MASTER THESIS

26-04-2018

Developing a Forecasting Model for the Power Production of Wind Turbines

Mike M. J. Hesselink

Industrial Engineering and Management

Production and Logistics Management

UNIVERSITY OF TWENTE.

DVEP ENERGIE

Intentionally left blank

UNIVERSITY OF TWENTE.

Developing a Forecasting Model for the Power Production of Wind Turbines

Publication date
26-04-2018

Student
Mike M. J. Hesselink
Industrial Engineering and Management
Track: Production and Logistic Management

Supervisory committee:

University of Twente
Dr. R. A. M. G. Joosten
Dr. ir. L. L. M. van der Wegen

DVEP Energie
Joost Frank
Maarten Hofhuis
Maarten Vinke

University of Twente
Faculty of Behavioural, Management and Social sciences
Drienerlolaan 5
7522 NB Enschede
www.utwente.nl

DVEP Energie
Jan Tinbergenstraat 110
7550 AW Hengelo
+31 (0)74 376 62 43
info@dvep.nl

UNIVERSITY OF TWENTE.

Intentionally left blank

Management Summary

De Vrije Energie Producent (DVEP) is an energy supplier and Balance Responsible Party (BRP) in the Dutch energy market. A BRP is responsible for buying and selling energy in advance on behalf of the customers in its portfolio. Each day DVEP is responsible for forecasting the energy production and usage of its entire portfolio for each hour of the next day. The forecast at 9:00 in the morning is used, so the forecast horizon is 15-38 hours ahead. Due to the volatile intraday market, an inaccurate day ahead forecast can be very costly. A large part of the portfolio of DVEP consists of wind power producers. Currently, DVEP buys the day ahead wind power forecast from Company X. This forecast is believed to be inaccurate, which is very costly. Therefore, the research goal is to develop a day ahead forecasting model for the power production of wind turbines of DVEP producers that is more accurate than Company X. The main research question is:

"***How to develop a model that is able to translate day ahead weather forecasts into power production forecasts for wind turbines of DVEP producers with higher accuracy than Company X?"***

Literature suggests that weather forecasts are essential for our forecast horizon. According to our literature review, we should use causal models such as regression to describe the relationship between historical weather forecasts and historical production data for each producer separately. Day ahead weather forecasts provide values for average wind speed, average wind direction and average temperature per hour. However, historical day ahead weather forecasts (hindcasts) are only available for the second half of 2017. Therefore, we describe the relationship between historical weather measurements and historical production data using regression models.

For each producer, we have data for the total production of all its wind turbines for each hour. Historical weather data are not available at the wind turbine site, so data from KNMI weather stations are used. For each weather station we have historical data of average wind speed, average wind direction and average temperature for each hour. A selection of 10 producers is included in the study based on location and total rated power. The aggregated rated power of all 10 producers is 45.65 MW. The average distance of producers to the closest KNMI weather station is 8.6 km. We use historical weather measurements and hindcasts from the KNMI stations of Vlissingen, Hupsel and Lelystad. We found statistical evidence that wind speed, wind direction and temperature forecasts are biased in Vlissingen, Hupsel and Lelystad. Therefore, we adjust the day ahead weather forecast for the bias, before we insert it into the regression models. The parameters of the regression models are estimated using historical weather measurements.

Based on literature review and data analysis, we develop 21 regression models using 3 different sets of predictors. We start by using only wind speed as predictor; after this we add temperature and lastly, we add wind direction. Each predictor set has 7 regression models; we illustrate the regression models with all three predictors in Table 1.

Model	Function
Log	$\log P = \log\alpha + \beta\log v + k\log T + \lambda\log\cos(\frac{\theta-\delta^*}{c^*})$
3rd degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3) \times \frac{k}{T} \times \left(\lambda\cos\frac{\theta-\delta}{c}\right)$
4th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4) \times \frac{k}{T} \times \left(\lambda\cos\frac{\theta-\delta}{c}\right)$
5th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4 + a_5 v^5) \times \frac{k}{T} \times \left(\lambda\cos\frac{\theta-\delta}{c}\right)$
Exponential	$P = \left(P_r\left(1+\left(\frac{\beta}{v}\right)^{\alpha}\right)^{-y}\right) \times \frac{k}{T} \times \left(\lambda\cos\frac{\theta-\delta}{c}\right)$
Logistic 4	$P = \left(\alpha\left(\frac{1+me^{-\frac{v}{\tau}}}{1+ne^{-\frac{v}{\tau}}}\right)\right) \times \frac{k}{T} \times \left(\lambda\cos\frac{\theta-\delta}{c}\right)$
Logistic 5	$P = \left(d + \left(\frac{a-d}{\left(1+\left(\frac{v}{f}\right)^{b}\right)^{g}}\right)\right) \times \frac{k}{T} \times \left(\lambda\cos\frac{\theta-\delta}{c}\right)$

Table 1: Regression models using wind speed, temperature, and wind direction as predictors.

We select a top 3 regression models based on the accuracy using historical weather data. The 4th degree Polynomial, 5th degree Polynomial and the Logistic 4 model are the most accurate models. All three models are most accurate using wind speed, temperature and wind direction as predictors. Out of these three models, the 4th degree Polynomial model has the best day ahead forecast accuracy.

Model	Standard Error of Regression (S) in kW	Root Mean Squared Error (RMSE) in kW	Mean Absolute Error (MAE) in kW	Normalized Mean Absolute Percentage Error (NMAPE)
Company X	3,338	3,335	2,407	5.3%
4th degree Polynomial	4,166	4,162	3,092	6.8%

Table 2: Aggregated day ahead forecast error for Company X and our best model for the second half of 2017.

The Root Mean Squared Error (RMSE) of the day ahead forecast of Company X is smaller than the RMSE of the 4th degree Polynomial model for all 10 producers. This indicates that Company X is more accurate for each producer individually. Table 2 illustrates that Company X is more accurate for the aggregated day ahead forecast of all 10 producers as well. All four performance indicators have smaller errors for Company X than for the 4th degree Polynomial model. The aggregated day ahead forecast of Company X is 685 kW more accurate per hour on average; this is equal to 1.5% of the aggregated rated power.

We conclude that Company X has a more accurate day ahead forecast than the newly developed 4th degree Polynomial model for all producers included in the study. The aggregated day ahead forecast of Company X is also more accurate than the aggregated forecast of the 4th degree Polynomial model. Therefore, we recommend DVEP to keep outsourcing the day ahead forecast to Company X for the time being.

Preface

With this report my time as a student at the University of Twente comes to an end.

I want to thank DVEP for giving me the opportunity to conduct this interesting research project. During my graduation period I got to know many new colleagues. I immediately felt at home at the Supply Department of DVEP and want to thank my colleagues for the many laughs during my time there. Especially the skiing trip to Winterberg was really memorable. Special thanks go out to Joost Frank, Maarten Hofhuis and Maarten Vinke at DVEP for their supervision during this research project.

Furthermore, I want to thank my supervisors at the University of Twente for their patience and guidance during this project. I am grateful for the guidance of Reinoud Joosten and the interesting conversations we had throughout the research project. I want to thank Leo van der Wegen for his critical points during the last phase of my research. This was really helpful and contributed to the end result.

I made some amazing friends during my study period, without them this period in my life would not have been so nice. Finally, I would like to thank my family for the unconditional support during my study years. They have always encouraged me and supported me in finishing my master’s thesis in these last months.

I hope you will enjoy reading this report.

Hengelo, April 2018

Mike Hesselink

Table of Contents

List of Figures

List of Tables

1. Introduction

In the first chapter we provide a brief description of De Vrije Energie Producent. Secondly, we describe the research context in Section 1.2. After this, we provide a problem description in Section 1.3. In Section 1.4, we formulate the research objective and research questions. Lastly, we discuss the research scope and report outline in Sections 1.5 and 1.6.

1.1 Introduction to DVEP

De Vrije Energie Producent (DVEP) from Hengelo is one of the fastest growing energy suppliers in the Netherlands. DVEP offers a wide variety of services in the energy industry, among which the supply and resupply of energy. In 2003, the organization was founded as a one-man company, after growing steadily for 14 years the company had approximately 70 employees in August 2017. In September 2017 the company was bought by UGI corporation, which is an LPG distribution company headquartered in USA with extensive operations in Europe (DVEP Energie, 2017). With approximately 13,000 employees, UGI is big international player in the LPG industry. UGI bought DVEP to have a foothold in the Dutch energy market.

DVEP trades on Dutch and German energy markets and wants expand to other countries in Europe. In addition, it also trades on energy related markets like gas. DVEP is a Balance Responsible Party (BRP), which means one of its main responsibilities is managing the usage and production of energy for energy suppliers and customers in its portfolio. DVEP is responsible for buying and selling energy on the markets on behalf of these suppliers and customers. This can involve long term deals, which usually have a fixed price per hour over a timespan of months or years, or short term (intraday) deals over a timespan of an hour or a couple of hours. Long term deals involve buying energy for a period of at least a month, long in advance for a fixed price. This is mostly done for customers with a high energy usage like municipalities or organizations, since they want to limit risk of price fluctuations. A large part of the expected energy usage is bought in advance to reduce the risk of adverse price movements. Short term deals involve eliminating energy imbalance during the day. Besides energy, DVEP is also BRP in terms of gas for some customers. However, energy is its core business.

DVEP has a wide variety of customers in its portfolio, ranging from municipalities to football stadiums. DVEP is an energy supplier and BRP for other energy suppliers. Most customers consume energy; DVEP has to estimate how much energy is consumed on an hourly basis. Besides energy users, the company has many energy producers in its portfolio as well. Most producers focus on sustainable energy like solar power, wind power, biomass, cogeneration and hydropower. Just like DVEP has to estimate the energy consumption of its portfolio, it has to estimate the energy production on an hourly basis as well.

1.2 Research context

For each day, DVEP has to hand in an estimate of the energy usage and production per hour for the following day in the form of an auction. As input for the auction, forecasts are used to estimate the hourly energy usage and production of its entire portfolio. After all BRPs have handed in their auctions, the spot market operator, APX, determines the market price for each hour of the next day based on demand and supply.

The day ahead market prices, called APX prices, are based on expected demand and supply. However, during the next day the actual demand and supply can be extremely different. TenneT, the Dutch grid operator, has to maintain the grid stability of 50 Hz. This means TenneT monitors the grid intensively to maintain the balance. This energy imbalance is regulated using imbalance prices. Each minute of the day, TenneT releases the predicted imbalance prices in real-time, Figure 1.1 shows the imbalance prices in €/MWh up to 10:00 for a random day.

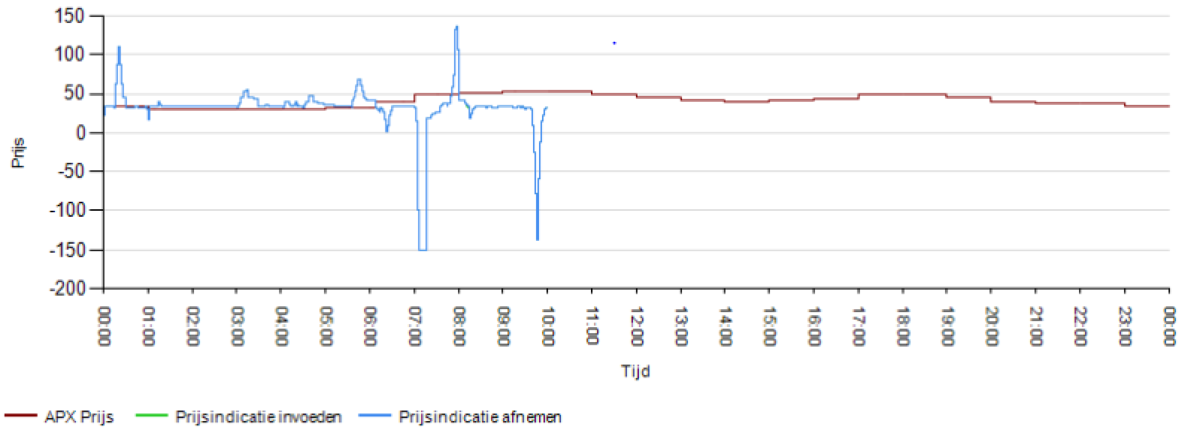

Figure 1.1: Imbalance prices up to 10:00 and APX prices (€/MWh) for a random day.

The red line indicates the APX price, which is determined day ahead. The blue line indicates the predicted imbalance prices, these are determined by demand and supply each minute. Figure 1.1 shows that the energy imbalance market can be very volatile, with predicted prices ranging from €-150 to €140 per megawatt-hour (MWh) within one hour. The actual imbalance prices are determined per 15 minutes after the hour has passed.

The position taken by DVEP each hour is mainly determined by the day ahead auction. Figure 1.2 illustrates a buy scenario for Hour Y and a sell scenario for Hour X at the auction.

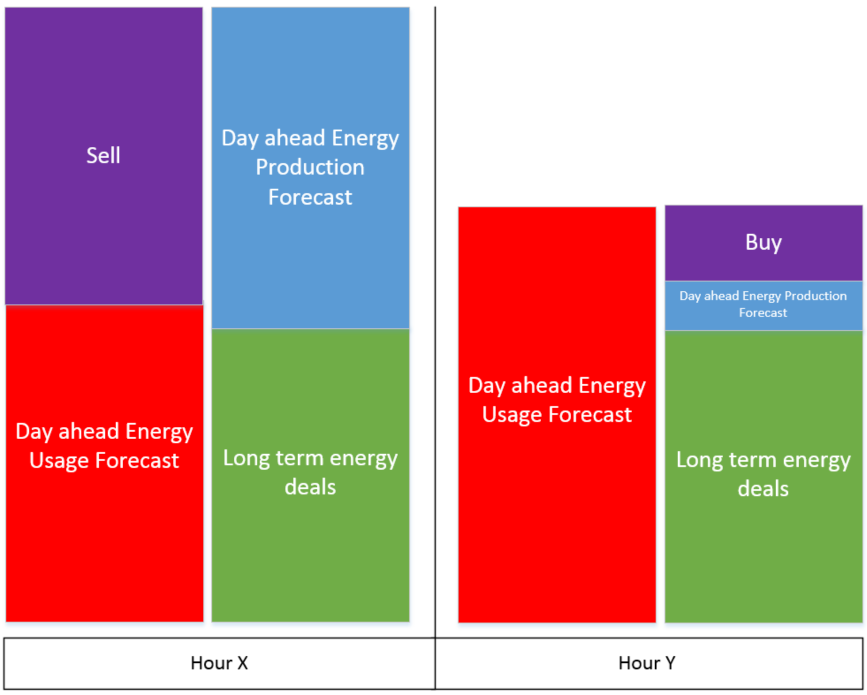

Figure 1.2: Overview of buy and sell scenario on the day ahead auction.

Each auction contains all hours for the next day. During each hour an energy usage forecast (red) and an energy production forecast (blue) are used. The energy usage forecast determines the expected usage of all customers in the portfolio. DVEP has hedged a large amount of the usage with long term deals (green) to limit the risk of price fluctuations. The energy production forecast determines the expected production of all producers in the portfolio of DVEP. Together with the long term deals, this determines the expected amount of energy DVEP has during an hour. Depending on the energy usage and production forecasts per hour, energy is bought or sold on the auction.

The day ahead auction mainly determines the position you take for each hour of the day. During the next day, this position can be altered up until 5 minutes before the hour starts. On the intraday market, traders can buy or sell energy for upcoming hours. Some parties have excess energy based on intraday forecasts, while other parties have shortages. By selling energy to each other, they can alter their position before the hour starts. This can lower risk, because the amount that is bought/sold is not traded using imbalance prices, but the price agreed upon by both parties. As can be seen in Figure 1.1, imbalance prices can be very volatile.

According to DVEP, energy usage forecasts are quite accurate and do not form a problem. Production forecasts however, do form a problem since DVEP believes these are inaccurate. DVEP has many types of producers in its portfolio, among which solar power and wind power. Especially the wind power production forecasts are important for DVEP, since it has a lot of wind power in its portfolio.

1.3 Problem description

Currently DVEP buys a wind production forecast from a third party. We call this party 'Company X' out of confidentiality. The forecast from Company X shows the energy production in kilowatt-hour per hour of the day. While making the auction for the day ahead, the forecast at 9:00 is used, since the input for the auction has to be delivered before 12:00. Ideally, DVEP would like to use the forecast from 11:00. However, a time buffer for technological issues is necessary, so the 9:00 forecast is the most recent forecast that can be used. Normally, DVEP uses the exact production forecast from 9:00 of Company X in the auction. This can be seen in Figure 1.3.

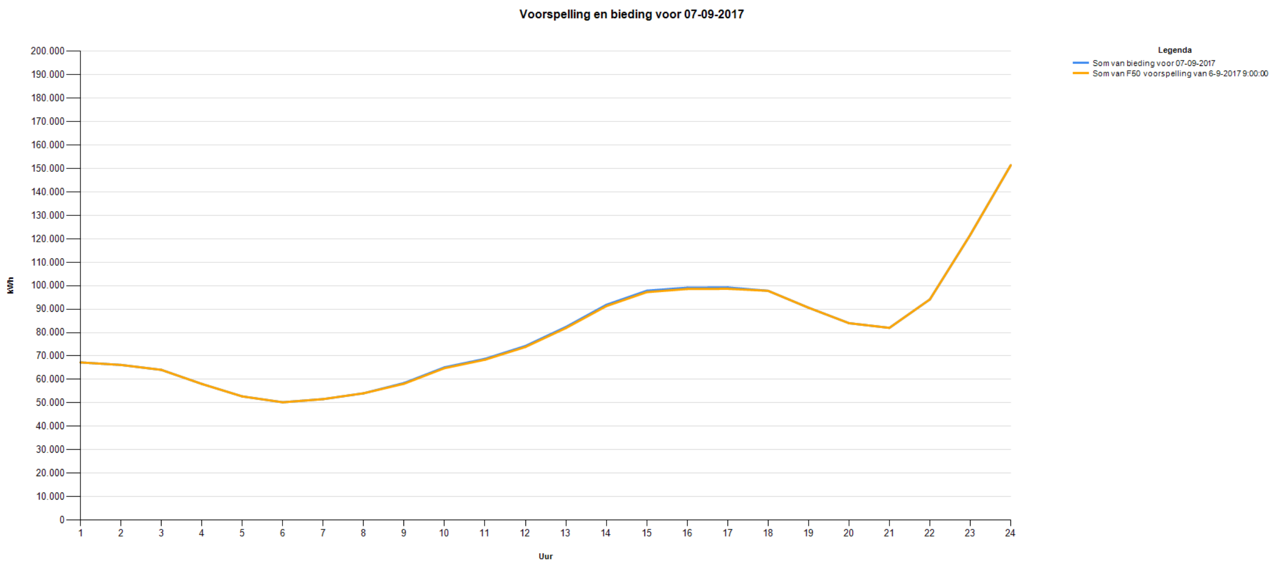

Figure 1.3: Day ahead aggregated energy production forecast for wind turbines at 9:00.

The aggregated day ahead forecast at 9:00 is illustrated by the yellow line in Figure 1.3. The aggregated wind production is illustrated by the blue line, this is used in the auction that DVEP has placed before 12:00. Usually this coincides with the day ahead forecast, since DVEP uses the forecast. Therefore, the blue line is hard to see in Figure 1.3.

DVEP is not satisfied with the day ahead wind production forecast of Company X. At this moment, the company has no idea how accurate the day ahead forecast of Company X is in comparison to competitors. However, DVEP thinks it can be improved and wants to develop an in-house forecast model to improve forecast accuracy and lower costs. The desired output of the model is a graph that shows the expected energy production (kilowatt-hour) for each hour of the next day (see Figure 1.3).

In addition to the day ahead forecast, DVEP also uses intraday forecasts so their traders have up-to-date information about the expected energy production for the upcoming hours. Since weather forecasts change during the day, the expected energy production of wind turbines changes also.

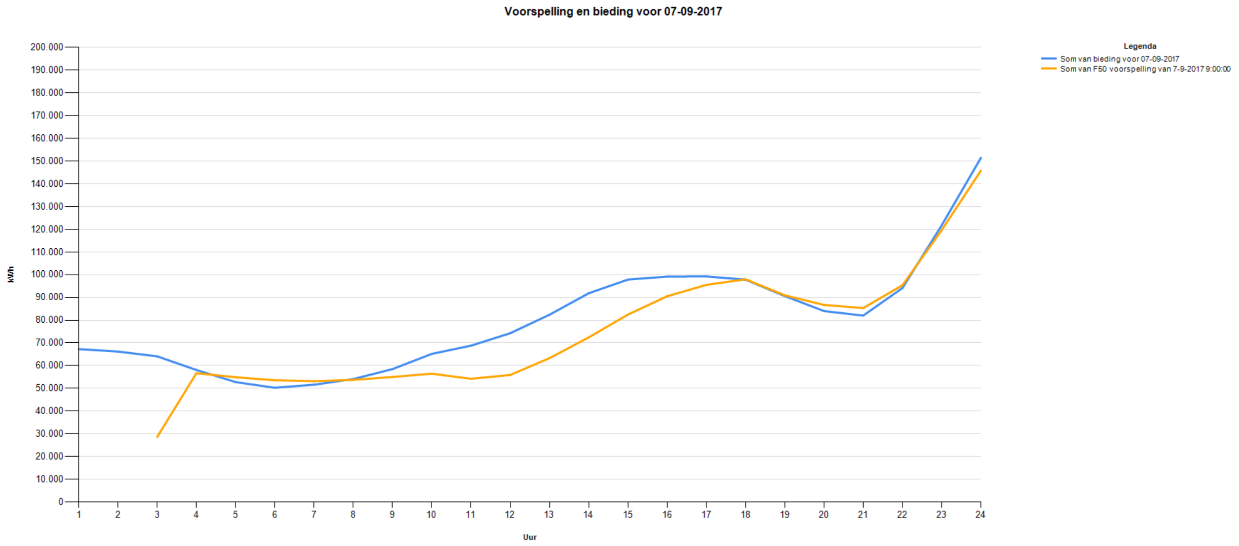

Figure 1.4: Intraday aggregated energy production forecast for wind turbines at 9:00.

Figure 1.4 shows the intraday forecast at 9:00, this forecast is based on more recent weather data. While comparing the day ahead graph in Figure 1.3 and the intraday graph in Figure 1.4, we see a difference in expected energy production (yellow line), while the auction (blue line) remains the same. To minimize the difference between the auction and the actual production, DVEP wants a more accurate day ahead forecast. Therefore, DVEP would like to develop a forecasting model that is more accurate than the current one. This model should be able to translate weather forecasts into expected wind production for DVEP wind turbines.

The core problem for this project is that the current day ahead forecast for power production of wind turbines is believed to be inaccurate. DVEP currently buys this forecast from Company X, which is costly. DVEP thinks that a forecasting model can be developed that is more accurate than Company X. DVEP is especially interested in the time horizon between 15-38 hours ahead, since this is the time horizon that is used for the day ahead auction.

1.4 Research objective and questions

The research goal is to develop a day ahead forecasting model for the power production of wind turbines of DVEP producers. This model should be more accurate than the model that is currently used. After development of the model a comparison should provide insight into which model is most accurate.

To assist DVEP with the accuracy of day ahead wind power production forecasts, we answer the following research question:

"***How to develop a model that is able to translate day ahead weather forecasts into power production forecasts for wind turbines of DVEP producers with higher accuracy than Company X?***"

To answer this research question, the following sub-questions are addressed during the project:

1. Which factors influence the power production of wind turbines according to the literature?

(a) Which weather conditions influence the power production of wind turbines?

(b) Which turbine characteristics influence the power production of wind turbines?

(c) Which site-related factors influence the power production of wind turbines?

To determine which factors influence energy production of wind turbines, we conduct 3 literature reviews. First, we look into which weather conditions have an impact on the power production. After this, we look into wind turbine design to determine which turbine characteristics influence power performance. Lastly, we look into site-related factors.

2. What is known in literature on day ahead forecasting power production for wind turbines?

(a) Which methods are used in literature for forecasting power production for wind turbines?

(b) How can forecast accuracy be measured and the forecast model be validated?

We conduct 2 literature reviews to see which forecasting methods are used in literature for the energy production of wind turbines. We assess which forecasting method is most appropriate for the time horizon we wish to forecast. Also, we look at how to validate forecast models and how to measure forecast accuracy.

3. What is the current situation at DVEP with respect to data?

(a) Which data are available and how are these measured?

(b) Which producers should be selected for model testing?

(c) What are the characteristics of the available data?

Here, we look at the available data and describe how this data was measured. We make a selection of producers that are included in the project scope. We select KNMI stations throughout the Netherlands located near producers of DVEP. We analyze production and weather data and describe how we clean the data. Also, we check whether there is a bias in the weather forecasts.

4. Which forecasting approach should result in the most accurate day ahead forecast according to the data patterns and literature review?

In Chapter 4 we describe our solution design. We introduce the forecasting approach, which is based on the reviews of existing literature and the data that is available. We present the forecasting models that we test and how we expand these to include more predictors. We describe which performance indicator we use to select the most accurate model. Also, we discuss which algorithm we use to solve the minimization problems at hand.

5. **Which day ahead forecasting model is most accurate in production forecasts and how accurate is this model in comparison to Company X?**

In Chapter 5 we select the 3 most accurate regression models based on historical weather and production data. After this, we use historical weather forecasts for all 10 producers to see which model has the most accurate production forecast using day ahead weather forecasts as input. We do this for each producer separately, as well as for the aggregated selection of 10 producers. This enables us to compare the day ahead forecast accuracy of Company X and the models developed in this project. This leads to conclusions and recommendations in the final chapter.

1.5 Research scope

For every research project, a scope should be defined. When investigating problems, other problems may come to the surface. It is very tempting to investigate these problems as well. However, we should stick to the problem at hand. Furthermore, we are dependent on the data that are available. Therefore, we have to draw a line; we do this by defining the research scope:

- The forecast horizon is 15-38 hours ahead, using weather forecasts from 9:00 in the morning as input to forecast power production for each hour of the next day.
- The forecasting model focuses on accuracy in terms of production. The goal is to minimize the financial risk by being as accurate to the realized production as possible. We do not look into the financial implications of the forecasting model.
- The selection of producers should cover at least 10% of the total rated power of DVEP.
- We assume every turbine has storm detection and ice detection.
- We focus on 3 KNMI weather stations to obtain weather data.
- We exclude producers with multiple sources of production (e.g. solar AND wind power), since we cannot distinguish the production of multiple sources under the same EAN (unique connection code).
- We only include producers that have a contract between 01-01-2015 and 01-01-2019.
- We exclude producers with long term downtime between 01-01-2015 and 01-01-2019.

The most important choices regarding the scope of this project are listed above. Throughout the project, the scope is further defined.

1.6 Report outline

In Chapter 2, we conduct a review of the existing literature to answer Sub-questions 1 and 2. First, we look at which factors influence power production of wind turbines. Secondly, we look at what is known about power production forecasting for wind turbines in the literature and how accuracy is to be measured. In Chapter 3, the current situation at DVEP is described. We describe what data are available and how the data was measured. A producer selection is made that is used for data analysis. Production and weather data are analyzed and data are cleaned. Next, in Chapter 4 we discuss our general approach and we introduce the forecast models that are used. We describe how we select the best model that is compared to the model of Company X. Also, we discuss which optimization algorithm is used for parameter estimation. In Chapter 5, we conduct an analysis of the results. First, we look at the effect of adding predictors to the forecasting models. Secondly, we determine which forecasting models are most accurate using historical data. Lastly, we look at the day ahead accuracy of the top 3 models in Chapter 5 using day ahead weather forecasts. Finally, we finish this project with a conclusion and recommendations in Chapter 6. Here, we also give some suggestions for further research and discuss the limitations of the research.

2. Review of Literature

In this chapter, we conduct a literature review to obtain the information that is necessary to develop a forecasting model for the energy production of wind turbines. We answer Sub-question 1 in Sections 2.1, 2.2, and 2.3. Firstly, the wind resource is researched in Section 2.1, since this is the driving power behind the energy production of wind turbines. Secondly, in Section 2.2 we look into the types of wind turbines that are used in practice. After this, we look into the effect of turbine characteristics and site-related factors on the energy production of wind turbines in Section 2.3. Next, we answer Sub-question 2 in Section 2.4 up until Section 2.7. In Section 2.4 we review existing literature to see what forecasting methods are used in the literature. In Section 2.5, we discuss wind turbine power curve modeling techniques that are most promising. Afterwards, in Section 2.6 we show how these models can be expanded with the help of regression analysis. Lastly, in Section 2.7 we discuss how forecasting performance can be measured.

2.1 Wind resource

The winds of the world are unpredictable, intermittent, fickle in speed and direction, and are occasionally extremely strong. This poses a challenge to predict the effectivity of wind energy systems. To do so, we need to understand the wind's behavior.

2.1.1 Wind speed variability during different timescales

The wind speed variability can have a big impact on energy production of a wind energy application. When considering variations in wind speed in time, conventional practices use four time categories (Lynn, 2012; Manwell, McGowan & Rogers, 2009):

- Inter-annual.
- Annual.
- Diurnal (time of day).
- Short-term.

We briefly discuss each category and its implications for wind turbines.

Inter-annual

The wind resource at a particular site differs from year to year. For example, a coastal site in the Western Europe may experience a series of strong 'autumn gales' one year, but not during the next (Lynn, 2012). So a single year's wind speed measurement, although widely used to assess a site's potential, may not always give an accurate picture. Inter-annual variations of up to 5% in average wind speed are pretty common. These variations in wind speed lead to even bigger variations in power output, as we demonstrate later on. The ability to estimate the inter-annual variability at a site is almost as important as estimating the long-term mean wind at the site. Manwell et al. (2009) state that it takes meteorologists approximately 30 years of data to determine long-term values of weather or climate, and that it takes at least five years to arrive at reliable average annual wind speed at a given location. Lynn (2012) adds that climate change can have an influence in the future as well. Who can tell what will happen to the world's wind patterns over the coming decades?

Annual

Most locations experience substantial variations in wind speed during the year. In Figure 2.1, average monthly values for a UK site are presented. The dots represent average values for a single year, the vertical red bars show the range of values recorded over a 10-year period (Lynn, 2012). Figure 2.1 indicates a substantial inter-annual as well as seasonal variation in wind speed at the UK site. We see that the autumn and winter months tend to be most windy, summer months are the calmest. These annual variations are important when it comes to assessing wind energy production in relation to seasonal energy demand.

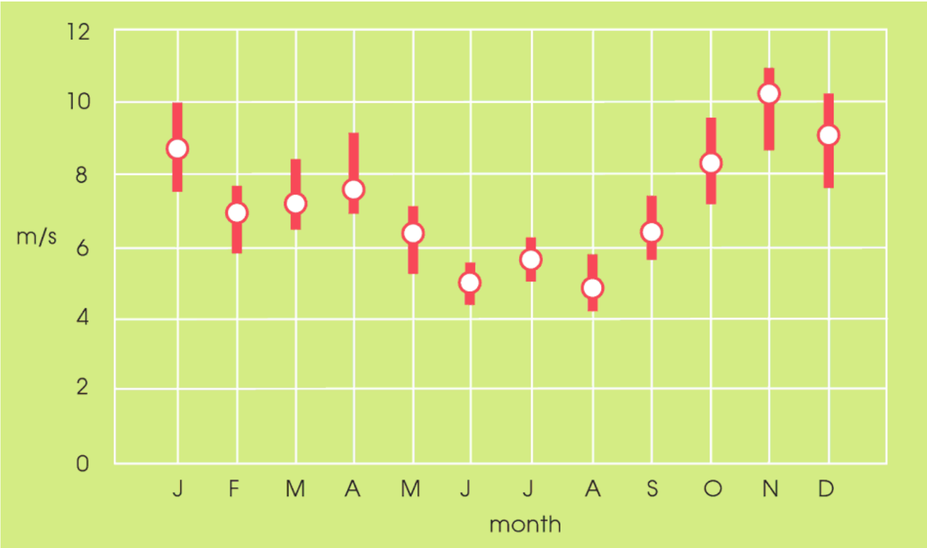

Figure 2.1: Average monthly wind speed and variation at a UK site (Lynn, 2012).

Diurnal (time of day)

In temperate latitudes, large wind variations can occur on a diurnal or daily time scale. This type of wind speed variation is due to differential heating of the earth's surface during the daily radiation cycle (Manwell et al., 2009). A typical diurnal variation is an increase in wind speed during the day with lowest wind speeds during the hours between midnight and sunrise (Lynn, 2012). Daily variations in solar radiation are responsible for diurnal wind variations in temperate latitudes over flat land areas. According to Manwell et al. (2009), the largest diurnal changes occur in spring and summer, and the smallest in winter. Diurnal variation may also vary with location and altitude. At mountainous areas the diurnal variation may be very different than at flat areas. This variation can be explained by the mixing or transfer of momentum from the upper air to the lower air (Manwell et al., 2009).

Short-term

Short-term wind speed variations include turbulence and gusts. These variations are usually measured over time intervals of ten minutes or less. Ten-minute averages are typically determined using a sample each second. Figure 2.2 shows a typical plot of wind speed during a short period of time. It is generally accepted that variations in wind speed with periods from a second to ten minutes have a stochastic character and are considered to represent turbulence (Manwell et al., 2009). Turbulence can be thought of as random wind speed fluctuations imposed on the mean wind speed, this is discussed later in this chapter. Besides turbulence, gusts also contribute to short-term wind speed variations. A gust is a discrete event within a turbulent wind field, gusts can be characterized by their amplitude, rise time, maximum variation and lapse time.

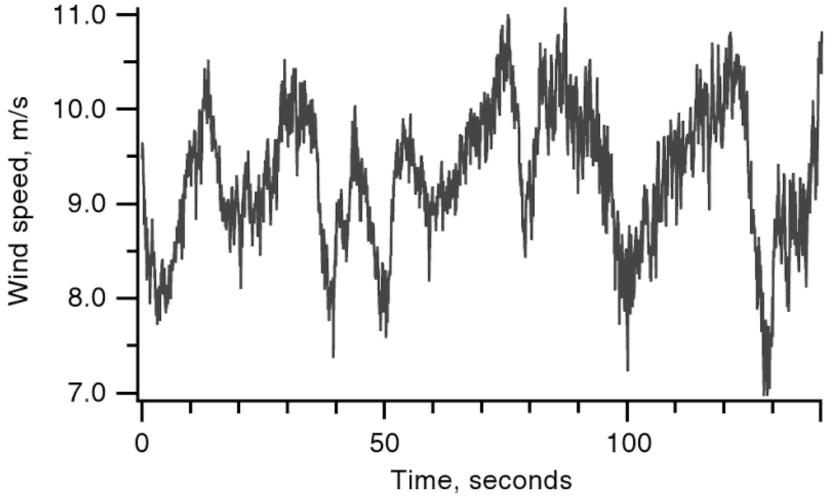

Figure 2.2: Typical plot of wind speed for a short period of time (Manwell et al., 2009).

When considering wind energy applications for a given location, all these time categories have to be taken into account. From long-term wind speed prediction to maximum load calculations due to gusts or turbulence, a wide variety of implications of wind speed variations need to be considered over time.

For wind energy applications, knowledge of wind behavior is of particular importance to successfully utilize the kinetic wind energy. While short-term behavior of wind is of significance with regard to the structural strength and control function of a wind turbine, the long-term characteristics of the wind have relevance with regard to the energy yield (Hau, 2013). The long-term characteristics of the wind can only be determined by using statistical surveys over many years.

2.1.2 Wind speed probability distribution

While year-to-year variation in annual mean wind speeds is hard to predict, wind speed variations during the year can be well characterized in terms of a probability distribution. In the literature it is widely found that the Weibull distribution gives a good representation of the variation in hourly mean wind speed over a year at many typical sites (Burton et al., 2001).

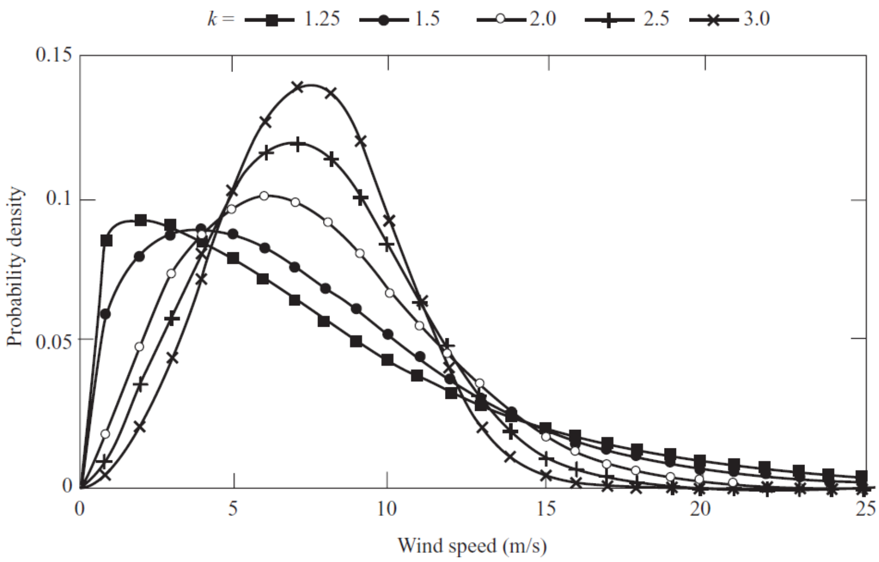

Figure 2.3: Weibull Probability density function of wind speed using different shape parameter values (Burton et al., 2001).

The shape parameter *k* determines the shape of the distribution. A special case of the Weibull distribution is the Rayleigh distribution with *k*=2, this is a fairly typical value for many locations (Burton et al., 2001). On real sites the shape parameter *k* varies from about 1.5 to 2.5. A value of 1.5 is typical for offshore sites, over land the factor reaches values up to 2.5 or somewhat above (Hau, 2013). Offshore locations typically have a longer tail, because there is less surface friction offshore. The Weibull distribution can be used to estimate annual production, the Danish manufacturer Vestas uses a shape parameter of *k*=2 to estimate annual production of its turbines (Vestas, 2017). For short term production estimates, the Weibull distribution is not very useful.

2.1.3 Wind speed at different altitudes

One of the most important factors with respect to the utilization of wind energy is the increase in wind speed with altitude. The moving air masses have less friction against the earth's surface as the altitude increases. The range up to where the wind is undisturbed is between 600 and 2000 m above ground, depending on the time of day and atmospheric conditions (Hau, 2013). This is called the atmospheric boundary layer. The area of the boundary layer closest to the ground is called the Prandtl layer, where flow conditions are dominated by the friction with the earth's surface. In meteorology the area above the Prandtl layer is called the Ekmann layer. The influence of friction is less dominant in this layer, while wind direction is influenced by Coriolis forces due to the earth's spin. Above the Ekmann layer, geostrophic winds flourish since there is no surface friction, and there are large influences by Coriolis forces.

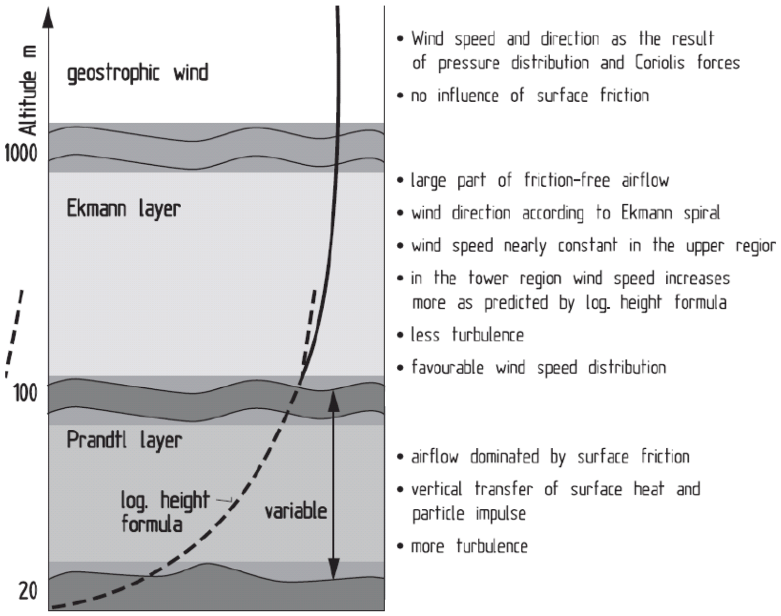

Figure 2.4: Atmospheric boundary layer with Prandtl layer and Ekman layer (Hau, 2013).

The height of Prandtl layer varies with the meteorological conditions. During the night, it is only 20 to 50 m thick, whereas during the day it is between 50 and 150 m thick. A rotor hub height of 60 m is in the Prandtl layer for approximately 30% of the annual hours whereas a hub height of 100 m this is only about 7%. Therefore, the wind conditions of large turbines are extensively influenced by the characteristics of the Ekman layer (Hau, 2013).

2.1.4 Turbulence and gusts

Turbulence is characterized by chaotic changes in wind speed and pressure on a relatively fast time-scale, typically less than ten minutes. Turbulence is mainly generated by two causes, namely friction with the earth's surface and thermal effects (Burton et al., 2001). Friction with the earth's surface can be thought of as flow disturbances caused by hills and mountains or man-made structures. Thermal effects cause air masses to move vertically due to differences in temperature and density of the air. These two effects are often interconnected. Turbulence is a complex process, in order to describe it, it is necessary to take into account the temperature, pressure, density and humidity as well as the motion of the air itself in three dimensions (Burton et al., 2001).

Wind gusts are big, short-term fluctuations in wind speed. Whereas the long-term fluctuations in wind speed are significant to the power output and energy yield of a wind turbine, the loads are marked by short-term fluctuations in wind speed (Hau, 2013). The extreme wind speeds must be taken into consideration for the fatigue strength and loads, although they may occur rarely. Wind gusts and turbulence are especially interesting for fatigue and maximum load calculations, not for power output and energy yield because of their short-term nature (Hau, 2013).

2.2 Wind turbine design

Mankind has been trying to use the wind to its advantage for a long time. The oldest windmill in recorded history is the so-called Persian windmill. It was first described around 900 AD and is a drag-driven windmill with a vertical axis of rotation (Schaffarczyk, 2014). Drag-driven means the windmill generates its power by using drag force, which has the same direction as the wind. Later, in 1279 the Dutch windmill appeared which represented a milestone in technological development. The axis of rotation changed from vertical to horizontal. From an aerodynamic point of view, the Dutch concept began a movement toward lift-driven wind turbines instead of drag-driven (Schaffarczyk, 2014). Lift force refers to forces perpendicular to the wind direction. Today, a wide variety of wind turbines are used. Wind turbines can either rotate about a horizontal or vertical axis, therefore wind turbines are classified as Horizontal Axis Wind Turbines (HAWTs) or Vertical Axis Wind Turbines (VAWTs). HAWTs are the dominant design principle in wind energy technology today, since this design has a higher power output (Hau, 2013). Therefore, we take a closer look at HAWTs.

2.2.1 Horizontal axis wind turbine

Within the HAWT classification, there are a lot of variations. These variations include the number of blades, arrangement of rotor, variable/constant speed, blade pitch control and yawing options. According to Schaffarczyk (2014), standard HAWTs have the following properties:

- Horizontal axis of rotation.
- Three bladed.
- Driving forces mainly from lift.
- Upwind arrangement of rotor; tower downwind.
- Variable speed/Tip-speed ratio (TSR) control.
- Blade pitch control after rated power is reached.

Figure 2.5 shows a schematic arrangement of a HAWT. The components and their configuration are typical for a standard HAWT (Hau, 2013).

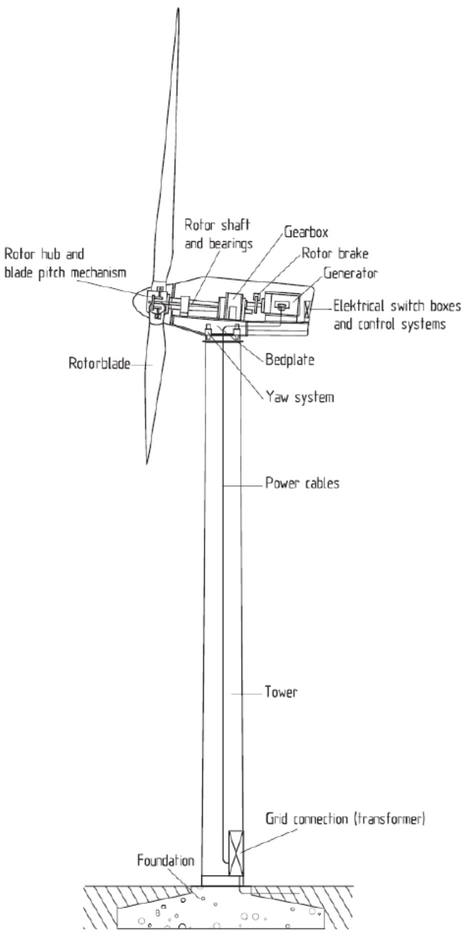

Figure 2.5: Schematic arrangement of a typical HAWT (Hau, 2013).

Most turbines have a hub height between 40 and 120 m, in extreme cases the height can go up to 180 m. Rotor blades range in length from 20 to 80 m and most turbines have 3 rotor blades. Turbines are built with a rated power of up to 8 MW today. Offshore wind turbines are usually larger than onshore ones; this typically leads to a higher rated power. Today's onshore turbines range up to 120 m normally and rarely exceed 3 MW.

If a turbine wants to capture the full power of the wind, it has to be oriented to the wind direction correctly. Wind direction is constantly measured and the yaw system makes sure the horizontal axis of rotation is perpendicular to the wind direction. There are three different yawing methods (Hau, 2013):

- Yawing by aerodynamic means (wind vanes or fan-tail wheels).
- Active yawing with the help of a motorized yaw drive.
- Free yawing of rotors located downwind.

Most modern wind turbines use a motorized yaw drive since wind vanes or fan-tail wheels are not able to move massive tower heads of big turbines and locating rotors downwind leads to a big power loss due to disturbed airflow (Hau, 2013). Locating rotors downwind means the tower is facing the wind direction instead of the rotors, this 'backward' configuration leads to a disturbance in airflow at the rotor blades which leads to power loss. Therefore, most rotors are located upwind of the tower.

Tip-speed ratio (TSR) control makes sure the turbine can operate at variable speeds. TSR is the ratio between the tip speed of the blade and the wind speed, TSR is related to efficiency, the optimum varies with blade design (Hau, 2013). TSR control is used to ensure the ratio between the speed of the tip of the blade and the wind speed is kept at a constant, optimal rate. This is done to achieve maximum efficiency.

The last property which we discuss is power control, in case of strong winds it is necessary to waste part of the excess energy to avoid damaging the wind turbine in high wind speeds. All wind turbines are therefore designed with some sort of power control. There are two different ways of doing this safely on modern wind turbines:

- Blade pitch control.
- Stall control.

The most effective way of adjusting the aerodynamic angle of attack, and thus the input power, is by mechanically changing the rotor blade pitch angle (Hau, 2013). The rotor blade is turned about its longitudinal axis with the aid of actively controlled actuators with this method. This is not only done for safety measures, but also to maintain maximum output after rated power is reached.

When a turbine does not have blade pitch control, rotor blades have a fixed angle which is called stall control. Stall controlled wind turbines have blade designs that create turbulence on the side of the blade not facing the wind when the wind speed increases. As the actual wind speed increases, at some point the rotor blade starts to stall, which prevents it from reaching dangerously high speeds.

2.3 Power output of wind turbines

A wind turbine has to capture as much of the wind's power as possible and convert it efficiently into electricity. This is done by converting kinetic energy of the wind into electrical energy. The performance of a wind turbine depends crucially on the conditions at a particular site including the wind's average speed and variability (Lynn, 2012). To see what factors influence the power output of wind turbines, (Lynn, 2012) starts by considering a well-known equation of fluid mechanics:

$$P = \frac{1}{2}\rho A v^3 \tag{2.1}$$

where:

P = power in W

ρ = air density in kg/m^3

A = area of the intercepted airstream in m^2 (swept area of rotor blades)

v = wind velocity in m/s

Equation 2.1 is used to calculate the available kinetic wind power. We see that the available wind power increases with the air density, the area of the intercepted airstream and the wind velocity. Especially the wind velocity has a big impact due to its cubic relationship with power. To illustrate its impact, a doubling in wind velocity leads to an eight times higher available wind power. Air density and the swept area of the rotor blades have an influence as well.

In Equation 2.1, the available wind power can be calculated. However, the power that is extracted by wind turbines is smaller. There are fundamental limitations to rotor efficiency that prevent wind turbines from converting 100% of the available wind power. Therefore, Burton et al. (2001) added a power coefficient to the equation, resulting in Equation 2.2:

$$P = \frac{1}{2}\rho C_p A v^3 \tag{2.2}$$

where:

C_p = power coefficient (fraction of the available wind power that may be converted by the turbine into mechanical work)

The power coefficient has a theoretical maximum value of 59.3% (Betz limit) due to the principles of conservation of mass and momentum of the air stream, though in practice lower peak values are reached (Burton et al., 2001). Incremental improvement in the power coefficient are constantly sought by detailed design changes in wind turbines. However, these changes only lead to a modest increase in power output. Major increases in power output can only be achieved by increasing the swept area of the rotor or by locating the wind turbines on sites with higher wind speeds (Burton et al., 2001).

2.3.1 Turbine characteristics and power output

A cause of reduced output is rotor yawing with the wind direction. Yawing is the process of aligning the rotor with the wind direction. This is done to use the wind to its highest potential. Even with sensitive yawing a certain loss is unavoidable. Various investigations have shown a loss of about 2 to 3% in energy yield of the turbine with a correctly operating yawing mechanism (Hau, 2013). Losses increase when there are frequent wind direction changes on site.

Another cause of reduced output can be explained with the C_p– λ performance curve. Here, C_p is the power coefficient and λ, the tip-speed ratio (TSR). The TSR is the ratio between the speed of the tip of the rotor blade and the wind speed.

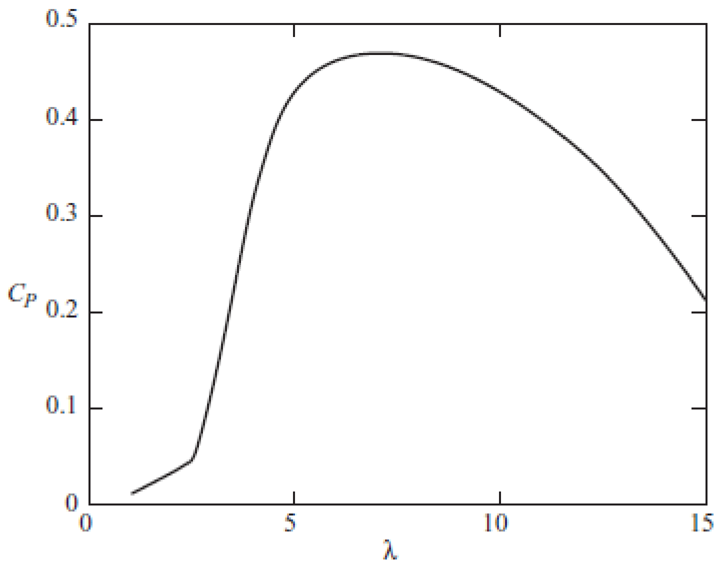

Figure 2.6: C_p – λ performance curve for a modern three-bladed turbine showing losses (Burton et al., 2001).

The first thing to note is that the maximum value of C_p is only 0.47, which is smaller than the Betz limit, achieved at a TSR of 7 (Burton et al., 2001). To have maximum efficiency, it is crucial that the TSR is kept at this constant rate. The fact that this value is considerable smaller than the Betz limit is due to stall, tip and drag losses among other losses.

To limit the power loss, most modern turbines operate at variable speed, this is done by TSR control. TSR control monitors the speed of the rotor so it can be continuously adjusted such that the TSR remains constant at the level which gives the maximum C_p. This significantly increases the efficiency of the turbine and leads to a higher output than turbines operating on constant speed (Burton et al., 2001).

An option that also affects the power output is blade pitch control. A change in angle of attack can have a big impact on the power output. Blade pitch control is also used to regulate the TSR, thus is connected with TSR control. Active pitch control is necessary to maintain a constant, optimal TSR after rated wind speed is reached (Burton et al., 2001).

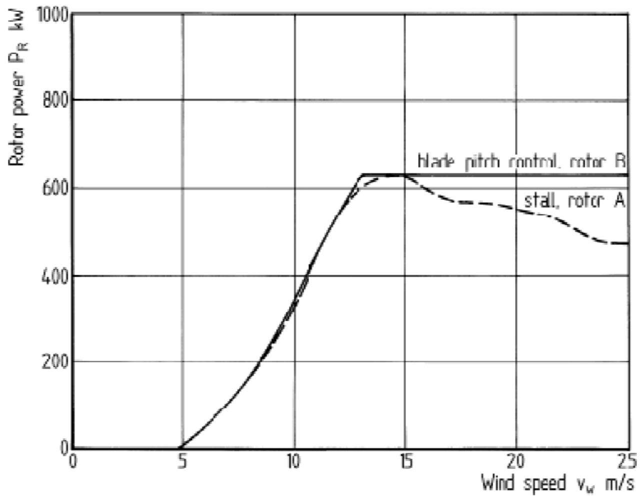

Figure 2.7: Power curve of blade pitch control versus stall control (Hau, 2013).

The pitch angle should continuously be adjusted after rated power is reached to maintain the highest efficiency. This is where blade pitch control distinguishes itself from stall control. After rated power is reached blade pitch control is able to maintain optimal TSR so rated power is achieved at a wider wind range than using stall control. Figure 2.7 illustrates that stall controlled turbines are less efficient at high wind speeds.

Obviously, wind speed affects the power output of the turbine. However, wind can reach tremendous speeds, leading to dangerous situations. To prevent turbine damage, the blades can be feathered and the turbine is turned off, this happens when cut-out speed is reached. This means only a certain range in the wind speed domain can be utilized (Lynn, 2012).

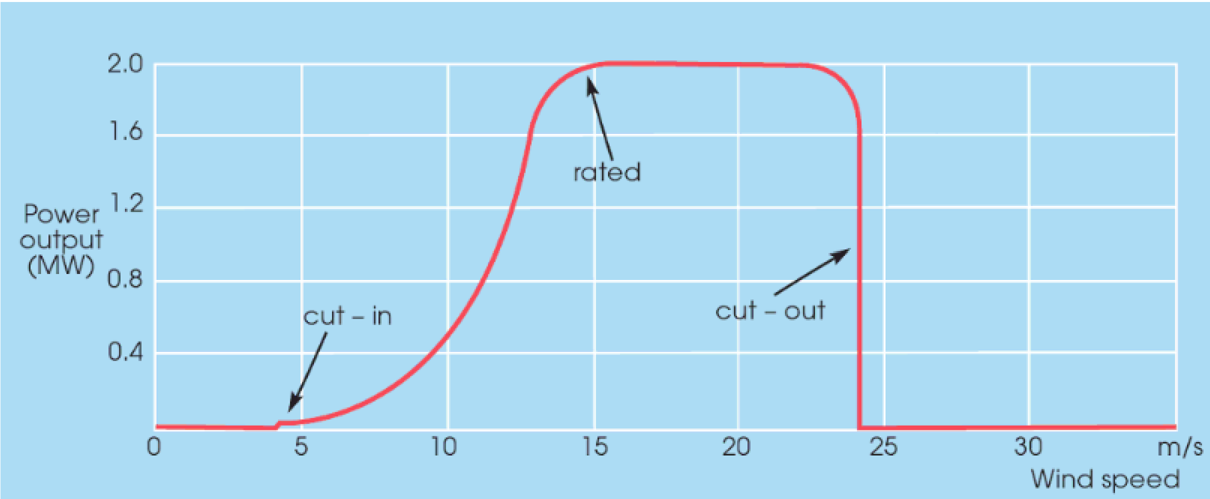

Figure 2.8: Theoretical power curve for a standard 2 MW turbine (Lynn, 2012).

Besides the range in wind speed above the cut-out speed, the lowest wind speeds can also not be utilized. This is due to the fact that the consumption of the turbine is higher than the energy output, which results in a negative yield. Therefore, turbines only start operating after a certain cut-in speed has been reached. After the cut-in speed, the power output rises until the rated power is reached where it ideally will remain until the blades are feathered, and the turbine is shut off. Cut-in and cut-out speeds can vary depending on design type and environmental factors (Lynn, 2012). The newest turbine designs from the German manufacturer Enercon have cut-out speeds between 28-34 m/s (Enercon, 2015). However, older turbines have lower cut-out speeds.

2.3.2 Site-related influences on power output

The density of the air has an influence on power output and varies with both elevation and temperature. Cold air at sea level is considerably denser than warm air at high upland sites (Hau, 2013). This is illustrated in figure 2.9:

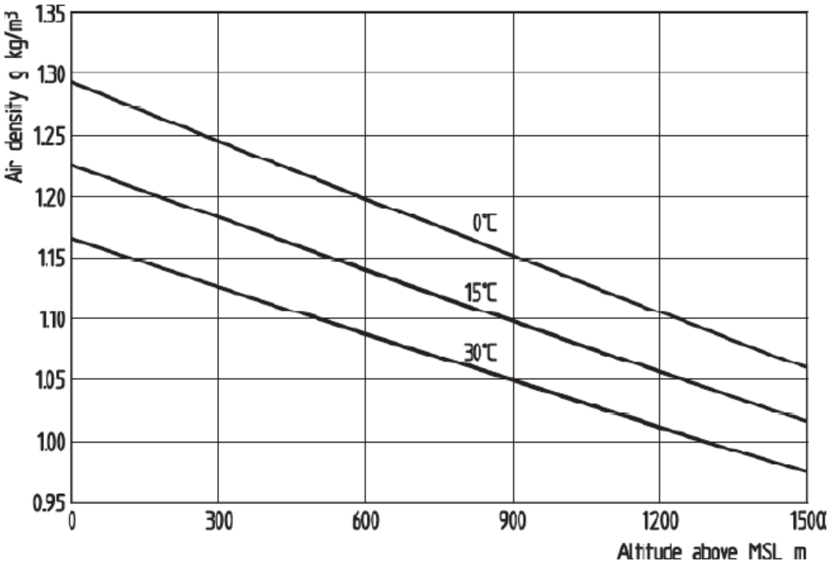

Figure 2.9: Air density as a function of the geographic altitude and temperature (Hau, 2013).

Air density decreases when temperatures increases from 0 °C. The density is largest at mean sea level (MSL), the decrease in air density is already noticeable at a few hundred meters, as well as the change in the temperature range between summer and winter, so that its influence on turbine performance cannot be neglected (Hau, 2013). This is supported by Lynn (2012), who states that a turbine produces more power during winter than midsummer, in winds of the same speed. Large manufacturers such as Enercon and Vestas assume a standard air density of 1.225 kg/m^3 in their power curves (Enercon, 2015; Vestas, 2017).

As a turbine extracts energy from the wind, it leaves behind a wake with reduced wind speeds and increased levels of turbulence (Burton et al., 2001). Another turbine operating in this wake or deep inside a wind farm will suffer and produce less energy. This is especially the case for offshore turbines, where other turbines are the only obstacles. For onshore turbines, buildings, trees or other objects in the vicinity of the turbine can have a large influence on the power output.

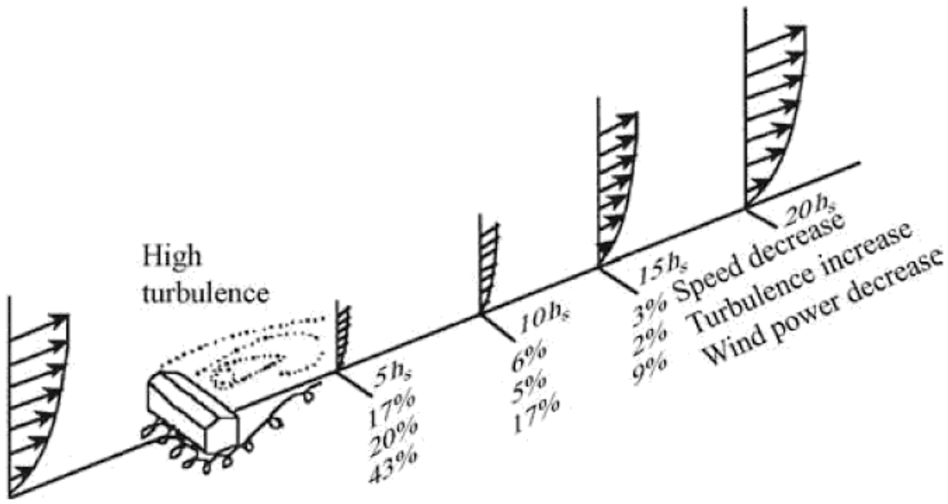

Figure 2.10: Wind speed, power and turbulence effects downstream of a building (Manwell et al., 2009).

In Figure 2.10 the change in available power and turbulence is illustrated in the wake of a sloped-roof building. At a distance of 5 times the height of the building (5 h_s), the wind power is decreased by 43% mainly due to an increase in turbulence and a decrease in wind speed. At larger distances the turbulence reduces and wind speeds increase again which results in a smaller loss in wind power (Manwell et al., 2009). Besides buildings and other manmade objects, wooded inland regions and mountainous areas have an impact on the annual energy yield as well (Hau, 2013). It is difficult to estimate flow conditions in complex terrain in detail. The flow field is affected by topographic shapes such as slopes or depressions. Depending on wind direction, wooded inland areas cause variable vertical wind shears (Hau, 2013). Seasonal changes need to be accounted for also, during summer the trees have a larger collection of leaves in comparison to the winter, which affects wind flow differently. Each location has its own specific air flow conditions, which have to be examined carefully when estimating annual energy yield. However, from the point of view of the practical operation of the wind turbine, the influence of turbulence on the energy yield is, as a rule, not severe (Hau, 2013).

Apart from the turbulence of the wind, other weather-related factors can influence the power output of wind turbines also. Primarily, icing of the rotor blades at temperatures below zero can alter the aerodynamic profile of the blades significantly (Hau, 2013). However, due to safety reasons the turbine has to be turned off so there is no sense in taking its influence on the power curve into consideration. The influence of snowfall or long-lasting rain can have a more practical significance. According to Hau (2013), recent studies have shown that the surface roughness of the rotor blades changes due to the rain, which can result in power losses.

Another factor that influences the surface roughness of the rotor blades is soiling. After a certain operating period, rotor blades exhibit soiling phenomena (Hau, 2013). The dirt on the surface of the blades is produced after long periods of dryness and high temperatures in summer. During this time there are more dust and insects in the air, which can stick to the blades. Soiling is not only dependent on the weather, but also on the site. Extreme conditions are observed in desert-like conditions. A prolonged operation with badly soiled rotor blades leads to a great loss in energy yield (Hau, 2013).

2.4 Forecasting wind power production

In this section we look into which forecasting method is best suited for this project. We discuss some promising wind power forecasting approaches used in literature and what type of input data are used for these approaches. We focus on day ahead wind power forecasting, since that is the forecasting horizon that we use in this research.

2.4.1 Forecasting approaches for wind power production

When considering forecasting methods for power production of wind turbines, a classification of forecasting horizon needs to be made. Forecasting serves different purposes for different time-scales, for 8 hours-ahead the main purpose is real-time grid operations, while multiple-days-ahead one of the main purposes is maintenance planning (Wang et al., 2011). Table 2.1 shows the classification of wind power forecasting and its applications:

Time-scale	Forecast horizon	Applications
Immediate-short-term	8 hours-ahead	- Real-time grid operations - Regulation actions
Short-term	48 hours-ahead	- Economic load dispatch planning - Load reasonable decisions - Operational security in spot market
Long-term	Multiple-days-ahead	- Maintenance planning - Operation management - Optimal operating cost

Table 2.1: Classification of wind power forecasting and its applications (Wang et al., 2011).

In this research, we use wind power forecasting for operational security in the spot market. Besides classification based on the prediction horizon, wind power forecasts can also be classified based on their methodology. Here, the physical approach, statistical approach or a combination (hybrid approach) can be taken.

Short term wind power forecasting requires predictions of meteorological variables from Numerical Weather Prediction (NWP) models as input. The physical and statistical approaches differ in how they translate meteorological predictions into power production forecasts.

The physical approach focuses on the description of air flow around the turbine and uses the manufacturer’s power curve for estimating power production. The core idea of the physical approach is to refine the NWPs by using physical considerations about the terrain such as roughness, orography and obstacles (Wang et al., 2011). The manufacturer’s power curve is used to translate the refined NWP data into power production forecasts.

The statistical approach is based on a vast amount of historical data to capture the relation between meteorological forecasts or historical meteorological measurements and historical power production. It does not use physical considerations at the turbine site (Wang et al., 2011). The hybrid approach combines the physical and statistical methods and tries to use the advantages of both methods.

According to Giebel (2003) and Landberg et al. (2003), the various forecasting approaches can be classified according to the type of input. This is illustrated in Figure 2.11 and Table 2.2:

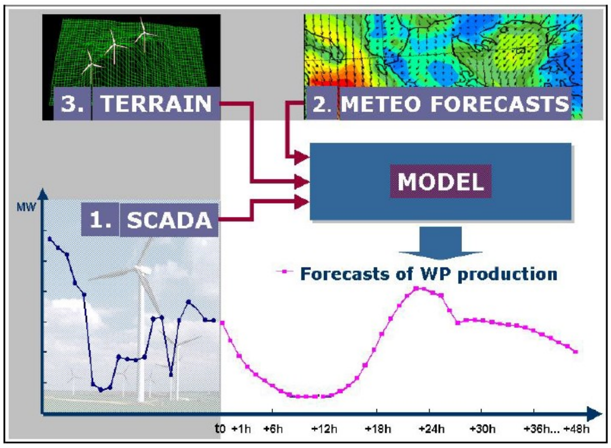

Figure 2.11: Input sources for forecasting wind power production (Giebel, 2003).

The various forecasting approaches can be classified according to the type of input that is used. All models involving Meteo Forecasts have a horizon that is limited by the NWP model (usually 48 hours). Models that use online production data use Supervisory Control and Data Acquisition (SCADA) systems. Models that use terrain specific data use information about terrain complexity, obstacles, orography and turbine specifications to enhance the forecast accuracy.

Input	Approach	Horizon
1	Statistical	< 6 hours
2	Physical/statistical	> 3 hours
2 + 3	Physical	> 3 hours
1 + 2	Statistical	-
1 + 2 + 3	Combined	-

Table 2.2: Forecasting approach with different input data (Giebel, 2003).

Table 2.2 illustrates that different approaches should be taken for different input data from Figure 2.11. The horizon at which good results can be achieved also differs for each approach and input combination. The approach should be chosen according to the data that are available and the horizon to be forecast (Giebel, 2003).

Statistical approach

Statistical prediction methods include linear and non-linear regression models, but also autoregressive and black-box type models. Black-box type models include most Artificial Intelligence (AI) based models like Artificial Neural-Networks (ANN) and Support Vector Machines (SVM) (Foley et al., 2012). These are called black-box models, because not even the designers can analyze what is happening inside the model. The linear and non-linear regression models involve estimating parameters based on historical data, it is essential to choose the right meteorological variables and use suitable models. Lastly, some statistical methods include an autoregressive part. Methods such as Auto-Regressive Integrated Moving Average Model (ARIMA), the Box-Jenkins methodology and Kalman filters are used (Wang et al., 2011). However, this requires the use of online data from SCADA systems. The use of online data with autoregressive methods improves the forecasts up to 6 hours ahead (Giebel, 2003). However, in the short-term horizon (up to 48 hours) wind speed forecasts, the influence of meteorological predictions becomes more important and the use of NWP models becomes essential (Foley et al., 2012).

Physical approach

Physical models tailor the predictions from NWP models to the turbine site by using a detailed description of the terrain. The use of 3D Computational Fluid Dynamics (CFD) models allows physical models to accurate compute the air flow at the turbine site (Lange & Focken, 2006). Along with the manufacturer's power curve, this leads to power production forecasts. Most physical models use Model Output Statistics (MOS) to avoid systematic forecasting errors and to correct the predicted power output of the manufacturer's power curve (Foley et al., 2012; Giebel, 2003). MOS can be used to avoid systematic forecasting errors in production forecasts. It involves the use of historical weather predictions and historical power production to adjust the manufacturer's power curve.

According to Giebel (2003), sub-models for orography and surface roughness were not always able to improve the results. However, the use of MOS was deemed useful. A large influence regarding the power curve was found. The theoretical power curve given by the manufacturer and the power curve found from the data proved to be rather different in many cases. Even the power curve estimated from different years showed strong differences. Nevertheless, the largest influence on the forecast error originated from the NWP model itself (Giebel, 2003).

2.5 Power curve modelling techniques

Power curve modelling techniques are used to model the relationship between wind speed and wind power production. This is a form of simple regression, since only one predictor is used. Wind speed conversion to wind power through Wind Turbine Power Curve (WTPC) modelling is a key pillar of any wind power prediction model (Marciukaitis et al., 2017). The easiest way to do this is to use theoretical (manufacturer's) wind power curve. However, in most cases this leads to additional errors due to differences in theoretical and real-life wind power measurement data. Many different mathematical modeling techniques for WTPC are available. Literature classifies these techniques into parametric techniques and non-parametric techniques (Lydia et al., 2014).

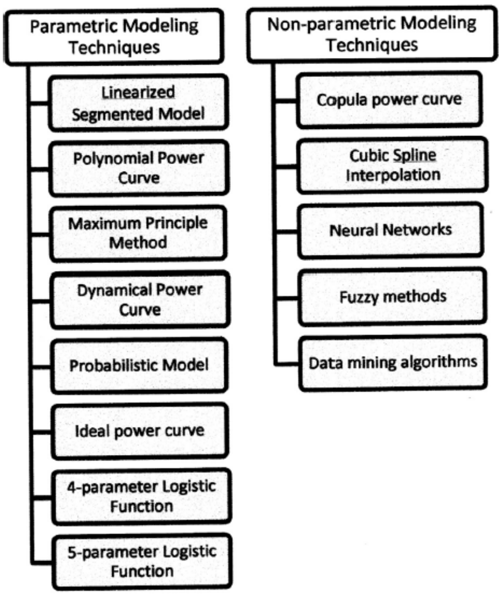

Figure 2.12: WTPC modelling techniques (Lydia et al., 2014).

Each turbine has a different power curve depending on model type and environmental factors like orography, site turbulence and complexity of terrain. Therefore, accurately modelling the power curve for power output prediction is essential (Marciukaitis et al., 2017). Figure 2.11 shows an overview of techniques, these are not all techniques that are available in the literature. We highlight some techniques that show promising results according to the literature. Parametric techniques are mostly used in the physical approach, while non-parametric techniques are often used in statistical approaches. First, we focus on the parametric modeling techniques, after this we discuss the non-parametric techniques.

2.5.1 Parametric techniques

Parametric techniques are based on solving mathematical expressions. These techniques are often used in the statistical approach to estimate the power curve. Some techniques are only able to calculate a part of the power curve, which is shown in Equation 2.3. The actual power output, *P(v),* can be expressed as given below (Carrillo et al., 2013):

$$P(v) = \begin{cases} 0, & v < v_{ci}, v > v_{co} \\ q(v), & v_{ci} \le v \le v_r \\ P_r, & v_r \le v \le v_{co} \end{cases} \qquad (2.3)$$

where:

v = wind speed in m/s

v_{ci} = cut-in wind speed in m/s

v_{co} = cut-out wind speed in m/s

v_r = rated wind speed

P_r = rated power

Here, *q(v)* is the variable region between the cut-in speed and the rated speed at which rated power is reached. This distinction has to be made, since some techniques focus on approximating this part of the power curve instead of the entire curve. The most typical mathematical equations for representing *q(v)* are the polynomial power curve, exponential power curve and approximate cubic power curve (Carrillo et al., 2013). All of the equations listed in this subsection, except for the approximate cubic power curve, are used for curve fitting, which means the parameters have no physical meaning.

Approximate cubic power curve

The cubic power curve is estimated by assuming the power coefficient (C_p) is equal to the maximum value of the effective power coefficient ($C_{p,max}$) of the turbine type. The term effective means that electrical and mechanical losses are included in this coefficient. The resulting equation is:

$$q(v) = \frac{1}{2}\rho A C_{p,max} v^3 \qquad (2.4)$$

This equation is similar to Equation 2.2. To be able to calculate the resulting power output, the air density, area of the swept rotor and the maximum power coefficient have to be known. Of course, the entire power curve can also be calculated using this equation, whether or not this impacts the results negatively is not certain. The approximate cubic power curve showed the best results according to Carrillo et al. (2013) and Lydia et al. (2014). However, Thapar et al. (2011) argue that models based on Equation 2.2 are cumbersome and are not suitable for accurately calculating hourly energy production.

Polynomial power curve

Polynomial functions can be used to approximate both the non-linear part of the power curve as well as the entire curve. Which part of the curve is estimated depends on the degree of the polynomial, the polynomial function is expressed as follows:

$$P(v) = a_0 + a_1 v + a_2 v^2 + a_3 v^3 + \cdots + a_n v^n \tag{2.5}$$

Here, n is the order of the polynomial and a_n are the parameters of the polynomial function to be estimated. Among the polynomial functions, the quadratic (*n*=2) power curve showed the worst results when estimating $q(v)$ (Carrillo et al., 2013). The ninth-order polynomial showed the most promising results when estimating the entire curve $P(v)$ (Lydia et al., 2014).

Exponential power curve

Exponential functions are used in literature to estimate the power curve. A lot of adaptations of these kinds of functions are used. Recently, Marciukaitis et al. (2017) used the following function to estimate the entire curve:

$$P(v) = P_r\left(1 + \left(\frac{\beta}{v}\right)^{\alpha}\right)^{-y}, \qquad \alpha,\ \beta,\ y > 0 \tag{2.6}$$

Here, β, α, and k are positive parameters which have to be estimated. A lot of different other exponential functions have been used in literature, this function yielded the best results after cross-validation according to Marciukaitis et al. (2017). They claimed that this model outperforms the polynomial and approximate cubic power curve functions.

Logistic power curve

The shape of the power curve can be approximated by using a logistic expression with varying parameters. Lydia et al. (2013) experimented with four and five parameter logistic expressions successfully. The four parameter logistic function is expressed as follows:

$$P(v) = \alpha\left(\frac{1+me^{-\frac{v}{\tau}}}{1+ne^{-\frac{v}{\tau}}}\right) \tag{2.7}$$

Parameters $\alpha, m, n,$ and τ have specific ranges giving the function favorable results. The five parameter logistic function is expressed as follows:

$$P(v) = d + \left(\frac{a-d}{\left(1+\left(\frac{v}{f}\right)^{b}\right)^{g}}\right), \qquad f, g > 0 \tag{2.8}$$

The five parameter logistic function showed the best results of the parametric functions (Lydia et al., 2013). However, this method was not compared to the exponential, polynomial or approximate cubic power curve. It did outperform some non-parametric techniques like neural networks, fuzzy logic and data mining algorithms.

2.5.2 Non-parametric techniques

Several non-parametric techniques are used to find the relationship between the input wind speed data and output power. We highlight the techniques that are most widely used in the literature and show the most promising results. Most of these techniques are used in the statistical approach. These techniques are far more complex than their parametric counterparts, therefore we only give a short description of each.

Artificial neural networks

An Artificial Neural Network (ANN) is an information-processing model simulating the biological nervous system (Lydia et al., 2014). It has the capacity to derive meaning from complicated and imprecise data and extracts patterns and trends that are too complex to be identified by humans. Lydia et al. (2014) mentioned three ANNs that were widely used, Generalized Mapping Regressor (GMR), feed forward Multi-Layer Perceptron (MLP) and a General Regression Neural Network (GRNN).

Fuzzy methods

Fuzzy logic is a multi-valued logic which deals with approximate reasoning. Lydia et al. (2014), who made a comprehensive review on WTPC modeling techniques, distinguishes three types of fuzzy methods, fuzzy cluster center method, fuzzy c-means clustering and subtractive clustering. Fuzzy cluster center method clusters data using a clustering algorithm, the accuracy of the model increases with the number of clusters. The performance of the fuzzy cluster center method is the best out of the fuzzy methods.

Data mining algorithms

Data mining is the process of analyzing data present in huge databases and extracting valuable information and patterns. For most wind farms, huge volumes of data are available which presents opportunities for the application of data mining algorithms (Lydia et al., 2014). Non-parametric models of a WTPC have been obtained using five data mining algorithms, random forest, Multi-Layer Perceptron (MLP), M5P tree, boosting algorithm and k-Nearest Neighbor (k-NN). The last algorithm mentioned yielded the best results.

2.5.3 Summary of WTPC modeling techniques

Models based on the basic concept of power available in the wind, like the approximate cubic power model do not give accurate results. Models based on the historic wind speed-power data of a wind turbine using curve-fitting techniques perform better. These models include the polynomial, exponential and logistic power curve models. Out of the polynomial functions, the ninth-degree polynomial had the most accurate results. The non-parametric models give accurate results as well; however, these are not desirable because they are complex to implement due to their underlying algorithms. Lydia et al. (2013) used four optimization algorithms for parameter estimation with logistic parametric models. These algorithms included a genetic algorithm (GA), evolutionary programming (EP), particle swarm optimization (PSO) and differential evolution (DE). The five parameter logistic function got the best results when using the DE algorithm. According to Lydia et al. (2013), the five parameter logistic function using the DE algorithm outperforms the non-parametric techniques. It is not clear whether or not the logistic power curve techniques outperform the polynomial or exponential models.

2.6 Regression models

In Section 2.5, we discussed a variety of WTPC models, the parametric models are a form of simple regression. Regression models are causal models that assume the variable to be forecast (dependent variable) is somehow related to other variables (independent variables or predictors). These relationships take the form of a mathematical model, which can be used to predict future values of the dependent variable. Depending on the nature of the relationship, the forecaster may develop a linear or a nonlinear mathematical model (Hoshmand, 2009). In this section, we discuss several regression models: simple linear regression, multiple linear regression and nonlinear regression.

2.6.1 Simple linear regression

In case of simple linear regression, we are interested in the relationship between one predictor (X) and the dependent variable (Y). The value of the predictor is used to predict the value of the dependent variable, this is called a bivariate relationship (Hoshmand, 2009). For example, an economist might be interested in the effect of personal income (independent, X) on customer expenditure (dependent, Y). The simplest model to describe the relationship between variable X and Y is a straight line, which is called a linear relationship. The linear relationship between the two variables X and Y can be expressed with a simple linear equation:

$$Y = a + bX + \varepsilon \tag{2.9}$$

where:

Y = dependent variable

X = independent variable

a = regression constant

b = regression coefficient

ε = error term

An error term is added, because most observations will not be on the regression line. The error term captures the difference between the observed value of Y and the predicted value of Y. The parameters a and b are estimated so the average error term is zero. To calculate the parameters used in the regression model, the least squares method can be used. This is a standard method in regression analysis called least squares regression. The objective of the least squares method is to minimize the Sum of Squared Errors (SSE). This can be done by minimizing the following objective function:

$$SSE = \sum_{i=1}^{N} (Y_e(i) - Y_a(i))^2 \tag{2.10}$$

In Equation 2.10, $Y_a(i)$ is equal to the actual value of Y for observation *i*, $Y_e(i)$ is the predicted value of Y for observation *i* of the used model. The advantage of this method of parametrization is its simplicity. Most statistical packages use optimization algorithms to calculate the parameters that minimize the loss function shown in Equation 2.10. To evaluate how well the regression model fits the observed data, the coefficient of determination (R^2) can be used. This is a statistical measure that indicates the percentage of the total variance that can be explained by the model (Hoshmand, 2009). However, we should exercise caution with the interpretation of the coefficient of determination. The simple linear regression model cannot always be used. The model is based on assumptions that must be met before we can properly interpret the R^2 statistic. For simple linear regression the following assumptions must be met:

- Normality of errors.
- Linearity.
- Homoscedasticity.
- Independence of errors.

Normality requires the errors to be normally distributed with a mean of zero. Linearity means that the relationship between the dependent and independent variable is linear. Also, the regression equation should be linear in the parameters. The assumption of homoscedasticity requires that the variation around the line of regression is constant for all values of X. This can be checked by plotting the residuals (errors) against all values of X. Lastly, the independence of errors assumption requires that in the population, the residuals should be independent for each value of X (i.e. the residuals may not show autocorrelation). Violations of independence especially arise in time-series regression models. The residuals can be plotted against time to check whether this assumption is valid, if this is not conclusive, a Durbin-Watson test can be conducted.

2.6.2 Multiple linear regression

In the previous section, we looked at regression using a single predictor. However, in practice a lot of dependent variables can be explained using multiple predictors. This can be done by expanding the simple linear regression model into multiple linear regression. Multiple linear regression allows us to include more information in the model (Hoshmand, 2009). However, this does not necessarily make the regression model more accurate. The regression equation is quite similar to that of simple linear regression:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \cdots + b_nX_n + \varepsilon \tag{2.11}$$

where:

Y	= dependent variable
$X_1, \dots, X_n$	= independent variables
$a, b_1, \dots, b_n$	= regression coefficients
ε	= error term

The method for parameter calculation is the same as for simple linear regression. Least squares can be used to calculate the regression coefficients. The R^2 statistic is interpreted similarly as with simple linear regression, but now tells us about the amount of variance that is explained by several predictors instead of one. However, one should exercise caution with the interpretation of this statistic since several assumptions must be met. Violations of these assumptions may present difficulties when using a regression model for forecasting purposes (Hoshmand, 2009). Multiple linear regression has one more assumption than simple linear regression, because we are dealing with multiple predictors. The extra assumption states that there should be no multicollinearity between the predictors. The predictors should not be correlated to each other. The other assumptions of normality of errors, linearity, homoscedasticity and independence of errors should also be met when using multiple linear regression.

2.6.3 Nonlinear regression

Nonlinear regression extends linear regression for use with a much larger and more general class of functions. Almost any function that can be written in closed form can be incorporated in a nonlinear regression model. Unlike linear regression, there are very few limitations to the way parameters can be used in the functional part of a nonlinear regression model (Bates & Watts, 2008). In practice, a lot of relationships between dependent and independent variables cannot be described properly with a function that is linear in the parameters. A nonlinear model is any model of the basic form:

$$Y = f(X_i; b_i) + \varepsilon \tag{2.12}$$

where:

Y	= dependent variable
$X_1, \dots, X_n$	= independent variables
$b_1, \dots, b_n$	= regression coefficients
ε	= error term
$f(X_i; b_i)$	= nonlinear function with predictors X_i and parameters b_i

Equations 2.8, 2.9 and 2.10 are examples of functions that are nonlinear in both the variables and the parameters. Therefore, the linearity assumption of linear regression is not met and we should not use the R^2 statistic to assess the goodness of fit. Polynomial functions shown in Equation 2.5 are linear in the parameters. Therefore, we can use linear regression as long the assumptions for linear regression are met. If this is not the case, we should be cautious with the interpretation of the R^2 statistic.

When using the method of nonlinear least squares, the way in which the unknown parameters in the function are estimated is conceptually the same as it is in linear least squares regression. Parameters are calculated so the total SSE is minimized. However, the major cost of moving to nonlinear least squares regression is the need to use iterative optimization procedures to compute the parameter estimates (Bates & Watts, 2008). With functions that are linear in the parameters, the least squares estimates of the parameters can always be obtained analytically, this is not the case for nonlinear models. The use of iterative procedures requires the user to provide starting values for the unknown parameters before the software can begin the optimization. The starting values must be reasonably close to the as yet unknown parameter estimates or the optimization procedure may not converge. Bad starting values can also cause software to converge to a local minimum rather than the global minimum (Bates & Watts, 2008). In Subsection 2.5.3, we discussed some optimization algorithms that were used by Lydia et al. (2013).

If the nonlinear model can be transformed into a linear model, the user should always try this first and use linear regression instead of nonlinear regression (Hoshmand, 2009). When the assumptions for linear regression hold true for the transformed model, the R^2 statistic can be properly interpreted for the transformed model. However, transforming the model back to its original state means we cannot interpret the R^2 statistic that belongs to the transformed model (Frost, 2014). This is because the underlying assumptions for R^2 are not true for the original nonlinear model. Frost (2014) advocates the use of the Standard Error of Regression (S), also called the standard error of estimate by Hoshmand (2009). This statistic can be used for both linear and nonlinear regression models. We discuss this statistic in Section 2.7.

2.7 Measuring forecast accuracy

The most important criteria for assessing the accuracy of a forecasting model is the model accuracy. In this section we provide a list of accuracy metrics that is commonly used by researchers. In this list, $Y_a(i)$ is the actual value of Y of the ith observation, $Y_e(i)$ is the expected value of Y forecasted by the model, $\overline{Y}_a$ is the mean value of the actual observations of Y, *N* is the total number of observations and *p* is the number of parameters. The list consists of the Relative Error (RE), Mean Absolute Error (MAE), symmetric Mean Absolute Percentage Error (sMAPE), Normalized Mean Absolute Percentage Error (NMAPE), Root Mean Squared Error (RMSE), the Coefficient of Determination (R^2) and lastly, the Standard Error of Regression (S).

$$RE = \left|\frac{Y_e(i)-Y_a(i)}{Y_a(i)}\right| \times 100\% \tag{2.13}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|Y_e(i) - Y_a(i)| \tag{2.14}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|Y_e(i)-Y_a(i)|}{Y_a(i)} \times 100\% \tag{2.15}$$

$$sMAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|Y_e(i)-Y_a(i)|}{(|Y_e(i)|+|Y_a(i)|)/2} \times 100\% \tag{2.16}$$

$$NMAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|Y_e(i)-Y_a(i)|}{max_{i=1}^{N} Y_a(i)} \times 100\% \tag{2.17}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Y_e(i) - Y_a(i))^2} \tag{2.18}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(Y_e(i) - Y_a(i))^2}{\sum_{i=1}^{N}\left(Y_a(i) - \overline{Y}_a\right)^2} \tag{2.19}$$

$$S = \sqrt{\frac{\sum_{i=1}^{N}(Y_e(i) - Y_a(i))^2}{N-p-1}} \tag{2.20}$$

The first metric in the list show the errors for a single observation, so this is not suited to measure the accuracy for a total sample size (Hyndman & Koehler, 2006). MAE is suited for the entire sample size and is very easy to interpret. However, this metric is scale-dependent, meaning that the metric will have higher outcomes as the scale (maximum power output) increases. RMSE has the same disadvantage, since it is scale-dependent as well. RMSE is also more sensitive to outliers than MAE due to its squared error, which led some researchers to recommend against the use of RMSE in accuracy evaluation (Hyndman & Koehler, 2006).

Percentage errors like MAPE have the advantage of being scale-independent (Hyndman & Koehler, 2006). This makes that they are frequently used to compare forecast performances across different data sets. However, MAPE has the disadvantage of being undefined if $Y_a(i) = 0$ for any observation or being extremely skewed if $Y_a(i)$ is close to zero. MAPE also has the disadvantage of putting a heavier penalty on positive errors than on negative errors. To avoid this, the sMAPE can be used. However, according to Hyndman and Koehler (2006) the sMAPE is not as symmetrical as their name suggests. For the same value of $Y_a(i)$, sMAPE gives a heavier penalty when forecasts are low compared to when forecasts are high. NMAPE has the advantage of showing the mean percentage compared to the maximum actual value, this makes the metric desirable since it is simple and easy to interpret (Hyndman & Koehler, 2006).

The coefficient of determination, R^2, expresses the fraction of variance that can be explained by the model. R^2 is a statistic that gives information about the goodness of fit of a model. For example, in regression, R^2 is used to indicate how well the regression line fits the data, a value of 1 indicates a perfect fit, which means that 100% of the variance can be explained by the model. However, we should note that R^2 can only be used with linear regression models. The assumptions of the linear regression model should be met, otherwise the interpretation of R^2 can lead to misleading conclusions. R^2 should not be used for nonlinear regression models (Frost, 2014; Spiess & Neumeyer, 2010).

Lastly, we discuss the standard error of regression, *S*, which is also called standard error of estimate (Hoshmand, 2009). In contrast to R^2, *S* can be used for both linear and nonlinear regression. According to Frost (2017), the standard error of regression is superior to the coefficient of determination for both linear and nonlinear regression. The *S* statistic is an absolute measure of the typical distance that the data points fall from the regression model. *S* is measured in the units of the dependent variable. The standard error of regression is interpreted like any other standard deviation. It means that if the dependent variable is distributed normally around the regression plane, approximately 68% of the values of the dependent variable fall within a range of ± S (Hoshmand, 2009). Furthermore, approximately 95% of the values fall within ±2S. This means that if the error terms are normally distributed with a mean of 0, then the statistic *S* can be used to calculate a 95% prediction interval (Frost, 2017). Frost (2017) prefers the standard error of regression over the coefficient of determination, because it is better at evaluating the precision of the predictions.

Model validation

To validate a regression model, the dataset can be split into training data and test data. Training data are used to estimate parameters and the test data are used to evaluate the model accuracy. When calculating the forecast accuracy, always use test data that were not used when computing the forecasts (Hyndman, 2014). If there is a big difference in accuracy between the training data and the test data, then we are probably overfitting the model to the training data.

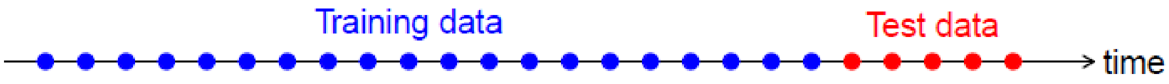

Figure 2.13: A time series divided into training- and test data (Hyndman, 2014).

The size of the test data is typically 20% of the total sample, although this value depends on the sample size and the forecast horizon (Hyndman, 2014). The size of the test set should be at least as large as the forecast horizon.

In case of a small sample size or a short time series, we do not want to split the data since the conclusions we draw from the forecast accuracy measures are not very reliable due to the small data set. To avoid this problem, cross-validation can be used. A lot of types of cross-validation are available, they all have the same underlying idea. The entire dataset is split into training- and test data several times. Each time a different part of the dataset is used as training- and test data, cross-validation combines the measure of fit to derive a more accurate estimate of model performance. If the sample size is large, there is no need to use cross-validation. The dataset can simply be split into training and test data.

2.8 Conclusion

We conducted a literature review to attain information on how to develop a wind turbine power production forecasting model. We thoroughly reviewed literature about the wind resource, wind turbine design, power output of wind turbines, forecasting of wind power production and how we should measure forecast accuracy. We conclude the following:

- Wind speed variation is very different depending on the timescale and location.
- Hourly wind speed variation in winter months is larger than during the summer. This is probably due to temperature decreases during winter months. Wind speed variations during the year can be characterized with a Weibull distribution, which can be used to estimate total annual production for wind turbines.
- Short-term (10 minutes or less) wind speed variations include turbulence and gusts. Wind gusts are of importance when considering peak load calculations. Turbulence is a complex process, which is hard to measure or predict. The effect of turbulence on the power production is, as a rule, not severe.
- Most modern horizontal axis wind turbines have three blades, TSR control, blade pitch control and have a yawing mechanism.
- The theoretical power output of a wind turbine is affected by the air density, swept area of the rotor, power coefficient and the wind speed at hub height. Air density is governed by the air pressure, temperature and altitude at a certain location. The power coefficient is a variable that is different depending on wind turbine design and wind speed, wind turbines with TSR control typically have a greater power coefficient at high wind speeds.
- In practice, the power output is affected by site-related influences like surrounding buildings, trees or wind turbines. Also, the presence of mountains or hills has an influence on the power output. Therefore, wind direction can have a big influence on the power output depending on

the site. Maintenance has an influence as well, especially when turbines have been operational for many years bad maintenance (e.g. soiled rotor blades) has a negative impact on the power output.

- Each wind turbine has a unique power curve; the theoretical power curve provided by the manufacturer is usually not very accurate.
- The forecasting approach should be chosen according to the forecast horizon and the available data. For a forecasting horizon of up to 6 hours, autoregressive approaches with online data from SCADA systems are recommended.
- For a forecasting horizon of up to 48 hours, meteorological predictions become essential and weather forecasts from NWP models should be used in combination with causal models such as regression or learning approaches such as neural networks. Regression has the advantage of being able to use multiple predictors and is desirable due to its simplicity.
- Parametric WTPC modeling is a form of simple regression, this can either be linear or nonlinear regression. WTPC modeling is a key pillar in wind power forecasting, since it models the relationship between wind speed and wind power output. More predictors can be added to the model, which can improve the accuracy of the prediction.
- MAE and RMSE are good scale-dependent error indicators; RMSE punishes larger errors more severely than MAE. NMAPE can be used as a scale-independent error indicator. The standard error of regression is very similar to the RMSE, the difference being a punishment of the number of parameters.
- When interpreting the R^2 of linear regression models, we should check whether the underlying assumptions are met. The standard error of regression can be used for linear and nonlinear regression models. Nonlinear regression models do not rely on underlying assumptions.

3. Current Situation

In this chapter we answer Sub-question 3. In Section 3.1 we look at the data that are available. In Section 3.2, we select the producers that are placed in the project scope. These producers are used to test the accuracy of the forecasting models. After this, we analyze the historical weather data for three KNMI weather stations in Section 3.3. The production data are compared with the historical weather measurements in Section 3.4. In Section 3.5, we assess the accuracy of the weather forecasts and we check whether the forecasts are biased. Lastly, in Section 3.6 we collaborate how and why the data are cleaned before ending with a conclusion.

3.1 Available data

The data available determines the method to be used. Therefore, we evaluate which data are available with regard to weather, turbine characteristics and energy production. This concerns historical production data and historical production forecasts from Company X, for weather data we also have forecasts and historical forecasts (hindcasts).

3.1.1 Weather data

To get historical weather data from the Netherlands, DVEP uses data from 15 weather stations located throughout the country. These are Koninklijk Nederlands Meteorologisch Instituut (KNMI) weather stations, the following locations are used by DVEP:

- Amsterdam.
- Beek.
- Berkhout.
- De Bilt.
- De Kooy.
- Deelen.
- Ell.
- Gilze.
- Groningen.
- Hoogeveen.
- Hupsel.
- Lelystad.
- Marknesse.
- Twenthe.
- Vlissingen.

Each location, based on the coordinates given by KNMI, is indicated in Figure 3.1 with a yellow marker (KNMI, 2000). For each location the hourly average wind speed (m/s), wind direction (°), temperature (°C), sum of rainfall (mm) and sum of radiation (J/cm^2) are available. To measure the wind speed, an anemometer is used, at each location wind speed is measured at a height of 10 m. Wind speed is measured over periods of ten minutes with a sample each second. For wind direction, the last ten minutes of each hour are used to calculate the average wind direction. Here 360° represents the north, 90° the east, 180° the south and 270° represents the west. When the wind speed is zero, there is no wind direction, which means there is no value available. The average temperature is measured at a height of 1.5 m and calculated the same way as wind speed and direction. For rainfall and radiation the hourly sum is used. (KNMI, 2014)

Figure 3.1: KNMI weather stations used by DVEP located in the Netherlands (KNMI, 2000).

For the weather forecasts, a supercomputer is used that calculates hundreds of alternate predictions using four NWP models, combining NWP models results in a forecast that is as reliable as possible. However, weather remains unpredictable, especially as the forecast horizon increases. Forecasts are available for each hour of the current day and a day ahead. Here predictions for average wind speed, direction, temperature and the sum of rainfall and radiation are given for each hour. Atmospheric pressure is not available with the forecast. The forecast locations are the KNMI weather stations listed above, the NWP models are downscaled to the same height as the historical measurements from the stations, which is 10 meters.

To be able to compare the accuracy of Company X and the new models, we need to use historical forecasts (hindcasts). Company X used day ahead weather forecast from 9:00 to predict the energy production, we have to use day ahead weather forecasts from 9:00 as well to predict the energy production with the new model. For example, if we want to compare the day ahead forecast accuracy of Company X on and the new model on 2-7-2017, we have to use historical weather forecasts from 9:00 on 1-7-2017. Luckily, the weather hindcasts are available for each day at 9:00. This means that for each day we know what the forecast was at 9:00 the previous day, we need the hindcasts from 9:00 because they were used with the auction. Unfortunately, we only have hindcasts for the second half of 2017.

3.1.2 Wind turbine characteristics

Each wind turbine or wind farm is connected to a connection point which has a unique EAN code. For each connection point we know the number of turbines and the total rated power. Most connections only have a single turbine, some have multiple (wind farms). Also the location and hub height are known. For most turbines, the brand and model are known as well. We have information about the length of the rotor blade and the swept area. However, we do not know whether the turbine is pitch- or stall controlled or whether the turbine has tip-speed ratio control. Whether the turbine has a yawing system is also not known, however, most modern turbines do. All turbines are HAWTs and have three blades. All turbines under DVEPs supervision are located onshore in the Netherlands, the province Flevoland hosts most of them. Friesland is second when it comes to the number of turbines and hosts a lot as well. The provinces Groningen, Noord-Holland, Zuid-Holland, Zeeland, Drenthe, Gelderland and Limburg are the remaining locations. The majority of the sites are located near coastal areas for efficiency purposes.

3.1.3 Production data

Each connection point that is connected to the electricity grid has a unique code, the EAN code. Wind turbines are connected to the grid with an EAN code. All turbines in a wind farm are connected to the same EAN. For each EAN, the number of turbines and the summed production in kilowatts per 15 minutes is known, this can be used to calculate the summed production per hour. If a producer has a wind farm, we do not know the production per wind turbine. We only have the summed production data per 15 minutes.

Production data are available from the start of the contract for each producer. Producers are only placed in the scope if they have a contract between 1-1-2015 and 1-1-2019. The day ahead predicted power production from Company X is available also.

In Table 3.1 we illustrate an example of what a dataset looks like when we connect historical weather measurements and historical production measurements.

Date	Time	Average wind speed (m/s)	Average wind direction (°)	Average temperature (°C)	Production
1-1-2015	00:00 – 01:00	…	…	…	…
1-1-2015	01:00 – 02:00	…	…	…	…
…	…	…	…	…	…
…	…	…	…	…	…
31-12-2016	22:00 – 23:00	…	…	…	…
31-12-2016	23:00 – 00:00	…	…	…	…

Table 3.1: Example of dataset with historical weather measurements and historical production measurements.

There are some problems with the data that should be addressed. Wind turbines have storm detection, which means the wind turbine shuts down in case of extreme wind speeds. This results in low production values when there is a very strong wind. The same goes for ice detection, at temperatures around the freezing point, icing of the rotor can occur. Ice detection shuts down the turbine when icing occurs, this leads to low production values. It is hard to see when storm or ice detection kicked in. Besides storm and ice detection, wind turbines are also subject to failures and maintenance. In case of a failure, the turbine is not operational until it is repaired, which can take days or weeks depending on what is broken. Maintenance usually means a turbine is not operational for several hours. Failures and maintenance are not logged, so we do not know when or whether they occurred. We address these problems in Section 3.6 by cleaning the data.

Another problem with the production data is caused by multiple sources of production. Some producers have solar panels and wind turbines. When these are connected to the same EAN, we cannot distinguish between solar and wind production. However, this problem can easily be avoided by removing these producers from the research scope.

3.1.4 Overview of available data

In this subsection we provide an overview of available data. This helps us in choosing our approach in Chapter 4.

Data	2015-2017	Second half of 2017
Historical weather measurements from KNMI stations	X	X
Historical production measurements from producer	X	X
Weather hindcasts		X
Company X production hindcasts	X	X
Weather hindcasts used by Company X		

Table 3.2: Overview of available data per period of time.

Table 3.1 illustrates the historical weather and production measurements are available for each hour. The weather hindcasts and production hindcasts from Company X are available for each hour as well. However, Table 3.2 illustrates that weather hindcasts are only available for the second half of 2017. We do not have data for the weather hindcasts that are used by Company X.

For each producer we have the following information:

- Number of turbines.
- Rated power per turbine.
- Hub height and location of turbine.
- Brand and model type.
- Length of rotor blade.
- Swept area of rotor.

3.2 Producer and weather station selection

DVEP has more than 200 wind power producers. However, for most producers we do not have sufficient data, since their contract is either too short or in the wrong period of time. When excluding the producers based on their contract period, approximately 50 producers remain.

Unfortunately, we do not have wind measurements at hub height and at the location of the turbines. Therefore, the KNMI stations are used for the historical weather measurements. To make sure the historical measurements are as accurate as possible, we sort the producers according to their Euclidean distance to the closest KNMI station. Out of locations that are closest to a certain KNMI station we make a selection of 10 producers. The selection is based on the rated power and the geographical location as well. Ideally, we want producers from KNMI weather stations dispersed throughout the country so we can compare several locations. Also, the sum of the rated power should be at least 10% of the total rated power of all producers to make sure the selection is representative. Factoring in all these choices, the following selection was made:

Producer	City	Number of turbines	Rated power per turbine in kW	Total rated power in kW	Nearest KNMI weather station	Distance in kilometers	Type Turbine	Swept area in m²
1	Nieuw- en Sint Joosland	2	900	1800	Vlissingen	6	Vestas NM900	2124
2	Aalten	8	2000	16000	Hupsel	14	Enercon E-82	5281
3	Lelystad	2	3000	6000	Lelystad	10	Enercon E115	10515
4	Lelystad	2	1000	2000	Lelystad	6	NEG Micon NM1000	2827
5	Lelystad	1	900	900	Lelystad	6	NEG Micon NM52	2140
6	Zeewolde	1	900	900	Lelystad	9	NEG Micon NM900	2124
7	Zeewolde	1	1000	1000	Lelystad	8	NEG Micon NM1000	2827
8	Swifterbant	6	1650	9900	Lelystad	14	Vestas V66	3421
9	Zeewolde	1	850	850	Lelystad	7	Vestas V52	2124
10	Nieuw- en Sint Joosland	7	900	6300	Vlissingen	6	NEG Micon NM900	2124

Table 3.3: Selection of 10 producers used for analysis.

We focus on three KNMI weather stations, namely those located in Lelystad, Vlissingen and Hupsel. The total rated power of this selection is 45.65 MW, which is approximately 12% of the rated power of all DVEP producers. A mix of wind farms (multiple turbines) and single wind turbines was chosen so these can be compared. All turbines in the selection are located onshore. DVEP feels that this selection is representative for the entire portfolio.

3.3 Weather data

In this section, we look into the wind speed and wind direction at Lelystad, Vlissingen and Hupsel in 2015 and 2016. The weather data are gathered from KNMI weather stations and contains average hourly values. The values for wind speed are rounded by KNMI to bins of 0.5 m/s (e.g. 0.5, 1, 1.5, 2, etc.), the values for wind direction rounded to bins of 5° (e.g. 5°, 10°, 15°, etc.) with a maximum of 360°.

3.3.1 Wind speed

We expect wind speeds to be different at each location. For example, locations near the coast probably have a higher average wind speed than inland locations. The wind speed in 2015 and 2016 at the three KNMI stations is distributed as follows:

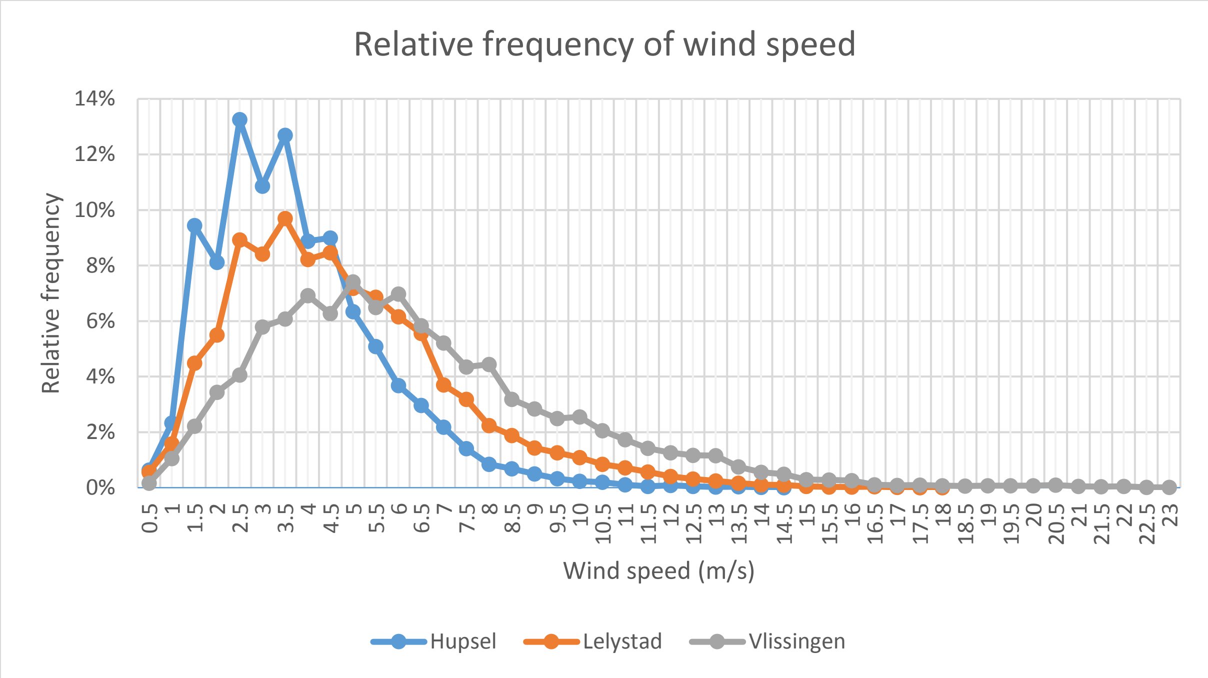

Figure 3.2: Wind speed distribution in Hupsel, Lelystad and Vlissingen in 2015-2016.

The three distributions all resemble a Weibull distribution, which they should according to the literature. We are not interested in the values of the shape parameter *k* for each location, since we do not want to estimate the annual production of a producer. We simply look at the relative frequency of wind speed at different locations to see whether it meets our expectations.

The distributions do differ per location, especially at the higher wind speeds. What stands out is that Vlissingen's distribution has that longest tail and is placed more to the right, which indicates a higher wind speed more often. Wind speed distribution in Lelystad has a shorter tail than Vlissingen and is placed more to the left. However, Hupsel has the shortest tail and is distributed more strongly towards the lower wind speeds than the other two locations. When considering the geographic locations of the three weather stations, this is logical. Vlissingen is located in Zeeland near the coast, which generally has a strong wind. Lelystad is located next to the IJsselmeer, here the wind can prevail as well, although to a smaller extent than near the North sea coast. Hupsel is located in Gelderland which is more inland, therefore it is only logical that the wind is generally weaker than at the other locations.

3.3.2 Wind direction

For the wind direction, KNMI shows average wind direction per hour in degrees from 0° to 360°. Here 0° and 360° indicates that the wind came from the north, 90° the east, 180° the south and 270° the west. The distribution of wind direction for Lelystad, Vlissingen and Hupsel are as follows:

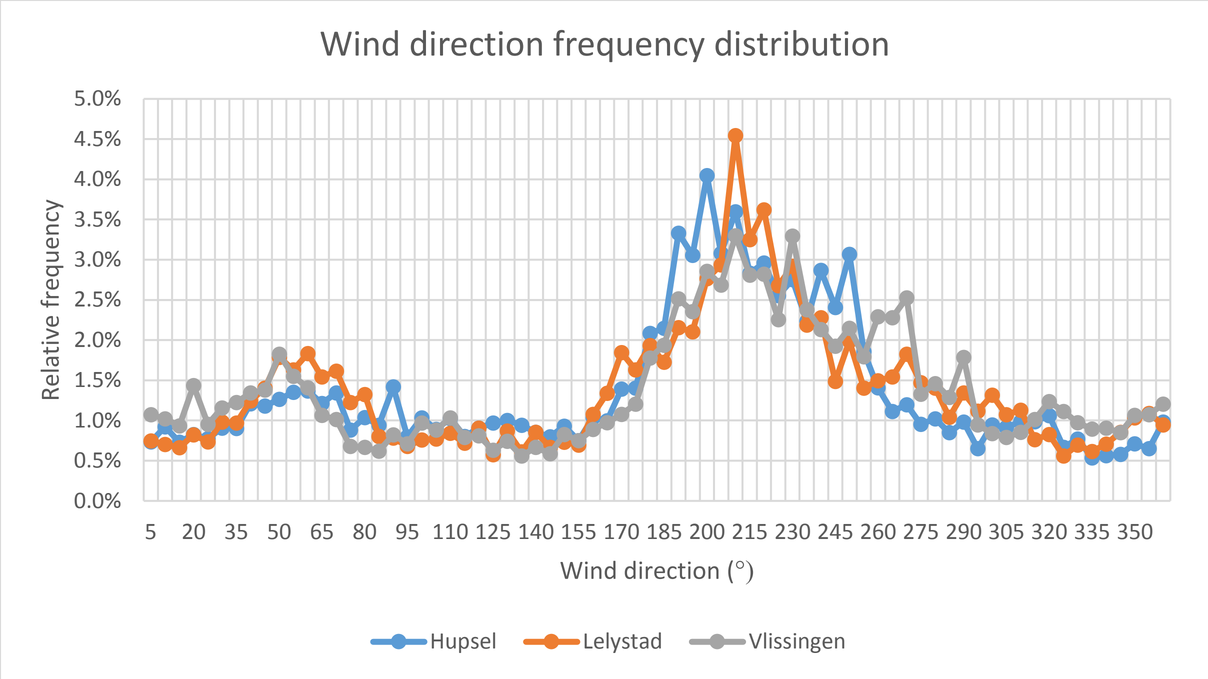

Figure 3.3: Wind direction distribution in Hupsel, Lelystad and Vlissingen in 2015-2016.

When comparing the distribution of wind direction for the three locations, we see the same pattern emerging. The biggest peak is between 180° and 250° for all locations. Which means the wind most often comes from the southwest. For Vlissingen and Lelystad this was expected, since they are located near the North sea and the IJsselmeer. Hupsel on the other hand, is located in the east of the Netherlands, but still shows the same pattern.

3.3.3 Wind speed per direction

Now that we know the wind the prevailing wind direction for each location, we want to know whether there is a particular wind direction that has stronger winds. For this purpose, we look at the average wind speed per wind direction at Lelystad, Vlissingen and Hupsel.

The average wind speed in Vlissingen is highest out of the three locations, which was expected due to its location. Hupsel has the lowest average wind speed across all wind directions due to its inland location. As far as a pattern goes, we see that the average wind speed is highest where the frequency is highest as well. This pattern is clearly visible for Vlissingen and Lelystad, wind is generally stronger when it comes from the south-west. For Hupsel the pattern is also present, however, the difference with other wind directions is smaller.

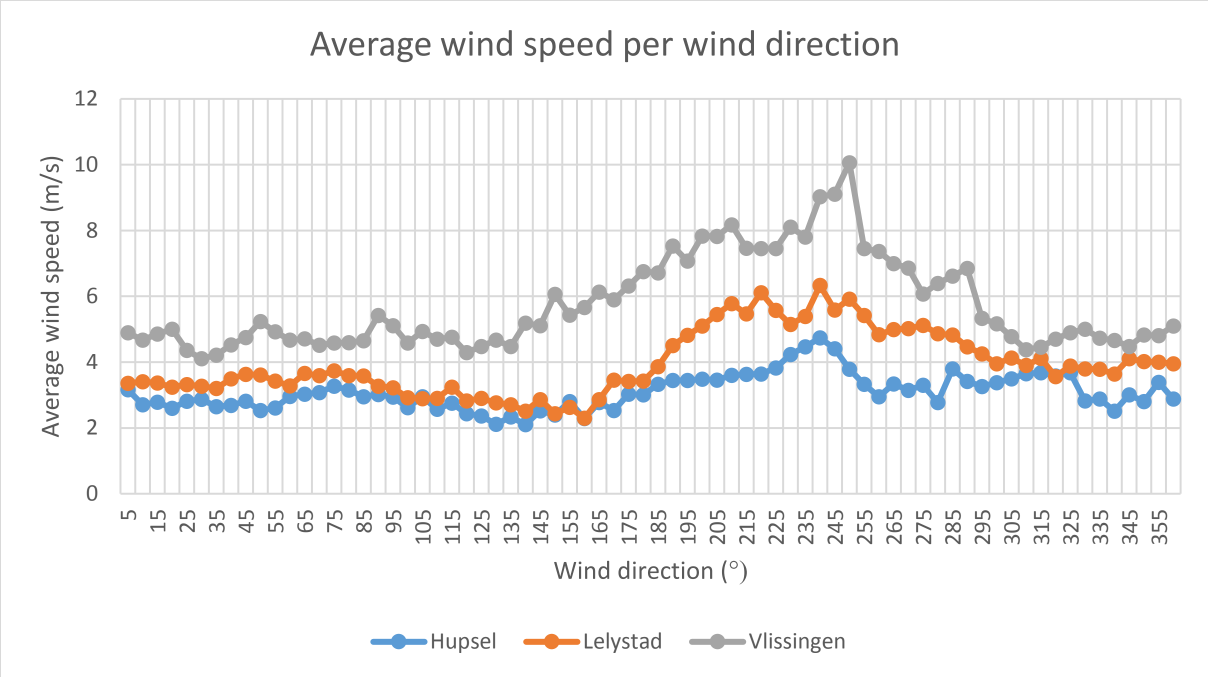

Figure 3.4: Average wind speed per wind direction in Hupsel, Lelystad and Vlissingen in 2015-2016.

3.4 Production data

In this section, we examine the production data of 2015 and 2016 for three producers, one close to Lelystad, Vlissingen and Hupsel. For each producer, we have production data for each quarter of the hour of the day. We choose Producers 1, 2, and 3 from Table 3.3 for the comparison of production data. The production data are transformed into hourly production data by adding the production per quarter. This enables us to compare the weather data with the production data. First, we analyze the production data with respect to wind speed. After this, we look into the production per wind direction. Lastly, we examine the relationship between temperature and production.

3.4.1 Wind speed versus production

At each location we look at the minimum, maximum and average production as percentage of rated power for each wind speed to see if anything stands out in the dataset. Table 3.4 illustrates the minimum, average and maximum production as a percentage of the rated power for each wind speed of Producer 3. Producer 3 is located 10 km from Lelystad and has two wind turbines of 3000 kW, which results in a total rated power of 6 MW. The tables for Producers 1 and 2 show a similar pattern, these tables can be found in Appendix B.

Wind speed (m/s)	Minimum	Average	Maximum
0	0%	4%	34%
0.5	0%	3%	40%
1	0%	5%	63%
1.5	0%	10%	92%
2	0%	15%	100%
2.5	0%	19%	100%
3	-1%	24%	101%
3.5	-1%	28%	102%
4	-1%	35%	102%
4.5	-1%	38%	100%
5	-1%	45%	102%
5.5	-1%	50%	102%
6	-1%	56%	102%
6.5	0%	61%	102%
7	0%	64%	102%
7.5	0%	66%	102%
8	0%	66%	102%
8.5	0%	72%	102%
9	0%	72%	102%
9.5	0%	72%	102%
10	0%	73%	102%
10.5	30%	74%	102%
11	0%	76%	102%
11.5	0%	67%	102%
12	-1%	73%	102%
12.5	0%	76%	102%
13	45%	76%	102%
13.5	46%	72%	102%
14	0%	66%	101%
14.5	0%	45%	100%
15	0%	30%	50%
15.5	0%	41%	100%
16	0%	40%	51%
17	0%	33%	99%
18	96%	96%	96%
19	95%	95%	95%

Table 3.4: Minimum, average and maximum production as a percentage of the rated power for Producer 3 using raw data.

The production is negative sometimes, this happens around the cut-in speed most often. This is due to the fact that wind turbines consume power when they are starting up. When the wind speed crosses the cut-in speed threshold, but then drops again, the production can be negative during an hour. At higher wind speeds the production can be negative as well. This happens when the wind speed is higher than the cut-out speed, but temporarily drops so the turbine is activated shortly, which costs power.

Some peculiar values are the large maxima at the wind speeds around 0 m/s. A production of 34% at an average wind speed of 0 m/s is impossible. This means there is an error in the data, this is probably due to the fact that the average wind speed at hub height was different than at the KNMI weather station. This would explain the large maxima at wind speeds 0,5 and 1 m/s as well. The production often exceeds the rated power for Producer 3. However, this is not a problem since the rated power is just an indication of the production at full capacity.

At the mid-ranges of wind speed, around 8 m/s, the minima are often 0%. This can have several reasons. Firstly, wind turbines are scheduled for maintenance once in a while. This means they are not operational and the production is zero. Usually, maintenance only takes a couple of hours. Secondly, wind turbines are subject to random failures. Some failures can be fixed quickly and the turbine is operational in a couple of hours. However, some failures take weeks to fix. Lastly, when temperatures drop below 0 °C, icing of the rotors can occur which forces the turbines to be shut down. Unfortunately, we do not have historical data about failures, maintenance or icing. Therefore, we do not know when this occurred. Also, at wind farms we do not know how many turbines had failures, maintenance or icing at a certain moment. This is due to the fact that multiple wind turbines are connected to one EAN. When maintenance is scheduled at a wind farm, usually only one turbine is scheduled at a time. Failures and icing have the same problem, we do not know how many turbines are experiencing problems at a certain time. For example, 1 turbine has downtime while 3 are operational.

At wind speeds of 18 or 19 m/s, the minimum, average and maximum are the same since this only happened once during the two years of training data. Some average wind speed values, for example 16.5, are missing because they never occurred in 2015 and 2016. This is illustrated in a scatterplot in Figure 3.5:

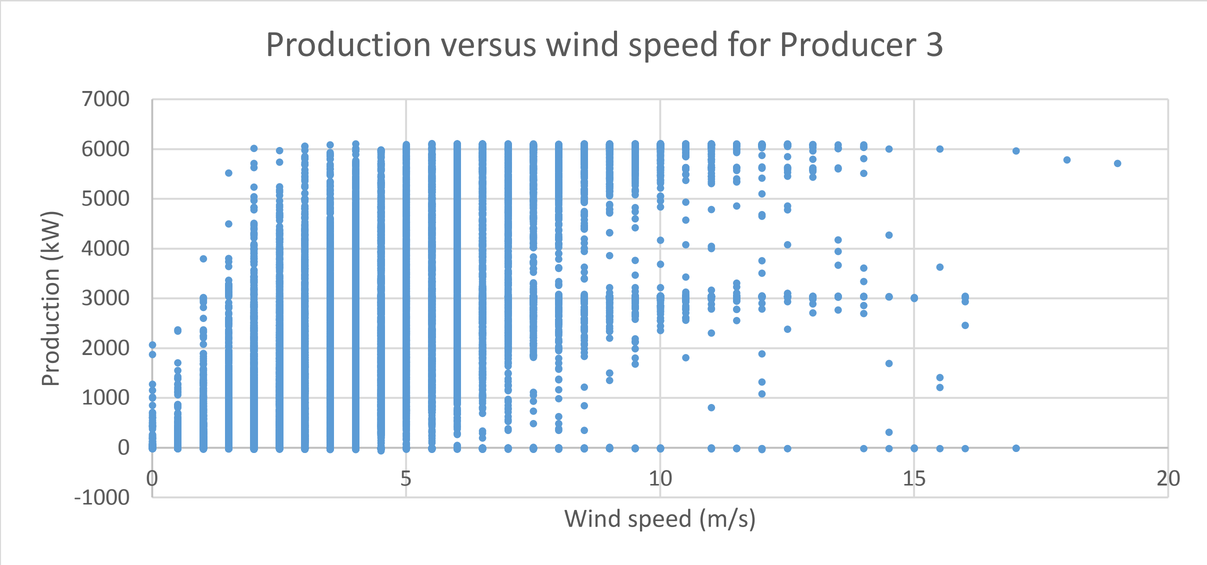

Figure 3.5: Scatterplot of production versus wind speed in 2015-2016 for Producer 3 with 2 turbines with 6 MW total rated power.

For wind speed of 10 m/s or higher, the scatterplot is more dense around 3000 kW than at 2000 or 1000 kW. This is probably because one turbine is operating at full capacity, while the other has downtime due to maintenance, failure, icing or windstorms.

When wind speeds drop below cut-in speed or rise above cut-out speed, turbines shut down. We do not know when or whether this happened during an hour. This can also occur several times during an hour. When wind speed reaches operational range again, the turbine starts again which costs power. The timing of these events is important, when this happens in the last minute of the hour this has little impact. However, when a turbine produces during the first 10 minutes and is shut down during the rest of the hour, summed production can be very low while the average wind speed during the hour was high.

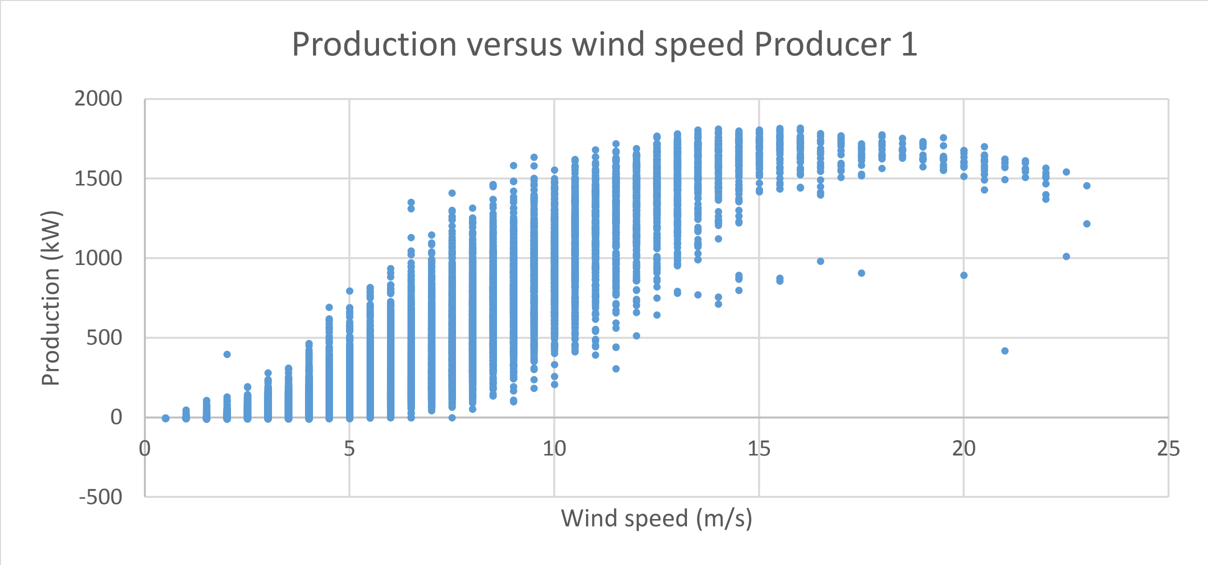

Figure 3.6: Scatterplot of production versus wind speed in 2015-2016 for Producer 1 with 2 turbines with 1.8 MW total rated power.

Figure 3.6 shows a different pattern. During wind speeds of 10 m/s or higher, both turbines of Producer 1 have not been shut down simultaneously for longer than an hour during 2015-2016. This was not the case for Producer 3. We can also see that Producer 1 has a more reliable pattern for almost all wind speeds, the variance is smaller across the entire wind speed spectrum. This shows that wind turbine power curves can be very different depending on factors like location, type and maintenance. We refer to Appendix B for the scatterplot of production versus wind speed for Producer 2.

3.4.2 Wind direction versus production

For the production per wind direction, we check whether there are some wind directions that have a consistently poor performance. We do this by comparing the average wind speed per wind direction with the average production per wind direction. If there are no surrounding obstacles that block the wind, we would expect to see the same pattern for the production per wind direction as for the wind speed per wind direction. We assume that all turbines have a yawing system, so a poor performance in a certain wind direction should be caused by obstacles or the wake effect. The wake effect causes the air flow to be disturbed for turbines inside a wind farm, which lowers the overall production of the wind farm. If turbines of a wind farm are positioned in the form of a line, we should see reduced production in opposite directions, for example, 90° and 270°. Unfortunately, we have no information about the positioning of the turbines.

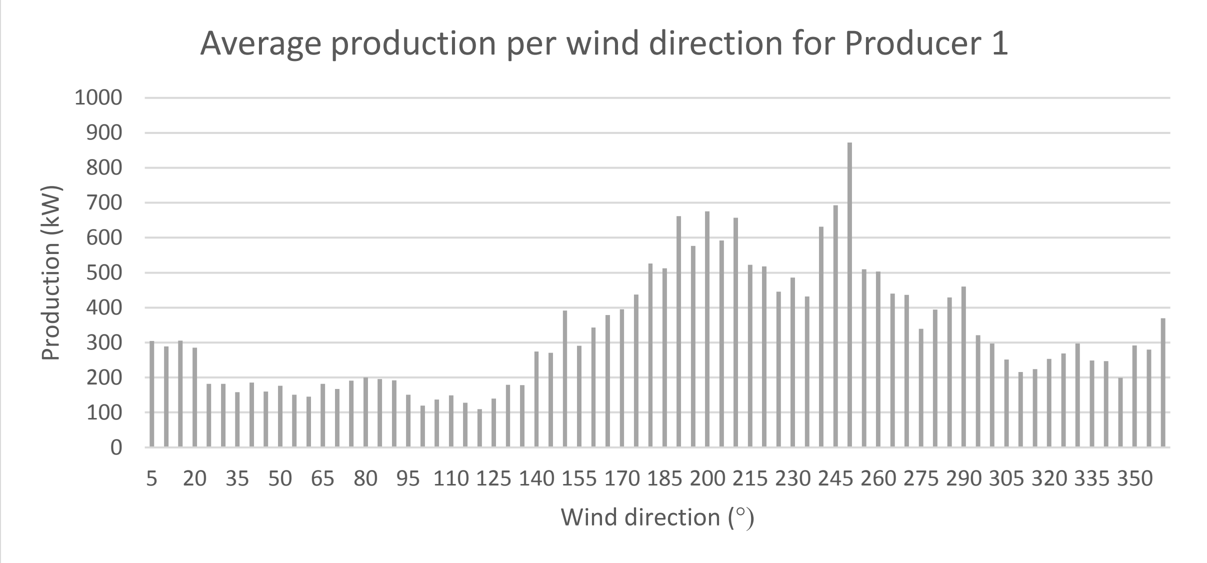

Figure 3.7: Average production per wind direction in 2015-2016 for Producer 1 with 1.8 MW total rated power.

When comparing Figure 3.7 with the pattern from Vlissingen from Figure 3.4, we see roughly the same pattern occurring. The theoretical cubic relationship between wind speed and production would suggest that peaks in average wind speed would result in magnified peaks in production. At wind directions around 250° we can see a peak in production and wind speed, the peak for production is magnified, which is supported by the theoretical relationship. For Producer 1, there is no reason to believe there is an obstacle in the vicinity of the turbines. Around 120°, we see a depression in average production. However, this depression is accompanied by a low average wind speed.

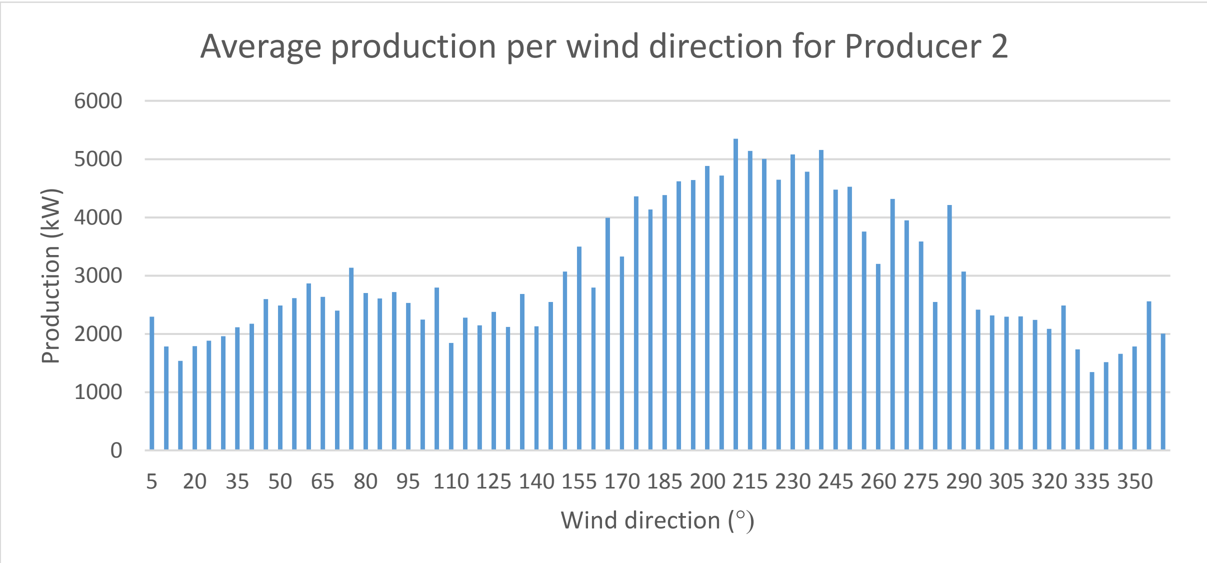

Figure 3.8: Average production per wind direction in 2015-2016 for Producer 2 with 16 MW total rated power.

Figure 3.8 shows a similar pattern as the average wind speed per wind direction for Hupsel in Figure 3.4. Around the wind direction of 120°, there is a low average production while average wind speed is not lower around this direction. Around 300-330°, we see the same decrease in production. This could also be due to an obstacle at the site, since this is a wind farm of 8 turbines, the wake effect would be a likely cause. The graph for Producer 3 shows a depression in production at a wind direction 160°, a little drop in wind speed at this direction would not account for such a big drop in production. However, if this drop was caused by an obstacle, the surrounding wind directions 155° and 165° would be affected as well. For all other directions, the wind speed and production patterns show similarities for Producer 3, so there is no reason to believe the production of Producer 3 suffers from obstacles. The graph for Producer 3 can be found in Appendix C.

Figures 3.7, 3.8 and C.1 indicate that wind direction can have an effect on the production of wind turbines. For each producer this effect is different, because each location is different. Production patterns of Producers 1 and 3 show no signs of obstacles reducing production. Producer 2, however, does have some wind directions with lower production than would be expected.

3.4.3 Temperature versus production

Temperature is related to the air density, this relationship can be expressed using the ideal gas law:

$$\rho = \frac{p}{R_{dry}T} \tag{3.1}$$

where:

ρ = air density in kg/m³

p = atmospheric pressure in Pascal

T = absolute temperature in Kelvin

R_{dry} = specific gas constant for dry air in J/(kg·K) (=287,058)

As the temperature decreases, the air density rises. When air density rises, the production of turbines should rise according to the theoretical relationship introduced in Equation 2.1. Unfortunately, we do not have data for the atmospheric pressure, so we cannot calculate the air density. However, the temperature can be used as a proxy, since it is clearly connected to air density. In Appendix I, the correlation matrices for Producers 1 ,2 and 3 are illustrated. For Producers 1 and 2, the correlation between temperature and production is -0,139 and -0,104 respectively. Producer 3 has a correlation of 0,025, all three correlations are significant at a 95% confidence level. According to the theoretical relationship in Equation 3.1, the correlation should be negative, which is the case for Producers 1 and 2. In addition to correlations, we also look at the average production plotted versus temperature.

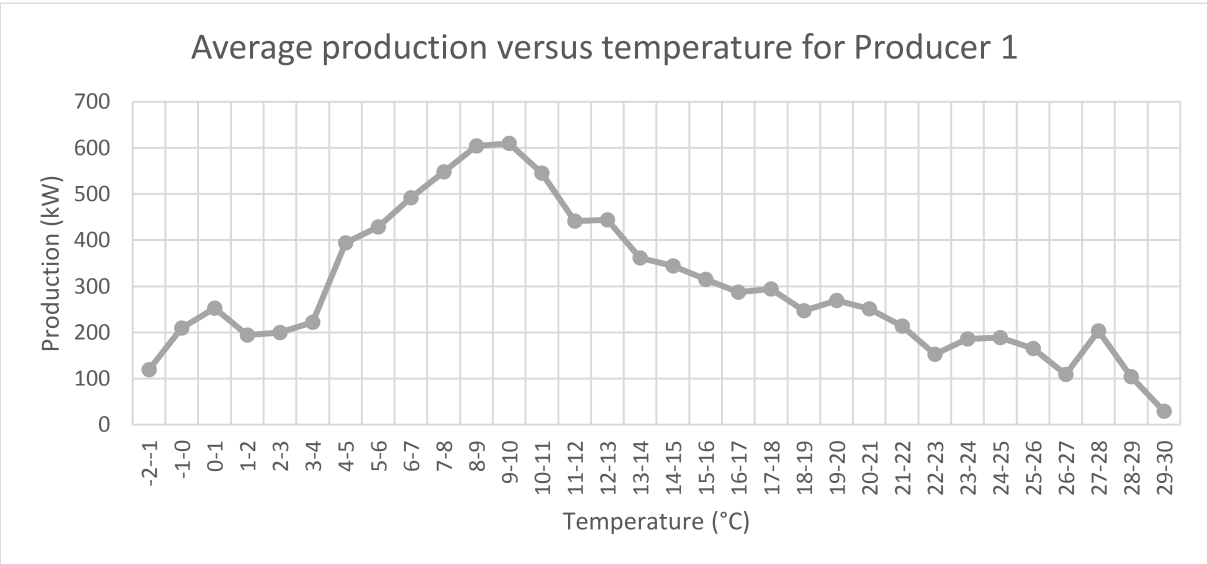

Figure 3.9: Average production versus temperature in 2015-2016 for Producer 1 with 1.8 MW total rated power.

To see whether temperature has an influence on the production we should see whether there is a relationship between temperate and wind speed. Therefore, we look at the graph of average wind speed versus temperature.

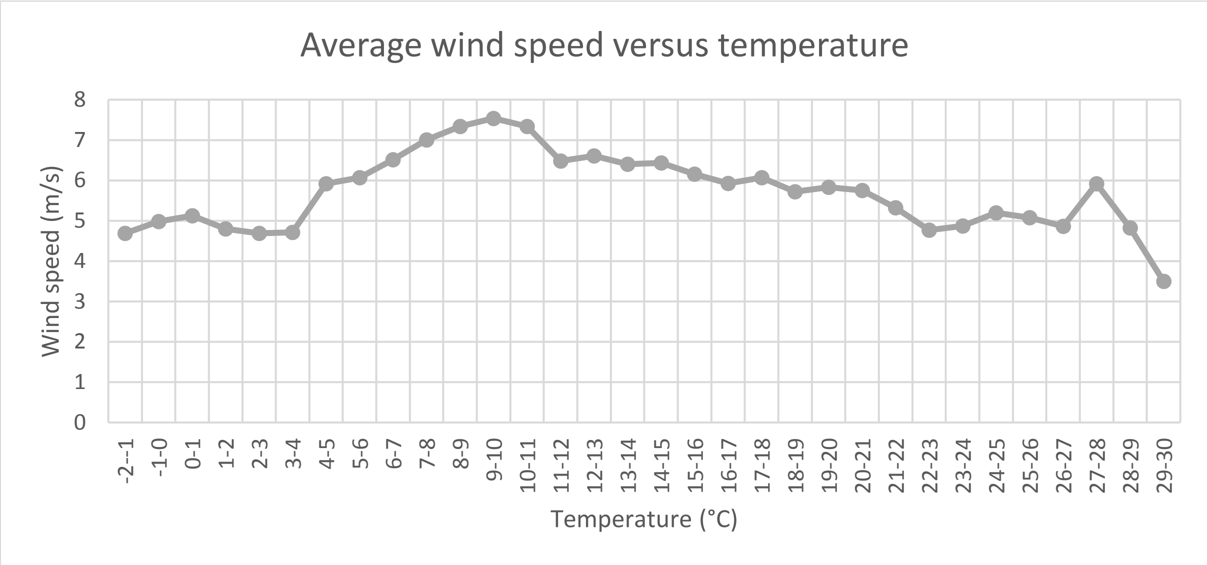

Figure 3.10: Average wind speed versus temperature in 2015-2016 in Vlissingen.

The temperatures with the highest average production are also the temperatures with the highest average wind speed. At temperatures around or below the freezing point we see a decrease in production, the average wind speed is lower at these temperatures as well. However, this would not account for such a decrease in production. The most likely cause of this decrease is the combination of low wind speeds and icing. During high temperatures, the wind speed decreases a little, while production decreases far more. This effect is supported by the theoretical relationship described in Equation 3.1. Producers 2 and 3 show the same pattern, these graphs can be found in Appendix D.

3.5 Weather forecast accuracy

In this section, we look at the weather hindcast data. Hindcasts are historical forecast, this means they are forecasts which were made in the past. We only have hindcast data from the last 6 months of 2017. We look at day ahead hindcasts from 9:00 in the morning, these historical forecasts have a forecast horizon of 15-38 hours ahead. We use this horizon, since DVEP uses the same forecasts for the auction. We want to see whether there is a systematic error (bias) in the weather hindcasts. Therefore, we compare the hindcasts with the historical weather measurements. We do this for Vlissingen, Hupsel and Lelystad. For each location, we evaluate the forecast error in wind speed, wind direction and temperature per hour. These measurements are chosen, because in Section 3.4 we could see they affected the production. The forecast error is calculated by subtracting the actual measurement from the forecast value, so a negative error means the measurement was greater than the forecast. We use statistical tests to see whether the average forecast error is significantly different from zero. We do not test the average error for each hour separately, since this would multiply the number of tests by 24. Although, we do look at the average forecast error for each hour. Lastly, we look at the forecast accuracy across the forecast horizon for each location. If a bias is present, we can adjust weather forecast values to increase the production forecast accuracy.

3.5.1 Vlissingen

In Section 3.3, we saw that the wind speed in Vlissingen is generally higher than at the other locations. This should make forecasting wind speed more difficult, the absolute values are higher, which probably results in a larger variance. However, maybe the NWP models that are used in the forecasts are better at forecasting coastal sites. We are most interested in the error in wind speed forecasting, since this is the key predictor for wind turbine power production. However, we would like to see whether there are systematic errors in wind direction and temperature forecasting as well.

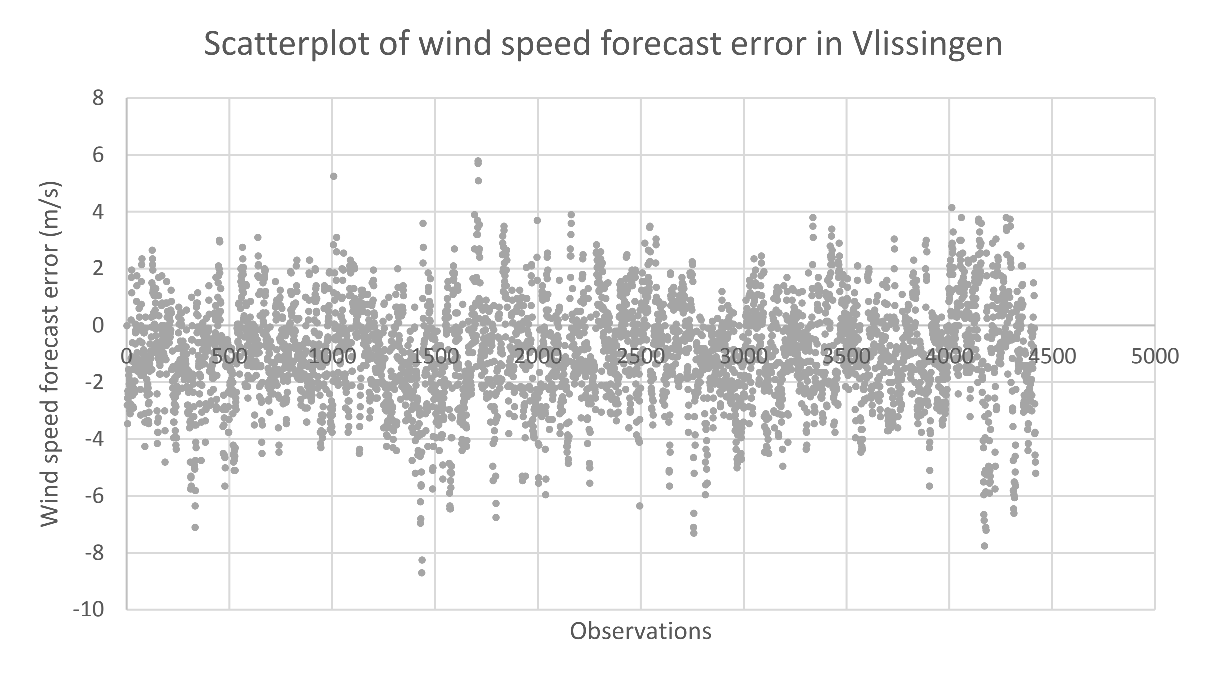

Figure 3.11: Scatterplot of wind speed forecast errors in Vlissingen during the last 6 months of 2017.

Ideally, the forecast errors should have a mean of 0. Looking at Figure 3.11, we can see that the errors are more concentrated below the x-axis, indicating the average error is smaller than zero. The T-test in Appendix J confirms that at a 95% confidence level, the mean of wind speed forecast error in Vlissingen is smaller than 0 (p=0.000). This means there is a negative bias in the wind speed forecasts in Vlissingen. The mean wind speed forecast error in Vlissingen is -0.99 m/s, meaning the forecast is structurally ± 1 m/s too low.

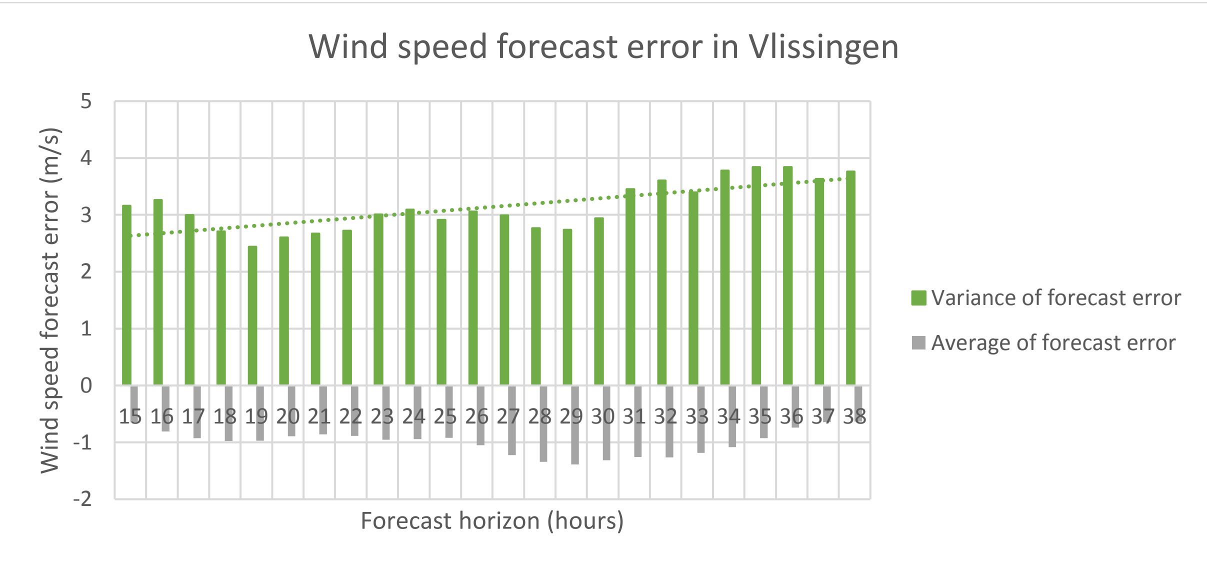

Figure 3.12: Variance and average of wind speed forecast error in Vlissingen during the last 6 months of 2017.

In Figure 3.12, the variance and average forecast error are plotted for the entire forecast horizon. We can see that for the entire forecast horizon, the average error is below zero. When looking at the variance of wind speed forecast errors, we would expect the variance to increase when the forecast horizon increases. The trend line of the variance confirms this.

Now, we look at the wind direction forecast error in Vlissingen. While evaluating the forecast error for wind direction, we have removed observations between 350-360° and between 0-10°. These observations give misleading results, since 360° and 0° both represent the same wind direction. For example, a forecast of 355° and a measurement of 5° would give a forecast error of 350°, while the actual error is only 10°. Removing these misleading values resulted in Figure 3.13.

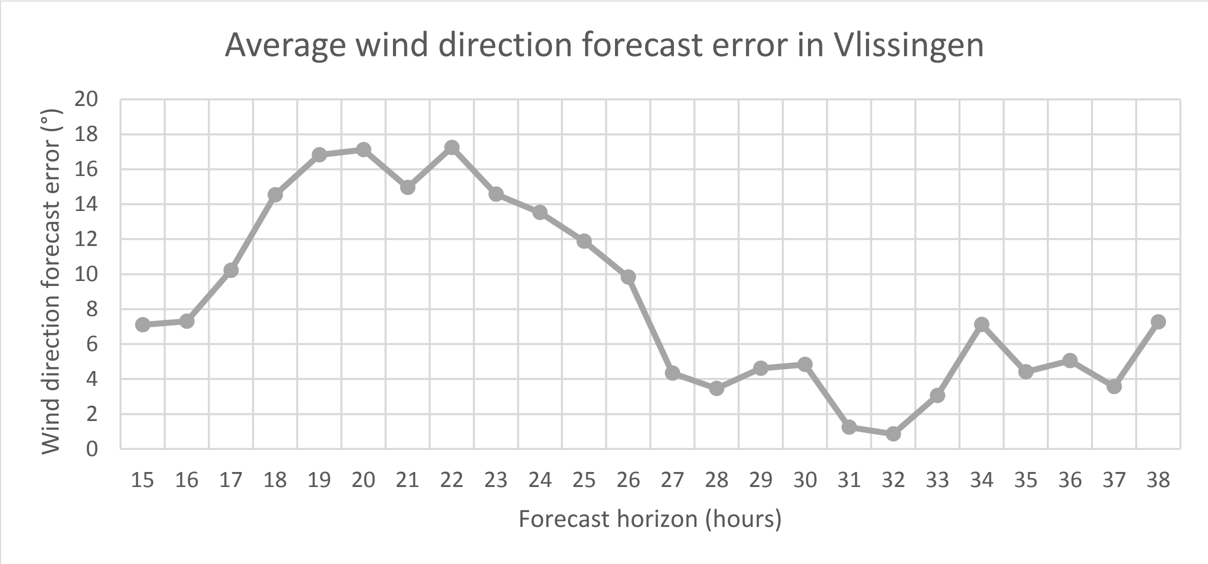

Figure 3.13: Average wind direction forecast error in Vlissingen during the last 6 months of 2017.

We do have reason to believe there is a bias in the wind direction forecast, since the entire forecast horizon has a positive average forecasting error. There is statistical evidence that the mean wind direction forecast error in Vlissingen is greater than 0. At a 95% confidence level, the value of p=0.000, which means there is a probability of at least 95% that the wind direction forecast error in Vlissingen is greater than 0. The average error is 7.87°, meaning the forecast is systematically too high. We have no idea why the wind direction forecast error is smaller as the forecast horizon increases. We would expect the error to increase as the horizon increases.

The average forecast error for temperature in Vlissingen is illustrated in Figure 3.14.

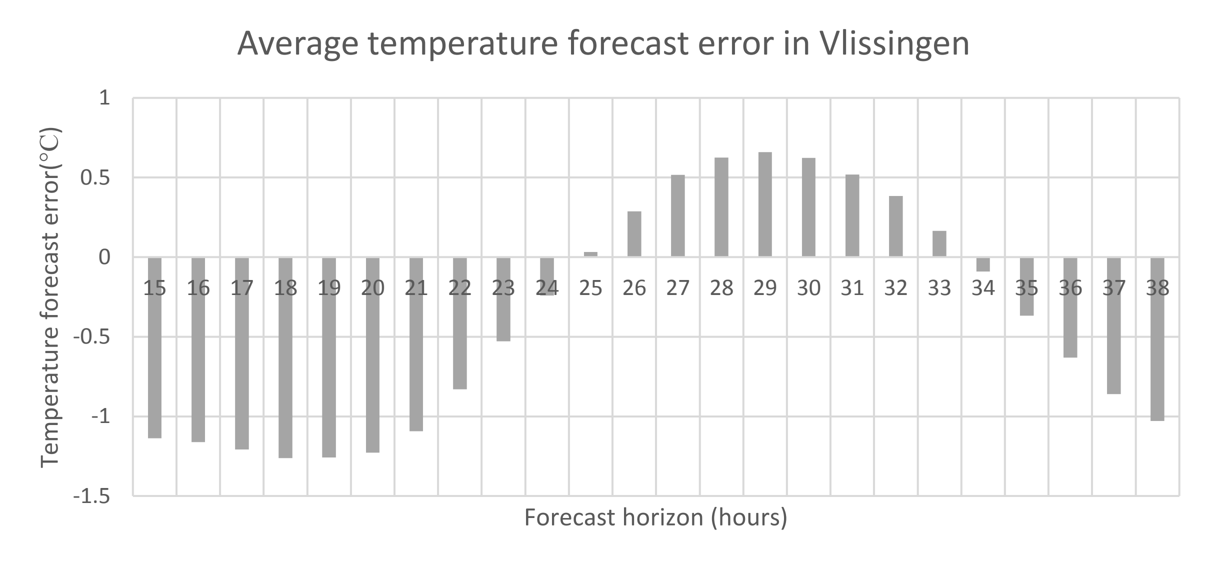

Figure 3.14: Average temperature forecast error in Vlissingen during the last 6 months of 2017.

The average error for temperature does show a systematic pattern. However, the average error is not positive or negative over the entire horizon. We have no explanation for the pattern in the temperature forecast error in Vlissingen. The forecast error has an average of -0.38. This average is significantly smaller than 0 with p=0.000 at a 95% confidence level.

3.5.2 Hupsel

Hupsel is located in Gelderland, this is far more inland than Vlissingen. In Section 3.3, we already saw that the wind speed pattern was different. Now, we look into the forecast pattern, to see whether there is a systematic error.

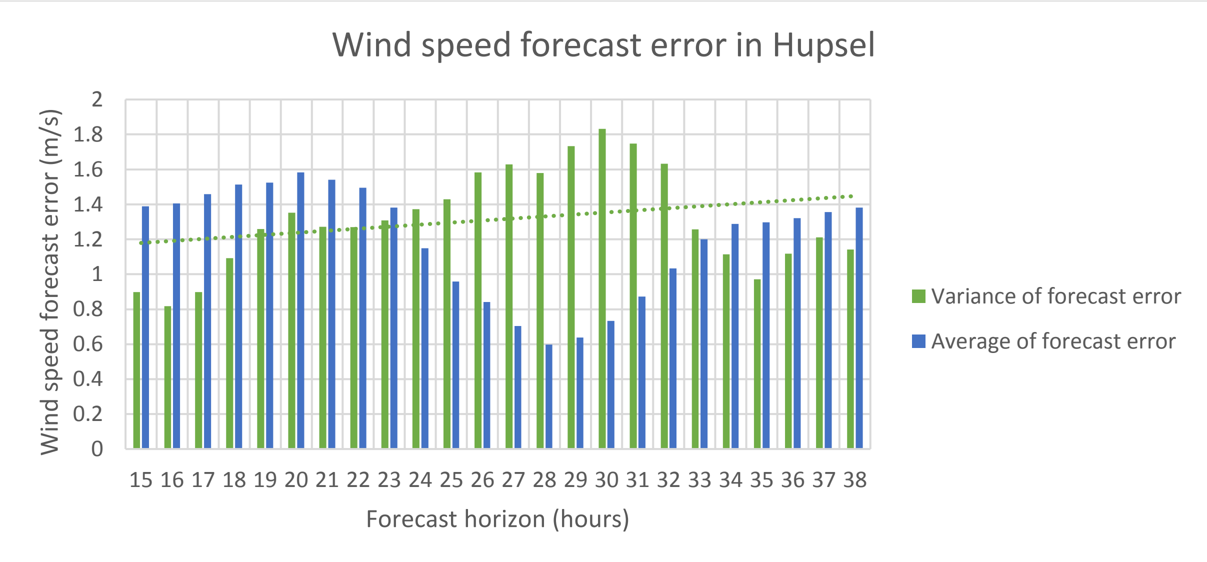

Figure 3.15: Variance and average of wind speed forecast error in Hupsel during the last 6 months of 2017.

The average wind speed forecast error is positive for the entire forecast horizon. This could indicate a bias. The average forecast error is 1.19 m/s, this average is significantly greater than 0 with p=0.000 (95% confidence level).

The variance of wind speed forecast error is smaller than in Vlissingen. This can be expected, since the absolute wind speed is higher in Vlissingen. Larger absolute values increases the overall variance, a larger variance means that it is harder to predict. The trend line indicates that the forecast error variance increases as the forecast horizon increases. We also saw this in Vlissingen. This is logical, because weather forecasts tend to get worse as the forecast horizon increases. Surprisingly, the forecast error does not increase as the horizon increases.

Looking at the forecast error for wind direction in Hupsel, we see the following pattern emerge:

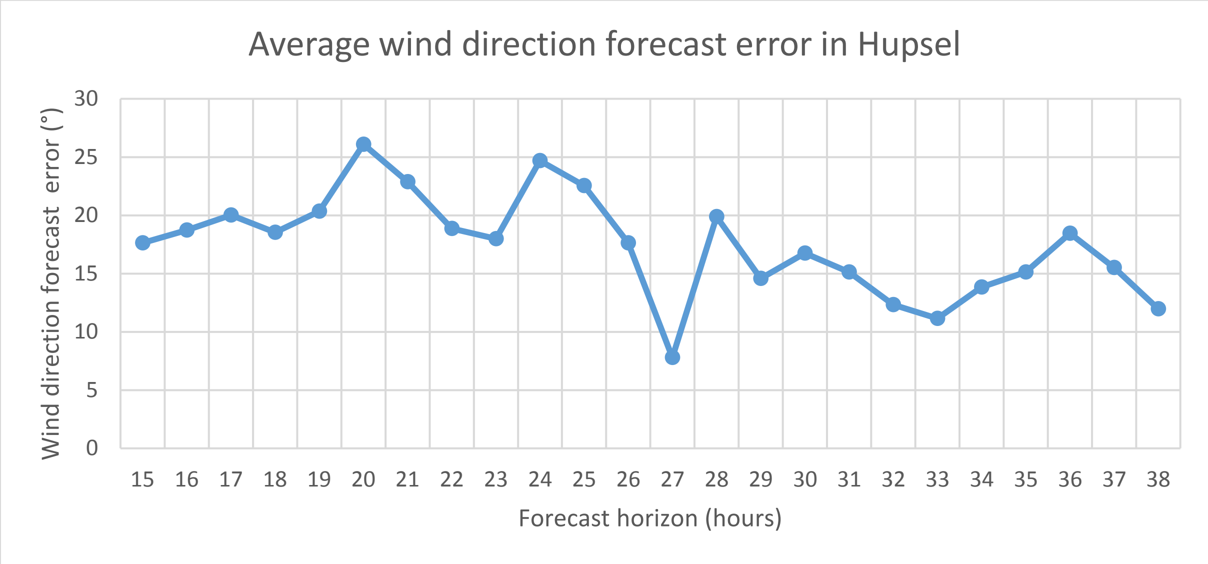

Figure 3.16: Average wind direction forecast error in Hupsel during the last 6 months of 2017.

Again, we removed the observations between 0-10° and 350-360° to prevent misleading results. In Figure 3.16, we can see that the entire forecast horizon contains positive average wind direction forecast errors. The suspected bias is confirmed by the T-test in Appendix J. The average wind direction forecast error in Hupsel of 16.96° is significantly greater than 0 with p=0.000 (95% confidence level).

The graph for average temperature forecast error in Hupsel looks roughly the same as in Vlissingen, we cannot explain the pattern. The average temperature forecast error in Hupsel of -0.06 °C is significantly different from 0 with p=0.002 (95% confidence level). This means there is a small negative bias in the forecast. We refer to Appendix E for the graph.

3.5.3 Lelystad

Lelystad is located near the IJsselmeer, which generally has a stronger wind than Hupsel, but not as strong as in Vlissingen. We expect this would lead to a smaller variance of forecast errors than in Vlissingen, but a larger variance than in Hupsel.

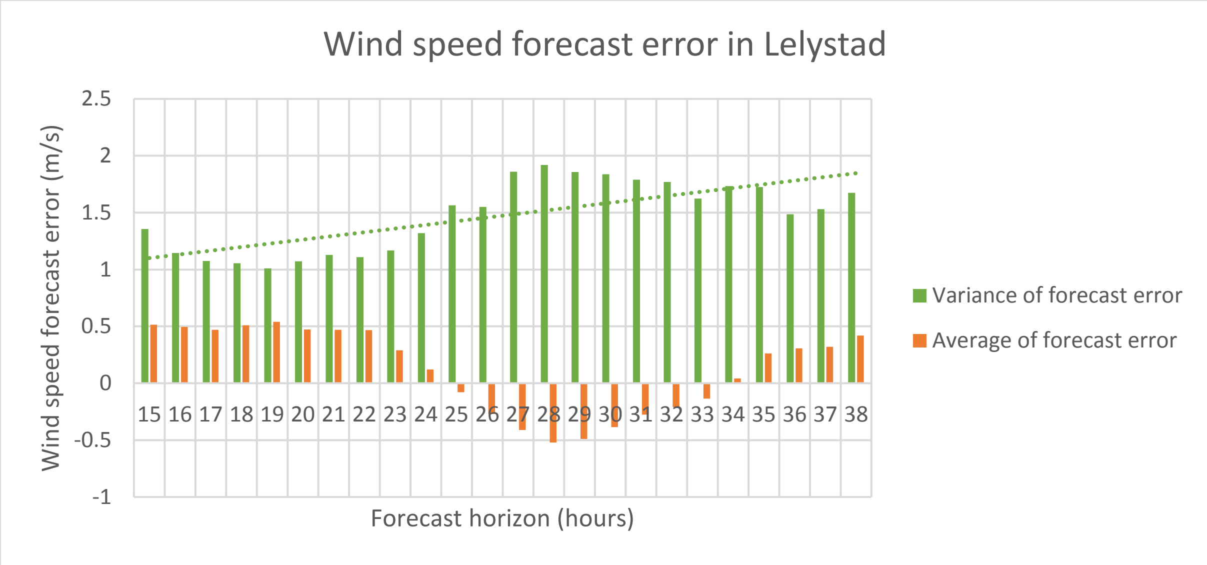

Figure 3.17: Variance and average wind speed forecast error in Lelystad during the last 6 months in 2017.

Figure 3.17 shows that the variance of wind speed forecast errors in Lelystad is approximately the same as in Hupsel. The variance does increase as the forecast horizon increases. With respect to a systematic error, the average error is concentrated around zero, so there is no reason to believe there is a systematic error. The T-test shows that the average wind speed forecast error in Lelystad of 0.12 m/s is significantly greater than 0 with p=0.000 (95% confidence level). This indicates there is a small positive bias in the wind speed forecast in Lelystad.

When looking at the wind direction forecast error for Lelystad, we removed observations between 0-10° and 350-350° again.

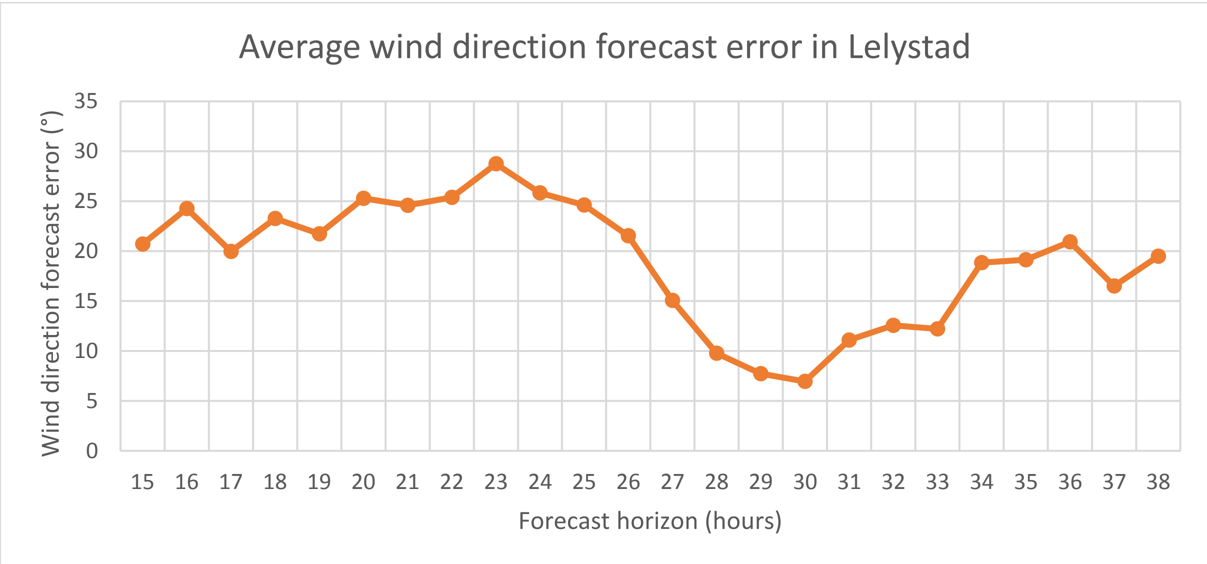

Figure 3.18: Average wind direction forecast error in Lelystad during the last 6 months in 2017.

Like in Vlissingen and Hupsel, the entire forecast horizon has a positive average wind direction error. Also, the error decreases near the forecast horizon of 25 hours, just like in Vlissingen, we cannot explain this. The average wind direction forecast error in Lelystad is 18.36°, this is significantly different from 0 with p=0.000 (95% confidence level).

Lastly, we look into the temperature forecast error in Lelystad. This graph can be found in Appendix E. The average error is concentrated around zero, just like with the temperature forecast in Vlissingen and Hupsel, there is no reason to believe there is a bias in the forecast. However, the T-test shows there is a small bias in the temperature forecast in Lelystad. The average of 0.09 °C is significantly different from 0 with p=0.000 (95% confidence level).

3.5.4 *Summary of bias analysis*

In this subsection, we provide a summary of the bias analysis of the previous subsections. The values displayed in Table 3.5 are obtained with a one-sample T-test using the statistical package SPSS. For the output of the T-tests, we refer to Appendix J.

Day ahead forecast	Average error (Forecast – Actual)	Significantly different than 0? (95% confidence level)
Wind speed in Vlissingen	-0.99 m/s	Yes
Wind speed in Hupsel	1.19 m/s	Yes
Wind speed in Lelystad	0.12 m/s	Yes
Wind direction in Vlissingen	7.87°	Yes
Wind direction in Hupsel	16.96°	Yes
Wind direction in Lelystad	18.36°	Yes
Temperature in Vlissingen	-0.38 °C	Yes
Temperature in Hupsel	-0.06 °C	Yes
Temperature in Lelystad	0.09 °C	Yes

Table 3.5: Summary of forecast biases in Vlissingen, Hupsel and Lelystad.

All average day ahead forecast errors are significantly different than 0, indicating the forecasts are biased. Especially the wind speed bias in Vlissingen and Hupsel can have serious implications for the power production forecast of wind turbines near these locations. In Chapter 4, we discuss how we use the biases to improve the accuracy of the day ahead power production forecast.

3.5.5 *Wind speed forecast accuracy*

In the previous subsections, we have confirmed there are biases in the weather forecasts in Vlissingen, Hupsel and Lelystad. We did this by looking at average forecast errors. Negative and positive errors cancel out, which affects the average forecast error. We used this to determine whether the forecasts were biased. To measure forecast accuracy however, we should look at absolute or squared errors instead of average errors. We would like to assess wind speed forecast accuracy, because this is the key predictor for wind power production. To assess the wind speed forecast accuracy, we can use RMSE from Equation 2.18 in Section 2.7, because at the three locations we would like to compare, the scale (wind speed in m/s) is the same.

For each location we would expect the RMSE to increase as the forecast horizon increases. Vlissingen should have the biggest RMSE, since the wind speed is generally higher there. By this logic, Lelystad would have the second biggest RMSE and Hupsel the smallest.

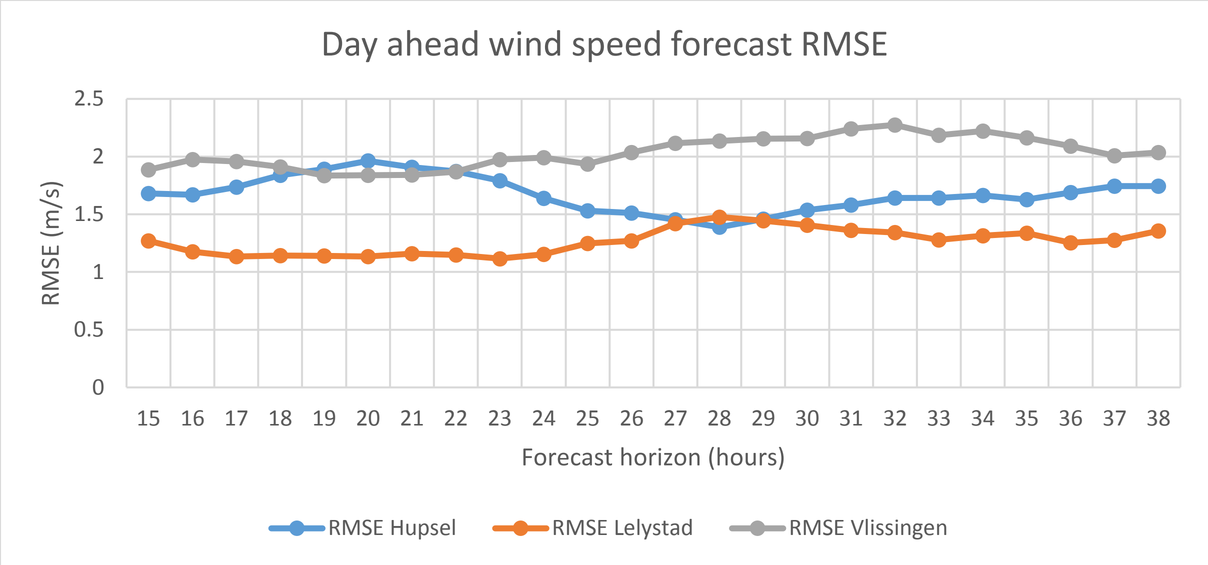

Figure 3.19: Root Mean Squared Error (RMSE) for day ahead wind speed forecasts in Vlissingen, Hupsel and Lelystad during the last 6 months in 2017.

In Figure 3.19, we can see that in Lelystad and Vlissingen the RMSE increases slightly as the forecast horizon increases. In Hupsel however, RMSE actually decreases a little. Our expectation about Vlissingen is confirmed, Vlissingen has the biggest RMSE. To our surprise, Lelystad has the smallest RMSE instead of Hupsel. This means the NWP models were better at predicting the wind speed in Lelystad than in Hupsel and Vlissingen during the second half of 2017.

3.6 Cleaning the data

In Sections 3.1 and 3.4 we discussed some problems with the historical production data that should be addressed. Table 3.4 and Figure 3.5 in Section 3.4 illustrate that there is a lot of variation in production for each wind speed. This can be due to discrepancies in wind speeds at the turbine hub height and the KNMI station or due to disturbing factors like windstorms, icing, maintenance and failures. To capture the relationship between historical wind speed and historical production as accurately as possible, we should address these problems.

Some observations should be removed, since they were affected by the disturbing factors that are mentioned above. For example, maintenance at a turbine resulted in a production of 0 kW for several hours while the average wind speed was 10 m/s during these hours. Due to the maintenance, the relationship between wind speed and production was affected. We want to exclude these observations so we can accurately capture the relationship between historical wind speed and production. Also, producers are obligated to report scheduled maintenance in advance and report expected repair times after failures. DVEP can adjust the day ahead production forecast for each producer accordingly. Windstorms and icing are indicated by day ahead weather forecasts. DVEP adjusts the day ahead production forecast if these weather conditions are predicted. Because DVEP can adjust the day ahead production forecast that is used for the auction, DVEP wants a production forecast that excludes windstorms, icing, maintenance and failures. Therefore, we try to exclude all these observations. Here a distinction between producers with single turbines and multiple turbines (wind farms) has to be made.

3.6.1 Single turbine

To avoid problems with single turbines shutting down during parts of an hour, we exclude observations where the production was smaller than or equal to 0 kW during a quarter in the hour. To clarify, observations are only included if the production was above 0 kW during each quarter of the hour. This removes historical data from failures and maintenance completely. Icing usually takes a couple of hours or days depending on the weather, so this will be completely removed as well. Downtime around cut-in and cut-out speeds will be mostly removed also. The only cases that slip through the cracks is when the turbine shuts down and starts up again within a quarter of an hour. This rarely happens and does not have a big impact according to DVEP.

3.6.2 Multiple turbines

For producers with multiple turbines connected to the same EAN, the data cleaning process is not as straightforward. This is because we do not know the number of turbines that are shut down at a certain time. When filtering the same way as with single turbines, we are only able to remove simultaneous downtime of all turbines. However, in case of maintenance or failures, most often not all turbines have simultaneous downtime. Maintenance of a wind park is scheduled one turbine at a time and failures rarely happen simultaneously. Each turbine in a wind farm has separate wind speed detection as well, which means some turbines can be active while other turbines in the same wind farm are shut down due to wind speeds below the cut-in threshold or above the cut-out threshold. The same goes for ice detection, this is monitored for each turbine separately as well.

To circumvent these problems with multiple turbines we conduct an ‘extra round’ of data cleaning. The data are cleaned using a lower and upper bound production for each wind speed. For each wind speed bin (bin size of 0.5) we calculate the average production per hour. The lower- and upper bound are constructed as follows:

$$LB_v = Average\ Production\ for\ wind\ speed\ v - \frac{1}{4} \times Rated\ Power \tag{3.2}$$

$$UB_v = Average\ Production\ for\ wind\ speed\ v + \frac{1}{4} \times Rated\ Power \tag{3.3}$$

For each wind speed, we count the number of observations, if this is lower than 10, we remove all observations for this wind speed. This is done to make sure the average production per wind speed is reliable, in order to get the appropriate bandwidth. For constructing the lower and upper bounds, we looked at several power curves. We felt that ± 25% would filter out most of the polluted data without removing too much clean data. The problem with data cleaning with multiple turbines is that it is impossible to distinguish the clean from the polluted data.

3.6.3 Data after cleaning

After cleaning the raw data, the polluted data due to downtime should be removed. This would mean that at operational wind speeds, the production should not be zero. Since Producer 3 has the most variance in its power curve according to Figure 3.5, we look at the power curve of Producer 3 after the data has been cleaned.

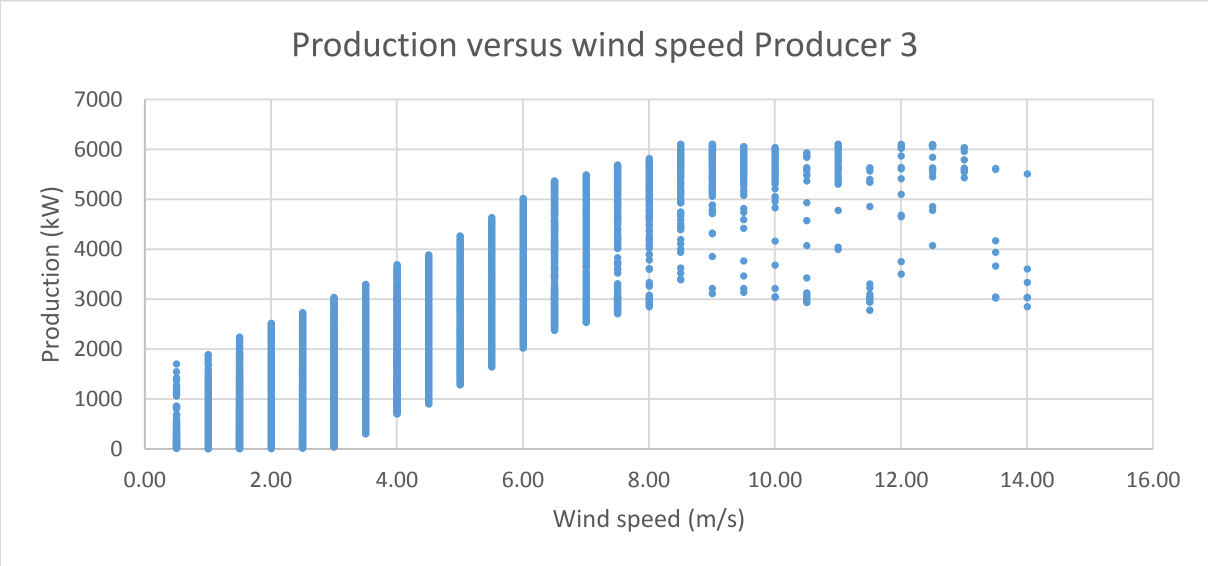

Figure 3.20: Scatterplot of production versus wind speed of Producer 3 with 6 MW total rated power.

Producer 3 has two turbines of 3 MW resulting in 6 MW total rated power. Cleaning the data was able to remove most of the polluted data. However, we still see some observations around 3 MW at high wind speeds. This probably indicates that one turbine was fully operational, while the other had been shut down. The cause of one turbine being shut down could be icing, maintenance, failure or a storm. This illustrates how difficult it is to clean the data properly, we are able to remove most of the polluted data. However, we will never be able to remove 100% without removing a large amount of clean data as well. We could lower the bandwidth around the average, this would remove more polluted data. However, most additional data that is removed will be clean data.

The graphs with cleaned data for Producers 1 and 2 can be found in Appendix F.

3.7 Conclusion

In this chapter, we evaluated which data are available. We analyzed the data to see if there were patterns to be found. We compared weather data in Vlissingen, Hupsel and Lelystad and looked at production data for a producer close to each location. This revealed some problems in the datasets, which we tried to remove by cleaning the data. We also looked at hindcasts to see how accurate the weather forecasts are and to see if the weather forecasts are biased. All of this leads to the following main conclusions:

- For 15 weather stations throughout the Netherlands, we have historical (measured) weather data for average hourly wind speed, wind direction, temperature, rainfall and radiation.
- Historical weather forecasts (hindcasts) are only available for the second half of 2017.
- For each producer of DVEP we have historical data for the quarter hourly production.
- A selection of 10 producers is made with a total rated power of 45.65 MW, which is approximately 12% of the total portfolio of DVEP.
- Wind speed distributions in 2015-2016 in Vlissingen, Hupsel and Lelystad are very different. Vlissingen generally has higher wind speeds more often than Lelystad and Hupsel, Hupsel has the lowest wind speeds according to the wind speed distributions.
- Wind direction shows the same patterns for all three locations. Southwest winds prevail at all locations. Wind speeds are also greater when the direction is southwest.
- Production data shows that power curves of different producers are very different. The raw data are affected by failures, maintenance, icing and windstorms. Cleaning the data makes

sure that most effects are mitigated. However, it is hard to distinguish polluted data from clean data.

- For some wind directions, we see a depression in production without a depression in wind speed. This could indicate obstacles and suggests that including wind direction could improve the forecasting model.
- All weather hindcasts are biased. Especially wind speed and wind direction forecasts are biased. Temperature only has a small bias at each location. In Chapter 4 we discuss how we use the bias to improve the weather forecasts.
- Wind speed forecasts in Vlissingen have the largest RMSE, which means the wind speed forecast in Vlissingen is most inaccurate. Lelystad has the most accurate wind speed forecasts. The forecast error does not increase as the forecast horizon increases.

4. Solution Design

In this chapter, we answer Sub-question 4: *Which forecasting approach should result in the most accurate day ahead forecast according to the data patterns and literature review?* First, define the general forecasting approach in Section 4.1. In Section 4.2 we make adjustments to the models that were proposed by the literature. We discuss how we add temperature and wind direction to the models, to see if this increases their accuracy. Section 4.3 discusses how we select the top 3 regression models that are introduced in Section 4.2. After this, in Section 4.4 we discuss how we select the best model and how we compare the accuracy of this model with the accuracy of Company X. Lastly, we discuss which optimization algorithm is used to estimate the parameters that optimize each model in Section 4.5.

4.1 General forecasting approach

In this section we discuss the choices we make with regard to our general forecasting approach. Our general forecasting approach describes how we translate day ahead weather forecasts into day ahead production forecasts for wind turbines. As mentioned in Chapters 2 and 3, we choose our approach based on the available data and the forecast horizon.

The forecast horizon for this project is 15-38 hours ahead due to the auction. If we recall, the auction requires DVEP to deliver a day ahead production forecast before the deadline at 12:00 (noon). Weather forecasts from 9:00 are used to create a time buffer of 3 hours for technical issues. The day ahead production forecast requires a summed production forecast of all producers in the portfolio of DVEP for each hour. The first hour (00:00 – 01:00) of the day ahead forecast is 15 hours ahead; the last hour (23:00 – 00:00) is 38 hours ahead, therefore the forecast horizon is 15-38 hours ahead.

In Chapter 2 we saw that if we want to forecast power production up to 48 hours ahead, it is essential to use weather predictions from NWP models. This means we need weather predictions up to 38 hours ahead, since this is our maximum forecast horizon. Fortunately, these weather forecasts are available. To translate weather forecasts into production forecasts, we distinguished two approaches in Chapter 2, the physical approach and the statistical approach.

The physical approach inserts weather forecasts into 3D models with detailed information of the turbine site to accurately describe the air flow at the turbine site. The manufacturer's power curve is then used to calculate the predicted power production. The statistical approach uses statistical methods like regression models or Artificial Intelligence (AI) approaches to capture the relationship between historical meteorological measurements or weather forecasts and historical power production as accurately as possible. Since the manufacturer's power curve is not available and we do not have detailed information of the turbine site, we have to use the statistical approach.

In Chapter 3 we looked at which data are available. In Table 3.2 an overview of the available data illustrates that historical production and weather measurements are available for 2015-2017. Weather hindcasts are only available for the second half of 2017. Historical weather measurements contain hourly average values for wind speed, wind direction and temperature. Weather hindcasts contain predicted values for wind speed, wind direction and temperature for each hour.

In the statistical approach we can use either historical weather measurements or weather hindcasts to capture the relationship between production and meteorological variables. Ideally, we would like to capture the relationship between weather hindcasts and production, since the weather forecasts are used as input to predict power production. However, we believe a period of 6 months is too short to accurately capture this relationship. Therefore, we use historical measurements from 2015-2016 to try to capture the relationship between historical weather measurements and historical power production as accurately as possible.

To evaluate which model can most accurately predict day ahead power production, we need to use weather hindcasts as input instead of historical weather measurements. To account for the discrepancy between the weather measurements and weather hindcasts, we evaluated the weather hindcasts in Chapter 3. We found statistical evidence that the weather hindcasts are biased. Especially wind speed and wind direction forecasts are biased. We calculated the average forecast error over the forecast horizon of 15-38 hours ahead (not for each hour separately). To improve the weather hindcasts, we can adjust the value by adding the average forecast error. For example, the wind speed forecast per hour in Vlissingen is on average 0.99 m/s too low. Therefore, we add 0.99 m/s to the forecasted value. We adjust the weather hindcasts for wind speed and wind direction biases. Temperature forecasts only have a very small bias, which is why we do not adjust the temperature forecasts.

In Chapter 2, we saw that we can use several methods to capture the relationship between historical weather measurements and power production. These methods include causal models like regression models, AI methods like Artificial Neural Networks (ANNs). Regression is the simpler approach and is more understandable than AI approaches. Also, regression is able to use multiple predictors. In Chapter 3, we saw that this comes in handy, since we want to see the effect of adding temperature and wind direction as predictors. Therefore, we use regression models to capture the relationship between historical weather measurements and power production.

Eventually, our general forecasting approach can be summarized in 3 steps:

Step 1: Estimate the least squares parameters of a regression model using historical weather measurements and historical production measurements.

Step 2: Adjust the weather forecasts for the bias found in Chapter 3.

Step 3: Insert the bias adjusted day ahead weather forecasts into the regression models (using the least squares parameters we estimated with the historical measurements) to translate weather forecasts into power production forecasts.

We use historical weather measurements to estimate the least squares parameters, because we only have weather hindcasts from the second half of 2017. We believe this period is too short to accurately capture the relationship between weather hindcasts and historical power production.

4.2 Regression models

In this section, we introduce the regression models we use to capture the relationship between the historical weather measurements and historical production data. In Section 2.5, we introduced several WTPC modelling techniques. Parametric WTPC modelling is a form of simple regression using only wind speed, this model can be expanded by adding temperature and wind direction, making it multiple regression. Parametric WTPC modeling can be linear and nonlinear, nonlinear models have more freedom in the functions that can be used. Nonlinear regression also has no underlying assumptions. However, nonlinear regression requires the need of an iterative optimization algorithm which can converge to a local optimum instead of the global optimum. Linear regression can be solved with a Linear Programming (LP) solver and ensures a global optimum is reached. Although, linear regression has some assumptions that should be met before it can be used.

In Subsection 4.2.1 we start with the parametric WTPC modelling techniques we introduced in Section 2.5. These models only use wind speed as predictor, in Subsection 4.2.2 we add temperature as predictor. Lastly, in Subsection 4.2.3 we add wind direction to the models.

4.2.1 Predictor: wind speed

The WTPC modelling techniques that showed promising results are listed in Table 4.1.

Model	Function
Approximate cubic	$P = \frac{1}{2}\rho A C_{p,max} v^3$
N-th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + \cdots + a_n v^n)$
Exponential	$P = \left(P_r\left(1 + \left(\frac{\beta}{v}\right)^{\alpha}\right)^{-y}\right)$
Logistic 4	$P = \left(\alpha\left(\frac{1 + me^{-\frac{v}{\tau}}}{1 + ne^{-\frac{v}{\tau}}}\right)\right)$
Logistic 5	$P = \left(d + \left(\frac{a - d}{\left(1 + \left(\frac{v}{f}\right)^{b}\right)^{g}}\right)\right)$

Table 4.1: WTPC modelling techniques.

All models except the approximate cubic model only use wind speed as predictor. The approximate cubic model uses air density (ρ) as well. However, we have no data about the air density. Also, we have no data about the power coefficient ($C_{p,max}$). To be able to use the approximate cubic we should make some adjustments. Air density should be removed, since we do not have sufficient data about this. We replace the power coefficient with parameter α. This parameter also captures the swept rotor area and the $\frac{1}{2}$ in the equation. These parts can be kept in the equation, however, this would only make α greater by a factor of $\frac{1}{2}A$. According to the literature, models based on the cubic relationship between wind speed and power are cumbersome. Therefore, we introduce parameter β. The theoretical relationship between wind speed and power would suggest the value of β is close to 3. All these changes lead to Equation 4.1.

$$P = \alpha v^{\beta} \tag{4.1}$$

We can use a log transformation on this equation to make it linear in the parameters so we can use linear regression. The advantage of linear regression is that we do not need an iterative optimization algorithm and we are sure that we can find the global optimum. However, we should check whether the assumptions of linear regression are met for the transformed model. The log transformation of Equation 4.1 results in Equation 4.2.

$$\log P = \log \alpha + \beta \log v \tag{4.2}$$

After the parameters have been calculated, $\log P$ can be transformed back using Euler's number, *e*, which results in predicted values for the power production. We shall call the log transformation of the adapted approximate cubic model the 'log model'.

The polynomial model can be used for linear regression as well. We can simply calculate the values for v, v^2, v^3, etc. and insert these as an independent variable. For the polynomial models we only use the

3rd, 4th and 5th degree polynomial models, because we believe that a higher degree polynomial would use too many parameters, keeping in mind that we want to add temperature and wind direction as well.

The exponential, logistic 4 and logistic 5 models cannot be used for linear regression, which means we need an iterative optimization algorithm. In Section 4.5 we discuss which algorithm we use for nonlinear regression.

The adjustments we make lead to the following regression models using only wind speed as predictor.

Model	Function
Log	$\log P = \log \alpha + \beta \log v$
3rd degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3)$
4th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4)$
5th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4 + a_5 v^5)$
Exponential	$P = \left(P_r \left(1 + \left(\frac{\beta}{v}\right)^{\alpha} \right)^{-y} \right)$
Logistic 4	$P = \left(\alpha \left(\frac{1 + m e^{-\frac{v}{\tau}}}{1 + n e^{-\frac{v}{\tau}}} \right) \right)$
Logistic 5	$P = \left(d + \left(\frac{a - d}{\left(1 + \left(\frac{v}{f}\right)^{b} \right)^{g}} \right) \right)$

Table 4.2: Regression models using wind speed as predictor.

4.2.2 Predictors: wind speed and temperature

To include more information in the regression models, we add temperature as a predictor. In Chapters 2 and 3 we saw that temperature is related to air density. Unfortunately, we do not have historical data or forecasts for air density, so we add temperature as a proxy. The theoretical relationship between air density and production suggests we should add temperature multiplicatively to the models. In Equation 3.1, we described the theoretical relationship between air density and temperature. Following this relationship, we add temperature to the nonlinear models in the following fashion:

$$P = f(v) \times \frac{k}{T} \tag{4.3}$$

where:

P = power in W

$f(v)$ = function with wind speed v as variable

T = temperature in K

k = parameter for temperature

In Equation 4.3, $f(v)$ represents the functions in Table 4.2 except for the log model. We expect the value of $\frac{k}{T}$ to be around 1.225 at 15 °C, since this is the standard air density at mean sea level for 15 °C. This means the value of k should be approximately 350. For the log model, we should add temperature differently, because adding temperature like this is not possible in linear regression. In the log model, we add temperature additively:

$$\log P = \log \alpha + \beta \log v + k \log T \tag{4.4}$$

We do not use $\log \frac{1}{T}$, because using temperature in Kelvin leads to small values for $\frac{1}{T}$. The log of small barely changes, therefore we use $\log T$. Using temperature in Celsius in not an option, because the log of a negative number does not exist, which means we would have to remove all observations with subzero temperatures.

Adding temperature additively has the disadvantage that the effect of temperature is the same for all wind speeds. Whether the wind speed is 0 m/s or 20 m/s, the effect of temperature remains the same. This is not the case according to the theoretical relationship between air density (by extent temperature) and production. Therefore, we do not add temperature to the polynomial models additively. Adding temperature like this for the log model enables us to use linear regression, so we will find a global optimum. Whether this optimum is better than the optima of the nonlinear models is to be seen.

The adjustments we make lead to the following regression models using wind speed and temperature as predictors.

Model	Function
Log	$\log P = \log \alpha + \beta \log v + k \log T$
3rd degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3) \times \frac{k}{T}$
4th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4) \times \frac{k}{T}$
5th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4 + a_5 v^5) \times \frac{k}{T}$
Exponential	$P = \left(P_r \left(1 + \left(\frac{\beta}{v}\right)^{\alpha}\right)^{-y}\right) \times \frac{k}{T}$
Logistic 4	$P = \left(\alpha \left(\frac{1 + me^{-\frac{v}{\tau}}}{1 + ne^{-\frac{v}{\tau}}}\right)\right) \times \frac{k}{T}$
Logistic 5	$P = \left(d + \left(\frac{a - d}{\left(1 + \left(\frac{v}{f}\right)^{b}\right)^{g}}\right)\right) \times \frac{k}{T}$

Table 4.3: Regression models using wind speed and temperature as predictors.

4.2.3 Predictors: wind speed, temperature, and wind direction

Adding wind direction is not as straightforward as temperature. This is due to the fact that wind direction has a different scale than temperature. Wind direction is measured between 0° and 360°, here 0° and 360° represent the same point, which complicates things. Also, the relationship between wind direction and production is different than the relationship between temperature and production.

A higher temperature should theoretically mean a lower production. For wind direction, this is more dynamic. There is little to no information available on wind direction modelling, so we have to be innovative.

Some wind turbines have a wind sector at which they generally have a high production, this peak is usually accompanied by depressions on both sides next to it. Figures 3.7 and 3.8 illustrate this. Especially for wind farms, like Producer 2, where the wake effect can negatively affect the production, including wind direction could improve the accuracy. For Producer 2, we saw signs that the wake effect negatively impacts the production in opposing wind directions (180° difference in wind direction). The production peak ('optimal' wind direction) was located in between these wind directions. Meaning we have a peak with depressions on both sides next to it. To model a production peak, and depressions on both sides next to the peak, we can use a periodic function. The sine and cosine function come to mind.

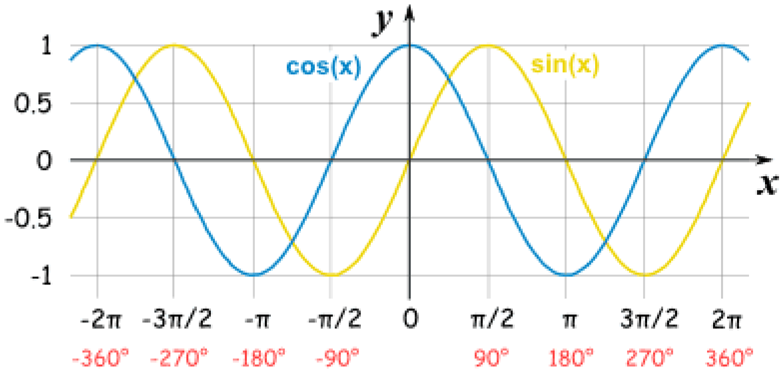

Figure 4.1: Sine and cosine functions between -2π and 2π.

The cosine function has the advantage that cos(0) = 1. This is useful for modeling a peak, therefore we use the cosine function to incorporate the wind direction in the model. We add wind direction multiplicatively just like temperature. This results in Equation 4.2:

$$P = f(v) \times \frac{k}{T} \times \left(\lambda \cos \frac{\theta - \delta}{c}\right) \tag{4.2}$$

where:

λ = parameter for wind direction

θ = wind direction in radians

δ = parameter that indicates 'optimal' wind direction in radians ($0 < \delta < 2\pi$)

c = parameter that determines the range of the cosine function ($c > 0$)

The parameter for the optimal wind direction, δ, should be between 0 and 2π, because 2π in radians is equal to 360°. This means the optimal wind direction should be between 0° and 360°, which is logical. We let the model choose the optimal wind direction for each producer, it should choose the wind direction with the highest average production. The parameter c is used to limit the range of the cosine function, we only want to utilize the cosine function between $-\frac{1}{2}\pi$ and $\frac{1}{2}\pi$, since we do not want values smaller than 0. We illustrate this in Figures 4.2 and 4.3.

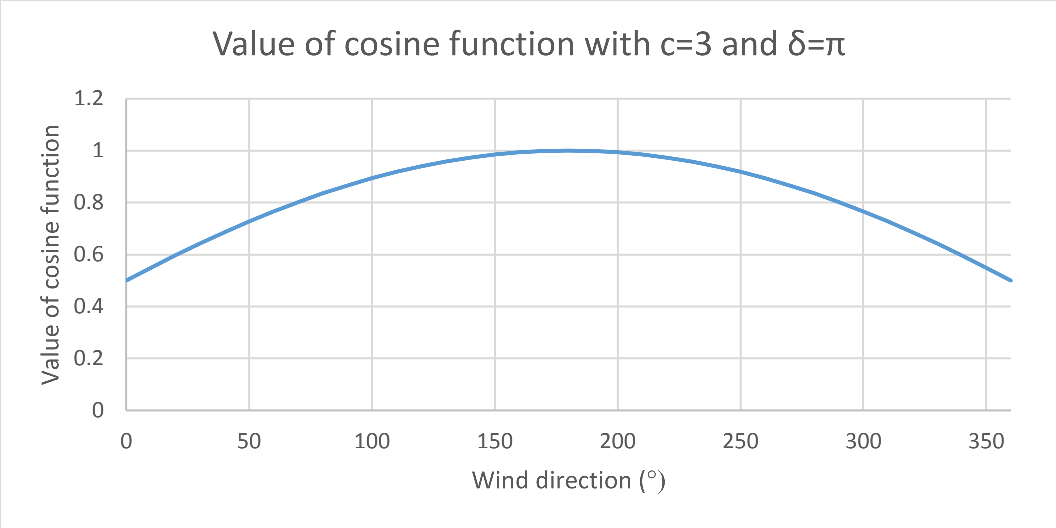

Figure 4.2: Value of cosine function for all wind directions using example of c=3 and δ=π.

In Figures 4.2 and 4.3 the optimal wind direction chosen by the model is 180° (π in radians), meaning the value of the cosine function is equal to 1 at 180°. When c=3, the value of the cosine function is 0.5 at 0° and 360°. In Figure 4.3, we change the value of c to 10. This results in values of 0.92 at 0° and 360°.

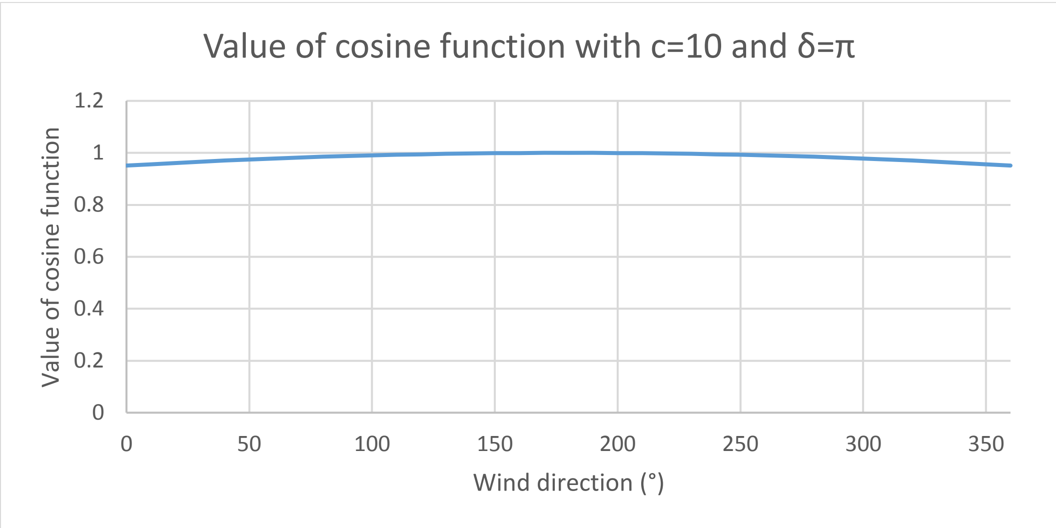

Figure 4.3: Value of cosine function for all wind directions using example of c=10 and δ=π.

In the example illustrated in Figures 4.2 and 4.3 we see that increasing the value of c reduces the effect of wind direction on the predicted production. For values of c > 20 the effect of wind direction is practically eliminated, since the value of the cosine function is between 0.99 and 1 for all wind directions. We use c as a parameter, which means the model 'chooses' to what extent wind direction is used as predictor. If wind direction does not influence the production for a certain producer, the value of c should increase so the effect of wind direction is eliminated.

For the Log model, we add wind direction differently to be able to use linear regression:

$$\log P = \log \alpha + \beta \log v + k \log T + \lambda \log \cos(\frac{\theta - \delta^*}{c^*}) \qquad (4.3)$$

We cannot use δ and c as parameters. The equation should be linear in the parameters, which means the parameters inside the cosine function should be replaced. We can do this by simply using trial and error to find the 'optimal' direction ,δ^*, instead of using parameter δ. The same can be done for the parameter c. We simply fill in values for δ^* and c^* and look at which values result in the best outcome. The parameter λ is used to determine how much the optimal wind direction should be rewarded. The production of the Log model can be calculated by reversing the log transformation using the irrational number e, this results in the following equation:

$$P = \alpha v^{\beta} \times T^{k} \times \cos(\frac{\theta-\delta^*}{c^*})^{\lambda} \quad (4.4)$$

The adjustments we make lead to the following regression models using wind speed, temperature, and wind direction as predictors.

Model	Function
Log	$\log P = \log \alpha + \beta \log v + k \log T + \lambda \log \cos(\frac{\theta - \delta^*}{c^*})$
3rd degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3) \times \frac{k}{T} \times \left(\lambda \cos \frac{\theta - \delta}{c}\right)$
4th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4) \times \frac{k}{T} \times \left(\lambda \cos \frac{\theta - \delta}{c}\right)$
5th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4 + a_5 v^5) \times \frac{k}{T} \times \left(\lambda \cos \frac{\theta - \delta}{c}\right)$
Exponential	$P = \left(P_r \left(1 + \left(\frac{\beta}{v}\right)^{\alpha}\right)^{-y}\right) \times \frac{k}{T} \times \left(\lambda \cos \frac{\theta - \delta}{c}\right)$
Logistic 4	$P = \left(\alpha \left(\frac{1 + me^{-\frac{v}{\tau}}}{1 + ne^{-\frac{v}{\tau}}}\right)\right) \times \frac{k}{T} \times \left(\lambda \cos \frac{\theta - \delta}{c}\right)$
Logistic 5	$P = \left(d + \left(\frac{a - d}{\left(1 + \left(\frac{v}{f}\right)^{b}\right)^{g}}\right)\right) \times \frac{k}{T} \times \left(\lambda \cos \frac{\theta - \delta}{c}\right)$

Table 4.4: Regression models using wind speed, temperature, and wind direction as predictors.

In Chapter 5 we evaluate the 21 regression models introduced in this section. We want to know which model using which predictors can most accurately predict the historical production. In Section 4.3 we describe how we select the most accurate regression models. The most accurate models are used for day ahead forecast accuracy evaluation, we describe this process in Section 4.4.

4.3 Selecting the top 3 regression models

In Chapter 5 we estimate the parameters of the regression models from the previous section using clean data from 2015-2016, for each model we use three sets of predictors. First, we only use wind speed, then we expand the model by adding temperature. Finally, we use wind speed, temperature, and wind direction as predictors. We illustrate the process of parameter estimation in Figure 4.4.

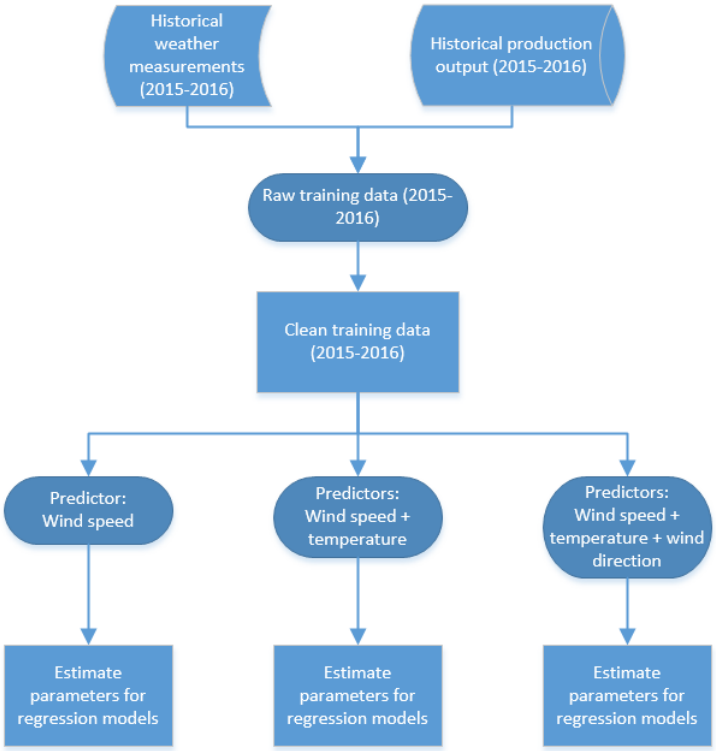

Figure 4.4: Parameter estimation process for Producers 1, 2, and 3.

We estimate the parameters for all regression models for Producers 1, 2, and 3, because doing this for all 10 producers would be too time consuming. For each producer we estimate parameters for 21 regression models, resulting in 63 models for Producers 1, 2, and 3 combined.

Producer 1 is located near Vlissingen, Producer 2 near Hupsel and Producer 3 near Lelystad. To narrow down the number of regression models, we select a top 3 which we use for all 10 producers. To prevent overfitting the models to the training data, we select the 3 most accurate models based on test data from the first half of 2017. As selection criterion we use the standard error of regression (S), because we use linear and nonlinear regression models. The Root Mean Squared Error (RMSE) and *S* are very similar, we choose *S* because it slightly punishes the number of parameters. For each regression model, we calculate the *S* for the training and test data. After we select the 3 models that are most accurate for Producers 1, 2, and 3, we estimate the parameters for these models for the remaining 7 producers.

4.4 Evaluating day ahead forecast accuracy

At this point in Chapter 5, we should know which 3 regression models, using which predictors, can most accurately capture the relationship between historical weather measurements and historical power production. However, we want to know which model can most accurately predict power production based on day ahead weather forecasts. So our next step is to insert the historical day ahead weather forecasts (hindcasts) into the top 3 regression models. In Figure 4.5 we illustrate the evaluation process for day ahead accuracy.

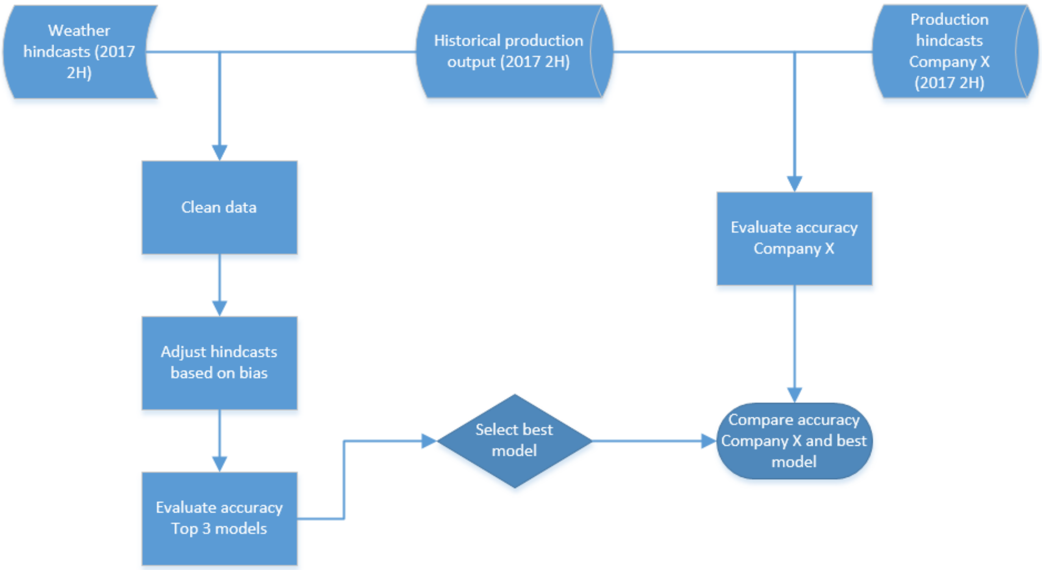

Figure 4.5: Day ahead accuracy evaluation process for Producers 1-10.

To evaluate the day ahead forecast accuracy for the second half of 2017, we use weather hindcasts as input and adjust the hindcast based on the average forecast errors in Table 3.5. We clean the production data by removing observations with production smaller than or equal to 0 for a quarter during an hour, just like we did for the training data of 2015-2016. We want to know the accuracy on clean data, since DVEP adjusts the day ahead forecast for maintenance, failures, icing and windstorms. So, we want to know which model is most accurate when these disturbing factors are removed.

After cleaning the data, the adjusted hindcasts are used in combination with the historical production data to evaluate the day ahead forecast accuracy for the cleaned data from the second half of 2017. We use the adjusted hindcasts to predict the production per hour. In Table 4.5 we illustrate an example of what our dataset looks like.

Date	Time	Hours ahead	Adjusted wind speed hindcast (m/s)	Adjusted wind direction hindcast (°)	Temperature hindcast (°C)	Production
1-7-2017	00:00 – 01:00	15	…	…	…	…
1-7-2017	01:00 – 02:00	16	…	…	…	…
…	…		…	…	…	…
…	…		…	…	…	…
31-12-2017	22:00 – 23:00	37	…	…	…	…
31-12-2017	23:00 – 00:00	38	…	…	…	…

Table 4.5: Example of dataset with weather hindcasts adjusted for bias.

We insert the adjusted weather hindcasts into the regression models to compute the production forecasts. For the top 3 models, we calculate the average standard error of regression over all hours (not for each hour separately) of the second half of 2017 for 10 producers. We select the best model based on the accuracy for Producers 1-10. To compare the best model with Company X, we use production hindcasts of Company X for the second half of 2017. We calculate the average RMSE over all hours of the second half of 2017 for Company X for Producers 1-10. We use RMSE as performance indicator instead of *S*, because we do not know the number of parameters for the model of Company

X. For all 10 producers, we compare the day ahead forecast accuracy of the best model and Company X for the second half of 2017.

Finally, in Chapter 5 we look at the aggregated day ahead forecast accuracy. The aggregated forecast is the sum of the production forecasts for Producers 1-10. We compare the accuracy of Company X and our best model by looking at the average RMSE and MAE for all hours. We also compare the RMSE and MAE for each hour in the forecast horizon separately (15 hours ahead, 16 hours ahead, etc.). We would expect the errors to increase as the forecast horizon increases for both models.

4.5 Optimization algorithm

In Section 2.6, we discussed the difference between linear and nonlinear least squares regression. Both type of regression models are optimization problems, since they involve a minimization of the SSE. However, solving this problem is easier for linear regression, an LP solver is able to find the global minimum for a linear regression model. We use the linear regression option in the statistical software package SPSS to estimate the parameters of the linear regression models.

For nonlinear regression, we need an iterative optimization approach. The reason that we need an iterative approach is that nonlinear optimization problems have local optimal solutions. The optimization algorithm can get stuck in a local optimum, while there are better solutions available.

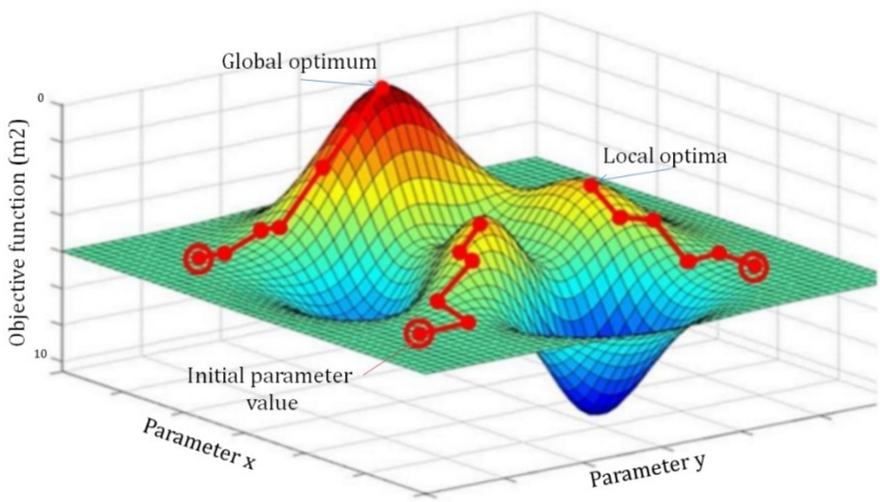

Figure 4.6: Global versus local optima in a simplified 3D representation.

Figure 4.6 illustrates the difference between a local and global optimum in a simplified 3D representation with only 2 parameters. We are looking to solve the optimization problem to optimality, however, the optimization algorithm is not able to distinguish whether a solution is a global optimum or a local optimum. The example in Figure 4.6 has only 2 parameters, which is fairly simple. When the number of parameters increases, the problem because more difficult very fast and the number of local optima increases. The user has to provide the algorithm with initial parameter values, providing good starting values can improve the solution that is found.

In Chapter 2, we saw that Lydia et al. (2013) used 4 optimization algorithms, namely a Genetic Algorithm (GA), Evolutionary Programming (EP), Particle Swarm Optimization (PSO) and Differential Evolution (DE). They found that DE had the best results. Unfortunately, we only have EP to our disposal. EP is a good algorithm, but it is relatively slow. Therefore, we have to find alternatives. Barati (2013) used a combination of Generalized Reduced Gradient (GRG) algorithm and Evolutionary Programming (EP). They found that combining these approaches increased the efficiency of the parameter estimation. We can use GRG to find a good initial solution and use EP to try and improve it. Both these algorithms are part of the Excel Solver package. GRG has a multistart option, which allows it to use a population of random start values. This increases the chance of finding a good initial solution (Frontline Solvers, 2018). GRG also allows us to set a lower and upper bound for the parameters, this confines the search area to a limited space, which reduces the search time.

The GRG algorithm is a gradient based method, which uses derivatives to find a local optimum. The algorithm starts at the initial values and makes adjustments to these values. By looking at the objective function, the algorithm knows whether it is moving in the right direction. The derivative values tell the algorithm whether a local optimum is reached. When using the multistart option, the algorithm starts again at random initial values and repeats this process. If the objective function is improved, the parameter values are saved as the current solution. The algorithm does not know when to stop, since it cannot determine whether a solution is a global optimum. Therefore, we can provide stopping conditions. We can adjust the population size, which determines how often the algorithm starts at random values. We can also use a convergence value, which tells the algorithm to stop when the objective function has not improved by a specific amount in the last 5 iterations (Frontline Solvers, 2018).

After finding a solution with the GRG algorithm, we use an evolutionary algorithm to try to improve the solution. Evolutionary algorithms apply the principles of evolution found in nature to solve the optimization problem. It relies on random sampling to find a population of solutions. Only one of these solutions is best. However, the candidate solutions are sample points in other regions of the search space. Now, the algorithm makes random changes (mutations) in members of the population, yielding new solutions which may be better or worse. The algorithm also performs cross-overs; this means it attempts to combine elements of existing solutions in order to create a new solution. Eventually, the algorithm performs a selection process in which the ‘most fit’ members of the population survive, and the ‘least fit’ members are eliminated. The selection process is the step that guides the evolutionary algorithm towards ever-better solutions (Frontline Solvers, 2018). Just like the GRG algorithm, the evolutionary algorithm does not know when to stop so we have to give some stopping conditions. We can adjust the population size, convergence value and we can set a maximum time without improvement.

The cost of estimating parameters using the GRG and evolutionary algorithms is that these methods do not provide parameter uncertainties. We only know which parameters minimize the objective function; we do not know the probable ranges these parameters take. Hu et al. (2015) used Monte Carlo simulation and a bootstrap method to estimate the parameter uncertainty. This is useful if we want to know the estimated parameter value that is most likely to be close to the unknown ‘real’ parameter value. However, in most of our models the parameters have no physical meaning. Therefore, the uncertainty of the parameters is of less importance. This is also due to the fact that in most equations the parameters influence each other. In most cases, a change in one parameter requires changes in other parameters to get diserable results. Especially when the model becomes more complex and the number of parameters increases, the uncertainty of a parameter tends to rise (Benke, Lowell & Hamilton, 2008). Since we use curve fitting equations which are fairly complex, we do not use Monte Carlo simulation or a bootstrap approach, like Hu et al. (2015). This means that parameter uncertainty is out of the scope of this project.

4.6 Conclusion

In this chapter, we defined our general forecasting approach. We introduced the regression models we use in Chapter 5. To include more information into the models, we made some model adjustments by adding temperature and wind direction as predictors. We also adjusted the approximate cubic model to be able to use a log transformation, which enables us to use linear regression. Furthermore, we illustrated how we select the best regression models in Chapter 5 based on the accuracy for Producers 1, 2, and 3. After this, we discussed how we evaluate the accuracy of the day ahead production forecast in Chapter 5. Lastly, we discussed the optimization algorithms that are used. All of this leads to the following conclusions:

- Our general forecasting approach can be summarized in 3 steps. In step 1 we estimate the least squares parameters of a regression model using historical weather measurements and historical production measurements. In step 2 we adjust the weather hindcast for the bias we found in Chapter 3. In step 3 we insert the adjusted weather hindcast into the regression models (using the least squares parameters we estimated in step 1) to translate weather hindcasts into production hindcasts.
- We use linear regression if possible, because linear regression ensures a global minimum is found and we do not need an iterative optimization algorithm. We use the linear regression option in the statistical package SPSS.
- We start with 7 regression models using only wind speed as predictor.
- Temperature is added as predictor to the 7 regression models as follows: $\times \frac{k}{T}$. k Resembles a parameter for temperature and T is the temperature in Kelvin. We divided k by the temperature, because as the temperature rises, the power output should decrease according to the theoretical relationship.
- Wind direction is added as predictor to the 7 regression models as follows: $\times \left(\lambda \cos\frac{\theta-\delta}{c}\right)$. We chose the cosine function, because it is a periodic function with a value of 1 at the peak. The parameter δ resembles the ‘optimal’ wind direction and dividing by the parameter c helps us confine the function to a specific region.
- We use 21 regression models in total, 7 using only wind speed as predictor, 7 using wind speed and temperature and 7 using wind speed, temperature, and wind direction as predictors. We use 6 linear models; the remaining 15 models are nonlinear.
- We estimate the least squares parameters for Producers 1, 2, and 3 for the 21 regression models using the clean training data from 2015-2016. The clean test data from the first half of 2017 are used to select the top 3 models. We estimate the least squares parameters for the top 3 models for the remaining 7 producers.
- The standard error of regression (S) is used as performance indicator for selecting the top 3 regression models. For the day ahead forecast accuracy we also use RMSE, MAE and NMAPE as performance indicators.
- We compare the day ahead forecast accuracy of the top 3 models. The accuracy of the best model is compared with the accuracy of Company X for each producer separately. We also compare the accuracy for the aggregated forecast of all 10 producers.
- We use the Generalized Reduced Gradient (GRG) algorithm in combination with the evolutionary algorithm to estimate the parameters that minimize the sum of squared errors. These methods do not provide parameter uncertainties.

5. Analysis of Results

In this chapter, we answer Sub-question 5: *Which day ahead forecasting model is most accurate in production forecasts and how accurate is this model in comparison to Company X?* First, we test the assumptions of the linear regression models in Section 5.1. In Section 5.2, we evaluate which regression model using which predictors can most accurately describe the relationship between historical weather measurements and historical production data. Section 5.3 discusses the effect of adding the predictors temperature and wind direction to the models. Also, we choose a top 3 regression models in Section 5.3. In Section 5.4, we evaluate the day ahead production forecast accuracy of the top 3 models after adjusting the weather forecasts for the bias found in Chapter 3. After this, we compare the accuracy of the best model with Company X's day ahead forecast for all 10 producers separately in Section 5.5. Next, we evaluate the day ahead forecast accuracy for the aggregated forecast in Section 5.6. Lastly, we end this chapter with a conclusion in Section 5.7.

5.1 Linear regression assumption testing

Before interpreting the results of the linear regression models by looking at the SPSS output, we should check whether the underlying assumptions of linear regression have been met. For each producer we have 6 linear regression models, so for Producers 1, 2, and 3 we have 18 linear regression models in total. We use the clean training data from 2015-2016, which has a sample size of 13,729, 15,003 and 11,702 for Producers 1, 2, and 3 respectively.

If we recall Chapter 2, simple linear regression has the following assumptions:

- Normality of errors.
- Linearity.
- Homoscedasticity.
- Independence of errors.

We test each assumption in this section.

Normality of errors

The normality of errors assumption requires the errors of the linear regression model to be normally distributed. Since we have a large sample size for each model (N>2000), we should use the Jarque-Bera test to test the normality of errors instead of the Shapiro-Wilk or Kolmogorov-Smirnov test (Thadewald & Buning, 2007). The Jarque-Bera (JB) test checks whether the sample data have the skewness and kurtosis matching a normal distribution. The Chi-square test with 2 degrees of freedom (X_2^2) can be used to test if the sample data are normality distributed at a confidence level of 1-α (Jarque & Bera, 1987). We use a confidence level of 95%. The JB test statistic should be close to 0 in case of normality. For more information on the JB test we refer to Appendix K.

Model	Predictor(s)	N	JB statistic	X_2^2 (α=0.05)	Normality?
Log	Wind speed	13729	1199.28	5.99	No
3rd degree Polynomial	Wind speed	13729	276.56	5.99	No
4th degree Polynomial	Wind speed	13729	204.97	5.99	No
5th degree Polynomial	Wind speed	13729	331.77	5.99	No
Log	Wind speed + Temperature	13729	1512.08	5.99	No
Log	Wind speed + temperature + wind direction	13729	1535.77	5.99	No

Table 5.1: Jarque-Bera test for normality of errors of linear regression models for Producer 1.

In Table 5.1 the normality of errors for all linear regression models for Producer 1 are tested. Since the JB statistics is greater than the critical point of the X^2 test at a 95% confidence level for all models, we have statistical evidence that the errors of all linear regression models for Producer 1 are not normally distributed. In Appendix K, we can see that the errors for all linear regression models of Producers 1, 2 and 3 are not normally distributed. This means the normality of errors assumptions for all linear regression models are violated.

Homoscedasticity

An underlying problem such as heteroscedasticity could result in errors not being normally distributed. We could check this by simply looking at a scatterplot of the dependent variable (production) and our key predictor (wind speed). We use Producer 1 as an example.

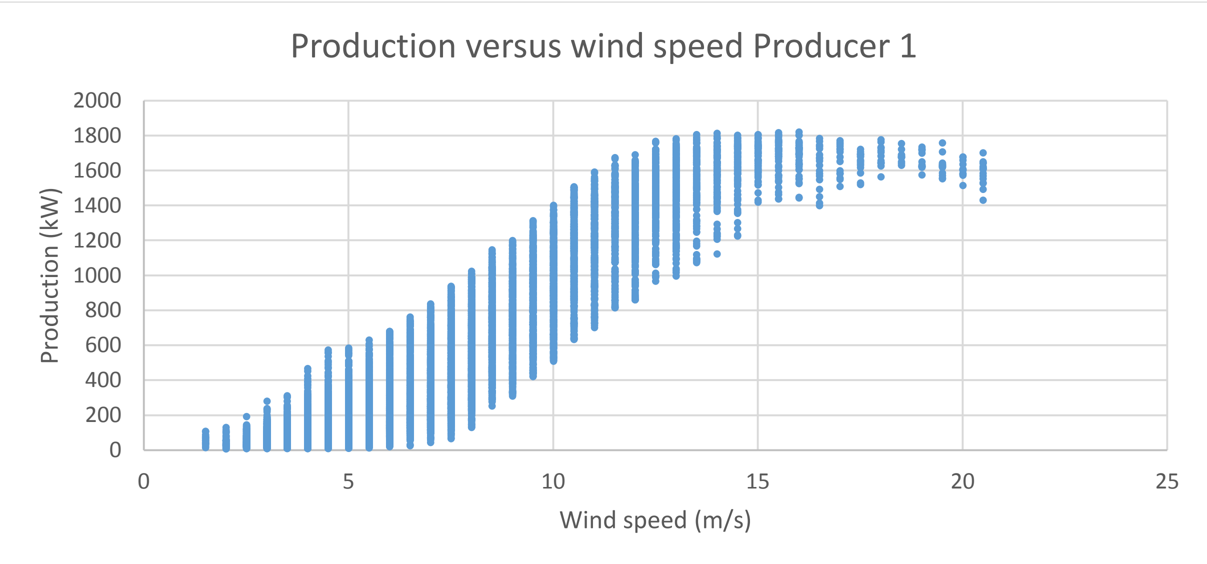

Figure 5.1: Scatterplot of production versus wind speed for Producer 1.

The assumption of homoscedasticity requires the variance in the dependent variable (production) to be more or less equal for all values of the predictor (wind speed). Figure 5.1 indicates that this is not the case for Producer 1. At wind speeds lower than 5 m/s, and higher than 15 m/s, the vertical spread of production is smaller than for the wind speeds in between. This could point to heteroscedasticity in the regression models. We check this for the 3rd degree Polynomial model of Producer 1 using only wind speed as predictor by looking at a scatterplot of the standardized predicted values versus the standardized residuals.

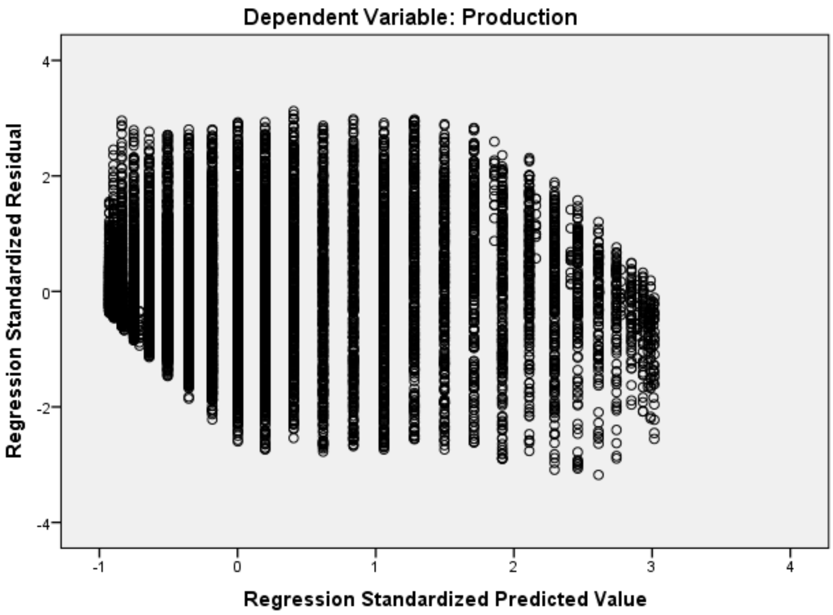

Figure 5.2: Scatterplot indicating heteroscedasticity for the 3rd degree Polynomial linear regression model of Producer 1using wind speed as predictor.

Figure 5.2 indicates heteroscedasticity since the standardized residuals are not equally spread out for all predicted values. For the standardized predicted values around -1 and 3 the variance in standardized residuals is much smaller than in the middle.

We have looked at scatterplots like Figure 5.2 for all linear regression models of Producers 1, 2, and 3. All models show signs of heteroscedasticity. We believe this is due to a difference in production variance for different wind speeds. The scatterplots of production versus wind speed for Producers 1, 2, and 3 all show smaller variances for low wind speeds and high wind speeds. Therefore, we conclude that for all linear regression models the homoscedasticity assumption is violated.

Linearity

The linearity assumption means the relationship between the dependent variable and the predictors should be linear. Also, the regression equation should be linear in the parameters. We made sure the regression equations are linear in the parameters. In Appendix I we illustrate the Pearson correlation matrices for Producers 1, 2, and 3. The linear correlation between wind speed and production is greater than 0.8 for each producer, meaning a strong linear relationship exists. Therefore, we believe the linearity assumption is met for all linear regression models.

Independence of errors

The assumption of independence of errors means the errors should not have serial correlation. We can use the Durbin-Watson test statistic to test for first order autocorrelation. We should make sure the data is sorted by time though. Again, we use the example of the 3rd degree Polynomial model for Producer 1 using only wind speed as predictor. We use SPSS to compute the Durbin-Watson (DW) statistic.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	,941[a]	,886	,886	157,76916	,638

Table 5.2: SPSS model summary including Durbin-Watson statistic for 3rd degree Polynomial linear regression model for Producer 1 using only wind speed as predictor.

The value of the Durbin-Watson statistic is 0.638 for the 3rd degree Polynomial linear regression model of Producer 1. In case of no serial autocorrelation, the DW statistic will have a value of 2. Rule of thumb for the Durbin-Watson statistic is that a value of smaller than 1 indicates that successive error terms are definitely positively correlated (King, 1983).When running all the linear regression models for each producer in SPSS, we included the DW statistic option. All values are smaller than 1, meaning the errors are positively correlated for all linear regression models of Producers 1, 2, and 3. Therefore, we assume the assumption of independence of errors is violated for all linear regression models.

Multicollinearity

Multiple linear regression has an additional assumption, which states that there should be no multicollinearity. This means that the predictors of the regression model should have no strong linear correlation. In Appendix I the correlation matrices for Producers 1, 2, and 3 are illustrated. We use wind speed, wind direction and temperature. In Chapter 3, we did see that wind speed and wind direction are definitely correlated. The matrices in Appendix I illustrate that the linear correlation between wind direction and wind speed is 0.167, 0.149, and 0.206 for Producers 1, 2, and 3 respectively. These correlations are statistically significant, but are too weak to cause serious problems with multicollinearity. The correlations between the other predictors are weak as well, so we conclude the multicollinearity assumption of multiple linear regression is met for all models.

When testing the assumptions of linear regression, some problems arise. None of linear regression models have normally distributed errors. Also, we suspect that all models are subject to heteroscedasticity and serial autocorrelation. Therefore, we should not use linear regression because some of the underlying assumptions are violated.

5.2 Accuracy using measured weather data

In this section, we look at which model is most accurate based on measured weather data. We want to know which regression model using which predictors can most accurately describe the relationship between historical weather data and production per hour. We do this by estimating the least squares parameters for each model using the cleaned training data from 2015-2016. We test the accuracy on a separate test dataset using clean data from the first 6 months of 2017. The standard error of regression is used as performance indicator. The standard error of regression is defined as:

$$S = \sqrt{\frac{\sum_{i=1}^{N}(Y_e(i)-Y_a(i))^2}{N-p-1}} \qquad (5.1)$$

where:

$Y_e(i)$ = predicted value of production for observation i

$Y_a(i)$ = actual value of production for observation i

N = number of observations

p = number of parameters

We calculate the standard error of regression for the entire clean training dataset from 2015-2016 and the entire clean test dataset from the first half of 2017. We do so for each combination of model and set of predictors. Each observation contains hourly average values. The training data contains more than 10,000 observations for each producer.

To prevent overfitting the model to the training dataset, we compute the standard error of regression for the test dataset as well. Normally, the performance of the models on the test data is worse than on the training data. If the performance is significantly worse, this indicates that we are overfitting the model to the training data.

For Producers 1, 2 and 3 we compare the standard error of regression on training and test data for each model using different predictors. The most accurate model has the smallest standard error of regression. In Tables 5.3, 5.4, and 5.5, the linear regression models are highlighted in red. We do not use these because the underlying assumptions are violated. All other models use nonlinear least squares regression. For each set of predictors, the most accurate model based on test data is highlighted in green.

5.2.1 Producer 1

The results for Producer 1 are illustrated in Table 5.3, Producer 1 has a combined rated power of 1.8 MW.

Predictors	**Wind speed**		**Wind speed + Temperature**		**Wind speed + Temperature + Wind direction**	
Model	**Training data (S)**	**Test data (S)**	**Training data (S)**	**Test data (S)**	**Training data (S)**	**Test data (S)**
Log	254.00	210.31	248.74	205.07	255.32	209.02
3rd Polynomial	157.77	161.31	155.00	158.85	155.01	159.00
4th Polynomial	157.02	161.09	154.87	158.62	154.19	158.98
5th Polynomial	155.75	160.01	153.18	157.22	152.87	157.29
Exponential	156.52	160.71	199.54	181.07	153.30	157.84
Logistic 4	155.74	160.07	152.85	157.45	152.86	157.64
Logistic 5	155.57	159.92	152.68	157.32	152.82	157.57

Table 5.3: Standard error of regression (S) for Producer 1 using different predictors.

When comparing the results of training and test data, we do not see a big difference in the accuracy for all models. For the best models per predictor configuration, which are colored green, the average decrease in *S* for the test data is only 2.7%. This indicates that we are not overfitting the models to the training data. The 5th degree Polynomial model using wind speed and temperature is most accurate. However, adding wind direction leads to practically the same accuracy (only a slight decrease).

For Producer 1, the nonlinear models clearly outperform the linear models. We should definitely include temperature as a predictor, since it improves 6 out of 7 models. Adding wind direction slightly decreases the accuracy for most models. However, the difference is negligible. This can be explained by a value of 200 for parameter c for the 5th degree Polynomial model using all 3 predictors (Table G.2 in Appendix G). A value of 200 means the cosine function is equal to 1 for all wind directions, so the effect of wind direction is eliminated.

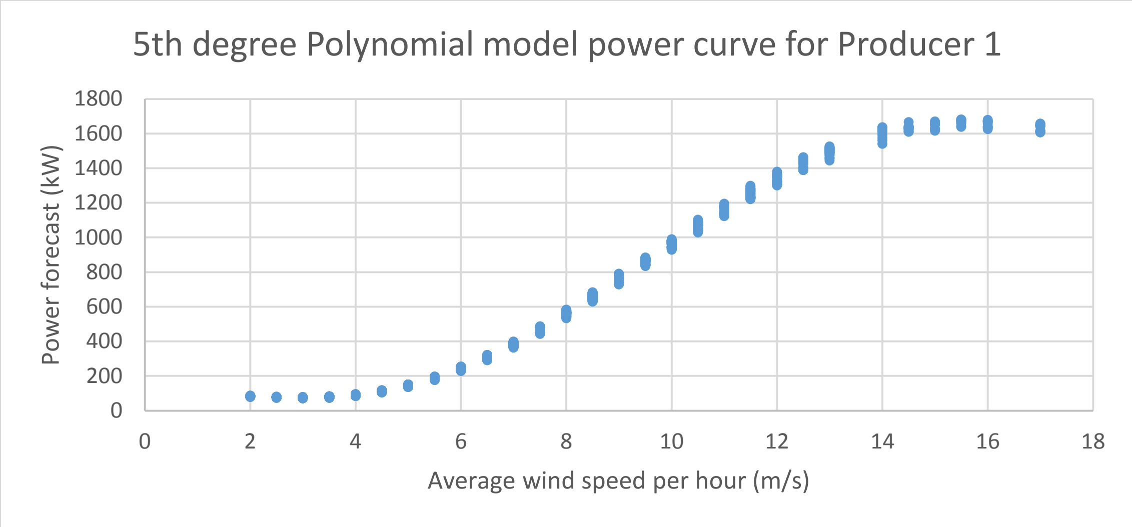

Figure 5.3: Power curve of 5th degree Polynomial model using all 3 predictors for Producer 1 with 1.8 MW rated power using test data.

Figure 5.3 illustrates the effect of temperature and wind direction is little for this model. The spread per wind speed (vertical spread) is small for the 5th degree Polynomial model.

5.2.2 Producer 2

The results for Producer 2 are illustrated in Table 5.4, Producer 2 has 8 turbines with a total rated power of 16 MW.

Predictors	Wind speed		Wind speed + Temperature		Wind speed + Temperature + Wind direction	
Model	Training data (S)	Test data (S)	Training data (S)	Test data (S)	Training data (S)	Test data (S)
Log	2194.95	1716.12	2088.62	1745.06	1949.01	1710.72
3rd Polynomial	1742.66	1867.04	1724.25	1858.85	1625.76	1718.22
4th Polynomial	1731.65	1846.19	1713.29	1837.21	1619.42	1704.63
5th Polynomial	1710.86	1869.69	1692.97	1859.88	1600.12	1719.59
Exponential	1774.36	1901.88	1754.25	1897.71	1654.21	1750.26
Logistic 4	1710.98	1874.17	1693.19	1864.38	1601.09	1722.67
Logistic 5	1712.99	1877.23	1695.29	1867.45	1603.22	1724.92

Table 5.4: Standard error of regression (S) for Producer 2 using different predictors.

The most accurate model for Producer 2 is the 4th degree Polynomial using all 3 predictors. For this model, the value of *S* increases with 5.3% on the test data. This is within acceptable range, which means we are not overfitting. Although we do not use the log model (assumptions violated), it is the second most accurate model based on the test data when using all 3 predictors.

In Figure 5.4, the power curve for the 4th degree Polynomial for Producer 2 using 3 predictors for shows a wide vertical spread. This indicates that the effect of temperature and wind direction is large. This can also be concluded when looking at the parameter values in Appendix G. The value of *c* is approximately 3 for the 4th degree Polynomial model with an ‘optimal’ direction of 181°. A low value of *c* indicates the effect of wind direction is big.

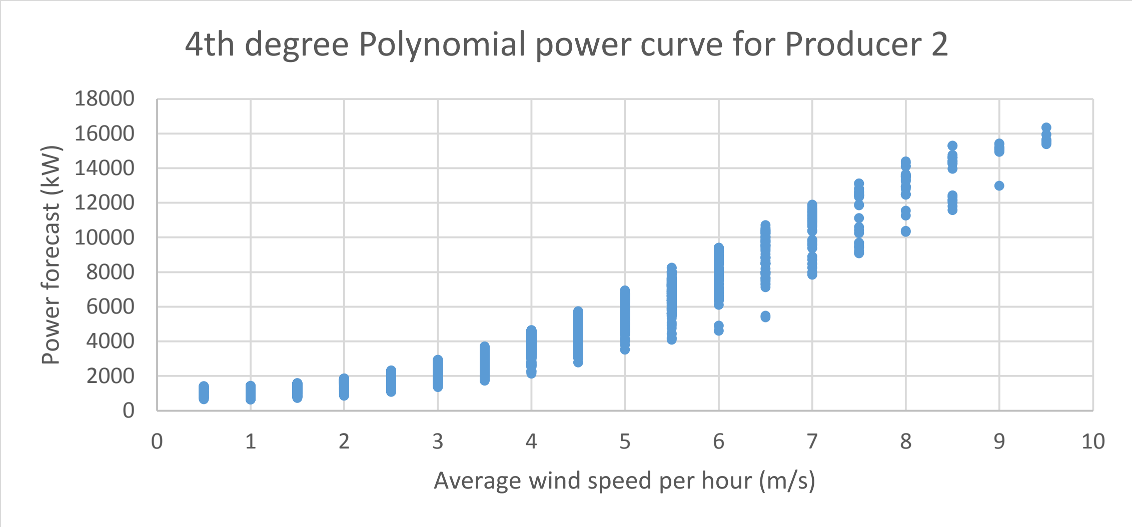

Figure 5.4: Power curve of 4th degree Polynomial model using all 3 predictors for Producer 2 with 16 MW rated power using test data.

Figure 5.4 illustrates that wind speeds do not exceed 9.5 m/s for the test data. This explains why the log model has a good performance for this producer. As wind speeds rise above 12 m/s, the accuracy of the Log model tends to decrease rapidly.

For Producer 2, we should definitely include temperature and wind direction in the regression model. Especially adding wind direction improves the accuracy for all models of Producer 2.

5.2.3 Producer 3

The results for Producer 3 are illustrated in Table 5.5, Producer 3 has 2 turbines with a combined rated power of 6 MW.

Predictors	Wind speed		Wind speed + Temperature		Wind speed + Temperature + Wind direction	
Model	**Training data (S)**	**Test data (S)**	**Training data (S)**	**Test data (S)**	**Training data (S)**	**Test data (S)**
Log	951.85	984.64	981.82	932.02	977.48	895.62
3rd Polynomial	809.35	819.51	807.94	804.73	808.05	805.27
4th Polynomial	808.47	820.07	806.95	805.21	804.35	806.52
5th Polynomial	807.70	818.60	806.13	803.66	803.65	804.83
Exponential	819.22	828.13	838.99	846.44	806.51	808.99
Logistic 4	811.54	829.36	809.95	814.71	807.33	816.26
Logistic 5	811.42	828.25	809.61	813.40	807.15	815.41

Table 5.5: Standard error of regression (S) for Producer 3 using different predictors.

For Producer 3, the difference between the training and test data is minimal, which indicates we are not overfitting. For some models, for example the 3rd degree Polynomial, the test data yield better results than the training data. The 5th degree Polynomial model using wind speed and temperature as predictors is the most accurate model based on the test data for Producer 3.

Adding temperature improved the accuracy on the test data for 6 out of 7 models. So we should definitely use temperature as a predictor for Producer 3. The addition of wind direction is less successful for Producer 3. For all models except the Exponential model, adding wind direction as predictor slightly decreases the accuracy on the test data. Looking at the parameter values of δ and c

in Table G.2 we see that the effect of wind direction is not eliminated. The optimal wind direction for the 5th degree Polynomial is 168° with c=6.76. This value of c means the cosine function gives values between 0.9 and 1.

5.3 Top 3 models using measured weather data

In the previous section, we compared the accuracy on training and test data for 21 models for Producers 1, 2 and 3. For almost all models, the accuracy on test data was not much worse than on the training data. For some models, it was even better. This indicates that we should not be worried about overfitting the models to the training data.

5.3.1 Effect of adding predictors temperature and wind direction

If we evaluate the effect of adding predictors, we observe that adding temperature improves the accuracy for 86% of the models. This suggests that we should definitely use temperature as a predictor. Adding wind direction improved only 48% of the models. However, all models for Producer 2 improved when adding wind direction. Models for Producers 1 and 3 practically have the same accuracy when adding wind direction, the accuracy decreases only slightly. To illustrate the effect of wind direction in the regression models, we compute the values of the cosine functions for the best model of each producer using wind direction. We insert the least squares parameter values in the cosine function for the 4th degree polynomial model for Producer 2 and the 5th degree Polynomial model for Producers 1 and 3. If we recall, the cosine function is defined as:

$$\cos\frac{\theta-\delta}{c} \qquad (5.2)$$

where:

θ = wind direction in radians

δ = parameter that indicates 'optimal' wind direction in radians ($0 < \delta < 2\pi$)

c = parameter that determines the range of the cosine function ($c > 0$)

In Figure 5.5 we illustrate the values of δ in degrees (°), because this is easier to interpret.

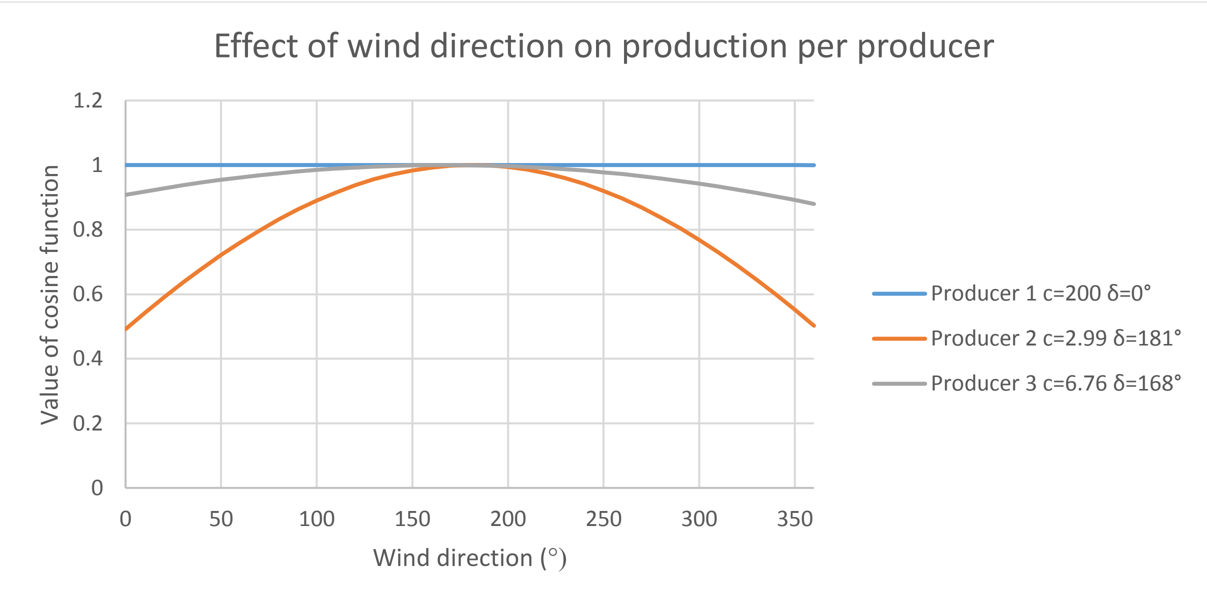

Figure 5.5: Effect of wind direction for the best regression model for Producers 1, 2, and 3.

In Figure 5.5 we can see the effect of wind direction is eliminated for Producer 1. For all wind directions the cosine function returns a value of 1. The optimal wind direction is located at 0°, due to the high value of c=200 we do not see a peak at this wind direction. Producer 3 has an optimal wind direction of 168°, this means wind directions near 360° are punished more heavily than wind directions near 0°. However, due to the reasonably high value of c=6.76 the value for the cosine function at 360° only decreases to 0.88, while the value at 0° is 0.9. This indicates the predicted production value is punished by wind direction with 12% (multiply by 0.88) at the most for Producer 3. The effect of wind direction on production is the greatest for Producer 2. A low value of c=2.99 and an optimal wind direction at 181° lead to a 50% punishment at wind directions 0° and 360°. In Chapter 3 we saw that Producer 2 experiences production losses at wind directions around 120° and around 300°. We would have expected the optimal wind direction to be at 210°, however this would punish the wind directions between 0-120° severely. A disadvantage of our method of wind direction modelling, is that the outer wind directions are most heavily punished. We cannot target a small wind sector without punishing neighboring wind sectors.

5.3.2 Selecting the top 3 regression models

To narrow the 21 models down to the best 3 models, which are used for all 10 producers, we have to choose models that are accurate on the test data for each producer. First, we have to decide which predictors to use. As described above, adding wind direction can be very lucrative and it does not really hurt the performance for any producer since the value of c acts as a 'defense mechanism'. Adding temperature definitely improves the accuracy of the model. Therefore, we use wind speed, temperature and wind direction as predictors. This narrows it down to 7 models. The 5th degree Polynomial model has the best performance in terms of accuracy for Producers 1 and 3, so we definitely include this model in the top 3. For Producer 2, the 4th degree Polynomial model is most accurate. This model also yields good results for the other producers, so the 4th degree Polynomial is also placed in the top 3. The Log model only performs well for Producer 2, but is the worst model by far for the other producers. Also, the linear regression assumptions for the Log model are violated. The Exponential model is the worst model for Producer 2, but performs well for Producers 1 and 3. The Logistic models perform well for all producers, which makes them desirable. Logistic 4 and Logistic 5 are quite similar in terms of accuracy for all 3 producers. The values of S are inconclusive in this case; therefore, we choose the model that has the fewest parameters, which is the Logistic model with 4 parameters. This completes our top 3 models:

Top 3 Models	Function
4th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4) \times \frac{k}{T} \times \left(\lambda \cos \frac{\theta - \delta}{c}\right)$
5th degree polynomial	$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3 + a_4 v^4 + a_5 v^5) \times \frac{k}{T} \times \left(\lambda \cos \frac{\theta - \delta}{c}\right)$
Logistic 4	$P = \left(\alpha \left(\frac{1 + me^{-\frac{v}{\tau}}}{1 + ne^{-\frac{v}{\tau}}}\right)\right) \times \frac{k}{T} \times \left(\lambda \cos \frac{\theta - \delta}{c}\right)$

Table 5.6: Top 3 models based on test data from Producers 1, 2 and 3.

5.4 Day ahead accuracy top 3 models

In the previous section, we selected the 3 most accurate regression models. These are all nonlinear regression models, which means we used the GRG and evolutionary algorithms to estimate the least squares parameters. We established that we should not be worried about overfitting, since the performance on test data was quite similar to the performance on training data for most models. Now that the top 3 models are selected, we know the regression models that most accurately describe the relationship between historical weather measurements and historical production data. We want to remind the reader that we are looking for a model that can translate day ahead weather forecasts into power production forecasts. Therefore, we should insert historical day ahead weather forecasts (hindcasts) into the regression models. This enables the regression models to translate day ahead weather forecasts into day ahead production forecasts. However, before we insert the hindcasts we should adjust the hindcasts for the bias we found in Chapter 3. In Chapter 3 we found statistical evidence that the weather hindcasts at each location (Vlissingen, Hupsel and Lelystad) are biased.

Day ahead forecast	Average error (Forecast – Actual)	Significantly different than 0? (95% confidence level)
Wind speed in Vlissingen	-0.99 m/s	Yes
Wind speed in Hupsel	1.19 m/s	Yes
Wind speed in Lelystad	0.12 m/s	Yes
Wind direction in Vlissingen	7.87°	Yes
Wind direction in Hupsel	16.96°	Yes
Wind direction in Lelystad	18.36°	Yes
Temperature in Vlissingen	-0.38 °C	Yes
Temperature in Hupsel	-0.06 °C	Yes
Temperature in Lelystad	0.09 °C	Yes

Table 5.7: Bias in historical weather forecast data for the locations Vlissingen, Hupsel and Lelystad.

We adjust the weather hindcasts by simply adding the bias to the forecasted value for each hour. For example, the forecasted value of wind speed is 7 m/s in Vlissingen for a particular hour. To adjust for the bias, we add 0.99 m/s since the forecast is structurally 0.99 m/s too low. This means the adjusted forecast value is 7.99 m/s. We apply this logic for wind speed and wind direction in all three locations. We do not adjust the forecasts for the temperature bias, since the temperature is only slightly biased at all three locations. Considering we use temperature in Kelvin in the regression models, the difference of adjusting the temperature can be neglected.

Now that we have adjusted the hindcasts, we insert the adjusted hindcasts into the top 3 regression models. We estimate the least squares parameters for the top 3 regression models for Producers 4-10 and insert the adjusted hindcasts. The hindcast data (see Table 4.5 for example) contains forecasted values of wind speed, wind direction and temperature for each hour (N=4418) of the second half of 2017.

To be able to select the most accurate day ahead forecast model, we calculate the standard error of regression over the entire hindcast dataset (N=4418). The forecast time horizon is 15-38 hours ahead, we do not calculate the standard error of regression for each hour since this would reduce our sample size to $\frac{4418}{24} = 184$ per hour. In Table 5.8 the standard error of regression for the adjusted hindcasts are illustrated.

Producer	4th degree Polynomial (S)	5th degree Polynomial (S)	Logistic 4 (S)
1	230.07	230.87	230.56
2	1751.10	1762.73	1768.91
3	1113.49	1102.92	1102.30
4	222.25	221.06	221.53
5	98.67	96.85	97.51
6	102.41	108.15	107.29
7	113.90	114.16	114.33
8	1256.13	1269.01	1268.54
9	109.54	112.28	112.86
10	621.34	620.83	621.56

Table 5.8: Standard error of regression (S) for adjusted hindcasts of the second half of 2017.

Using hindcasts instead of actual measured data introduces an error, which increases the value of *S*. For Producers 1, 2 and 3 we can see this increase. The values of *S* increased with 46%, 3% and 37% for Producer 1, 2 and 3 respectively. This shows that the error introduced by the weather forecast can be very different for each location.

For each producer, the model with the smallest standard error of regression is highlighted in green. Cleary, the 4th degree Polynomial had the best performance, since it is the best model for 6 out of 10 producers. Also, for the other 4 producers, the 4th degree Polynomial has a similar accuracy as the model with the highest accuracy. Therefore, the 4th degree Polynomial is our best model for the day ahead forecast.

The power curves for the 4th degree Polynomial model of Producers 1, 2, and 3 are illustrated in Appendix H. The power curve for Producer 3 is illustrated below as well.

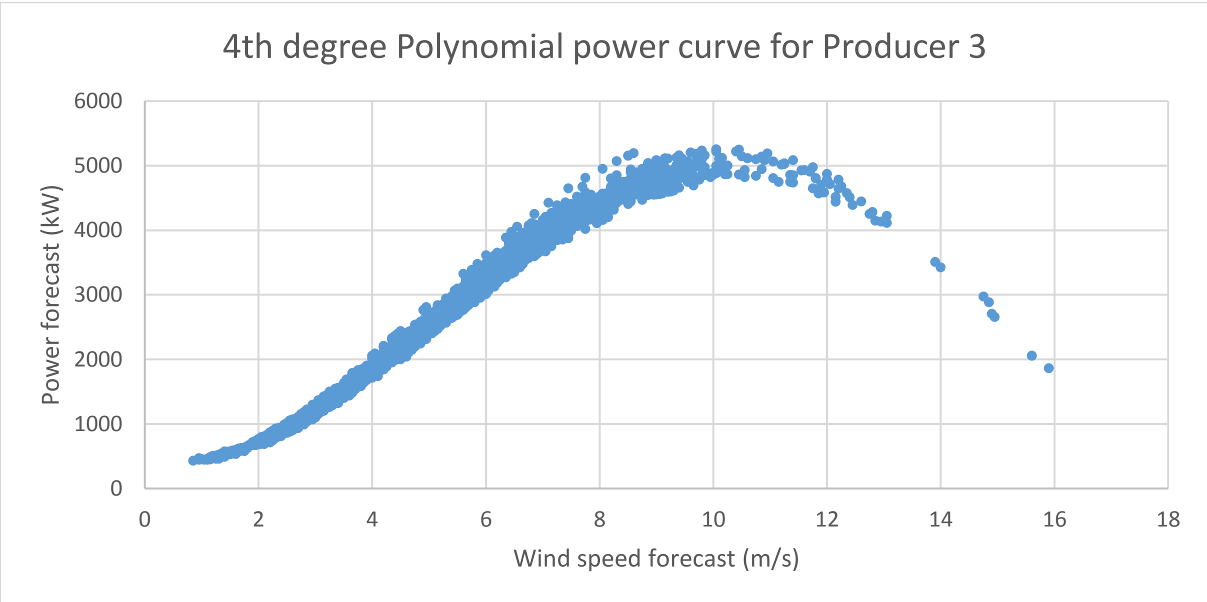

Figure 5.6: Power curve using bias adjusted hindcast data for Producer 3 with 6 MW rated power.

The power curve for Producer 3 is interesting, since the production decreases after rated power is reached. In Figure 2.7, we saw that stall controlled turbines have this property. Therefore, it is likely that Producer 3 has stall controlled turbines.

The standard error of regression for Producer 2 is much larger than for Producer 1. However, Producer 1 has a total rated power of 1.8 MW, while Producer 2 has a rated power of 16 MW. Therefore, we should use a scale-independent metric like Normalized Mean Absolute Percentage Error (NMAPE). If we recall, the NMAPE indicates the mean absolute error as percentage of the total rated power. The 4^{th} degree Polynomial has NMAPE values ranging from 6.8% for Producer 10 to 15.4% for Producer 3.

5.5 Day ahead forecast accuracy of Company X per producer

Now that we know the day ahead accuracy of our regression models, we compare the accuracy of our best model with the accuracy of Company X. We clean the production data by removing observations with production values equal to or smaller than 0. We do this because we want to remove observations that are affected by maintenance, failures, icing and windstorms. In Table 5.9, we compare the accuracy of Company X with the accuracy of the 4^{th} degree Polynomial model. To be able to compute the standard error of regression (S) for Company X, we need to know the number of parameters this model uses. We do not know this; therefore we use the RMSE as performance indicator instead *S.* The RMSE and *S* are very similar, the difference being a punishment of number of parameters for *S*. However, due to the large sample sizes (N>3000 for all producers) the values for RMSE and *S* barely differ.

Producer	RMSE of Company X using bias adjusted hindcast data	RMSE of 4^{th} degree Polynomial using bias adjusted hindcast data
1	193.78	229.82
2	1423.62	1749.12
3	824.91	1112.23
4	174.20	221.99
5	82.58	98.56
6	94.75	102.29
7	90.22	113.77
8	1042.57	1254.71
9	89.09	109.41
10	527.03	620.65

Table 5.9: Root Mean Squared Error (RMSE) for Company X versus the 4^{th} degree Polynomial model using bias adjusted hindcast data.

The most important result from Table 5.9 is that for all producers, Company X is more accurate than the 4^{th} degree Polynomial model. For some producers, the 4^{th} degree Polynomial was almost as accurate as Company X. However, it was not able to beat Company X. For Company X the NMAPE ranges from 5.4% for Producer 10 to 9.4% for Producer 3. For the 4^{th} degree Polynomial model, Producer 10 had the smallest percentage and Producer 3 had the largest percentage as well. This indicates both models struggle with accuracy for Producer 3. Producer 3 has the largest turbines by far, with a hub height of 138.5 m. The 4^{th} degree Polynomial does not include hub height in the forecast model, we are not sure if Company X does so. The discrepancy between wind speed forecasts at heights of 10 m and 138.5 m could be a reason why the 4^{th} degree polynomial is struggling for Producer 3. Since Company X struggles with Producer 3 as well, it is likely that Company X also does not use wind speed forecasts at hub height.

In Section 3.4, we saw that the variance of production is smaller for low wind speeds. This means that production should be easier to predict for low wind speeds. The NMAPE illustrates the average absolute error as a percentage of the rated power in Figure 5.7.

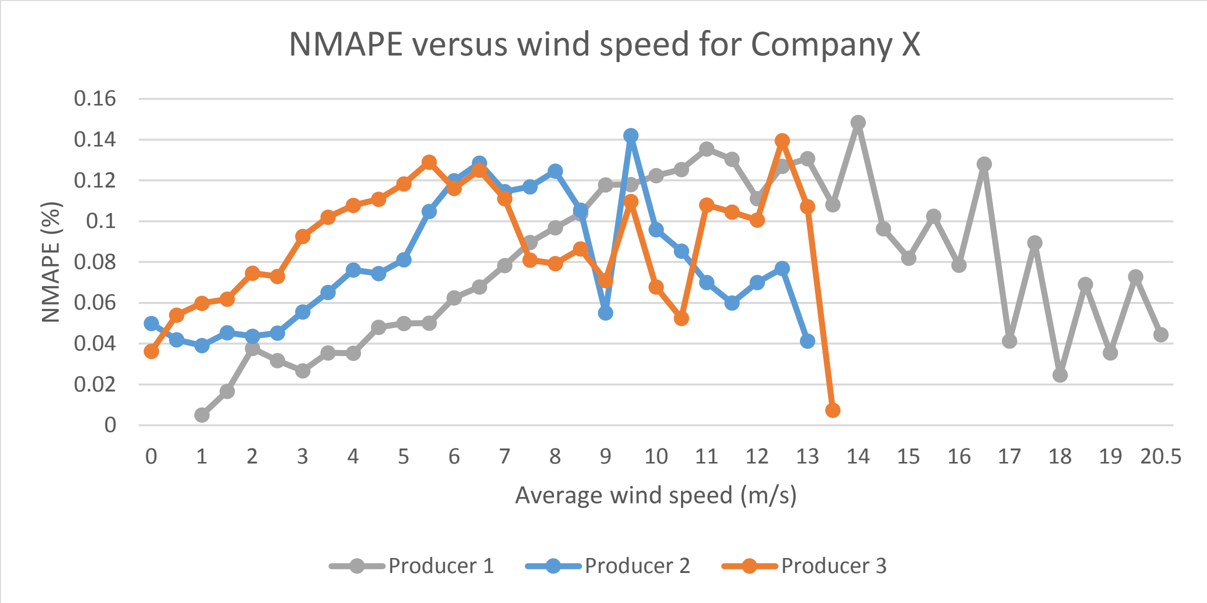

Figure 5.7: Normalized Mean Absolute Percentage Error versus wind speed for Company X for Producers 1, 2 and 3 in the second half of 2017.

For all 3 producers we see that the forecast error is generally smaller when wind speeds are low. As wind speeds increase, the NMAPE increases as well. The maximum NMAPE for Company X are reached at wind speeds of 14 m/s, 9.5 m/s and 12.5 m/s for Producers 1, 2 and 3 respectively. For all producers, NMAPE for Company X tends to decrease when maximum wind speeds are reached. In Chapter 3 we saw that the variance of production is generally smaller for low and high wind speeds, than for moderate wind speeds. This explains the elevated forecast error at moderate wind speeds for Company X.

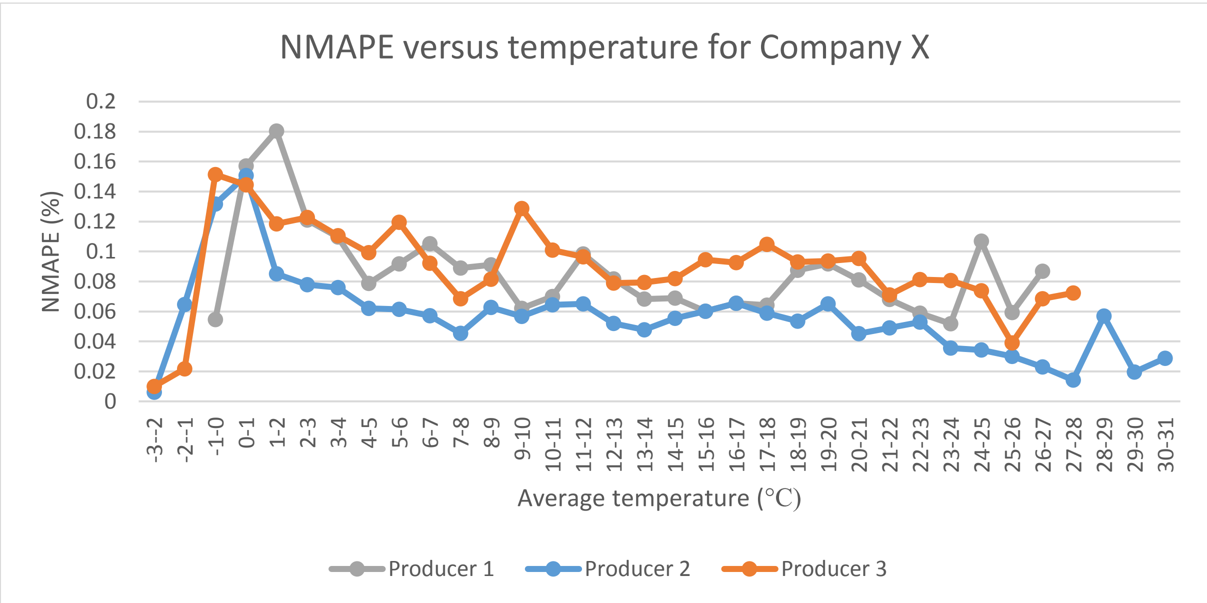

Figure 5.8: Normalized Mean Absolute Percentage Error versus temperature for Company X for Producers 1, 2 and 3 in the second half of 2017.

Company X has the largest NMAPE at temperatures around 0 °C for Producers 1, 2, and 3. Although we cleaned the data, we are not able to remove all instances of icing. Producers 1, 2 and 3 have 2, 8 and 2 turbines respectively. It is possible that one turbine experiences icing, while other another turbine of the same producer is operational. This could explain the NMAPE peak around 0 °C.

Furthermore, as temperatures increase, the NMAPE of Company X for all 3 producers decreases. In Appendix D and Figure 3.10, we saw that wind speeds tend to be lower when temperatures are below 5 °C. For these temperatures, Company X has a larger NMAPE, while wind speeds are generally lower at these temperatures. However, in Figure 5.7 we see that Company X has a smaller NMAPE for low wind speeds. So at low temperatures, the error of Company X is greater, while the wind speed is generally smaller. This should result in a smaller error instead of a larger error. Because this is not the case, it is likely that Company X does not include temperature in the forecast.

5.6 Aggregated day ahead forecast accuracy

In the previous section, we saw that for each producer Company X had a more accurate forecast than our best model. However, during the daily auction DVEP uses aggregated forecasts for their entire portfolio. The advantage of using aggregated forecasts is that negative and positive errors cancel each other out. Therefore, it is important to see how accurate both Company X and the 4th degree Polynomial are when aggregating the forecasts. In this section, we evaluate the day ahead accuracy by aggregating the production and day ahead forecasts of all 10 producers for each hour. The total rated power for the 10 producers is 45.65 MW. First, we evaluate the average forecast errors for the entire second half of 2017. After this, we evaluate the forecast error per hour of the forecast horizon. Lastly, we look at the forecast error per month.

The results for both models for the aggregation of 10 producers with a total rated power of 45.65 MW are illustrated in Table 5.10.

Model	Standard Error of Regression (S) in kW	Root Mean Squared Error (RMSE) in kW	Mean Absolute Error (MAE) in kW	Normalized Mean Absolute Percentage Error (NMAPE)
Company X	3,338	3,335	2,407	5.3%
4th degree Polynomial	4,166	4,162	3,092	6.8%

Table 5.10: Aggregated day ahead forecast errors with a total rated power of 45.65 MW in the second half of 2017.

For all 4 performance indicators, the model from Company X is more accurate, since the errors are smaller. The values of *S* and RMSE are almost the same, this is due to the large sample size. RMSE and *S* are greater than MAE, this can be expected because RMSE and *S* use squared errors, which punish large errors more severely. On average, Company X has an absolute error of 2,407 kW while the 4th degree Polynomial model has an absolute error of 3,092 kW on average. This means Company X is 685 kW per hour more accurate on average, the MAE is 22.2% smaller than the 4th degree Polynomial model. The RMSE of Company X is 19.8% smaller than the 4th degree Polynomial for the aggregated day ahead forecast. The NMAPE over the period of 6 months for Company X is 1.5% smaller than for the 4th degree Polynomial model.

Company X has a NMAPE of 5.3% and the 4th degree Polynomial has 6.8% for the aggregated forecast. In Section 5.4, we saw the percentages for Producers 1-10 for Company X ranged between 5.4% and 9.4%. For the 4th degree Polynomial the percentages ranged between 6.8% and 15.4%. For both models, the error decreased for the aggregated forecast , which is due to negative and positive errors canceling out.

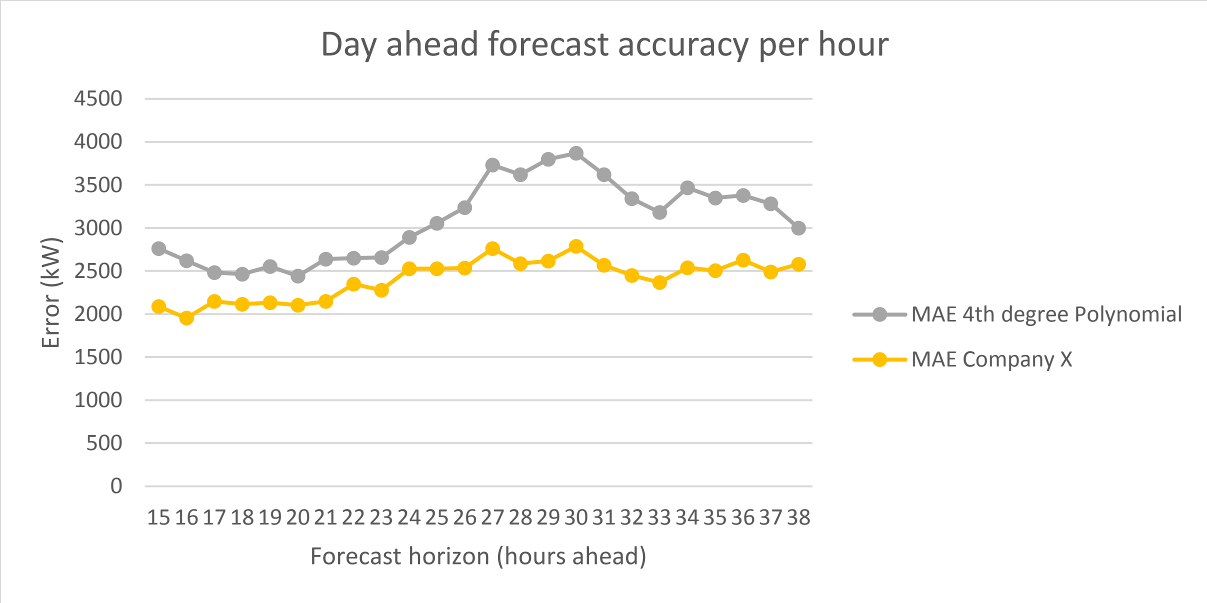

Figure 5.9: Aggregated day ahead mean absolute error per hour with a total rated power of 45.65 MW in the second half of 2017.

In Figure 5.9, we plot the MAE for both models for each hour in the forecast horizon. As the forecast horizon increases, the forecast error increases as well. This is supported by our findings in Section 3.5, which indicated a larger variance of the wind speed forecast error as the time horizon increased. Greater errors in wind speed forecasts lead to larger production forecast errors. Company X is more accurate than the 4th degree Polynomial model for the entire forecast horizon. Both forecast models seem to struggle between the forecast horizon of 25-33 hours ahead, the 4th degree Polynomial more so than Company X.

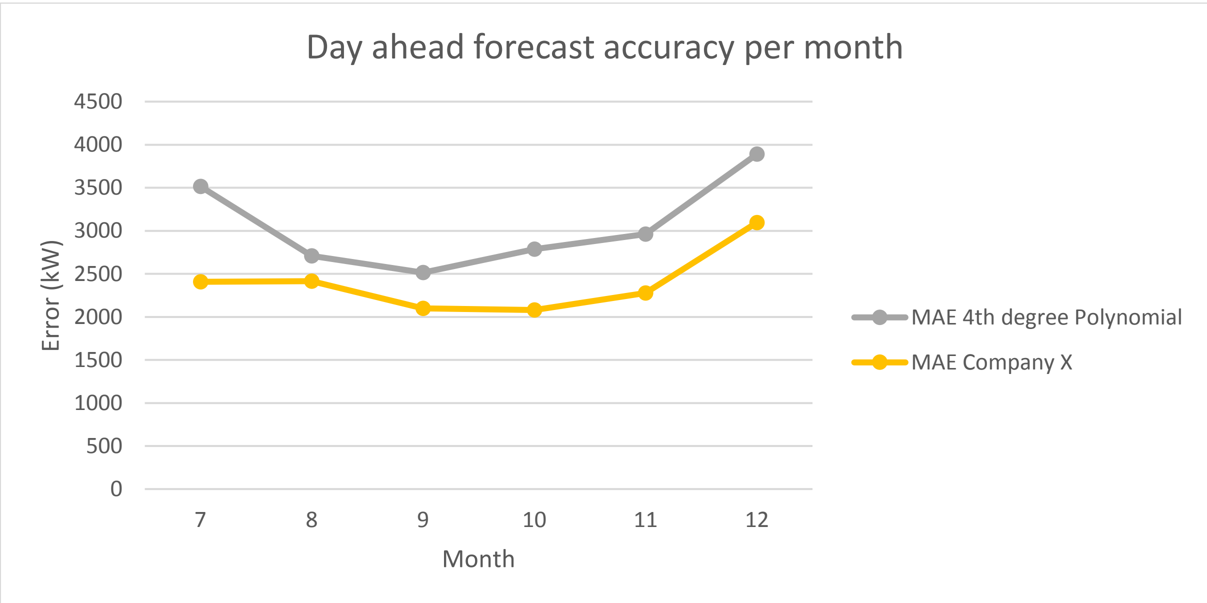

Figure 5.10: Aggregated day ahead mean absolute error per month for a total rated power of 45.65 MW in the second half of 2017.

If we look at the forecast error per month in Figure 5.10, we see that both models show the same pattern. Month 7 and especially month 12, have the largest error for both models. Month 12 usually has lower temperatures, in Figure 5.8 we saw this could explain the increased error for Company X. In month 12, the absolute error per hour for Company X was on average approximately 1 MW larger than in month 9 or 10. However, we must add that the 4th degree Polynomial model had difficulties the same months as Company X.

5.7 Conclusion

In this chapter, we analyzed the results of the research. First, we tested the assumptions of the linear regressions models. After this, we looked at which regression models using which predictors could most accurately describe the relationship between historical weather measurements and historical production data. For the 7 models selected in Chapter 4, we evaluated accuracy on training and test data using 3 predictor configurations per model. This resulted in 21 different model configurations, from which a top 3 was selected based on test data accuracy. For the top 3 models, we evaluated the day ahead forecast accuracy on all 10 producers using bias adjusted weather forecasts. The most accurate day ahead production forecast model was compared with Company X. We compared the accuracy of our best model and Company X for all 10 producers separately and aggregated. Chapter 5 leads to the following conclusions:

- The linear regression assumption of normality of errors is violated for all linear regression models of Producers 1, 2, and 3. Assumptions of homoscedasticity and autocorrelation are violated as well; this leads to inaccurate linear regression models we should not use.
- Adding temperature as a predictor improved the performance for 18 of the 21 models, which is 86%. Therefore, we use temperature as a predictor in our regression model.
- Adding wind direction is more producer-dependent. All models for Producer 2 improve when adding wind direction. Models for Producers 1 and 3 practically have the same accuracy when adding wind direction. Therefore, we use wind direction as a predictor in our regression model.
- The top 3 regression models based on historical weather and production data are the 4th degree Polynomial, 5th degree Polynomial and the Logistic 4 model. All 3 models use wind speed, temperature and wind direction as predictors.
- The most accurate day ahead forecasting model using bias adjusted day ahead weather forecasts is the 4th degree Polynomial model.
- The day ahead production forecasts of Company X are more accurate than the 4th degree Polynomial model. For all 10 producers the standard error of regression is smaller for Company X.
- For Company X the NMAPE per producer ranges from 5.4% to 9.4%. For the 4th degree Polynomial model these percentages are between 6.8% and 15.4%. For both models, Producer 10 has the smallest NMAPE and Producer 3 the largest.
- Company X has the smallest day ahead forecast error for Producers 1, 2 and 3 when wind speeds are low.
- Company X has the largest day ahead forecast error for Producers 1, 2 and 3 when temperatures are around the freezing point. Day ahead forecast errors for Company X decrease as the temperature increases.
- The aggregated day ahead forecast for all 10 producers of Company X is more accurate than the 4th degree Polynomial. During the last 6 months of 2017, Company X has a MAE of 2,407 kW per hour while the 4th degree Polynomial model has a MAE of 3,092 kW. This means that the average absolute error per hour of Company X is 685 kW smaller, which is 1.5% of the total rated power of 45.65 MW.
- Company X is more accurate for all hours in the 15-38 hours ahead forecast horizon.

We conclude that Company X is more accurate at day ahead power production forecasting than our 4th degree Polynomial model.

6. Conclusion and Recommendations

This final chapter concludes this research and answers the research question in Section 6.1. In Section 6.2 we discuss the limitations of our research. We propose several recommendations to DVEP in Section 6.3 and some suggestions for further research are given in Section 6.4.

6.1 Conclusion

The core problem for this project is that the current day ahead forecast for power production of wind turbines is believed to be inaccurate. DVEP currently buys this forecast from Company X, which is costly. DVEP thinks that a forecasting model can be developed that is more accurate than Company X. DVEP is especially interested in the time horizon between 15-38 hours ahead, since this is the time horizon that is used for the day ahead auction. The research goal is to develop a day ahead forecasting model for the power production of wind turbines of DVEP producers. This model should be more accurate than the model that is currently used. The research question we want to answer in this research is:

*"**How to develop a model that is able to translate day ahead weather forecasts into power production forecasts for wind turbines of DVEP producers with higher accuracy than Company X?"***

We conclude that the newly developed model using nonlinear least squares regression does not have a higher accuracy than the current model from Company X. The model from Company X has a higher day ahead forecast accuracy for all producers included in our research. Also, the aggregated day ahead forecast of Company X is more accurate for all hours in the 15-38 hours ahead forecast horizon.

The top 3 forecast models using historical weather and historical production data are the 4th degree Polynomial, 5th degree Polynomial and the Logistic 4 model. All 3 models use nonlinear regression with wind speed, temperature and wind direction as predictors. Wind speed is the most important predictor by far. Adding temperature multiplicatively as a predictor to the forecasting models improved the accuracy for 86% of the models. Adding wind direction multiplicatively as a predictor improved the accuracy for only 48% of the models. However, it is very producer-dependent. For Producer 2, adding wind direction as a predictor improves the accuracy, while the accuracy for Producers 1 and 3 is practically identical.

The forecast model that has the highest day ahead forecast accuracy is the 4th degree Polynomial model. For 60% of producers the 4th degree Polynomial model has the smallest standard error of regression (S). For the other 40% of producers the 4th degree Polynomial model has practically the same accuracy as the best performing model. The day ahead forecast of the 4th degree Polynomial is less accurate than Company X's forecast for all producers included in this research. Company X has Normalized Mean Absolute Percentage Errors (NMAPEs) per producer between 5.4% and 9.4%, while these percentages are between 6.8% and 15.4% for the 4th degree Polynomial. Producer 10 has the smallest NMAPE for both models, while Producer 3 has the largest NMAPE for both models. Nevertheless, Company X has a smaller NMAPE and Root Mean Squared Error (RMSE) for all producers.

The aggregate day ahead forecasts of Company X are more accurate. Company X has an NMAPE of 5.3%, while the 4th degree Polynomial has an NMAPE of 6.8%. The aggregated day ahead forecast from Company X has a Mean Absolute Error (MAE) that was 22.2% smaller than the 4th degree Polynomial. The RMSE of the aggregated day ahead forecast is 19.8% smaller for Company X.

The evaluation of the day ahead forecast error of Company X for Producers 1, 2 and 3 reveals that Company X is more accurate when wind speeds are low. This can be expected, due to low production values for low wind speeds. The evaluation of the forecast error versus temperature reveals that the forecast error increases as temperature decreases. For low temperatures, the wind speed is generally lower, which should result in smaller forecast errors. However, the forecast error increases as temperature decreases. This suggests that the forecast of Company X could be improved for low temperatures.

6.2 Limitations of our research

During our research several problems with the data arose, which indicates that this research has some limitations. These limitations are due to the lack of data or due to inaccurate data. Lacking data forces us to look for alternatives, like using temperature instead of air density as a predictor. Inaccurate data introduces noise to the input data, which hurts the accuracy of the forecast.

The first limitation concerns the historical weather data. The weather data are measured at KNMI weather stations located 8.6 km from the turbine site on average. This introduces errors in the weather data. Wind speed, wind direction and temperature could have been very different at the turbine site at the time of measurement. Wind speed and direction are measured at a height of 10 m. The hub height of wind turbines ranges from 36 to 138.5 m in this research. We did not adjust for hub height in our models, since the weather forecasts give wind speed and direction forecasts at 10 m height as well. The temperature was measured at 1.5 m and also forecast at this height. Temperatures at hub height are generally lower.

The second limitation of our research is related to the weather forecasts. Forecasts of air density or air pressure are not available. Literature suggests that the power production of wind turbines is connected to the air density. Air density is governed by air pressure, temperature and altitude. Unfortunately, we were not able to use air density as a predictor; we had to settle for temperature as a predictor.

We found statistical evidence that the weather forecasts used in this research are biased. Wind speed, wind direction and temperature forecasts for all 3 locations included in this research are biased. We adjusted the forecasts for the bias we found. A negative bias (systematic low forecast) was adjusted by increasing the forecast, a positive bias by decreasing the forecast. However, the fact that the weather forecast contains a bias is troubling.

Another limitation of this study is related to the cleaning of the data. By cleaning the data, we hoped to remove the effects of downtime due to icing, failures, maintenance and windstorms. Since a lot of producers have multiple turbines connected to one grid connection, we cannot distinguish the number of turbines that were operational during an hour. We tried to clean the data by removing outliers per wind speed bin. We believe this removed most of the disturbing factors mentioned above, but it also removed some data that were not affected by these factors. It is impossible to distinguish whether the data were affected by these factors or not, especially because average wind speed was measured at 8.6 km distance on average.

Thanks to the critical points of Maarten Vinke from DVEP, we found some unnecessary parameters in the nonlinear regression models. He pointed out that the parameter λ could have been left out and that the parameter k could have been fixed at 352.98375 so $\frac{k}{T}$ equals the standard air density (1.225 kg/m^3 at 15 °C) at mean sea level. If we use the 3rd degree Polynomial model as example, these changes would lead to the following equation:

$$P = (a_0 + a_1 v + a_2 v^2 + a_3 v^3) \times \frac{352.98375}{T} \times \left(\cos \frac{\theta - \delta}{c}\right) \quad (6.1)$$

We have calculated the parameters for this nonlinear regression model for producers 1, 2 and 3. The parameter values for δ and c did not change, values for a_0, a_1, a_2 and a_3 did change. For all models, the value of standard error of regression changed only very slightly and would have not lead to different decisions. However, the unnecessary parameters are troubling and should be removed in all models. Unfortunately, we only found this error at the end stages of this project.

Lastly, the parameters of the regression models were estimated using measured weather data instead of weather forecasts. This choice was made because we only have historical weather forecasts for the second half of 2017. We believe that this period is too short to accurately capture the relationship between the weather forecasts and production.

6.3 Recommendations

Based on the conclusion and limitations of the research discussed in Section 6.1 and 6.2, we give some recommendations to DVEP.

We recommend DVEP to keep outsourcing the production forecast for wind turbines to Company X for the time being. The current data infrastructure does not allow for an accurate forecast to be developed by DVEP. In order to develop an accurate forecast model in-house, weather data would have to be collected at the wind turbine sites at hub height. Also, weather forecasts should be specified to the wind turbine location, which means the weather forecast models should have to be used with (preferably) a small grid size. Weather forecasts should be adjusted to hub height and should include wind speed, air density and wind direction. However, since DVEP is interested in the aggregated forecast for the entire portfolio, this forecasting approach is not practical. Installing the equipment for all producers is time-consuming and very costly. After the equipment has been installed, additional set-up time is necessary because data needs to be collected. Because of the portfolio size of DVEP, developing a forecast model in-house would be impractical. Also, to what extent this improves the current forecasting method is uncertain. Therefore, we recommend that DVEP keeps outsourcing the production forecast for all wind turbines in its portfolio to Company X for the time being.

Unfortunately, we were not able to develop a more accurate day ahead forecast model than the model of Company X. However, when evaluating the day ahead forecast performance of Company X, we did see some room for improvement. The day ahead forecast performance of Company X could be improved for low temperatures. Generally, wind speeds are lower for low temperatures, which should be accompanied by smaller forecast errors. However, for Company X the forecast errors are larger as temperatures decrease. Therefore, we believe that Company X could improve its forecast performance for low temperatures. Literature suggests that using air density as predictor should improve the forecast performance. Furthermore, we encourage DVEP to inquire about the input of Company X's forecast model. Knowing the input variables enables DVEP to adjust the forecast more accurately.

We recommend DVEP to use the Root Mean Squared Error (RMSE) as the scale dependent performance indicator. The Mean Absolute Error (MAE) can be used additionally, because it is easy to interpret. The Normalized Mean Absolute Percentage Error (NMAPE) should be used as scale-independent performance indicator. When evaluating forecast performance, DVEP should inquire about the forecast horizon. Third parties tend to give forecast errors for their models without specifying the forecast horizon.

Lastly, we discourage DVEP to use the weather forecasts that were used in this research. Wind speed, wind direction and temperature forecasts are biased for all 3 locations used in the research. To prevent systematic forecast errors, we recommend that DVEP does not use the weather forecasts that were used in the research.

6.4 Suggestions for further research

In this research, we used measured weather data as training data. Another approach for further research would be to use hindcasts as training data. Unfortunately, we did not have enough hindcast data for a proper sample size.

The goal of this research was to develop an accurate forecast model to limit the energy imbalance. The financial aspect of the forecast was not included in the research. In some situations it can be lucrative to have imbalance, because market prices are favorable. Further research should be done to investigate when these situations occur and how this can be incorporated into a forecast model.

We proposed a new method of modeling wind direction in regression analysis. This yielded an improvement for one producer, while the accuracy of other producers practically remained the same. In Section 6.2 we pointed out that we used an unnecessary parameter with the wind direction modeling. Further research is needed into wind direction modeling in regression analysis to determine the best way to incorporate wind direction into forecast models.

Due to a lack of data, temperature was used in the forecast model instead of air density. According to the literature, air density should be used to predict power production of wind turbines. Therefore, we would have preferred to use air density as a predictor. In Section 6.2 we pointed out that we could have used a fixed value for the parameter k to insert the standard air density at mean sea level. Further research should be conducted to study the effect of using air density as a predictor and how to incorporate air density in a regression model.

Lastly, we suggest that further research is done with respect to intraday power production forecasts for wind turbines. We focused on a forecast horizon of 15-38 hours. The forecast horizon of 0-15 hours is not included. Further research can provide insight into how accurate these production forecasts are and which models should be used for this forecast horizon.

Bibliography

Allua, S., & Thompson, C. B. (2009). Hypothesis Testing. *Air Medical Journal*, 108-110.

Barati, R. (2013). Application of Excel Solver for Parameter Estimation of the Nonlinear Muskingum Models. *Korean Journal of Civil Engineering, 17*, 1139-1149.

Bates, D. M., & Watts, D. G. (2008). *Nonlinear Regression Analysis and Its Applications.* New York: John Wiley and Sons, Inc.

Benke, K. K., Lowell, K., & Hamilton, A. (2008). Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model. *Mathematical and Computer Modelling, 47*, 1134-1149.

Burton, T., Sharpe, D., Jenkins, N., & Bossanyi, E. (2001). *Wind Energy Handbook.* Chichester: John Wiley and Sons Ltd.

Carrillo, C., Obando Montano, A. F., Cidras, J., & Diaz-Dorado, E. (2013). Review of power curve modelling for wind turbines. *Renewable and Sustainable Energy Reviews, 21*, 572-581.

DVEP Energie. (2017, September 1). *De Vrije Energie Producent*. Retrieved from dvep.nl: https://www.dvep.nl/Over-ons

Enercon. (2015). *ENERCON product overview.* Aurich.

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

Frontline Solvers. (2018, Januari 5). *Excel Solver - Multistart Methods for Global Optimization*. Retrieved from Solver: https://www.solver.com/excel-solver-multistart-methods-global-optimization

Frost, J. (2014, March 6). *Why Is There No R-Squared for Nonlinear Regression?* Retrieved from Minitab.com: http://blog.minitab.com/blog/adventures-in-statistics-2/why-is-there-no-r-squared-for-nonlinear-regression

Frost, J. (2017, March 25). *Standard Error of Regression vs. R-squared*. Retrieved from statisticsbyjim.com: http://statisticsbyjim.com/regression/standard-error-regression-vs-r-squared/

Giebel, G. (2003). *The State-of-the-Art Short-Term Prediction of Wind Power: A literature Overview.* Roskilde: Riso National Laboratory.

Hau, E. (2013). *Wind Turbines: Fundamentals, Technologies, Application, Economics.* Heidelberg: Springer.

Hoshmand, A. R. (2009). *Business Forecasting: A practical approach.* New York: Routledge.

Hu, W., Xie, J., Chau, H. W., & Si, B. C. (2015). Evaluation of parameter uncertainties in nonlinear regression using Microsoft Excel Spreadsheets. *Environmental Systems Research, 4*. doi:10.1186/s40068-015-0031-4

Hyndman, R. J. (2014). *Measuring forecast accuracy.*

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting, 22*, 679-688.

Jarque, C. M., & Bera, A. K. (1987). A Test for Normality of Observations and Regression Residuals. *International Statistical Review, 2*, 163-172.

King, M. L. (1983). The Durbin-Watson Test for Serial Correlation. *Journal of Econometrics, 21*, 357-366.

KNMI. (2000). *Handboek Waarnemingen: Meetstation algemeen.* De Bilt: KNMI.

KNMI. (2014). *Het weer in Nederland: Toelichting maandoverzicht van het weer in Nederland.* De Bilt: KNMI.

Landberg, L., Giebel, G., Aalborg Nielsen, H., Nielsen, T., & Madsen, H. (2003). Short-Term Prediction - An Overview. *Wind Energy, 6*, 273-280.

Lange, M., & Focken, U. (2006). *Physical Approach to Short-Term Wind Power Prediction.* Berlin: Springer-Verlag Berlin Heidelberg.

Liese, F., & Miescke, K. (2008). *Statistical Decision Theory: Estimation, Testing, and Selection.* Berlin: Springer Berlin.

Lydia, M., Immanuel Selvakumar, A., Suresh Kumar, S., & Edwin Prem Kumar, G. (2013). Advanced Algorithms for Wind Turbine Power Curve Modeling. *IEEE Transactions on Sustainable Energy, 4*, 827-835.

Lydia, M., Suresh Kumar, S., Immanuel Selvakumar, A., & Edwin Prem Kumar, G. (2014). A comprehensive review on wind turbine power curve modeling techniques. *Renewable and Sustainable Engery Reviews, 30*, 452-460.

Lynn, P. (2012). *Onshore and Offshore Wind Energy.* Chichester: John Wiley and Sons Ltd.

Manwell, J., McGowan, J., & Rogers, A. (2009). *Wind Energy Explained: Theory, Design and Application.* Chichester: John Wiley and Sons Ltd.

Marciukaitis, M., Zutautaite, I., Martisauskas, L., Joksas, B., Gecevicius, G., & Sfetsos, A. (2017). Non-linear regression model for wind turbine power curve. *Renewable Energy, 113*, 732-741.

Schaffarczyk, A. P. (2014). *Introduction to Wind Turbine Aerodynamics.* Heidelberg: Springer.

Spiess, A., & Neumeyer, N. (2010). An evaluation of R-squared as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. *BMC Pharmacology, 10*. doi:10.1186/1471-2210-10-6

Thadewald, T., & Buning, H. (2007). Jarque-Bera Test and its Competitors for Testing Normality - A Power Comparison. *Journal of Applied Statistics, 1*, 87-105.

Thapar, V., Agnihotri, G., & Krishna Sethi, V. (2011). Critical analysis of methods for mathematical modelling of wind turbines. *Renewable Energy, 36*, 3166-3177.

Vestas. (2017, September 28). *V110-2.0 MW at a glance*. Retrieved from vestas.com: https://www.vestas.com/en/products/turbines/v110-2_0_mw#!aep

Wang, X., Guo, P., & Huang, X. (2011). A Review of Wind Power Forecasting Models. *Energy Procedia, 12*, 770-778.

Appendix A: KNMI weather stations in the Netherlands

Figure A.1: Map of KNMI weather stations in the Netherlands.

Appendix B: Raw production data for Producers 1 and 2

Wind speed	Minimum	Average	Maximum
0	0%	2%	16%
0,5	0%	2%	29%
1	0%	5%	40%
1,5	0%	7%	59%
2	0%	11%	58%
2,5	0%	13%	56%
3	0%	17%	75%
3,5	0%	19%	83%
4	2%	24%	84%
4,5	3%	31%	91%
5	4%	38%	93%
5,5	3%	46%	95%
6	9%	55%	100%
6,5	22%	65%	100%
7	19%	70%	100%
7,5	29%	73%	99%
8	18%	76%	100%
8,5	47%	81%	100%
9	44%	85%	100%
9,5	56%	90%	100%
10	74%	95%	100%
10,5	85%	96%	100%
11	93%	99%	100%
11,5	91%	97%	100%
12	88%	98%	101%
12,5	93%	97%	100%
13	89%	96%	100%
13,5	97%	99%	100%
14	98%	98%	98%

Table B.1: Minimum, average and maximum production per wind speed for Producer 2.

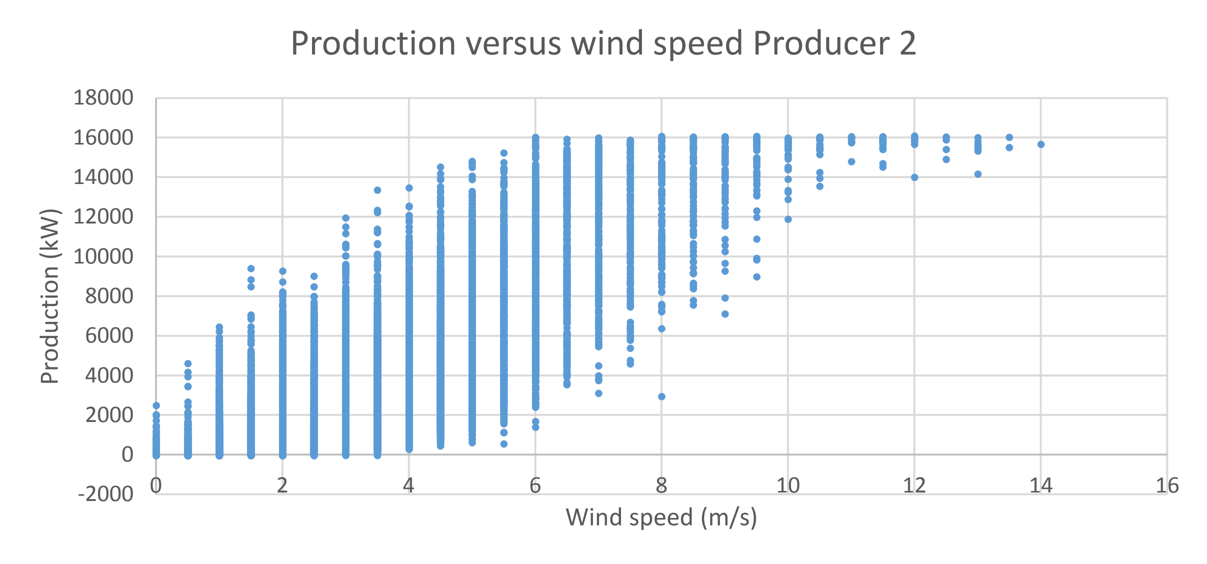

Figure B.1: Scatterplot of production per wind speed in 2015-2016 for Producer 2 with 16 MW combined rated power.

Wind speed	Minimum	Average	Maximum
0,5	0%	0%	0%
1	0%	0%	3%
1,5	0%	0%	6%
2	-1%	0%	22%
2,5	0%	1%	11%
3	0%	2%	16%
3,5	0%	3%	17%
4	-1%	4%	26%
4,5	0%	6%	39%
5	0%	8%	44%
5,5	0%	10%	45%
6	0%	13%	52%
6,5	0%	17%	75%
7	2%	22%	64%
7,5	0%	28%	78%
8	3%	32%	73%
8,5	8%	39%	81%
9	6%	42%	88%
9,5	10%	48%	91%
10	12%	53%	86%
10,5	23%	60%	90%
11	22%	64%	93%
11,5	17%	70%	96%
12	29%	72%	94%
12,5	36%	78%	98%
13	43%	80%	99%

13,5	43%	84%	100%
14	40%	86%	101%
14,5	44%	87%	100%
15	79%	92%	100%
15,5	48%	91%	101%
16	80%	93%	101%
16,5	55%	88%	99%
17	84%	92%	98%
17,5	50%	89%	96%
18	87%	93%	99%
18,5	91%	93%	97%
19	87%	92%	96%
19,5	86%	90%	98%
20	50%	86%	93%
20,5	79%	88%	95%
21	23%	80%	90%
21,5	84%	87%	90%
22	76%	82%	87%
22,5	56%	71%	86%
23	68%	74%	81%

Table B.2: Minimum, average and maximum production per wind speed for Producer 1.

Appendix C: Production versus wind direction

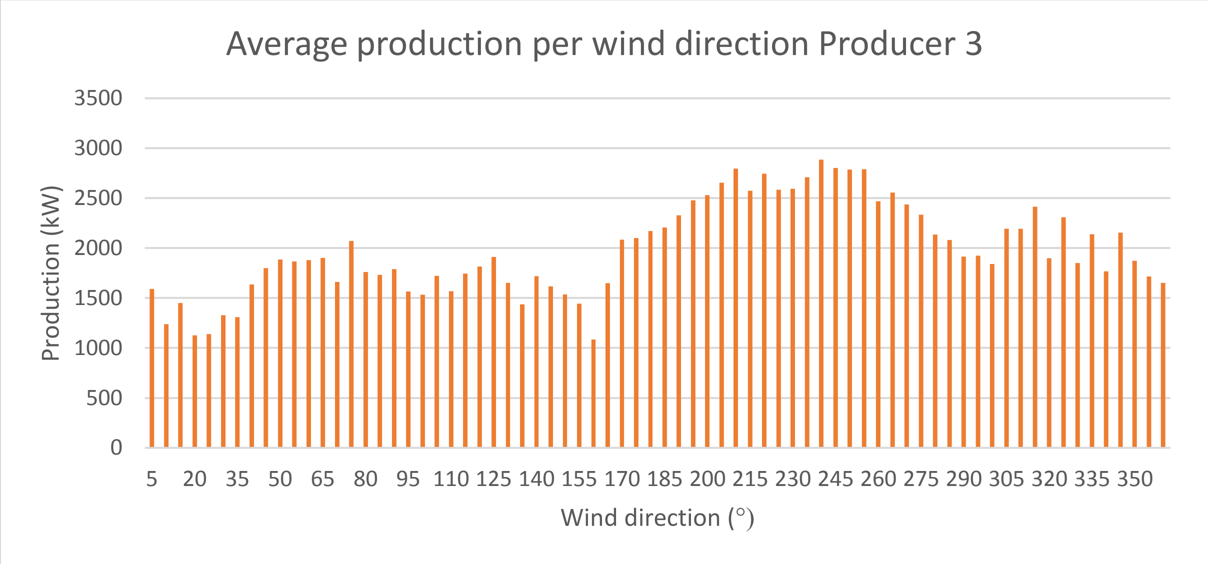

Figure C.1: Average production per wind direction in 2015-2016 for Producer 3 with 16 MW combined rated power.

Appendix D: Production and wind speed versus temperature

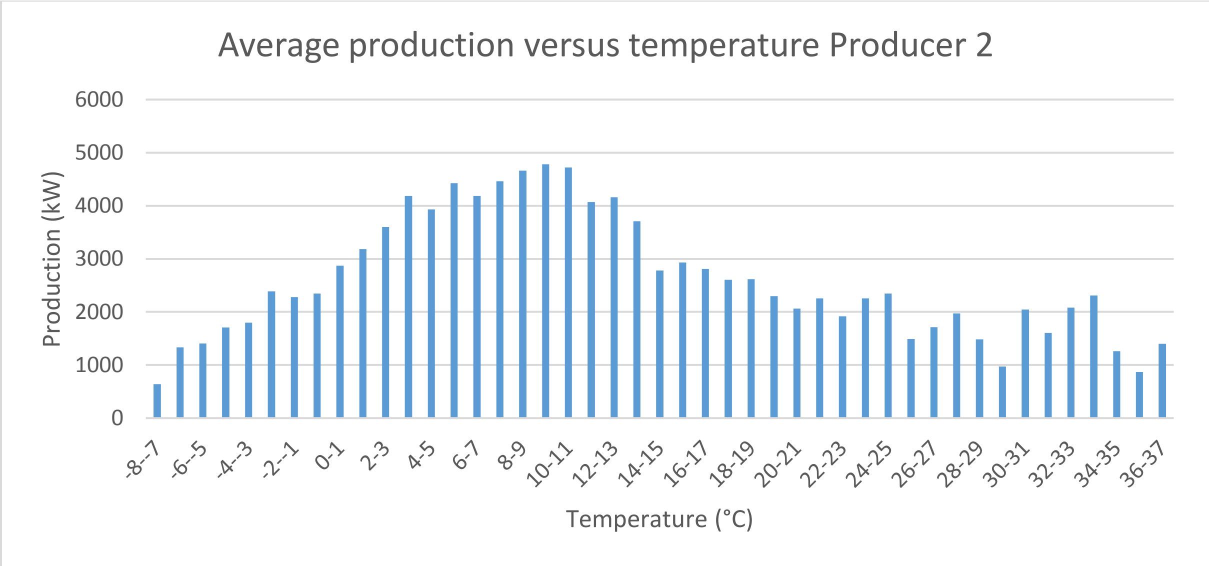

Figure D.1: Average production versus temperature in 2015-2016 for Producer 2 with 16 MW combined rated power.

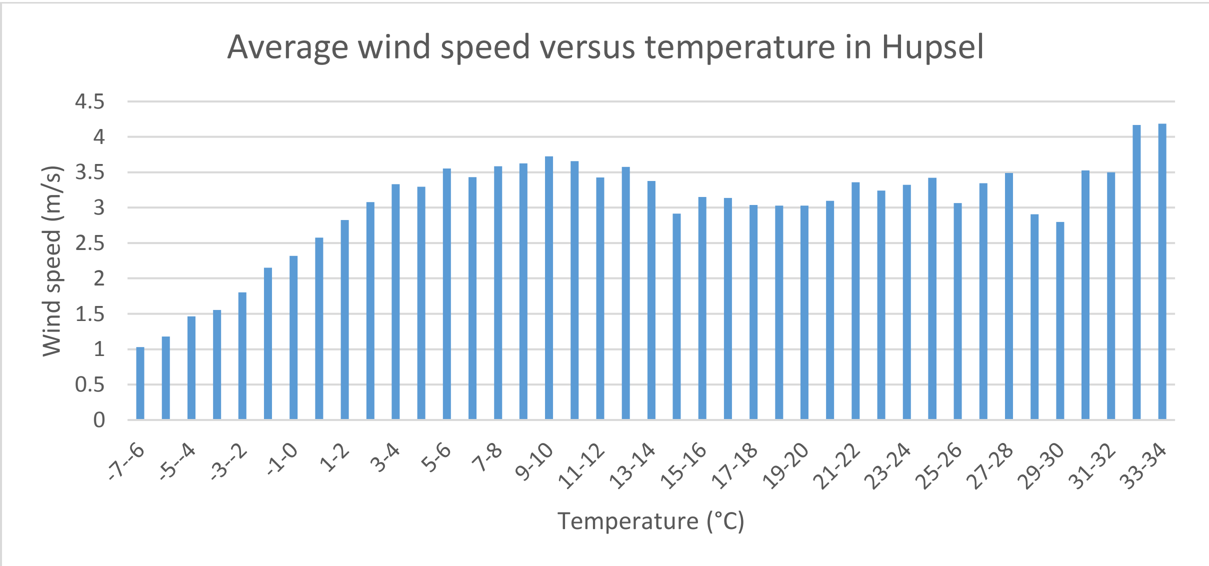

Figure D.2: Average wind speed versus temperature in 2015-2016 in Hupsel.

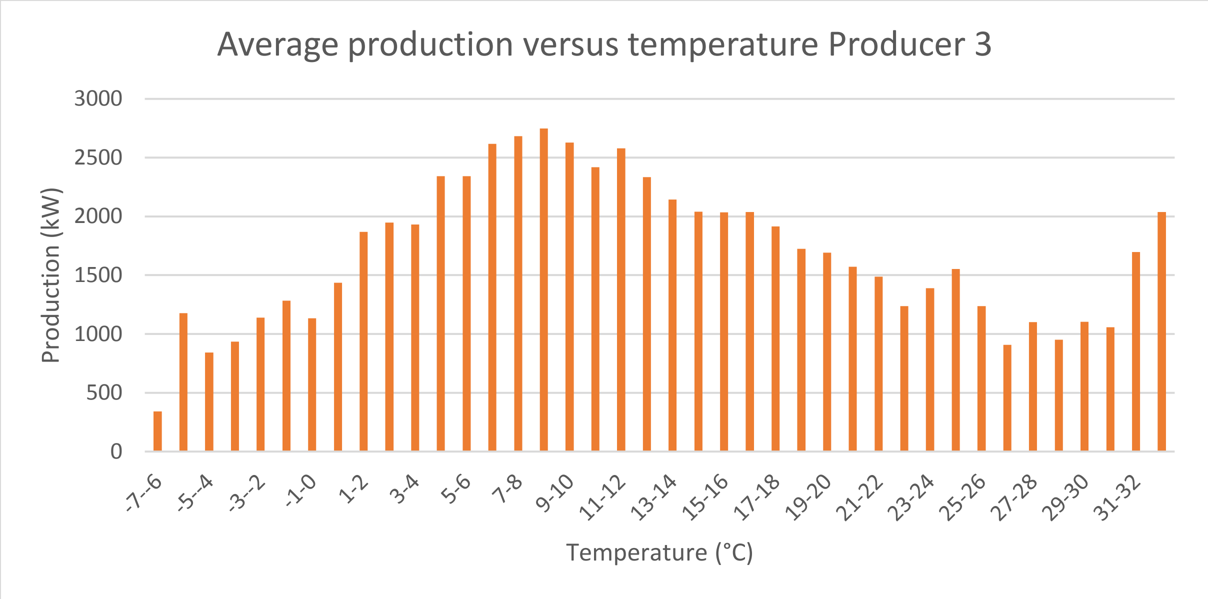

Figure D.3: Average production versus temperature in 2015-2016 for Producer 3 with 6 MW combined rated power.

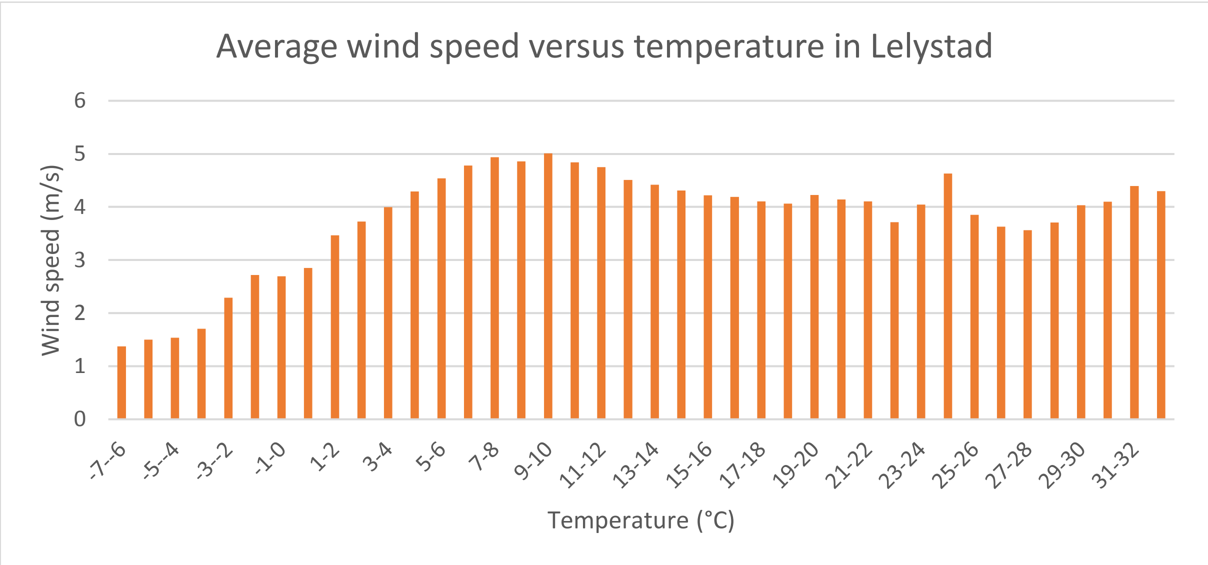

Figure D.4: Average wind speed versus temperature in 2015-2016 in Lelystad.

Appendix E: Weather forecast error

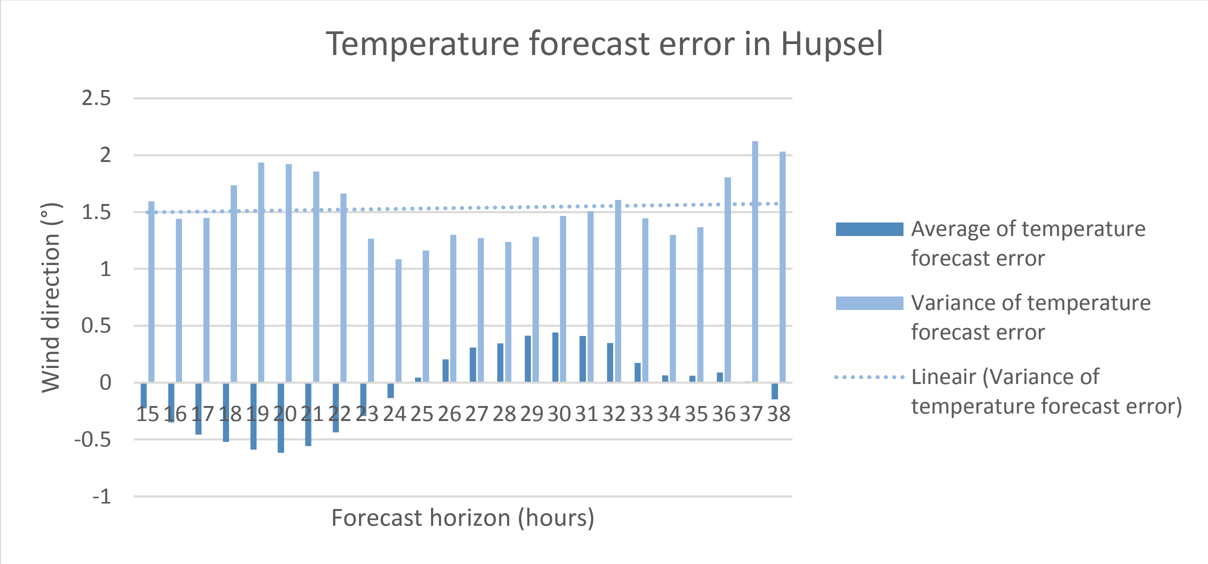

Figure E.1: Average and variance of temperature forecast error in Hupsel during the last 6 months of 2017.

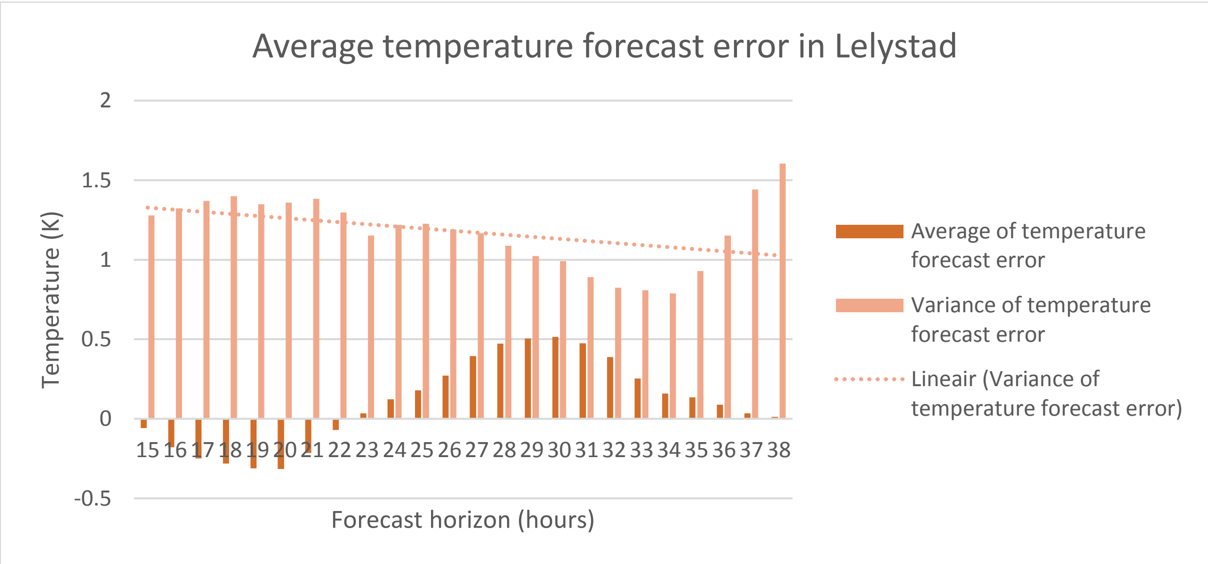

Figure E.2: Average and variance of temperature forecast error in Lelystad during the last 6 months of 2017.

Appendix F: Clean data power curves

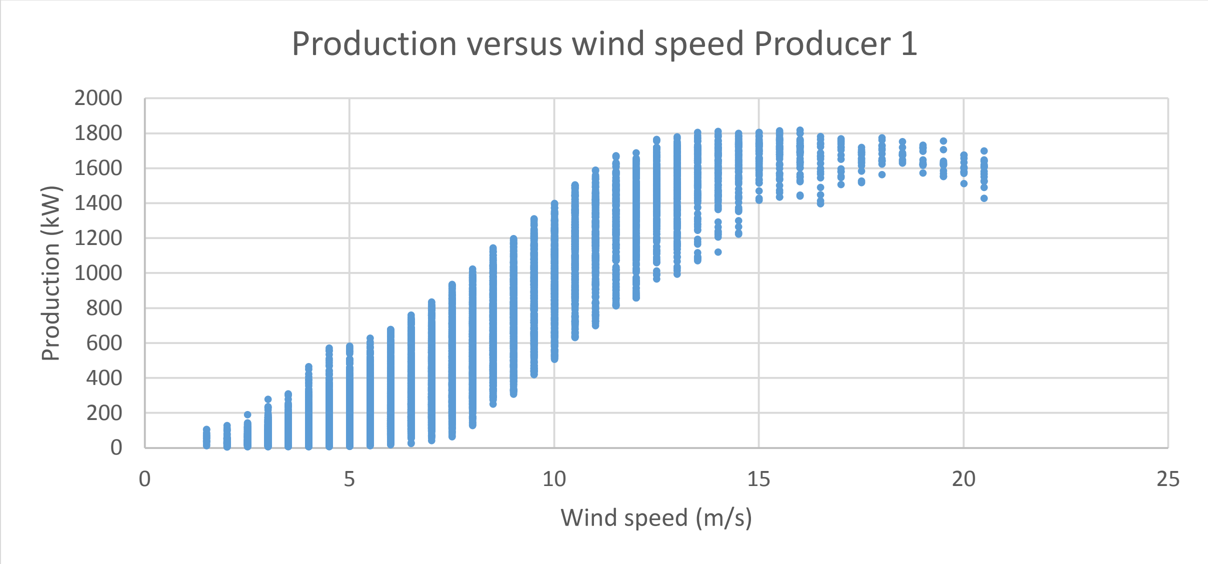

Figure F.1: Scatter plot of production versus wind speed for Producer 1 with 1.8 MW total rated power.

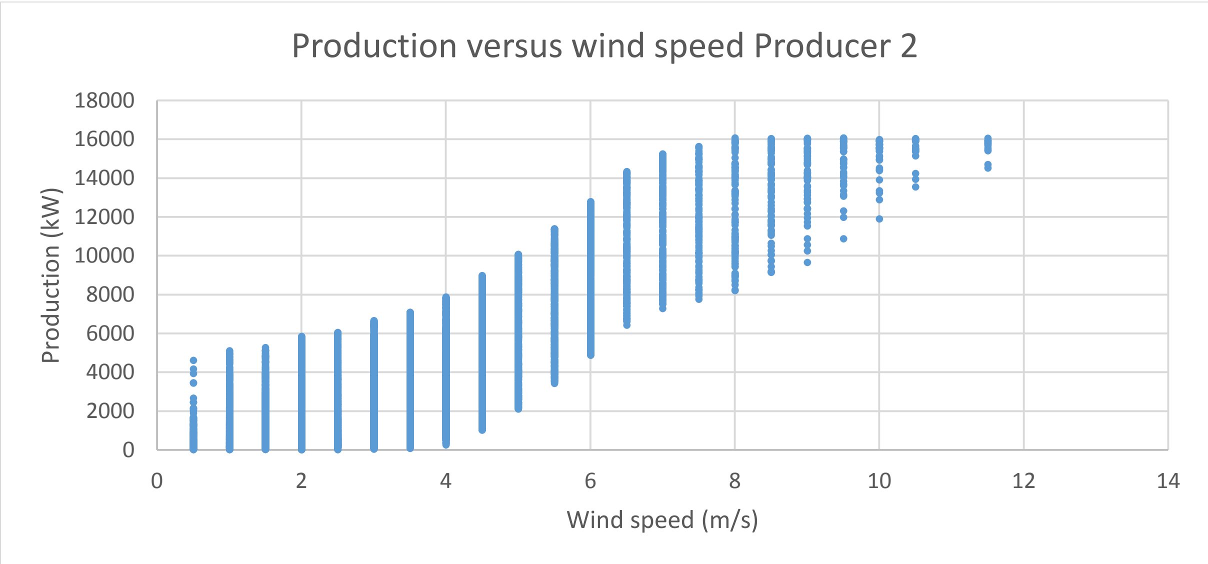

Figure F.2: Scatterplot of production versus wind speed for Producer 2 with 16 MW total rated power.

Appendix G: Regression parameter values

Producer	a_0	a_1	a_2	a_3	a_4	k	λ	δ	c
1	598,48	-333,08	61,02	-2,95	0,04	76,07	4,34	0,10	33,12
2	-9514,43	1615,67	-897,05	-295,00	26,70	-6,09	7,00	3,16	2,99
3	584,84	-342,39	243,97	-21,18	0,47	128,69	2,31	2,95	6,67
4	1012,47	-400,56	61,09	17,74	-1,23	427,02	0,12	1,78	15,87
5	1422,51	-460,85	11,16	36,12	-2,12	211,94	0,07	0,00	66,46
6	4,87	3,69	-2,35	1,06	-0,06	269,69	2,33	2,96	4,33
7	397,87	-149,07	20,89	7,81	-0,52	183,50	0,34	0,00	18,69
8	222,67	244,13	-91,89	27,48	-1,42	334,51	1,01	6,28	10,12
9	732,13	-450,09	96,64	-2,56	-0,11	255,81	0,35	3,49	4,03
10	3411,67	-1976,93	383,63	-16,44	0,16	209,13	0,66	0,00	9,05

Table G.1: Parameter values for the 4th degree Polynomial model.

Producer	a_0	a_1	a_2	a_3	a_4	a_5	k	λ	δ	c
1	4810,10	6120,8	-3948,67	859,80	-56,21	1,16	4,27	0,76	0	200
2	-751,51	2633,4	-1471,50	391,16	-39,88	1,38	521,20	0,92	3,17	3,02
3	4973,99	1230,8	790,68	280,38	-45,52	1,62	7,66	2,37	2,93	6,76

Table G.2: Parameter values for the 5th degree Polynomial model.

Producer	a	m	n	τ	k	λ	δ	c
1	1391,27	0,48	149,36	1,88	345,37	1,00	0	46,36
2	3909,59	6,91	98,76	1,23	729,02	1,60	3,17	3,02
3	1288,37	0,49	24,46	1,56	546,98	2,16	2,93	6,67

Table G.3: Parameter values for the Logistic 4 model.

Appendix H: 4th degree Polynomial power curves for day ahead weather forecasts

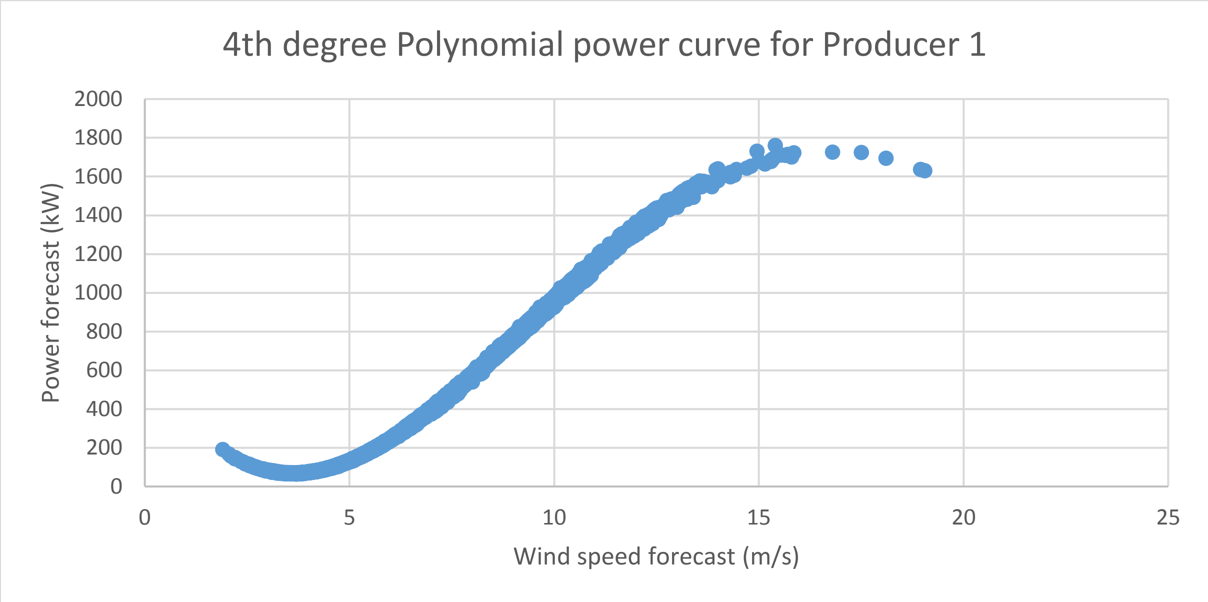

Figure H.1: Power curve using day ahead weather forecast for Producer 1 with 1.8 MW rated power.

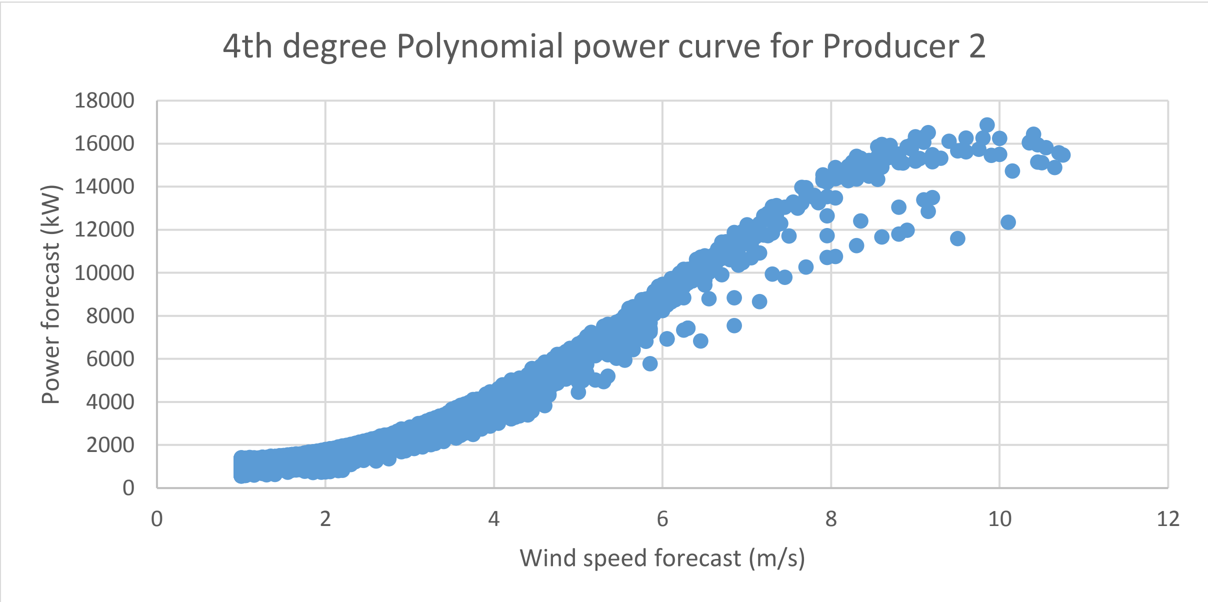

Figure H.2: Power curve using day ahead weather forecast for Producer 2 with 16 MW rated power.

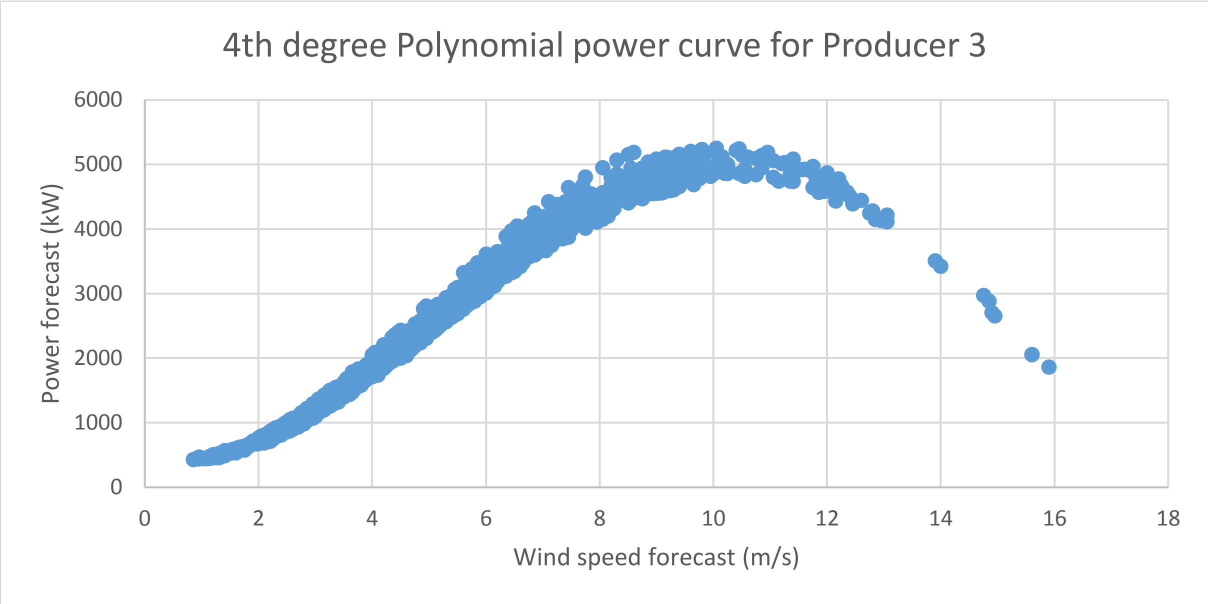

Figure H.3: Power curve using day ahead weather forecast for Producer 3 with 6 MW rated power.

Appendix I: Correlation matrices

In Figures I.1, I.2, and I.3 we use temperature in Kelvin and wind direction in radians. The correlation matrices show Pearson correlations obtained with the statistical software package SPSS.

Correlations

		WindSpeed	WindDirection	Temperature	Rainfall	Radiation	Production
WindSpeed	Pearson Correlation	1	,167**	-,046**	,089**	-,044**	,916**
	Sig. (2-tailed)		,000	,000	,000	,000	,000
	N	13729	13729	13729	13729	13729	13729
WindDirection	Pearson Correlation	,167**	1	,045**	,070**	,047**	,161**
	Sig. (2-tailed)	,000		,000	,000	,000	,000
	N	13729	13729	13729	13729	13729	13729
Temperature	Pearson Correlation	-,046**	,045**	1	-,009	,407**	-,139**
	Sig. (2-tailed)	,000	,000		,308	,000	,000
	N	13729	13729	13729	13729	13729	13729
Rainfall	Pearson Correlation	,089**	,070**	-,009	1	-,085**	,088**
	Sig. (2-tailed)	,000	,000	,308		,000	,000
	N	13729	13729	13729	13729	13729	13729
Radiation	Pearson Correlation	-,044**	,047**	,407**	-,085**	1	-,087**
	Sig. (2-tailed)	,000	,000	,000	,000		,000
	N	13729	13729	13729	13729	13729	13729
Production	Pearson Correlation	,916**	,161**	-,139**	,088**	-,087**	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	
	N	13729	13729	13729	13729	13729	13729

**. Correlation is significant at the 0.01 level (2-tailed).

Figure I.1: Correlation matrix for Producer 1.

Correlations

		WindSpeed	WindDirection	Temperature	Rainfall	Radiation	Production
WindSpeed	Pearson Correlation	1	,149**	,059**	,079**	,161**	,818**
	Sig. (2-tailed)		,000	,000	,000	,000	,000
	N	15018	15003	15018	15018	15018	15018
WindDirection	Pearson Correlation	,149**	1	,053**	,062**	,013	,104**
	Sig. (2-tailed)	,000		,000	,000	,120	,000
	N	15003	15003	15003	15003	15003	15003
Temperature	Pearson Correlation	,059**	,053**	1	,028**	,494**	-,104**
	Sig. (2-tailed)	,000	,000		,001	,000	,000
	N	15018	15003	15018	15018	15018	15018
Rainfall	Pearson Correlation	,079**	,062**	,028**	1	-,067**	,089**
	Sig. (2-tailed)	,000	,000	,001		,000	,000
	N	15018	15003	15018	15018	15018	15018
Radiation	Pearson Correlation	,161**	,013	,494**	-,067**	1	-,148**
	Sig. (2-tailed)	,000	,120	,000	,000		,000
	N	15018	15003	15018	15018	15018	15018
Production	Pearson Correlation	,818**	,104**	-,104**	,089**	-,148**	1
	Sig. (2-tailed)	,000	,000	,000	,000	,000	
	N	15018	15003	15018	15018	15018	15018

**. Correlation is significant at the 0.01 level (2-tailed).

Figure I.2: Correlation matrix for Producer 2.

Correlations

		WindSpeed	WindDirection	Temperature	Rainfall	Radiation	Production
WindSpeed	Pearson Correlation	1	,206**	,094**	,106**	,137**	,836**
	Sig. (2-tailed)		,000	,000	,000	,000	,000
	N	11713	11702	11713	11713	11713	11713
WindDirection	Pearson Correlation	,206**	1	,074**	,047**	,100**	,168**
	Sig. (2-tailed)	,000		,000	,000	,000	,000
	N	11702	11702	11702	11702	11702	11702
Temperature	Pearson Correlation	,094**	,074**	1	,034**	,481**	,025**
	Sig. (2-tailed)	,000	,000		,000	,000	,008
	N	11713	11702	11713	11713	11713	11713
Rainfall	Pearson Correlation	,106**	,047**	,034**	1	-,068**	,089**
	Sig. (2-tailed)	,000	,000	,000		,000	,000
	N	11713	11702	11713	11713	11713	11713
Radiation	Pearson Correlation	,137**	,100**	,481**	-,068**	1	-,041**
	Sig. (2-tailed)	,000	,000	,000	,000		,000
	N	11713	11702	11713	11713	11713	11713
Production	Pearson Correlation	,836**	,168**	,025**	,089**	-,041**	1
	Sig. (2-tailed)	,000	,000	,008	,000	,000	
	N	11713	11702	11713	11713	11713	11713

**. Correlation is significant at the 0.01 level (2-tailed).

Figure I.3: Correlation matrix for Producer 3.

Appendix J: T-tests of weather forecast errors

To see if there is statistical evidence of a bias in the weather forecasts, we conduct statistical tests. We are interested in whether the mean of a forecast error is significantly different from 0. The student's T-distribution and the normal distribution can be used for significance testing. Rule of thumb is to use the T-distribution if the variance of the population is not known (Liese & Miescke, 2008). We can calculate the variance of the sample size, but we do not know the true variance of the population. Therefore, we can use a one-sample T-test to test whether the mean is significantly different from 0 (Allua & Thompson, 2009). The hypotheses for the one-sample T-test are:

- H_0: Sample mean = 0
- H_1: Sample mean $\neq 0$

In this case our sample is the forecast error. First, we compute the forecast error:

$$\varepsilon_i = F_i - A_i \tag{J.1}$$

where:

ε_i = forecast error for observation i

F_i = forecast value for observation i

A_i = actual measurement for observation i

We use SPSS to calculate the sample means and the conduct the T-tests. We test the means at a 95% confidence level, which means that if the p-value is significant, we are 95% certain that the mean is different than 0. A statistically significant mean greater than zero indicates the forecast has a positive bias (forecasts are too high). A negative bias is indicated by a statistically significant mean smaller than 0.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
WindspeedErrorVlissingen	4417	-,9933	1,78102	,02680
WinddirectionErrorVlissingen	4327	7,8693	58,15953	,88415
TempErrorVlissingen	4417	-,3800	1,36243	,02050
WindspeedErrorHupsel	4411	1,1943	1,18390	,01783
WinddirectionErrorHupsel	4406	16,9602	63,38953	,95498
TempErrorHupsel	4411	-,0592	1,28371	,01933
WindspeedErrorLelystad	4417	,1223	1,26399	,01902
WinddirectionErrorLelystad	4396	18,3621	60,93155	,91900
TempErrorLelystad	4417	,0980	1,11442	,01677

One-Sample Test

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
WindspeedErrorVlissingen	-37,066	4416	,000	-,99330	-1,0458	-,9408
WinddirectionErrorVlissingen	8,900	4326	,000	7,86931	6,1359	9,6027
TempErrorVlissingen	-18,539	4416	,000	-,38004	-,4202	-,3399
WindspeedErrorHupsel	67,001	4410	,000	1,19434	1,1594	1,2293
WinddirectionErrorHupsel	17,760	4405	,000	16,96022	15,0880	18,8325
TempErrorHupsel	-3,061	4410	,002	-,05916	-,0971	-,0213
WindspeedErrorLelystad	6,431	4416	,000	,12231	,0850	,1596
WinddirectionErrorLelystad	19,981	4395	,000	18,36215	16,5605	20,1638
TempErrorLelystad	5,846	4416	,000	,09802	,0651	,1309

Figure J.1: SPSS output of One-Sample T-Test for means of weather forecast errors.

The forecast error of wind speed in Vlissingen has a mean of -0.9933. Figure J.1 shows that the p-value is 0.000 at a 95% confidence level. This means we can reject H_0 and assume that the mean is significantly smaller than 0. We can conclude that the wind speed forecast is negatively biased, meaning the forecast is systematically 0.99 m/s too low.

Figure J.1 illustrates that all p-values are smaller than $\alpha = 0.05/2 = 0.025$. This means that at a confidence level of 95%, all forecasts are biased.

Appendix K: Jarque-Bera test for normality

To test the assumption of normality or errors for the linear regression models we use the Jarque-Bera (JB) test. This test checks whether the sample data have the skewness and kurtosis matching a normal distribution (Jarque & Bera, 1987). The JB test is more reliable than the Shapiro-Wilk test and the Kolmogorov-Smirnov test for large sample sizes greater than 2000 (Thadewald & Buning, 2007). The Chi-square test with 2 degrees of freedom (X_2^2) can be used to test if the sample data are normality distributed (Jarque & Bera, 1987). The JB test statistic is defined by:

$$JB = \frac{N-k}{6}\left(S^2 + \frac{(C-3)^2}{4}\right) \tag{K.1}$$

where:

N = number of observations

k = number of predictors

S = sample skewness

C = sample kurtosis

A normal distribution has a skewness of 0 and kurtosis of 3. The closer the sample skewness and sample kurtosis are to these values; the closer the JB test statistic is to 0. We can use the X^2-test with 2 degrees of freedom to test whether the sample data are normally distributed with the JB statistic. The hypotheses for the Jarque-Bera test are:

- H_0: Sample data are normally distributed
- H_1: Sample data are not normally distributed

In this case our sample is the forecast error. First, we compute the forecast error:

$$\varepsilon_i = F_i - A_i \tag{K.2}$$

where:

ε_i = forecast error for observation i

F_i = forecast value for observation i

A_i = actual measurement for observation i

If the JB statistic is greater than the critical point of X_2^2 with α=0.05 (95% confidence level), we reject H_0 and assume that the sample is not normally distributed. For the log model, we test whether the errors of the log transformation are normally distributed. We test for normality of errors for all linear regression models for Producers 1, 2, and 3. The clean training data from 2015-2016 is used for each producer.

Model	Predictor(s)	N	JB statistic	X_2^2 (α=0.05)	Normality?
Log	Wind speed	13729	1199.28	5.99	No
3rd degree Polynomial	Wind speed	13729	276.56	5.99	No
4th degree Polynomial	Wind speed	13729	204.97	5.99	No
5th degree Polynomial	Wind speed	13729	331.77	5.99	No
Log	Wind speed + Temperature	13729	1512.08	5.99	No
Log	Wind speed + temperature + wind direction	13729	1535.77	5.99	No

Table K.1: Jarque-Bera test for normality of errors of linear regression models for Producer 1.

Model	Predictor(s)	N	JB statistic	X_2^2 (α=0.05)	Normality?
Log	Wind speed	15003	2643.33	5.99	No
3rd degree Polynomial	Wind speed	15003	398.94	5.99	No
4th degree Polynomial	Wind speed	15003	425.35	5.99	No
5th degree Polynomial	Wind speed	15003	416.79	5.99	No
Log	Wind speed + Temperature	15003	2335.40	5.99	No
Log	Wind speed + temperature + wind direction	15003	2961.55	5.99	No

Table K.2: Jarque-Bera test for normality of errors of linear regression models for Producer 2.

Model	Predictor(s)	N	JB statistic	X_2^2 (α=0.05)	Normality?
Log	Wind speed	11702	4131.80	5.99	No
3rd degree Polynomial	Wind speed	11702	134.65	5.99	No
4th degree Polynomial	Wind speed	11702	183.02	5.99	No
5th degree Polynomial	Wind speed	11702	206.51	5.99	No
Log	Wind speed + Temperature	11702	3922.26	5.99	No
Log	Wind speed + temperature + wind direction	11702	3984.78	5.99	No

Table K.3: Jarque-Bera test for normality of errors of linear regression models for Producer 3.

We reject H_0 for all linear regression models of Producer 1, 2, and 3. This means we have statistical evidence that the errors for all linear regression models are not normally distributed.