CHALMERS

An analytical tool for the evaluation of wind power generation

Master's Thesis within the Sustainable Energy Systems programme

NELLY FORSMAN

Department of Energy and Environment
Division of Electric Power Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2011

MASTER’S THESIS

An analytical tool for the evaluation of wind power generation

Master’s Thesis within the Sustainable Energy Systems programme

NELLY FORSMAN

SUPERVISOR:

Katharina Fischer

EXAMINER:

Torbjörn Thiringer

Department of Energy and Environment
Division of Electric Power Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2011

An analytical tool for the evaluation of wind power generation

Master's Thesis within the Sustainable Energy Systems programme
NELLY FORSMAN

Department of Energy and Environment
Division of Electric Power Engineering
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: + 46 (0)31-772 1000

Cover:
Two of Göteborg Energi's wind turbines at Risholmen. Photo: Nelly Forsman

Chalmers Reproservice
Göteborg, Sweden 2011

An analytical tool for the evaluation of wind power generation

Master's Thesis in the Sustainable Energy Systems programme
NELLY FORSMAN
Department of Energy and Environment
Division of Electric Power Engineering
Chalmers University of Technology

ABSTRACT

Based on a system previously developed for district heating by Göteborg Energi, a tool for monitoring and evaluation of wind power generation has been developed. With the tool it is possible to detect long-term trends of the turbine's performance.

Parameters influencing the power generation were identified to be wind speed, wind direction and air density. These were combined to create models which represent historical data from specific wind turbines and can be used as a reference signature. To get the signature independent of air density and comparable, measured data was normalized according to standard IEC 61400-12-1. With the tool, it is possible to add new data and compare it with the reference signature. If new data is deviating more than a specified value from the reference, an alarm is initiated. It was found that a method based on piecewise linear regression with seven pieces could be used to parameterize the data. The method of bins, where average values of power and wind speed in small intervals are interpolated to construct a curve, resulted in an even more accurate representation of the historical data. A third method for parameterization was investigated, polynomial regression. This method, however, returned lower accuracy than the other methods.

Three concepts to generate alarms in case of deviating wind-turbine performance were developed. Piecewise linear regression was combined with an alarm concept where alarms are initiated when the present data deviates more than a predefined threshold from the reference signature. Another alternative of parameterization and alarm generation returning satisfying results consists of the method of bins for both parameterization and detection of deviations. For this model, a larger amount of data is needed than for the other models. With all aspect taken into account the recommendation to Göteborg Energi was to apply the piecewise linear regression, 7 pieces, combined with constant limit method for alarm generation.

The models were developed using data from Göteborg Energi's wind turbines at Risholmen in the harbour of Gothenburg. These turbines have a rated power of 600 kW and are of the type Vestas V44.

Key words: wind power, monitoring system, performance monitoring

Ett analysverktyg för utvärdering av vinkraftproduktion
Examensarbete inom mastersprogrammet Sustainable Energy Systems
NELLY FORSMAN
Institutionen för Energi och Miljö
Avdelningen för Elteknik
Chalmers tekniska högskola

SAMMANFATTNING

Göteborg Energi har sedan 2006 tillbaka ett system för kontroll av fjärrvärme-leveranser. Ett liknade system för kontroll och övervakning av vindkraftsproduktionen efterfrågades. Vindhastighet, vindriktning och luftens densitet identifierades vara de parametrar som påverkar produktionen mest. Dessa parametrar kombinerades och användes i tre olika modeller där historisk data parametriserades för att användas som en referenssignatur. I modellerna kan ny data läggas in och jämföras mot referenssignaturen. Om den nyare datan avviker från referensen initieras ett larm.

En metod där styckvis linjär regression används för att parametrisera data visade sig kunna användas. En annan metod som också visade sig fungera bra var binmetoden. Denna metod bygger på interpolation av medelvärden för små delintervall. Noggrannheten för denna metod blev något högre. Den tredje metoden som utvecklades bygger på polynomregression, vilket dock visade sig ge lägre noggrannhet än övriga undersökta metoder.

För att kunna detektera fel och avvikelser utvecklades tre tillvägagångssätt. Styckvis linjär regression kombinerades med en metod där larm initierades om avvikelsen var större än en fördefinierad gräns för accepterade avvikelser. Den tidigare nämnda binmetoden visade sig också kunna användas för larmdetektering. Denna kombination kräver en större mängd data men visade goda resultat. Med dessa metoder är det möjligt att identifiera trender och utvärdera turbinens prestanda för en längre period. Med hänsyn tagen till alla olika aspekter blev rekommendationen till Göteborg Energi att tillämpa styckvis linjär regression, 7 linjer, och konstanta gränser baserade på δ för larmgenerering.

För att utveckla modellerna användes data från Göteborg Energis vindkraftverk på Risholmen i Göteborgs hamn. Dessa turbiner är alla av typen Vestas V44 och har maximal effekt på 600 kW. Mängden data tillgänglig för analys var begränsad och möjligheterna att tillämpa modellerna på andra typer av turbiner kunde inte testas..

Nyckelord: vindkraft, övervakningssystem, prestandaövervakning

Contents

Preface

The thesis has been carried out between May 2011 and October 2011 at Göteborg Energi at the department TK (Teknisk Kundkommunikation). The project was financed and commissioned by the department OE (Förnyelsebar el) under the auspices of Jonas Cognell.

The department OE is working with development of the wind power generation. Today, the generation cannot be supervised efficiently by any system at Göteborg Energi. In 2006, a master's thesis at TK resulted in a system for monitoring and quality control of district heating named Kasper. This system is now in use and contributes strongly to ensure the measurement quality for district heating. In the thesis project presented in this report, a similar system has been developed for wind power. A first version of Wind-Kasper has been implemented in parallel to this work by Therese Berge, Kentor.

The idea and the first initiative of the project were launched by Ola Jobring at TK, Göteborg Energi. During the work with the thesis, Ola has been the supervisor and has contributed with invaluable help and support. I would also like to thank Katharina Fischer at the Department of Energy and Environment, Division of Electric Power Engineering, who was the supervisor at Chalmers and assisted with qualified guidance and support. Thanks also to Professor Torbjörn Thiringer who has contributed with valuable specialist expertise.

Göteborg, October 2011

Nelly Forsman

Nomenclature

Roman letters

a	Parameter in the vector θ determing the logistic function $f(x,\theta)$
A	Rotor swept area (m^2)
b	Vector containing parameters determining a parameterized curve
C_p	Power coefficient
E	East
ENE	East-north-east
ESE	East-south-east
m	Parameter in the vector θ determining the logistic function $f(x,\theta)$
n	Parameter in the vector θ determining the logistic function $f(x,\theta)$
N	Number of data points, North
NNE	North-north-east
NNW	North-north-west
p	Atmospheric pressure (Pa)
p_{ref}	Reference atmospheric pressure (Pa)
P	Power output (W)
$\hat{P}$	Expected power (W)
P_n	Normalized power (W)
R	Gas constant (J/K kg), Rotor radius (m)
S	South
SSE	South-south-east, sum of squared errors (kW^2)
SSE_{tot}	Total sum of squared errors for all four wind directions (kW^2)
SSW	South-south-west
T	Air temperature (K)
T_{ref}	Reference temperature (K)
$T.I.$	Turbulence intensity
v	Wind speed (m/s)
v_n	Normalized wind speed (m/s)
v_r	Wind speed at reference height (m/s)
W	West
WNW	West-north-west
WSW	West-south-west
x	Independent variable, wind speed (m/s)

y Dependent variable, power (W), energy (MWh)

$\hat{Y}_x$ Expected value of the dependent variable for the independent variable x

z Height above ground level (m)

z_r Reference height (m)

Greek letters

α Wind-shear power law exponent

δ Standard error (kW)

δ_{tot} Weighted standard error for all wind directions (kW)

Δ Deviation in power from reference line (kW)

λ Tip speed ratio

ρ Density of air, 10 minute averaged air density (kg/m^3)

ρ_{ref} Reference air density for normalization (kg/m^3)

σ Standard deviation of the power (kW)

σ_v Standard deviation of wind speed variations (m/s)

τ Parameter in the vector θ determining the logistic function $f(x, \theta)$

ω Rotational speed of the wind-turbine rotor (rad/s)

θ Vector parameter determining the shape of a logistic function

Indices

i Index of data point

j Index of bin

k Index representing wind direction

1 Introduction

The wind power production is growing significantly and the importance of an efficient generation increases since unpredicted losses in production can cause substantial economic consequences. The development of multiple megawatt turbines placed in remote locations also requires better ways to control and monitor the performance of the turbines to optimize maintenance and production. The optimization of power output is complicated by the fact that it is a multiple-input, single-output signal. If the power production of the turbines can be supervised on distance, better control and optimization will be possible. Thereby, reductions in production, which might indicate a deteriorated condition of wind-turbine components, can be detected earlier and measures can be taken before any serious failure occurs. The economical revenue of this is increasing while the rated power of the turbines is increasing.

1.1 Problem description

Göteborg Energi, is planning for a reinforcement of wind power within the next years. The objective is to set up 100 new turbines until 2015 and to supply 500 GWh of wind energy in 2015. Therefore, there is a need to find an effective way to measure and monitor the performance of the wind turbines. A couple of years ago, a software for monitoring and quality control of district heating was developed by Göteborg Energi. The purpose of this software, named Kasper, was to ensure the quality of the measurements of district heating deliveries and to detect faulty measures. The software monitors the energy signature, that is the energy supplied to buildings versus outdoor temperatures, and compares the hourly average values with a reference signature. The system has been proven to be very efficient and reliable. It is therefore desirable to develop and adapt this system to be applied to wind power installations as well.

Currently, Göteborg Energi has eleven wind turbines running within the area of Göteborg. Five of these are situated at Risholmen, four are Vestas V44 and have a capacity of 600 kW each, and the last one is a Bonus Mk3 with rated power of 450 kW. The study has been focused on the four Vestas turbines, as the available data on those were most comprehensive and the fact that they have been operating during a longer period. Today, there is no effective performance-monitoring system for those and information concerning their production has to be gathered from different sources. A new monitoring system will also support the monthly evaluations of the production of the turbines as no such tool exists today.

A so called power curve is usually supplied by the turbine manufacturer. It shows how the specified type of turbine shall generate power under specified conditions in ideal terrain. This curve works as a reference of the performance from the manufacturer but as the terrain is barely ever ideal, its usefulness is limited in practice. It is therefore desirable to have an accurate turbine- and location specific reference signature representing the historic performance, to which the present performance can be compared.

1.2 Previous work

The first version of Kasper for district heating was developed in 2006 as a master thesis in statistical mathematics on behalf of Göteborg Energi. It was found that piecewise linear regression should be used to represent the reference signature (Munoz, 2006). The work contains a statistical analysis of data for the energy consumptions of 90 buildings in Göteborg from May 2004 to May 2005. The thesis also contains an error detection method where alarms are initiated if the deviation from the reference line is higher than 3.5 standard deviations of the data used for the reference. Based on the results from the thesis, a Java application was implemented.

During autumn 2009 and spring 2010, a second thesis was conducted, aiming to improve the current application Kasper. The result became a version of Kasper where the total difference between the actual supply of district heat and the expected demand during a longer period can be calculated (Lindqvist, 2010). The application makes it possible to evaluate energy saving actions and the energy efficiency of buildings.

1.3 Purpose

The purpose of the master's thesis is to develop the basis of an application for analysis and monitoring of wind power production. The system Kasper will be developed and adapted to handle energy signatures of wind turbines. With the application, it shall be possible to detect sudden performance degradations and short-term deviations e.g. due to measurement errors and signal faults. Also long-term changes and trends of the production shall be detected with the system. This is done through detection of deviations from a reference energy signature of the turbine. With these functions, the system can be used as an Early-Warning System and the information obtained will provide decision support concerning maintenance. It can also be used for evaluation of optimization trials e.g. regarding control functions.

1.4 Scope and objectives

Within the scope of the work is to identify the parameters influencing the wind power production. Those parameters shall be used when alternative methods for parameterization of historical data are developed. In order to enable detection of deviating data, procedures which are compatible with the parameterization methods, shall be developed. The methods and alarm procedures are combined to models were the wind power production from one period can be evaluated and compared with a reference period. For construction and testing of the models, Matlab is used.

The study is limited to data available at Göteborg Energi. Due to difficulties of obtaining valid data, the data used in the models is mainly limited to two periods of time, June 2008 to March 2009 and May 2010 to October 2010.

The study is focused on identifying a suitable model and algorithm for a monitoring and evaluation system which can be used at Göteborg Energi. The software implementation of the system is out of the scope of the work. As a result of the work, a recommendation to Göteborg Energi is given with respect to the possibilities of applying the existing system for district heating to wind power with only minor modifications in order to facilitate the implementation.

2 Background

The energy which can be extracted from a wind turbine is dependent on the wind speed. A basic introduction to the governing equations is given in section 2.1.

During the last years several different systems for monitoring and control of wind power performance have been developed, which are described in section 2.2. Many of those are based on mathematical models and are using large amounts of data for prediction of wind power production.

2.1 Theory

The power production of a wind turbine is theoretically determined by

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

where ρ is the density of the air, A is the rotor swept area, C_p is the power coefficient, (the ratio between the power output of the wind turbine and the power available in the wind, and v is the speed of the incoming air flow. From (2.1) it can be seen that the power output is a function of the cube of the wind speed, therefore, a doubled wind speed will lead to eight times higher power production. The power is directly proportional to the density of the air. The density is a function of air temperature and atmospheric pressure according to

$$\rho = p/RT \tag{2.2}$$

where p is the atmospheric pressure, R is the gas constant for dry atmospheric air, 287 J/K kg, and T is the temperature of air in K. This implies that the density will vary according to the local and temporal atmospheric conditions and thus also affect the power output from the turbine.

The value of C_p varies with the wind speed which implies that the power curve of a turbine is not just a function of the cubic wind speed. The power coefficient is dependent on the so called tip speed ratio, λ, given by

$$\lambda = \frac{R\omega}{v} \tag{2.3}$$

where ω is the rotational speed in rad/s and R is the radius of the rotor. With a fixed rotational speed of the rotor, λ will vary which leads to that also C_p will vary. The maximum value of C_p is usually around 0.45 while a theoretical limit of 16/27 (≈0.59) is determined by Betz's Law. At high wind speeds, around 18 m/s, the coefficient is approximately around 0.1. At these wind speeds, the wind turbine is extracting its rated power and the turbine is regulated to not extract more. In Figure 2.1 it can be seen that the Vestas V44 reaches its rated power of 600 kW at a wind speed of 16 m/s, hence called rated speed. The figure also shows that the power coefficient reaches its maximum of 0.43 at a wind speed of 8 m/s. All turbines, irrespective of size, have a cut-in speed and cut-out speed which are the limits at which the turbine will start generation and the wind speed at which the turbine is shut down. For a Vestas V44, the cut-in wind speed is 4 m/s, meaning that the turbine will start to generate electricity when the wind speed exceeds 4 m/s. It is interesting to note that, however, when the turbine is operating and the wind speed decreases, the turbine will stop generating at 3 m/s. Cut-out speed for the same turbine is 25 m/s.

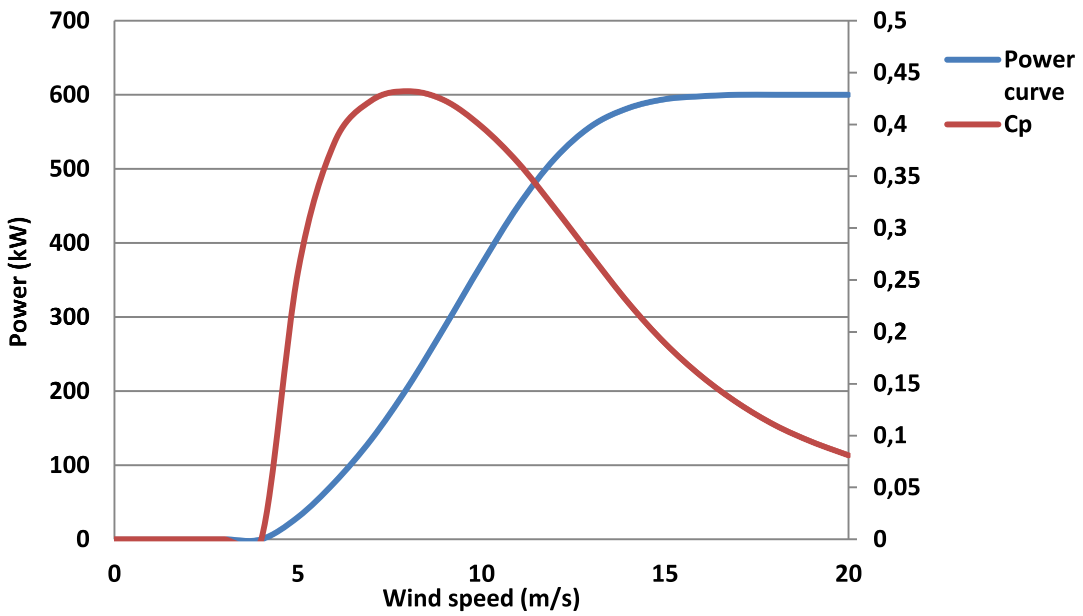

Figure 2.1 Power curve (blue line) and the power coefficient (red line) for Vestas V44-600 kW

2.2 Existing systems for wind turbine monitoring

Currently, there are several different systems on the market for monitoring wind turbines operation, condition monitoring systems (CMS). Most of these systems are based on vibrational analysis and thus use vibrations as an indicator of condition.

The main part to consider for condition monitoring systems (CMS) is the drive train: The gearbox consisting of several critical components such as bearings and gears, the main shaft, the main bearing and the generator (Nilsson & Bertling, 2007). SKF's condition monitoring system, WindCon, is based on vibrational analysis. Other techniques which can be used for condition monitoring are oil analysis, thermography, physical condition of materials, strain measurement, acoustic measurement, electrical effects, process parameters, visual inspections and performance monitoring. These techniques allow the maintenance strategy called condition based maintenance (CBM) to be applied on the turbines. This strategy implies that maintenance shall be conducted before a fault occurs as the CMS shall make it possible to predict impending faults.

Since the trend goes towards larger turbines being installed in more remote areas, the cost for maintenance actions is increasing. The most cost-effective strategy for maintenance differs for different components within the turbine. CBM has been proven to be effective and cost efficient for components such as gears and bearings (Andrawus, et al., 2009, Besnard & Bertling, 2010) Research has shown that this strategy can optimize the maintenance for a whole wind farm and thereby, the cost of maintenance can be reduced significantly (Tian et. al., 2011). A large project, "Advanced maintenance and repair for offshore wind farms using fault prediction and condition monitoring techniques", co-ordinated by ITES (2005) advocates that the condition-based maintenance is most beneficial for offshore wind turbines.

Other strategies that can be applied are time-based strategy, implying that maintenance is carried out at a predetermined time, and run-to-failure strategy (Nielsen & Sørensen, 2010). Those strategies can be more appropriate or more cost effective for some components within the turbine, e.g. in cases in which it is hard to predict an impending fault.

For monitoring of wind turbine performance, SCADA (Supervisory Control And Data Acquisition) system is used. This is used as a complement to CMS, which only provide monitoring of technical component condition and not of the turbine performance. SCADA systems have during a longer period been a vital part for the control system of wind turbine. These systems have during more recent times been developed to collect, monitor and log much larger amounts of data. This has improved the monitoring systems significantly. All large turbines that are being installed nowadays have sensors used to obtain large amounts SCADA data. The system is aiming to control and monitor the entire process and the number of parameters measured is usually more than fifty for each wind turbine. Parameters of interest in particular, in addition to the wind speed and active power, are usually bearing temperatures in the gearbox, gear oil temperature and temperature in the generator windings. The sampling frequency is usually 1 Hz but most SCADA systems store only 10-minute averaged values.

Göteborg Energi uses CMS for some of their turbines and some of the turbines at Risholmen are equipped with WindCon for vibrational analysis. Larger, more recently installed turbines all have CMS and are supervised through it. The turbine in Gårdsten in the northeast part of Göteborg is equipped with a CMS and selected data from this turbine is obtained via the manufacturer's web. Other turbines owned by Göteborg Energi, for instance the recently installed (2011-05-22) turbines at Töftedalsfjället in Dalsland also have CMS and these are being supervised by external staff at a remote monitoring center. SCADA data and CMS data from these turbines can currently be obtained through the web and by Vattenfall. There are plans for integrating the data from these turbines into the operation management system at Göteborg Energi so that also the performance of these turbines can be monitored in the Kasper application under development The remaining turbines owned by Göteborg energi can be performance monitored when connected to the server of the manufacturer, or for some turbines, when connected to Vattenfall's web server.

2.3 Mathematical models

The power curve of a wind turbine can be described by different mathematical models. In a study by Kusiak et. al (2009), three different models were compared; two parametric models and one non-parametric. The study showed that a least-squares parametric model and a non–parametric model based on a k-nearest neighbour (k-nn) algorithm gave the best result. The least squares method is often used for regression and data fitting and minimizes the sum of the squared errors, e.g. the squares of the deviations from the parameterized curve, Eq. (2.4). The data set, (x_i, y_i), (x_{i+1}, y_{i+1}), … , (x_n, y_n), consist of N data points where x is the independent variable and y is the dependent variable, in this case the wind speed and the power respectively. The sum of the squared errors is calculated by

$$\sum_{i=1}^{N}(y_i - f(x_i, b))^2 \tag{2.4}$$

where $f(x_i, b)$ is the function of the curve searched and b is the vector containing the parameters determining the curve. A logistic function was used to approximate the curve:

$$y = f(x, \theta) = a\frac{1+me^{-x/t}}{1+ne^{-x/t}} \tag{2.5}$$

In this equation, x denotes the wind speed measured at the turbine nacelle top, y is the power generated and $\theta = (a, m, n, \tau)$ is a vector parameter of the logistic function determining the shape.

In the k-nn model, the wind speed is used as a predictor and the power is used as a dependent variable. A residual control is used to eliminate residuals and thereby achieve a more accurate model. The models were designed from a test set of data from one period of time. The models were then compared with another set of data points and the result was satisfying and the models could be used as reference power curves for online monitoring.

Another method which is commonly used for data fitting is the maximum likelihood method (MLM). For the MLM, a known probability distribution function is used to find the parameters which maximize the probability (likelihood) of the measured data. For wind speed, the Weibull distribution is often used as this distribution has been proven to match the wind data well (Seguro & Lambert, 2000).

To construct a reliable model from data, automatic filtering of the data will be needed to eliminate erroneous data and thereby prevent their influences on the result. Sainz et. al (2009) propose a statistical technique of the least median squares combined with a random search to detect and reject erroneous data points. The least median square method is less sensitive to outliers than the least squares method as a single outlier can destroy the fitting using the least square method. The least median square method was applied to raw data and after filtering, a reference power curve could be constructed using an exponential model which yielded satisfying results. The exponential model took wind speed, wind direction and air temperature into consideration and was shown to work well throughout the whole range of wind speeds.

Another method investigated by Sainz et. al (2009), "the method of bins", is also working well, at least as far as there is sufficient data, covering the whole wind speed range. An advantage of this model is that the level of rejection of data can be adjusted easily, and thereby assure that no correct data will be rejected. With the above mentioned least median square method proposed by Sainz et. al. (2009), the rejection of data might be extensive. "The method of bins" is also described in the IEC 61400-12 standard "*Power Performance Measurements of Electricity Producing Wind Turbines"* (IEC, 2005) for construction of a power curve for a wind turbine: The wind speed range is split up into bins of 0.5 m/s for wind speed between 0 m/s and the cut-out speed. For each bin j, the average of the normalized wind speed is calculated using:

$$v_{n,j} = \frac{1}{N_j}\sum_{i=1}^{N_j} v_{n,j,i} \tag{2.6}$$

The normalized wind speed is the measured wind speed recalculated for a reference air density. The normalization procedure is described in Section 4.4 of this thesis. The average power and the standard deviation of the power in each bin are calculated according to

$$P_j = \frac{1}{N_j}\sum_{i=1}^{N_j} P_{j,i}, \tag{2.7}$$

$$\sigma_j = \sqrt{\frac{1}{N_j}\sum_{i=1}^{N_j}(P_{j,i} - P_j)^2} \qquad (2.8)$$

where $v_{n,j}$ is the normalized and averaged wind speed in bin j, $v_{n,\mathrm{j},i}$ is the normalized wind speed of data point i in bin j, N_j is the number of data point in bin j, P_j is the averaged power output in bin j, $P_{j,i}$ is the power of data point i in bin j and σ_j is the standard deviation of the power output in bin j.

Data sets deviating more than e.g. $5\sigma_i$ can be rejected and by repeating the procedure, recalculating$v_{n,j}$, P_j and σ_j the accuracy can be increased. The averaged values can then be used to create a power curve.

Catmull (2011) proposed to use SCADA data for the construction of Artificial Neural Networks (ANN) which then can be used for prognosis of component health condition. ANN is a nonlinear mapping system for identifying patterns in large amounts of data. From the study of Catmull (2011) it was concluded that this is one of the best methods so far for prediction of power from a wind turbine under certain conditions. The Kohonen Self-Organizing Map (SOM), which is an unsupervised learning ANN, can successfully be applied to SCADA data to detect an impending failure although it cannot provide information about the type and location of the problem (Catmull, 2011). The distance between the SOM and the point in space described by the state of the turbine is used as an overall condition indicator. An increasing distance indicates an impending fault.

3 Current system at Göteborg Energi

Göteborg Energi have developed a system for control and monitoring of district heating. Preceding the project described here, no dedicated system for monitoring of wind power existed at Göteborg Energi and the monitoring was considered complicated and insufficient.

3.1 Kasper

The primary purpose of the tool Kasper was to ensure the quality of the measurements of the deliveries of district heating, which previously had been deficient. With the program it is possible to identify faults and measurement errors. Since the start up, the program has made it possible to debit costumers afterwards when they had been paying for less than actually delivered, due to faulty measurements. Previously, it was not possible to debit afterwards and this led to losses of revenues where the measured consumed energy was lower than the actually supplied as the measurement error could not be detected.

The system creates energy signatures, one signature for a reference period where the measures are assumed to be correct and another energy signature for a period to be studied. The energy signature links a certain daily averaged temperature to a certain energy demand for a building by using historical data. The reference period is usually one year in order to cover the whole temperature range. A new energy signature can be created for another period and the periods can thereby be compared even though the temperature is diverging between the periods. From the data of daily heat deliveries (MWh) and data of daily averaged temperature (ºC), a regression line is determined, creating the energy signature. Figure 3.1 is showing the energy signatures of an example building, one for the reference period (light blue) and the other for a more recent period (dark blue). The horizontal axis is showing the temperature (ºC) and the vertical axis is showing the daily deliveries of heat. The energy signatures consist of four linear regression lines, calculated using the least square method. The positions of the four knots on the horizontal axis are predetermined and are the same for all buildings.

Each night, meter readings for delivered energy of district heating and daily averaged temperature for all properties within the district heating network are collected. Those values are compared to the reference signature and its expected delivery of district heating for the current temperature. The heat demand can differ significantly between different days since activity in the building might be different e.g. in an industrial building or school where the demand is reduced during the weekends. Due to this, Kasper can handle up to three different reference signatures and three different signatures of a studied period. Each value is compared with the closest energy signature.

There are two different kinds of alarm functions within the system: point deviation alarm and sliding alarm.

A point deviation alarm will be initiated when the difference, between the measured value,y_i, and the expected value $\hat{Y}_{x_i}$for the actual temperature x_i , exceeds a certain deviation

$$abs(\hat{Y}_{x_i} - y_i) > 3.5\sigma \qquad (2.9)$$

where σ is the standard deviation determined in the context of reference-signature calculation. The magnitude and the number of alarms during the last 90 days are counted by the system. The properties can be ranked according to the magnitude of the deviations measured during the last day or according to the number of alarms for a property during the last 90 days. In order to not ignore smaller properties, with a relatively low heat demand, the deviations can also be normalized against the reference value and ranked according to this.

The sliding alarm identifies deviations in the energy signature over the last 365 days compared to the energy signature of the reference period. Those deviations can be ranked according to the total deviation of energy deliveries during the period or by the rate at which the signature is changing. The worst alarms will be handled and investigated first. The deviations of signature might be caused by changes in the use of the building e.g. if the building has not been in use during a period. In cases when the changes cannot be explained with the use of the building, it is likely that something is wrong with the gauge and that more, alternatively less, heat has been delivered than what the meter has recorded. The meter is then replaced and the old one is sent to a laboratory where its accuracy can be controlled. If the meter has registered too low values of supplied heat during a longer period, Göteborg Energi can debit the costumer afterwards. The amount of supplied heat that has not been charged for can be estimated by the software through comparing the measured values to the reference line.

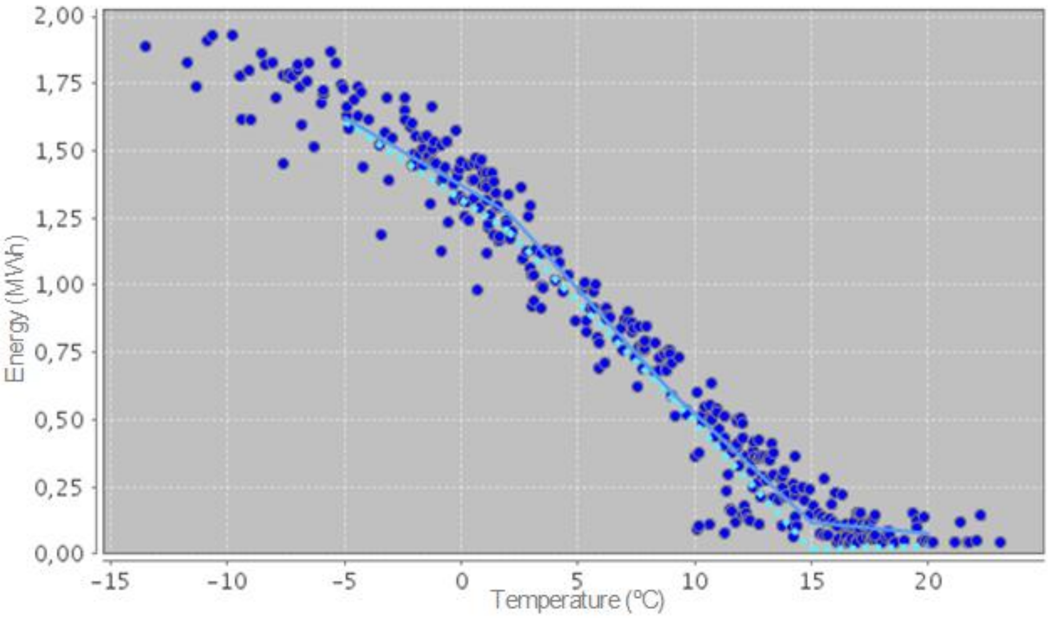

Figure 3.1 Graph of supplied energy at different temperatures for a certain building obtained with Kasper. Dark blue dots are measured values from the last year. The dark blue line is the regression line resulting from the data points, the line of light blue dots is the reference line obtained from historic data.

3.2 Available data for wind power generation

Many wind turbines owned by Göteborg Energi are today supervised by the turbine manufacturer or external staff. Operational data for these can be monitored from the manufacturer's server but has not been available for analysis in this study. Instead data from four wind turbines, all of them of the type Vestas V44, have been used for analysis. These are all placed at Risholmen in the western part of Göteborg. Figure 3.2 shows a map of this area. The turbines are named with Swedish names, Boel, Eivind Elida and Elin. Göteborg Energi owns a fifth turbine at Risholmen, Marta. This is a Bonus Mk 3 turbine with lower capacity, 450 kW instead of 600 kW and another control strategy, stall regulation instead of pitch regulation. The stall regulation strategy prevents the same method of normalization to be used and since all newer turbines are pitch regulated, the developed model has not been adapted to this control strategy.

Data for the wind power production of the four turbines is collected every second. From these data, minute average and hourly averages are created. All data, including the momentary values, are long-term stored in a server. A selected part of data is also available through an operational management system; momentary values from every ten minute from the two last weeks and hourly average from the start-up of measurements are available and can be presented in time history plots. The parameters available for each turbine are presented in *Table 3.1*. The wind speed for each turbine is measured by means of the anemometer on top of the nacelle, thus behind the rotor. Also the wind direction is measured at each wind turbine in order to enable yaw-angle control but currently, the signal is not stored. Instead data of wind direction measured at a wind mast located close to the turbines has to be used, “Measuring point” in Figure 3.2. Since the rotor is affecting the wind speed, it may be assumed that it is advantageous to use the wind direction measured at the measuring mast. This is also the case for wind speed. The wind speed downstream the rotor is always lower than upstream the rotor when the turbine is operating since energy is extracted from the wind over the rotor. Other parameters of interest, where no data measured at the nacelle have been available, are the air temperature and the atmospheric pressure. It can be assumed that the temperature and pressure will not vary significantly locally within the area of the five turbines, thus, ambient temperature and atmospheric pressure data collected at Rya heat plant located 5 km east of Risholmen has been used in this work.

During the investigation of data, it was found that the average values of wind direction had been derived incorrectly. Due to this, new values for minute average and hourly average had to be recalculated from the momentary values in order to make it possible to use the direction data. Measurement series for the pitch angles were found but it was concluded that these values were incorrect as they were just a fraction of the power produced and took unreasonable values. These series could therefore not be used for any analysis. Göteborg Energi is now aware of the problem and will measure and store the correct values in the future.

The five turbines considered in the study started to operate in 1996. However, data is not available until July 2008 for Boel and Eivind and until November 2008 for Elida, Elin and Marta.

Table 3.1 Data available for the wind turbines at Risholmen, Göteborg.

	Boel	Eivind	Elida	Elin	Marta	Wind mast	Rya HP
Power	X	X	X	X	X	-	-
Wind speed	X	X	X	X	X	X	-
Wind direction	-	-	-	-	-	X	-
Generator temp.	X	X	X	X	X	-	-
Gear bearing temp.	X	X	X	X	X	-	-
Gear oil temp	X	X	X	X	X	-	-
Pitch angle	-	-	-	-	-	-	-
Air temperature	-	-	-	-	-	-	X
Air pressure	-	-	-	-	-	-	X

Figure 3.2 Map of Risholmen showing the location and names of the wind turbines and the wind mast included in the study.

4 Identification of influencing parameters

The analysis of influencing parameters was performed using data from the database at Göteborg Energi. A literature review of the field was also conducted in cases where the available data was insufficient, and in order to support the conducted analysis.

The wind turbine should ideally be placed on a site without local topographic features and obstacles which can affect measurements. In other case, which is the case at Risholmen and most other sites, site calibration is needed. The criteria for this are stated in the IEC standard (2005) for wind power performance. Series for wind speed, wind direction, atmospheric pressure, air temperature and power output shall be measured.

4.1 Wind speed

As the power output is a function of the cube of the wind speed, this is the most important parameter determining the power. In Figure 2.1, the ideal power curve of a Vestas V44 provided by the manufacturer has been shown.

For obtaining correct measurements of wind speeds, the anemometer shall be placed between 2 and 4 rotor diameters from the turbine and it must not be placed in the wake of the turbine (Burton et. al., 2001). These criteria are not fulfilled by the mast at Risholmen; the distance between the anemometer and the closest turbine to be investigated, Eivind, is 380 m and thereby significantly more than 4 rotor diameters (176 m).

According to Cutler et al. (2011), an average of the measured wind speed at the nacelle and the wind speed at the mast gives the best fit to the power curve supplied by the manufacturer. Important when measuring wind speed for wind power purposes is that the measures are calibrated for the height of the nacelle since the wind speed will vary with the height above the sea. According to the IEC standard (2005), for test purposes and the elaboration of power curves, the wind speed shall be measured with a cup anemometer at a mast on the same height as the hub. Correction for height can in theory be done by taking into account wind shear (the increase of mean wind speed with height) by using the wind profile power law relation,

$$v = v_r \left(\frac{z}{z_r}\right)^{\alpha} \tag{4.1}$$

where v is the wind speed at height z, v_r is the known wind speed at the reference height z_r and α is a empirically derived coefficient, which usually takes on values around 0.14, but varies dependent on the terrain (Burton et al., 2001). The importance of measuring at the right height is increasing when the surface roughness is high. The surface roughness is dependent on the topography of the surroundings and if obstacles, such as buildings, are present in the proximity. A high surface roughness causes a high wind shear which implies that the horizontal wind speed is reduced at ground level and the nominal, undisturbed, wind speed is reached higher in the atmospheric layer (Peña, 2009).

Figure 4.1 shows the power generated by the turbines at Risholmen plotted against the wind speed measured at the wind mast. Figure 4.2 shows the power plotted against the wind speed measured at the nacelle of the turbines. When using data from the wind mast the deviations of the power are significantly higher. When data from the anemometer on top of the nacelle is used, a clearer shape of the curve is obtained. An analysis of the winds speeds showed that the wind speed at the mast was significantly higher than at the nacelle for all the four turbines, approximately 30 % higher at the wind mast.

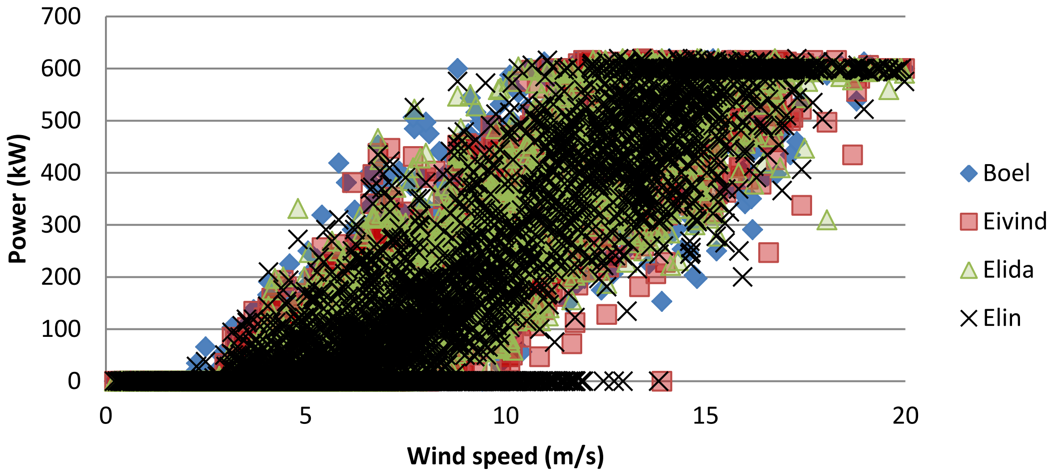

Figure 4.1 Power generated plotted against wind speed measured at the wind mast.

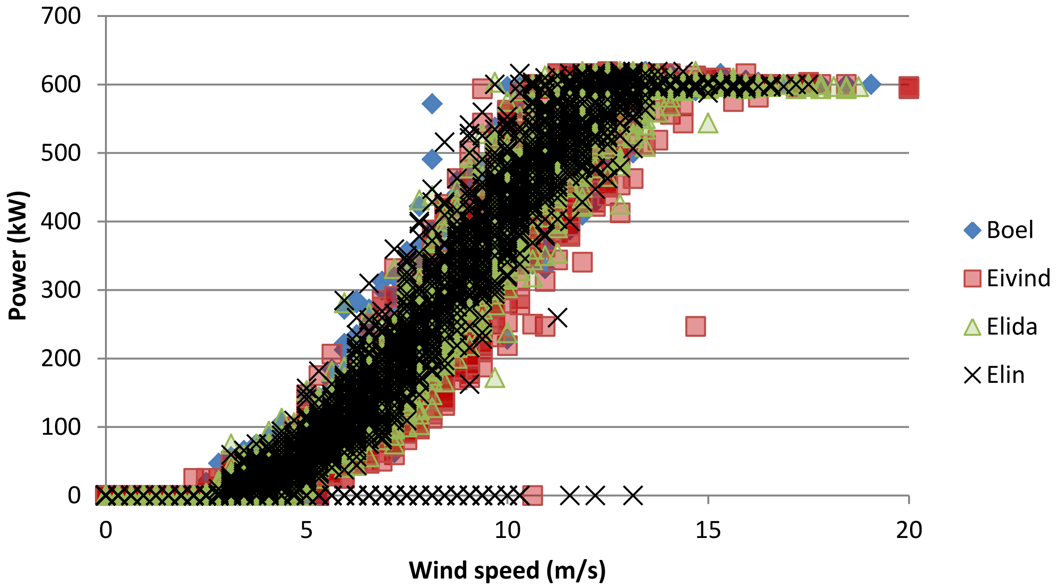

Figure 4.2 Power generated plotted against the wind speed measured at respectively turbine.

4.2 Wind direction

The wind direction range, measured from 0 – 360 degrees, was split up into twelve sectors as shown in Figure 4.3. The prevailing wind directions at the wind mast at Risholmen were found to be south-west and west (SSW, WSW and W, see Figure 4.3). These three wind directions are all four times more frequent than northerly winds. This is not surprising as south-west is the prevailing wind direction in this part of Sweden in general (Wern & Bärring, 2009).

Figure 4.4 shows that the rated power has not been obtained in all wind directions for Boel, e.g. north-north-east (NNE) where the highest measured power is about

500 kW. It can therefore be concluded that difficulties will occur when trying to parameterize the power curve for those directions, due to the lack of data. Rated power is mainly reached in south-westerly (SSW and WSW) and westerly (W) winds.

From Figure 4.5 it can be concluded that the wind direction influences the power output for a certain wind speed. The directions that yield best power output are south-south-west (SSW) and southeast, while north-westerly winds are least beneficial. This is displayed in the figure as the same wind speed yields a higher production for south-south-westerly (SSW) winds compared to e.g. north-north-westerly (NNW) winds. Its power curve is therefore shifted to the left.

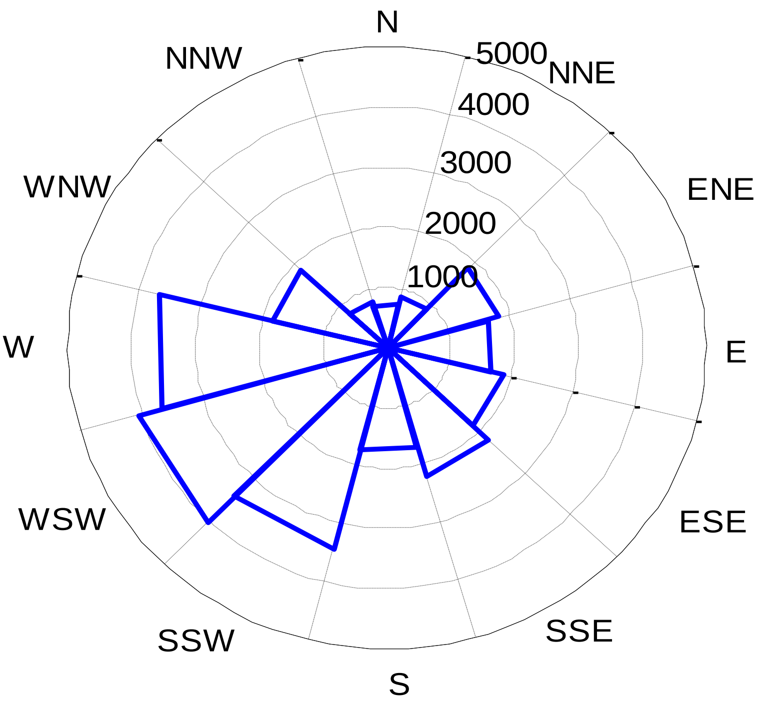

Figure 4.3 Distribution of wind frequency over twelve sectors during June 2008 -March 2009.

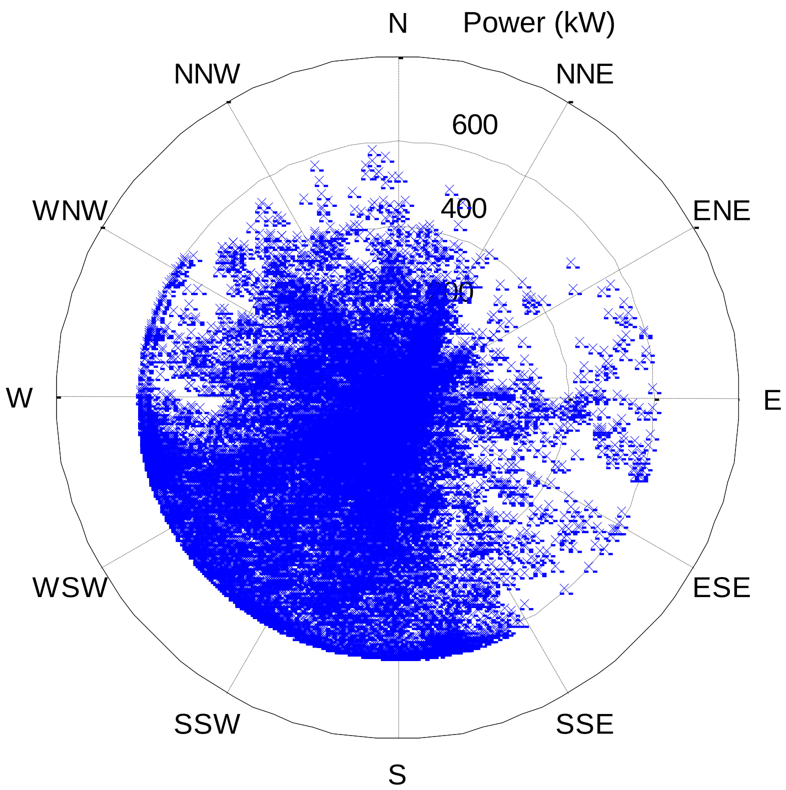

Figure 4.4 Distribution of wind directions of measured power from Boel during June 2008 until March 2009.

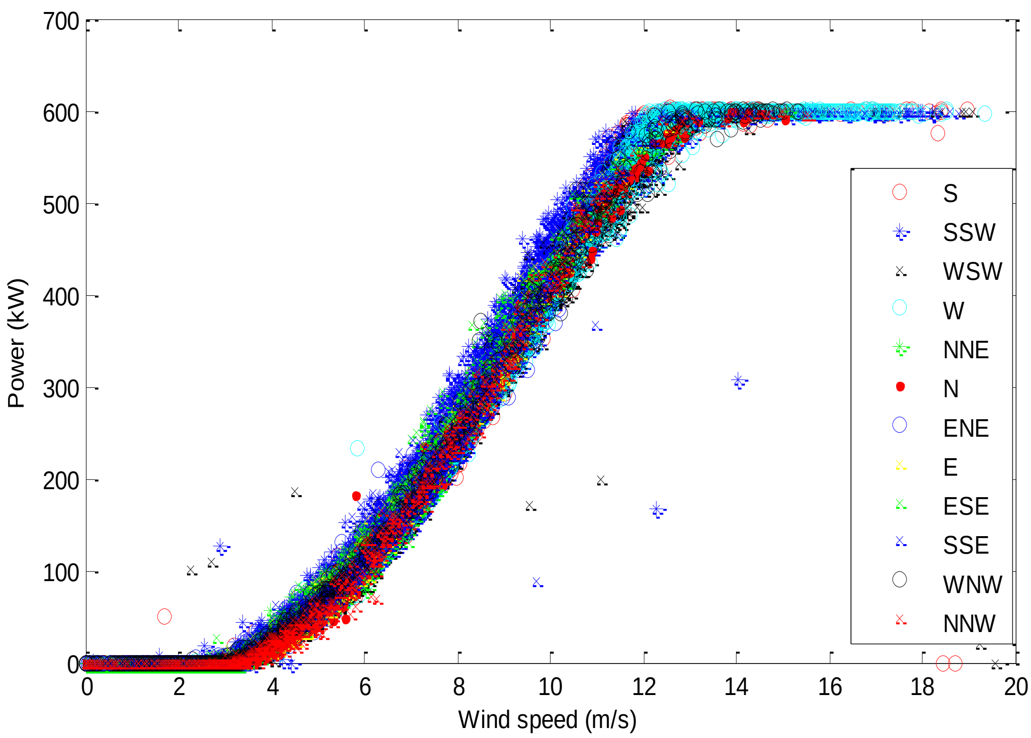

Figure 4.5 Power output for each wind direction from the wind turbine Boel.

4.3 Turbulence

Turbulence is characterized by unstable flow in which local speed and pressure change unpredictably, in contrast to laminar flow. This gives rise to fast fluctuations in wind speed. These fluctuations in wind speed can be caused by either friction with the earth's surface or by the impact of obstacles such as buildings or trees.

Downstream of a wind turbine, a wake will always occur since some energy is extracted by the turbine (Burton et. al, 2001). In the wake, the turbulence is high and this is a major factor when planning wind farms. The turbines affect each other and the turbines behind the first row will face a lower wind speed with higher turbulence than the first row. Power losses due to wake effects in wind farms are on average 10 % and the turbulence generated by wakes causes fatigue loads on the turbines and thus reduces their lifetime (Barthelmie, 2007). A distance of 2 – 4 rotor diameters between turbines is suggested as a minimum in wind farms in order to reduce these effects (Petersen & Madsen, 2004). This is sustained between all turbines at Risholmen, where the shortest distance between two turbines is 250 m, while the rotor diameter is 44 m.

Turbulence is measured in turbulence intensity,$T.I$, and has a percentage value.

$$T.I. = \sigma_v / v \tag{4.2}$$

Herein, σ_v is the standard deviation of the momentary wind speed variations, v_i, at a specific location over a specified period of time,

$$\sigma_v = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(v_i - v)^2} \tag{4.3}$$

where v is the mean velocity at the same location at the same period of time. In the ideal case of perfectly laminar flow, the value would be zero. The turbulence intensity may even be higher than 100 %. Typical values at hub-height onshore are 10-12 %, while offshore typical values are around 6-8 % (Barthelmie, 2007).

Gusts and turbulence affect the power generation negatively since it causes large variations in loads at the turbine. The surface roughness is lowest over the sea, at least in absence of high waves. Due to this, the turbulence intensity for the turbines placed at the shore of Risholmen will vary around the turbine. To quantify the turbulence, momentary data at ms resolution is required. This has not been available and no investigations could be done to further analyse those effects. The turbulence intensity can be assumed to be relatively constant for a certain wind direction at a given wind speed (Burton et al., 2001). By having sufficiently narrow wind direction sectors, the turbulence effects preventing the construction of a normalized power curve can be eliminated since these effects are constant for each direction.

4.4 Air density

The density of a gas is determined by the ideal gas law. An equation for this relation was introduced in section 2.1,

$$\rho = p/RT \tag{2.2}$$

where ρ is the density of the gas, p is the pressure inPa, R is the gas constant for dry atmospheric air, 287 J/K kg, and T is the temperature of the gas in K. Hence, the density will decrease with an increasing temperature and rise with an increasing pressure. This, in turn, implies that the power output is expected to be lower at higher temperatures as the power is a function of the density, see equation (2.1). In Figure 4.6, power output is shown for different temperature intervals and it can be clearly seen that higher temperatures yield lower power output. At the lowest temperature during the period (-16.4°C), the density of the air is approximately 17 % higher than for the highest temperature during the period (29.0°C), and thus shall the power output be 17 % higher for the lower temperature at a certain wind speed. The data used for Figure 4.6 is hourly data from three years.

The influence of the variations in atmospheric pressure, which is the other variable parameter in the gas law, was also investigated. The difference between the highest and the lowest pressure during the same period was 79.38 hPa. This corresponds to 8 % higher density for the higher pressure and an expected increase in power output of approx. 8%. When power is plotted against wind speed for different pressure intervals, Figure 4.7, no correlation between higher pressure and higher power is found, confirming that the temperature variation has a greater influence on the power output than the pressure variation.

At high temperatures, also the relative humidity affects the density of the air. The effects become significant at a temperature of around 30°C (IEC, 2005). The calculations for obtaining a density normalized regarding humidity as well are more complicated than the multiplicative correction necessary for temperature and pressure (IEC, 2005). It can be assumed that these effects can be neglected in Swedish climate as the temperature only exceeds 30°C a few times a year.

The IEC standard (2005) is suggesting that for active power controlled turbines, a normalized wind speed shall be calculated according to

$$v_n = v\left(\frac{\rho}{\rho_{ref}}\right)^{1/3} \tag{4.4}$$

where v_n is the normalized wind speed, v is the averaged wind speed measured during 10 minutes, ρ is the averaged density calculated according to Eq. (4.4) for data measured during 10 minutes and ρ_{ref} is the reference density. The measured power output can thereby be compared to its normalized wind speed.

For stall regulated turbines, normalized power shall be calculated (IEC, 2005), by means of

$$P_n = P\frac{\rho_{ref}}{\rho} \tag{4.5}$$

where $P_n$is the normalized power output and P is the averages power over 10 minutes. In Figure 4.8, the two ways of normalization, both normalized wind speed and normalized power, has been applied to data obtained at Boel, which is an active power controlled turbine. For simplicity reasons, reference condition was set to 10°C (283 K) and 1 bar (100,000 Pa).

$$\rho_{ref} = \frac{p_{ref}}{RT_{ref}} = \frac{100{,}000\,Pa}{287\,{}^{J}/_{kg\,K}\;283\,K} = 1.23\ \mathrm{kg/m^3} \tag{4.6}$$

It was found that the normalization of wind speed resulted in a slightly narrower and more accurate curve then for normalized power, Figure 4.8.

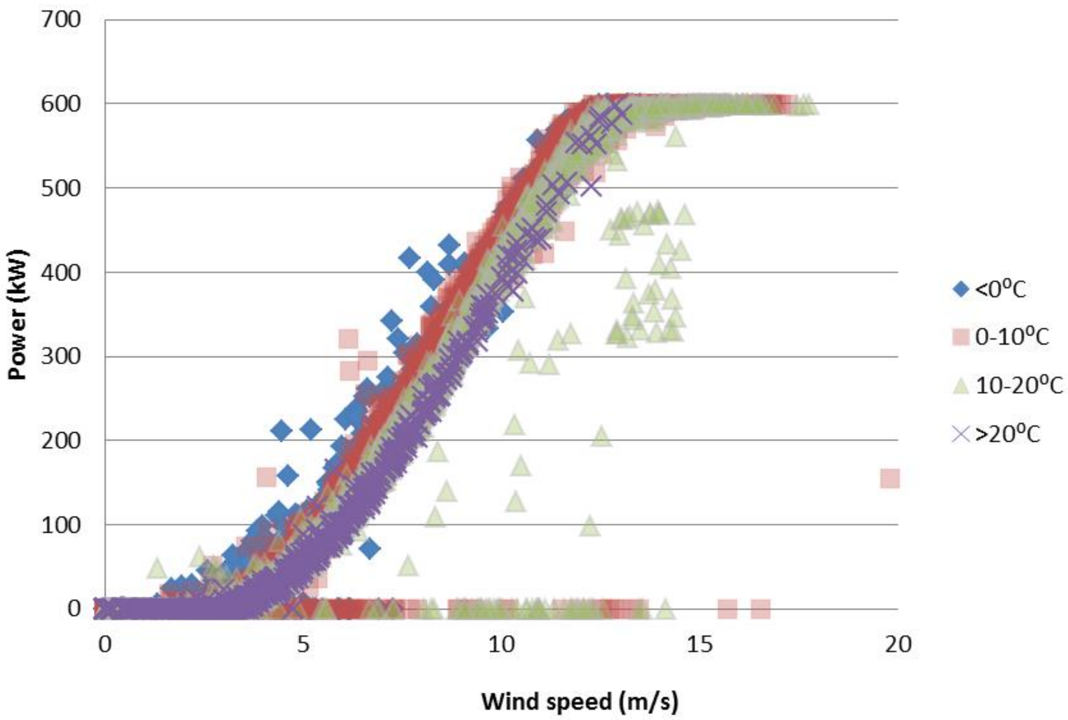

Figure 4.6 Power output from the wind turbine Boel for different temperature intervals.

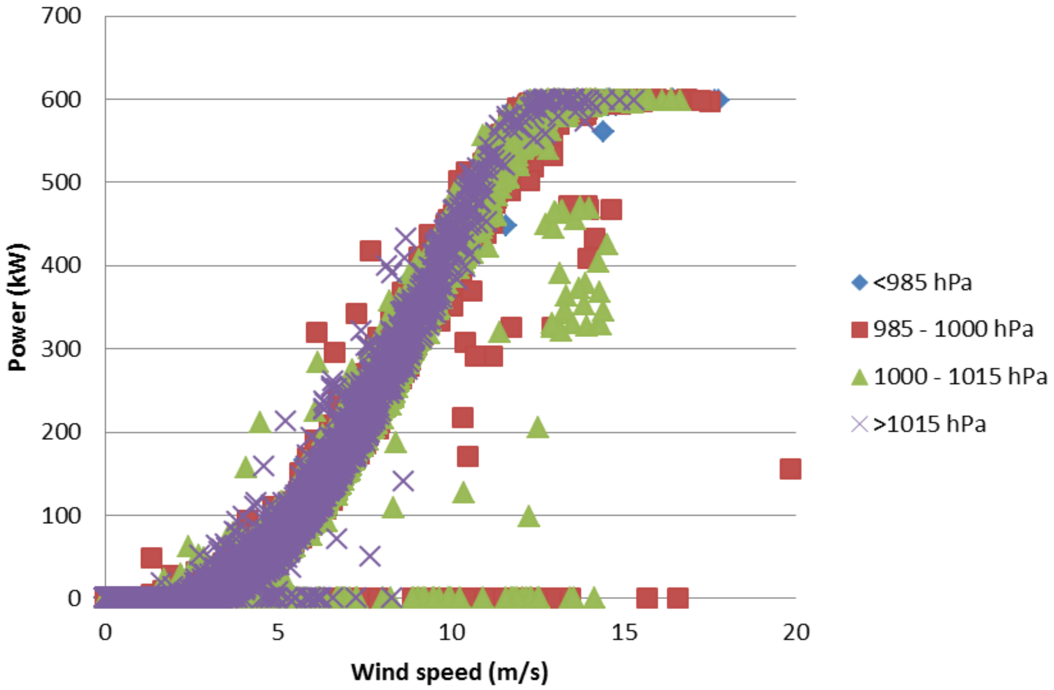

Figure 4.7 Power output for different pressure intervals.

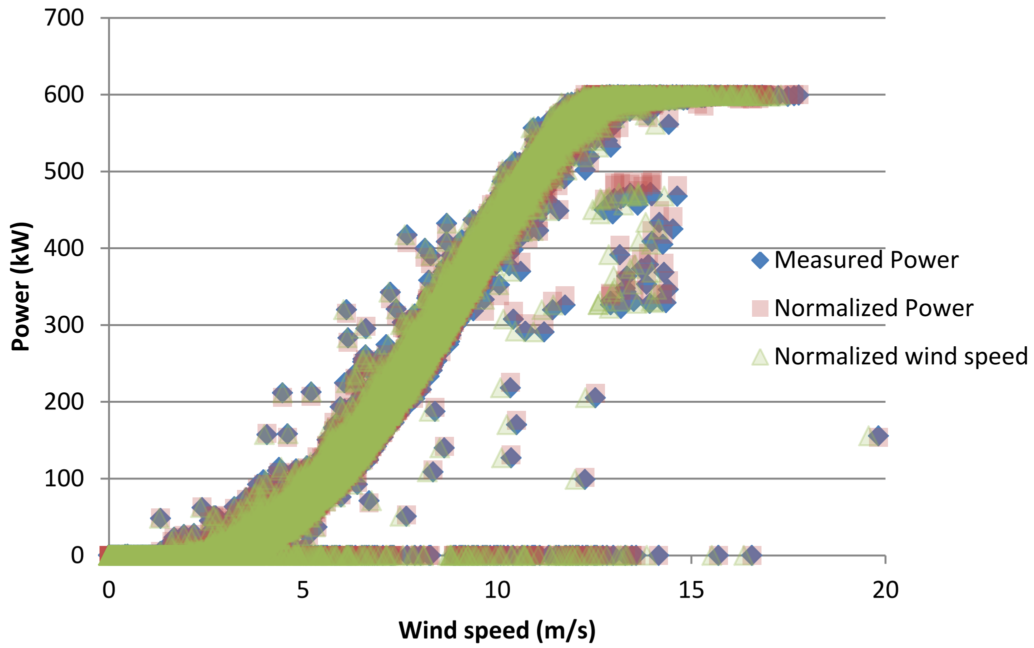

Figure 4.8 Power curves for measured power, normalized power and normalized wind speed.

5 Model development

The objective is to develop a tool where data of the production from one period of time can be compared with a data model representing a reference signature. The reference signature shall be a power curve created from data obtained during a period when the turbine is assumed to have operated without any disturbances or faults. In order to make data comparable, selection and normalization of data are needed as described in Chapter 4.

The data contain erroneous values due to e.g. gauge errors or data generated during start-up or shutting down of the turbine. Such values should not be included in the construction of the reference signature and therefore, a filtering method for the rejection of this data is needed. After the filtering and selection of data have taken place, a parameterization of data is carried out in order to obtain a reference signature. Three different methods for parameterization are investigated using Matlab. The data used when developing the methods was 10 minute average values. This data was created from minute average values, which in its turn was created from momentary values of 1 Hz. This is in line with IEC standard (2005) which states that the sampling rate has to be at least 1 Hz and that the data shall be presented as 10 minute averages. Alternative averaging periods of one minute and hour have been considered but were rejected as the amount of data would have become either too extensive or too small. It was also found that minute averages resulted in significant higher variation of power for each wind speed, which would aggravate the parameterization.

To compare data from two different periods, a signature based on data from another period shall be constructed using the same method as for the reference. Thereby, long term trends can be identified and an evaluation of the performance can be conducted. The tool shall also include a function to detect data deviating from the reference signature. When this occurs, an alarm shall be initiated and hence, degradation of turbine performance can be detected.

5.1 Model approach

In Chapter 4, it was found that several variables affect the power generation. From these results, the following variables were chosen to be included in the model:

- Wind speed
- Power
- Wind direction
- Air temperature
- Atmospheric pressure

To get as high accuracy as possible, all relations found will be combined in order to give a narrow range of acceptable power output for a certain wind speed.

IEC 61400-12 recommends using the wind speed measured at a wind mast but in Section 4.1 it was found that when the wind speed measured at the nacelle was used, the variations of power were significantly lower. Since the performance of the wind turbine will be compared to its own in the past, it is assumed to be of advantage to use the wind speed at the nacelle even though this is lower than the actual free stream wind speed at hub height.

It was found that the power output is different depending on wind direction. It was also found that high wind speeds are rare or even absent in some directions thus, for these directions a parameterization of the data cannot be fulfilled. Winds from the four sectors, S, SSW, WSW and W, were found to be the most frequent directions, more than half (54%) of the data was measured in these sectors. High wind speeds, and thus rated power, were also found in these sectors. Therefore, the models were designed to include data measured in these directions, creating separate signatures for the four directions. Due to the high frequency of these wind directions, it is considered sufficient to base the monitoring procedure on these directions only. An additional option, the monitoring based on data from all wind directions can be included in the application in order to visualize even the remaining directions.

The data processing can be divided into the following steps:

1. Validation to ensure that all variables exist for a given time
2. Selection of wind directions to be included in the analysis
3. Calculation of normalized wind speed, v_n, according to Eq. (4.6)

5.2 Detection and rejection of erroneous data

When looking at any of the figures showing the power versus wind speed, it is obvious that not all data should be included in the calculation of the reference signature. Only data representing normal operation should be included. There is data measured at high wind speed when the turbine has not been in operation, having a power output of zero. Such values will destroy attempts of creating an accurate reference signature using regression. As the values to be used are average values, the turbine may be in operation for a part of the time when a value is formed, and out of operation for the rest of the time period. This phenomenon is mainly present, and has effects, for wind speeds around the cut-out speed. In Figure 4.8, such data is found for wind speeds above the rated speed 16 m/s and is resulting in a power data lying between zero and the curve representing normal operation.

The simplest method for rejection of erroneous values is based on the "method of bins" which was outlined in Section 2.3. The method of bins was applied to data sets obtained after the data treatment outlined in Section 5.1. It was found, that an appropriate level for rejection was $5\sigma_j$, where σ_j is the standard deviation of the power data in bin j according to Eq. (2.8). Two curves, one upper limit, Eq. (5.1), and one lower limit, Eq. (5.2), serving as rejection thresholds were created by interpolation of the bin mean power plus respectively minus, $5\sigma_j$, versus the bin averaged normalized wind speed, $v_{n,j}$:

$$Upper\ limit = P_j + 5\sigma_j \tag{5.1}$$

$$Lower\ limit = P_j - 5\sigma_j \tag{5.2}$$

The suggested one-step filtering based on the method of bins appeared to be insufficient at high wind speeds where several data points are generated when the turbine only has been in operation for a part of the ten minutes period. The amount of data in each bin at high speed is limited, hence these erroneous values affect the average in the bin substantial and these points are not rejected even though they do not represent normal operation. To make sure that such data is rejected, an extra condition for rejection is included;

If $v_{n,i}$ > rated speed and P_i < (rated capacity − 100 kW) data shall be rejected.

This corresponds to 16 m/s respectively 500 kW for Vestas V44.

To further improve the data quality, the binning and filtration process is repeated for only that data which was accepted in the first filtration, using the same condition for rejection. The result after the whole filtering process of south-south-westerly winds is shown in Figure 5.1. In the figure, there is still one obvious erroneous data among the accepted values, v_n = 17.8 m/s, P =520 kW. A third filtering would not have rejected this data as no data in this bin was rejected in the second filtering and hence, the alarm threshold would have remained the same after a third filtration.

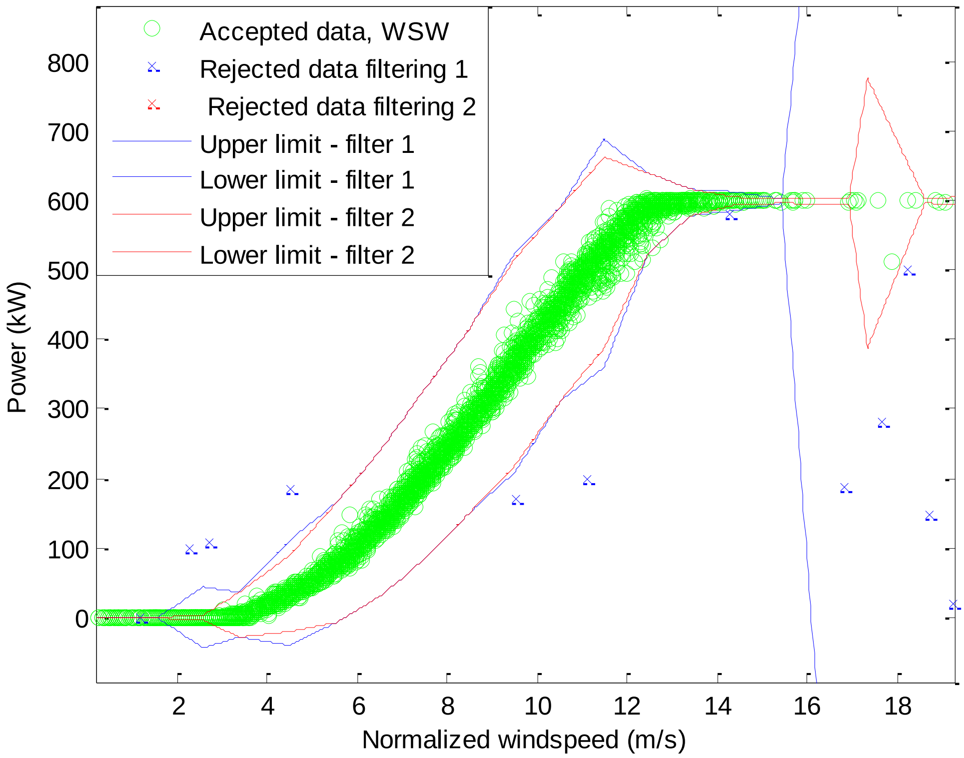

Figure 5.1 Accepted and rejected data after the iterative filtering process for data measured in SSW direction at Boel.

5.3 Parameterization of the reference signature

Three different methods have been investigated in order to create a suitable data model representing the reference signature:

- Piecewise linear regression
- Polynomial regression
- Linear interpolation based on the method of bins

In the current version of Kasper for district heating, the method used for the construction of the reference signature is piecewise linear regression, with three knots and four different lines. In order to reduce the modifications necessary for adaption to wind power, it would be beneficial if the same method could be used. Two different piecewise linear regressions, one with three knots and one including six knots, were compared to the two other alternatives to determine the most appropriate method. All methods were implemented using Matlab.

5.3.1 Piecewise linear regression – 4 pieces

The method used for district heating was adapted to the wind power data extracted after accomplished filtration. The regression is found using least-squares approximation. The algorithms compute the slope of each line and also the location of each knot at the vertical-axis (power). The locations of the knots at the horizontal-axis (wind speed) are set beforehand. The algorithms are solving an equation system of four equations, each representing one of the regression lines. It was found that a choice of 3.5 m/s, 7 m/s, 12 m/s and the maximum v_n within the amount of data as locations of the knots provided the best model fit. These values were chosen to fit as much as possible of conceivable data and to minimize the sum of the total sum of squared error, SSE_{tot}, for the four wind directions considered:

$$SSE_{tot} = SSE_S + SSE_{SSW} + SSE_{WSW} + SSE_W \qquad \text{with} \tag{5.3}$$

$$SSE_k = \sum_{i=1}^{n_k} \left(P_{k,i} - \hat{P}\left(v_{n,k,i}\right)\right)^2 \tag{5.4}$$

wherein these equations, $P_{k,i}$ denotes the measured power of data point i in wind direction k, n_k is the number of data points in wind direction k and $\hat{P}(v_{n,k,i})$ is the expected value from the created regression line at v_n in data point i for wind direction k. For the data used to create the regression in Figure 5.2, SSE_{tot}= 2,760,500 kW^2. To further get an overall measure of the goodness of fit and to get a measure to analyze how well the knots are compatible with all wind directions, a weighed deviation measure, δ_{tot}, over all wind directions was calculated,

$$\delta_{tot} = \frac{1}{n_s + n_{ssw} + n_{wsw} + n_w} \left(\delta_s n_s + \delta_{ssw} n_{ssw} + \delta_{wsw} n_{wsw} + \delta_w n_w\right) \tag{5.5}$$

where δ_k is a measure of deviation from the fitted curve of data measured in wind direction k, given by:

$$\delta_k = \sqrt{\frac{1}{n_k} \sum_{i=1}^{n_k} \left(P_{k,i} - \hat{P}\left(v_{n.k,i}\right)\right)^2} \tag{5.6}$$

Weighting of δ_{tot} allows δ_j of the most frequent wind direction to be most significant, hence the optimization reflects the total set of data.

From the figure, it seems like the poorest regression appears in the interval between the two first knots, 3.5 and 7 m/s, where the power curve is strongly non-linear. In this region, linear regression is insufficient. This is also were much data is measured and thus affecting SSE_{tot} most.

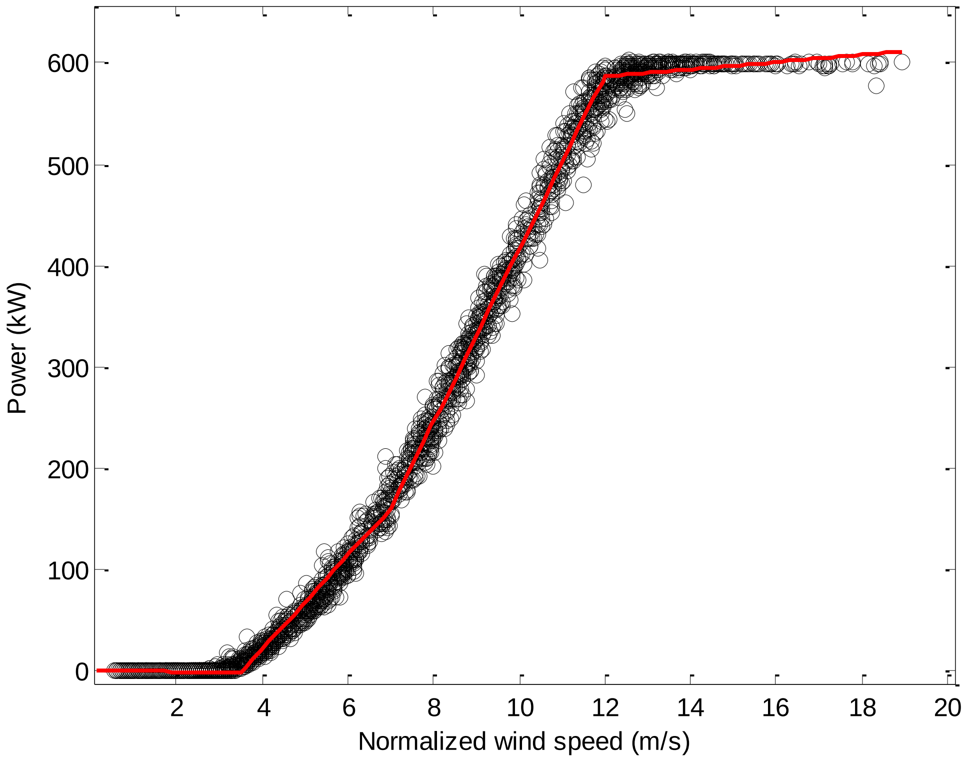

Figure 5.2 Four-lines piecewise linear regression of data obtained from Boel in southerly (S) winds measured between 2008-07-01 and 2009-03-15. Knots are set at 3.5, 7 and 12 m/s.

5.3.2 Piecewise linear regression – 7 pieces

It was assumed that, by increasing the number of pieces in the piecewise linear regression, the regression would be improved, thus a reduced SSE_{tot} and δ_{tot} would be achieved as well. With an increased number of pieces, the amount of data for each regression line is reduced; hence the sensitivity for erroneous data in the regression is increased. Shortage of data for each piece is another factor limiting the number of knots. It is desired that the same points for knots shall be compatible with all wind directions and for relatively small amounts of data. It was found that a number of seven pieces was most appropriate as it gives a satisfying approximation while it at the same time is compatible with smaller amounts of data. The algorithms for the four pieces linear regression were modified to generate a seven pieces, piecewise linear regression. Thereafter, the optimum location of the knots could be found using Eq. (5.3) and (5.5) as measures to compare different alternatives of knots. The results of the best alternatives found for knots are shown in Table 5.1. Figure 5.3 is showing piecewise linear regression of data from Boel with knots according to the third alternative in Table 5.1.

Table 5.1 Alternative locations of knots for a seven pieces linear regression.

Knot 1 (m/s)	Knot 2 (m/s)	Knot 3 (m/s)	Knot 4 (m/s)	Knot 5 (m/s)	Knot 6 (m/s)	SSE_{tot} (kW^2)	δ_{tot} (kW)
3.5	6	9.5	11.5	12.5	13.5	2,302,200	12.81
3.5	5.5	9	11.5	12.5	13.5	2,380,000	12.66
3.5	5	7	9	12	13.5	2,245,300	12.64

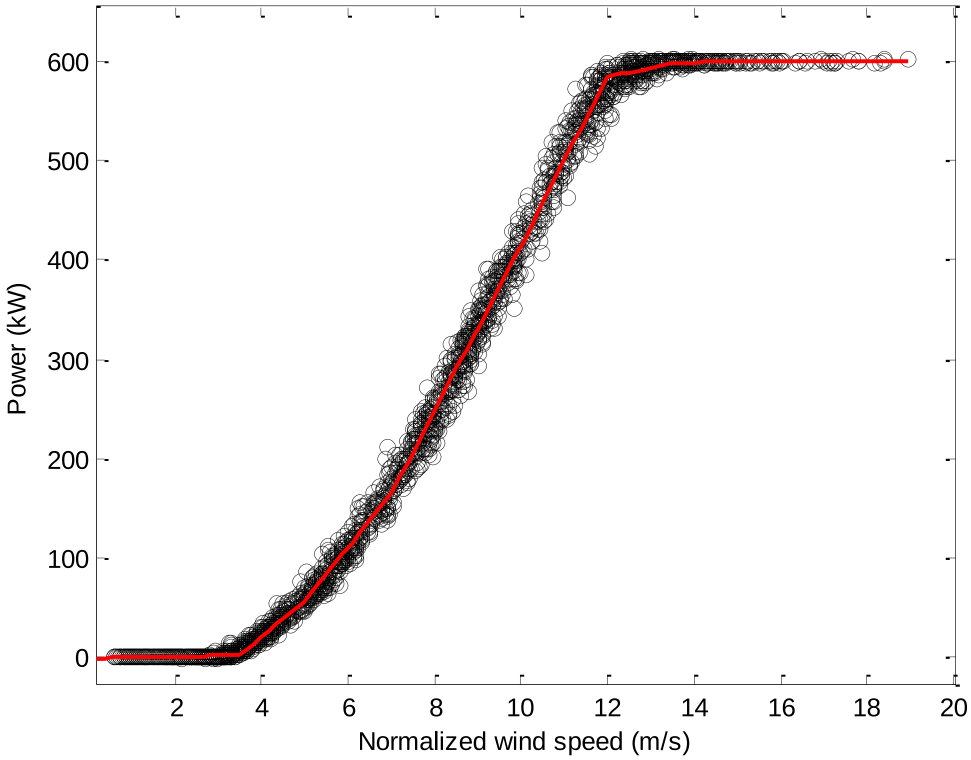

Figure 5.3 Seven pieces, piecewise linear regression of data measured in southerly (S) winds at Boel between 2008-07-01 and 2009-03-15.

The improvement compared to the four pieces regression, is mainly obvious in the regions 3 m/s-7 m/s and in the upper part of the curve. The reduction of SSE_{tot} when using seven pieces instead of four is 515,200 kW^2 (2,760,500 kW^2-2,245,300 kW^2). Also δ_{tot} is reduced, 12.64 kW compared with 14.13 kW when using four pieces regression.

5.3.3 Polynomial regression

Based on the form of the power curve, it was expected that one single third-order polynomial could fit the wind power data obtained for wind speeds in the non-horizontal parts of the curve, i.e. between 3 and 13 m/s for Vestas V44. Winds speeds below 3 m/s and above 13 m/s, do not need to be parameterized as the power in these regions are determined beforehand e.g. 0 kW and approximately 600 kW.

The data used previously, July 2008 – January 2009 for wind turbine Boel, was used again and the filter and detection method was applied in the same manner. A least-squares-based regression of the data was used to find the polynomial. The result is demonstrated in Figure 5.4.

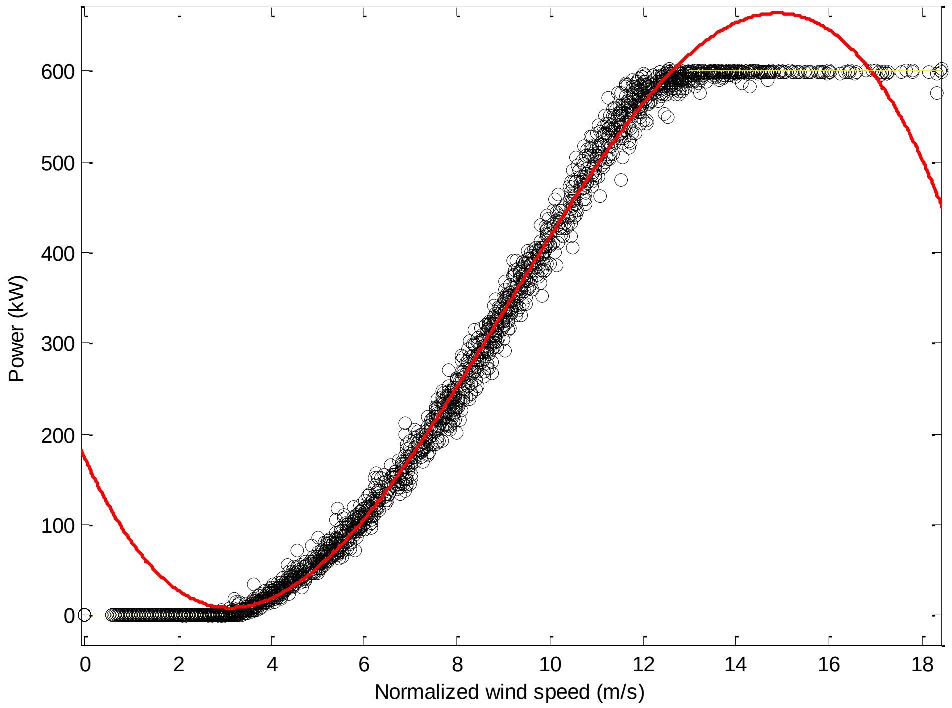

Figure 5.4 A polynomial of third order was used to parameterize the data obtained from Boel in southerly winds between July 2008 and March 2009.

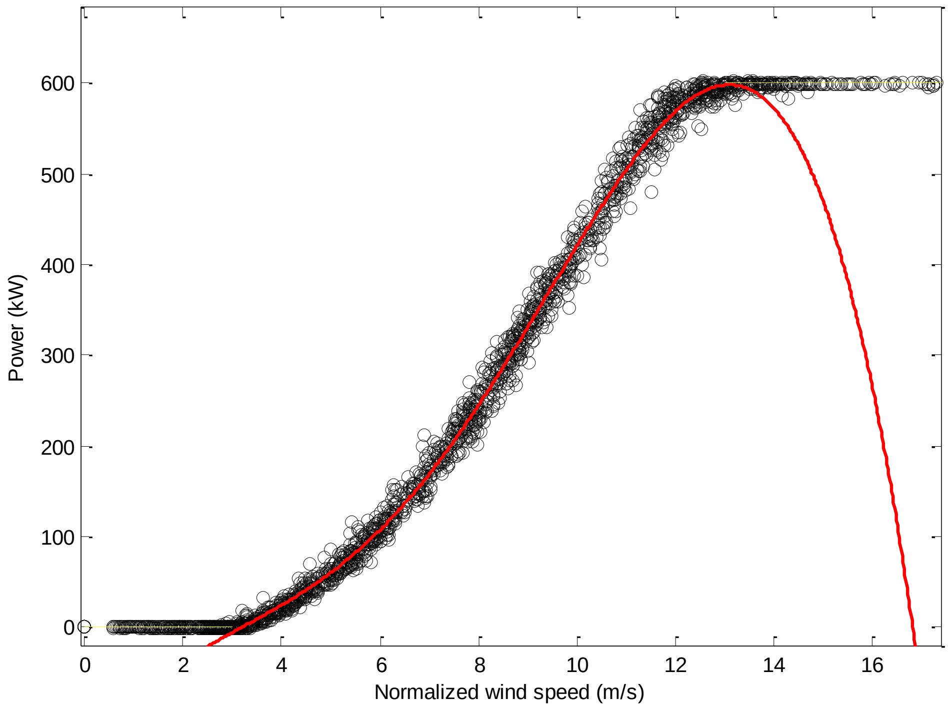

Figure 5.5 Fourth order polynomial used for regression of wind power data measured at Boel between July 2008 and March 2009.

The approximation turned out to be rather poor, $SSE_{tot} =$ 2,518,300 kW and δ_{tot} =15.59 kW^2. A trial using a fourth order polynomial instead was conducted; the result is shown in Figure 5.5. Both SSE_{tot} and δ_{tot} were slightly reduced, SSE_{tot} = 2,339,800 kW and δ_{tot} =14.10 kW^2. Even higher order polynomials were investigated, but the accuracy remained relatively constant and higher order polynomials were therefore rejected as alternatives for parameterization.

5.3.4 Method of bins

As suggested by the IEC standard (2005), the method of bins was used to parameterize the data in order to obtain a reference signature. After the filtering process, the bin average $v_{n,j}$ and P_j could be calculated from Eq. (2.6) and (2.7) again. Linear interpolation of the obtained mean values resulted in the reference signature curve found in Figure 5.6. Using this method with a bin width of 1 m/s provided values of SSE_{tot} = 2,233,200 kW^2 and δ_{tot} =12.62 kW.

According to IEC standard the width of each bin shall be 0.5 m/s. A modified model was developed, striving to meet the standard. In order to interpolate, at least one set of data is necessary for each bin. For some bins at high winds speeds, data was missing in some of the considered directions, S and WSW. For these cases, the bin width of the neighbouring bins was increased in order to enable a reference signature. This will not affect the accuracy since, above the rated wind speed; the expected power is relatively constant and independent of wind speed. A slightly higher accuracy was reached, SSE_{tot} = 2,164,500 kW^2 and δ_{tot} =12.40 kW.

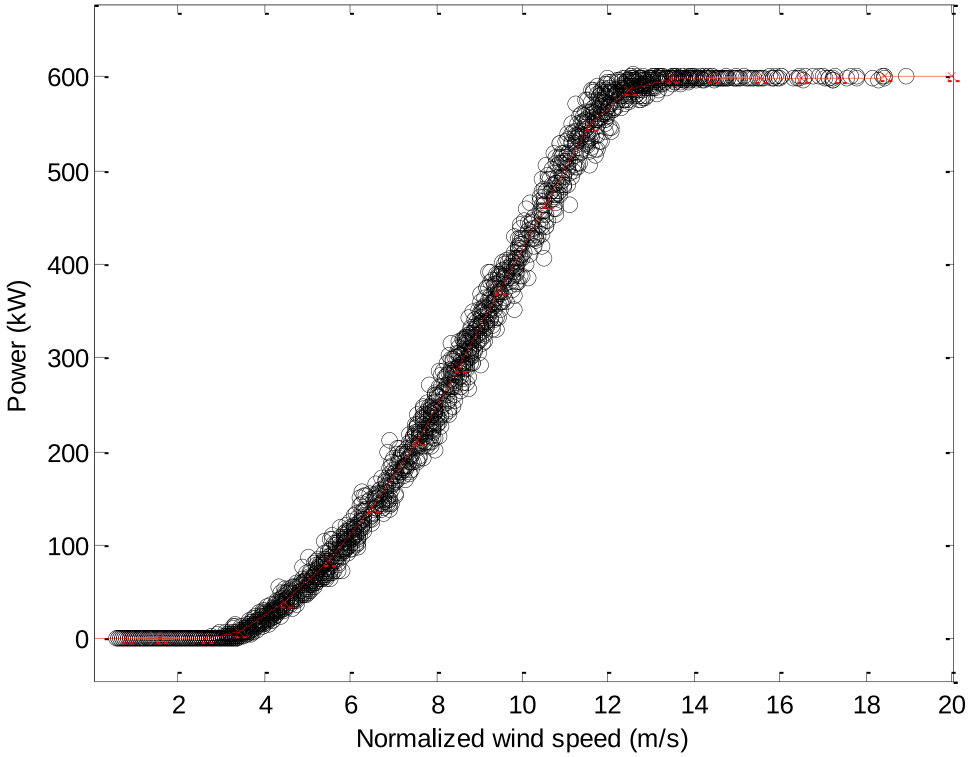

Figure 5.6 Parameterization of data for Boel, using the method of bins with bins of 1 m/s each. Data is measured in southerly (S) winds from July 2008 until March 2009.

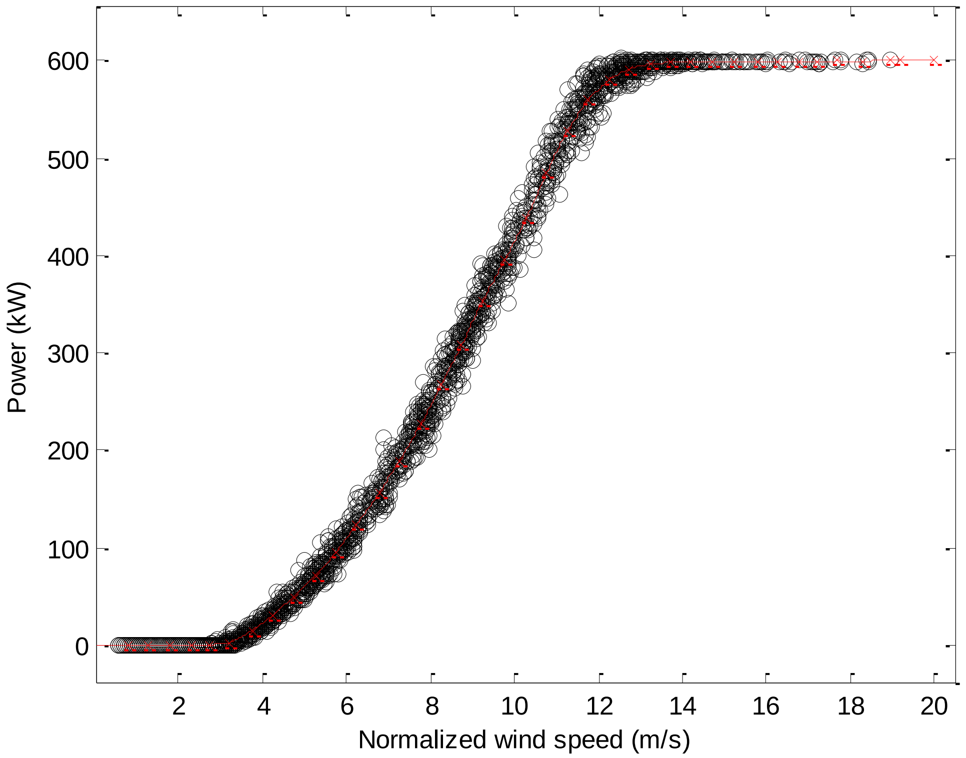

Figure 5.7 Parameterization of data for Boel, using the method of bins with bins of 0.5 m/s each. Data is measured in southerly (S) winds from July 2008 until March 2009.

5.4 Generation of alarms

In the application for district heating, several alarms can be initiated. Many of the alarm methods are based on point alarms, which detect points deviating more than 3.5σ from the reference line. By counting the number of point alarms during the last 90 days, or by calculating the total deviation in terms of energy of the point alarms, the detection of point deviations is an essential part of the system for controlling the district heat deliveries. A similar method for deviation detection used for the generation of alarms was therefore requested for wind power adaption as well. Depending on parameterization method, different detection methods can be used. Three main concepts were developed and applied to the parameterization methods, although some concepts are just compatible with one or two of the parameterizations.

The three concepts developed were:

1. Alarm limits based on δ , constant limits along the curve
2. Alarm limits based on piecewise δ, piecewise limits
3. Alarm limits based on method of bins

For each new data i added to the system, the deviation Δ_i is calculated according to Eq. (5.7). This is done irrespective of concept to be used:

$$\Delta_\mathrm{i} = \mathrm{P}_\mathrm{k,i} - \widehat{\mathrm{P}}(\mathrm{v}_\mathrm{n.k,i}) \tag{5.7}$$

Herein, $\mathrm{P}_\mathrm{k,i}$ is the measured power of data set i in wind direction k and $\widehat{\mathrm{P}}(\mathrm{v}_\mathrm{n.k,i})$ is the power of the reference signature at wind speed v_n in wind direction k. The filtering process will not be applied to the new data as erroneous data shall be detected and alarmed. However, alarms generated for data measured during start-up or shutting down are not desired as these points may influence the regression of the new data considerably. Therefore, some of these data points can be rejected by applying the additional criteria used in the filtration of the reference data, which was outlined in Section 5.2:

If $v_{n,j} > \text{rated speed and } P_j < (\text{rated capacity} - 100$ kW) data is rejected.

During the investigated period, no severe deterioration of the turbine occurred and therefore shall most of the data be considered as normal operation, hence just a few alarms are desired. To be able to compare the models and to analyse which model that is returning the most reasonable alarms, the thresholds were set to trigger some alarms even for the investigated data. If the accepted deviation is set too narrow, the number of alarm will be enormous which will thus prevent the possibilities to detect the really relevant deviations. The acceptable range shall neither be set too wide as this can prevent the detection of severe performance loss. A number of different levels of threshold were tested for each model.

In addition to the point alarms for monitoring of performance changes, performance degradation can be detected by comparing the signature of the newer data with the reference signature. The new signature is derived using the same method as for the reference signature. By comparing the two signatures, long-term trends of performance can be evaluated. If the number of alarm is high during one period, it can be expected that its signature will diverge from the reference.

5.4.1 Constant limits along the curve

In Kasper for district heating, the overall standard deviation along the entire range of temperature is used to define the limit for alarm generation. Aiming to resemble this, a similar alarm generation concept was developed for wind power. The developed concept is compatible with all parameterization methods.

After the wind direction separation, the deviation Δ_i is calculated for all data. From the data used to derive the reference signature, the deviation measure δ for each wind direction is calculated. An upper and lower limit for alarm initiating can thereby be determined as a multiple of δ. An appropriate level was, after testing, determined to be 4δ. The interval of accepted power, not causing an alarm, will thereby be constant along the whole wind speed range. The concept was applied to all of the three methods of parameterization. The same reference period and hence, also the same reference signatures, as in section 5.3 were used. The new data, added to the system was data from 2010-05-01 to 2010-10-01. In Figure 5.8, the concept is applied together with the piecewise linear regression method (4 pieces). The total number of alarms for the four wind directions became 25. Data points which generate alarms are marked with red in the figures and the limits for alarms are represented by the thin red lines. In Figure 5.9, the constant limit concept is applied together with the polynomial (3rd order) regression method.

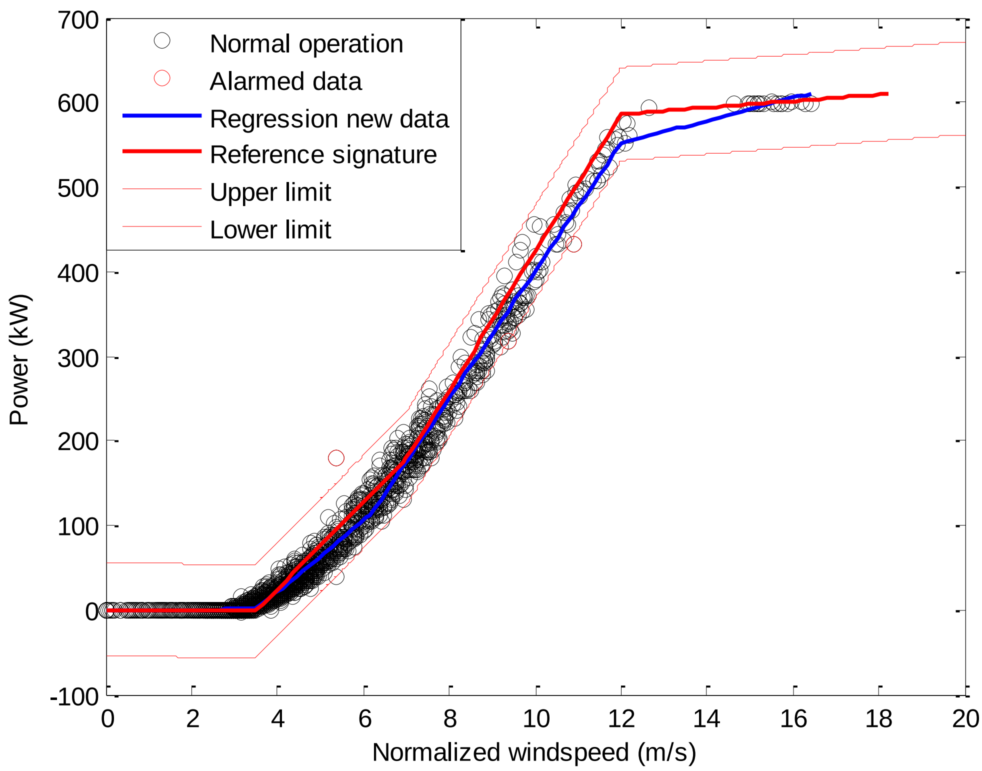

Figure 5.8 Piecewise linear regression (4 lines) with constant limit concept for alarm generation. The data displayed is measured in the SSW direction, May 2010 - October 2010. Data points outside of the upper and lower limit would cause an alarm.

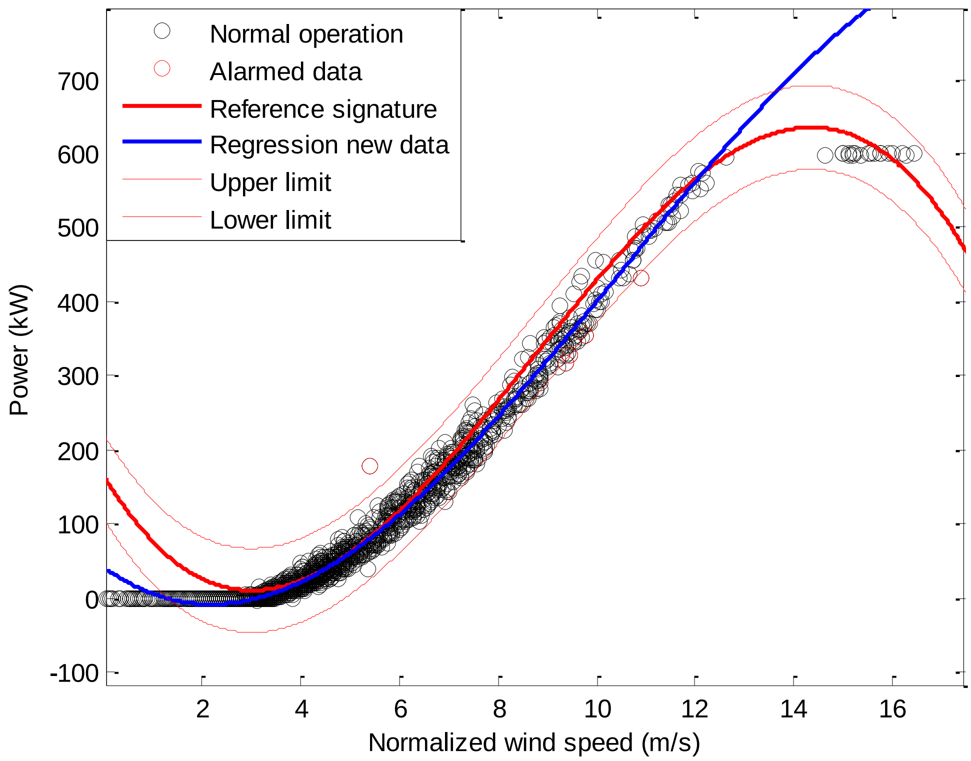

Figure 5.9 Polynomial (3rd order) regression with constant limit concept for alarm initiating. The data displayed is measured in the SSW direction, May 2010 - October 2010. Data points outside of the upper and lower limit are considered as alarm.

5.4.2 Piecewise limits

In some parts of the wind speed range, the difference between the highest and lowest power in the historical data is just a few kW, while in other regions the difference is tenfold. For the piecewise linear regression, it is possible to define different regions. Hence, the deviation measure δ can be calculated for each line. For regions with small differences, smaller deviations from the reference signature, Δ_i, will initiate an alarm and vice versa for regions with larger differences. It was found that the most appropriate level for alarm generation was 5δ, for both the four pieces regression and seven pieces regression. When using four lines for regression, hence four levels of alarms, alarms occurred in the regions around the knots, where the level of alarms was changed, for instance around 3.5 m/s in Figure 5.10. Several of those should be considered as normal operation but if the level for alarm is increased, the acceptable range is too wide in regions with higher δ. This problem was almost eliminated when the number of lines, and thus the number of different δ, was increased to seven.

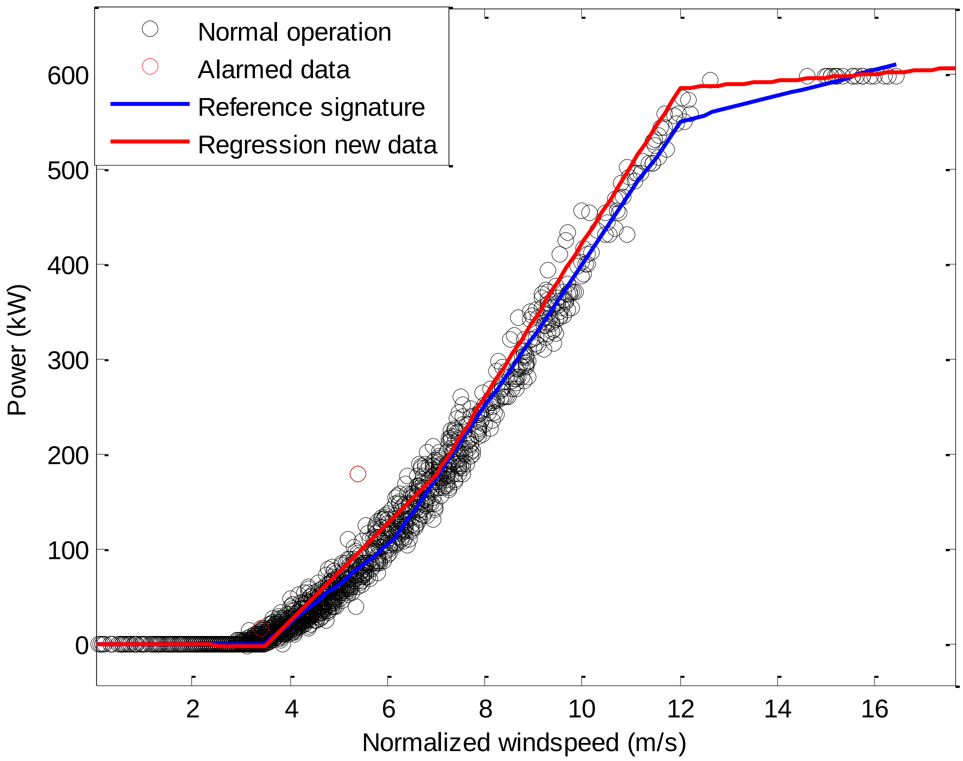

Figure 5.10 Piecewise linear regression (4 lines) with piecewise limits concept for alarm initiating with an alarm threshold of 5δ. The data displayed is measured in the SSW direction, May 2010 - October 2010.

5.4.3 Method of bins for alarms

When the method of bins is used for parameterization of the curve, the standard deviation of each bin can be used for detection of alarms. The upper and the lower limit for alarm are created by interpolation of the values calculated by the principles in Eq. (4.1) and (4.2). For alarm initiating although, it was found that the level for alarms could be reduced to $2.5\sigma_i$ when a bin width of 1 m/s is used. For bin width of 0.5 m/s, the standard deviation of each bin is small and the level of alarms generation has to be at least $4\sigma_i$. In some regions, σ_i is very small, especially below cut-in speed and above rated speed. Due to this, neither the upper nor the lower limit is visible for these regions in Figure 5.11. Alarms below cut-in speed are not of interest and can be neglected.

For the reference period, it is expected that a period long enough to cover all wind speeds, in all directions will be used. It is likely that a shorter period of time, for instance a couple of months during the summer, is desired to be investigated and compared to the reference signature. In these cases, it is not certain that the amount of data will be enough if a bin width of 0.5 m/s is used. When the regression line of the new data is created, at least one set of data measured in each bin is needed. As mentioned in section 5.3.4, in order to prevent lack of data in bins at high wind speeds, the number of bins above rated power can be reduced to one since the expected power in this region is constant. To ensure that a regression line can be obtained also for shorter periods of time, it is suggested that the regression line of new data added to the system shall be build up by bins of 1 m/s. When regression lines of data from May 2010 – October 2010 were constructed, it was necessary to use bins of 1 m/s and all data for wind speeds above 16 m/s was collected in one common bin. To ensure a signature over the whole wind speed range, an extra data set had to be added for each wind direction, $v_{n,i} = 20$ m/s and $P_i = 600$ kW. Even though the bin width for the new data is 1 m/s, it can still be compared with the reference signature constructed from bins of 0.5 m/s.

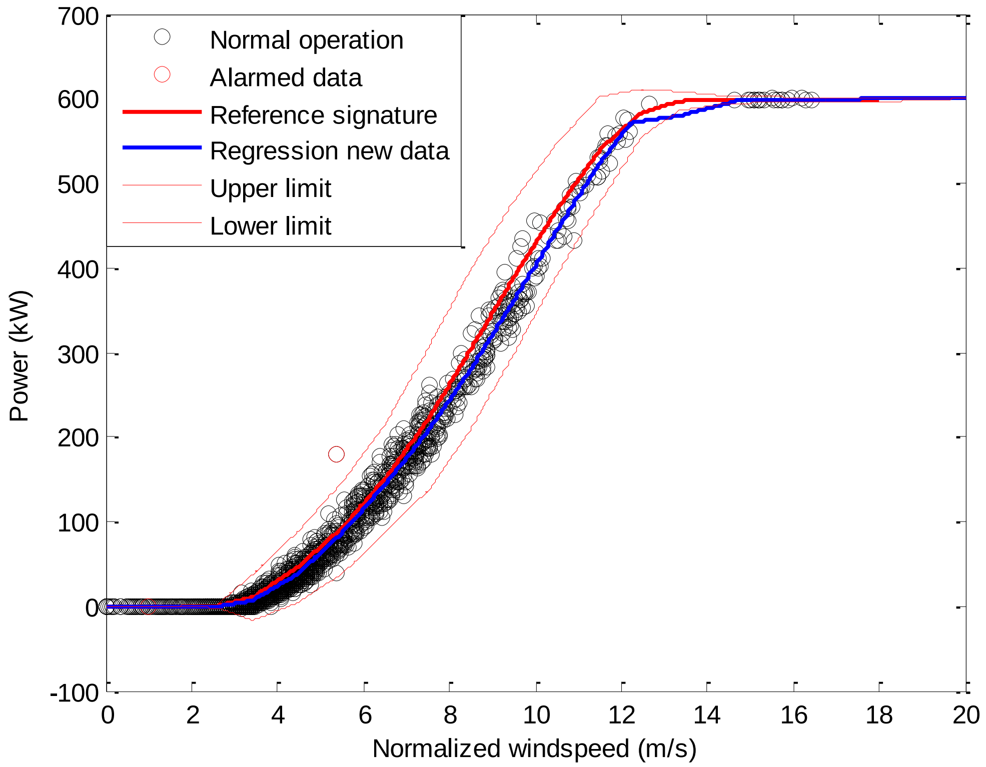

Figure 5.11 Method of bins for parameterization and method of bins for alarm generation with a threshold set to 2.5σ. The bin width is 1 m/s for both the reference signature and for the regression of new data up to 16 m/s. The width is thereafter enlarged for the new data. The data displayed is measured in the SSW direction, May 2010 - October 2010.

6 Comparison of the different models

The three different methods of parameterization presented in Section 5.3 were combined with the different concepts of alarm generation, resulting in ten different combinations of models. A survey of the results is found in Table 6.1. Diagrams of all the different combinations are found in Appendix 1.

Table 6.1 Survey of results for the different models developed.

Parameterization method	δ_{tot} (kW)	SSE_{tot} (kW^2)	Alarm concept	Alarm threshold	Number of alarms
Piecewise linear – 4 lines	14.13	2,760,500	Constant δ	4δ	25
			Piecewise δ	4δ	84
			Piecewise δ	5δ	28
Piecewise linear – 7 lines	12.64	2,245,300	Constant δ	3δ	129
			Constant δ	4δ	31
			Piecewise δ	4δ	172
			Piecewise δ	5δ	67
3rd order polynomial	15.59	2,518,300	Constant δ	3δ	76
			Constant δ	4δ	11
4th order polynomial	14.10	2,339,800	Constant δ	3δ	81
			Constant δ	4δ	21
Method of bins (1m/s)	12.62	2,233,200	Constant δ	3δ	52
			Constant δ	4δ	21
			Method of bins (1 m/s)	2.5σ	39
			Method of bins (1m/s)	3σ	21
Method of bins (0.5m/s)	12.40	2,164,500	Constant δ	3δ	118
			Constant δ	4δ	29
			Method of bins (1m/s)	4σ	19

6.1 Parameterization

To quantify the accuracy of fit for the methods of parameterization for the historical data, two measures were introduced, δ_{tot} and SSE_{tot}. From Table 6.1, it can be seen that the method of bins (0.5 m/s) resulted in the lowest δ_{tot} and SSE_{tot}. The bin method is very sensitive when the amount of data is limited. Since data for all bins is necessary, the total amount of data has to be extensive to ensure that data from normal operation is measured in every bin. With just a few values in each bin, the sensitivity to erroneous values increases as the divergent data will affect the average strongly. Due to this, usage of bins of 0.5 m/s is not suitable for construction of signatures for periods less than approximately one year. With data from the five month between May 2010 and October 2010, it was not possible to obtain a parameterization. Although it can still be used to construct the reference signature which shall contain more data, but then combined with another method for parameterization of new data for instance; bins of 1 m/s. Method of bins with bin width of 1 m/s resulted in the second best fit. With this method, it is possible to use the same method for both the reference signature and for new data of a shorter period. Another advantage of the method of bins is its simplicity; the method is easy to understand and does not contain any advanced mathematical calculations as it is based on average values and linear interpolation.

Except from the bin method, also piecewise linear regression with seven lines resulted in a high accuracy, just slightly poorer than by using bins of 1 m/s. This method is less sensitive to shortage of data as the intervals are increased. Because of the same reason, erroneous values will affect the parameterization less and it is easier to get a sufficiently good signature. The mathematics behind this method is more complex than the method of bins, although it is easily computed using Matlab.

Piecewise linear regression with four pieces would imply the least efforts for adaption of the current application of district heating to wind power since the mathematical algorithms already exist. To extend the number of knots, and thus the number of lines, to better fit the wind power data, should not cause any major difficulties.

6.2 Alarm generation concepts

The ten different models in Table 6.1 returns highly diverging number of alarms. The number can be modified by adjusting the alarm factor. Some models resulted in substantial amounts of alarms in some wind directions while at the same time, nearly no alarms for another direction. In these cases, it is hard to determine an appropriate alarm factor. This phenomenon occurred for the model using the method of bins for parameterization and the constant limit concept for alarms. For southerly (S) winds, no alarms occurred, Figure 6.1. In westerly (W) winds the number of alarms became 46, Figure 6.2. When looking at the figures, an alarm threshold of 2δ for southerly winds seems more appropriate, as the acceptable range of power is much wider for southerly compared to the range for westerly winds. It can thereby be concluded that the deviation measure δ for westerly winds during the reference period was considerably smaller than for southerly winds, 10.88 compared to 30.63 for southerly winds.

A large amount of alarms, 73, was obtained when 4 piecewise linear regression was combined with the piecewise limits concept for alarms. Many of these alarms are found around the knots as the acceptable limit in this point is not continuous. The alarm factor can be increased to reduce the number of alarms, but due to the rather poor parameterization, there will thereby be a risk that actual suspicious data will be considered as normal operation.

Even though the number of alarms in some wind directions is substantial, the regression line of the new data coincides fairly well with the reference signature in most cases. In all models, a minor degradation of performance could be traced in SSW direction; this can be seen in Figure 5.10.

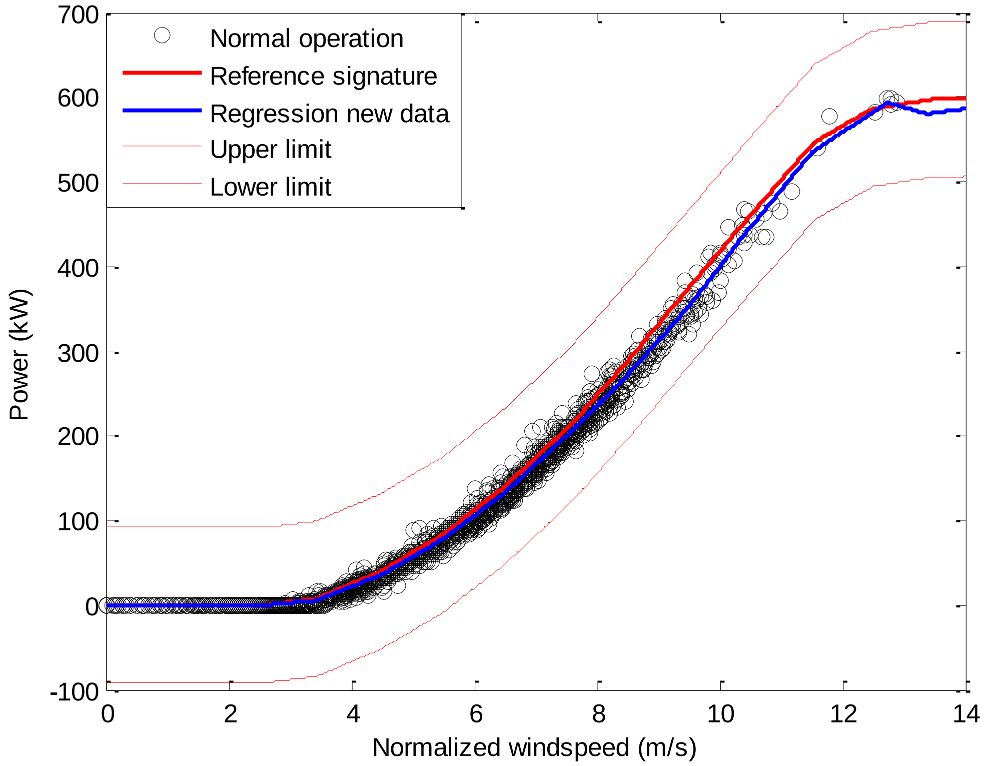

Figure 6.1 Result of the method of bins for parameterization and constant limits concept for alarm generation, using data measured in direction S. With an alarm factor of 3, no alarm is initiated.

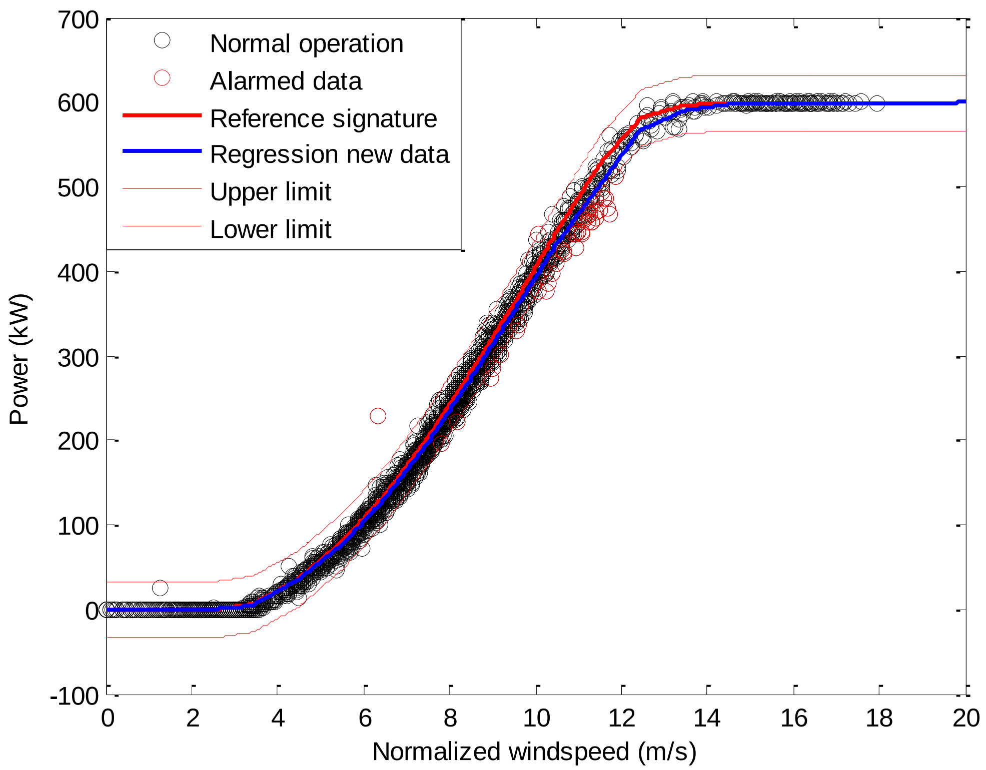

Figure 6.2 Method of bins for parameterization and constant limit concept for alarms of data measured in direction W. An alarm factor of 3 is used, resulting in 47 alarms

7 Analysis and discussion of alternative scenarios

Four different scenarios were outlined in order to evaluate the models' adaption possibilities to other turbines and other conditions. For other turbines, in other sites, data for temperature, atmospheric pressure or wind directions might be missing. It is therefore of interest to investigate if the models can be used under these circumstances. For a recently installed turbine, for which no historical data exists and a reference signature cannot be constructed, it is therefore of interest to find alternatives to reference signatures for new turbines. The investigated turbines at Risholmen, each with a capacity of 600 kW, are nowadays belonging to the smallest turbines in operation. New turbines being installed currently usually have a rated power above 1 MW. To make the developed system useful even in the future and for additional turbines, adaption to larger turbines were investigated. An evaluation of the filtering process and its effects on the models was performed.

7.1 Lack of temperature and pressure data

If data on temperature or pressure are missing at the site, normalized wind speed cannot be calculated. Instead, the measured wind speed which does not take into account the density of the air, has to be used. As can be expected, the deviation measure from the parameterized curve increased compared with when normalized data was used. This was the case for all methods. For instance, for un-normalized data, using piecewise linear regression (7 lines), $\delta_{tot} = 13.74$, compared with $\delta_{tot} = 12.64$ for normalized data. The level for initiation of alarm is thereby increased, in order to tolerate higher deviations.

When the reference signature, obtained using normalized wind speed, was compared to the signature for the same period, using non-normalized wind speed, the signatures appeared to be almost identical. In Figure 7.1, piecewise linear regression method (7 pieces), has been used to parameterize data, both normalized and non-normalized, measured from July 2008 to February 2009. During this period, the average temperature was 7°C; this is rather close to the temperature of 10°C which the wind speed has been normalized against. The effects of the normalization are therefore reduced as both temperatures above and below the average are represented during this period. When normalizing data measured in warmer periods, the signature is moved towards left and the opposite for colder periods. If data from e.g. a winter period, with an average temperature significantly lower than 10°C is analysed, one can expect the signature of the non-normalized wind speed to be above the one of normalized data. In Figure 7.2 data from October 2008 to February 2009 has been parameterized. The difference is small and barely visible in the figure. When analysing the data of the fit, it was found that the knots were slightly shifted upwards for the non-normalized data. From these results, it can be concluded that a sufficiently accurate parameterization can be obtained even without temperature and pressure data.

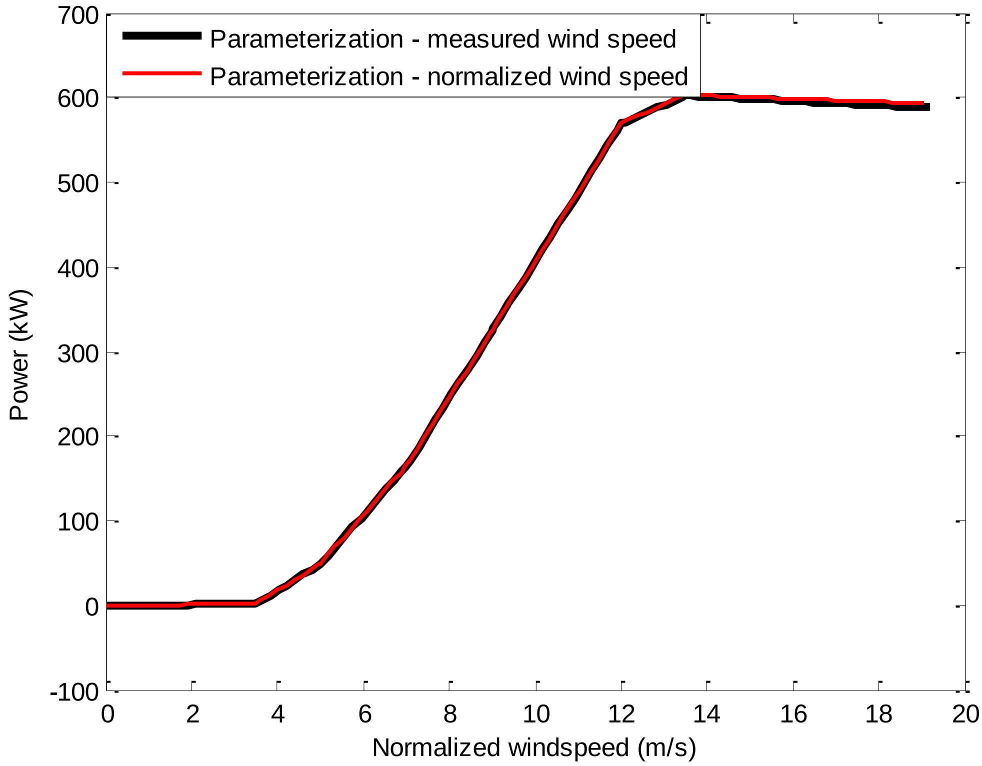

Figure 7.1 Piecewise linear regression (7 lines) of both data with normalized wind speed, and non-normalized wind speed.

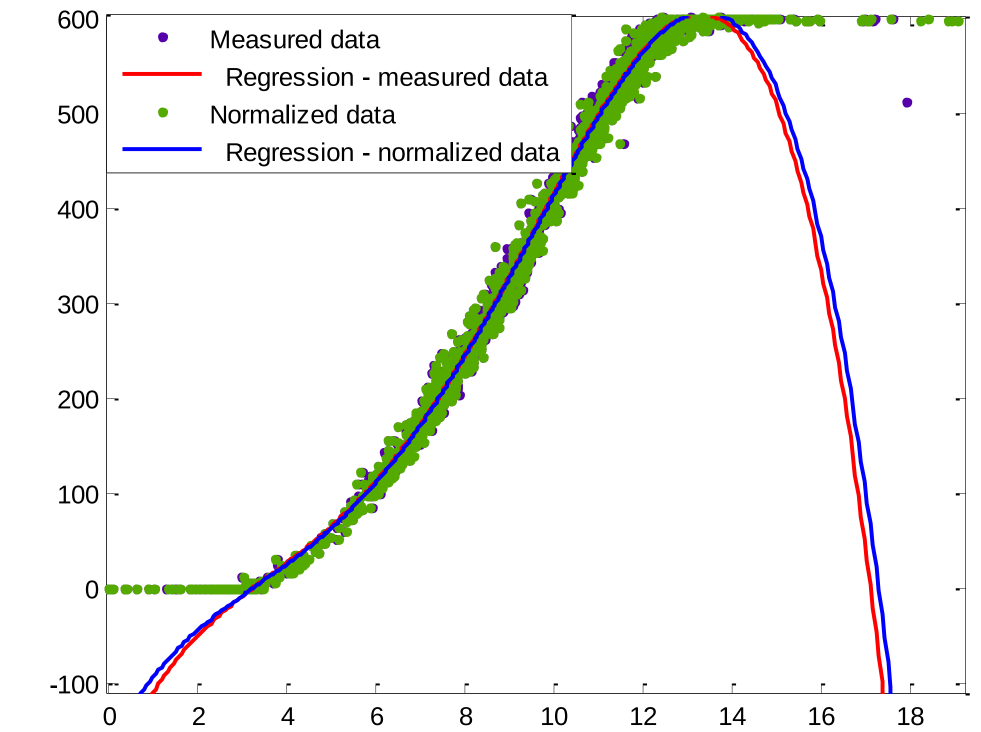

Figure 7.2 Polynomial (4th order) regression of both normalized data and non-normalized data measured during winter.

7.2 Lack of wind direction sensor

If turbines are located in areas where no wind mast is present, data for wind direction will not be available. In section 4.2, it was found that the power output is differentiated between directions and grouping of data according to direction was therefore included in all models. It would still be desirable that the developed model can be used even when wind direction data is unavailable.

The models were modified so all data, independent of the wind direction, was monitored in one figure. The filtering process remained the same. For all different methods, a parameterization could be achieved. The results are summarised in Table 7.1. Both a third order and fourth order polynomial regression of the data are demonstrated in Figure 7.3. When looking at Figure 7.3, the dense curve of data is wider than the direction dependent curves, it was therefore expected that the standard error δ for parameterizations independent of direction would be higher. When comparing Table 6.1 and Table 7.1, the opposite was found, δ is reduced for all methods. The explanation to this is probably that the wider range of power data is compensated by a large data density around the mean power. The sum squared error, SSE, increases in all cases, mainly because these parameterizations contain more data sets.

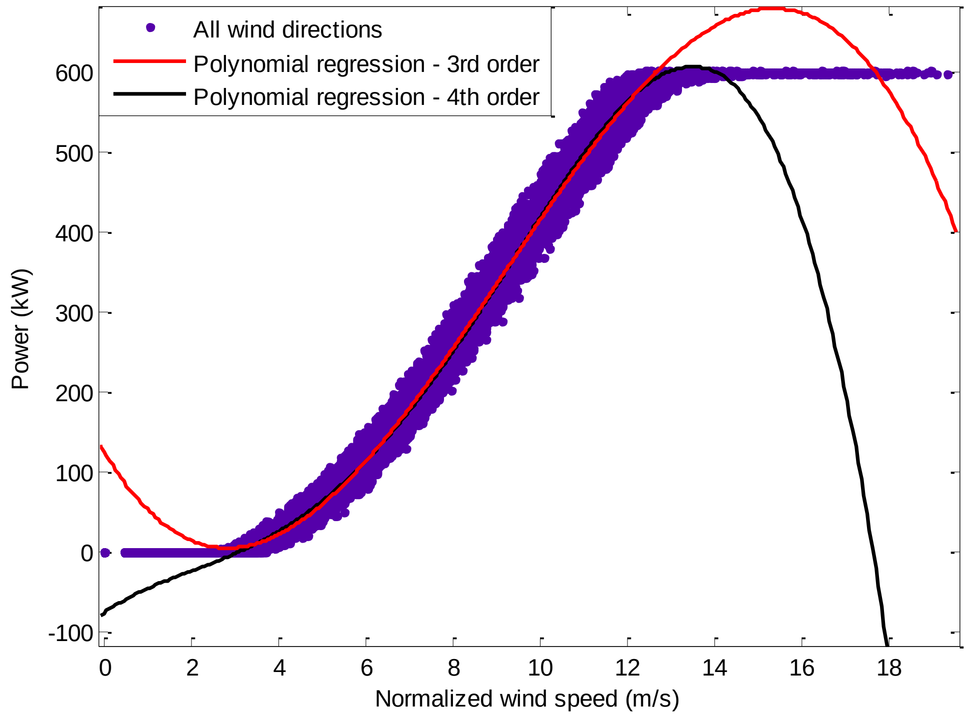

Figure 7.3 Parameterization of data, independently of wind direction.

Table 7.1 Results of parameterization for the models, in absence of wind direction data.

Parameterization method	δ	SSE
Piecewise linear – 4 line	13.76	4,683,000
Piecewise linear – 7 line	12.18	3,668,600
3rd order polynomial	15.13	3,846,000
4th order polynomial	14.61	3,587,000
Method of bins (1m/s)	12.18	3,667,300
Method of bins (0.5m/s)	12,03	3,577,200

7.3 Lack of historical data

During the first time of operation for a turbine, no historical data exists and therefore, no reference signature can be constructed. The new incoming data however needs to be compared to a reference. If other turbines of the same type already exist in the vicinity, data from these should be the best alternative to be used as a preliminary reference signature.

To see how the parameterized signatures of two turbines placed in the same area differ from each other, two signatures were created; one from data for Boel and one from data for Elin. Figure 7.4 shows that the performance of Elin was slightly better than for Boel at the same period of time. The difference is small and in this case, the signature from Boel could have been used as reference for Elin if no historical data for Elin would have been available.

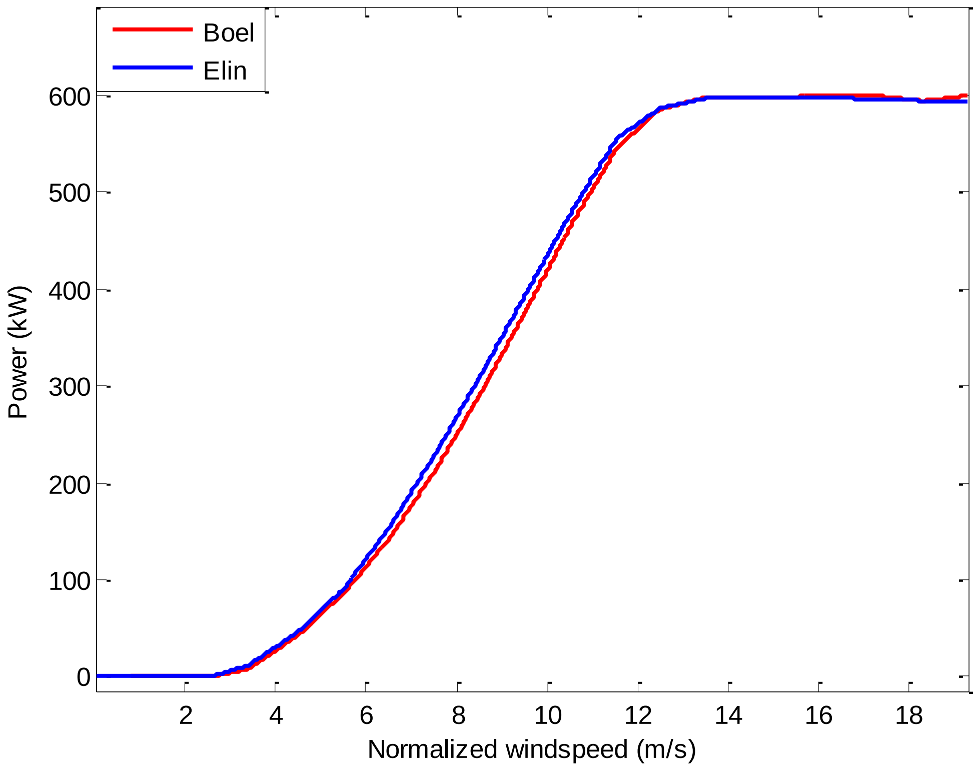

Figure 7.4 Comparison of the signature of Boel and the signature of Elin based on data measured during the same period The method of bins (1 m/s) was used to parameterize data measured in southerly (S) winds.

The turbine manufacturers usually supply a power curve, specifying a reference performance of the turbine. The curve supplied by Vestas for their V44 turbine is shown in Figure 2.1. It was assumed that such a curve could be used as an initial reference signature, when no other data is available. The power curve shall represent the power production of a turbine ideally located under certain standard condition, for instance air density ρ =1.225 kg/m^3. Ideal locations are rare and this requirement is in reality, most of the time, not fulfilled.

Figure 7.5 is showing the manufacturer's power curve and direction independent data from Boel which has been normalized to the standard air density ρ =1.225 kg/m^3. The curve from the measured data is higher than the manufacturer's specification. The performance of Boel seems to be remarkable better than the specification but this is not the case. The difference is mainly caused by the fact that the specification is based on wind speed measured according to IEC 61400-12 standard. Hence higher wind speeds than measured at the nacelle. If the measured power of the new turbine can be plotted against the wind speed measured at wind mast nearby, the manufacturer's power curve may be used as a preliminary reference signature. If only the wind speed measured at the nacelle is available, the manufacturer's power curve cannot be used to detect deviations of performance since the difference between the wind speed measured at the nacelle and the wind speed measured at a wind mast is too large.

The amount of data needed to get an accurate reference signature depends on the method used. Methods containing less separate intervals are least sensitive to the amount of data. When methods with smaller intervals are used, more data is necessary to ensure that each interval is containing a sufficient amount of correct data. The period of time needed to construct a reference signature of measured data is also strongly dependent on during which period of time the data is measured. During the winter, the wind variations are larger and the probability to measure wind speed in all ranges is therefore increased during these months (Boverket, 2009). Hence, a signature can be obtained with less data. For the development of alarm concepts, data from May to the beginning of October was used; this appeared to result in some difficulties for southerly winds where data of high wind speeds were insufficient. To even obtain a signature using the bin method, one data point at rated power had to be added. If data measured e.g. between November and March, would have been used instead, this problem would probably not have occurred. If a reference signature is desired recently after taken into operation, the piecewise linear regression with four lines can be used. This method has been proven to work with a small amount of data.

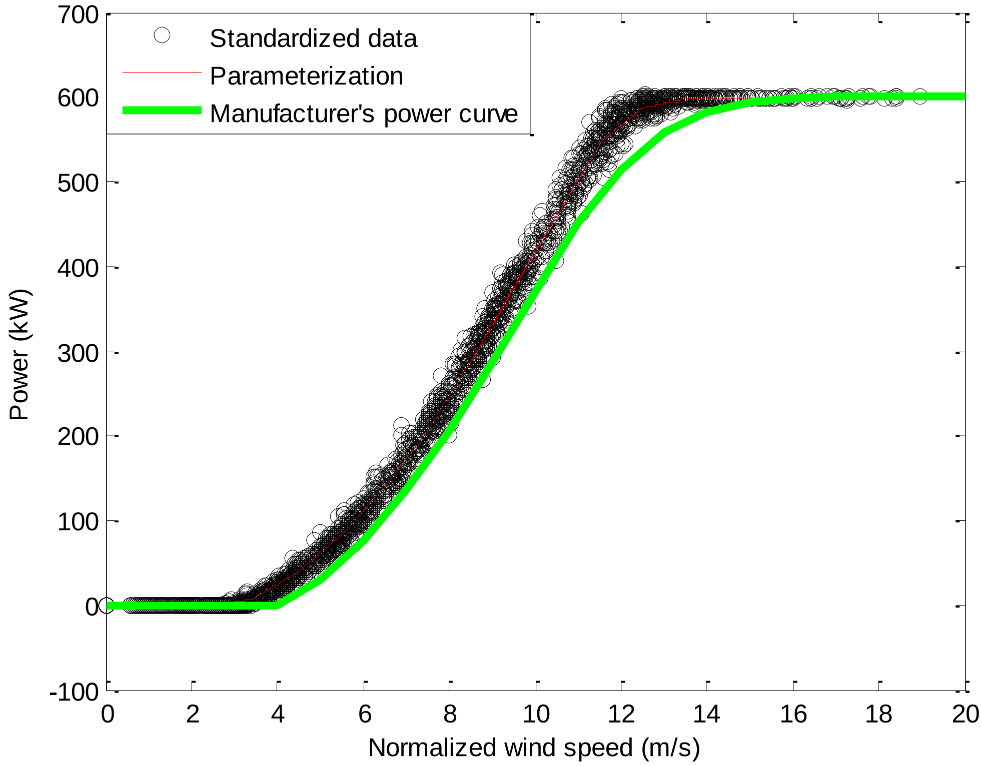

Figure 7.5 Parameterized data from Boel compared to the power curve supplied by the manufacturer.

7.4 Filtering sensitivity

The proposed filtering function is an easy way to reject occasional data which is caused by defect turbines, by meter reading errors or values measured during start-up or shutting down periods. For some models, the filtering is the process which requires most data because data in each bin is required. By improving the algorithms, it should be possible to solve the problem of missing data in occasional bins by interpolating of the neighbouring bins.

For larger quantities of data, the filtering process is working satisfactory. The appropriate level for rejection was determined to 5σ for both the first and second iteration. The second iteration is of less importance as the major part of rejection is done in first iteration. Several of the largest deviations are due to the shutting down and start up around cut-out speed, and these are rejected by the additional condition, not based on standard deviation. If the number of erroneous values within the data sample is increased, the second filtration's usefulness will rise and it should therefore be retained.

To analyse the consequences of the filtering, the deviation measure of the parameterization when the filtering process is left out was compared to the deviation measure when filtering is included in the model. The results are summarized in Table 7.2. For all models, both deviation measure δ_{tot} and the sum of the squared errors SSE_{tot} were increased.

Regarding the polynomial regression, data below 3 m/s and above 13 m/s is excluded. This reduces the effects of the filtering and the increase of deviation when left out is therefore small. Figure 7.6 is showing how the limits for alarms and how the parameterization are affected by exclusion of filter. The number of alarms initiated by the models decreased for all models, except from when the method of bins (1 m/s) was used for both parameterization and alarm detection. It shall be noticed, that the number of data points which shall be alarmed is higher since also the new data generated during start up or shutting down at high wind speeds shall now be detected as alarms. These points were previously rejected.

Table 7.2 Results of parameterization if no filtering process is conducted on the data before parameterizations.

Parameterization method	δ_{tot}	Increase δ_{tot}	SSE_{tot}
Piecewise linear – 4 lines	20.96	48.3 %	6,261,400
Piecewise linear – 7 lines	19,03	50.5 %	5,157,700
3rd order polynomial	16.63	6.7 %	2,873,800
4th order polynomial	16.04	13.8 %	3,251,200
Method of bins (1m/s)	17.83	41.2 %	4,497,900
Method of bins (0.5m/s)	17.72	42.9 %	4,403,300

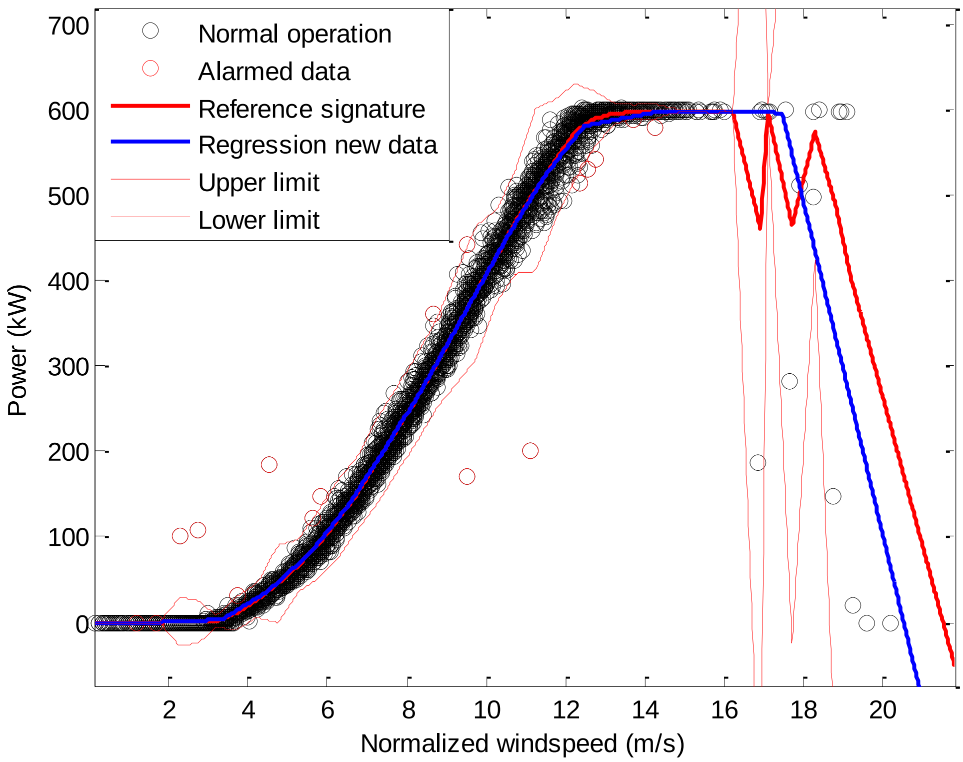

Figure 7.6 Method of bins applied to data which has not been filtered. Method of bins is also used for alarm initiation. Data is measured at Boel in WSW direction.

8 Conclusions

In order to find the best model for monitoring and detection of diverging data, three different methods for parameterization, namely piecewise linear regression, polynomial regression and linear interpolation based on the method of bins, were investigated. These were combined with three alternatives for alarm generation and resulted in ten different algorithms. The most appropriate model to be used depends on several different aspects and will be discussed in the following.

8.1 The tool and method

During the investigation of influencing parameters it was found that wind direction and air density affected the power output of the turbine. To eliminate the effect of varying air density, the wind speed used in the models was normalized according to the IEC standard (2005). Four of the most frequent wind directions were chosen to be monitored in the model namely; south (S), south-south-west (SSW), west-south-west (WSW) and west (W).

Comparative analyses of the methods investigated showed that the use of linear interpolation together with the method of bins, with a bin width of 0.5 m/s, resulted in the lowest standard error δ_{tot} and also the lowest sum of squared errors SSE_{tot}. This method was, however, also found to be the one that requires most data for parameterization, and its applicability is therefore reduced. An increased width of the bins does not raise the standard error considerably but it reduces the required amount of data and hence, its applicability increases.

The application of four pieces linear regression for parameterization was found to be insufficient. When the number of pieces was increased to seven, the accuracy of the fit was improved significantly and became comparable with that derived by the method of bins. An advantage of the piecewise method is that the presently used district heating algorithms easily may be modified and elaborated to include seven pieces and to be applied for wind power monitoring. Thereby operational synergies may be gained with respect to support. Polynomial regression is not recommended as its ability to create a satisfying parameterization was found to be limited.

Regarding concepts for alarms, the concepts based on differential alarm limits, i.e. the method of bins concept and piecewise standard error, was found to generate several incorrect alarms. In particular these occurred in regions of low δ. For the piecewise limit concept, many alarms were also initiated in the knots, where the limits are discontinuous. These methods are also more sensitive to erroneous data and lack of data, than methods with constant limits based on δ for all wind speeds.

If the number of alarms is increasing during a period, this may indicate an impending fault. The system can thereby serve as an early warning system. When the amount of recorded data is large enough to be parameterized, the wind power generation can be evaluated by comparing the regression line from one period to the reference signature. Long-term trends, and possible performance degradations, may thereby be detectable and, in combination with diagnostic measures, be utilised for condition based maintenance.

For decisions on which model to use, three important aspects have to be considered:

- The accuracy of the parameterization

- The minimum amount of data needed to enable a reliable evaluation
- The quality of available data.

With high quality data, in terms of few measurement errors, a model with shorter intervals can be applied. Poor quality will prevent high accuracy of the parameterization and in such cases, the piecewise linear regression of four pieces may be sufficient. If it is desired that short periods, less than one year, shall be used for evaluation, it is necessary to use either piecewise linear regression or polynomial regression. Taken into account the aspects mentioned above and also the implementation aspect, the recommendation to Göteborg Energi will be to apply the piecewise linear regression, with 7 pieces, combined with the constant limit concept for generation of alarms.

The developed filtering process requires data in each bin. Even if the model may work with a limited amount of data, the filter’s data requirement may introduce a limitation. If data is missing in some bins, the filtering process can be excluded to enable a parameterization using for instance piecewise linear regression. However, it is desirable that the reference period shall be long enough to cover all wind speeds and hence, this problem is eliminated.

Test analyses conducted without wind direction, temperature and air pressure data indicated that the models will generate reasonably accurate parameterization results and alarms also in case of lacking temperature and pressure data.

8.2 Future work

The data used for development of the models was mainly restricted to one turbine. To get a general analysis, the model should be tested with data from other turbines. The adaption possibilities to larger turbines should also be investigated as it is expected that larger turbines will dominate the wind power production in the future. With more data available, it will also be possible to investigate if data from one turbine can be utilised for modelling a reference signature for another turbine for which reference data is missing. The data investigated is generated during periods of normal operation, to better analyse the models, and especially the alarm concepts and the alarm thresholds, data representing deterioration should be tested.

For the investigated turbines, the number of measurement series is low compared with more recently installed turbines. With more data series available e.g. turbulence intensity, influences of other parameters can be investigated, and hence the model may be further refined.

In Kasper for district heating, methods for calculation of energy are implemented. These methods make it possible to quantify the total deviation in terms of energy (MWh) for a specific time period. A similar method for wind power could be developed as well. The consequences of performance degradation can thereby be quantified in terms of reduced electricity production and loss of revenues.

There are several functions not included in this thesis which have to be developed before Wind-Kasper can be taken in to operation. One essential part is to develop the data handling. Currently, pre-processing of the data is needed before it can be used in the models. Methods to automatically detect the most severe alarms should also be developed in order to facilitate the performance investigations.

9 References

Andrawus J., et al., (2009): The Selection of a Suitable Maintenance Strategy for Wind Turbines. *Wind engineering,* Vol 30, No 6, December 2009, pp 471-486.

Baumbach, G. (1991): Luftreinhaltung, Springer Verlag, Berlin, Heidelberg, New York

Barbounis, T.G., Theocharis, J.B. (2005): Locally recurrent neural networks for long-term wind speed and power prediction. *Neurocomputing*, Vol 69, No 4-6, January 2006, Pages 466-496

Barthelmie R. J. et al., (2007): Modelling and measurements of power losses and turbulence intensity in wind turbine wakes at Middelgrunden offshore wind farm, *Wind Energy,* Vol. 10, No. 6, November/December 2007, pp. 517–528

Besnard F., Bertling L. (2010): An Approach for Condition-Based Maintenance Optimization Applied to Wind Turbine Blades. *IEEE Transactions on Sustainable Energy*, Vol 1, No 2, July 2010, pp. 77 – 83.

Boverket. (2009): *Vindkraftshandboken – Planering och prövning av vindkraftverk på land och i kustnära vattenområden* (The wind power handbook – Planning and review of wind power onshore and in coastal sea districts. In Swedish), Boverket, Karlskrona, Sweden.

Burton T., et al. (2001): *Wind energy handbook*. John Wiley and sons, Ltd, West Sussex, England.

Catmull S. (2011): Self-Organising *Map Based Condition Monitoring of Wind Turbines*. RES Offshore, Kings Langley, United Kingdom

Cutler N. et al. (2011): Using nacelle-based wind speed observations to improve power curve modeling for wind power forecasting. *Wind Energy,* Published online May 2011.

Garlick W. G. et. al., (2009): A model-based approach to wind turbine condition monitoring using SCADA data. 20th International Conference on Systems Engineering, 8-10 September 2009, Coventry University, Coventry, UK

IEC (2005): *Wind Turbines – Part 12-1: Power Performance Measurements of Electricity Producing Wind Turbines*, IEC 2005-12-16, IEC 61400-12, International Electrotechnical Commission.

Kusiak A. et. al. (2009): On-line monitoring of power curves. *Renewable Energy,* Vol 34, No 6, June 2009, pp. 1487-1493

Kusiak, A., Li, W. (2010): Virtual Models for Prediction of Wind Turbine Parameters. *Transaction on energy conversion,* Vol. 25, No. 1, March 2010, pp. 245-252.

Lindqvist, P., (2010): Verktyg för utvärdering av energieffektiviseringar, baserat på effektsignaturanalyser (Tool for evaluation of energy efficiency, based on heat power signature analyses), Bachelor Thesis, Department of Computer Science and Engineering, University of Technology, Göteborg, Sweden, 2010.

Munoz, M. (2006): *Kvalitetsstudie för fjärrvärmemätare* (Quality study of district heating meters), Master Thesis. Department of Mathematical Science , Chalmers University of Technology, Göteborg, Sweden, 2006.

Nielsen, J. J., Sørensen, J. D., (2010): On risk-based operation and maintenance of offshore wind turbine components. *Reliability Engineering and System Safety,* Vol 96, no 1, January 2011, pp. 218-229.

Nilsson, J., Bertling, L., (2007): Maintenance Management of Wind Power Systems Using Condition Monitoring Systems—Life Cycle Cost Analysis for Two Case Studies. IEEE Transactions on Energy Conversion, Vol 22, no 1, March 2007, 223-229

Pedersen, T. et. al. (2002): *Wind turbine power performance verification in complex terrain and wind farms*. Risø National Laboratory, Risø-R- 1330(EN), Roskilde, Denmark, 2002

Petersen, L., E., Madsen, H., P., (2004): Wind farms. *Encyclopedia of Energy,* June 2004, pp. 449-463.

Peña A., (2009): *Sensing the wind profile.* Ph.D. Thesis. Department of Geography and Geology, Faculty of Science, ISBN 978-87-550-3709-0, Copenhagen, Denmark, 2009, pp. 34.

Sainz, E. et al. (2009): Robust filtering for the characterization of wind turbines: Improving its operation and maintenance. *Energy Conversion and Managaement*, Vol 50, pp. 2136- 2147.

Seguro, J. V., Lambert, T. W., (2000): Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis. *Journal of Wind Engineering and Industrial Aerodynamics,* Vol. 85, Issue 1, March 2000, pp 75-84

Wern, L., Bärring, L. (2009): Sveriges vindklimat 1901-2008, Analys av förändring i geostrofisk vind, SMHI, Meteorologi, No. 138/2009

Tian, Z., et. al. (2011): Condition based maintenance optimization for wind power generation systems under continuous monitoring. *Renewable Energy,* Vol 36. No 5, May 2011, pp. 1502-1509.

Figures