Stockholms universitet

Wind Power Forecasts

Gonzalo Aponte Navarro

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2017:9
Matematisk statistik
Juni 2017

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Bachelor Thesis **2017:9**
http://www.math.su.se

Wind Power Forecasts

Gonzalo Aponte Navarro*

June 2017

Abstract

In this thesis, we will apply time series analysis to make wind speed forecasts and then compute the corresponding wind energy that may be produced. The model used in this report is linear regression with ARMA-errors. The linear part of the model is intended to capture the seasonal effects present in the data and is composed by Fourier terms. A comparison between a model fitted to hourly average wind speeds and daily average wind speeds will be performed to check which produces the smallest forecasting errors.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: gonzalo.navarro@outlook.com. Supervisor: Mathias Lindholm, Filip Lindskog.

Contents

1 Introduction

Renewable energy has been a subject which has drawn much attention during the last decades. In efforts to halt global warming, several countries have invested in energy sources in order to reduce carbon emissions. Unlike other energy sources, wind energy production depends on the weather conditions which may vary throughout the day or throughout the seasons.
The subject of wind energy production estimates can be approached in different ways. Many such methods are based on Numerical Weather Prediction (NWP) models. These models typically generates forecasts up to 48h [4]. The NWP's use data from large areas and of many physical variables, and require supercomputers to run them. There are many statistical methods that rely on much less information that either model wind speeds or wind power [5], and can be run on ordinary PC's.

1.1 Aim

The purpose of this report is to predict wind energy production with the use of time series analysis. A model will be fitted to wind speed data and then wind speed forecasts will be converted into the corresponding wind power. We will apply methods that have been used in previous studies and analyze wind speed data in different time scales (hourly and daily average wind speeds).

1.2 Other studies

There are several studies that have used different variants of ARMA-models to forecast wind speeds. A common feature is that many of these models are designed to handle certain seasonal variations that often are present in wind speed data.
Such an example is the seasonal ARMA-GARCH model applied by J. W. Taylor et al. in [3] to model daily average wind speeds. J.S. Benth et al. [4] used a similar model to both daily average wind speeds and average wind speeds for every three hours. The seasonal parts of both these models were Fourier series of finite order.

Seasonal ARMA-models which are fitted to smaller time scales (hourly or three-hourly wind speeds) are typically adjusted to daily seasonal effects in wind data, known as diurnal variation. Such examples can be found in [1], [2] and [4]. In all these cases except for in [1], Fourier terms have been the method of removing diurnal and annual seasonality. Instead, in [1] the diurnal seasonality has been removed by subtracting the average wind speeds of the corresponding hour of the day. Brown et al. [1] has proposed Fourier terms have been proposed as a preferred method since it reflects the continuous nature of wind speed data and typically requires few parameters to be estimated.

David C. Hill et al. [2] have noted that diurnal seasonality differs for different seasons of the year, and have for this purpose fitted different Fourier terms for each season.

2 Theoretical framework

Most of the terminology comes from the works of Tsay (Analysis of financial time series) [6].

2.1 Stationarity

The models used in this study will be rely on the weak stationarity assumption. Let $\{z_t\}$ be a time series. The weak stationarity assumption says that for all t it holds that $E(z_t) = \mu$ is constant and that $Cov(z_t, z_{t-l})$ is only dependent on l.

2.2 ARMA models

ARMA is the abbreviation for autoregressive moving-average models. This is a class of models which are defined as follows. Let $\{z_t\}$ be a weakly stationary time series with mean μ, and let $\tilde{z}_t = z_t - \mu$. Also let $\{\varepsilon_t\}$ be a white-noise process. A white noise series is such that it is independent, identically distributed and has finite mean and variance. Let $\{\varepsilon_t\}$ have variance σ_ε^2 and mean 0. The following equation describes an ARMA-model of order p, q (which are positive integers).

$$\tilde{z}_t = \sum_{i=1}^{p} \phi_i \tilde{z}_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{2.2.1}$$

We will henceforth use the notation ARMA(p,q) for an ARMA-model of order p, q.

2.3 Fourier Series

To remove seasonal effects we will make use of Fourier series of finite order. Let a_k, b_k, f be real constants. Let $S(t)$ be a Fourier series of finite order as defined below.

$$S(t) = a_0 + \sum_{k=1}^{K} a_k cos\,(2\pi k f t) + b_k sin\,(2\pi k f t) \tag{2.3.1}$$

$S(t)$ is then a periodic function with fundamental frequency f. The function $S(t)$ is thus a sum of several periodic functions, each with frequencies kf, where k is a positive integer.

2.4 ACF and PACF

The autocorrelation function (ACF) is a measure of the correlation between the observations (known as serial correlation). Seasonal effects may be visually detected by regular peaks in the plot of the ACF. Serial correlation is a property of ARMA-processes and depending on the order of the ARMA-process its ACF may exhibit certain patterns. The ACF is defined as the following. Let $\{z_t\}$ be a weakly stationary time series. Then the ACF ρ_l of lag l is the following.

$$\rho_l = \frac{Cov(z_t, z_{t-l})}{\sqrt{Var(z_t)Var(z_{t-l})}}$$

Under the stationarity assumption, for all t and l it holds that $Var(z_t) = Var(z_{t-l})$ and therefore it holds that

$$\rho_l = \frac{Cov(z_t, z_{t-l})}{Var(z_t)}. \tag{2.4.1}$$

The ACF together with the partial autocorrelation function (PACF) are tools which together may be used to determine the orders p,q of an ARMA(p,q)-model. Unfortunately, it becomes increasingly difficult to identify the patterns for ARMA(p,q)-models when both p,q are greater than 1. For models where either p or q equals 0 the ACF is more easily identifiable.
The partial autocorrelation function (PACF) is more difficult to define. First, let ϕ_{kj} be the jth coefficient of the AR(k) process. We begin by noting that for an AR(k)-process the following holds true [7].

$$\rho_j = \phi_{k1}\rho_{j-1} + ... + \phi_{k(k-1)}\rho_{j-k+1} + \phi_{kk}\rho_{j-k} \quad j = 1, 2, ..., k \tag{2.4.2}$$

This relation can be expressed as a linear system of equations. First we define the following vectors and matrices.

$$\phi_k = \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \vdots \\ \phi_{kk} \end{bmatrix}, \quad \vec{\rho}_p = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_k \end{bmatrix}, \quad P_k = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & 1 \end{bmatrix}$$

$$P_k\phi_k = \rho_k \tag{2.4.3}$$

The coefficients of ϕ_k are the partial autocorrelations. They can be estimated by substituting the theoretical autocorrelations by their estimates $\hat{\rho}_l$.

$$\hat{\rho}_l = \frac{\sum_{t=l+1}^{T}(r - \bar{r})(r_{t-l} - \bar{r})}{\sum_{t=1}^{T}(r_t - \bar{r})^2}, \quad 0 \leq l < T - 1,$$

where $\bar{r} = \frac{1}{T}\sum_{t=1}^{T} r_t$.

2.5 Box-Ljung test

The Box-Ljung test is another way of checking whether serial correlation is present in the observations. This tool will be used to check if the residuals $\{\varepsilon_t\}$ of the ARMA-model (2.2.1) are serially correlated. According to the assumptions, the residuals should have no serial correlation. If serial correlation is present in the computed residuals $\{\hat{\varepsilon}_t\}$, this is indicative that the model is inappropriate. Let the autocorrelation estimates be denoted by

$\hat{\rho}_l$ and let T be the sample size. The Box-Ljung test statistic $Q(m)$ is the following.

$$Q(m) = T(T+2)\sum_{l=1}^{m} \frac{\hat{\rho}_l{}^2}{T-l}$$

Let's assume that $\{\hat{\varepsilon}_t\}$ are i.i.d. - meaning that no serial correlation is present. Under certain conditions, which we will assume to be fulfilled, the statistic $Q(m)$ is asymptotically $\chi^2(m)$-distributed [6]. The Box-Ljung test statistic will be used to test the following hypothesis $H_0 : \rho_1 = ... = \rho_m = 0$ against the following alternative hypothesis.

$$H_a : \rho_i \neq 0, \quad i \in \{1, ..., m\}$$

The choice of m might affect the performance of the test. For seasonal data, it is recommended by Rob J. Hyndman [13] to select $m = min(2h, T/5)$ where h is the period of seasonality and T is as above.

2.6 Akaike's Information Criterion

The choice of the order of an ARMA-model is not always obvious, and in practice one often has to select a model among a list of candidates according to some principle. We will try to apply the parsimony principle and pick a model which fits well to the data and has as few parameters as needed. Akaike's information criterion (AIC) is a measure of how well a model fits to data. It is defined as follows.

$$AIC = -2log(L(\hat{\theta})) + 2q$$

q is the number of parameters in θ, and $\hat{\theta}$ denotes its estimate which maximises the likelihood L. AIC penalizes a high number of parameters and for this reason it suits our purposes. A low AIC indicates a good fit, and a high AIC indicates a poorer fit.

3 Data

The data used in this study has been obtained from the Western Wind Data Set available on NREL's (National Renewal Energy Laboratory) website.

The Western Wind Data Set is a re-creation of the weather for western U.S. the years 2004-2006 and is intended for use in studies of power production estimation. This is a very large data set which contains smaller data sets for different locations. For each location there is, besides wind speed, also a re-creation of the power output of 10 wind turbines (of type Vestas, V90 3 MW). The power output was based on a model that mimics the randomness found in the power output for actual wind turbines.

4 Analysis

We will fit a model similiar to that of Hill [2] and Benth [4]. First we will fit a model to daily average wind speeds (DAWS) and then for hourly average wind speeds (HAWS). The next step will be to check the performance of each model and compare them on the same time scale (daily average wind speeds). We will then find a function of wind speed which gives the power output of ten wind turbines. Finally we calculate the power output and compare them to the actual power output.

4.1 Data distribution

In previous studies, data has been transformed to approximate a normal distribution. In a study by Hill [2], the transformation was needed because of the model identification procedure that was used. Hill proposed that detrending the data, by removal of the annual and diurnal effects, is enough to transform data approximately normal. Other transformation methods that have been used are Box-Cox transformations [4] and power transformations [1].

In this study we will apply the power transformation method used by Brown [1]. We base this decision on two main reasons. By choosing an appropriate power transformation of the observations, forecasted wind speeds are guaranteed to be positive. This is important because the ARMA-model does not assume that the time series is positive. We also note that this is necessary if one decides to use the ARMA-model for simulation, simply because there are no negative wind speeds. Secondly, the parameters of the ARMA-model will be estimated by the Maximum Likelihood method (ML) which assumes that the innovations are i.i.d. normally distributed.

First the data is transformed using several values of m. Let X_t denote the wind speed at time t. The power transformation has the following notation.

$$X_t^m = \begin{cases} X_t^m, & \text{if } m \neq 0 \\ log(X_t), & \text{if } m = 0 \end{cases}$$

The next step is to calculate d_m - a measure of the symmetry of the distribution.

$$d_m = \frac{mean - median}{scale}$$

As the scale parameter, Brown et al. have used the interquartile range (IQ). A value close to zero suggests an almost symmetric distribution, which is what we desire. In figure 1 we have made a qq-plots of the DAWS and HAWS compared to the normal distribution. The figure is a plot of the empirical quantiles versus the theoretical quantiles of the normal distribution. A straight line suggests that data is close to being normally distributed.

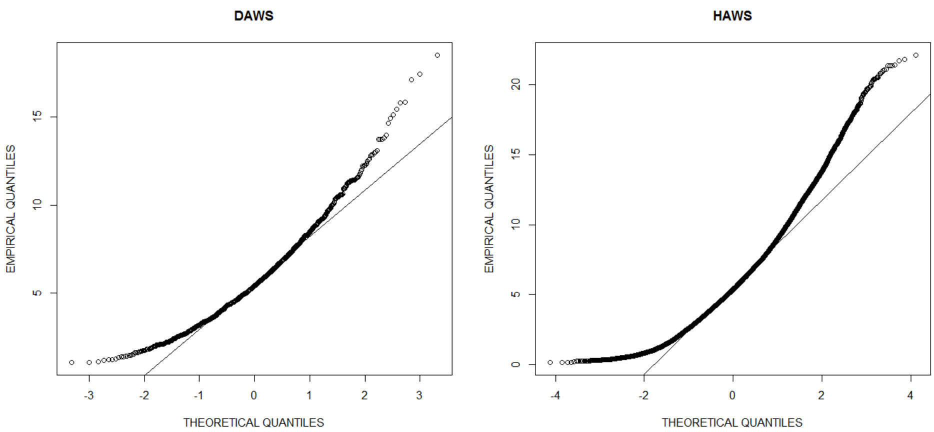

Figure 1: QQ-plot of DAWS to the left and HAWS to the right, compared to the normal distribution.

We see that both DAWS and HAWS are positively skewed. By applying different power transformations we obtain the values of d_m shown in figure 2. We see that the best values of m are 3/8 for DAWS and 1/2 for HAWS.

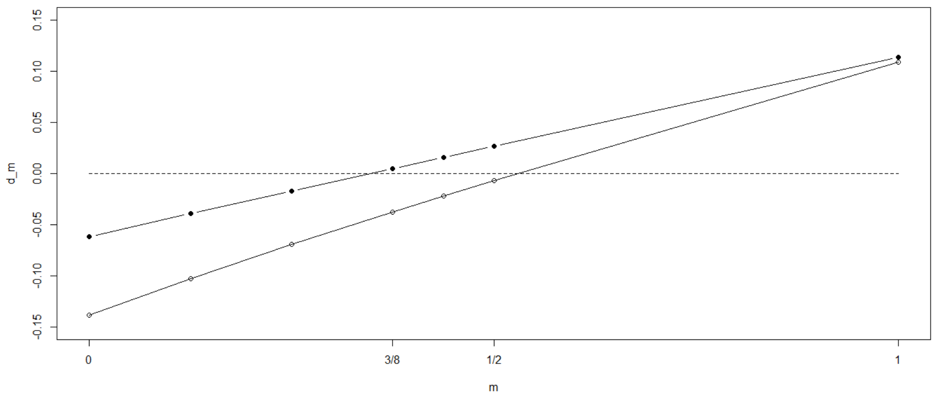

Figure 2: Values of the symmetry statistic d_m for the power-transformations of DAWS (in black) and HAWS (in white).

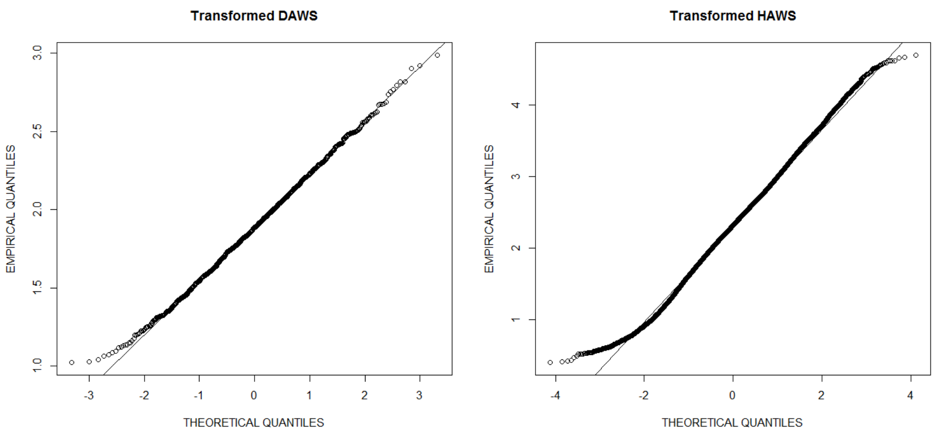

Figure 3: QQ-plot of transformed DAWS to the left and transformed HAWS to the right, compared to the normal distribution.

4.2 Seasonal effects

Wind flows from high pressure to low pressure regions, and the pressure of the air is affected by temperature changes. Regular changes in temperature can give rise to regular changes in wind speed - sea and land breezes are such examples [8]. For these reasons, several seasonal effects and cyclical patterns may be present in the data.
In figure 4 we see the autocorrelation-plots for the transformed DAWS and HAWS. There is a very clear annual seasonality in both cases. This is observed by the regular peaks by the lags corresponding to the yearly length, 365 for the DAWS and 8760 for the HAWS.

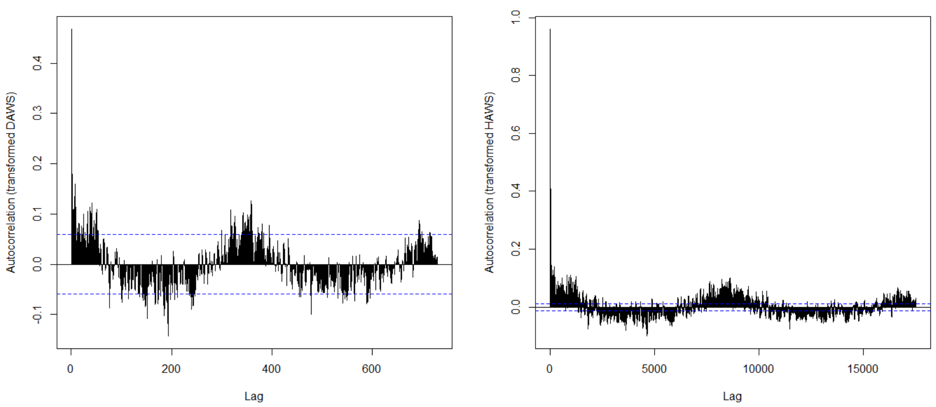

Figure 4: Autocorrelation plots for transformed DAWS to the left and transformed HAWS to the right.

In figure 5 we see the ACF-plot for HAWS up to a smaller number of lags. We see that the ACF has high values close to the multiples of 24, which may be indicative of a diurnal seasonal effect.

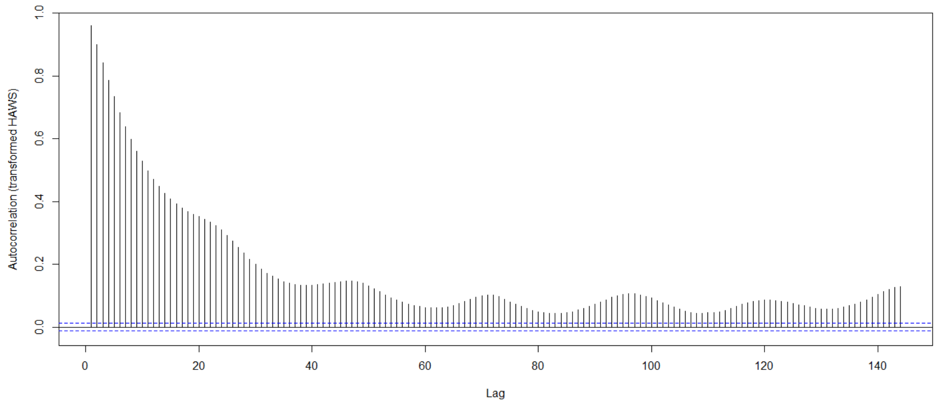

Figure 5: Autocorrelation plot for transformed HAWS.

4.3 Model description

Our model consists of a seasonal part S_t and an ARMA error-term e_t. Let X_t be the wind speed at time t. The model is given below.

$$X_t = S_t + e_t \tag{4.3.1a}$$

$$e_t = \sum_{i=1}^{p} \phi_i e_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{4.3.1b}$$

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{4.3.1c}$$

The parameters of the seasonal part S_t will be fitted by Ordinary Least Squares estimates. The parameters of the ARMA-model of the residuals e_t will be estimated separately by the Maximum Likelihood method.

5 Results

The ARMA-part of each model will be selected by the lowest AIC and seasonal part will be chosen such the residuals ε_t are devoid of seasonal effects.

5.1 DAWS

In order to eliminate the seasonal effects we will perform a multiple linear regression to our data set. Let $x_{k,t} = cos(2\pi f t)$ and $y_{k,t} = sin(2\pi f t)$, where f denotes the frequency corresponding to the annual seasonality. The model below is intended to eliminate the annual seasonal effect.

$$X_t = a_0 + \sum_{k=1}^{K} a_k x_{k,t} + b_k y_{k,t} + e_t$$

The equation above is a linear system of equations, and is therefore such that the Ordinary Least Squares method can be applied for parameter estimation. One condition required to compute the Ordinary Least Squares estimates is that the variables $x_{k,t}$, $y_{k,t}$ are linearly independent, or almost uncorrelated. According to [9] it holds that the variables $x_{k,t}$ and $y_{k,t}$ are uncorrelated, and so we should not expect any problems to compute our estimates.
In other studies leap days have been removed in order to keep the annual frequency corresponding to 365 days as the annual period [2], [4]. In this study we will instead let the annual frequency f correspond to the average year length 365.25, as recommended by Rob J Hyndman [14]. Specifically, we will let $f = 1/365.25$. The estimates are found in table 2.
In figure 6 we find the ACF and PACF of the residuals e_t. We note that the annual effect is removed, and conclude that $K = 1$ is enough for this purpose. There are two high peaks for the ACF at lags 1 and 2, as well as a large peak at lag 1 for the PACF.

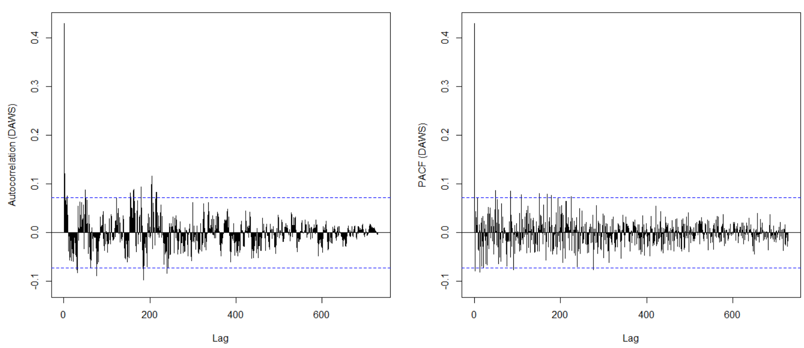

Figure 6: ACF and PACF for the residuals of the deseasonalized DAWS.

We will now fit an ARMA-model to the residuals e_t. We will bound the values p and q from above by 5 and choose the ARMA(p,q)-model with the lowest AIC. Let X_t be the transformed DAWS. Our model is then the following.

$$X_t = a_0 + a_1 cos\,(2\pi f t) + b_1 sin\,(2\pi f t) + e_t \tag{5.1.1a}$$

$$e_t = \sum_{i=1}^{p} \phi_i e_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{5.1.1b}$$

$$\varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{5.1.1c}$$

We assume the ε_t to be i.i.d.. The ARMA(2,0)-model has the lowest AIC, and thus it will be selected. The parameter estimates of the seasonal part are shown in table 2 and those from the ARMA-part are found in table 3. We examine if the residuals ε_t of the ARMA-model are i.i.d. and normally distributed. In figure 7 we see the ACF and QQ-plot of the residuals. The empirical quantiles do not deviate much from the line and so the normality assumption is valid. In the ACF the residuals do not appear to be serially correlated.

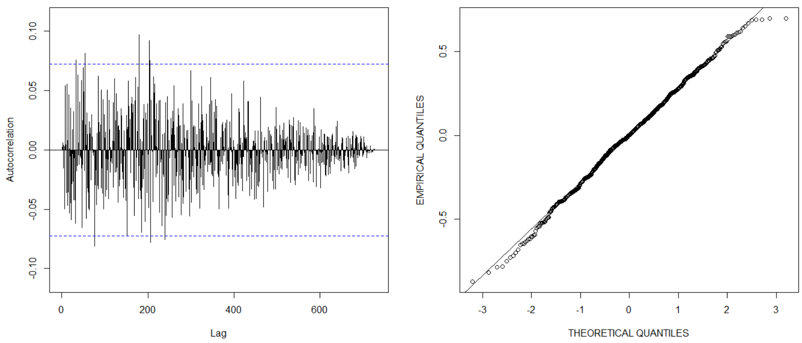

Figure 7: Dataset: DAWS. ACF (left) and QQ-plot (right) for the residuals of the ARMA-model

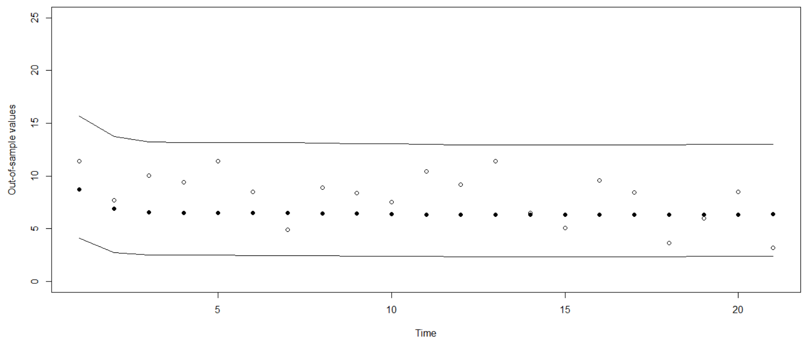

Figure 8: DAWS-forecasts. The lines are 95% prediction intervals, the white circles are out-of-sample values and the black points are predictions.

5.2 HAWS

In the ACF of the transformed HAWS (figure 5) we see a regular pattern of peaks every 24 lags. We conclude that there is diurnal seasonality. For this reason we will fit the following model to eliminate both the annual and diurnal variation.

$$X_t = a_0 + S(t) + e_t$$

$$S(t) = S_1(t) + S_2(t)$$

$$S_i(t) = \sum_{k=1}^{K_i} a_{k,i} cos(2\pi f_i t) + b_{k,i} sin(2\pi f_i t),$$

where $i = 1, 2$ and $f_1 = 365.25 \cdot 24$, $f_2 = 24$ are the frequencies corresponding to the annual and diurnal seasonal effects. We begin by letting $K_1 = 1$ and $K_2 = 1$. Again, we use the Ordinary Least Squares to estimate the parameters of the seasonal part. In figure 9 we see that the annual effect has been eliminated, while the diurnal effect appears to remain. We proceed by fitting an ARMA-model to the residuals e_t. We compare the AIC of the ARMA(p,q)-models of orders $p \leq 13$, $q \leq 5$ and find the model with the lowest AIC to be ARMA(1,2).

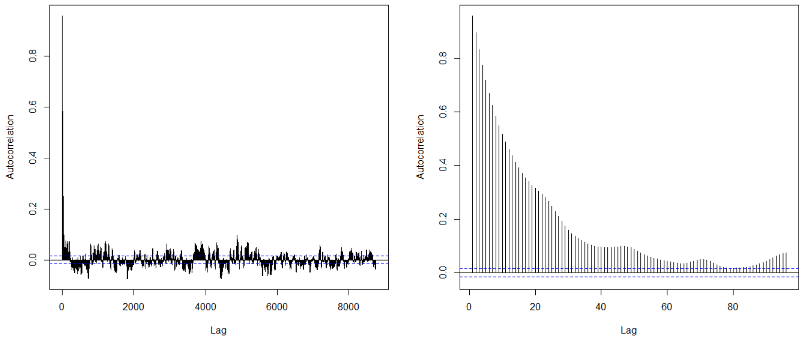

Figure 9: HAWS. The ACF of the residuals after annual and diurnal terms have been fitted (order 1,1)

In the ACF of the residuals for the ARMA-model (figure 10) we observe that diurnal seasonality has not been eliminated. For this reason we will try other orders K_1, K_2 of the seasonal function $S(t)$ and fit the corresponding ARMA-models with the lowest AIC.

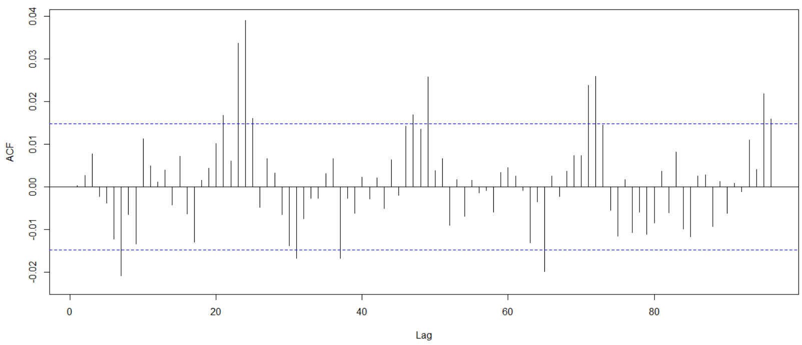

Figure 10: HAWS. The ACF of the ARMA(1,2)-residuals.

Rather than inspecting whether the residuals are serially correlated by examining the ACF we have opted for performing a Box-Ljung test on the residuals. The results are presented in table 4, where we have used the notation FOU(i, j) for a model with orders K_i, K_j for the seasonal part $S(t)$. The Box-Ljung test reject the absence of serial correlation, and by inspecting the ACF's we see that there are peaks at multiples of 24, so the diurnal seasonality has not been erased for these models.

5.3 Forecasting the wind power

The wind turbine used in this study is of the model Vestas V90. It only operates at certain wind speeds. The smallest speed at which the turbine generates power is called the cut-in wind speed (V_I) and the largest velocity is called the cut-off wind speed (V_O). The wind turbine is designed to produce constant power, known as the rated power (P_R), between the so-called rated wind speed (V_R) and the cut-off wind speed (V_O). In theory [8], the following expression holds for this kind of wind turbine:

$$P(V) = \begin{cases} P_R \frac{V^n - V_I^n}{V_R - V_I^n}, & \text{if } V_I \leq V \leq V_R \\ P_R & \text{if } V_R < V \leq V_O \\ 0 & \text{otherwise,} \end{cases} \tag{5.3.1}$$

where $P(V)$ is the power output at the velocity V and $n$is known as the power proportionality. The values of the different wind speeds V_I, V_R, V_O corresponding to the Vestas turbine are found in table 1. Ideally n equals 3, but in practice this expression does not apply since the speed of the wind passing through the turbine cannot be measured precisely [8]. This occurs because the wind measurements are usually made at a distance from the actual wind turbines. Other sources of deviation from the theoretical expression are explained by variations in air density [8].

There are different ways of forecasting the wind power produced by a wind farm. The method in this study is to use a deterministic model for the relation between wind power and wind speed, as in the study by Brown [1]. Since we have wind power data from 10 wind turbines, we could compare our wind power forecasts to actual observations. It does however become problematic for various reasons. It is unrealistic to assume that all wind turbines produce the same amount of power for a given wind speed. In other words, multiplying the power function (2.9.1) by 10 may not suffice. The problem lies in finding an adequate deterministic function which can estimate the total power produced by ten wind turbines corresponding to the wind speeds.

S. Mathew [8] has proposed a method to determine the power proportionality of one wind turbine. By calculating the correlation of the observed wind power and the wind power $P(V)$ of with power proportionalities n Mathew chooses n that gives the highest correlation. For practical reasons, we will perform a similar selection of n, but instead we will choose n such that the correlation between $10P(V)$ and the observed power output is as high as possible. Since the power curve gives the power produced for instantaneous wind velocity, we will study the correlation of the ten-minute observations of the power output and $10P(V)$.

In figure 11 we see the correlations plotted against the power proportionality. $n = 2$ gives the highest correlation, and so we select the corresponding power curve for our forecasts, which is plotted in figure 12. The wind power forecasts of the DAWS-model are shown in figure 13.

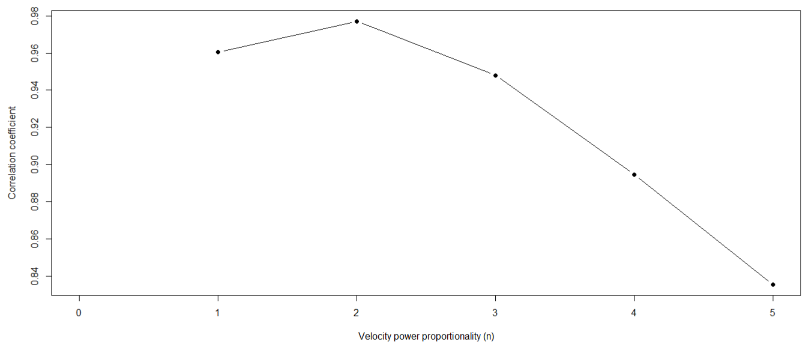

Figure 11: Correlations plotted against the power proportionality n.

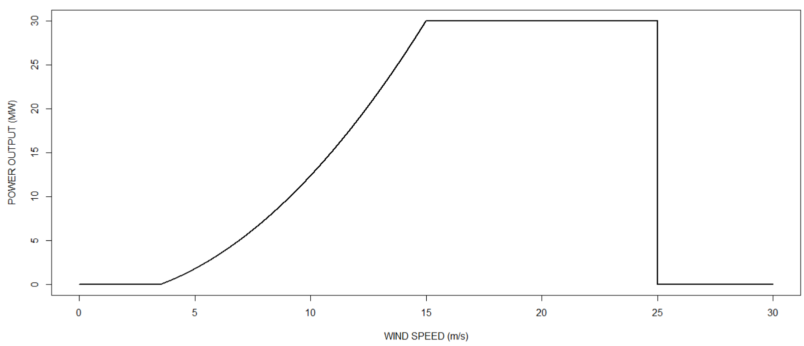

Figure 12: Power curve, $n = 2$.

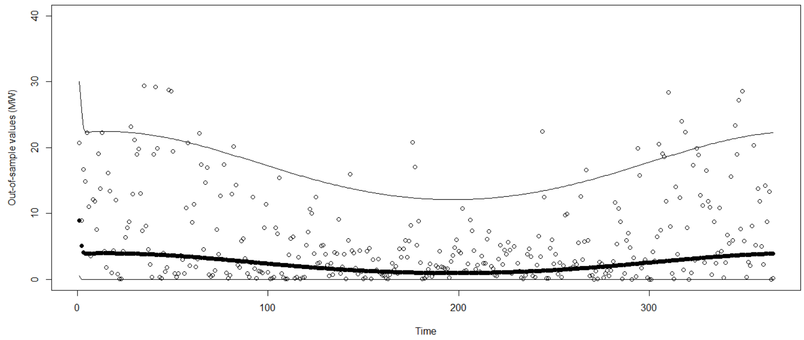

Figure 13: Daily Average Power-forecasts, forecasted by the DAWS-model.

5.4 Discussion

We will now examine how well the DAWS- and HAWS-models are at forecasting. We begin by defining the Root Mean Square Prediction Error (RMSE) as follows. Let $\hat{X}_t$ be the forecast of the wind speed X_t.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(\hat{X}_t - X_t)^2}$$

This quantity measures how large the prediction errors are. The HAWS-model we will choose for this comparison is the FOU(1,1)+ARMA(1,2)-model. This model is chosen because has the fewest numbers of parameters. In order to compare the DAWS-model with the HAWS-model we will transform the hourly forecasts back into the units of the original measurements. The hourly forecasts are then averaged corresponding to each day. The results are presented in table 5. We see that the RMSE-values for the HAWS-model are higher than those of the DAWS-model. Similar results were found by Benth [4], who suggests that predictions based on HAWS are more inaccurate than that for DAWS due to hourly wind data being much more noisy.

We were unable to eliminate the diurnal seasonal effects from the HAWS data. In the study by Hill [2] the diurnal component was found to vary among the four seasons of the year, with Spring and Summer having a stronger diurnal component than Winter and Autumn. One could fit a

model similar to that used by Hill by fitting a different diurnal model for each season of the year. The problem with such a model, as stated by Hill, is that there can be abrupt changes where the seasons meet.
Another problem with the method of our study is how the parameters have been estimated. We have opted for estimating the seasonal parameters by the Ordinary Least Squares method, and the ARMA-parameters by the Maximum Likelihood method separately. According to Pankratz [10] this could lead to several unwanted consequences, such as the estimators not having minimum variances or that the forecasts become inaccurate. This problem could be eliminated by performing a joint estimate of the parameters, but that requires knowledge of more advanced methods.

Acknowledgements

I would like to thank my supervisors Filip Lindskog and Mathias Millberg Lindholm. Thank you for your support.

6 Appendix

6.1 A1 Tables

Operating data	
Rated power	3.0 MW
Cut-in wind speed	3.5 m/s
Rated wind speed	15 m/s
Cut-out wind speed	25 m/s

Table 1: Vestas V90 specifications

Parameter	
a_0	1.87135
a_1	0.03424
b_1	0.12567
R^2	7.791%

Table 2: DAWS dataset. OLS-estimates for the seasonal function.

σ_ε^2	ϕ_1	ϕ_2
0.0815	0.4660	-0.0789

Table 3: DAWS dataset. Estimates of the ARMA-parameters

Model	P-value (box-test)
FOU(1,1)+ ARMA(2,2)	0.0004
FOU(1,2)+ ARMA(1,3)	0.0020
FOU(1,3)+ ARMA(1,3)	0.0017
FOU(2,1)+ ARMA(2,2)	0.0005
FOU(2,2)+ ARMA(2,2)	0.0030
FOU(2,3)+ ARMA(1,3)	0.0020

Table 4: HAWS dataset.

	1-day ahead	**3-day ahead**	**5-day ahead**	**7-day ahead**
DAWS-model	2.703379	2.634034	3.346102	2.994439
HAWS-model	3.488577	3.031512	3.579653	3.184506

Table 5: RMSE of daily wind forecasts.

References

[1] Barbara G. Brown, Richard W. Katz, Alan H. Murphy *Time Series Models to Simulate and Forecast Wind Speed and Wind Power* Journal of climate and applied meteorology, 1984.

[2] David C. Hill, David McMillan, Keith R. W. Bell and David Infield *Application of Auto-Regressive Models to U.K. Wind Speed Data for Power System Impact Studies* IEEE Transactions on sustainable energy, vol. 3, no. 1, january 2012.

[3] J. W. Taylor, P.E. McSharry *Wind Power Density Forecasting Using Ensemble Predictions and Time Series Models* IEEE transactions on energy conversion. 24. 775-782. 2009.

[4] J. S. Benth and Fred Espen Benth *Analysis and modelling of wind speed in New York* Journal of applied statistics, vol. 37, no.6, June 2010, 893-909.

[5] Lars Landberg, Gregor Giebel, Henrik Aalborg Nielsen, Torben Nielsen and Henrik Madsen *Short-term prediction - an overview* Wind Energy, 2003.

[6] Ruey S. Tsay *Analysis of financial time series* Third Edition, 2010.

[7] George E. P. Box, Gwilym M. Jenkins and Gregory C. Reinsel *Time series analysis - forecasting and control* Fourth edition, 2008.

[8] Sathyajith Mathew *Wind Energy - Fundamentals, Resource Analysis and Economics* 2006.

[9] Wilks, D.S. *Statistical methods in atmospheric sciences* New York, Academic Press (1995).

[10] Alan Pankratz *Forecasting with Dynamic Regression Models*

[11] Vestas V90 3.0 MW Available:
http://nozebra.ipapercms.dk/Vestas/Communication/Productbrochure/V9030MW/V9030MWUK/

[12] Rob J. Hyndman *Thoughts on the Ljung-Box test* [Online]. Available: http://robjhyndman.com/hyndsight/ljung-box-test/

[13] Rob J. Hyndman *Forecasting with long seasonal periods* 2017 [Online]. Available: http://robjhyndman.com/hyndsight/longseasonality/

[14] Rob J. Hyndman *Seasonal periods* 2017 [Online]. Available: http://robjhyndman.com/hyndsight/seasonal-periods/