible Party (BRP) in the Dutch energy market. They are responsible for buying and selling energy on behalf of their customers. To do so, they have to nominate the expected electricity production and consumption of their portfolio for every hour of the following day. Hereby, it is crucial to predict these two volumes as precisely as possible. Forecasting the demand side is rather straightforward. The production side, however, is much more complicated for a wind based portfolio because of uncertainty. A bad forecast can become costly due to imbalance costs and it is thus desirable to have the wind power forecast as precise as possible. Being as precise as possible, however, is not always the most beneficial strategy, as profitable imbalance prices may be harvested otherwise. This is the topic of the second part of the project. Combined, this translates into the following research question:
Which model, based on historical market and weather data, can provide the most accurate and profitable day-ahead electricity bidding when considering a wind-based electricity portfolio?
From literature we find that normalized bias (NBIAS), normalized mean absolute error (NMAE) and normalized root mean square error (NRMSE) are appropriate measures to assess the forecast performance. The research at hand is done for a share of DVEP's wind portfolio, which includes 15 wind parks and a capacity of 55.1 MW and considers the period from 01-07-2018 till 01-07-2019. We firstly compare the two current forecasters and find that they have very similar results, with an NMAE of 7.47% and 7.13% for Forecast 1 and 2, respectively. When this wind power portfolio is considered as a whole, we can not state that both forecasts are significantly different. Besides, when we take the wind speed or day hour into account and rearrange the data based on this, it can be concluded that Forecast 2 outperforms Forecast 1. Wind direction and temperature are also tested, but deliver less explicit results. This is also substantiated with findings in case of big forecast difference between Forecast 1 and 2.
After this, we develop other strategies based on Forecast 2 to determine the best bidding volume for the day-ahead bidding. In order to find the ideal bidding volume that deviates from a strategy of zero imbalance, earlier research stressed that forecasting of prices is crucial. However, due to the characteristic of the Dutch energy market being a dual pricing market, this is a very difficult task. To find the ideal bidding strategy for the day-ahead market, we use two approaches: The point forecasts and the probabilistic forecast. The point forecasts includes next to the above mentioned forecasts also an average of both. For the probabilistic strategy we use empirical distributions of both historical prices and production data related to the forecast volume of Forecast 2. These distributions create scenarios for which the bidding volume is optimized: This is done without restriction, but also with restrictions due to VaR and ES. From the probabilistic strategies the strategy of VaR 0 with a dependent price/production resampling was found to be the best, however this approach was still not better than the day-ahead bidding of the point forecasts, from which Forecast 2 was the best. We conclude that it is wise to use Forecast 2 as input for the day-ahead bidding instead of the currently used Forecast 1.

Acknowledgements

With this thesis, my time as a student of my master Industrial Engineering and Management comes to an end. I look back at almost seven years of studying at the University at Twente, firstly at the Bachelor Advanced Technology and later at this master. When thinking back, I remember a lot of experiences that let me grow as a person. Not only at the university, but also in my board year, in committees and social life. This time lies now behind me, but I am eager to take the next steps in my career.

I want to thank DVEP for allowing me to conduct research on this interesting topic. It was a great time with many laughs but also hard work at the team of Supply, where I felt to be a full member of the team from day one. I would like to thank my supervisors here at DVEP, especially Maarten Hofhuis, to introduce me into the world of the Dutch energy market, Maarten Vinke, for the interesting talks about modelling and Bart Hollema, to explain to me the specialities of short term energy trading.

Next to that, my gratitude goes out to Reinoud Joosten, for his supervision of my master's project and interesting conversations throughout this process, and Abhishta, for his contribution to this report.

Last but not least, I want to thank my family for their unconditional support throughout my entire study time and Sieta, for supporting me in the difficult moments of the last months and giving helpful input with the review of my thesis.

Jonas Amtsfeld, Enschede, December 2019

Contents

List of Figures

List of Tables

Abbreviations and mathematical symbols

Abbreviations

APX	Amsterdam Power Exchange
BRP	Balance Responsible Party
ES	Expected Shortfall
$NBIAS$	Normalized Bias
$NMAE$	Normalized Mean Absolute Error
$NRMSE$	Normalized Root Mean Squared Error
PTU	Program Time Unit, 15min for imbalance market, 1h for spot market
VaR	Value at Risk

Symbols

$\hat{E}_{t+k\|t}$	Prediction of electricity, forecasted at time t
λ^{APX}	APX spot price
λ^{buy}	Imbalance price for buy (or: consume)
λ^{sell}	Imbalance price for sell (or: feed)
E^*_{t+k}	Realized electricity production at t+k
$e_{t+k\|t}$	Error of period t+k, forecasted at moment t
E^b_{t+k}	Electricity bidding volume for PTU t+k
I^C_{t+k}	Cost of imbalance for PTU t+k
k	Leadtime
p_{insta}	Installed capacity
R^2	Coefficient of determination

R_{t+k}	Revenue for PTU t+k
t	Moment of prediction
t_r	Temporal forecast resolution

1

Introduction

This chapter starts with a short introduction of the problem owner, De Vrije Energie Producent (DVEP). After that, we introduce the context of the problem at hand. In Section 1.3 the problem is decribed in more detail, such that in Section 1.4 the research objective and questions can be presented. We conclude the first chapter with the scope of the report (Section 1.5) and its outline (Section 1.6).

1.1 Introduction to DVEP

DVEP is a Dutch energy company based in Hengelo, providing electricity trading possibilities to small electricity-producing companies (wind farms, solar parks, greenhouse farmers) as core business. It was founded in 2003 as a one-man company and since then DVEP has been growing steadily. In the mid of 2017, DVEP had 70 employees. At the end of that year, DVEP was bought by the American LPG distribution company UGI International, as an entry possibility to the European market.

DVEP trades its energy portfolio on the Dutch energy markets but is also active on the German, Belgian and French ones. As a Balance Responsible Party (BRP), one of the main responsibilities for DVEP is to balance production and consumption of its electricity portfolio. They ensure that the energy produced by clients is sold as profitably as possible and on the other hand, the consumed energy of other clients is bought under good conditions. This can include long-term deals, with a lifetime from months to years, up to trades on intraday basis, which are cleared up till five minutes before the hour starts. While long term deals have the goal to reduce the risk of high price fluctuations, short term intraday deals are needed to balance out differences between expected and realized production and consumption. To achieve this, DVEP has its trading desk from which they are active on four different energy markets: The longterm market, spot market for the day-ahead trading, the over-the-counter market for intraday deals and the imbalance market. The different markets will be explained in more detail in Section 2.1.

Clients for DVEP are, as mentioned before, both producing and consuming parties. Producers for the electricity portfolio are wind, Combined Heat and Power Plants (CHP), bio-energy and solar energy, which are all sustainable energy sources. On the other hand, DVEP delivers energy to different organizations, like municipalities or schools.

1.2 Research context

We all have experienced a situation that our weather app tells us that it is rainy outside, but in fact, the sun is shining bright. While it only causes some annoyance for us, it can mean high losses for companies depending on the weather, like DVEP. They are highly dependent on the performance of the wind power forecasts when estimating their clients electricity productions. These forecasts are used to bid the hourly production volume at the day-ahead market. Based on the demand and offer of all BRPs, the electricity prices for each hour of the following day are determined, the so called APX spot prices. However, as a certain volume was nominated, the BRPs are obliged to deliver and consume this exact amount: DVEP has to ensure the balance of its portfolio. For a BRP like DVEP, this can be a difficult task because many producers in their portfolio produce electricity with sun and wind energy. These energy sources are highly dependent on the weather, which is even with just one day ahead very difficult to forecast. In consequence, there can be a big difference between forecast and production.

To ensure a working electricity grid, it is crucial that the grid remains stable: Production and consumption have to match. The grid operator, which is TenneT in the Netherlands, is responsible for this and is thus constantly monitoring the grid. To balance out the differences between forecasted and realized production, TenneT makes use of the imbalance market. Depending on the market situation TenneT issues the prices for feed-in and consumption at the imbalance market every 15 minutes, such that the balance of the market is ensured at all times. This can result in prices that range from -75 € per MWh to 175 € per MWh in just one hour (or even bigger differences).

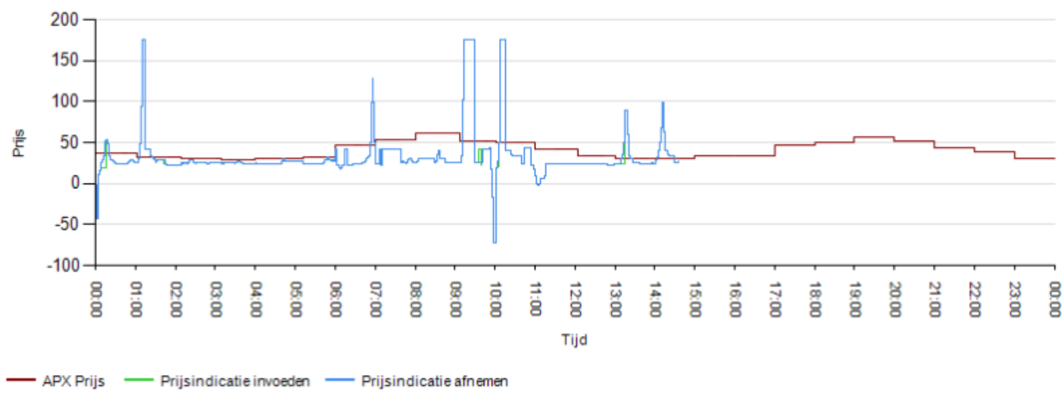


Figure 1.1: The imbalance of 20-09-2019.

Figure 1.1 shows the imbalance price for 20-09-2019 with the characteristic price spikes, where the red line represents the APX price, the day-ahead price from 19-09, green corresponds with the price for feed-in and blue is the price for consumption. The green line is almost not noticeable since those two prices are very often the same. A highly positive price corresponds to an under-production, which can be caused by less production than expected but also by much more demand than forecasted. On the other hand, a highly negative price corresponds to an overproduction of electricity, which is caused by the opposite effects. Those two situations can be caused by imprecise weather forecast or unforeseen downtime of electricity plants.

For every hour of the day ahead, a auction for electricity takes place: To illustrate the buy and sell

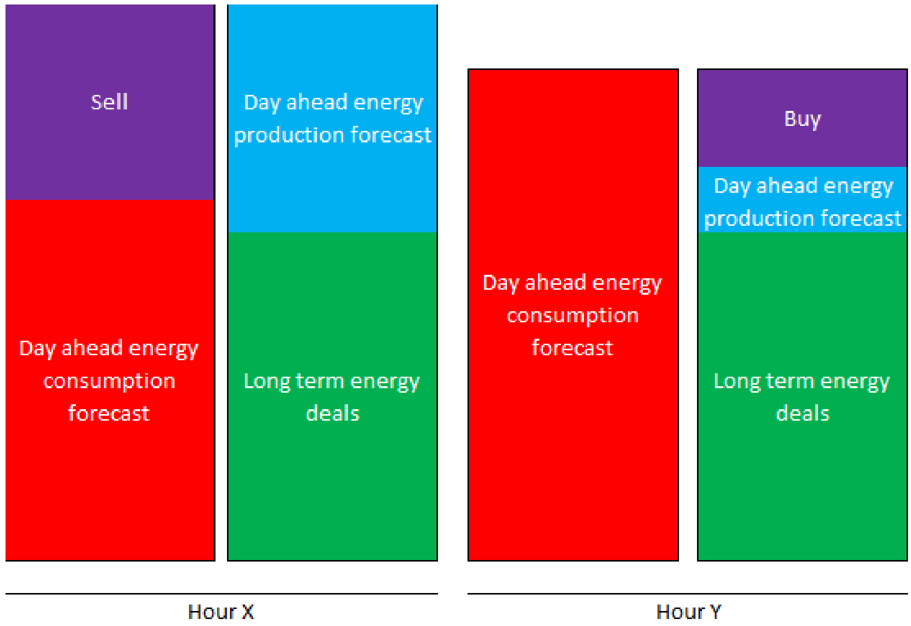


Figure 1.2: Schematic of buy and sell scenarios for the day ahead auction.

scenarios we can consult Figure 1.2, where hour X represents a sell situation and hour Y represents a buy scenario on the spot market. For each hour, a forecast of energy consumption (red) and production (blue) are used to determine the expected consumption and production volumes. As can be seen, a great part of the consumption is hedged with long term deals (green) to reduce the risk of price variability. With these parameters known, we can understand the situations in both hours: In hour X, the long term deals together with the forecast production volumes exceed the expected energy consumption of this hour. This means that at the spot market the expected remaining energy is sold. However, in hour Y, the long term deals plus the expected electricity production is less than the needed volume based on the consumption forecast. As consequence, for this hour additional volumes will be bought in at the day-ahead market. One could easily argue, that hour X can be highly beneficial, while a situation like in hour Y is undesirable. However, we need to keep in mind that the red and blue boxes are only forecasts. While the consumption forecast is quite accurate, the production forecast can be off the real production volume. When we consider the hour X and a situation where the energy production is much lower than the forecast, there is a difference between consumption and production. Assuming there is no intraday market, the remaining volume to fill up the production block thus needs to be bought at the imbalance price. However, when we see Figure 1.1, the imbalance prices are highly volatile thus sometimes not desirable to rely on. Prices may get highly positive or negative, such that the extra buy of electricity can be beneficial, when prices are below the APX spot price, or disadvantageous, when the imbalance prices are higher than the APX spot price. To be less dependent on the imbalance market, the traders try to clear the outstanding positions at the intraday market. This is a market, which is open until five minutes up to delivery for the local market (60 min before delivery for the European market). However, we notice also that in certain moments an over or underproduction could even be more beneficial. In the case of high imbalance prices, we theoretically want to produce more electricity than the forecast volume. In this case, we can sell the surplus volume at the intraday market or the imbalance market for a higher price than at the day-ahead market.

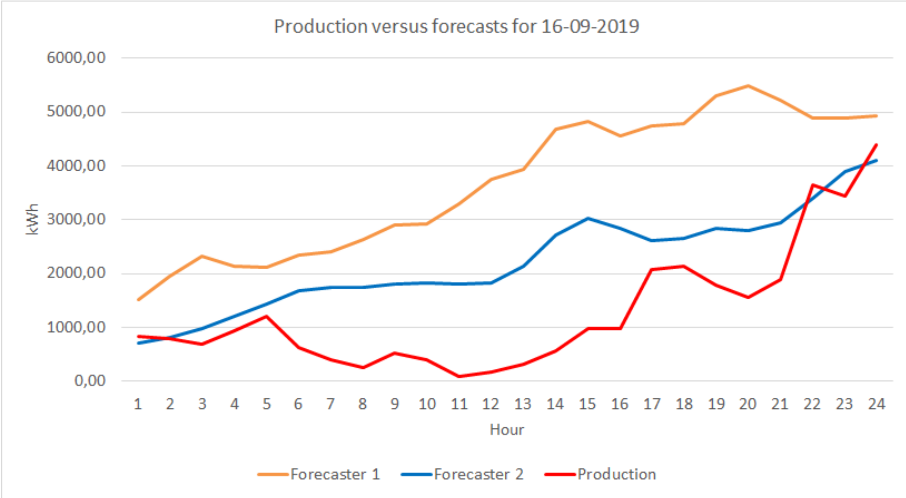


Figure 1.3: Production versus forecast for 16-09, based on portfolio.

This is only a short description of the situation we are dealing with, the following report will dive deeper the different situations.

1.3 Problem description

The performance of the forecast is crucial for an energy trader with a substantial amount of wind power producers. To forecast wind energy production, DVEP has contracts with two different forecasters. Until now, there is no measure of the performance of the two different forecasts and Forecast 1 is always used. This is because this forecaster is providing the forecasts already for several years, while Forecaster 2 is providing forecasts just since June 2018. However, it is not known, if this forecaster is indeed better than the other forecaster.
In Figure 1.3 we show the forecast production against the actual production for 16-09-2019. Both forecasts have been made on 15-09 at 09:00. Here we can already see the difference between the two forecasters. The red line represents the realized production, while yellow and blue refer to Forecast 1 and 2, respectively. Based on this example, which reflects only one day, the forecast of Forecaster 2 would have been better than Forecaster 1 at all times of the day. DVEP seeks to have this comparison standardized, such that they can tell which of the two forecasts is forecasting more precise and under which conditions this is the case. We predict that wind directions and wind speeds influence the forecast performance. Next to that, also the forecast horizon can be a source of difference. Currently, the bidding for the day-ahead market is almost always the same as the Forecaster 1. Only in cases that Forecaster 2 is deviating from this forecast significantly, a different bidding volume is chosen. However, it can be the case that this is not always the most profitable strategy. In case of very low forecast volumes, it can be profitable to bid even less volume at the day-ahead market. In this way, we can reduce the risk of big losses in case of high imbalance. On the other hand, in case of remaining volume, this can be sold at the imbalance market or intraday market. It can be concluded that the bidding strategy for the day ahead market is currently heavily relying on the experience of the trader.

1.4 Research objective and questions

The problem description can be translated into research objectives and questions: The final objective of this research is to develop a model with which we can make day-ahead electricity biddings more accurately and profitable. To achieve this goal, several topics have to be clarified and understood:

The Dutch electricity market and the impact of sustainable energy sources on it have to be understood, such that we can make statements regarding bidding strategies later on.

Next, we have to find performance metrics for the present forecast models and apply these both in the general case as well as for the different conditions of weather and forecast horizon. The performance metrics are based on scientific literature and adapted for the case at hand.

With this knowledge, we want to find a model to make statements about the relation between the APX price and the imbalance price and finally formulate a rule regarding bidding volumes in particular situations. Again, this has to be applied to the Dutch market and in particular the portfolio of DVEP.

While a risk-adapted bidding strategy will be incorporated in the model, it remains to be decided to which extent a risk analysis of the model will be included. The main question that needs to be asked is the following:

- Which model, based on historical market and weather data, can provide the most accurate and profitable day-ahead electricity bidding when considering a wind-based electricity portfolio?

To answer this main question, several other questions have to be asked:

- A.1: How is the Dutch energy market organized and what is the impact of wind energy on the market?
- A.2: Which measures are appropriate to evaluate the performance of wind power forecasts?
  - A.2.1: What metrics are proposed by literature to estimate the error of wind power forecasts?
  - A.2.2: Which of the wind power forecasts available for DVEP is the best based on the proposed metrics?
- B.1: What is the influence of weather-specific or other performance-influencing conditions and can we, with the choice of two different forecasts, find the ideal forecast depending on different conditions?
  - B.1.1: Which conditions can influence the performance of a wind forecast?
  - B.1.2: Which conclusions can we draw regarding the optimal forecaster depending on the before determined conditions?
- C.1: Which model is most appropriate to predict the relationship of the APX price and the imbalance price, and which variables are necessary for this?
  - C.1.1: What models are proposed by literature to estimate the optimal bidding strategy for day-ahead electricity trading and how can these be applied to the Dutch energy market?
  - C.1.2: What are the possible risks of this model?
  - C.1.3: Can this model improve the performance of day-ahead trading in the case of DVEP?

1.5 Research scope

It is apparent that the electricity market is a complicated field, especially now that uncertain power sources like wind energy become more important. This is why it is impossible in the frame of a master's thesis to discuss the research objectives comprehensively and ultimately, but certain assumptions have to be made and the scope needs to be limited.
It is not our goal to discuss the technical characteristics of forecasting methods, but only apply outcomes of the two forecasts DVEP uses for their biddings. These two forecasts come from two different commercial parties DVEP has contracts with. We do not consider the whole wind portfolio of DVEP, but 16 wind parks in different locations of The Netherlands, which we chose based on location. The data considered comes from these 15 wind parks between 01-07-2018 till 01-07-2019. We have to set the transactions at the intraday markets aside, since historical price data at this market is very hard to gather due to the Over The Counter characteristic of the market. There is no set price like at the other two markets. As a consequence we assume for this report that every imbalance is settled at the imbalance market, while in reality the traders have still the possibility to reduce the imbalance at the intraday market.

1.6 Report outline

After this introduction, we continue with a discussion of relevant literature. The goal of the literature study is to answer the Research Questions A.1, A.2.1 and C.1.1 and prepare for the other questions. In Chapter 3 the current situation at DVEP is examined. This includes the explanation of the selected wind parks and based on this, general historical wind data are analyzed. Based on this, we compare the forecast performance of both Forecasters under different conditions to answer the Research Questions A.2.2 and B.1. After this, we also describe the market data of the previous year, which is needed to set up a possible solution for Research Question C.1.1. This is explained in chapter 4. The results of the developed simulations are explained in Chapter 5 and with this, we can finally answer Research Question C.1.3. The thesis is finalized with Chapter 6, where we conclude the research and give based on this, recommendations. This chapter also consists of propositions for further research and points the limitations of this project out.

2

Theory

In this chapter we introduce the literature in this topic. To begin with, we explain the Dutch energy market in more detail and seek to answer research question A.1. In the following we answer question A.2.1 regarding appropriate methods for error measurements. The last part deals with optimal bidding strategies and proposed options by the literature, which answers question C.1.1

2.1 Dutch energy market

Tanrisever et al. (2015) investigated the Dutch electricity market and the impacts of the deregulation on the market. Like most of the other European electricity markets, the Dutch market has a liberalized form since the 1998 Electricity Act, such that customers and suppliers have more freedom in buying and selling electricity. This has led to a more reliable, sustainable and efficient electricity market. Instead of one organization responsible for the whole vertical supply chain, the chain is now split up into different entities. It is not the scope of this report to discuss the different entities of the Dutch electricity in detail. However, the different markets will be introduced to understand the different clearing possibilities for a balance responsible party like DVEP. Figure 2.1 shows a good scheme of the markets and participators per market (TenneT, 2019a).

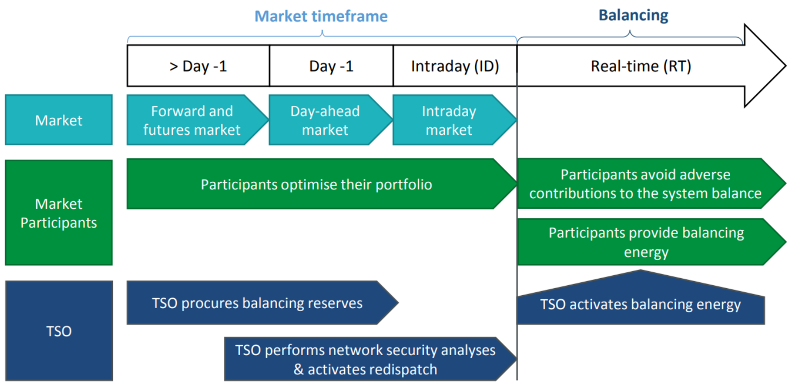


Figure 2.1: The different time frames of the wholesale electricity market (TenneT 2019).

We can say that the market is separated in three different markets, which all serve a different purpose: Forward and Futures markets concentrate on long term deals to ensure price stability for both buyer and seller and hedge possible risks. This market is not influenced by wind power forecasts, thus not in the scope of this report.
The next market closer to the moment of clearing is the day-ahead market: On the day-ahead

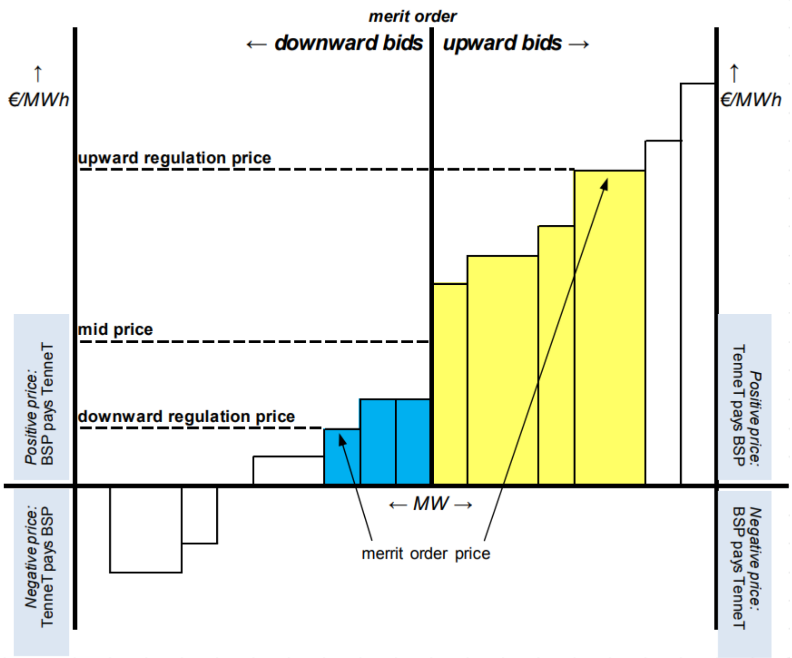


Figure 2.2: The bidding ladder to determine the upward or downward regulation prices (TenneT 2019).

market, electricity can be bought or sold for each hour of the following day. At the day-ahead market, this hour is the shortest possible time unit for trading, which is the program time unit (PTU). This market is particularly interesting to us, as it is highly influenced by the forecasts for wind and solar energy. It is important to realize that the communicated orders for the following day are binding, thus the market participant is required to match their bidding with their final production or consumption. Until 12:00 at the day before delivery (Day-1) the market participants are required to place their anonymous buy and sell orders which are then matched such that at 12:55 the prices are published by the transmission system operator (TSO). In the Netherlands, this is TenneT. These contracts are traded on the Amsterdam Power Exchange (APX).
As it is very unlikely that the bidding volume of the day before matches the actual production volumes, the market participants can adjust their positions at the intraday market. Here, electricity can be bought and sold up until 5 minutes before the physical delivery of the electricity, such that one can adjust according to new information. The goal is here to reduce the imbalance between the bidding of the day before and the actual productions. Also in cases of beneficial intraday prices, the traders might decide to trade here. The positions are cleared over the counter between the market participants.
If these intraday trades do not result in a complete balance in the market, which occurs very often, the TSO uses the imbalance market to ensure balance. On the imbalance market, all market participants are required to buy or sell the volumes they differ from the forecast volume. The prices at the imbalance market are issued by the TSO and based on the upward/ downward bids of the balancing service provider (BSP), which have a reserve volume to counteract imbalance in the market (TenneT, 2019b).
Figure 2.2 shows how the prices are determined: In case of upward regulation, upward bids are ordered depending on their marginal price, the highest bid needed to ensure grid balance is then

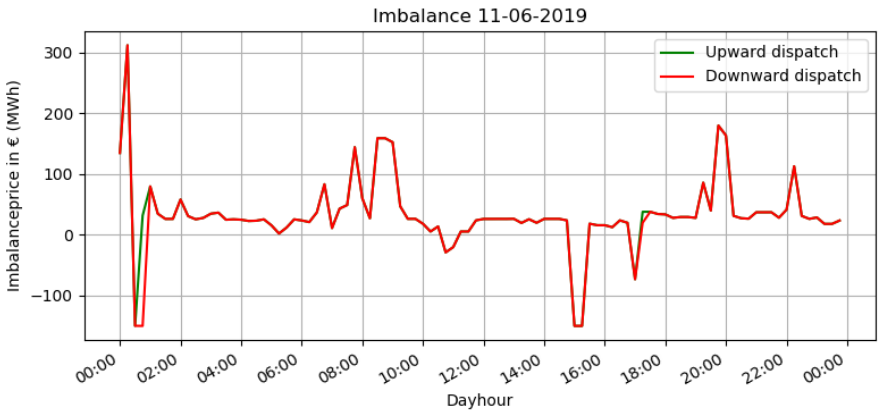


Figure 2.3: The imbalance price for 11-06-2019.

the imbalance price. The shortest time unit at the imbalance market, however, is 15 minutes, such that the PTU at the imbalance market is 15 min. It is important to notice, that the bidding volumes at the day-ahead market are done for one hour, this means there can be four different imbalance prices issued for the same bidding volume of an hour.

To regulate the production, the TSO handles four different regulation states:

- **Regulation state 0:** No up- or downward regulation is applied.
- **Regulation state +1:** Only upward regulation is applied. This implies an underproduction in the respective PTU.
- **Regulation state -1:** Only downward regulation is applied. This implies an overproduction in the respective PTU.

In cases where both upward and downward regulation takes place, the development of the balance delta determines the state of regulation. The balance delta is determined as difference between activated upward bids and activated downward bids:

- When the balance deltas within a PTU continuously increase or is constant, **regulation state +1** applies.
- When the balance deltas within a PTU continuously decrease or is constant, **regulation state -1** applies.
- **Regulation state 2:** When the balance deltas both increase and decrease in the same PTU, regulation state 2 is applied.

In the Netherlands, a dual imbalance pricing is applied, which means that in regulation states 0,-1 and +1 the same prices for both feed and consume are used, while in regulation state 2, the prices for those situations differ (TenneT, 2019b).

Figure 2.3 shows the imbalance price of 11-06-2019, which displays the high variance in prices within a very short period of time. It is important to consider that imbalance prices are handled in different ways by the national grid operators. We have to differentiate between single and dual pricing systems, which have a great influence on the ideal bidding strategies. The Dutch electricity market handles a dual pricing system, while Germany and Spain, for example, handle a single pricing system. A single price implies that prices for feed-in and consume are the same, while the dual system can have different prices for feed-in and consume in the same imbalance PTU (Bal, 2013). Unlike other European countries, there are no restrictions regarding the feed and consume

prices with respect to the APX price.
According to Mulder and Scholtens (2013), the impact of wind energy on the day-ahead price is still rather low. In 2013, the price was mainly correlated to the marginal costs of gas-fired power plants, when the share of wind turbines was 0.88% of the Dutch energy production. In 2019, the share of wind energy increased to 1.7% (CBS, 2019). This shows that wind energy has an increased volume, but compared to conventional energy it is still very small. They conclude, however, with a further increase of renewables, it can happen that the prices will be driven by weather conditions and scarcity in peak supply (Mulder & Scholtens, 2013). Due to lack of more recent literature on this topic, it is likely that this level has not been reached yet.

2.2 Forecast models

A forecast is an essential part of the decision making on the electricity market. A BRP uses forecasting in the first place to predict their production and demand, but also predictions of the different prices can be made. Here, especially the production side is of great interest due to the high shares of wind and solar power. Since solar electricity production is not considered here, we concentrate on models to forecast the power output of wind turbines. Also, the demand side needs to be forecast, as well as market and imbalance prices.
In the following, we present a short introduction to wind forecasting methods. We do not dive into the forecasting method for the other variables. Although there are complex forecasts, also naive forecasts can predict wind power well. They are used as benchmark model for advanced techniques. Pinson (2018) gives as example the random guess, where for each PTU a random value between 0 and the maximum capacity is chosen, and the persistence approach, where the forecast of each PTU is the latest measured value. Even though these forecasts look not smart, they are still difficult to beat by more advanced techniques. Wang et al. (2011) classify the more advanced forecasts into two groups of methods, physical approach and statistical approach, and three time horizons, immediate-short-term, short-term and long-term forecast.
The **physical approach** is based on lower atmosphere or numerical weather prediction and uses weather forecast data like temperature. The data is provided by a meteorological service and is then transformed for the specific wind turbine into expected wind power output.
The **statistical approach**, on the other hand, does not consider meteorological conditions. Using artificial intelligence and time series analysis, the forecasts are obtained.
The **immediate-short-term forecast** considers forecast horizons until 8 hours ahead and is needed for real-time grid operations and regulatory actions. These forecasts are generally based on the statistical approach.
The forecast horizon of the **short-term forecast** includes the day ahead and is used for dispatch planning and operational security. This is the most important forecast to predict the day-ahead volumes.
The **long-term forecast** looks several days ahead and is needed for applications like maintenance planning. They consider usually the physical approach with numerical weather prediction.
Although slightly different, all methods include the following steps:
Firstly, the wind speed is determined, with which the wind power predictions can be made. As last step regional forecasts are made by up- or downscaling (Foley, Leahy, & McKeogh, 2012).

2.3 Measuring forecast models

When working with highly weather-dependent electricity production like wind or solar, the usage of forecasting models is very important. However, as we see in Figure 1.3, this forecast can never exactly predict the actual production. This is why it is necessary to assess the quality of the forecast models extensively.
In general, we are interested in the prediction error for each lead time $t+k$, when predicted at time t, which is the difference between the predicted and realized value. We define this here as:

$$e_{t+k|t} = E^*_{t+k} - \hat{E}_{t+k|t} \tag{2.1}$$

When applied to the wind power forecasting, this means E^*_{t+k} is the realized energy production at $t+k$ and $\hat{E}_{t+k|t}$ is the prediction of electricity, forecast at time t (Madsen et al., 2006).
However, in this way, only the error for every lead time $t+k$ can be captured. In order to measure the overall error of the prediction, we want to consider all t in the time horizon.
To start with, we want to find the model bias, which can be seen as a trend of the predictor. To capture the model bias, we calculate the mean of the error for each horizon over the whole evaluation period. The model bias for itself is scale dependent, which makes it difficult to compare different wind parks. This is why we normalize the error measure. Possibilities are to use the installed production capacity or the measured production power. However, the latter is not feasible in this case, as zero or negative production is possible.

$$\begin{aligned} BIAS &= \frac{1}{N_T}\sum_{t=1}^{N} e_{t+k|t} \\ NBIAS &= \frac{BIAS}{p_{instal}} \end{aligned} \tag{2.2}$$

Where N_T refers to the number of prediction errors for each look-ahead time k for the considered time horizon and p_{instal} refers to the production capacity of the respective wind park.
A positive NBIAS implies an underestimation, while a negative NBIAS signifies that, on average, the forecast was higher than the realized production. A bigger absolute value of NBIAS means that there is a big systematic error, while with a NBIAS close to 0, no trends are detectable. The NBIAS however, tells hardly anything about the predictors performance, because it is averaging all prediction errors. NBIAS of 0 does not directly imply a perfect forecast.
This is why it is appropriate to use the normalized mean absolute error (NMAE) and normalized root mean squared error (NRMSE) (Madsen et al., 2006), (?, ?).

$$\begin{aligned} MAE &= \frac{1}{N_T}\sum_{t=1}^{N} |e_{t+k|t}| \\ NMAE &= \frac{MAE}{p_{instal}} \end{aligned} \tag{2.3}$$

$$\begin{aligned} RMSE &= \sqrt{\frac{1}{N_T}\sum_{t=1}^{N} {e_{t+k|t}}^2} \\ NRMSE &= \frac{RMSE}{p_{instal}} \end{aligned} \tag{2.4}$$

The MAE and RMSE give more information about the performance since they both use absolute values, positive and negative estimations cannot cancel each other out. While the BIAS and MAE are regarded as first-moment error measure, thus associated directly with the production of the wind farm, is the RMSE a second-order estimator of the error. This means it deals with the variance of the prediction error and give larger effects to larger prediction errors (Madsen et al., 2006). In this way, the RMSE is useful to detect a forecasting model with big outliers. The results of these measures are easy to interpret, a bigger MAE or RMSE imply a larger error. Due to the squaring of errors, it can be expected that the RMSE will lead to bigger error measures.
Kariniotakis et al. (2004) performed research on the impact of on-site characteristics on power prediction model performance. They selected six different wind parks in Germany, Spain, Denmark and Ireland. The wind parks were located at different distances from the shoreline and different heights and terrain. Based on location we chose two of the wind parks as a comparison for the wind parks in our portfolio: The German wind park was located 8km from the shoreline of the Baltic sea, while one of the Danish wind parks was in the close proximity of the shoreline of the North sea. The size of the German wind park was 1MW, while the Danish wind park was bigger, with an installed power of 21 MW. The MAE of these two wind parks was both around 10% of the nominal power (Kariniotakis et al., 2004). Other wind parks in the article were located in a more difficult terrain, which resulted in lower prediction performance. It is important to consider the year of publication, such that it can be expected that forecasts have improved since then because of better wind turbines and more precise computing models. However, it gives the indication that a MAE below 10% should be expected for the wind parks of our study

2.3.1 Influence of weather conditions on forecast performance

Next to the general performance of the forecasts, we are also interested in the performance under specific weather conditions like certain wind directions or higher wind speeds. When investigating literature on this, it got clear that this topic can only be discussed in a broad manner, as only few research papers were found on this topic. Next to that, the exact forecast models used for the wind power forecasts here at the company are not known. However, it can give a good indication about possible influences of weather conditions which can be validated later on in this study.
Draxl (2012) discussed the influence of wind speeds on the forecast performance of a mesoscale model. Although they consider the forecast of wind speeds, this can also be used as a metric for the wind power forecast. They found out there is a dependence of the error measure on the forecast wind speed. With wind speeds higher than 10 m/s, the forecast is likely to overpredict the wind speed, while with low wind speeds (under 5m/s) an underprediction appeared to be more likely. When considering the RMSE, this is less for low winds compared to high winds.

2.4 Optimal bidding strategies

When discussing the optimal bidding strategies for a wind power dominated portfolio, we need to define the problem at first.
We know that every market participant has to issue their expected energy production for every hour of the next day. However, we also know that the TSO issues the imbalance prices for every PTU, which is 15 mins. This is why we want to know the revenue of the market participant for every PTU $t + k$. The t refers to the moment of bidding, while k means the leadtime, which can be 13 to 36 hours. The revenue for the PTU $t + k$ depends on the bidding volume E^b, the spot price λ^{APX} and the imbalance costs I^C. The formula can be found in Equation 4.1. The market participant can influence the revenue by issuing the optimal bidding volume V^b, which influences the imbalance costs. The APX price is not known at the moment of bidding and we assume the condition of price taking. This implies that the price can not be influenced by our bidding volume. This assumption can be justified with the fact, that when considering the day-ahead market of August 2019, the portfolio considered accounts for 0.37% of the total traded volume.

$$R_{t+k} = E^b_{t+k}\lambda^{APX} + I^C_{t+k} \tag{2.5}$$

The imbalance cost depend on the sign of the imbalance and can be defined as the following (The subscript $t + k$ is omitted for clarity.):

$$I^C = \begin{cases} \lambda^{sell}(E^* - E^b), E^* > E^b \\ \lambda^{buy}(E^* - E^b), E^* < E^b \\ 0 \qquad\qquad\qquad\quad , E^* = E^b \end{cases} \tag{2.6}$$

The first row refers to a moment of downward regulation, which implies positive imbalance, while the second row refers to a upward regulation, a negative imbalance. For the theorical case, that E^* and E^b are the same, it is clear that the imbalance cost are 0.
This shows that we have to deal with four different uncertainties, the realized production, the spot price at the APX, as well as the two imbalance prices for sell and buy, λ^{sell} and λ^{buy}, respectively.

2.4.1 Uncertainties

To start with, it is important to investigate the uncertainty of power production. With the wind power forecasts at hand we can indeed make a sound approximation about the expected production, but is has to be clear that the forecast is never exactly true. Usaola and Angarita (2007) analyzed the distribution of power production depending on predicted value. When plotting the frequency of occurrence for different power levels, they found out that for low or high predictions, the shape of the frequency distribution of the real production is similar to exponential, while in the medium range, the distribution is more Gaussian. Next to that, also the forecast horizon was of influence: With longer time between the forecast and realization, the probability density function tends to flatten out.
The uncertainty of the electricity prices is the other big factor. Moreno et al. (2012) state that the modelling of prices is crucial for the revenue model. Especially the ability of forecasting imbalance

prices determines the goodness of the model. Still, several articles use known prices or average imbalance prices. However, as imbalance prices can highly differ, the results from these approaches may widely differ from reality.
Bueno et al. (2010) tried to optimize the revenue for the trader on the intraday market and identified the imbalance price as highly variable and difficult to forecast parameter. Based on findings on the hourly imbalance prices throughout one year, they were able to make a heuristic approach to forecast the imbalance price. This was due to the fact that there is a daily pattern recognizable. Based on this, a mean imbalance value for each day hour was used. However, it remains to be validated if this is also the case in the Dutch energy market. Next to that, the paper is from 2010, when liquidity in the intraday market was still low. Since 2018, the European markets are interconnected, which results in higher liquidity and a lowering in variance of prices and thus difference throughout the day.
Chaves-Ávila et al. (2014) investigated the impact of different imbalance rules on European energy markets and forecast the different prices using Seasonal Autoregressive Integrated Moving Average (SARIMA). With this model, weekly and daily seasonality can be well captured. They can also forecast the day ahead, positive and negative imbalance prices in the Dutch energy market with a MAE of 4.95%, 31.35% and 34.11%, respectively.

With the knowledge how others have dealt with the different uncertainties of electricity bidding, we can introduce propositions made by academics how to determine ideal bidding volumes. Pinson et al. (2007) distinguishes between two general approaches: point predictions or probabilistic approaches.

2.4.2 Point Predictions

A point prediction strategy can be seen as base line of bidding strategies. Given a look-ahead time $t + k$, they estimate the average power output between $t + k$ and $t + k - 1$. This implies that it is reasonable to forecast the wind energy produced in this period as product of the average power production by the temporal forecast resolution t_r. Depending on the forecast horizon, the resolution can range from 15 min to 1 h. However, for the application in power system management or trading, the time resolution is usually sampled to 1h (Pinson, 2006)

$$\hat{E}_{t+k|t} = \hat{p}_{t+k|t} t_r \tag{2.7}$$

$\hat{E}_{t+k|t}$ and $\hat{p}_{t+k|t}$ refer to the energy and power forecast, respectively, depending on their issued time t for the lead time $t + k$ (Pinson et al., 2007). When there is no more further information about the future wind production, this is the volume E^b that will be bid in the day ahead market for the PTU $t + k$:

$$E^b = \hat{E}_{t+k|t} \tag{2.8}$$

We consider the current approach at DVEP as a point prediction.

2.4.3 Probabilistic Forecast

Instead of assuming that E^* is a given fact that needs to be predicted as good as possible with $\hat{E}_{t+k|t}$, we can understand this problem also in a probabilistic way. In this way we see E_{t+k} as a

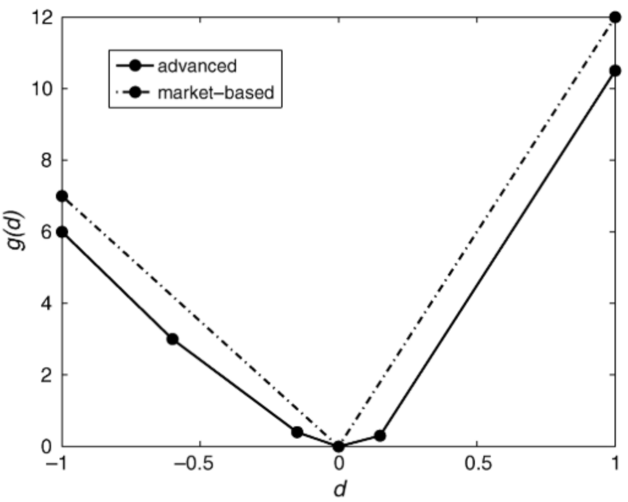


Figure 2.4: Example of two loss functions (Pinson et al., 2007).

random variable, where E^*_{t+k} is one possible realization of this variable.

Pinson et al. (2007) compared the trading results of point prediction versus probabilistic forecast on the Dutch energy market in year 2002: A simple persistence forecast is indeed the worst performing strategy, while using advanced forecasting techniques like fuzzy NN predictions increase the revenue. However, the best results were obtained when four loss functions could be defined, which include quarterly averages for both upward and downward dispatch prices. It was defined that the ideal volume was the bidding volume with the least imbalance. This can be translated in a loss function, which is defined as function that is strictly increasing, when the imbalance is unequal to zero. This is because the market participant can not expect to gain from imbalance. See Figure 2.4 as an example of the lossfunction, where the market-based function refers to average buy or sell imbalance prices and the advanced function reflects the sensitivity of a market participant on volume deviations, thus its risk appetite. On the x-axis, the imbalance is displayed in a normalized manner, while on the y-axis, the perceived loss is shown.

Based on this, they proposed two optimization situations:

- Minimization of imbalance costs, thus increasing the revenue.
- Reduction of maximal loss. In case of unpredictable weather conditions, it can be more beneficial to improve the worst possible scenario.

When the comparison of the naive forecast and advanced trading strategy is made, the persistence method realized 79.1% of the revenue of the perfect prediction and the advanced trading method accomplished 92.1% compared to the perfect prediction.

Chaves-Ávila et al. (2014) used the forecasts of the different prices to formulate an improved bidding strategy as well. Compared to bidding the expected strategy, the models incorporating the price forecast improved the average income per hour by 18%. It is important to mention that this result also includes the intraday market trading.

Zhang et al. (2012) use the assumption of normally distributed hourly wind power output. With this assumption three different models for the Spanish day-ahead market are made: They propose three different models: expected profit-maximization (EPS), chance-constrained programming-based strategy (CPS) or multi-objective bidding strategy (ECPS). Here, the EPS yields the highest revenues, however we have to notice that this is also the riskiest strategy.

Eransus (2016) applied a bidding strategy for the spanish market based on forecast imbalance length, thus whether the imbalance is positive or negative. The forecast was again made with a SARIMA and in 66% of the hours the sign of imbalance could be forecast. This strategy was compared to a point forecast 7% could be saved when only at the day ahead market is nominated. Zugno et al. (2013) considered the Nord Pool market and came up with to possible strategies: Expected Utility Maximization (EUM) and the restricted EUM. The EUM can be seen as risk neutral strategy, which can deviate heavily from the point forecast. The restricted EUM however is a compromise of those two, where the constraint can be in the decision space or probability space. A constraint in decision space means that the bidding volume may not deviate more than a defined percentage from the point forecast. On the other hand, the constraint in the probability space constrains the bids with a imbalance ratio. With this models it was found that the contrained strategies (±20%) delivered the best result.

2.5 Risk assessment in energy trading

As we consider methods to improve the bidding volume at the day ahead market, it is important to look at ways to counteract the risk of high losses. The spot price of electricity as well as the imbalance price are highly volatile, such that risk of high losses can be very high when the market has evoloved in an opposite way to the forecast. Here we introduce methods to adapt for possible risks of high losses, which are mentioned in literature.
According to Moreno et al. (2012), most articles consider Value at Risk (VaR) or Expected Shortfall (ES) as risk constraining parameter in the optimization.
The VaR is generally defined as maximum loss over a given time horizon, at a pre-defined confidence. A one month $VaR_{95\%}$ of €100,000 thus means that we are 95% certain that the maximum loss is no more than €100,000. The ES takes the mean of the interval from inf until $VaR_{95\%}$ and in this way also considers the very extreme values (Risk.net, 2018).
Based on this Moreno et al. (2012) recommends the ES as parameter for stochastic optimization. In order to include the ES in the bidding strategy, an accepted threshold of losses is defined and added as constraint to the maximization problem for each hour.

3

Current situation

In this chapter, we present the current situation at DVEP regarding the questions raised. This includes the answer to the research questions of B.1 about the performance of the wind power forecasts and the influence of weather conditions on this. Next to that, we introduce data necessary to come up with a model for subsection C.

3.1 Wind portfolio and selected wind parks

As this report is written, DVEP has a portfolio of 131 active wind parks in the Netherlands. These wind turbines have an installed power of 315 MW in total. One can see the location of these wind parks in Figure 3.1, where it becomes clear that the majority of parks are located in the western part of the country.

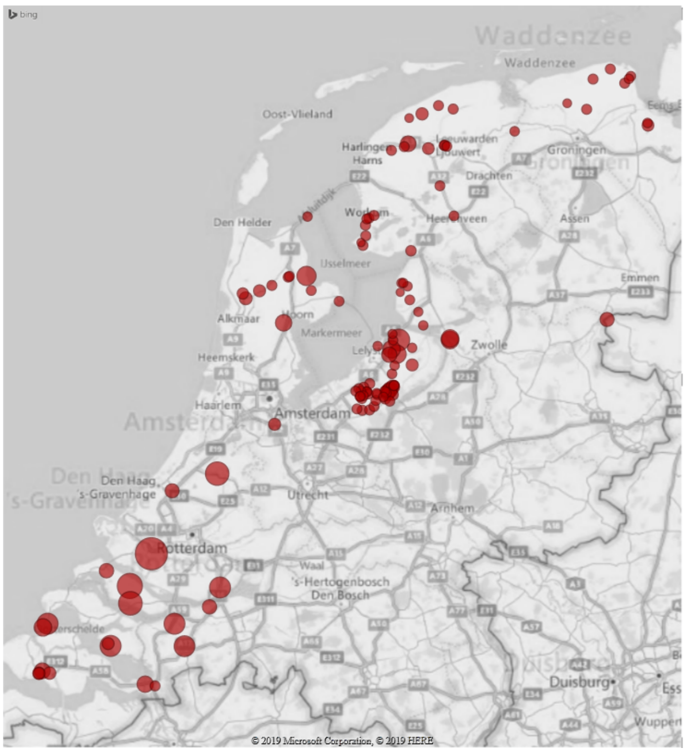


Figure 3.1: Locations of the wind parks of DVEP's clients.

Table 3.1: Selected wind parks.

	Province	Installed power (MW)	Weatherstation
Wind park 1	Friesland	4.5	De Kooy
Wind park 2	Flevoland	4	Lelystad
Wind park 3	Flevoland	2	Lelystad
Wind park 4	Flevoland	6	Lelystad
Wind park 5	Friesland	2.13	De Kooy
Wind park 6	Zeeland	2	Vlissingen
Wind park 7	Zeeland	1.75	Vlissingen
Wind park 8	Zeeland	6	Vlissingen
Wind park 9	Zeeland	2.3	Vlissingen
Wind park 10	Zeeland	9.2	Vlissingen
Wind park 11	Zeeland	5	Vlissingen
Wind park 12	Flevoland	4	Lelystad
Wind park 13	Friesland	1	De Kooy
Wind park 14	Friesland	1	De Kooy
Wind park 15	Flevoland	4.2	Lelystad
		Total: 55.1	

From these wind parks, we have chosen 15 wind parks in the three provinces of Zeeland, Flevoland and Friesland to execute the analysis. They are listed in Table 3.1 and were selected based on the following conditions:
It was important that the parks had a contract between 01-07-2018 and 01-09-2019 to have a complete training and test set. Next to that, the sizes of the wind parks should be as diverse as possible to reflect the whole portfolio as well as possible. Another important condition is that the clients chosen produce electricity only from wind and not from solar or biomass. The total volume installed of the selection is 55.1 MW.

3.1.1 Selected data

As stated, we chose data from 01-07-2018 till 31-08-2019. From which 01-07-2018 till 30-06-2019 are treated as training set and 01-07 till 31-08-2019 are used as test set. This had two reasons: Firstly because there are no forecast data of Forecaster 2 before this date. This was particularly important for Research questions A.2 and B.1.
Furthermore, on 13-06-2018 the intraday cross-border market XBID was introduced. This means that orders at the intraday market can be matched with any other similar order submitted by market participants in any other participating country. The participating countries from 2018 are Austria, Belgium, Denmark, Estonia, Finland, France, Germany, Latvia, Lithuania, Norway, The Netherlands, Portugal, Spain and Sweden (EPEX SPOT, 2018). The consequence of this cross-border market is an increase in liquidity in the intraday market, which improves the efficiency of the electricity markets. According to the traders here at DVEP, this has resulted in a decrease of variance in imbalance prices.

3.2 Forecasts

DVEP has both wind power forecasts and weather forecasts at its disposal. For the wind power forecast, two different parties forecast the production of each particular wind park up to 3 days ahead, from which the forecasts up to 38 hours are used for the day ahead forecasting. Due to confidentiality, these forecasters are named Forecaster 1 and Forecaster 2 in this report. While Forecaster 1 has been used since the first years of DVEP, Forecaster 2 has been recently added in June of 2018 to increase confidence about the forecasts. The forecasts used for the following day are received at 09:00 h and prepare the bidding for the day-ahead market. The output of both forecasts is the wind power production for each wind park individually in MWh per day hour. Both forecasts have a short to long term forecast horizon and thus use a physical forecast method. The weather forecast, on the other hand, is used to tell more about the general weather conditions. This includes wind speed and direction, as well as temperature, precipitation and, radiation. At 15 weather stations by the KNMI weather forecasts are made based on different weather models. As we have selected three areas with wind parks, we use three weather stations close to the wind parks to retrieve the weather data.

- Zeeland: Vlissingen.
- Flevoland: Lelystad.
- Friesland: De Kooy.

3.3 Historical data wind

In order to discuss the performance of the forecasts, we have to analyse the historical data, under the expectation that future situations will be similar to the past. Figures 3.2 and 3.3 give an impression of the distributions of both the production of the wind parks and the wind speed.
Figure 3.2 shows the distribution of the total production of the portfolio, which makes clear that 75% of all hours have a production of less than 21.9 MWh. The maximum production for one hour was found to be 55.1 MWh, which is the installed power of the portfolio.
While the distribution of the production can be fitted to a negative exponential distribution, the wind speed distribution can be fitted to a Weibull distribution, as can be seen in Figure 3.3. The mean wind speed was found to be 5.4 m/s. The distributions are in line with findings from the literature (Pinson, 2006). Here, we create one wind speed for each hour from 01-07-2018 till 01-07-2019 by taking the mean of the measurements of the three weather stations. It is important to consider that the measurements at the weather stations are taken at a height of 2m, while the hub height of the wind turbines is between the 40 and 135m depending on the type. In order for most wind turbines to produce electricity a minimum wind speed of 2.5 m/s is necessary, while the maximum wind speed for the most wind turbines is 25 m/s.
A power curve with individual data for several different wind turbine types in the portfolio can be found in Appendix A, Figure A.1.
However, there were no measurements of wind speeds higher than 25m/s in our time horizon. This does not mean that they did not occur, due to the difference in the height of the weather station compared to the hub height. According to the weather station measurements, the maximum wind speed was reached at the station in Vlissingen with 22 m/s, which could mean that at hub height

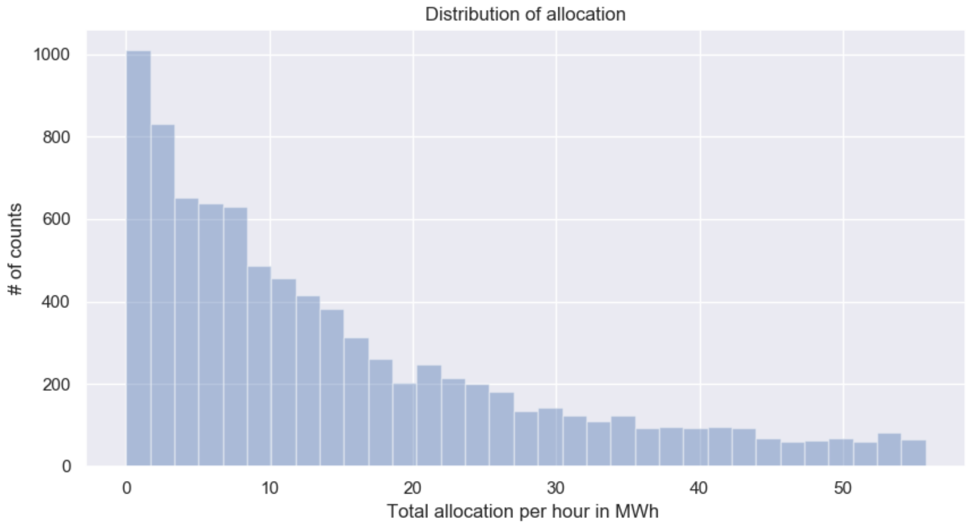


Figure 3.2: Histogram of the production.

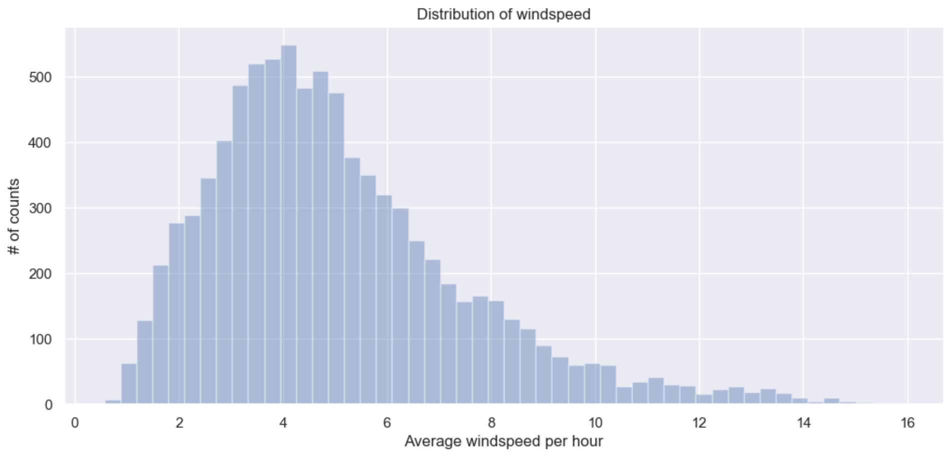


Figure 3.3: Histogram of the wind speeds.

25m/s was exceeded. This pitfall of measurements should be kept in mind. As you can see, this is not displayed in Figure 3.3 since we chose here for a mean wind speed of the three stations.
The correlation of the wind speed can be found in table 3.2 and as expected the Pearson coefficients indicate a correlation between the stations, with a higher correlation between De Kooy and Lelystad. This is due to the geographical proximity (62km) of these two stations, while Vlissingen is much further away (De Kooy: 183km; Lelystad: 175km). This was the reason, why we decided that a mean wind speed is appropriate, the individual histogram per wind park can be found in Appendix A, Figures A.2, A.3, A.4.
Figure 3.4 shows the distribution of occurrence of the wind directions. The wind direction is measured by the weather stations in degree, where 0° and 360° correspond to North, while 90° corresponds to East. The other directions are accordingly. To simplify, all wind direction within the interval ± 45° account to the corresponding wind direction, such that all wind directions from 315 ° till 45° correspond to North, for example.
As can be expected, the most frequent wind directions are West and South, which corresponds

with the geographical location of the Netherlands.

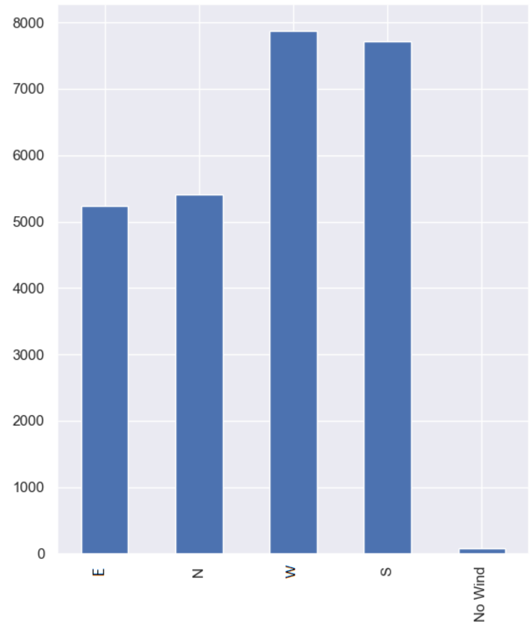


Figure 3.4: Barplot of the wind directions.

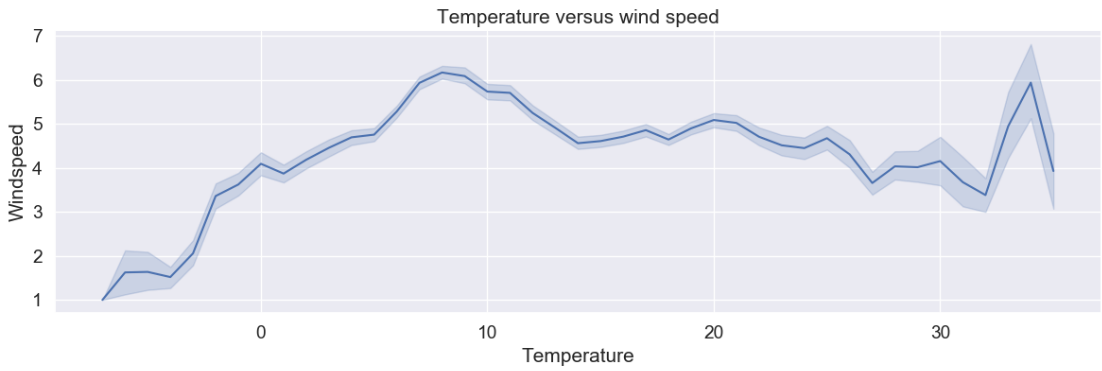


Figure 3.5: Distribution of mean wind speed versus temperature.

Figure 3.5 shows the mean wind speed per hour for the different temperatures for the previous year. As expected, low temperatures lead to low wind speeds and around 9 °C a maximum is found. After that, wind speeds slightly decrease with one remarkable peak at temperatures above 30 °C. This peak is no wrong measurement, but is validated with weather data by KNMI, but still can be seen as a extreme value. We can predict a plot of error metrics, which follows the plot closely.

Table 3.2: Pearson correlation coefficients of wind speeds of the different weather stations.

	De Kooy	**Lelystad**	**Vlissingen**
De Kooy	1.00	0.86	0.73
Lelystad		1.00	0.75
Vlissingen			1.00

3.4 Forecast performance

One goal of this report is to compare the performance of the two forecasts available at DVEP. There is the hypothesis that one of the two forecasts can perform better under certain weather conditions. However, this is not yet quantified. To quantify that, we use the error measures, introduced in the theory Section 2.3: Normalized BIAS (NBIAS), normalized mean absolute error (NMAE) and normalized root mean squared error (NRMSE). With the NBIAS we seek to find structural trends of under or over estimation for the respective condition. With the NMAE we try to express the performance of the forecast, as it takes away the signs of the forecast error. The NRMSE has the same approach as the NMAE, with increasing the weight of extreme outliers due to the use of a square.
The calculations are performed for both each individual wind park of our selection as well as for the overall portfolio, depending on four different conditions, which were assumed influence the wind power forecasting performance:

- Wind speed.
- Day hour.
- Wind direction.
- Temperature

Wind speed is crucial in the forecasting of wind power, as the wind is the force to propel the wind turbine, thus producing the power. In consequence, this means that it is very important to have the wind speed correctly forecast. It is interesting to see if one of the forecasts outperform the other at certain wind speeds.
The day hour condition is included, as it can be expected that forecast accuracy drops with an extended forecast horizon. This means small mispredictions will be amplified at the longer horizon.
There is the impression that at certain wind directions, there is a bigger difference between the two forecasts than at other directions. This is why also the wind direction has been included in the set of conditions.
Next to that, Hesselink (2018) suggests in his master thesis at DVEP, that Forecast 1 has worse performance under colder circumstances. This is why we also add this parameter in the list of conditions.

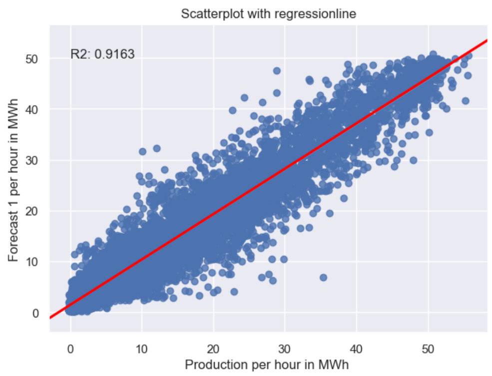
3.4.1 Overall results

Table 3.3: Results without any condition (from 01-07-2018 till 01-07-2019).

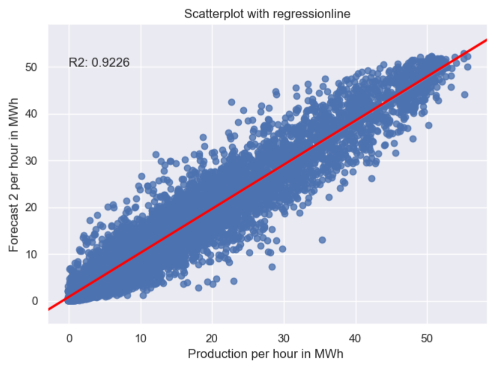
	Forecast 1	**Forecast 2**
NBIAS	0.0034	0.0012
NMAE	0.0747	0.0713
NRMSE	0.1131	0.1094

In Table 3.3 the results are presented we found without the use of any condition. This shows that both forecasters performed very similar, with the Forecast 2 slightly better. Both have almost no bias in their forecast and a NMAE of 7.47% for Forecaster 1 and 7.13% for Forecaster 2. The

NRMSE is 11.31% and 10.94%, respectively, such that the difference of both remains the same compared to the NMAE. With this, we can conclude that both forecasters have a comparable pattern.



(a) Forecast 1.



(b) Forecast 2.

Figure 3.6: Scatterplot of bidding versus production per hour.

Figure 3.6 shows the production on the x-axis versus the forecast on the y-axis, where Plot 3.6a refers to Forecast 1 and Plot 3.6b to Forecast 2. In both plots, the linear regression line is added to compare the outliers of both forecasts. To do so, the R^2 (coefficient of determination) was calculated, with which we can evaluate how close the data points lie at the regression line. An R^2 of 1 means that all points lay on the regression line. Here, Forecast 1 has an R^2 of 0.916, while the R^2 of Forecast 2 is slightly higher, 0.923. From this, it follows that both forecasters have only a few outliers from the theoretical linear relationship of production and Forecast 1 and Forecast 2, respectively. Next to that, it can be seen that at both ends the scatter points of forecast and production are very close to the regression line, while in the middle the distance is bigger. This implies bigger errors at medium power output.

As the values are very close likely, both forecasts are statistically the same. In order to test this, we use the Mann-Whitney U test to test whether they are indeed statistically the same. The Mann-Whitney U test is a non-parametric test to check if two independent data sets have the same distribution (Brownlee, 2018). The null hypothesis states that the distributions of both samples are the same. We use this test because we can see in Figure 3.2 that the distribution of production is far from being normally distributed and it can be assumed that the forecast follows a similar distribution. We can make this statement since the error measures are very low. Furthermore, both forecasts are independent, which rules out the Wilcoxon test. Based on the Mann-Whitney U test, we can say with a confidence of 95% that based on the individual wind parks, the forecasts are significantly different, the p-value is below $\alpha = 5\%$, thus we have to reject the null hypothesis. However, when considering the portfolio as a whole, with one forecast volume per PTU, we fail to reject the null hypothesis, both forecasts are thus the same.

Next, we checked if the errors of both forecasts can be called significantly different. To do so, we compared normalized e_{t+k} of Forecast 1 with the one of Forecast 2, which could be seen as comparison of both NBIASs. Next to that, we executed the same comparison for the normalized

$|e_{t+k}|$ and normalized $e_{t+k}{}^2$, which correspond to NMAE and NRMSE, respectively.
The conclusions of these statistical test are the following: Based on the normalized e_{t+k} we can not conclude that the distributions of the forecasts errors are significantly different. However, when we test for the normalized $|e_{t+k}|$ and normalized $e_{t+k}{}^2$, we can conclude based on the Whitney Mann U test, that the populations have different distributions.
To be able to give recommendations which forecast to use based on this yearly data, we also considered the situations were Forecaster 1 and 2 differed significantly in prediction. On average the difference was 2.7 % of the installed capacity (or 1.504 MWh), however, the 5% biggest differences were on average 10.1% (or 5.566 MWh). At this 5% biggest differences, we compared the NMAE of both forecasters. In these situations of big difference, the NMAE of Forecaster 1 is 9.6% while Forecaster 2 has a NMAE of 8.1%.

3.4.2 Results depending on wind speed

In Figure 3.7 we see the errors of the prediction depending on the wind speed. We have excluded the extreme wind speed values as they disturbed the graph with high error values, which made the remaining data points difficult to identify. This step can be justified as the number of wind speeds above 19 m/s is very small as we could see in the histogram of Figure 3.3.

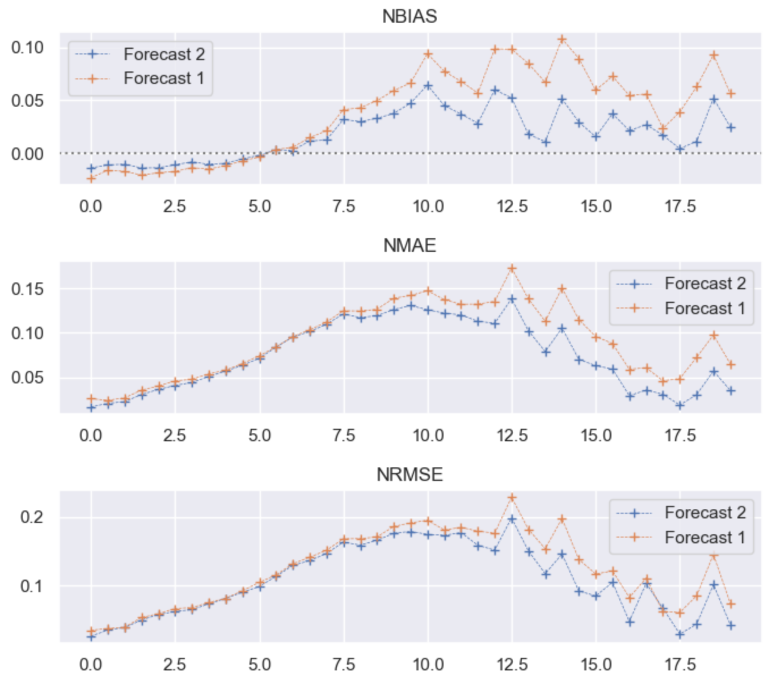


Figure 3.7: Results of the error measures with respect to wind speed. On the y-axis we have the relative error, on x-axis the wind speed.

When we analyze the first graph, it is clear that up to a wind speed of 10 m/s at the respective weather station, the NBIAS of both forecasts is very similar. From having a slight overestimation under 5m/s, the NBIAS proceeds to have an underestimation until 10m/s. This underestimation

remains, although it fluctuates slightly with Forecast 2 being closer to 0 than Forecast 1. The fluctuations can be explained with the few number of situations where wind speeds exceed 10 m/s. Of the 8760 records, only 469 (or 5.3%) exceed the 10 m/s. Another explanation can be found, when we take a look at the power curves of wind turbines, which can be found in Appendix A, Figure A.1. At higher wind speeds the power curves flatten out, such that deviations in the wind speed forecast have only a minor influence on the wind power forecast. This can explain the decrease of the NBIAS at higher wind speeds.
With the second graph we can give an answer regarding the performance of the forecasts: Up to a wind speed of 10 m/s, the NMAE is increasing, which is around 15%. This increase is of linear sort and according to our expectations: With higher wind speed, the deviations of the forecast versus the actual production get bigger, what results in a decreased performance of the forecast and higher NMAE and NRMSE. However, the better error measures for wind speeds exceeding 10 m/s do not coincide with this reasoning. But when the power curve is considered again, the results can be explained: The power curve is flattening at this wind speeds, which makes errors less impacting. We assume that similar reasoning then for the NBIAS can be given here. It would be interesting, how this graph would look like in case the number of data points would be evenly distributed over the range of wind speeds. Based on the progress of the graph from 10 to 12.5 m/s, it can be expected that both NMAE values remain to differ a little and the error level is slightly decreased or stays constant.
The RMSE shows a very similar pattern as the NMAE, only slightly shifted to higher error values, as expected. This is why we do not analyze this plot in further detail.

3.4.3 Results per day hour

Figure 3.8 shows the result for the three different error measures under the condition day hour. We define these day hours as the following: hour 0 refers to 00:00 till 01:00, while the last hour is hour 23, from 23:00 till 24:00 In the first graph, we see the NBIAS of both forecasts, which are close to each other and follow the same pattern. Until hour 10, they both slightly overestimate the productions, so the forecast volume is too big, after that the production is underestimated. This remains until hour 19, from which the NBIAS of Forecaster 1 is almost at 0, while for Forecaster 2 is slightly negative. With these results we can tell something about the bias of the forecast for the different hours, however, we can not make a judgement about the performance of the forecast since the individual results per hour are averaged. In order to tell more about the performance itself, we can consult the NMAE and NRMSE. Therefore it is also logical that the error for NMAE can be in the range of 6.5 % - 8.5 %, while the NBIAS has an error between -0.5% and 2.0 %. Based on the NMAE, we can state that Forecast 2 is outperforming Forecast 1 at all times apart from hours 2 to 4. We can make this statement since Forecast 2 is always closer to 0 than Forecast 1, and thus closer to the realized production. The third metric supports this statement by taking extreme deviations more into account. But, apart from slightly higher errors, the same pattern is identifiable. As NMAE and NRMSE are increasing throughout the day, we could state, that with a longer forecast horizon, the accuracy of both forecasts decreases. However, this does not explain the increase in accuracy after hour 18. In order to explain this drop, we can consult the distribution of average wind speeds throughout the day (Figure 3.9). At the hours of daylight, from approximately hour 8 until hour 18, the average wind speed is higher than at night time.

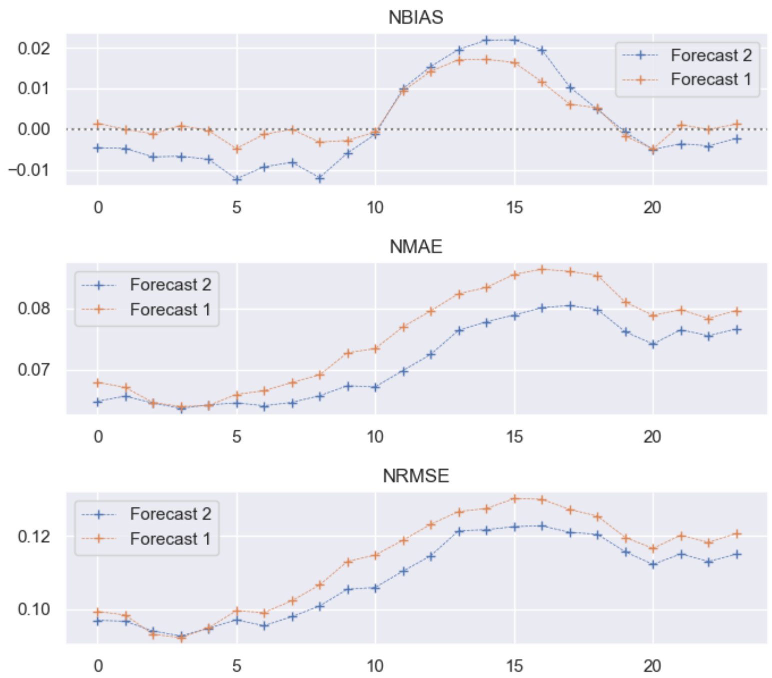


Figure 3.8: Results of the error measures with respect to day hour. On the y-axis we have the relative error, on x-axis the day hour (0 refers to 00:00 till 01:00).

From this, it follows that the increased error values at daylight are mainly caused by the relatively higher wind speeds in these hours. It needs to be stressed, that this only applicable for the average cases: From Figure 3.9 we can find the maximum mean speed to occur in hour 14, which is 6m/s, where the slope is steeper compared to lower wind speeds. Despite that, when the wind speeds increase further, the accuracy again increases, as we already noted in the previous subsection. However, when we disregard the upshift during the day hours, we can still observe an increased error with the plots of NMAE and NRMSE. This error is caused by the increased lead time and according to the expectation. A difference compared to the wind speed plots can be found in the magnitude of errors for both NMAE and NRMSE: For the day hours, they are much closer to the values obtained from the general case without any restriction.

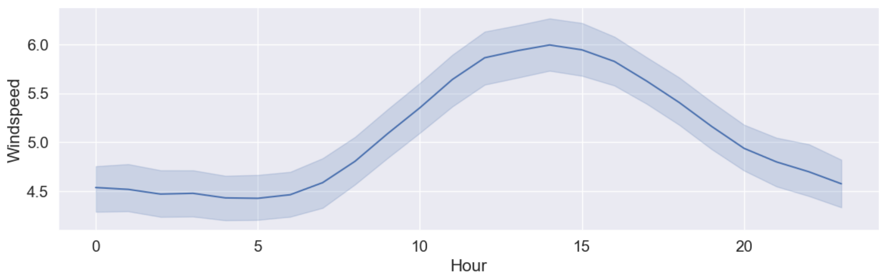


Figure 3.9: Windspeeds per day hour, 95% confidence interval around it.

3.4.4 Influence of wind direction

Figure 3.10 shows the different error measures when grouped by the wind direction. It can be seen that the values differ only little from the results without any condition (Table 3.3). However, it can be noticed that going from East to West, the NMAE as well as NRMSE increased from 6% to 9% (and from 10% to about 14%). Next to that, are the error metrics of both forecasters very close together, only East and West show the Forecast 2 in favor slightly. Thus we can not assign one forecaster to be significantly better at certain wind directions compared to the other one. The findings are in line with the statements made above. The NBIAS is only very small, about ± 0.5%

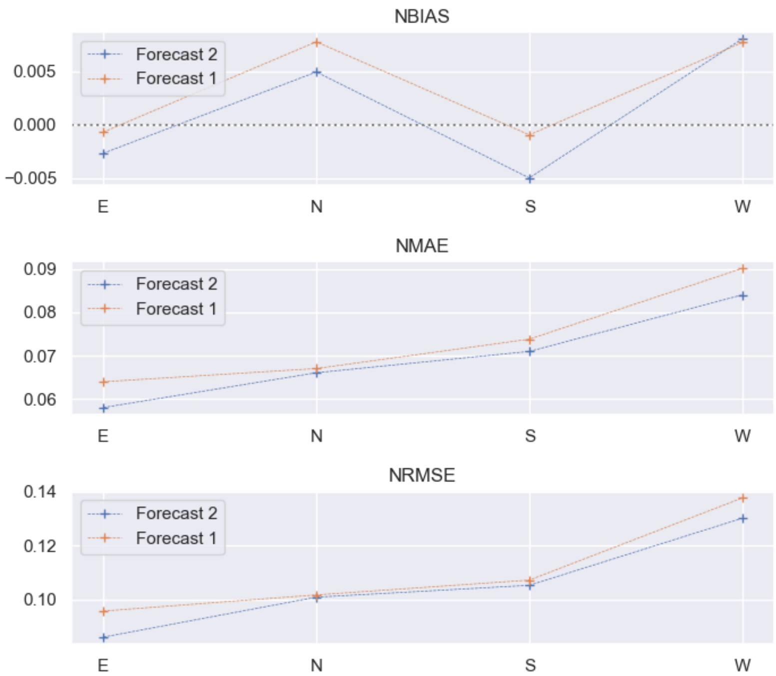


Figure 3.10: Results of the error measures with respect to wind direction. On the y-axis we have the relative error, on x-axis the wind direction.

3.4.5 Influence of temperature

Figure 3.11 shows the accuracy of the forecasts when grouped by temperature. When investigating the NBIAS, it gets clear that at low temperatures below 7 °C there is an overestimation of production, where after the trend is very close around zero, such that no specific error can be found. The outliers above 33 °C can be explained by the low amount of scenarios. An underestimation in this case of 6% is still in the range of reasonable values, when we compare them to the findings depending on wind speed.

When considering the NMAE, it gets apparent that with low temperatures, as well as high temperatures, the accuracy increases. This can be reasoned with the fact that in these situations, wind speeds are generally lower due to weather. As expected earlier, NMAE and NRMSE follow

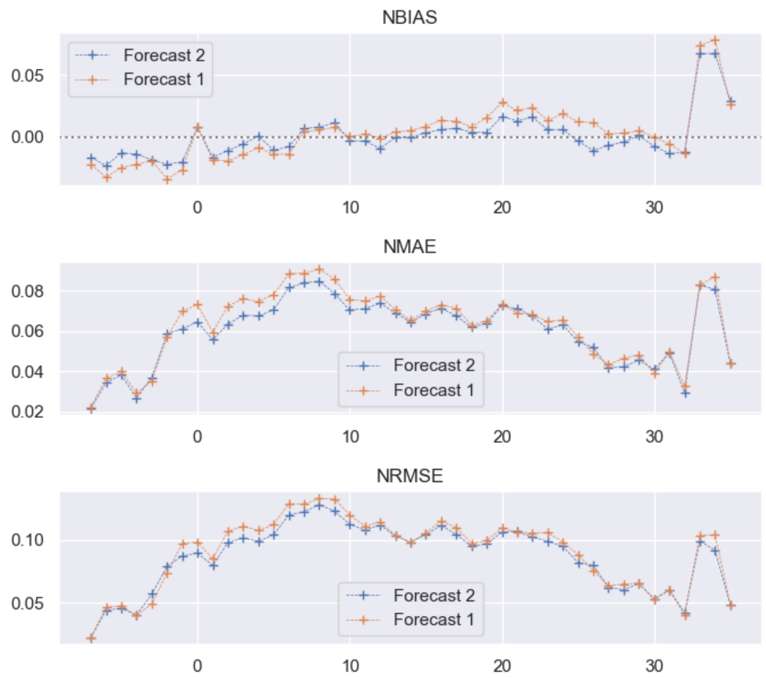


Figure 3.11: Results of the error measures with respect to temperatures at the weather station.

closely the plot of wind speed versus temperature, Figure 3.5. This made also the outlier at high temperatures predictable. However, this condition does not reveal a better performance of one of the two forecasts. For all three error metrics, the differences between Forecaster 1 and 2 are very low.

3.4.6 Individual windparks

Next to the overall wind power forecast performance, we realized that also the performance of the forecast at individual wind parks can be very interesting. While the overall numbers show error measures that are in line with the findings from earlier research, it became apparent that some individual wind parks did score worse. Worse translates in this case especially to fluctuating performance of the error measures and significantly higher errors. When there is a high fluctuation in the error values, it is difficult to detect trends in the forecast. This makes it difficult to extract these trends from the results. A high prediction error compared to the overall error implies some difficulty which is characteristic for the wind park. However, it needs to be investigated, what makes these wind parks score so badly. This can be due to the wind park or the forecaster and it is desirable to reduce this error.

From the wind parks of the portfolio, the **Wind Park 15** shows a pattern which is different from the rest of the portfolio: In Figure 3.12*(NBIAS)* we see that Forecast 1 is constantly underestimating production, while Forecast 2 overestimates production. However, in the overall portfolio the forecasts follow much more the same pattern. The same holds for the absolute error of the wind power forecast, while it is also 3-5 %-points higher then the overall NMAE, it is also fluctuating quite heavily. However, when we investigate the performance with respect to the wind speed, no

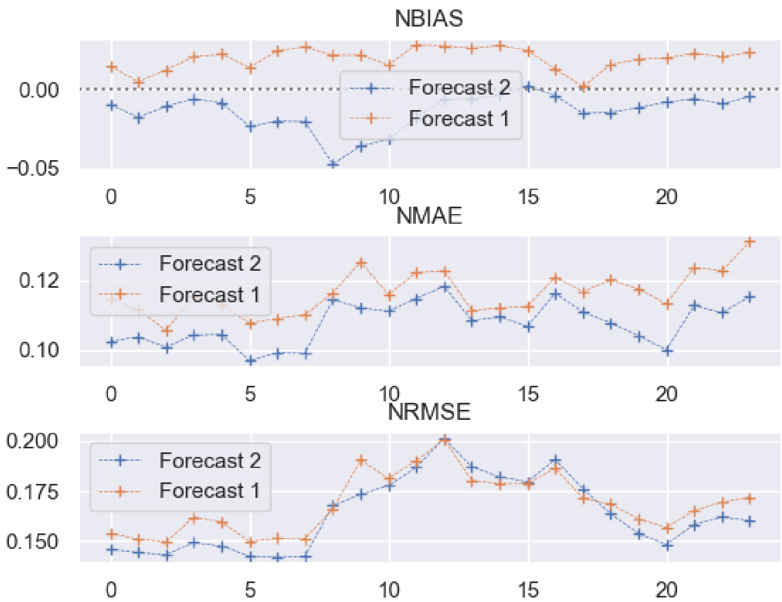
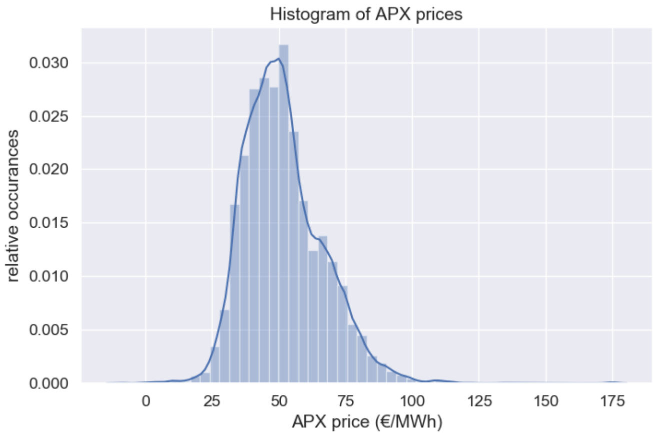


Figure 3.12: Results of the Wind Park 15 regarding dayhour.

such significant deviations can be found. All three error measures are in the same range and display the same patterns.

3.5 Market data

We now have an idea about the performance of the two different forecasters for DVEP. However, another important part of the current situation is the analysis of the current market situation. In Chapter 2.1 we have introduced the Dutch energy market and how it is composed. In the following, we explain the composition of both APX price and imbalance prices. Figure 3.13a shows the distribution of the APX price of 2018/2019. We can see that it can be fitted with a normal distribution which is slightly positively skewed. When investigating the statistics, we note that the mean price is 51.6 €/MWh and the standard deviation is 14.8 €/MWh.
In Figure 3.14 we show the mean APX prices on a daily, weekly and monthly basis. We notice that during the day, there are two peaks at hour 9 and hour 19. These are the moments when people go to work or return home. The night hours have the lowest prices, while during the day, the prices are situated between those two values. The second graph shows the progression of the average price throughout the week. Based on this, we can state that working days show higher prices, while the price drops in the weekend. This price difference, however, is smaller than throughout the day. Similar price differences can also be found when comparing the monthly average prices. Here we find months March, April, May and June to have the lowest prices, while September to January have much higher prices. Remarkably, that both July and August do not show such low prices (compare with Table 5.1 to see prices of 2019). This can be explained with the market situation back than: Coal and CO_2 prices were very high, which increased the marginal cost for electricity production with coal and low water levels in rivers created supply shortages of power plants.
Moreover, it is also interesting to check the distribution of the imbalance prices feed and consume.



(a) Spot price overall

(b) Spot price with peak/ offpeak distinction

Figure 3.13: Histogram of the APX spot price from 01-07-2018 till 01-07-2019.

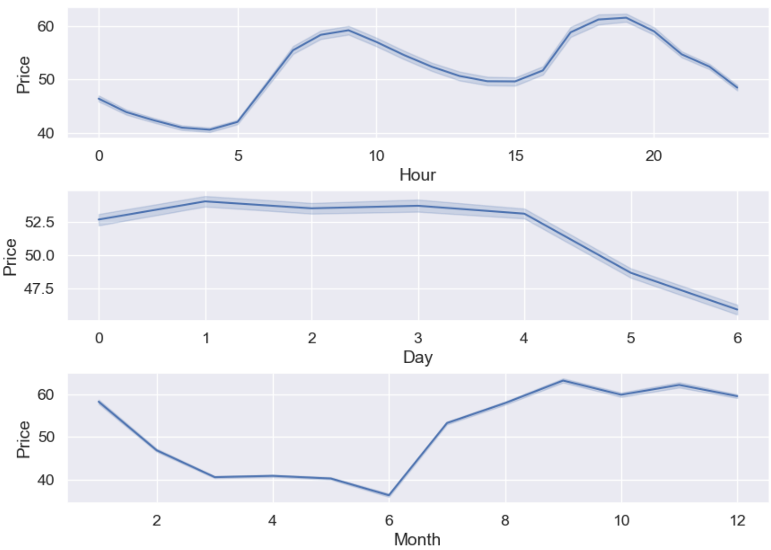
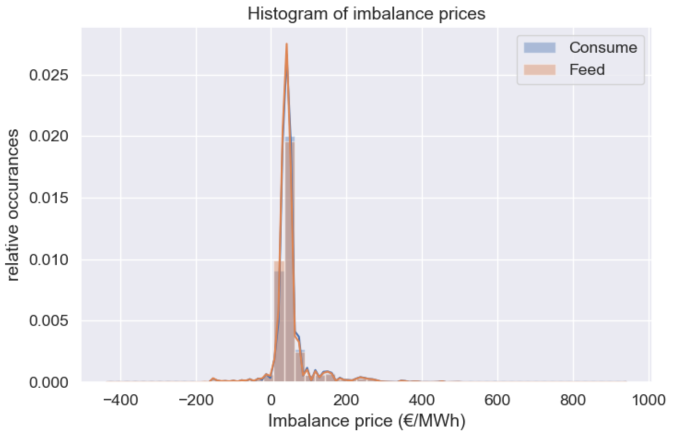
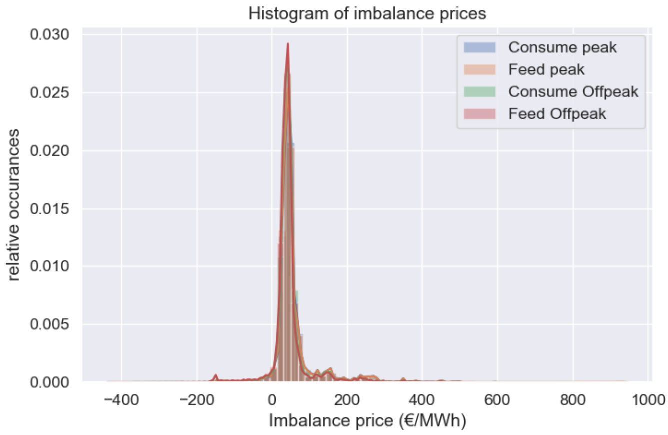


Figure 3.14: The mean APX price, versus dayhour, weekday (0=monday till 6=sunday), and month. With 95% confidence interval.



(a) Imbalance price overall.



(b) Imbalance price with peak/ offpeak distinction.

Figure 3.15: Histogram of the imbalance price from 01-07-2018 till 01-07-2019.

The mean imbalance price for consume and feed are 52.1€/MWh and 50.1€/MWh, respectively, but both values can take extreme values of -429.95 €/MWh and 936 €/MWh. In Figure 3.15a the distribution of the imbalance prices can be found. Although the statistics show a high standard deviation, the plot is very narrow, which can be explained with the extreme values. This makes it unfeasible to fit a theoretical distribution. Based on the difference in demand of electricity between workdays and weekend and day- and nighttime, we add peak and offpeak hours, which we define as the following:
Peak hour:

- workdays between 08:00 and 20:00.

Offpeak hour:

- weekend and public holiday, whole day.
- workdays between 20:00 and 08:00.

To analyse the influence of this distinction, the distributions of spot price and imbalance price were plotted depending on peak/ offpeak hours.
In Figure 3.13b we can see that the spot price of the peak hours are slightly shifted to the right, such that there is a distinction between the distributions of the two prices noticeable.
However, when investigating the histogram of the imbalance prices in Figure 3.15b, again with and without distinction between peak and off peak hours, this differentiation is really difficult to make. This finding can be substantiated with Figure 3.16, which shows the boxplot of feed imbalance prices throughout the day. Due to the many outliers, the deviation in the mean prices per hour is hardly recognizable. We expect that the consume imbalance shows a similar boxplot.

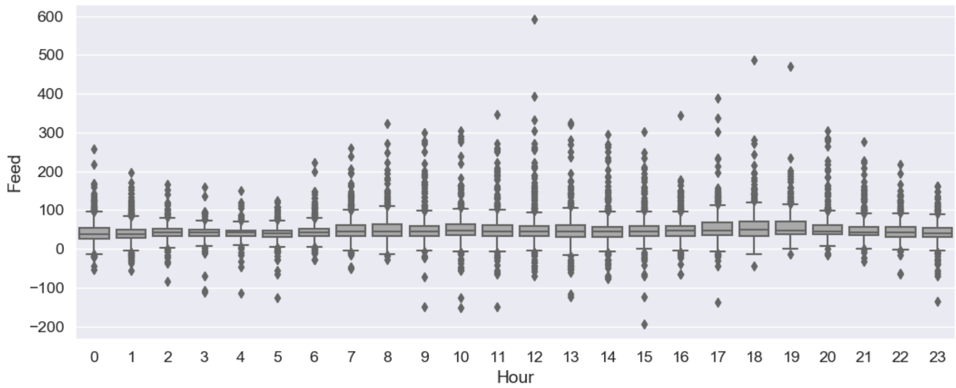


Figure 3.16: Boxplots of the feed imbalance price, throughout the day.

3.6 Conclusion

This chapter discussed the performance of two different wind power forecasters to find out, which of the two is preferred as a forecaster to base the day ahead bidding volumes on. After this, we investigated the influence of different conditions on the performance of the respective forecaster. We answered Research Question A.2.2 and B.1. Next to that, the obligatory analysis about the current situation has been carried out, in the first place to assist the Research Questions that are answered here, in the second place to prepare for the solution of Research Question C.1.
Based on the chapter we can draw the following conclusions regarding the research questions:

- When the forecasts are compared on a yearly scale, the results are statistically indifferent from each other, when the whole portfolio is considered. When we compare the individual results of the wind parks, this can not be concluded. However, based on situations of the 5% greatest differences between both forecasters we can conclude that Forecaster 2 has a lower prediction error.
- Wind speed, day hour, wind direction and temperature were used to analyze the influence of these parameters on the wind power forecast. It can be concluded that both wind speed, as well as day hour, have an influence on the forecast performance: With increasing wind speed the error of prediction increases, but as the wind turbine reaches a maximum power output at a wind speed characteristic for the wind turbine, the error reduces.
  The forecast accuracy reduces throughout the day but also reflects the distribution of wind speed over the day very well. This is because during daylight hours the wind speed is higher compared to night hours.
  The distinction between wind direction did not deliver better predictions for certain wind directions.
  When the temperature is used to split up the error metric, a similar pattern is recognized then for the day hour: There is a lower error for low and high temperatures, however, this is caused by the fact that wind speeds are lower in this situations. The temperature is thus only implicitly responsible for these results.
- Based on the statements before, we can conclude that in situations of wind speeds above

10 m/s Forecast 2 is slightly better than Forecast 1. When we base our conclusions on the day hours, we can conclude the same: Forecast 2 outperforms Forecast 1. When the wind direction is used as a condition, both forecasters perform similar to the yearly scale and no distinction between both forecasters can be made. The conclusions of the condition based comparison of the forecasters are thus in line with the conclusion made before: It can be advised to use Forecaster 2 as the leading forecaster.

- We found remarkable results for one wind park when we consider the day hour. Reason for that can be the location of the wind park, but this was not investigated further. Still, this shows that it can be wise to monitor the performance of wind power forecasts also on wind park level, to find wind parks which are complex to forecast.

4

Possible solution

In this chapter of the research, we have the goal to find an approach for the ideal bidding volume at the day-ahead market. With this knowledge, we can answer the research questions of section C. Currently, the volume of Forecast 1 is always used as bidding volume, with small deviations in case of big differences between Forecasts 1 and 2. Here, we have the goal to improve the revenues of the day-ahead market by using historical data and also data from the second forecast. It is important to realize, that forecasted volume does not necessarily need to be the bidding volume. This is why we distinguish between forecasted volume, which is the production volume that is forecasted by the two forecasters for the respective hour of the next day, and the bidding volume, which is the volume of electricity the trader bids in at the day-ahead market. The following bidding volume strategies have been tested:

- Perfect forecast.
- Persistence forecast.
- Forecasts 1 and 2.
- Average of both forecasts.
- Trend correction.
- Optimization based on historical data.

In the following, we dive deeper into the methods of the different optimization techniques. The perfect, as well as the persistence forecast, are used as benchmark strategies, where the perfect forecast is a forecast of zero imbalance, which means it is an only theoretically achievable forecast. The persistence forecast, however, is a commonly used naive forecast method used for comparison. This means the complete day-ahead bidding will be done with the last measured production volume, which is in this case the data of 11:00 of D-1 (Pinson, 2006) In this way we can analyze if the proposed bidding strategies improved this naive strategy. Next to that, we use the forecasted volumes of both forecasters as bidding volume.
Another option is the average of both forecasts, such that mispredictions of one or another forecast can be damped by the other one. By averaging the two different forecasts, we seek to improve the revenues by reducing imbalance volumes. The same approach is used when correcting the forecast with the bias. All these methods do not consider the day-ahead and imbalance prices but try to improve the performance by reducing the imbalance cost. In the last method, price data are also included by creating multiple scenarios based on historical data.

4.1 Profit structure in the electricity market

As we recall from Section 2.4, the revenue per PTU is determined by the following formula:

$$R_{t+k} = E^b_{t+k}\lambda^{APX} + I^C_{t+k} \tag{4.1}$$

The imbalance cost I^C can be found with the following (The subscript $t+k$ is omitted for clarity.):

$$I^C = \begin{cases} \lambda^{sell}(E^* - E^b), E^* > E^b \\ \lambda^{buy}(E^* - E^b), E^* < E^b \\ 0 \qquad\qquad\qquad\quad , E^* = E^b \end{cases} \tag{4.2}$$

Figure 4.1 shows an exemplary revenue graph with parameters which can be found in Table 4.1. These values are not based on actual data, but should only represent the different situations.
The x-axis shows the downsampled bidding volume in one PTU of 15min. The actual bidding has to be done per 1 hour, but as the imbalance prices are issued per 15min, we consider this PTU. This volume can range from 0 MWh to a maximum of 13.77 MWh. This is based on the maximum installed capacity of 55.1 MW. On the y-axis, the revenue is presented, which is calculated with Equation 4.1 where we note the discontinuities at the actual production for this quarter. These kinks are caused by the dual pricing characteristic of this PTU. This makes clear that the ideal bidding volume will be either minimum or maximum installed power. In case the APX spot price lies between λ^{sell} and λ^{buy}, the ideal volume is the production volume, as it can be easily read in Figure 4.1. However, this situation occurs only rarely, compared to the instances were both imbalance prices are found on one side of the APX spot price. In situations in which the feed equals the consume imbalance price, there is indeed no discontinuity at the actual production volume. This happens only in the regulation state 2, which occurs much less frequent than the other states.

Table 4.1: Parameters for graph in Figure 4.1.

	APX$>\lambda^{buy}>\lambda^{sell}$	**APX**$<\lambda^{sell}<\lambda^{buy}$	$\lambda^{sell}<$**APX**$<\lambda^{buy}$
Realized production	2.308 MWh	4.218 MWh	4.218 MWh
APX	28.98 €/MWh	25.98 €/MWh	28.98 €/MWh
Imbalance buy	24.34 €/MWh	32.34 €/MWh	32.34 €/MWh
Imbalance sell	17.51 €/MWh	29.51 €/MWh	12.51 €/MWh

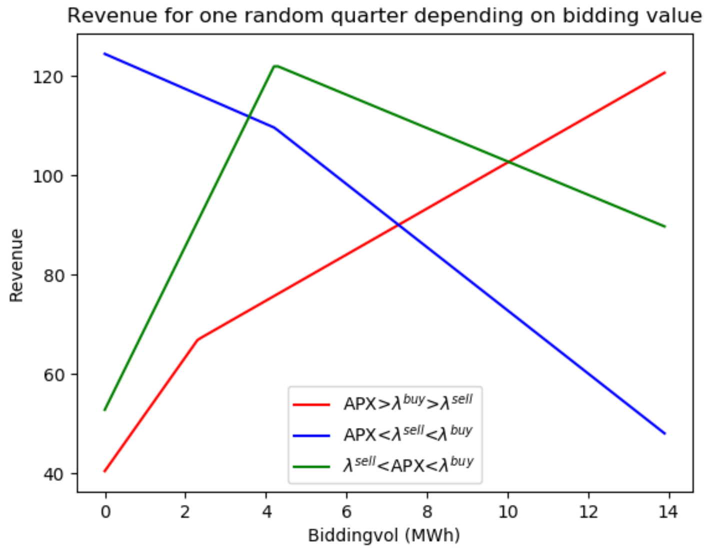


Figure 4.1: Revenue of one random quarter depending on bidding volume.

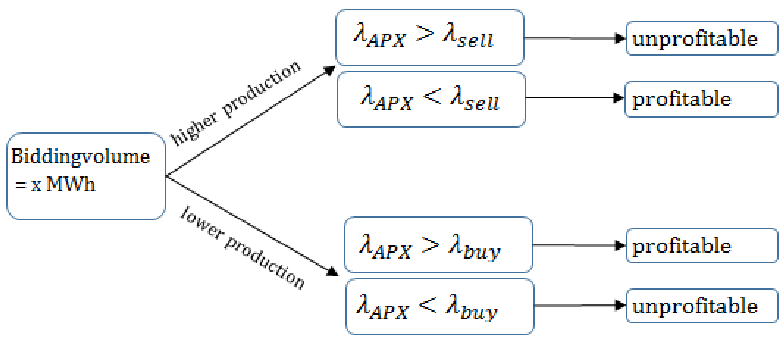


Figure 4.2: Different scenarios in case of higher or lower production versus the bidding volume.

4.1.1 Relation imbalance and APX price

Figure 4.2 shows the consequences of a higher or lower production compared to the bidding volume. With this, we can understand better which situations are more profitable. When we want to understand this graphic, it is important to keep in mind that both the APX price, as well as the imbalance prices sell and buy are uncertain and not known at the moment of bidding. The scenario of higher production can occur due to higher production compared to bidding volume. In such a scenario there is a surplus of electricity that needs to be sold on the imbalance market. To make an extra profit here, λ^{sell} has to be higher than the λ^{APX}. Else the profit is lower compared to a situation where all positions are sold at the day-ahead market, although the price at the imbalance market is still positive. In case the imbalance price gets negative due to high overproduction, this will cause a loss at the imbalance market.

Similar to the higher production scenario, the lower production scenario can also occur due to a forecasting error. In this case, we seek the possibility to buy the outstanding volume at the

imbalance market for a lower price than the APX price from the day-ahead market. In this situation, the most ideal situation is a negative imbalance, as we get even paid to take electricity from the grid.
When we look at the imbalance data of the last year, it gets apparent that there were more moments that the imbalance price is lower than the APX spot price. We assigned -1 to PTUs if the imbalance price is lower than the APX price, whereas we assign +1 to PTUs if the imbalance is higher than the APX price. These situations are grouped into bins of 0.5MWh, as can be found in Figure 4.3. A value of -0.5 means thus that in only 25% of the PTUs, the imbalance price was higher than the APX price. This holds for both the feed as well as the consume imbalance. However, the consequences are different for both cases: The feed imbalance price is needed in cases that the production is higher than expected. Given an imbalance price lower than the APX price is less profitable, because the surplus energy has to be sold at a lower price. The consume imbalance price refers to situations of lower production, such that extra volume has to be bought in order to meet the bidding volume of the day-ahead bidding. Here, a lower price at the imbalance market is thus profitable. This finding suggests thus that it is, on average, better to bid less to harvest the lower consume prices.

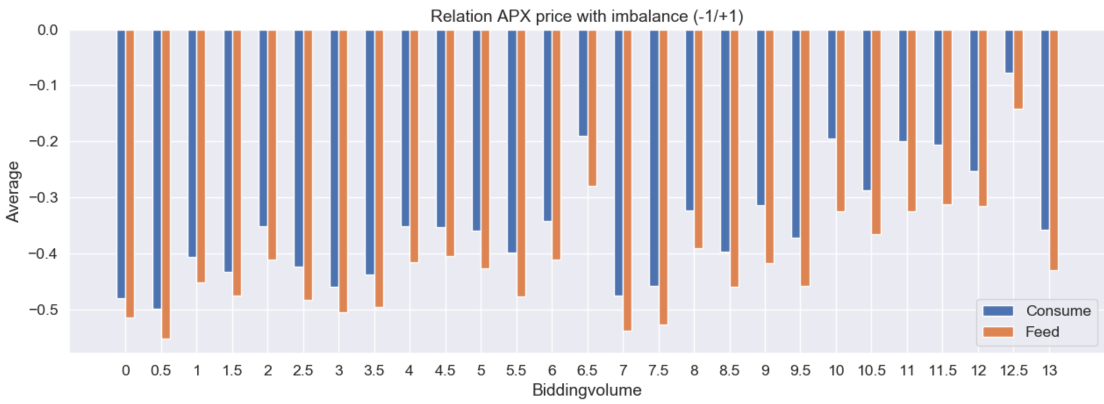


Figure 4.3: Different scenarios in case of higher or lower production versus the bidding volume.

However, the disadvantage of the underlying assumptions made in Figure 4.3 is the fact that it does not take the absolute difference of imbalance and APX price into account, a difference of 2 €/MWh has the same weight as a difference of 300 €/MWh. To account for this, we averaged the difference between APX price and imbalance price for bins of 0.5 MWh. Figure 4.4 shows the results of this analysis. With this result, it is much more difficult to observe trends about the correlation between APX price and imbalance price. For example, the bin of 6.5 MWh indicates that the imbalance prices in this bin are higher compared to the APX price, while the bin 7 MWh indicates the very opposite. Next to that, the standard deviation was measured to be around 50 MWh for all bidding volumes, which shows the high variation in difference between spot and imbalance price.

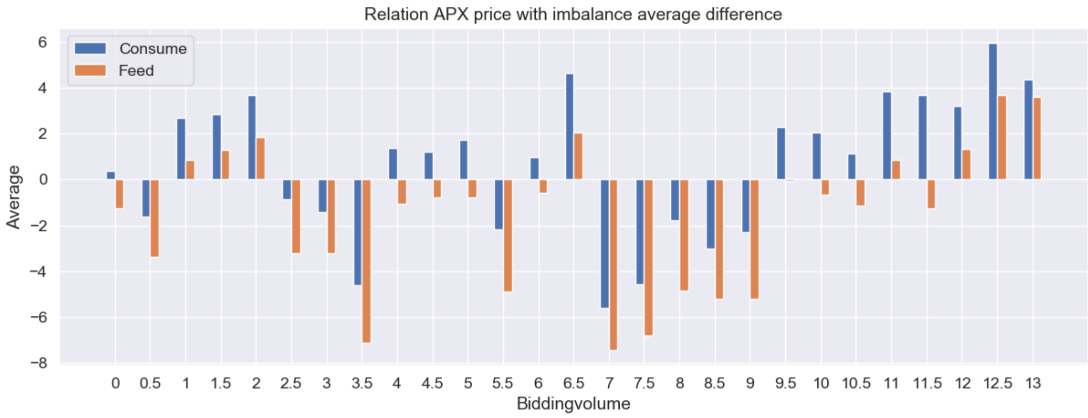


Figure 4.4: Different scenarios in case of higher or lower production versus the bidding volume.

4.1.2 Relation portfolio imbalance and prices

Next to the relationship between imbalance prices and the APX prices, it is also interesting to investigate the relation between the imbalance of the portfolio and the prices. With the imbalance of the portfolio, we mean the over- or underestimation of the energy production of wind parks in the portfolio. This is expected to be in line with the overall imbalance of the market since it can be assumed that all forecasters in the market rely on similar weather data to make their forecasts. A correlation between the two variables can influence the bidding strategy, as random resampling of price and production data may no longer be applicable. When comparing the regulation state issued by the grid operator (see for this Section 2.1) and the different imbalance prices, we notice a weak correlation of 0.38 for regulation state and consume and 0.32 for regulation state and feed-in. However, when we check the correlation of the portfolio imbalance and the imbalance prices, these coefficients are lower: 0.11 and 0.11, respectively. This low correlation is unexpected since based on experience, there is a greater relationship between the two variables. But can be explained with the number of data points: As there are so many, no sharp linear relationship could be found However, the experience of higher relationship may be based on extreme values: In case the forecast is deviating from the production volume with more than 3 MWh positively or negatively, the correlation is found to be 0.408 and 0.405 for the consume price and feed price, respectively. These situations are remembered more easily than situations with low deviations. However, these deviations of more than 3MWh happened only in 569 PTUs of the 35,040 PTUs in total.
As these occur so infrequently, it is a justifiable assumption, to treat the portfolios imbalance and the prices independently from each other. An improvement of the model for later on may be the addition of dependence.
Figure 4.5 shows the surface plot of production, forecast and spot price and it gets apparent that it is difficult to detect correlations between spot price and the other two parameters. However, it can be well noticed, due to the spikes, that at the low end of the forecast/production, the most data points are found here. Next to that, is the surface outside of the diagonal rough, which is an indication of only few data points. In Appendix B, Figures B.1 and B.2 the same surface plots are shown with respect to the two imbalance prices. Here, it is striking in the slightly turned plot of the consume imbalance price, that in situations of underprediction the imbalance prices appear to

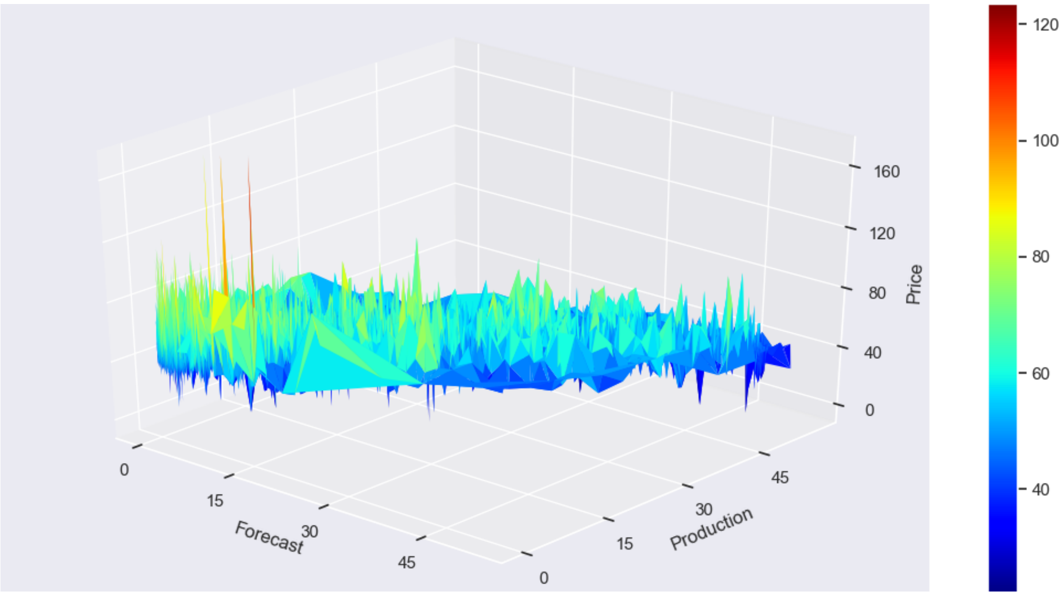


Figure 4.5: Surface plot of forecast, production and price.

be lower, while in the opposite case imbalance prices tend to be higher. This is in line with the expectation of general demand and offer. It substantiates the findings of correlation in case of high differences between forecast and prediction since these situations are well notable on the graph.

4.2 Bidding strategy

4.2.1 Point prediction

The first and most simple strategy is to bid the level of the point prediction. The goal of this method is to forecast the production volume as well as possible, such that imbalance costs can get reduced to a minimum. As a baseline, we used the perfect forecast: The perfect forecast means the theoretic bidding volume which is the actual production volume found in hindsight. This could then be compared with the realized revenues of the other bidding strategies.
As DVEP currently has two forecasters, it was a simple task to compare the results of these two forecasters. Finding the revenues in these two cases was not difficult: We had to apply function 4.1 with the forecasted volumes of Forecaster 1 and 2, respectively. This was done for both July and August 2019, as those were the most recent at the moment of simulation.
A slightly modified approach was the use of average volumes: Here, we used the average power forecast volume of both forecasters as bidding volume. In this way, we could counteract situations were production lays between both forecasters. When both forecasters are on the same side of the real production, so they have either over- or underestimated the production, this approach will result in worse results for the more accurate forecaster and better results for the less accurate forecaster. It remains to be found out, which of the characteristics of this strategy is more important.
As an additional naive approach, we also added the persistence method. This is in fact a forecasting method, which uses the last measured production value as forecast value and can be used as a comparison of the other strategies.

4.2.2 Probabilistic forecast

Instead of treating the production volume E^* as a true given, we can approach the realized production also as a probabilistic distribution, which depends on the forecasted volume $\hat{E}_{t+k|t}$. Figure 4.6 shows the distribution of produced electricity given different forecast volumes. The approach that delivered this graph is described in the following.

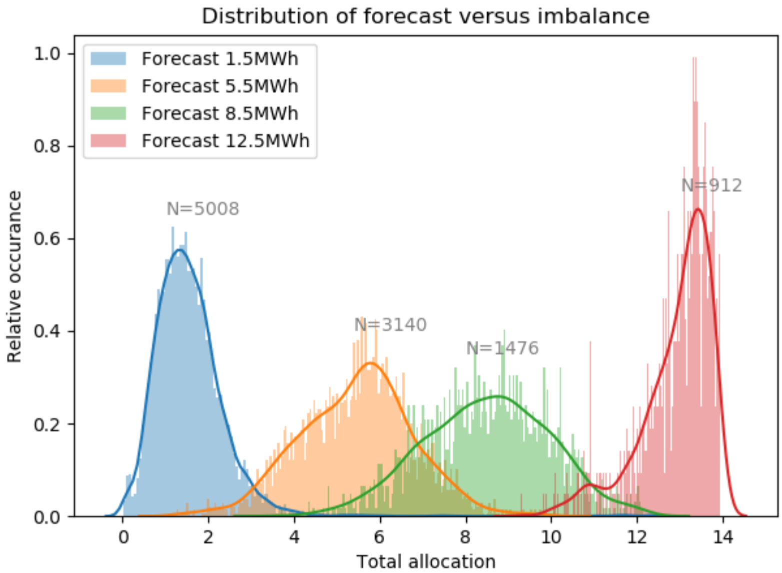


Figure 4.6: Distribution of production given a forecast.

Since it was found in chapter 3 that Forecast 2 is outperforming Forecast 1 in terms of accuracy, we have chosen to use Forecast 2 as reference forecast. We can see in Figure 4.6 that the distributions differ quite significantly with changing forecast. The reason for this can be found in the technical characteristics of a wind turbine: On the left side, it is bounded due to a minimal needed wind speed to create electricity from wind energy. On the right, the production is bounded by the maximum allowed wind speed for wind turbines, as well as the fact that above a threshold of 10-14 m/s, additional wind speed does not translate into higher electricity production. This made it not useful to use a statistical distribution for the possible production volumes but made empirical distributions more applicable.

Next to the distribution of the production depending on the forecast, we also investigated the distribution of the spot price depending on the forecast. This can be seen in Figure 4.7 and less distinction is notable. However, when forecasts are higher the skewness of the distribution of prices is less in comparison to the 1.5MWh case.

Since we wanted to create a probabilistic distribution of production values depending on the forecasted volume, it was important to define this selection. The selection needed to be on the one hand wide enough to capture all relevant production values, but on the other hand, should be narrow enough such that unrealistic values are disregarded. Since we have seen different accuracies of the forecast depending on the wind speed, we considered using an error measure to determine the width of the interval as a reasonable assumption.

The applicable production volumes were found using the flow, which is presented in Figure 4.8:

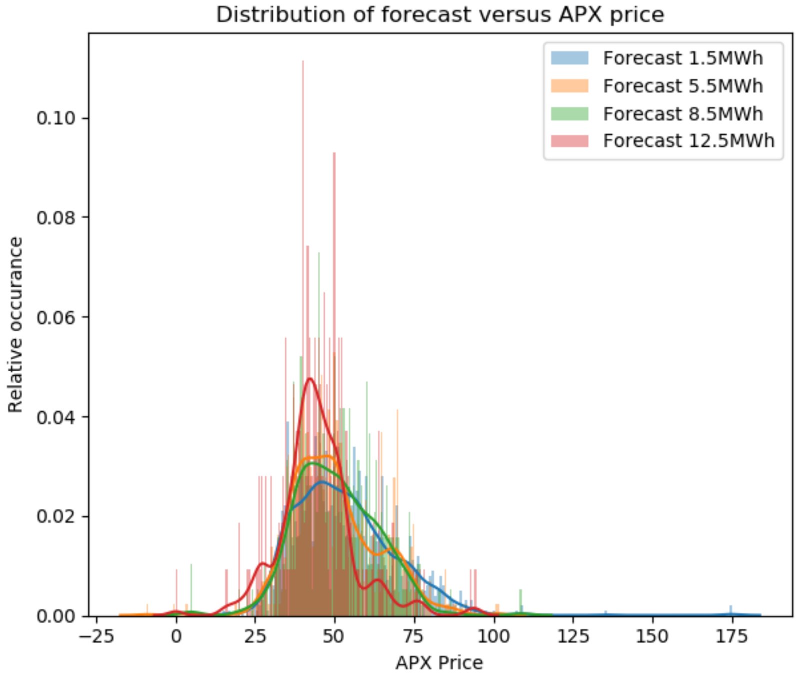


Figure 4.7: Distribution of prices given a forecast.

The starting point formed the forecasted wind power production by Forecast 2, from which we build an interval of ± 0.1 MWh. We considered all hours from 01-07-2018 till 01-07-2019 and from these, all hours with a forecasted volume between the before defined interval were selected. With this selection, we calculated the normalized mean absolute error (NMAE) of the forecast versus the realized production. We used this error measure to define the interval of possible values: The NMAE was multiplied with a factor 10 in order to ensure that the intervals do differ significantly depending on the forecast value. This approach ensures that the population throughout the forecast range remained relatively stable: With higher forecast volume, there are fewer data points, but also the prediction error is higher. In this way, the interval of possible forecast volumes was relatively larger when forecast accuracy is lower. Based on this interval we selected all scenarios applicable for the given forecast, where we assigned the number of possible values as Q.

In Subsection 3.4.2 we have seen that the forecasts tend to overestimate the wind power production with increasing wind speed. This made it useful to correct the forecast of the bias. In the following, we call this strategy "Trend Correction". To find the ideal bidding volume based on trend correction, the last step was the resampling of possible production values for N times. Based on these N samples we could find the mean production value. Now, the ideal bidding volume could be calculated using Equation 4.1. This step was repeated M times. To bring this optimization problem into a form of an equation, we can express it as the following:

$$Ideal\ Volume = \frac{1}{M}\sum_{j=1}^{M}(\max_{E^b}\frac{1}{N}\sum_{i=1}^{N}(E^b\lambda_i^{APX} + I_i^C))_j \tag{4.3}$$

Equation 4.3 contains Equation 4.1 but adds the mean of the N resamples as well as the mean of the M iterations. With this strategy, we sought to decrease the cost of imbalance, however as we disregard the price uncertainties, we could not utilize these in a beneficial way.

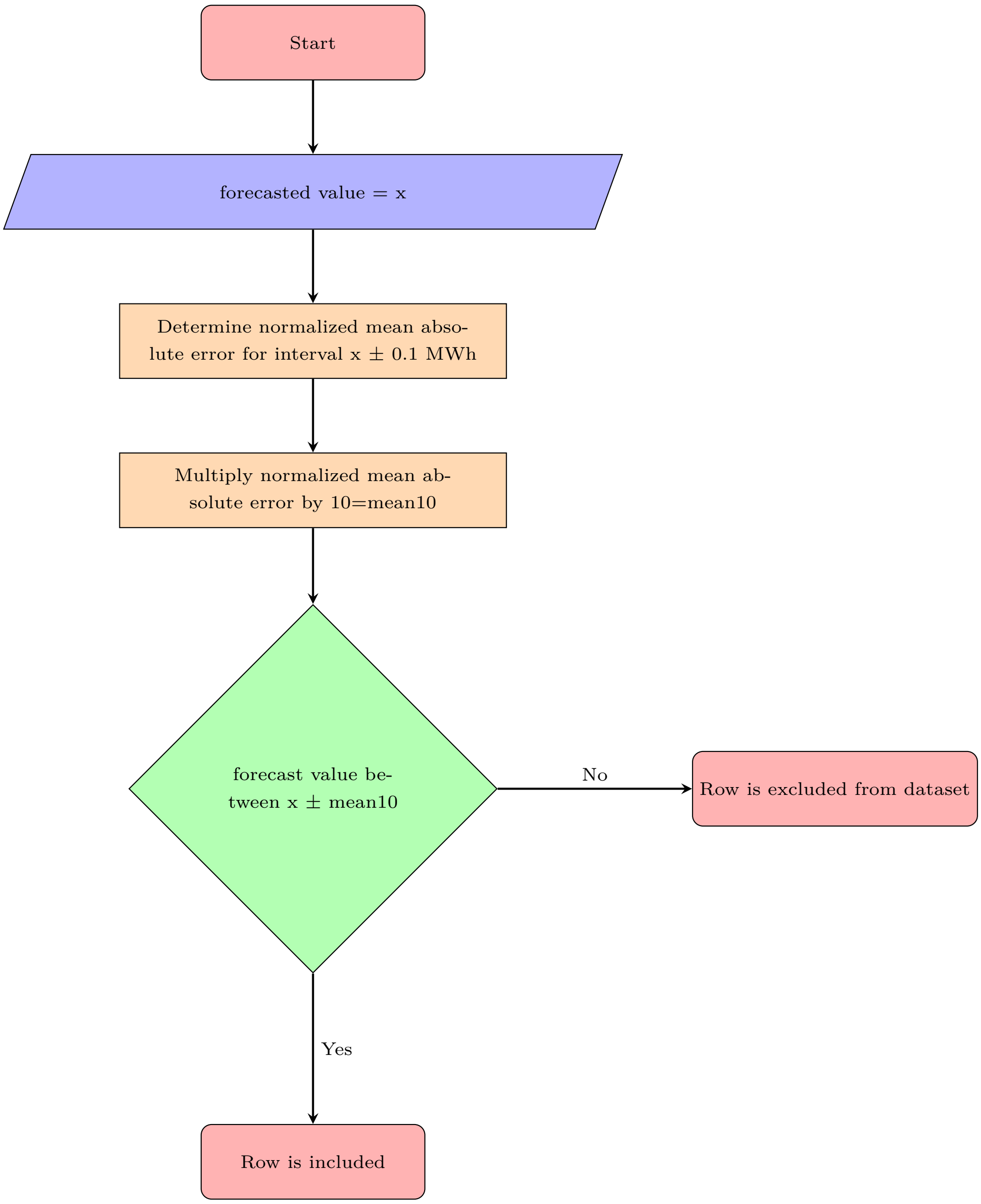


Figure 4.8: The flowchart for the choice of the acceptable rows, depending on forecast x.

At this point we call the number of resamples N since this value was changed for different simulations: The first simulations were done with a fixed N of 10,000. Also, the number of iterations M is interchangeable, therefore we defined it here as a variable. We used the iterations to ensure that the resampled data sets represented the real population as good as possible.
The fourth strategy can be seen as the most complicated one, as it considered all uncertainties: Production, APX price and imbalance price. In order to achieve this, we did to not only resampled the possible production volumes but also resampled the accompanying price values of this production volume. Since we resampled the prices independently from the production volumes, we ended up with Q^2 possible combinations of production and price values. We started with using an N of 10,000 resamples independent from the number of possible combinations.

4.3 Risk constrained bidding strategies

It is clear that the strategy of bidding all or nothing can be highly risky when the uncertainties have been forecasted wrongly. In order to account for this, it is wise to add a risk constraint. We chose to use Value at Risk (VaR) and Expected Shortfall (ES). We used the VaR 95% and ES 95% and omit the 95% in the following, as this parameter is kept constant. The two metrics are added as a boundary at the lower end of hourly revenues. This means that a bidding volume has only been accepted as possible bidding volume when the VaR (or ES) of the simulated hourly revenue was less negative than the VaR, which was defined beforehand as a measure of risk appetite. We used a VaR of -500,-100 and 0 to understand the influence of an increasing VaR on the optimal bidding volume and the consequnces on the revenues. For the ES, we chose to simulate for ES 1500 and 500. These values were chosen to be higher, as the ES considers the complete tail of the distribution. In this way, the bidding volume will be closer to the forecasted volume by the forecaster and wrong forecasts have a less severe impact on the results.

4.4 Additional modifications of solutions

During testing, some shortcomings of the strategies introduced above became evident, which is why we applied some modifications. It got apparent that when using a fixed number of resamples N, forecast volumes with many data points were only poorly represented. This is why we also checked the influence of using a flexible resample size of $N = Q^2 * 0.01$. While in the first simulations we did not consider the differences in prices during the day, we did so in another attempt of simulations. In order to do this, we assigned peak and offpeak to the different hours and only resampled from the respective hours. A third modification was done based on the relation of production and prices. Although the correlation values were found to be low, we add a simulation where production and the accompanying prices are resampled together. Indifferent of the data points chosen and the resampling, the calculation of the ideal bidding volume were the same: The ideal bidding volume for one iteration of the simulation is the volume with which the highest mean revenue can be achieved. This was found by iterating from 0 to maximum installed power. When we recall Figure 4.1, we can state that it is likely that the optimum bidding volume with this strategy will either be no volume or the maximum installed power.
In order to reduce the big imbalance, the last addition was made for constraining the risk, where we added a boundary for possible bidding volumes. In this way, bidding values are closer to the forecasted volume, such that the risk of disadvantageous imbalance gets reduced.

5

Results

In this chapter we present the results of the solution approaches proposed in Chapter 4, with which we seek to answer the research questions of part C. To test the strategies, which we developed in the previous chapter, we use data of July and August 2019. The first part considers the point predictions and slight modifications of these. After which the results of the probabilistic approach are shown as well the other additions.

5.1 Descriptive statistics

Before we show the results of the bidding strategies of July and August it is wise to present the characteristics of these very months on the energy market. In this way, it is easier to compare both months. Table 5.1 shows the important statistical numbers about the two months for us.

Table 5.1: Figures about the months of test.

		July	**August**
APX	mean	39,54 €/MWh	37,44 €/MWh
	SD	8,34	9,42
Imbalance feed	mean	37,85 €/MWh	38,89 €/MWh
	SD	45.45	45.54
Imbalance consume	mean	39.15 €/MWh	40.38 €/MWh
	SD	46.54	46.79
Production	Total	5518 MW	8865 MW

The prices of both month are in the expected area and differ only slightly from each other. In July 1 MWh of electricity was €2 more expensive than in August with a €1.1 lower standard deviation. The imbalance prices were in August a little higher. However, a big difference was the production volume: In August, the wind portfolio at hand produced in total 3347 MW more than in July.

5.2 Point predictions

In Table 5.2 we see the results of the point predictions for months July and August. To interpret the results better, also two error measures are added to the table. The revenues per PTU are shown as the mean revenues per MWh. We consider here the PTU of the imbalance market, which is 15min. Next to that, the percentage of the perfect prediction is shown to assess the monetized performance of the strategy.

Table 5.2: Results in €/MWh for point predictions.

	July				August			
	NBIAS	**NMAE**	**Revenue**		**NBIAS**	**NMAE**	**Revenue**	
Perfect	0%	0%	38.17		0%	0%	34.60	
Persistence	2.47%	12.08%	37.44	98.1%	0.05%	17.77%	33.53	96.9%
Forecast 1	-1.62%	4.73%	35.05	91.8%	-0.59%	4.91%	32.45	93.8%
Forecast 2	-1.26%	4.47%	35.11	92.0%	-0.71%	4.93%	32.37	93.6%
Average	-1.44%	4.45%	35.08	91.9%	-0.65%	4.68%	32.32	93.4%

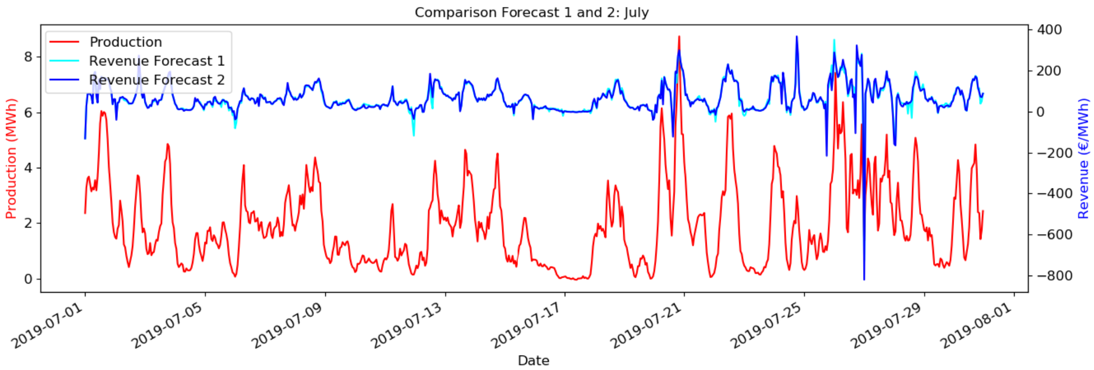


Figure 5.1: The resulting mean revenues for Forecast 1 and 2 in every hour of month July and the mean production volumes.

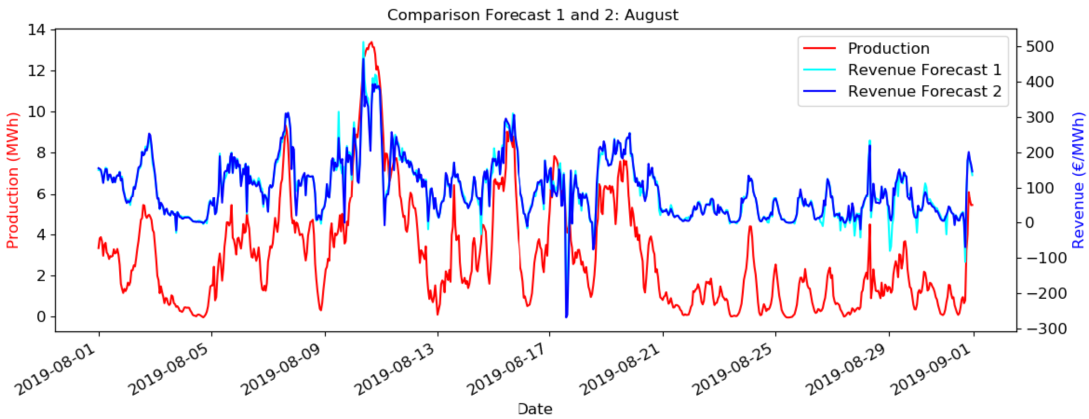


Figure 5.2: The resulting mean revenues for Forecast 1 and 2 in every hour of month August and the mean production volumes.

We only discuss the remarkable results here. First of all, we notice that the revenue per MWh was higher in July compared to August. This can be explained by the lower production of wind energy this month. It is striking that NMAE of both months are very similar, where July has a slightly lower error measure, which means none of both months can be declared to be more difficult to forecast. But still, the revenues of Forecast 2 compared to the perfect forecast are in July lower than August, 92.0% and 93.6%.
We can explain this with the weighted imbalance prices of these months: The cost of 1 MWh in imbalance in July was with 48.06 €/MWh higher than in August with 44.8 €/MWh. The weighted feed-in price, on the other hand, was similar with 30.86 €/MWh and 30.39 €/MWh, respectively, such that in July it was relatively more disadvantageous to sell at the imbalance market because the difference APX and feed was higher than in August. In consequence, this means, imbalance was more expensive in July.
For a better understanding, Figures 5.1 and 5.2 are added show the results of Forecast 1 and 2 for July and August, respectively. In these figures, we handle mean production and revenues per hour to reduce the noise in the plots. Here is interesting to see that peaks in production do not necessarily translate into a significantly higher revenue for this PTU. Consider for this, for example, 21-07-2019, where the production has a high peak, but this can not be seen in the revenues of these PTUs.
One approach to improve the revenue by decreasing cost of imbalance, was the average approach. This resulted indeed in a slightly lower NMAE. The NMAE is 0.02% and 0.25% lower in the two months when compared to Forecast 2. However, this higher accuracy did not improve the monetized performance of this strategy compared to Forecast 2.
This might make it useful to also consider the NBIAS: In both months, the production is slightly overestimated, with the average of course between the two values. However, both error measures cannot explain the differences in revenue. This can be since all bidding volumes are very similar. Based on just these two months and the data available in Table 5.2, we could not conclude which of the three strategies should be used. The difference between the highest and lowest revenue results in an overall revenue difference for July is €331.1 and €265.9 for August.
A remarkable result is that of the strategy of persistence, which only had been added for comparison reasons and was expected to only have poor results. Although this is a very simple forecasting method, it leads clearly to the best monetized results for these two months close to the perfect forecast. On the other hand, the error measures were significantly worse for this strategy. This substantiates the statement that naive forecasters are still difficult to beat for more advanced techniques.

5.3 Probabilistic forecasts

While the point prediction strategies are highly dependent on the forecasts, we tried to use historical data to find more ideal bidding volumes by utilizing imbalance cost. Table 5.3 shows the results of the seven different approaches we considered. The trend correction has the goal, to reduce the imbalance by resampling historical production values. The remaining six approaches also consider the prices: No risk measure has no further restrictions, while VaR -500,-100 and 0 has the restriction, that the 5% worst quarter of the resampling has no higher loss than these three

Table 5.3: Results in €/MWh of the probabilistic strategies.

	July				**August**			
	NBIAS	**NMAE**	**Revenue**		**NBIAS**	**NMAE**	**Revenue**	
Perfect	0%	0%	38.17		0%	0%	34.60	
Trend correction	-1.93%	4.92%	35.12	92.0%	-1.07%	5.32%	32.28	93.3%
No risk measure	-25.03%	34.25%	31.265	81.9%	-27.89%	39.77%	28.617	82.7%
VaR -500	-15.15%	24.14%	30.8	80.7%	-19.59%	23.52%	30.68	88.7%
VaR -100	-3.58%	12.71%	32.53	85.2%	-6.00%	22.51%	31.84	91.0%
VaR 0	0.41%	9.19%	33.21	87.0%	-1.01%	22.14%	32.22	93.1%
ES -1500	-22.22%	31.35%	29.67	77.7%	-26.95%	24.06%	29.85	86.3%
ES -500	-6.25%	15.35%	32.5	85.1%	-8.59%	22.70%	31.13	90.0%

values, respectively. The ES -1500 and -500 uses a similar approach, but with the use of expected shortfall, which is more sensitive to the distribution of the tail.

We can see that the error values are much higher compared to the point prediction approach for all strategies apart from trend correction and VaR 0. These two strategies have a similar NBIAS as the point predictions, however, the NMAE is slightly higher for the trend correction and much higher for VaR 0.

However, it is also important to consider that for these strategies, apart from the trend correction, it was not the goal to predict the production volume as precise as possible, but to exploit the imbalance prices in the best possible way. This resulted in the most extreme results for the strategies No risk measure and ES -1500. In these cases, the optimal bidding volumes were found to be either close to 0 or the maximum installed power. Only in cases where the APX prices lay between the imbalance price of feed and consume, the ideal volume was closer to the actual production volume. Recall for this Section 4.1.

This resulted in high imbalance, which could result in big revenue, but also resulted often in a big loss. See for this Figure 5.3, the quarterly results of the No risk measure strategy, where many negative peaks can be found in the last half of the month and only few positive peaks. When investigating the imbalance cost of these quarters, it gets apparent that in these quarters the imbalance cost, as well as the bidding volume, were very high, while the realized volumes are much lower than forecasted. This created an underproduction and the outstanding volume needed to be bought back on the imbalance market for prices above 100 €/MWh. The positive spikes, on the other hand, were caused by negative prices, which made it profitable to buy back the outstanding volume on the imbalance market.

With a higher VaR and ES, the errors reduced in both months and the highest revenues were found to be with the strategy of VaR 0. This can be explained by the fact that this restriction reduced the very negative situations. The overall worst results can be explained by the following:

We only use the volumes forecast by Forecaster 2 as a selection criterion of possible production volumes and prices. This means that in case the forecast was too high, the selection of possible values created a distribution with even higher production volumes. For price data which are not extreme, this can be a good approach, however, when extreme imbalance data occur this approach is prone to fail. This is noticeable in Figure 5.3. Next to that, only small differences between APX and imbalance price will cause the decision of an all or nothing bidding. Then, the imbalance is very high and the risk of forecasting the prices incorrect is very high as well. This can get severe

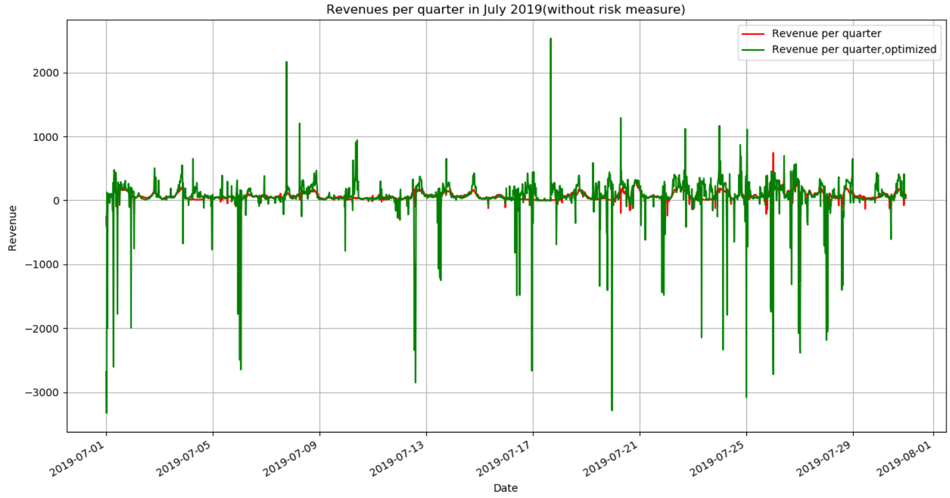


Figure 5.3: The resulting revenues for every quarter in month July, for strategy No risk measure.

Table 5.4: Results in €/MWh for modified probabilistic strategies.

	July			
	NBIAS	**NMAE**	**Revenue**	
Perfect	0%	0%	38.17	
VaR -500 peak/ offpeak	-15.32%	23.41%	28.76	75.3%
VaR -500 peak/offpeak, restricted	-9.85%	17.9%	30.39	79.6%
VaR -100, more repetitions	-3.51%	13.02%	32.42	84.9%
VaR -100, peak/ offpeak, more repetitions	-3.44%	13.09%	31.94	83.7%

in situations were the imbalance prices shoot very high. We can recall for that purpose Figures 4.3 and 4.4: When considering the scenarios that imbalance prices are lower than APX prices, this might indeed be a good solution. However, we notice with 4.4 that when we consider the mean difference between APX and imbalance prices depending on the forecast volume, this statement is weaker.
Still, due to resampling at least 10,000 times this should have been counteracted. This means, that another reason can still be the fact that dependence of price and production is more existent than anticipated and measured earlier.

This means there should be room for improvement of the results: To improve the results of the models shown in Table 5.3, we added for the VaR -500, the peak/ offpeak criterion and the restriction that the bidding volume may not be more than 300% of the forecast volume. Next to that, we altered the number of resamples, N, such that it depends on the number of possible values, and add the peak/ offpeak criterion. This was also done for the VaR -100 case. The results of August were omitted here, as the outcomes of July were not as good as expected and no improvements compared to the initial probabilistic strategies of VaR -500 and VaR -100 were realized. This is why we did not expect that August would show results that outperform the outcome of the point forecasts. When we consider the VaR -500 case, it is clear that the addition

Table 5.5: Results for price/production combined.

	July				August			
	NBIAS	**NMAE**	**Revenue**		**NBIAS**	**NMAE**	**Revenue**	
Perfect	0%	0%	38.17		0%	0%	34.65	
VaR -100	-2.06%	11.48%	32.93	86.3%	-3.85%	15.87%	31.75	91.8%
VaR 0	1.88%	8.18%	33.55	87.9%	1.11%	11.99%	32.36	93.6%

of peak and offpeak hours did not improve the result compared to using all data indifferent of peak hours. It is difficult to give a reasoning for this, but it is likely that the reduction of possible values had a opposite effect. The restriction of possible bidding volumes by 300% of the forecast volume and the addition of the peak/offpeak delivered a result comparable to the VaR -500 without any restriction.
The usage of more N did result in slightly lower results than the comparable simulation with a fixed N of 10,000 repetitions. It is important to note that the number of iterations M were reduced in this case from 10 to 2, because of computation time. While 10 x 10,000 resamples were simulated in 60 mins for the whole month July, it took 16h to simulate the bidding volumes for the month July when the number of resamples was depending on the number of possible values. Since the results were even worse than in the case of fewer resamples and the duration of the simulation was significantly longer, it was decided to not repeat the other situations with an increased N.
As last modification the prices and production were resampled together: Table 5.5 shows the results of the simulations where price and production are resampled together. It gets apparent that these results are better than the simulation results of resampling price and production separate from each other. The reasons for the improved results can be twofold: On the one hand, it can be because there is more correlation between prices and production than expected earlier. On the other hand, it can be caused by the better representation of distribution since the number of possible scenarios is significantly lower compared to the strategies above.

5.4 Conclusion

This chapter had the goal to answer Research Question C.1, whether it is possible to find a model that could improve the bidding volumes for the day ahead market and what risks are connected to this model. It turned out that for these two months a very simple approach, the persistence strategy led to the best results when only the day-ahead market is considered and imbalance is settled at the imbalance market. However, as this approach does not consider the wind power forecasts at all, we can not advice to rely on this method. Apart from this approach, the approaches that had as goal the reduction of imbalance scored best: These were the strategies Forecast 1, Forecast 2, Average and Trend correction. The strategies, which resampled both possible production and prices turned out to lead to worse results. Additional modifications of the strategies with peak/ offpeak hours, restricting the bidding volumes from excessive overforecasting as well as an increased number of resampling did not increase the revenues as expected.
The big risk of this model is that we chose to increase the revenue instead of reducing the losses. Due to the lack of forecastability of imbalance prices, this leads to potential high losses. This leads to a situation that only three possible bidding volumes are possible, where the choice is only

governed by a random selection of production and price values. Although a risk measure like the VaR reduces the extreme bidding choices, the results are still worse than the currently used point prediction approach.

Based on this, we have to say that the proposed strategies did not improve the revenue of the energy trader, but big restrictions on the bidding volume only caused the results to be close to the forecasted volume. Next to that, the strategies should be evaluated over a longer time and the results should be validated with more data. The current test set lies in the summer month and the naive predictor persistence gave the best results, which is unlikely to be true for the whole year.

6

Conclusion

The initial question for this research was if one of the two currently available forecasts at DVEP is outperforming the other. Is this the case in general, or only in certain conditions, and how can this be measured? Based on this, we went further in the optimization of the bidding volumes of the day-ahead market for a wind-based electricity portfolio and formulated the following Research Question:

> *Which model, based on historical market and weather data, can provide the most accurate and profitable day-ahead electricity bidding when considering a wind-based electricity portfolio?*

To find the most accurate forecaster, we used the metrics normalized BIAS (NBIAS), normalized mean absolute error (NMAE) and normalized root mean squared error (RMSE). When the one year data for Forecast 1 and 2 was tested with these metrics, we found that both forecasts are statistically the same. This was for the case that the portfolio was forecasted as a whole. When the wind parks were forecasted individually, we found that the forecasters provide significantly different forecasts. The NBIAS of both were close to zero, indicating, on average, a minimal underprediction, the NMAE was found to be 7.47% and 7.13%, for Forecast 1 and 2, respectively. The NRMSE were in line with this findings, without indications of big additional outliers. With these findings we can conclude that based on accuracy, Forecast 2 is the better forecast. This result is substantiated with the consideration of situations of big differences between both forecasters. When the 5% biggest differences are analysed, again Forecast 2 was most accurate.
After that, external conditions were investigated and how these influence the earlier conclusions regarding the best forecast. We examined the influence of wind speed, day hour, wind direction and temperature on the choice of the most accurate forecast. Again, similar results were found, assigning Forecast 2 to be the better forecast: We found wind speed and day hour to be influencing the performance of the forecasts, where again Forecast 2 was better compared to Forecast 1. The wind direction has some influence on the performance of the forecasts and shows a little accuracy advantage of Forecast 2 for wind directions from West and East. The temperature has only minor influence on the performance and only little distinction between the two forecasts could be found.
While these results focused mainly on the accuracy of the forecasts, the revenue was considered in the second part of the research. Different point forecasts and probabilistic forecasts were tested to find the ideal bidding strategy leading to the highest revenue per PTU in the July and August. We only considered the day-ahead market and consequently the imbalance market. We started with the comparison of the results based on Forecasts 1 and 2. The difference was only minimal, 0.26% and 0.02% when considering the NMAE. In monetary terms, the difference was €331.09 and

€265.95 for July and August, respectively. With regard to the other strategies, that resampled both production and price data, the bidding strategy “VaR 0 with price/production combined” lead to the highest revenue of those, which was 87.9% and 93.6% for July and August, of the revenue of the theoretical perfect forecast. This means, although it was proposed otherwise earlier in the research, keeping price and production dependent was the best approach. Next to that, we found the main factor of the unsatisfactory results to be the risk of extreme prices, which are very difficult to predict. However, we have to stress that these results were still below the results of both Forecasts 1 and 2. For July, Forecast 2 scored best with 92%, while in August Forecast 1 was best with 93.8%. The strategy of persistence was kept outside, because this forecast is very simple and not justifiable to use due to its simplicity. This leads to the conclusion that currently a bidding strategy based on a probabilistic forecast does not improve the revenue when considering only day-ahead bidding. Increasing the restriction of acceptable bidding volumes by different approaches improved the results, however not in such a way that the point forecast approach could be outperformed.

6.1 Limitations of the research

It is important to consider that this research also has its limitations. When we consider the first part of the research, a limitation can be found in the weather stations. They are only in the proximity of the wind parks, but not at the exact locations. Also, the wind speed measurements are done at a height of 2 m, while the wind turbines are at a height of 40 to 135 m. It is thus likely that measurements are prone to be too low.

More limitations can be found in the second part of the research. During the project, it became obvious that forecasting the different prices is a crucial part for the probabilistic forecasting. There, a decision has to be made, whether one is satisfied with a simplification of the prices by using average values, for example, or whether reality should be followed closely. The last part includes advanced forecasting of the prices. However, we chose to consider all different uncertainties without using an advanced forecasting technique. We assume that this can be an explanation of the disappointing results of the probabilistic forecasts.

It should also be mentioned that the two months for testing were difficult ones. Due to the seasonal characteristic of those months, it might be difficult to extrapolate them to the whole year.

Another limitation can be found in the algorithms. We chose to build the simulations ourselves in Python and likely it could have been more efficient. In this way, we restricted the number of resamples. As last point, we need to mention the amount of data. As the data was limited to one year, it was difficult to choose more distinctive filters in the algorithm. For example, we only made the attempt to differentiate between peak and offpeak, but also additional filters like months could be considered.

6.2 Recommendations

Finally we can make recommendations to DVEP about the further work with wind power forecasts and day-ahead market biddings. We advise DVEP to use Forecast 2 for the day-ahead electricity bidding. Although both forecasts’ accuracy is very similar, in moments of high wind speed,

Forecast 2 outperforms Forecast 1. Next to that, Forecast 2 is on average better than Forecast 1 in situations where the difference between both forecasts is very big. These findings give us the confidence to advise Forecast 2 as leading forecast and drop Forecast 1 as day-ahead forecaster.
In order to assess the performance of the forecast, no matter which of the two will be used, it is very important to add a continuous measurement of all forecasts in the portfolio. Currently, this is available in monetary terms, but to keep the insights it is important to measure the forecast accuracy by NMAE and NBIAS and NRMSE, based on the reasoning done in this report. Hereby it can be wise to have such a performance monitor of every month for the whole portfolio. Next to that, it can be a good addition to monitor also the individual wind parks and highlight those that have a big difference to the overall measure. With these individual wind park metrics DVEP can find more easily whether a wind park is generally difficult to forecast and adapt contracts with the respective client accordingly due to higher risk of imbalance.
In case we want to continue the research about optimizing bidding volumes, it can be wise to alter the validation set: For this project we used the most recent months, which however might have seasonal characteristics. These make it difficult to extrapolate over the whole year. In order to overcome this shortcoming it can be a choice to select a validation set throughout the year, a so-called aselect test validation set. These data points have to be deleted from the training set.
When we want to continue with the strategic approach to determine the bidding volume, independently from the finally chosen strategy, we advise to create a direct link between the database and the model. For this research, the data were withdrawn manually from the database, which is inflexible, slow and prone to mistakes. Currently, DVEP handles data processing mainly in MS Excel. When the step should be made to a more quantitative modelling approach, it can be advised to switch to Python, C or R.

6.3 Further research

Although this research discussed the topic in detail and was able to answer the proposed Research Questions, there is still more research needed to fully understand the topic. As mentioned before, the imbalance prices and their forecasting are still not fully understood, which is a topic for a master's thesis in itself. Incorporating more advanced methods to forecast prices can be an option to improve the results. This includes also a deeper understanding of the factors that influence the prices. We discussed this only briefly in this research.
Another option is the usage of more filters to find the possible values for resampling: Think about using only data of the same week day or same month.
Next to that, it can be worthwhile to consider this probabilistic forecast for the intraday market. It has been left out in this research, as no usable historical price data was available. If the data were to be used, a lot of processing needs to be done.

References

Bal, R. (2013). *Development of the imbalance of the Dutch electricity grid* (Unpublished master's thesis). Utrecht University, The Netherlands.

Brownlee, J. (2018). *15 statistical hypothesis tests in python (cheat sheet).* Retrieved 22-10-2019, from `https://machinelearningmastery.com/statistical-hypothesis-tests-in-python-cheat-sheet/`

Bueno, M., Moreno, M. , Usaola, J., & Nogales, F. J. (2010). Strategic wind energy bidding in adjustment markets. In *Proceedings of the universities power engineering conference* (pp. 1–6).

CBS. (2019). *Share of renewable energy up to 7.4 percent.* Retrieved 24-06-2019, from `https://www.cbs.nl/en-gb/news/2019/22/share-of-renewable-energy-up-to-7-4-percent`

Chaves-Ávila, J. P., Hakvoort, R. A., & Ramos, A. (2014, 5). The impact of European balancing rules on wind power economics and on short-term bidding strategies. *Energy Policy*, *68*, 383–393.

Draxl, C. (2012). *On the predictability of hub height winds* (PhD thesis). Technical University of Denmark.

EPEX SPOT. (2018). *Cross-border intraday: Questions answers.* Retrieved 09-09-2019, from `https://www.epexspot.com/document/40068/XBID%20Q%26A`

Eransus, F. J. (2016). A bidding strategy for minimizing the imbalances costs for renewable generators in Spanish power markets. Retrieved from `http://www.ucm.es/fundamentos-analisis-economico2/documentos-de-trabajo-del-icaeWorking`

Foley, A. M., Leahy, P. G., & McKeogh, E. J. (2012). Current methods and advances in forecasting of wind power generation. *Renewable Energy*, *37*, 1–8.

Hesselink, M. (2018). *Developing a forecast model for the power production of wind turbines* (Unpublished master's thesis). University of Twente, The Netherlands.

Kariniotakis, G., Martí, I., Casas, D., Pinson, P., Nielsen, T. S., Madsen, H., & Giebel. (2004). What performance can be expected by short-term wind power prediction models depending on site characteristics? In *Proceedings of the European Wind Energy Conference* (pp. 1–9).

Madsen, H., Pinson, P., Kariniotakis, G., Nielsen, H. A., & Nielsen, T. S. (2006, 3). Standardizing the Performance Evaluation of Short-Term Wind Power Prediction Models. *Wind Engineering*, *29*(6), 475–489.

Moreno, M. , Bueno, M., & Usaola, J. (2012, 12). Evaluating risk-constrained bidding strategies in adjustment spot markets for wind power producers. *International Journal of Electrical Power and Energy Systems*, *43*(1), 703–711. doi: 10.1016/j.ijepes.2012.05.059

Mulder, M., & Scholtens, B. (2013). The impact of renewable energy on electricity prices in the Netherlands. *Renewable Energy*.

Pinson, P. (2006). *Estimation of the uncertainty in wind power forecasting* (Doctoral dissertation, École Nationale Supérieure des Mines de Paris). Retrieved from `https://pastel.archives-ouvertes.fr/pastel-00002187`

Pinson, P. (2018). *Renewable energy modelling and forecasting.* University Lecture. Retrieved from `http://pierrepinson.com/?page_id=913`

Pinson, P., Chevallier, C., & Kariniotakis, G. N. (2007, 8). Trading wind generation from short-term probabilistic forecasts of wind power. *IEEE Transactions on Power Systems*, *22*(3), 1148–1156.

Risk.net. (2018). *Value-at-risk (var).*

Tanrisever, F., Derinkuyu, K., & Jongen, G. (2015). Organization and functioning of liberalized electricity markets: An overview of the Dutch market. *Renewable and Sustainable Energy Reviews*, *51*, 1363–1374.

TenneT. (2019a). *Annual market update 2018.* Retrieved 24-06-2019, from `https://www.tennet.eu/fileadmin/user_upload/Company/Publications/Technical_Publications/Dutch/Annual_Market_Update_2018_-_Final.pdf`

TenneT. (2019b). *Imbalance pricing system.* Retrieved 04-09-2019, from `https://www.tennet.eu/fileadmin/user_upload/SO_NL/ALG_imbalance_pricing_system.doc.pdf`

Usaola, J., & Angarita, J. (2007). Bidding wind energy under uncertainty. *2007 International Conference on Clean Electrical Power, ICCEP '07*, 754–759.

Wang, X., Guo, P., & Huang, X. (2011). A review of wind power forecasting models. In *Energy procedia* (Vol. 12, pp. 770–778). Elsevier Ltd.

Wind Turbine Models. (2019). *Compare power curves of wind turbines.* Retrieved 18-10-2019, from `https://en.wind-turbine-models.com/powercurves`

Zhang, H., Gao, F., Wu, J., Liu, K., & Liu, X. (2012). Optimal bidding strategies for wind power producers in the day-ahead electricity market. *Energies*, *5*(11), 4804–4823.

Zugno, M., Jõnsson, T., & Pinson, P. (2013, 9). Trading wind energy on the basis of probabilistic forecasts both of wind generation and of market quantities. *Wind Energy*, *16*(6), 909–926.

A

Appendix A

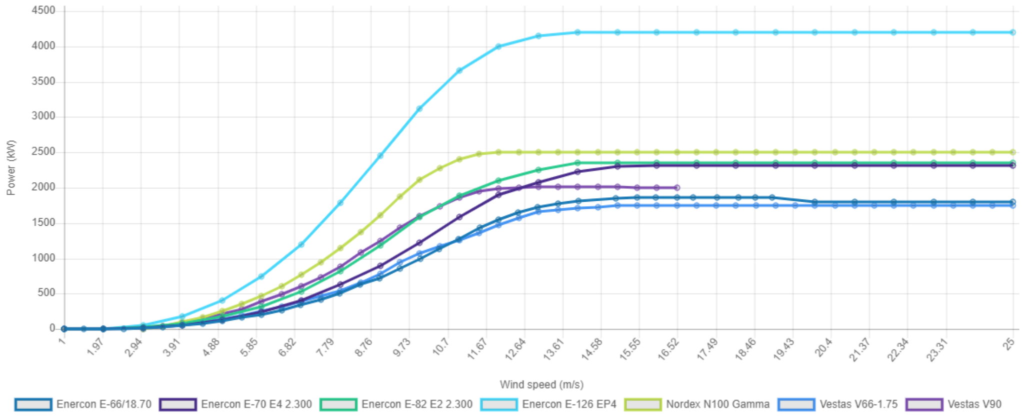


Figure A.1: Powercurve of some windturbines in the portfolio (Wind Turbine Models, 2019).

.

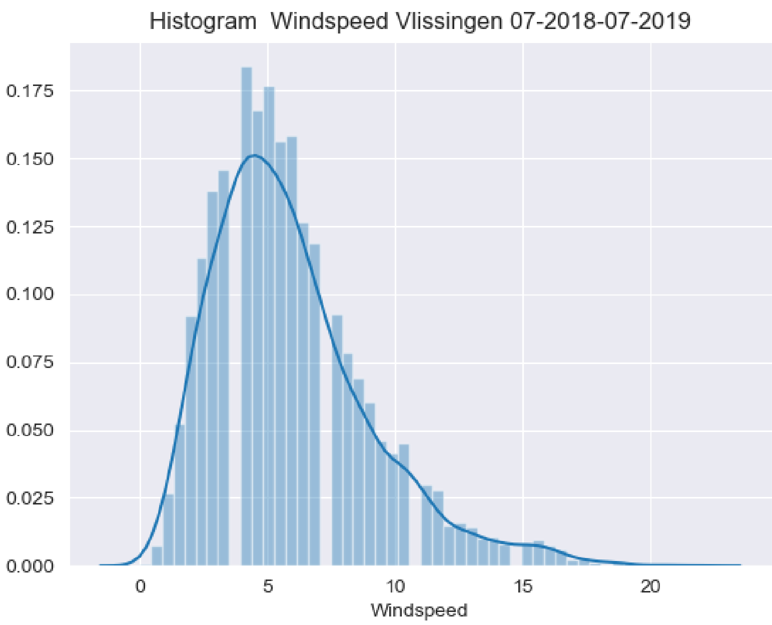


Figure A.2: Histogram of the wind speeds in Vlissingen.

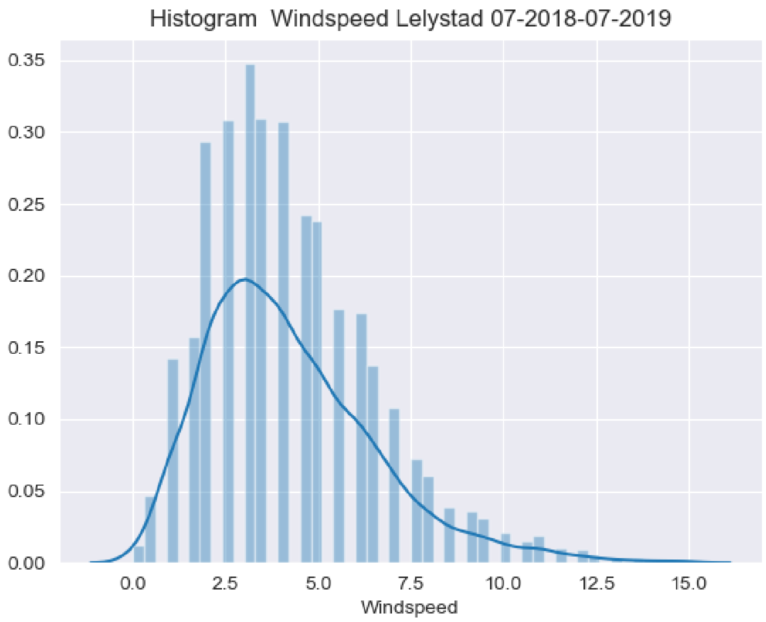


Figure A.3: Histogram of the wind speeds in Lelystad.

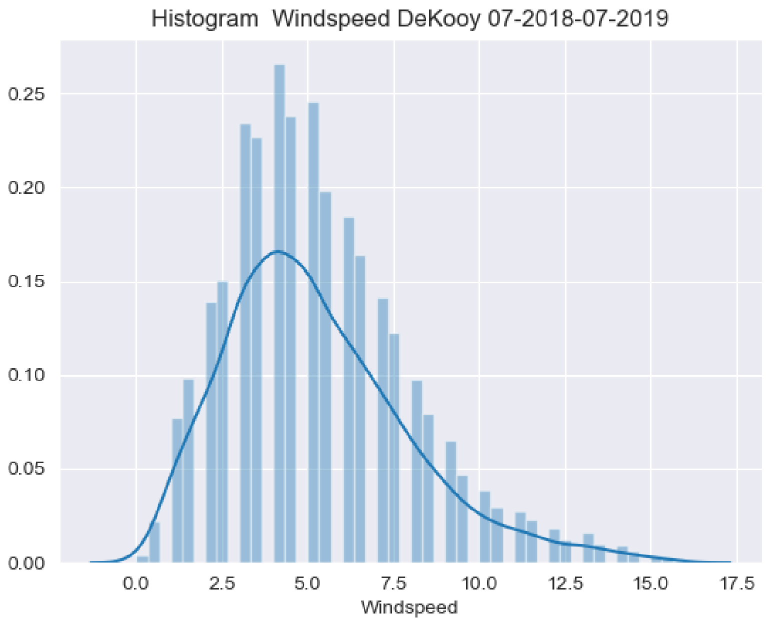


Figure A.4: Histogram of the wind speeds De Kooy.

B

Appendix B

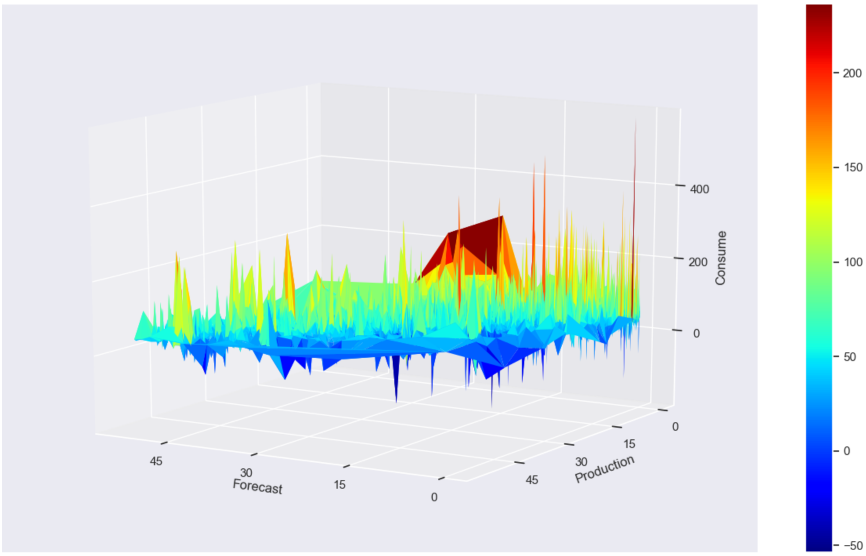


Figure B.1: Surface plot with consume imbalance.

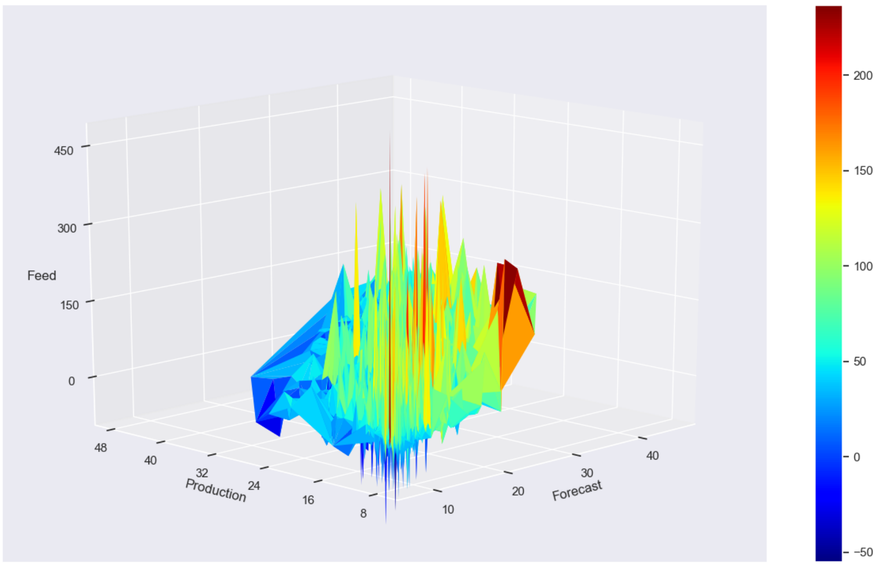


Figure B.2: Surface plot with feed imbalance, turned.

C

Appendix C

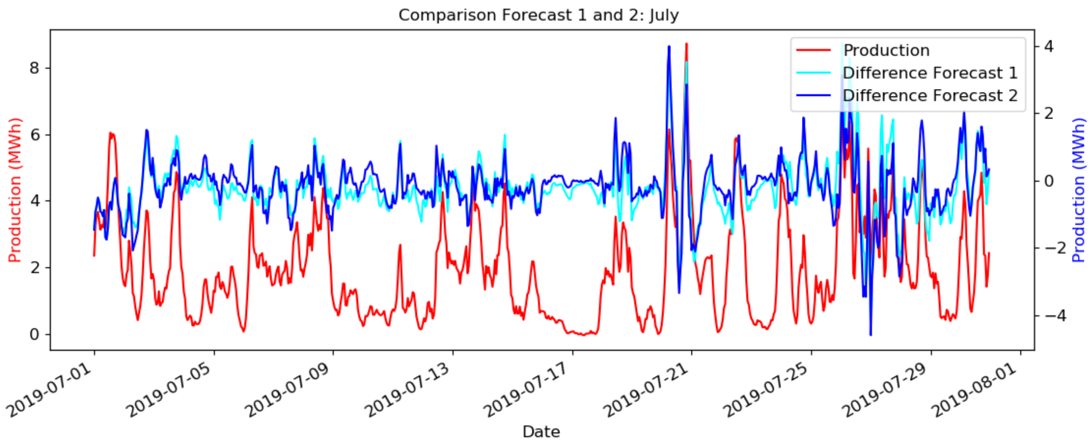


Figure C.1: The resulting mean difference for Forecast 1 and 2 compared to production in every hour of month July and the mean production volumes.

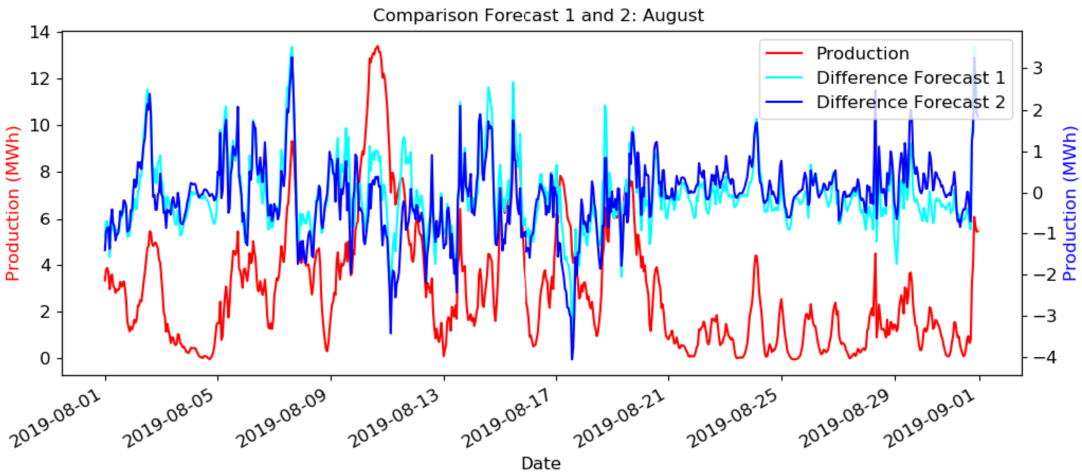


Figure C.2: The resulting mean difference for Forecast 1 and 2 compared to production in every hour of month August and the mean production volumes.