PERFORMANCE OF LONG-TERM CORRECTION AT WIND POWER DEVELOPMENT

AN ANALYSIS OF RESULTS FROM DIFFERENT REANALYSIS DATA SETS

Master of Science Thesis in the Master Degree Program Industrial Ecology for a Sustainable Society

In collaboration with Triventus Consulting AB

JAKOB SONDEREGGER

Department of Energy and Environment

Division of Electrical Power Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden, 2011

PERFORMANCE OF LONG-TERM CORRECTION AT WIND POWER DEVELOPMENT

AN ANALYSIS OF RESULTS FROM DIFFERENT REANALYSIS DATA SETS

JAKOB SONDEREGGER

Department of Energy and Environment

Division of Electric Power Engineering

CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden, 2011

Examiner: Ola Carlson, Electrical Power Engineering

PERFORMANCE OF LONG-TERM CORRECTION AT WIND POWER DEVELOPMENT

AN ANALYSIS OF RESULTS FROM DIFFERENT REANALYSIS DATA SETS

JAKOB SONDEREGGER

Department of Energy and Environment

Division of Electrical Power Engineering

Chalmers University of Technology

SE-41296 Göteborg

Sweden

Telephone: +46 (0) 31-772 1000

Göteborg, Sweden 2011

ABSTRACT

The purpose of this thesis is to compare how different reanalysis data sets affect the results of long-term wind estimations in wind energy development projects. This is done by applying a Measure-Correlate-Predict (MCP) method to four different sets of measured data.

The method is the same as used for linear regression in WindPRO, and the reference series used are MERRA, NCAR and ERA-Interim. The testing will be done using between one to 24 months of measurements on the specific site, making it possible to also detect if there is any seasonality to consider within the model.

The investigation shows that the data sets give different results at different sites. Therefore it is hard to determine if one is better than the other. However, according to the results it is clear that the newer ERA-Interim data set gives results with more or less the same errors as the other two. This suggests that the series can be used as a trustworthy source but further tests needs to be done before this is applied in real normal year corrections. Moreover, it is clear that the uncertainties are lowest when using 10 to 12 months of measured site data, at the same time as the uncertainties sometimes increase when using between 13 to 24 months.

Keywords: Normal year correction, Measure Correlate Predict, MERRA, NCAR, ERA-Interim

ACKNOWLEDGEMENTS

I would like to express my gratitude to the following

Triventus Consulting: For giving me the opportunity to do this thesis and helping me with guidance along the way.

Emil Svensson: For the good times in the office and the collaboration with the programming of MATLAB.

Fredrik Pettersson: For guidance through the databases.

Ola Carlson: For assistance and helping me finish the thesis.

Ana Martha: For your inspiration and the tips regarding the thesis.

My family: For your support throughout my studies.

Göteborg, September 2011

Jakob

TABLE OF CONTENTS

LIST OF FIGURES

1 INTRODUCTION

The diffusion of renewable technologies is important and desirable to meet the current environmental and climate challenges. For instance, the European Union has set a target to increase the share of renewables to 20% of the overall energy consumption by 2020 [1]. More specifically, Sweden is expected to increase the production of electricity from wind power from approximately 4 to 11 TWh until 2020 [2]. These targets together with economical interests drive a rapid development of wind power in Sweden at the moment.

To successfully plan and develop wind farms, it is crucial to have good knowledge about the long-term wind conditions at the target site. The most important parameter is the long-term mean wind speed, since this determines if the project will be economically feasible during its lifetime of 20-25 years.

In order to determine if a site is appropriate for development of wind power, there is a need for reliable wind speed data over a long period of time. Since it is unrealistic in practical situations to exactly measure the mean wind speed during the period data is needed, the short time measurement has to be normal year corrected to describe the wind conditions over a long period of time.

The method used for normal year correction is to perform wind speed measurements on the specific site during a shorter period of time and thereafter apply a Measure-Correlate-Predict (MCP) method to estimate the mean wind speed for a longer period. To do this a long time reference series that can be correlated to the specific site is needed. This reference series can be reanalysed data that is created by different meteorological institutes.

The objective of this thesis is to compare and evaluate the results of one specific MCP model with three different reference series. The model used is the same as used for linear regression in WindPRO[1], and the reference series used are MERRA, NCAR and ERA-Interim. The testing will be done using between one to 24

[1] WindPRO has not been used in this thesis. The analysis was done with the aid of MATLAB, where a program was created to allow for more freedom when testing the data.

months of measurements on the specific site, making it possible to also detect if there is any seasonality to consider within the model.

In section 2, the theory needed to create the program in MATLAB will be described. Section 3 will present the different data sets and how these are handled. This will be followed by section 4 where the implementation of the model is done together with a description of how the end results have been obtained. Section 5 presents the results, which is followed by section 6 where the results are analysed and discussed. Finally, section 7 draws the conclusions of this thesis.

2 THEORY

In this section the theory needed to perform the analysis is presented.

2.1 MCP

As mentioned earlier it is important to have reliable wind speed data over a long period of time. Since it is unrealistic in practical situations to exactly measure the mean wind speed during the period data is needed, a short time measurement is corrected to a normal year. The normal year is then used to describe the wind conditions over a long period of time.

Measure-correlate-predict (MCP) methods are used to estimate wind speeds and directions at a target site where wind power is assessed for development. These methods use two sets of in-data. To begin with a series of measured wind speeds and directions from the target site during a period of time (usually one year) is needed. In addition to this, a reference series from a much longer period needs to be obtained.

The target site data is usually retrieved from an anemometer that is erected at the site that needs investigating. The reference data, however, can be global reanalysed data reaching back to 20 or even 30 years in time. This data can be found at different meteorological institutes. More about target site and the reference series can be found in section 3.1.

A schematic overview of the MCP procedure can be seen in figure 2.1.

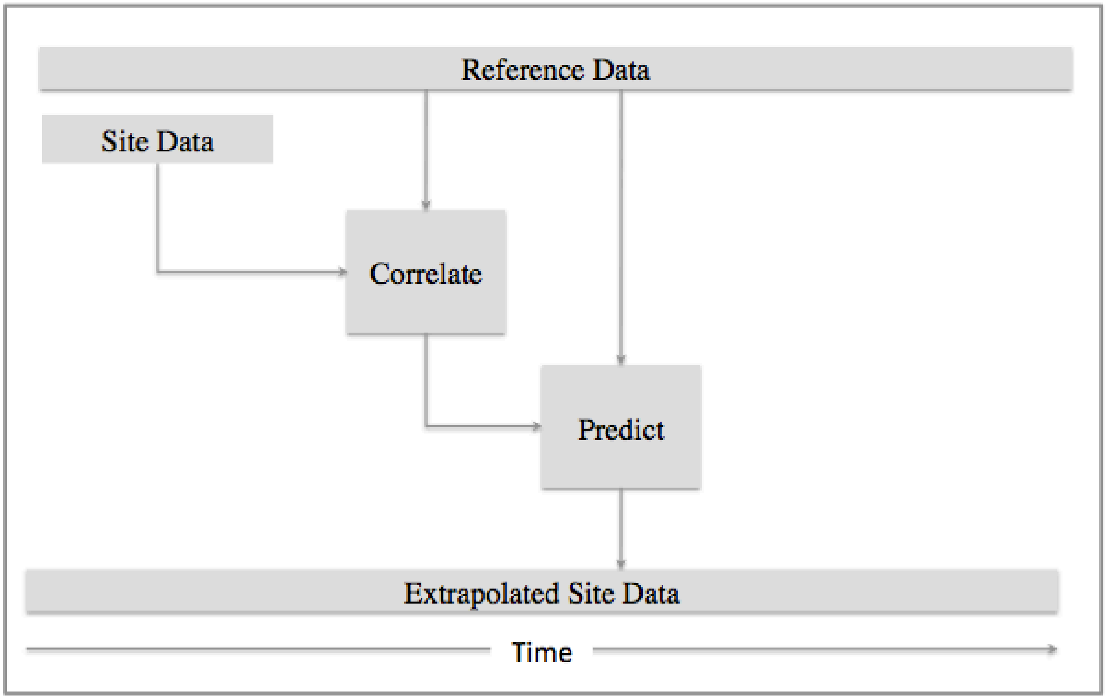

Fig 2.1 *Scheme of the MCP procedure.*

As can be seen in figure 2.1, the two series needs to be correlated. There are a number of available statistical methods that are used for this correlation procedure. These include linear regression, the Weibull scale method, the index method and the matrix method. In this study a linear regression model with residuals will be used since it is less complex than the matrix method but at the same time more accurate than the index-method and the Weibull scale method. Also, the objective of the paper is to compare different types of reference series and therefore the choice of statistical method is of less importance as long as the same is used.

2.2 LINEAR REGRESSION

Linear regression is in general a simplified model where a function is fitted to the available data. The function can be described by equation 1 [3].

$$y = \alpha + \beta x \qquad (1)$$

The unknown parameters α and β have to be determined with the aid of the data available. The most common method for this is called the least square method, where β can be determined with the aid of the following equation [3].

$$\beta = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}}{\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}} \qquad (2)$$

When α and β have been determined, usually the correlation coefficient (-1≤r≤1) is calculated in order to explain if the estimated line describes the correlation between the input data well or bad. A value of 0 describes that there is no correlation between x and y and 1 or -1 describes that there is a linear, or negative linear, correlation. This coefficient can be determined with the following equation [3].

$$r = \frac{\sum_{i=1}^{n} x_i y_i - \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n}}{\sqrt{\left(\sum_{i=1}^{n} x_i^2 - \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{n}\right)\left(\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}\right)}} \qquad (3)$$

2.3 RESIDUAL ANALYSIS

Normally when performing the regression there is a scatter of the points around the line. In order to determine if the linear model describes the situation in a good way, a residual analysis can be performed. The residuals are defined as

$$e_i = y_i - \overline{y_i}, i = 1,2,...,n \qquad (4)$$

where y_i is an observation and $\overline{y}_i$ is the corresponding fitted value [5]. A plot of the residuals can then tell many things about the model and how this fits the real situation. Some residual patterns that can normally be the result of a plot can be seen in figure 2.2.

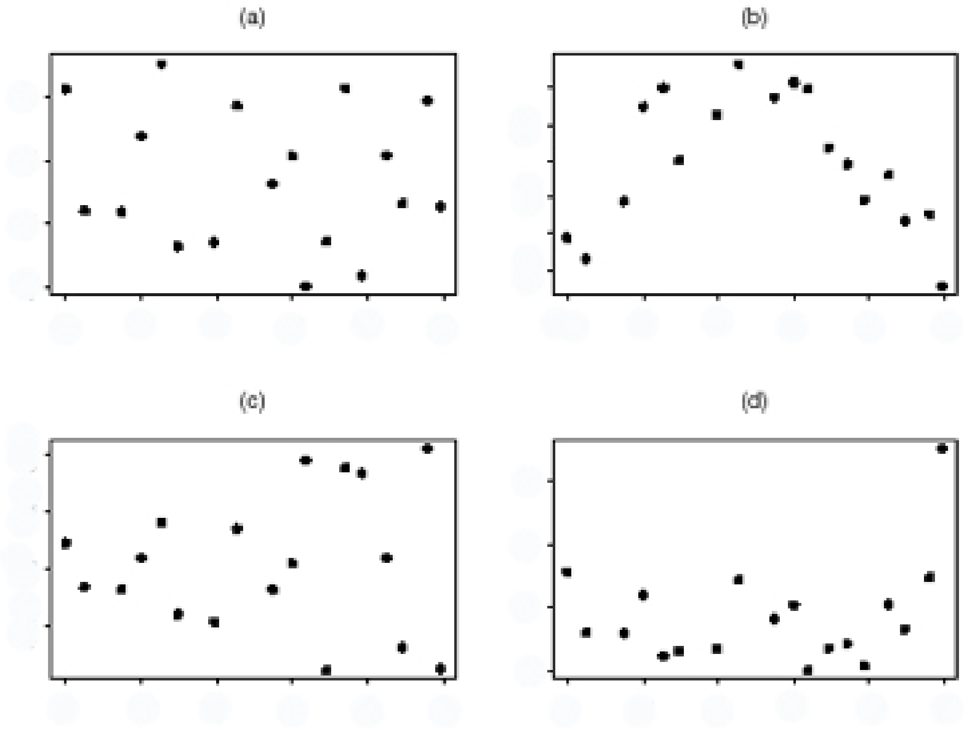

Fig 2.2 *Patterns for residual plots: (a) satisfactory; (b) nonlinear; (c) funnel; (d) underestimation with outlier [5].*

The pattern in (a) is desirable. Here, a residual plot can be seen with no systematic pattern, where the random errors have a constant variation and zero mean. The other three examples show cases where this is not true and if the residual plot in this analysis would show any of these patterns the model will not describe the real situation in a good way.

Residual plots that look like (b) indicate a model inadequacy, which means that higher order terms should be added to the model [5]. If residuals appear as in (c), the variance of the observations may be increasing with time or magnitude [5]. Finally, in (d) an outlier can be seen and this can be a result of a bad data series that still contain non-valid data points.

2.4 INTERPOLATION

The reference series are available at certain geographical coordinates distributed in a defined grid. Since the target site often is somewhere between these coordinates the available reference data can be interpolated to more accurately describe the target site.

There are a number of different interpolation methods available to generate values that are located outside the already predefined locations. In an evenly distributed grid, the bilinear interpolation is a commonly used method. In this method four points that surrounds the target site are the reference points, see fig 2.3 [6].

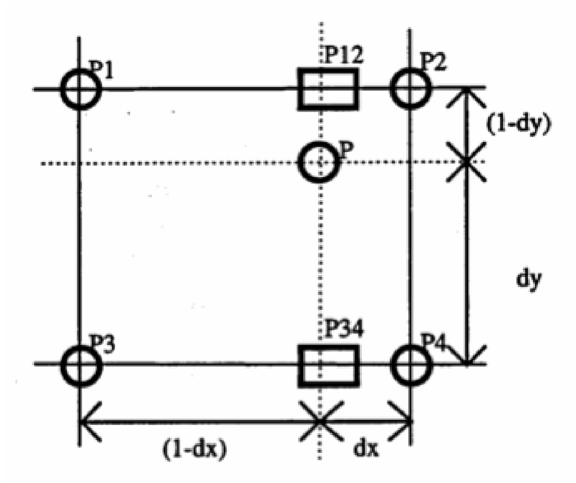

Fig 2.3 *Schematic picture of bilinear interpolation [6].*

The bilinear interpolation for point P is done in three steps. First, the points P12 and P34 are linearly interpolated with the aid of the following equations.

$$P12 = dx \cdot P1 + (1 - dx)\, P2 \qquad (5)$$

$$P34 = dx \cdot P3 + (1 - dx)\, P4 \qquad (6)$$

The third step is then to interpolate the value of P with the aid of the two earlier interpolated values P12 and P34.

$$P = dy \cdot P12 + (1 - dy)\, P34 \qquad (7)$$

2.5 U AND V COMPONENT OF THE WIND

The reference series are broken into its two horizontal components. The "U" component represents the east-west component of the wind while the "V" component represents the north-south component [7].

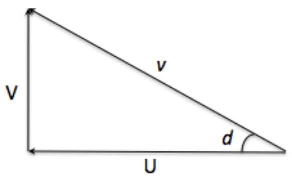

Fig 2.4 *U and V components of a wind vector.*

From this the actual wind speed can easily be calculated using Pythagoras theorem.

$$v = \sqrt{U^2 + V^2} \tag{8}$$

The correct direction of the wind is a bit more complicated to obtain, since the quadrant the wind vector end up in will depend on whether the U and V components are positive or negative, see fig 2.5.

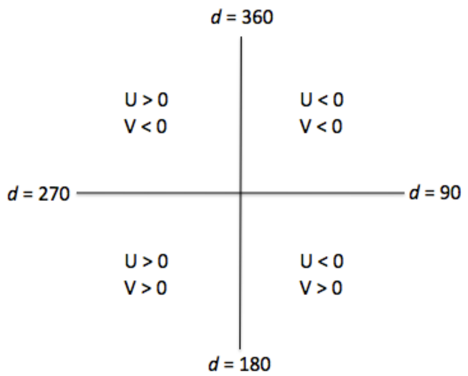

Fig 2.5 *Quadrant where the wind direction belong is depending on value of U and V component [7].*

Another thing that complicates the matter is that these components are measured in the direction it is blowing towards, while this analysis needs the direction it is blowing from. This means that when looking at fig 2.6 all the < and > signs become the opposite compared to what is normally true.

Keeping this in mind the direction can be calculated with the following formula.

$$d = \arctan\left(\frac{U}{V}\right) + \theta \tag{9}$$

Where θ depends on the following statements [7]:

- θ = 0 if U < 0 and V < 0.
- θ = 180 if V >= 0.
- θ = 360 if U >= 0 and V < 0.

3 DATA

This section will describe the data used in the analysis. It will describe the reference series as well as the site series. Finally, the data series have been plotted in a Weibull probability plot in order to show the wind speed distribution of the winds.

3.1 REFERENCE SERIES

Atmospheric reanalysis data is available from a number of different climatological institutes. They consist on the synthesis of worldwide observational data by an atmospheric model into a global three-dimensional grid [8]. These types of data series are useful when predicting the future wind climate at a specific site since they consist of data from 20-30 years back in time.

The first dataset of this type was produced during the 1990’s by the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR). Until 2009 this was the only freely available dataset for commercial use. Now, however there are more datasets available. Among these are the Modern Era Retrospective-analysis for Research and Applications (MERRA), published by NASA, Climate Forecast System Reanalysis (CFSR), published by NCEP. The newest dataset is the ERA-Interim that has recently been released by the European Centre for Medium Range Weather Forecasts (ECMWF). See Table 3.1 for more info.

Table 3.1. *Summary of the reanalysis datasets used in the analysis [9,10,11].*

Dataset	Institution	Release year	Time span	Time resolution	Vertical level
NCAR	NCEP	1995	1948-present	6 h	0.995 sigma
MERRA	NASA	2009	1979-present	1 h	0.995 sigma
ERA-Interim	ECMWF	2011	1989-present	6 h	1000 hPa

The first two series in the table has been provided by Triventus Consulting AB, while the ERA-Interim data set has been retrieved from the ECMWF database MARS. This was done with the aid of a python script, and the result was a GRIB file, that had to be translated with the aid of the program dgrib. GRIB is the format used by the meteorological institutes of the world to transport and manipulate weather data and is the foundation of the forecasts we see around us in our daily life [12].

Something important to keep in mind is that both MERRA and NCAR are specified with sigma levels. This means that the wind speed is calculated at the exact desired level over land. This is not the case for ERA-Interim however; since this service is not available for free download. Instead the pressure levels have been used and this means that a certain level above sea has been used instead.

3.2 TARGET SITE SERIES

In this study wind measurement series from four different sites have been analysed. These sites are Älvsborgsbron, Näsudden, Oskarshamn and Ringhals and can be seen in figure 3.1.

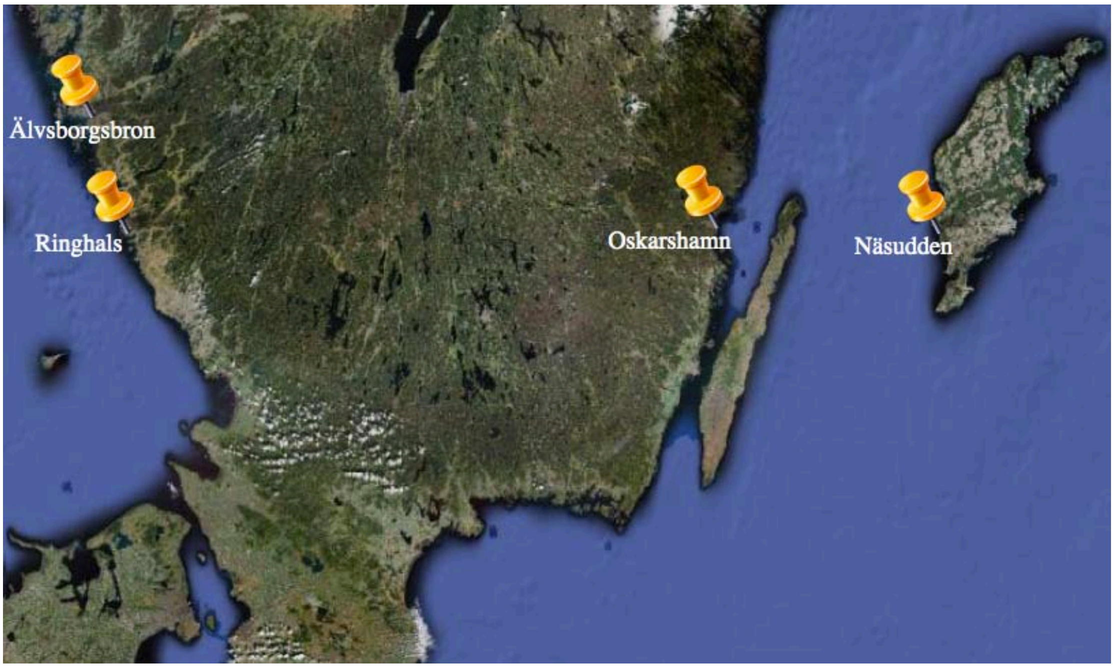

Fig 3.1. *Map of south of Sweden with the target sites.*

As can be seen in table 3.1 the series have different years from when they were measured. This is especially worth noting in the case of Näsudden where the period reaches from 1981 to 1997. This means that this series combined with the ERA-Interim reference series only have seven years of concurrent data compared to the other series at the same site have fourteen.

Table 3.2. *Summary of the measured site series used in the analysis.*

Site	Latitude	Longitude	Mean speed	No. of Years
Älvsborgsbron	57.69 N	11.90 E	6.04 m/s	9 (2002-2010)
Näsudden	57.07 N	18.20 E	7.86 m/s	14 (1981-1997, except 1984, 1990, 1991)
Oskarshamn	57.27 N	16.43 E	6.85 m/s	9 (2001-2009)
Ringhals	57.26 N	12.11 E	8.00 m/s	9 (1999-2007)

3.3 WIND DISTRIBUTIONS

The wind distribution at the different sites together with the corresponding reference series can be seen in the figures that follow.

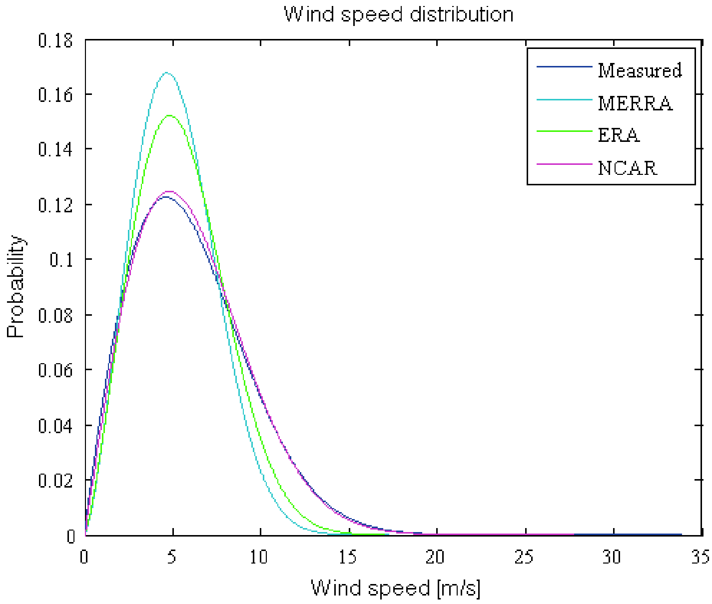

Fig 3.2. *Weibull distribution of the site and reference series at Älvsborgsbron.*

Something worth noticing in the figure 3.2 is that there are remarkable differences in the wind distribution between the reference series and the site series (except from NCAR). This makes it relevant to draw the conclusion that estimating the behaviour of the wind solely on the reference series gives too large uncertainties. This is due to that parameters adjusted to fit one site cannot be unchanged to the next site [13]. This is why a measurement combined with a reference series is needed to get a more representative description of the wind at a specific site.

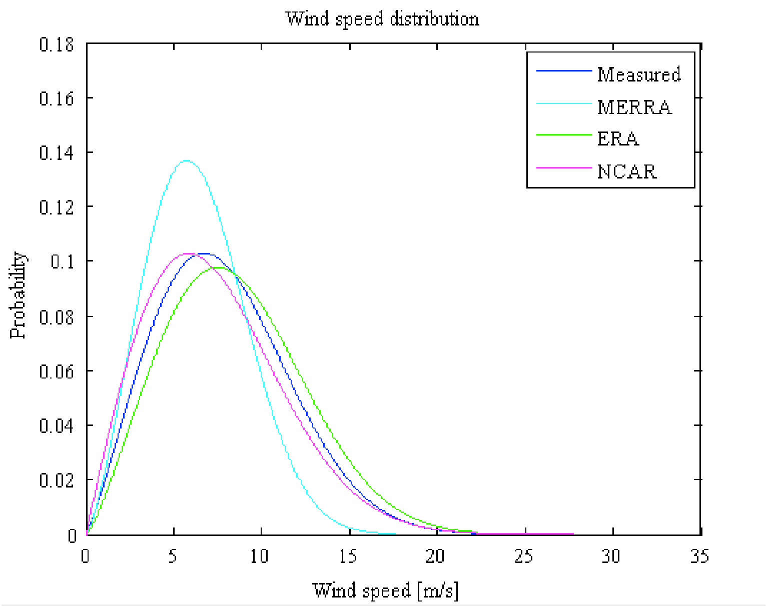

Fig 3.3. *Weibull distribution of the site and reference series at Näsudden.*

Here as well, a large difference between the reference series and the measured series can be noticed. One thing that is especially interesting is the difference in mean speed of the ERA-Interim series, which amounts to 8.55 m/s, and MERRA that is down at 6.37 m/s. Also here the distribution that fits best to the measured series is NCAR, although it is not as obvious as in the case of Älvsborgsbron.

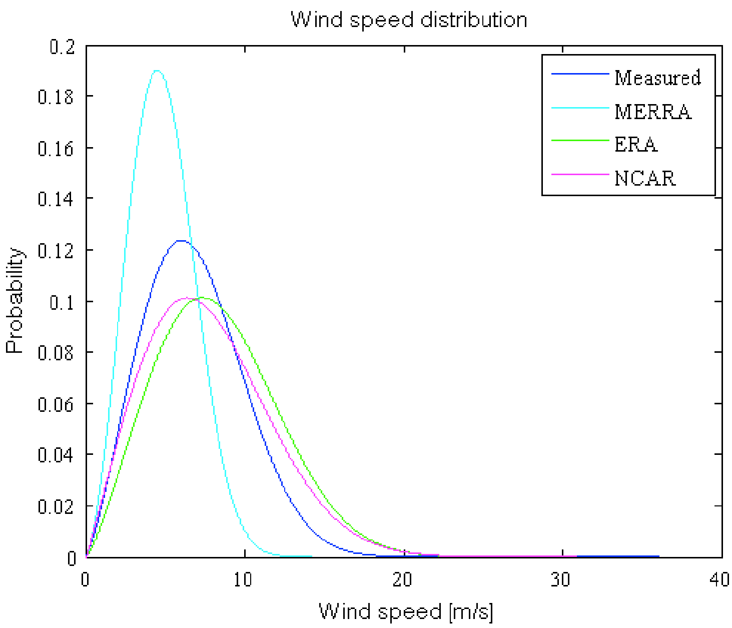

Fig 3.4. *Weibull distribution of the site and reference series at Oskarshamn.*

Also here it is interesting to see the difference in mean wind speed. At this site the difference is even bigger than at Näsudden with MERRA calculated to be 4.85 m/s while NCAR amounts to 7.76 m/s and ERA-Interim reaches as high as 8.33 m/s.

As at the previous sites NCAR seems to best describe the wind distribution also at Oskarshamn.

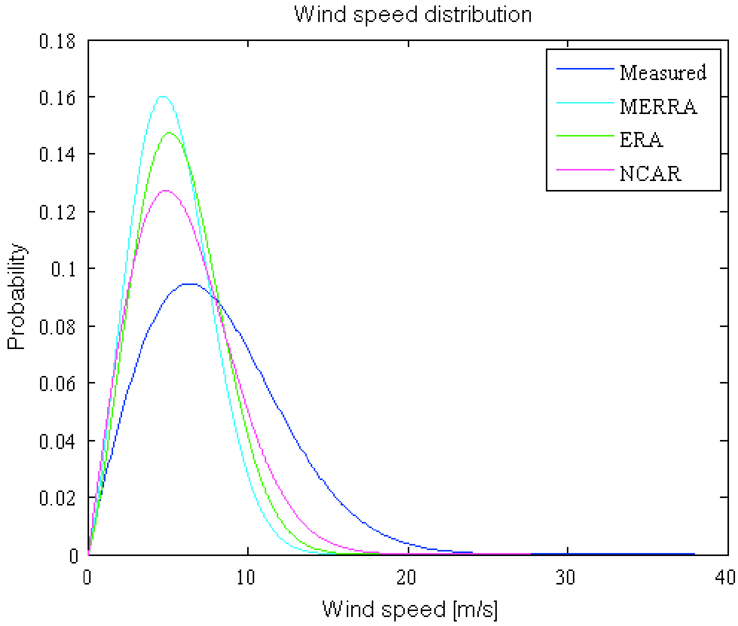

Fig 3.5. *Weibull distribution of the site and reference series at Ringhals.*

Also at this site it seems like NCAR describes the real situation best. However, all series have a much lower mean wind speed ranging from 5.8 m/s to 6.1 m/s, while the actual wind speed is 8.00 m/s.

4 WIND DATA ESTIMATION

This section will describe how the analysis has been carried out, starting with how the large amount of data needs to be handled. It continues with the linear regression model and ends with a brief explanation of how the artificial series have been compared to the real measured series.

4.1 DATA HANDLING

In order to correlate the series and to perform a long time estimate of the wind speeds and directions on the target site, the series has to be modified to fit each other. To begin with the site series needs to be controlled for incorrect data points. This is a time consuming but necessary part of the analysis since inaccurate data points affects the statistical method and the end result. For example points that give the value zero is a typical case when the anemometer might have been frozen.

The corresponding points of the reference series is then deleted so only the points that describe the wind speeds and directions at the same time are left.

Both series are then sorted after the reference series wind directions in 360 different one-degree sectors. However, to get sufficient number of data points in each sector a larger window of 30 degrees is used. Therefore, the result of the sorting is 360 bins with wind data in 30-degree sectors [14].

4.2 THE REGRESSION MODEL

Once the data is in the form that is desirable, the creation of an artificial series can begin.

The regression model is done for both wind speed and wind direction in all 360 sectors. While the wind speeds are modelled with the aid of the ordinary least square method, the wind veers/directions are simply modelled as a constant mean value of the difference between the reference direction and the site direction (see fig 4.1 and 4.2).

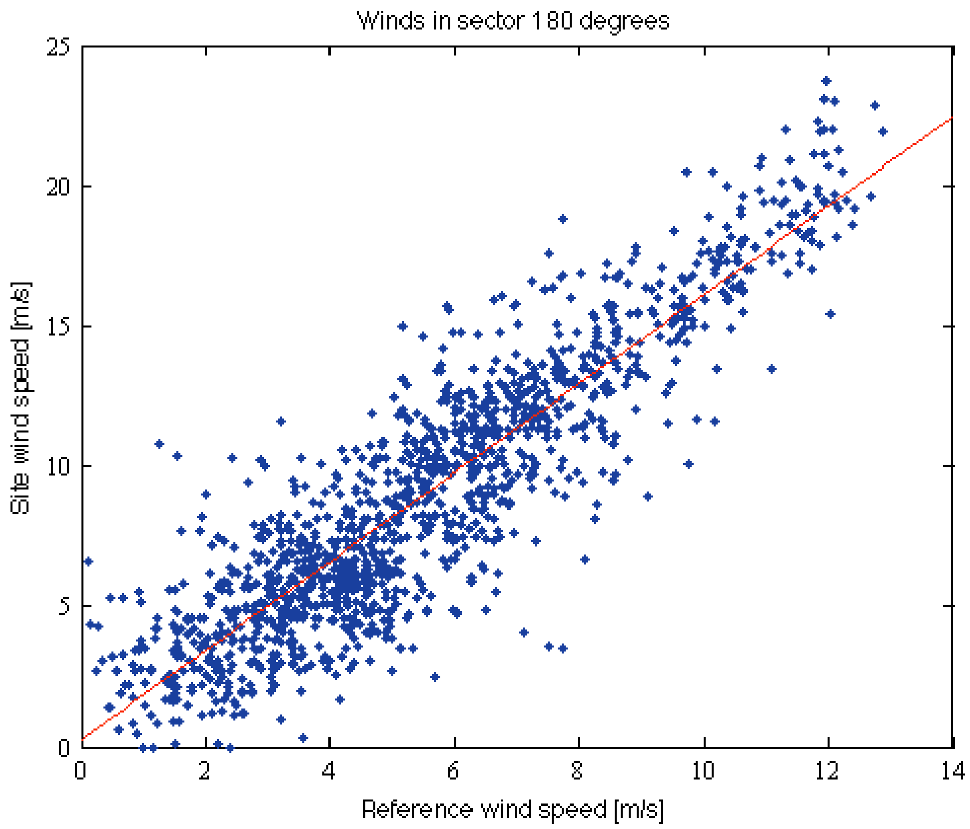

Fig 4.1. *Linear fitting of wind speed at Ringhals in sector 180 degrees.*

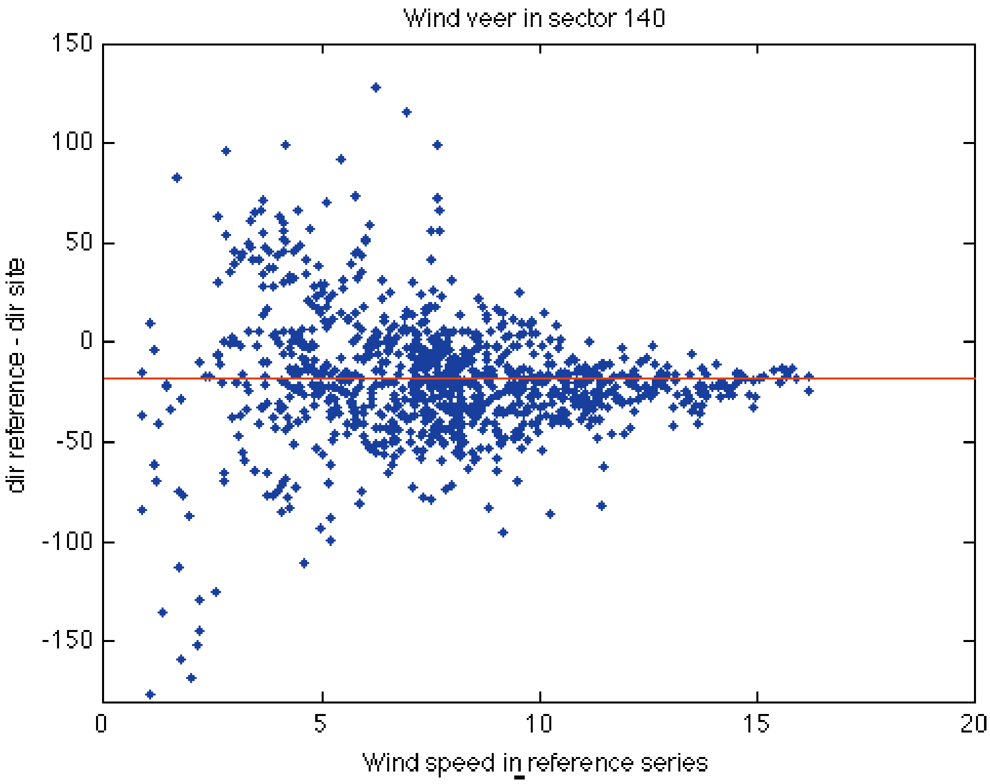

Fig 4.2. *Constant fitting of wind veer at Älvsborgsbron in sector 140 degrees.*

At the same time as the regression, the correlation and the residuals are calculated together with their standard deviation, variance and mean values.

According to EMD International A/S the distribution of the random errors can be assumed to follow a zero mean Gaussian distribution [14]. This has also been validated with this model and an example of the residuals in one sector can be seen in fig 4.3.

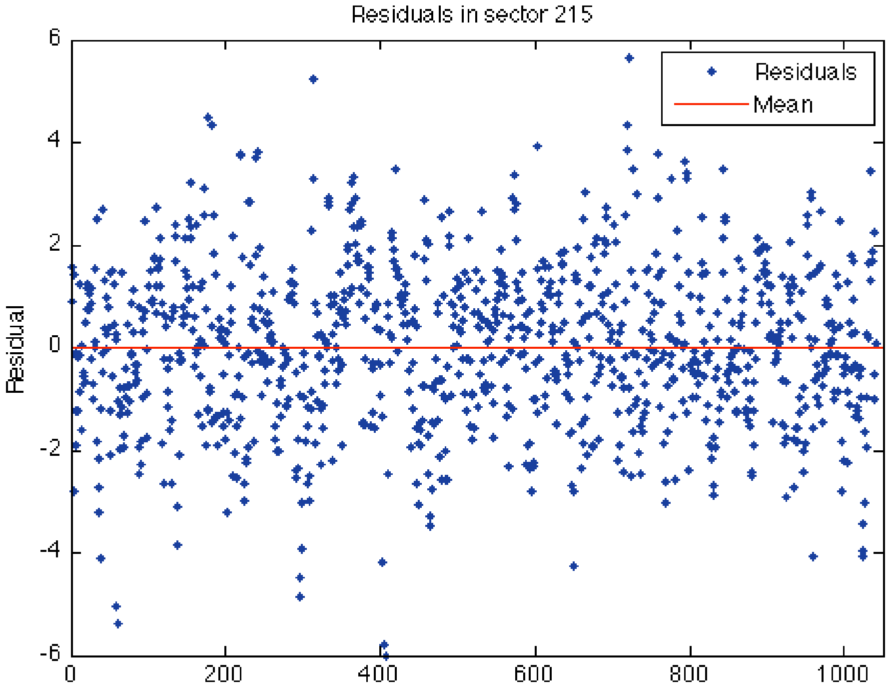

Fig 4.3. *Plot of wind speed residuals at Älvsborgsbron in sector 215 degrees.*

Therefore, when creating the artificial series, these residuals are included. The model will then look like the following:

$$y = \alpha + \beta x + e \tag{10}$$

Here, e represents a random Gaussian distributed error [14]. This is relevant to include since the points above the regression line contains significantly more energy than the points below, due to the third power relationship between wind speed and energy [14].

This gives 360 equations of the same kind as equation (10) that later can be used together with the reference series to create the artificial series.

Worth noting is that the residuals are important when creating the artificial series. However, when comparing the mean wind speed and not the energy production the zero mean residuals will have no effect on the end result.

4.3 EVALUATION OF RESULTS

When the artificial series has been created it is time to evaluate how good this series represents the real situation.

The testing has been done with measured site data ranging from 1 to 24 months. Since the outcome of the artificial series varies depending on what year, and how long measurements has been used, it is reasonable to compare more than just one result.

When using between 1 to 12 months, there is only one option of year each time the model is performed. Therefore, an artificial series is created for all years that measured data exists. When using only one month January is used, when using two months January and February are used and so on. This means that the months that are used together always are connected in reality as well making it possible to see more seasonal changes. Another option could have been to randomise all different combinations of months. This would demand more computer power, since the tests at Näsudden, when using between 13 to 24 months, took approximately one hour to run each.

When using between 13 to 24 months there are many more options of how to combine different years and the result varies with every option. Thus, one artificial series of every possible combination has been created. For all these artificial series a mean wind speed has been calculated. For example if a series between 2002 and 2010 is available there are 73 possible combinations of years.

This has to be compared with the actual long-term wind speeds at the site and the error is calculated as follows:

$$E_V = \frac{\overline{V}_{pred} - \overline{V}_{act}}{\overline{V}_{act}} \times 100, \qquad (11)$$

where $\overline{V}_{pred}$ is the predicted mean wind speed and $\overline{V}_{act}$ is the actual mean wind speed. The error is then presented in percent.

5 RESULTS

In this section the results will be presented in two graphs for each site investigated. In these graphs three series are compared, one for each reference series. For each site the mean wind speed error and the absolute mean wind speed error is presented. The y-axis represents the error in percent, while the x-axis represents how many months of measured site data that has been used in the creation of an artificial series.

Since it is most common to do these types of analyses with 12 months of data, a table with the error at this point is also presented.

5.1 ÄLVSBORGSBRON

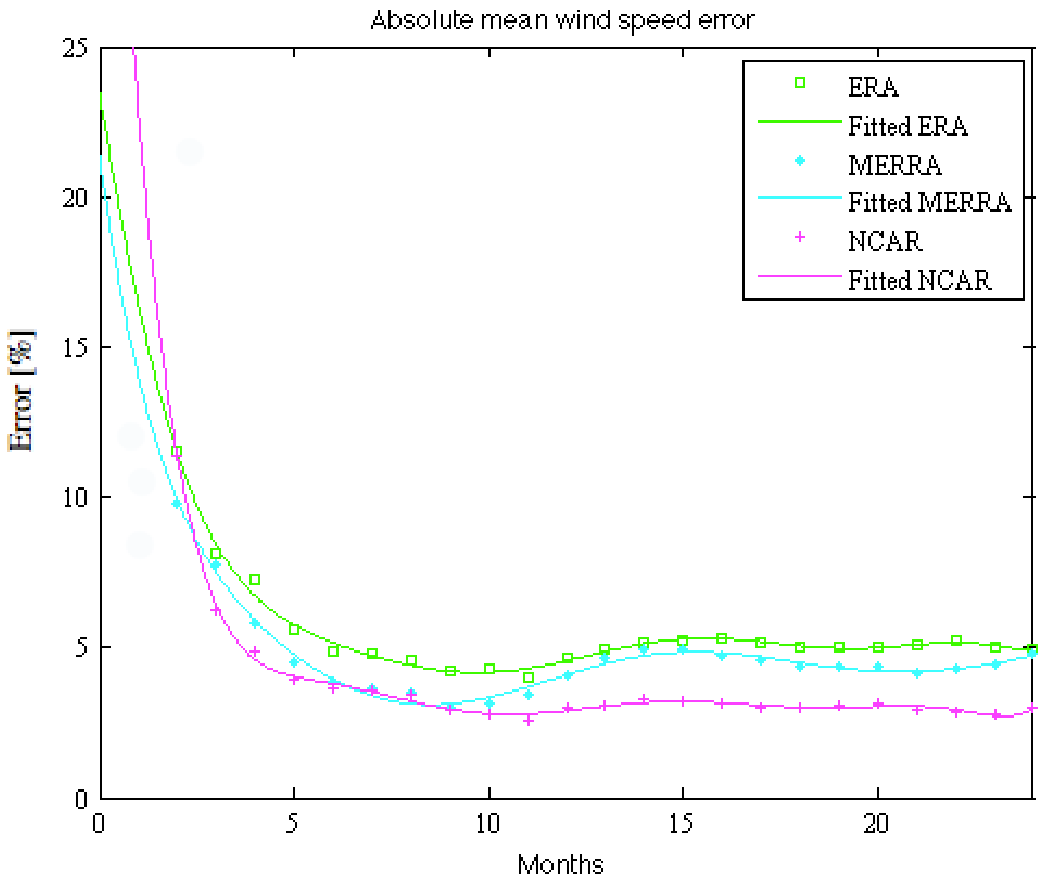

Fig 5.1 *Estimated absolute mean wind speed error at Älvsborgsbron.*

As stated in table 3.2 the actual mean speed during 2002 to 2010 at Älvsborgsbron is 6.02 m/s.

In figure 5.1 it can be seen that the predicted mean wind speed error is reduced rapidly when increasing the number of months of measured data to between 10 to 12. Using less data than this when doing an MCP on this site will probably give large uncertainties. When using more than 12 months the results are hardly improved, if not even worse in some cases. This can be related to a seasonal dependence and to the fact that if not using a complete year the artificial series might give an over-, or under-estimation of the mean wind speed. This is due to the fact that some months are normally windier than others. Therefore, depending on which months are included in the analysis, the artificial series might give a larger error than if using one complete year.

Worth noting is that NCAR is the series that starts by giving the largest error while when increasing the number of months it reduces the error most. This reference series seems to be best suited of the three for this site.

Also, it can be seen that ERA-Interim and MERRA result in more fluctuations when using more than 12 months of measured data.

As can be seen in table 5.1 NCAR gives the best estimation at this site, followed by MERRA and last ERA-Interim.

Table 5.1 *Error in prediction when using 12 months of reference series at Älvsborgsbron.*

	MERRA	ERA-Interim	NCAR
Error [%]	4.1	4.6	3.0

It is also interesting that when plotting without the absolute mean error it is clear that all series underestimates the wind speed at the site, as can be seen in the following figure.

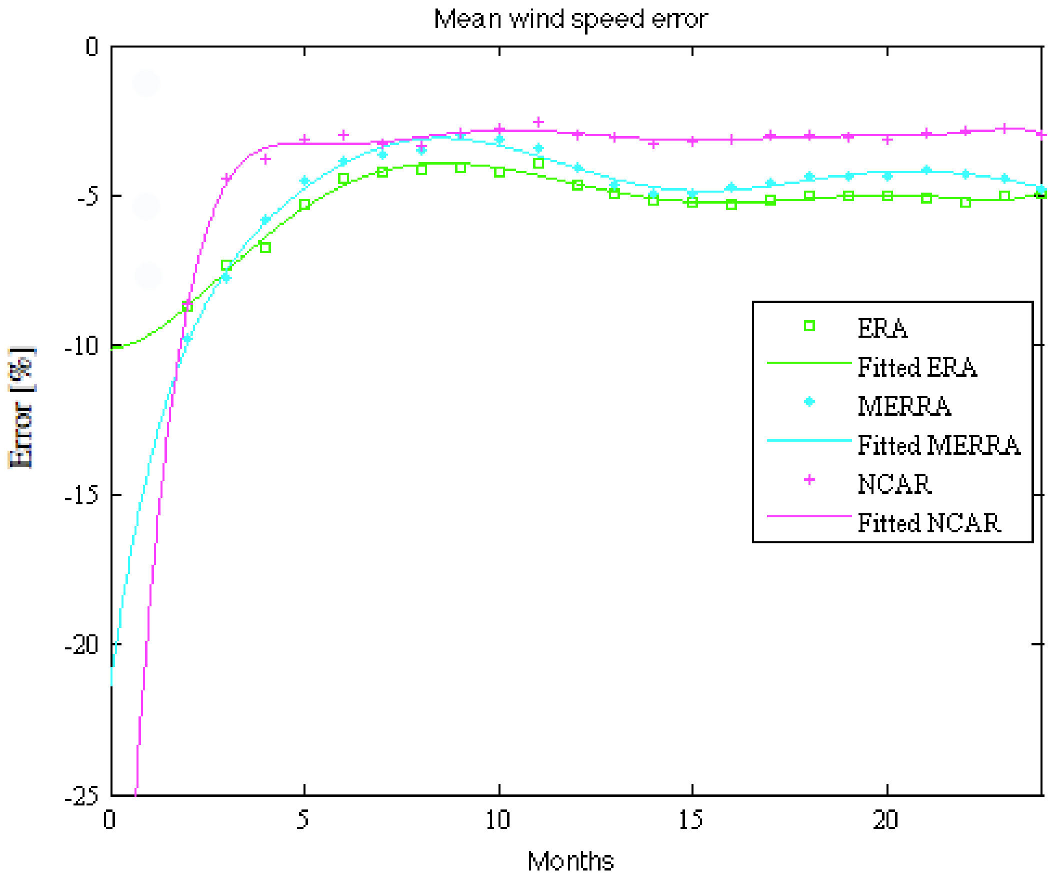

Fig 5.2 *Estimated mean wind speed error at Älvsborgsbron.*

5.2 NÄSUDDEN

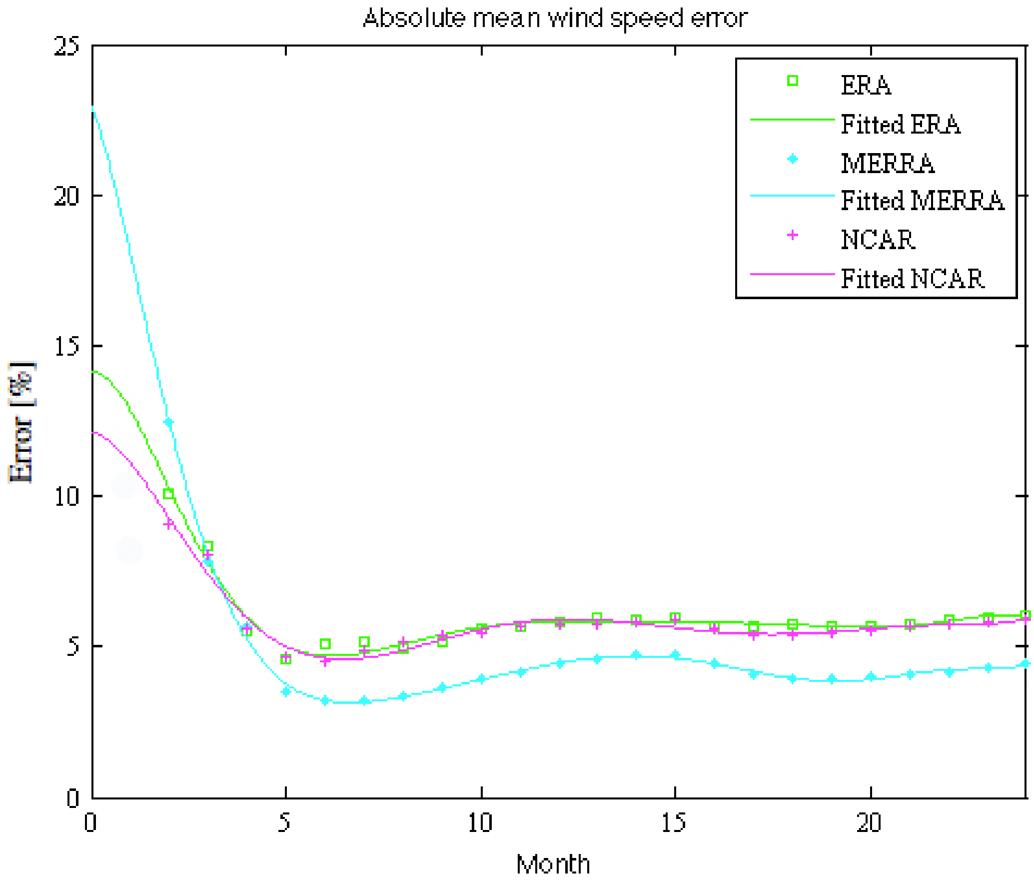

Fig 5.3 *Estimated absolute mean wind speed error at Näsudden.*

The mean wind speed at Näsudden for the years used in the analysis was 7.86 m/s.

The results from Näsudden show the same general pattern as at Älvsborgsbron. The most obvious difference is that at this site MERRA is the reference series that gives the best estimations, while NCAR and ERA-Interim are very similar.

Something interesting is, that even if it is a very small change, it seems like the best estimations are made when using 6 to 7 months of reference series and that the errors then increase a little bit up to 12 months and thereafter are more or less constant.

The series that gives most seasonal change is MERRA, while the others are fairly constant after 12 months. Table 5.2 presents the errors at 12 months of used reference series. It can be seen that ERA-Interim and NCAR give very similar results, while MERRA clearly gives the best result.

Table 5.2 *Error in prediction when using 12 months of reference series at Näsudden.*

	MERRA	ERA-Interim	NCAR
Error [%]	4.4	5.8	5.7

Also at this site all series underestimate the actual wind speed.

Important to notice is that it seems like NCAR gives a very good result when only using one month and that with ERA the results are not improved when using more site data in the analysis. This is however due to the fact that when using only a few months of site data there is a big difference in the results. Some estimations underestimate, while others overestimate the wind speed. When comparing this to the real situation and taking the mean values of these estimations the result sometimes end up very close to zero, while in fact the result is very uncertain.

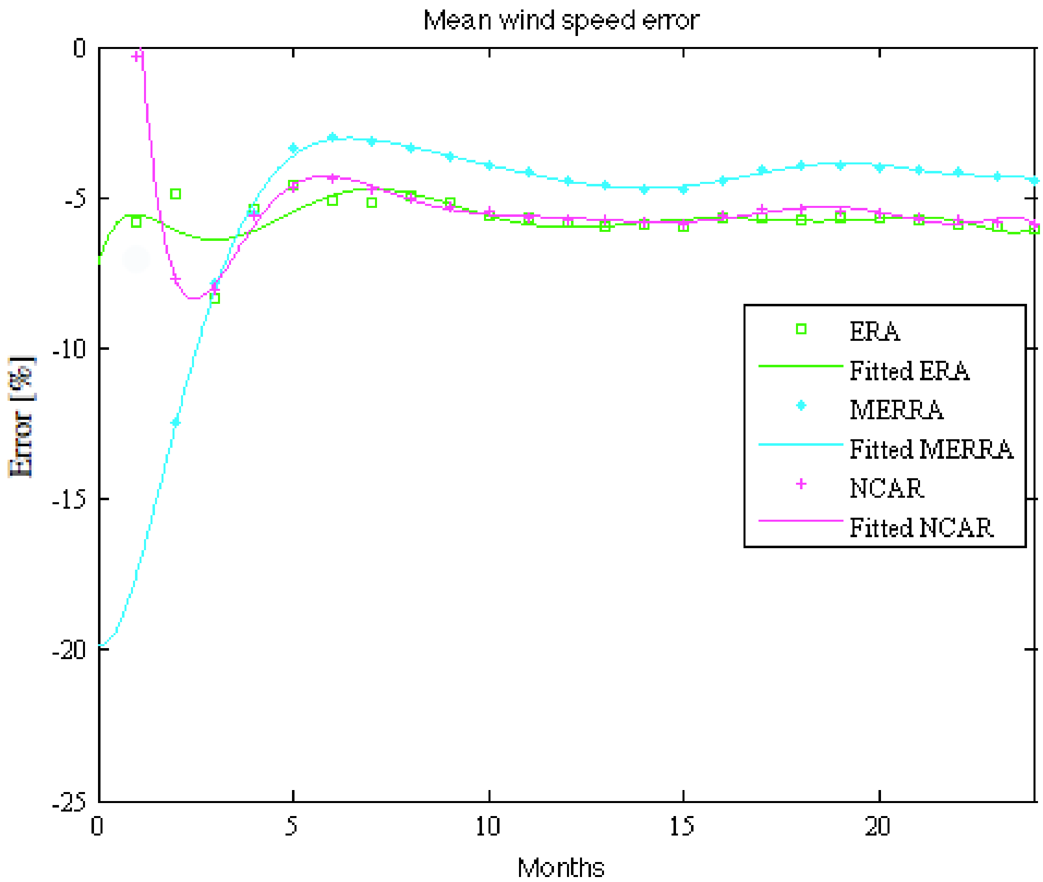

Fig 5.4 *Estimated mean wind speed error at Näsudden.*

5.3 OSKARSHAMN

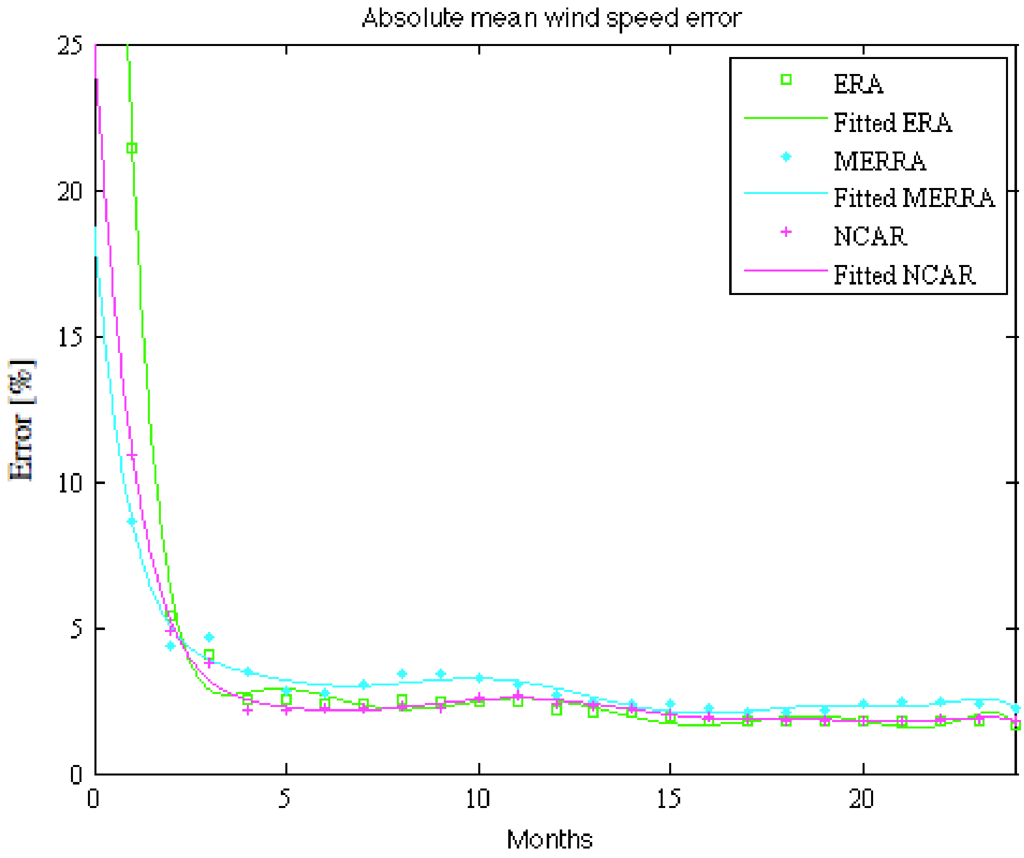

Fig 5.5 *Estimated absolute mean wind speed error at Oskarshamn.*

The predicted mean wind speeds are closer to the real value at Oskarshamn compared to the earlier presented sites. Although there is not much difference, ERA-Interim is the best reference series after twelve months.

Some tendencies towards an increased error after twelve months can also be seen. Otherwise this site give less fluctuations when using more than one year of measured data.

What is surprising is how equal results all the reference series gives. As can be seen in the table that follows the results are clearly better than at the two earlier sites, with ERA-Interim giving the best result.

Table 5.3 *Error in prediction when using 12 months of reference series at Oskarshamn.*

	MERRA	ERA-Interim	NCAR
Error [%]	2.7	2.2	2.4

This site gives an interesting result when plotting the mean wind speed error. Only MERRA underestimates the wind while both NCAR and ERA are very close to no error or an overestimation.

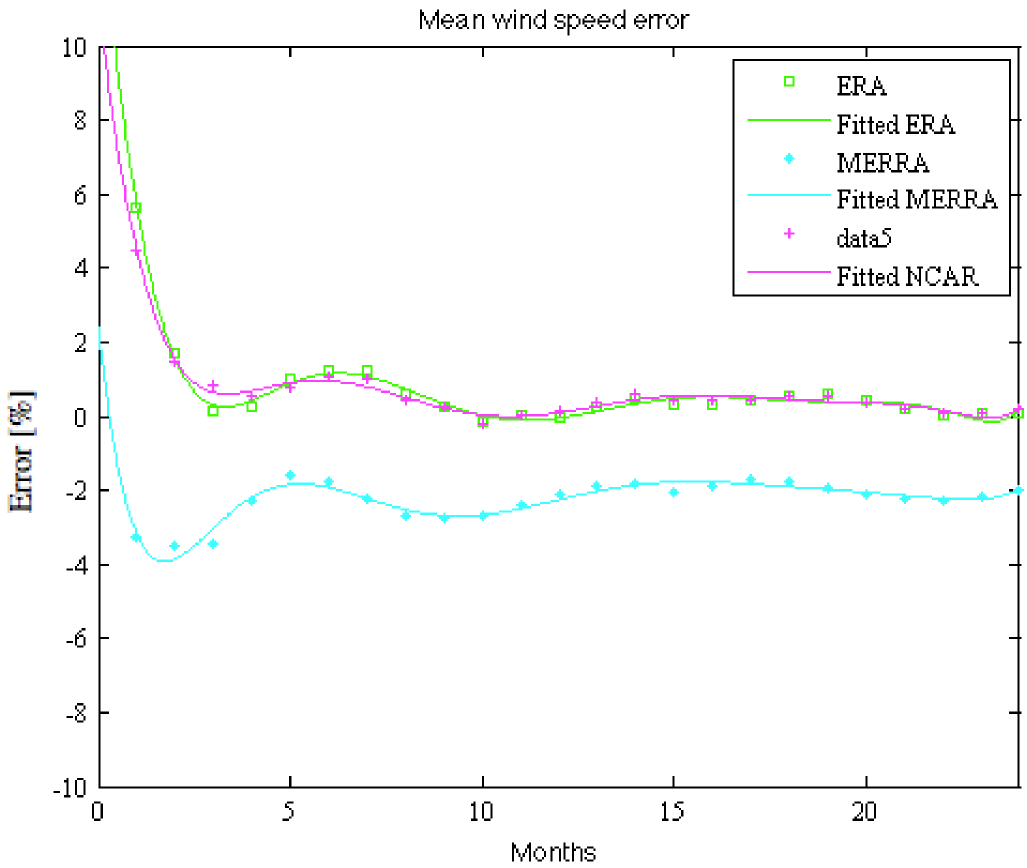

Fig 5.6 *Estimated mean wind speed error at Oskarshamn.*

5.4 RINGHALS

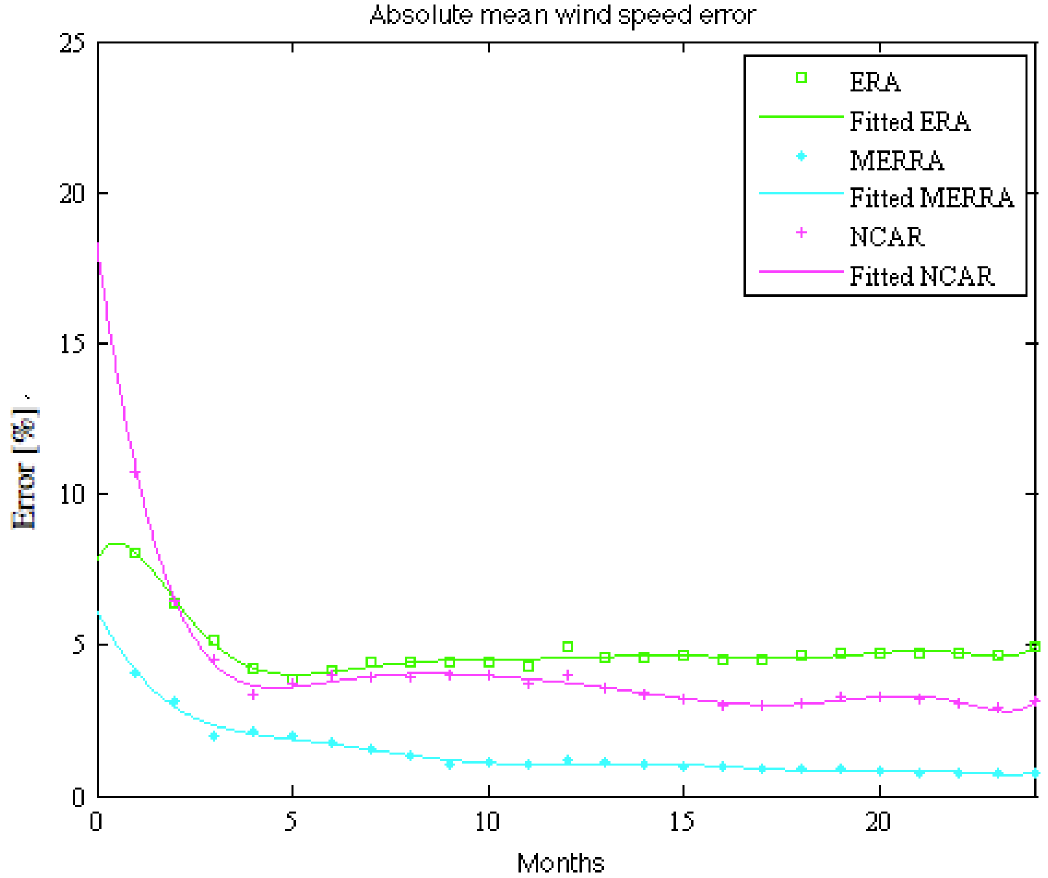

Fig 5.7 *Estimated absolute mean wind speed error at Ringhals.*

When using MERRA at this site the best value of all estimations is obtained. As low as 1.2 % error estimation over a nine-year period, while the other two series give a more uncertain result.

It is interesting to see how big differences the different series gives. This cannot be seen at any other site, although the same pattern is obvious at this site as well.

Table 5.4 *Error in prediction when using 12 months of reference series at Ringhals.*

	MERRA	ERA-Interim	NCAR
Error [%]	1.2	4.9	4.0

This site also gives similar results as at Oskarshamn when plotting the mean wind speed error. However, here it is the other way around, with MERRA being the series that estimates wind speeds close to zero or even positive values. NCAR and ERA on the other hand underestimates the winds.

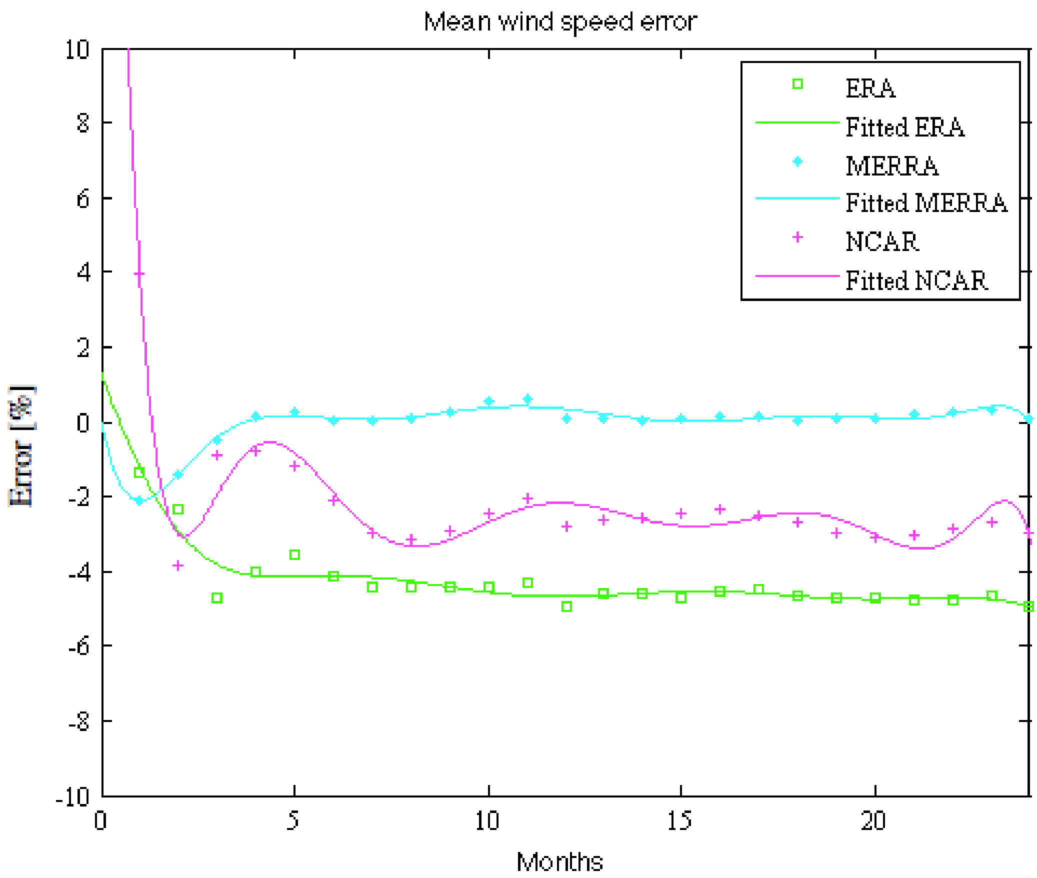

Fig 5.8 *Estimated mean wind speed error at Ringhals.*

5.5 SUMMARY

In this section a summary of the results when using twelve months of site data is presented. This is done to get an overview of how well the different reference series estimate the wind condition at each site.

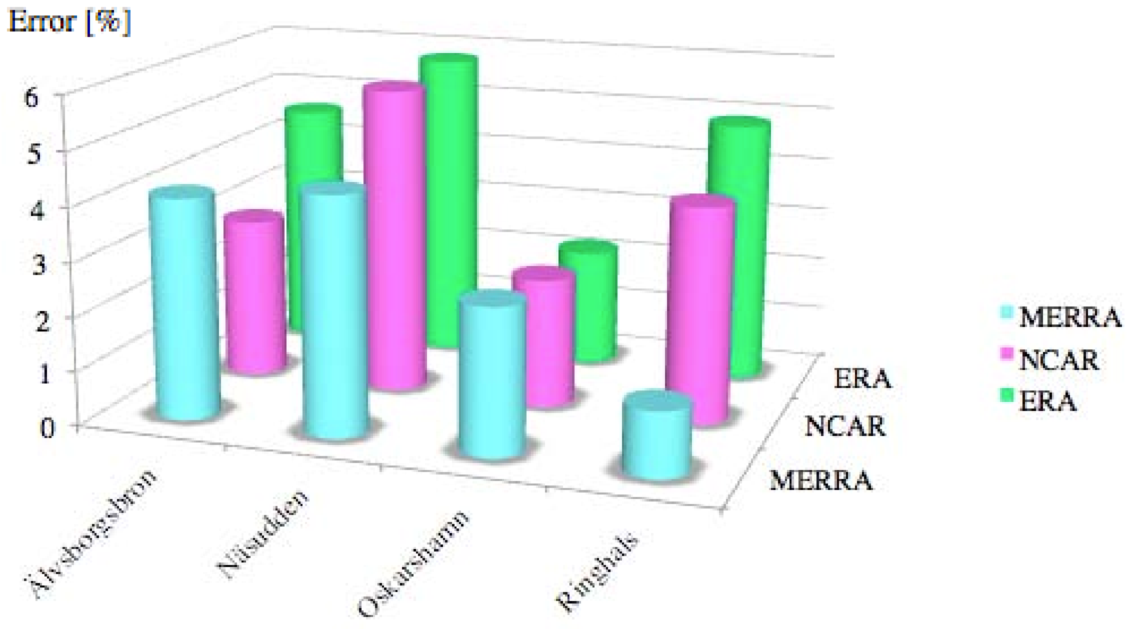

Fig 5.9 *Results at all the target sites when using twelve months of site series.*

6 DISCUSSION

This section presents some general observations and discusses around the uncertainties in the analysis.

6.1 GENERAL OBSERVATIONS

It is obvious that all reference series give results that show the same pattern. There is a clear reduction in error when increasing the number of months of used site series to somewhere in between 10-12 months.

There are tendencies towards increased errors when not using complete years of the site series at all sites.

Moreover, it is hard to determine if one series is better than the other, since they all give the best result at one site each, although ERA-Interim is only better with a minor marginal at Oskarshamn, where all series give very good results. Figure 5.9 also show the difficulty of finding one series that is better than the other.

Finally, it seems like a very general conclusion can be that most series most times underestimates the actual wind conditions at the sites. This might also be a general problem for the regression model, but this observation needs more investigation before it can be a conclusion.

6.2 UNCERTAINTIES

There are a number of uncertainties within this study that might have affected the end result.

To begin with the length of the concurrent data available for the different sites vary. This might have an impact of the results since longer concurrent series gives a more reliable result. Worth noting is that the ERA-Interim reference series together with the measured series at Näsudden only have seven concurrent years while the other two reference series have fourteen.

Another important parameter that needs to be lifted is that for the free version of ERA-Interim, which has been used in this thesis, the sigma levels of the U and V winds are not available for downloading. However, this is available when using the series in commercial matters in exchange for a fee. This might increase the uncertainty since the pressure levels had to be used. Easily explained it can be said that the difference between these levels is that the sigma level specifies the height over land at a specific site, while the pressure levels indicate the height above sea level. However, the differences are in reality more complex but this will not be analysed in this thesis.

Furthermore, there is a difference between the reference series where MERRA has data every hour, while ERA-Interim and NCAR only have available data every six hours. In this study for NCAR and ERA-Interim the concurrent points have been used. Another way of doing it could be to take the mean wind speed between these six-hour points and thereby a more accurate result might have been obtained. This is however an uncertain statement it self, since this also might give worse results. Nevertheless, it would change the result in one way or the other.

7 CONCLUSIONS

The purpose of this thesis was to compare how different reanalysis data sets affect the results of long-term wind estimations in wind energy development projects.

The investigation in this paper shows that all three tested reference series are trustworthy sources when performing an MCP. Although, it is impossible to draw a conclusion of which one is the best. They all show different qualities at different sites.

All series seem to generally underestimate the wind speeds at the sites. This can also be a problem with the regression model and needs to be investigated further.

Worth noting is that NCAR and MERRA are used commercially today when performing estimations of the mean wind speed over a long period of time. This suggests that also the newer series ERA-Interim is a source that should be further investigated. As mentioned earlier, ERA-Interim is not available for commercial use for free and cannot be downloaded specified on the sigma levels. This gives an excellent opportunity for further investigation. Using the ERA-Interim series with exact sigma levels, which would be done when using the series commercially, might give even better results and is definitely worth investigating further.

REFERENCES

1. EC 2009, Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC, European parliament.
2. Swedish Energy Agency, 2011, Energiförsörjningen I Sverige år 2030.
3. Pettersson E., 2003, Matematisk Statistik, Matematik Litteratur I Göteborg.
4. University of Leicester, Online Statistics, 2011. Available at http://www.le.ac.uk/bl/gat/virtualfc/Stats/regression/regr1.html 2011-05-12.
5. Montgomery D., Peck E., Vining G., 2006, Introduction to linear regression analysis. Wiley Series in Probability and Statistics.
6. Bergman J., 2001. Siktanalys, Master Thesis at Luleå Tekniska Universitet.
7. Convective Development, 2011. Available at http://www.aprweather.com/pages/wind.htm 2011-06-13.
8. Liléo S., Petrik O., 2010. Investigation on the Use of NCEP/NCAR and NCEP/CFSR Reanalysis Data in Wind Resource Analysis.
9. Berrisford P., Dee D., Fielding K., Fuentes M., Kållberg P., Kobayashi S. and Uppala S., 2009, ERA Report Series, The ERA-Interim Archive, Version 1.0.
10. Saha S., et al, 2010, The NCEP Climate Forecast System Reanalysis. Available at journals.ametsoc.org/doi/pdf/10.1175/2010BAMS3001.1 2011-04-20.
11. Global Modelling and Assimilation Office, 2008. File specifications for MERRA Products.
12. Grib.us 2011. Available at http://www.grib.us/ 2011-06-28.
13. Nilsson E., Bergström H., Från mätt vind till vindklimat, 2009, Elforsk rapport 09:03.
14. EMD International A/S 2010, An Introduction to the MCP Facilities in WindPRO.