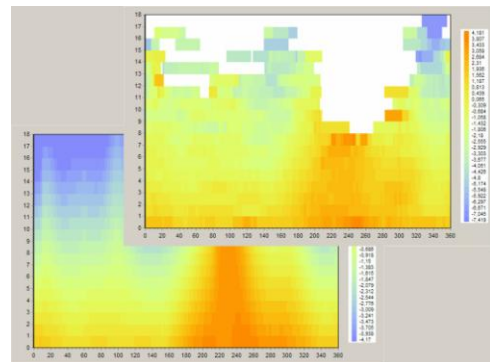
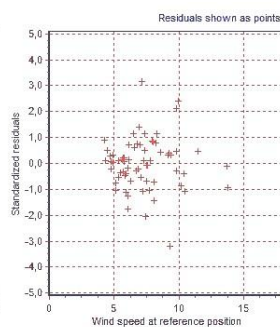
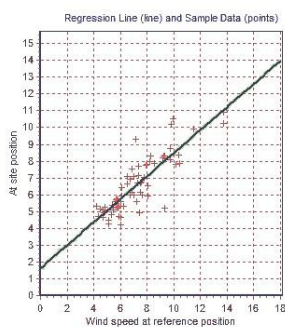
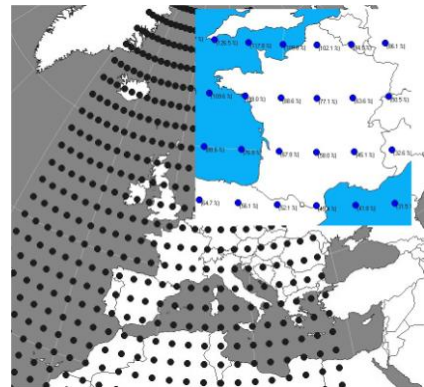


WindPRO / MCP

MCP - Measure-Correlate-Predict

An Introduction to the MCP Facilities in WindPRO



EMD International A/S

Niels Jernes Vej 10, DK-9220 Aalborg, Denmark

www.emd.dk or www.windpro.com

phone +45 9635 4444, fax. +45 9635 4446

e-mail: windpro@emd.dk



Published by:

EMD International A/S
Niels Jernesvej 10
9220 Aalborg Ø
Tel: +45 9635 4444
Fax: +45 9635 4446
Email: emd@emd.dk
web: www.emd.dk

Author

Morten Lybech Thøgersen, M.Sc.,

Co-authors

Per Nielsen, M.Sc.
Thomas Sørensen, M.Sc.
Lasse Ulrik Svenningsen, Ph.D.
Mads V. Sørensen, M.Sc.

Document history

18-08-05	1st draft – Morten Lybech Thøgersen, M.Sc., mt@emd.dk (NCAR/NCEP data extraction only)
20-11-05	2 nd draft – Added description of Regression MCP and added more lengthy introduction
28-11-05	3 rd draft – Matrix Method MCP description, Weibull Scale Description, Wind Index Description
29-01-07	1 st released version
14-06-07	Added description on R^2 (coefficient of determination) in the Regression MCP chapter
16-06-10	Updated to reflect the WindPRO 2.7 release. Removed section NCEP/NCAR data (now obsolete as this is now found in the online-services manual).
25-06-10	Added description of 'r' and 's' in MCP-Predict tab (Section 6-4)
28-01-13	Added description on comparing 'r' and 's' between different methods (Section 6-4-1)

Front cover

The front cover shows WindPRO logo. Upper right picture shows the availability of the NCEP/NCAR data in the European region. Lower left figure shows linear regression MCP analysis, with fit and residuals from the analysis. Lower right figure shows sample surface and model surface for the Matrix MCP service.

Copyright

This document is copyrighted EMD International A/S. Unauthorized use is prohibited.

The document may be used as reference manual only by registered users WindPRO or by users with a valid time limited registration key to the WindPRO software system. Other use requires a written permission from EMD International A/S.

Table of Contents

1. Introduction	1-1
About the Application of MCP	1-1
A Conceptual Model for the MCP-framework	1-2
The MCP facilities in WindPRO	1-3
2. Regression MCP Methods	2-1
Introduction	2-1
The Regression Model	2-2
Wind Speed and Wind Veer Regression Model	2-2
Estimation of the Regression Parameters	2-2
The Distribution of the Residuals, e	2-3
R^2 : Coefficient of Determination	2-5
Calculating the Long Term Corrected Data	2-6
3. Matrix Method MCP	3-1
Introduction	3-1
Matrix MCP modelling	3-1
Sample Data, Sample Distributions and Statistical Moments	3-2
Polynomial Model for the Statistical Moments	3-3
Parametric Model for the Joint Wind Speed-Up and Wind Veer Distribution	3-4
Calculating the Long Term Corrected Data	3-5
4. Wind Index MCP	4-1
Introduction	4-1
The power curve used in the Wind Index Calculation	4-1
Wind Speeds in the Wind Index Calculation	4-2
Calculating the Wind Index	4-2
Wind Index Correction	4-3
An Example	4-3
Validating the Wind Index Correlation	4-4
Calculating a LTC Wind Statistic based on a Wind Energy Index	4-5
5. Weibull Scale MCP	5-1
Introduction	5-1
The Weibull Scale Method	5-1
Doing the Weibull Scale MCP	5-1
6. Measure & Correlate: Evaluating the Data for the MCP-Analysis	6-1
Introduction	6-1
First Step	6-1
Suggested Evaluations of the Long-Term Reference Data	6-1
Evaluation of the Correlation	6-1
Considerations Concerning the Sample Averaging Period	6-3
7. References & Literature	7-1

1. Introduction

This document gives an introduction to the Measure-Correlate-Predict tools (MCP) that are included with WindPRO. The MCP module in WindPRO consists of the following sub-modules that were first introduced with the release of WindPRO 2.5 in 2005:

1. Linear Regression MCP
2. Matrix method MCP
3. Weibull Scale MCP
4. Wind Index MCP



Semi-offshore turbines at Frederikshavn, Denmark.

The Measure-Correlate-Predict toolbox in WindPRO (MCP) enables the user to calculate long term corrected wind data directly in WindPRO. The MCP module provides not only a direct access to different MCP-methods, but also provides reporting through overview-reports and detailed reports for each of the methods available.

Data that is used by the MCP-methods are contained in the WindPRO Meteo object(s). This object is the data container for wind data saved as either time series data, table data or Weibull distribution parameters. Most MCP-methods require two overlapping time series, each holding a concurrent time series for the reference position and site position respectively. The long-term reference data could be either Weibull data, table data or time series data.

The result from the MCP analysis is – typically – one or more new long term corrected Meteo-Object(s). This meteo-object is located at the exact same position as the one holding the site data – and the intention is, that this long term corrected object should be used in further analysis, e.g. in WAsP, PARK or other WindPRO modules.

Long term reference wind data is available directly from within WindPRO as a part of the WindPRO online services, see more in the WindPRO documentation available from the WindPRO knowledge base at <http://help.emd.dk/knowledgebase> [14]. Typically, the NCEP/NCAR reanalysis data are applied for this purposes, as well as any data from local SYNOP or METAR stations.

1.1 About the Application of MCP

MCP is the abbreviation for Measure-Correlate-Predict techniques, which is widely in use for establishing a long-term wind statistic using limited wind data from the current site and long-term data from a more-or-less nearby site.

The goal of any estimation of a long-term wind statistic is to establish a transfer model between the available short-term or long-term wind data and the long-term statistic on the prediction site. The transfer model can be grouped into (at least) four different types:

1. Physical models (e.g. CFD flow models)
2. Statistical models
3. Empirical models
4. Other (combinations of the above, e.g. like WAsP)

MCP-models may belong to any one of the categories or a combination hereof, indicating that the application of MCP-models has a very wide scope:

Some MCP-models operate on large timescales – like index correction methods, where monthly data typically are used. Other MCP-models tries to decode a one-to-one relationship between wind speeds and wind

1. Introduction to Measure-Correlate-Predict

directions on site and on the reference mast, calling for high quality measurements. In some situations, MCP-methods are applied in order to correct the lack of ability for a model to take long-term variation into account. This is the case when using local short-term site data in a WAsP analysis. Unfortunately, not all methods will perform equally well in all situations, calling for the user to get acquainted to the performance and limitations of the individual methods.

In general, the applied MCP method modifies one or more of the following descriptive data:

1. Wind energy index
2. WTG energy index (measured production from the WTGs)
3. Weibull A-parameter + Weibull k-parameter
4. Wind speed and possibly also wind direction

In WindPRO methods working on (1), (3) and (4) are implemented.

1.2 A Conceptual Model for the MCP-Framework

A conceptual model for the measure-correlate-predict model is shown in Figure 1 below. From this figure it is seen that three potential different measured datasets are input to the MCP-model:

- Reference data: Long data
- Reference data: Short term data (concurrent data)
- Site data: Short term data (concurrent data)

The long term data could be any type of distribution, e.g. Weibull data, table data (joint wind distribution) or time series data. The concurrent datasets are typically required to be time series data. Not only measurements on wind speeds and wind direction could be used as input to advanced MCP models but also temperature differences could prove useful. However none of the models in WindPRO currently uses other input than wind speed and wind direction.

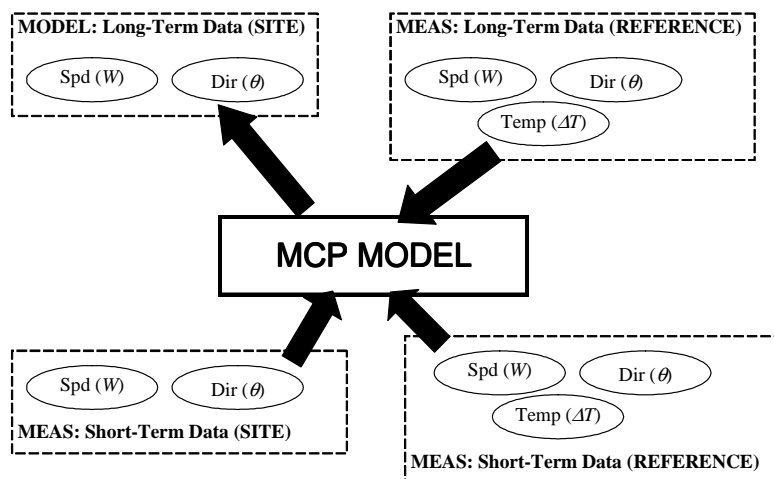


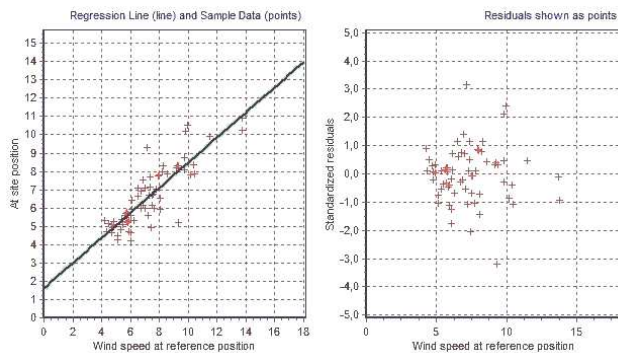
Figure 1: Conceptual Model for Measure-Correlate-Predict Models.

Note: When doing MCP calculations in WindPRO, the input data are always internally treated as three different input sources. However, in many cases the long term reference data and the concurrent reference data series may come from the same time series data source. In such a situation, WindPRO will automatically select the same series for both input sources.

1.3 The MCP facilities in WindPRO 2.5

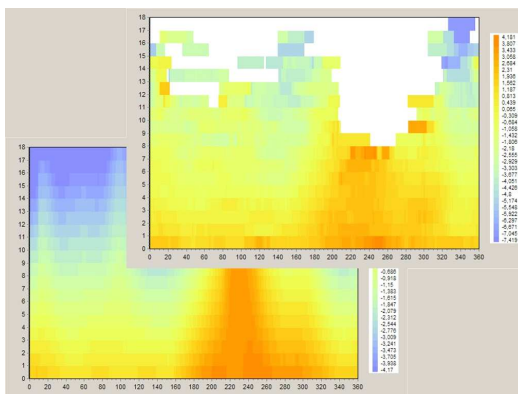
Below is a list of sub-sections briefly outlining the MCP facilities, which was introduced from WindPRO 2.5 and improved with later WindPRO releases.

1.3.1 Linear Regression MCP



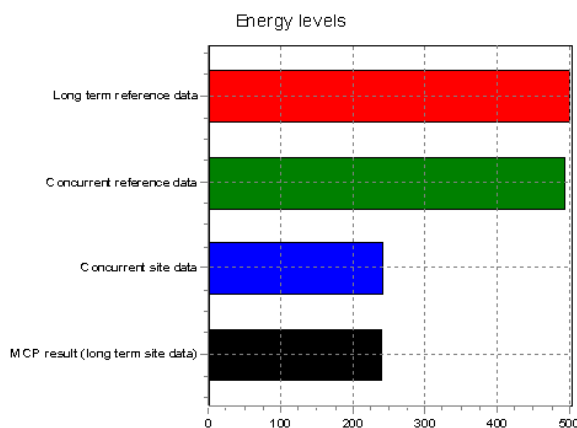
The (Linear) regression tool enables the user to inspect the fit directly through an animated graph. If the fit is not satisfactory, a wide range of parameters may be fine-tuned to provide a better fit. The regression tool is not limited to linear regression, but also higher order polynomials may be used in modelling wind speeds and wind veer.

1.3.2 Matrix Method MCP



The matrix method in WindPRO models the changes in wind speed and wind direction through a joint distribution fitted on the ‘matrix’ of wind speed bins and wind direction bins. The user may choose to either use polynomials fitted to the data statistics or – where appropriate - to use the measured samples directly when doing the matrix MCP.

1.3.3 Weibull Scale MCP



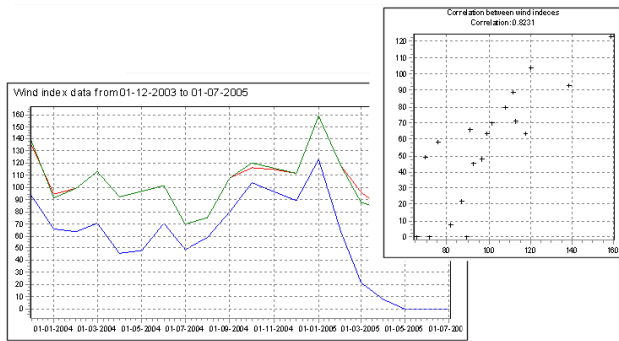
The Weibull Scale method is a very simple empirical method, which does its manipulation directly on the Weibull form and scale parameters (A,k) as well as on the frequency distribution.

The Weibull method has the advantage, that it will match the nature of the wind at most places, but beware that application of this method should be done with caution on locations with significant non-Weibull distributions.

1. Introduction to Measure-Correlate-Predict

1.3.4 Wind Index MCP

Period	Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Mean	55.4	94.4	63.5	46.4	26.4	24.0	35.1	24.5	58.7	79.8	104.0	-	90.9
Standard deviation	36.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	-	0.0
Minimum	0.1	65.8	63.4	22.0	7.5	0.1	0.1	0.1	58.7	79.8	104.0	-	88.9
Maximum	123.1	123.1	63.7	70.9	45.3	46.0	70.1	48.9	58.7	79.8	104.0	-	92.9
Number of samples	19	2	2	2	2	2	2	2	1	1	1	0	2



The index correlation method is a method typically making the MCP analysis by using monthly averages of the energy yield, thus disregarding the directional distribution of the winds. Even though this method may seem rather crude and primitive when comparing to other more advanced MCP methods, it has its advantages in stability and performance – even in the cases where other MCP methods seem to fail.

The Wind Index MCP method in WindPRO offers the opportunity to calculate the wind indexes using real power curves from the wind turbines included in the wind turbine catalogue in WindPRO. Also a generic power curve may be chosen.

2. Regression Measure-Correlate-Predict Methods

2.1 Introduction

The regression MCP method holds the traditional linear regression MCP analysis as well as other - more general - regression models using polynomials of other orders. Specializations of the polynomial-fitting methods are also included, provided in order to comply with methods used or suggested by other companies. One of the specializations is here an MCP method forcing the regression line through the origin (0,0). However, this option should be used only with caution as it typically provides a significantly poorer fit to the data than the methods where a non-zero intersection with the y-axis is allowed.



Wind Farm in Southern Europe

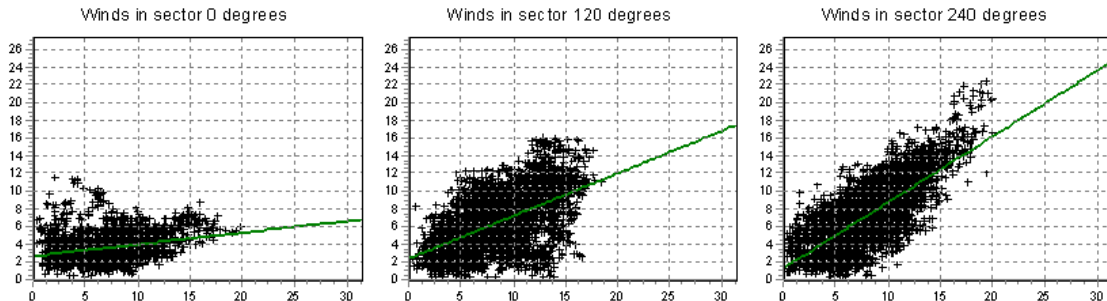


Figure 1: Linear fitting of wind speeds (x : winds at reference position, y : winds at site position)

The regression MCP methods in WindPRO are improved over a traditional linear regression analysis, as a model for the distribution of the residuals is also included. This model allows the regression MCP method to capture the energy content in the MCP corrected site wind distribution much better than regression models without this option. Experience has shown, as much as 10% energy can be erroneously lost in the long-term correction if the model is run without this option.

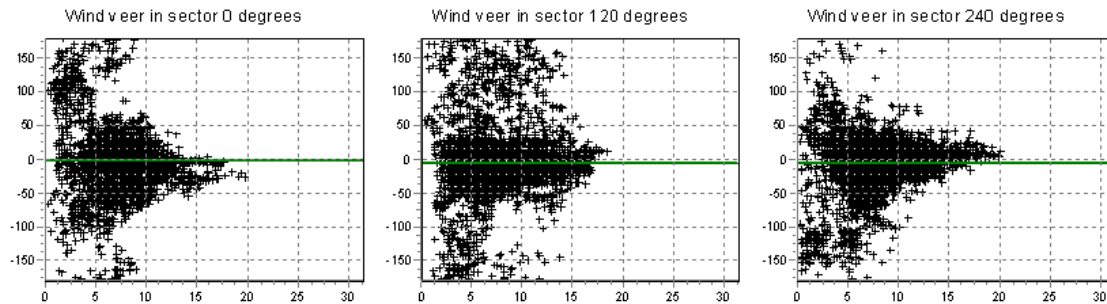


Figure 2: 'Constant' fitting of the wind veer (x : winds at reference position, y : veer from reference to site).

The Figure 1 and Figure 2 show examples on the fitting of wind speed and wind direction respectively. Note, that a very significant scatter is present. WindPRO models this scatter through the residuals in the linear model, which are included as zero-mean Gaussian random variables conditioned on the wind direction and also on wind speed. Especially for the wind speeds, it may be relevant to include a model of the scatter (residuals). This is due to the fact, that the measured points above the regression line represent significantly more energy than the points below the line. This is due to the third power relationship between wind speed and energy and the non-linearity of the WTG power curves. Failure in including the scatter will

2. Regression Measure-Correlate-Predict Methods

erroneously remove as much as 5-10% of the measured energy, though applying the scatter where a Gaussian model does not apply may add too much energy.

2.2 The Regression Model

Regression modeling, where only one independent (x) and one dependent (Y) variable is present, is based on the following equation:

$$Y = f(x) + e$$

where Y is the dependent variable
 x is the independent variable
 $f(x)$ is the regression model
 e is a random error (residual)

The regression model could be polynomials of any order or other models, but traditionally a linear model is assumed, as this model has been found to give reasonable fits for wind energy estimation. In the case of a regression MCP analysis, the independent variable could be the wind speed measured at the reference position. The dependent variable (Y) is then the wind speed at the local WTG site position.

3.2.1 Wind Speed and Wind Veer Regression Model

In WindPRO, the regression models are made as models for wind speed and wind veer. The wind veer is taken as the difference in wind direction at the site mast and the wind direction at the reference mast. All analyses are made conditioned on the wind direction on the reference mast, i.e. the fitted regression parameters are functions of reference site wind direction. The binning is made with one-degree resolution in the angular variable and a 1.0 m/s resolution in the wind speed variable. However, data are loaded into these bins from larger window - typically a 30 degree window is used – otherwise some bins may suffer from missing data.

The following regression models are currently available in WindPRO:

<i>Model description</i>	<i>Model equation, $f(x)$</i>
No model	$Y = x$
Constant	$Y = \beta_0$
Linear - 1 st order polynomial	$Y = \beta_1 \cdot x + \beta_0$
Linear regression through (0,0)	$Y = \beta_1 \cdot x$
Second order polynomial	$Y = \beta_2 \cdot x^2 + \beta_1 \cdot x + \beta_0$
Second order polynomial through (0,0)	$Y = \beta_2 \cdot x^2 + \beta_1 \cdot x$

The regression model named ‘No model’ actually just transfers the measured data from the reference position to the site position, i.e. if no model is chosen for the wind veer model, then the directional distribution for the site will be the same as for the reference site. Please note that the modelling of the residuals should be disabled if such a transfer is desired.

2.2.2 Estimation of the Regression Parameters

The regression parameters are estimated through a least squares algorithm. The least squares method calculates the summed squared difference between the estimated response and the actual measured response:

2. Regression Measure-Correlate-Predict Methods

$$SS_R = \sum_{i=1}^n [Y_i - f(x_i)]^2$$

The least squares method then chooses the set of regression parameters in $f(x)$, that minimizes the squared sum, SS_R . Note, that for the linear model, analytical expressions are available for calculating the optimal parameters enabling a quite fast calculation with this model, see S.M. Ross [11]. However for polynomials of other orders, a general optimization algorithm is used instead, resulting in longer calculation times. The optimization algorithm is the Amoeba method, e.g. found in Press et al [13].

2.2.3 The Distribution of the Residuals, e

As expressed earlier, the regression model yields,

$$Y = f(x) + e$$

where e is a random error.

The distributions of the random errors may, according to Ross [11], reasonably be assumed to follow a zero mean Gaussian distribution, $e \sim N(0, \sigma)$. However, in many cases, data shows that a better assumption is to model the residuals conditioned on both the wind direction and the mean wind speed. This model is also available in WindPRO, where the residuals are still modelled as Gaussian, but with the mean and standard deviation conditioned on both wind speed and wind direction. Note that caution is needed when choosing a residual model: such a model is needed in the MCP-modelling in order to give the right energy levels in the new MCP-corrected time series. Currently (2010), the default model in WindPRO 2.7 is to model residuals for wind speeds conditioned on both wind direction and wind speed, see the examples in Figure 5.

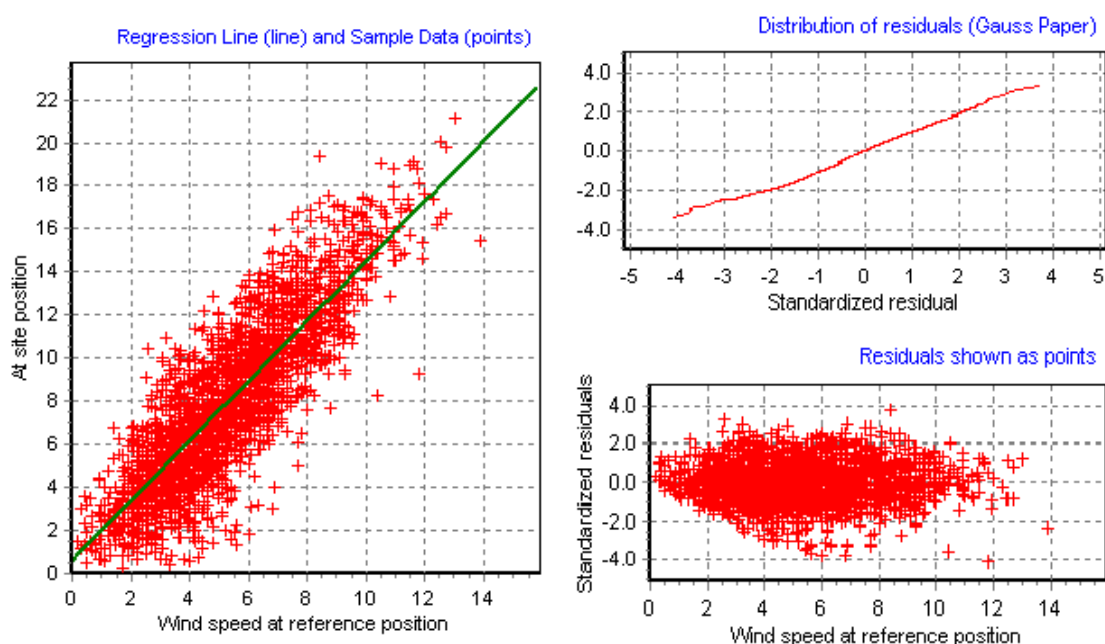


Figure 3: Wind Speed Fit (Left) with the Distribution of the Residuals (plot captured from WindPRO). To the right is shown the modelling of the mean and standard deviation of the residuals.

When comparing the distribution of the residuals, it is convenient to consider the Standardized Residuals. This is a normalization procedure made to bring the residuals on a common scale:

2. Regression Measure-Correlate-Predict Methods

$$\frac{Y_i - f(x_i)}{\sqrt{SS_R / (n - 2)}}$$

where

Y_i is the measured response for the i -th sample (x_i)

$f(x_i)$ is the model prediction for the x_i sample

n is the total number of samples

SS_R is the sum of squares of the residuals

Inspecting the Figure 3 where the wind speeds is plotted, it is seen, that the distribution of the residuals follows a Gaussian distribution quite well. This is easiest seen from the top-right figure showing the distribution of the residuals plotted on Gauss paper, i.e. a straight line in this graph and the distribution is a Gaussian distribution. The lower right figure shows, that the residuals seem to be reasonably independent on the reference wind speed (the standard deviation seems quite constant regardless of the reference wind speed).

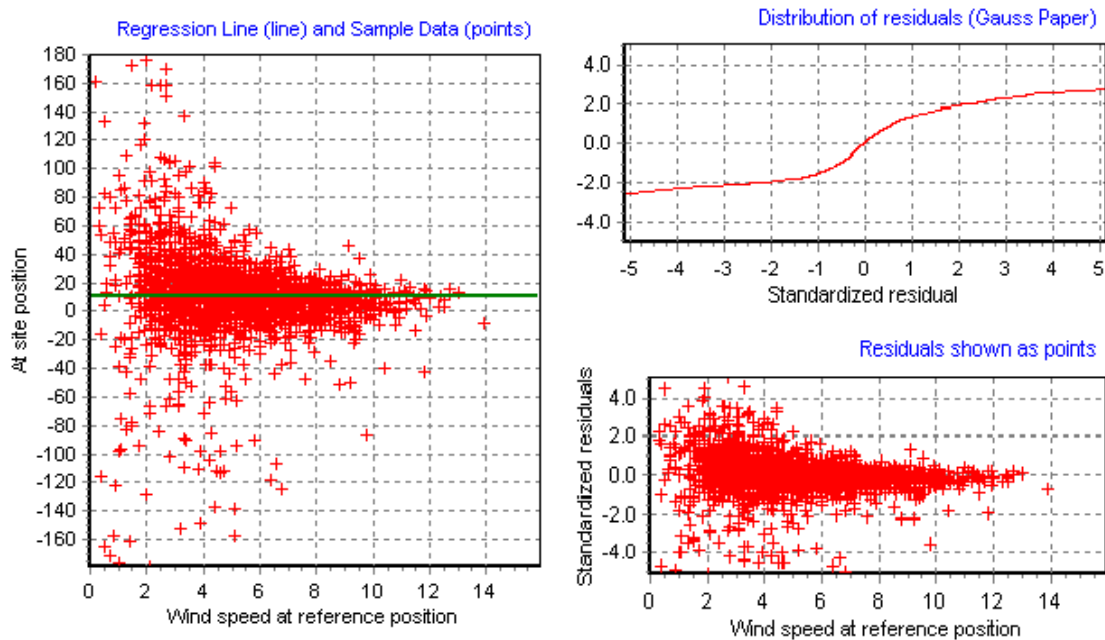


Figure 4: Wind Veer Fit with the Distribution of the Residuals.

The Figure 4 shows the fit to the wind veer distribution. By inspecting the Gaussian paper plot it is seen that a Gaussian distribution (straight line) is a quite raw approximation in this case. Also, when inspecting the lower right graph with the plot of the residuals, it is also noticed that the wind veer values at the low reference site wind speeds has a significant higher standard deviation than the ones at high wind speeds. This suggests that a revised model should be refined to include a residual random variable with standard deviation conditioned not only on the reference site wind direction but also the reference site wind speed. This model is actually also included with WindPRO as an 'Advanced Gaussian' model. An example of this modelling is shown in the Figure 5 where the blue line is the conditioned mean value and the green lines are the mean plus/minus one standard deviation. As an alternative, one could investigate the modelling options in the Matrix Method MCP; these are also able to catch such effects.

2. Regression Measure-Correlate-Predict Methods

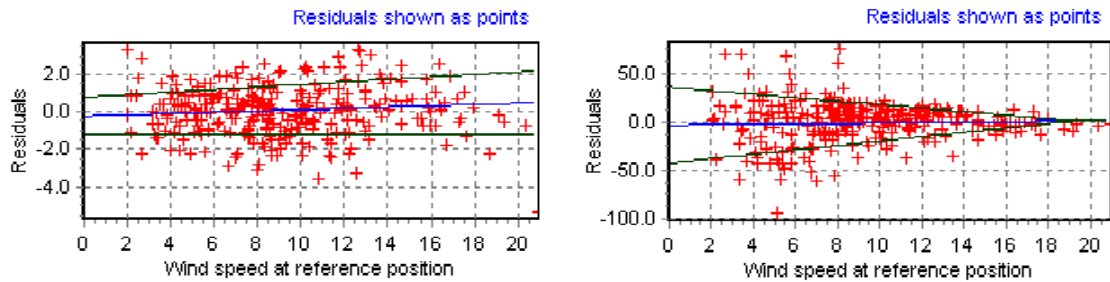


Figure 5: Example of the Advanced Gaussian Residual Model. Left: Residuals of Wind Speed. Right: Residuals of Wind Veer.

2.2.4 R^2 : Coefficient of Determination

The coefficient of determination, R^2 , is a statistic used for interpreting the proportion of variability in the data set that is accounted for by the selected statistical model. The R^2 gives some information about the goodness of fit of the chosen model. The R^2 is defined as

$$R^2 = 1 - \frac{SS_E}{SS_T}$$

$$SS_T = \sum_i (Y_i - \bar{Y})^2 \quad SS_E = \sum_i (Y_i - \hat{Y}_i)^2$$

where SS_T is the total sum of squares
 SS_E is the sum of squared errors
 Y_i is the sample considered
 \bar{Y} is the mean of the samples
 \hat{Y}_i is the response value for the sample i - calculated using the chosen regression model

When considering a linear regression model, $Y=A \cdot x+B+e$, then the R^2 value is actually an indication of whether the linear model is an improvement in the fit rather than the model $Y=A+e$ (i.e. using a constant model only). Again, e is a random error (residual). Please note, that by definition, the R^2 always relates the chosen model to the constant model.

Typically, the value of R^2 is in the range from 0 to 1. However - in the case where a regression model with a fixed intercept is used (e.g. when forcing a line through (0,0)) - then the R^2 value may in fact become negative. This happens when SSE is larger than SST. This is only an indication that the chosen model does not fit the data as well as the 'constant' model.

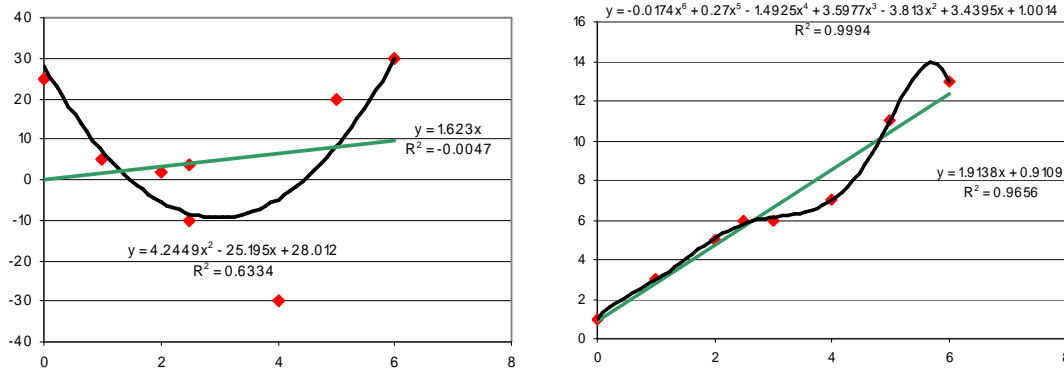
The R^2 values are often assessed like this:

- $R^2 = 1.00$: Indicates that all response variation is explained by model chosen (because of a perfect fit, so $SS_E=0$)
- $R^2 = 0.00$: The model chosen performs like the 'Constant' model
- $R^2 = 0.75$: About 75% of the variation in the response variable is explained by the model. The 25% remaining is explained by inherent variability, unknown or hidden variables.

The value $R = \sqrt{R^2}$ is called the index of fit. It is commonly used as an indication of how the model fits the data. The index of fit is also named the sample correlation coefficient, because - in the case of a linear model - it is a natural estimator of the correlation coefficient, ρ . If not using the linear model, then this relationship does not hold.

2. Regression Measure-Correlate-Predict Methods

In the two graphs below different models have been fitted to sample data, using the Microsoft Excel spreadsheet tool. In the graph to the left, a linear curve with intercept in (0,0) is fitted. Note that the R^2 is negative, implying that a constant model would be a better choice. Also shown is a 2 order polynomial fit with $R^2 = 0.63$. In the graph to the right, a linear regression fit with a high R^2 is shown ($R^2 = 0.97$). Also shown is a 6-th order polynomial fit with $R^2 = 0.999$. This high R^2 does not necessarily mean that the 6-th order polynomial is a (physical) better model, because the variability in the response may actually be caused by some random component inherent in the sample data.



2.3 Calculating the Long Term Corrected Data

The long term corrected meteorological data are calculated using Bootstrap and Monte-Carlo simulation techniques, i.e. probabilistic methods enabling generation of the long term corrected wind distribution through an ‘artificial’ time series. For details on Monte-Carlo simulation, see [15]. For an introduction to the Bootstrap, see [12].

An outline of the calculation and simulation procedure is as follows:

1. Fit the wind speed model with the corresponding residuals (conditioned on the reference wind direction, i.e. actually in 360 bins from 0 to 359 degrees).
2. Fit the wind veer model with the corresponding residuals (also conditioned on the reference wind direction, binned from 0 to 359 degrees)
3. Given the long term reference wind distribution, simulate out a random sample of wind direction and wind speed, (W_{ref}, θ_{ref}) .
4. Using the wind speed model in (1), calculate a sample of the on site wind speed data: $W_{site} = f(W_{ref}) + e$. Note that f is the regression model and e is a realization of the residual random variable.
5. Using the wind veer model in (2), calculate a sample of the on site wind direction: $\theta_{site} = \theta_{ref} + f(\theta_{ref}) + e$. Note that f is the (wind veer) regression model and e is a realization of the residual random variable. Also note, that the mathematical formulation is slightly different from the wind speed model, as we are here modelling the wind veer and thus need to add the reference wind direction.
6. Repeat 2-5 until the number of samples is ‘sufficiently’ large. This is typically the case when the number of ‘artificial’ samples has reached the number of samples expected in the long term reference period. If a time series is available for the entire reference period it is preferable to pick the actual measurements sequentially from the reference in (3). The artificial time series is then complete when all reference measurements have been transformed.
7. From the sample distribution generate table data and fit Weibull table data.

2. Regression Measure-Correlate-Predict Methods

The simulation of samples from the long term reference distribution may (step 3) may be done using data from either the fitted Weibull data, the distribution table or from a long term reference time series. It is recommended to use the data from the distribution table or time series, as the Weibull data are subjected to modelling errors.

When modelling from Weibull or table data, then no correlation structure in the generated time series is preserved. Thus only the resulting distribution tables should be compared to the actual measurements and used in further analysis.

If however the reference data input is a time series a correlation between measured site data and predicted site data is a measure of the success of the prediction.

3. Matrix Method MCP

3.1 Introduction

The matrix method in WindPRO models the changes in wind speed (speed-up) and wind direction (wind veer) through a joint distribution fitted on the ‘matrix’ of wind speed bins and wind direction bins.

The parallel period of measured wind data is used to calculate the set of the transfer functions, used for transferring wind speeds and wind directions from the reference site to the site position. Since real measurements will suffer from data missing in bins in the dataset, this method needs a way to fill out missing input bins. In WindPRO, polynomials are fitted to the statistics of the sample data enabling this interpolation/extrapolation. The user may choose to either use polynomials fitted to the data statistics or, where appropriate, to use the measured raw samples directly when doing the matrix MCP.



Wind Turbines at Göteborg, Sweden

A basic assumption of the matrix method is that the long term site data (wind speed and direction) can be expressed through the simultaneous measurements of on-site data and reference site data. Actually this relationship is basically modeled through a joint distribution between the two variables wind speed-up and wind veer. How this joint distribution is modeled should actually depend on the data in question, but experience using the WindPRO MCP, suggests that a combination of binned sample distributions and a modeled joint Gaussian distribution seem to work quite well.

The transfer model, given as a conditional distribution, is actually the key distribution in the generalized matrix method. The distribution gives the relationship between the site wind climate and the reference wind climate. *When applying the matrix method this conditional distribution is stipulated to hold regardless of the time frame considered.*

3.2 Matrix MCP Modelling

The Matrix MCP in WindPRO models the wind speed-up and the wind veer as functions of the wind speed and wind direction on the reference site, see Figure 1 and Figure 2. Two modelling options are available in WindPRO, either to use the measured samples themselves (through a resampling technique) or alternatively to use the fitted polynomial model (actually through a joint Gaussian random variable). If no sample data is available in a particular bin, then the polynomial model is always used.

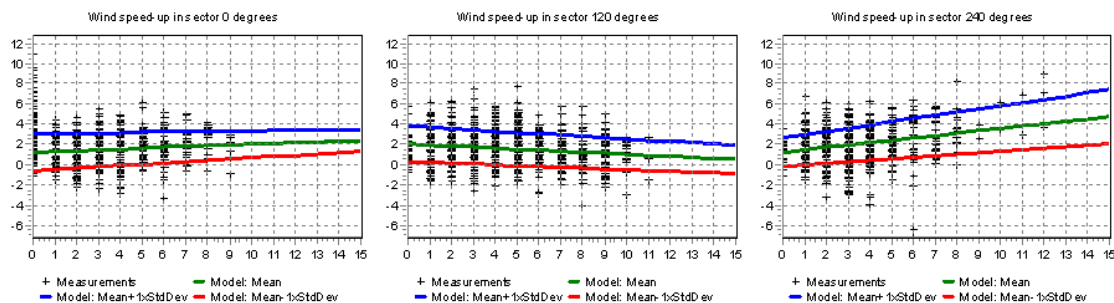


Figure 1: Sample data and first order model for the wind speed-up (x : wind at reference, y =speed-up).

3. Matrix Method MCP

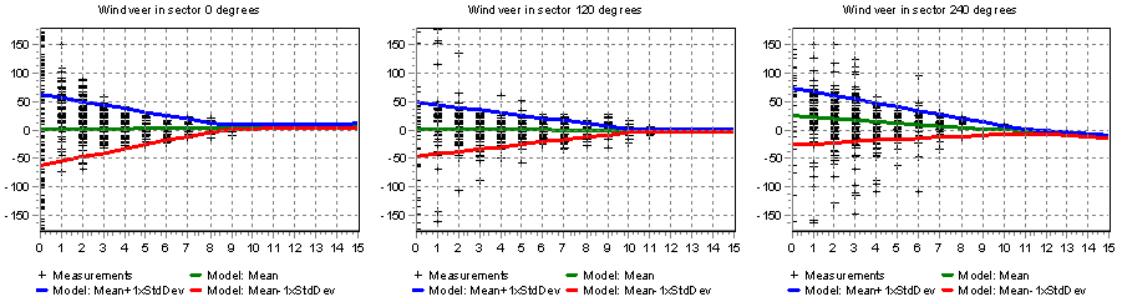


Figure 2: Sample data and first order model for the wind veer (x : wind at reference, y =wind veer).

The model is based on the joint distribution of the measured wind speed-ups and wind veers. Thus, for each measured sample we have to calculate/measure pairs of the two quantities (a pair is data with identical timestamps):

$$\Delta u = u_{site} - u_{reference}$$

$$\Delta \theta = \theta_{site} - \theta_{reference}$$

where Δu is the wind speed-up
 u_{site} the wind speed at the site position
 $u_{reference}$ the wind speed at the reference position
 $\Delta \theta$ is the wind veer
 θ_{site} the wind veer at the site position
 $\theta_{reference}$ the wind veer at the reference position

The joint distribution of $f_{\Delta u, \Delta \theta}$ is then modelled conditioned on the wind speed and the wind direction on the reference site. This joint distribution is represented as either through the samples (bootstrap model) or through a joint Gaussian distribution. In the case of the joint Gaussian distribution, the distribution parameters – mean, standard deviation and correlation - are modelled through polynomials of any order.

3.2.1 Sample Data, Sample Distributions and Statistical Moments

When the data has been measured and a match between the short-term site data and the short-term reference data has been established, then the samples are sorted into bins with the resolution 1 m/s and 1 degree. Since a 1-degree angular resolution is too small in most cases, it is possible to feed in data from a larger window, typically pre-set to around 30 degrees. The result from this binning is a set of joint *sample distributions* of wind veer and wind speed-up. Since the data are binned with wind speed and wind direction, these sample distributions are said to be conditioned on the mean wind speed at the reference position and the wind direction on the reference position. The sample distributions calculated are used directly in a bootstrapping technique (see Efron & Tibshirani [1]) when doing the Matrix MCP calculation.

Based on the sample distributions, the following sample statistics are calculated for the wind veer and the wind speed:

- Mean value
- Standard deviation
- Skewness
- Kurtosis
- Correlation

An example is shown in Figure 3 (left column).

3. Matrix Method MCP

3.2.2 Polynomial Model of the Statistical Moments

In order to enable interpolations and extrapolations into bins where no data are present, we choose to parameterize a model fitted to the sample distribution statistics. This parametric distribution is represented by the two first statistical moments and the correlation, and it is assumed that a joint Gaussian distribution is a reasonable distribution assumption (see the section 4.2.3 below).

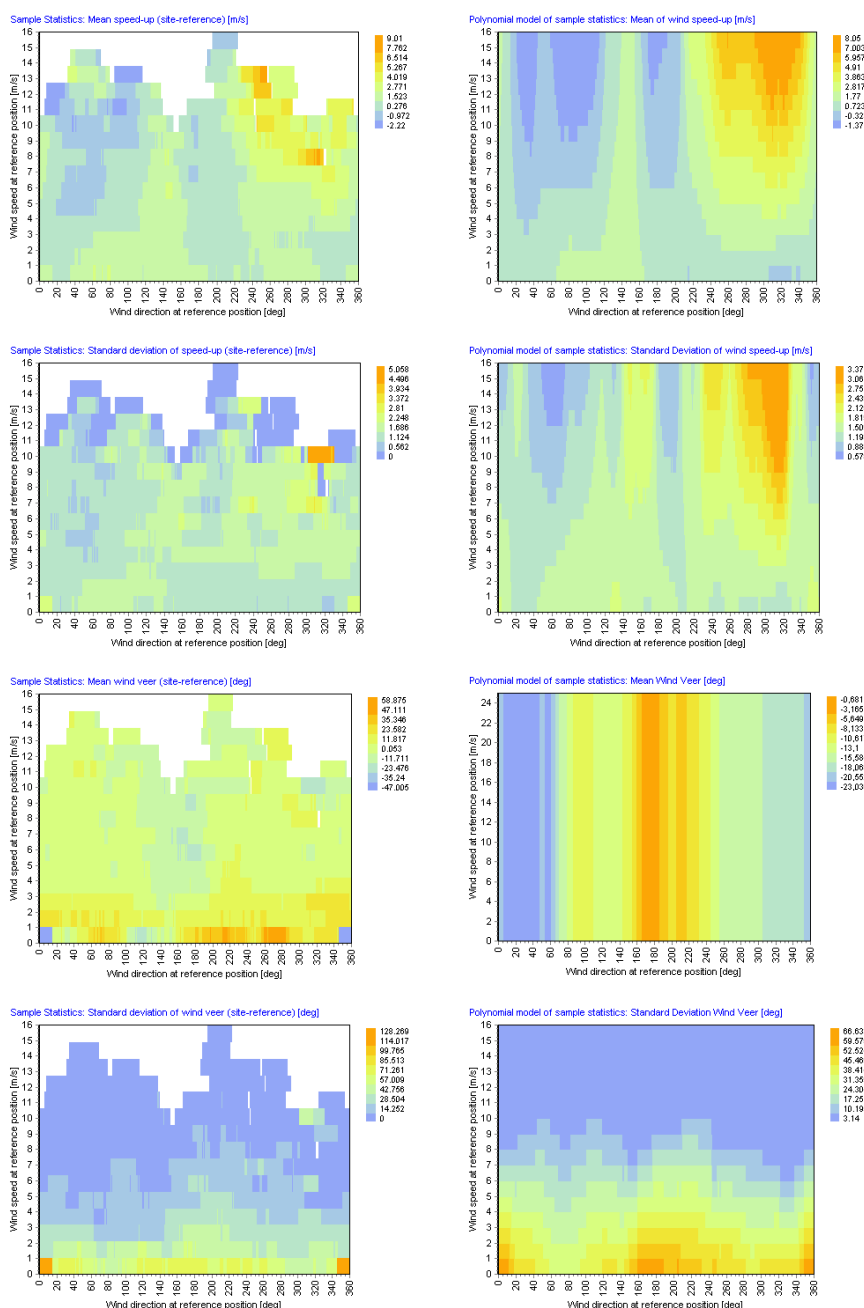


Figure 3: Left: Sample Statistics, Right: Model of the Statistics Moments.

The mean, standard deviation and correlation are now modelled as 'slices' of polynomial surfaces:

3. Matrix Method MCP

$$P | (U_{reference}, \theta_{reference}) = \sum_{i=0}^n a_i (\theta_{reference}) U_{reference}^i$$

where P denotes the sample statistical moment (or correlation) considered
 n is the order of the polynomial
 a_i is the polynomial coefficients (which are also functions of $\theta_{reference}$)

An example of this modeling is shown in Figure 3 (right side). It is seen, that the surface model seems to capture the trends in the sample data quite well (left side of the same figure). In this case 1st order polynomials are used in all cases except Mean wind veer where a 0th order polynomial is used.

3.2.3 Parametric Model for the Joint Wind Speed-up and Wind Veer Distribution

As mentioned earlier, it is assumed that the joint distribution of wind veer and wind speed is adequately modelled by a joint Gaussian distribution. This bivariate probability density function is given by:

$$f_{\Delta u, \Delta \theta}(\Delta u, \Delta \theta) = \frac{1}{2\pi\sigma_{\Delta u}\sigma_{\Delta\theta}\sqrt{1-\rho^2}} \cdot \exp \left[-\frac{(\Delta u - \mu_{\Delta u})^2 / \sigma_{\Delta u}^2 - 2\rho(\Delta u - \mu_{\Delta u})(\Delta \theta - \mu_{\Delta \theta}) / \sigma_{\Delta u}\sigma_{\Delta\theta} + (\Delta \theta - \mu_{\Delta \theta})^2 / \sigma_{\Delta\theta}^2}{2 \cdot \sqrt{1-\rho^2}} \right]$$

In the Figure 4 below, two examples of the joint wind veer and speed-up distribution are shown. This particular example was calculated using the following mean, standard deviation and correlation:

Wind speed-up: $(\mu_{\Delta u}, \sigma_{\Delta u}) = (0.3 \text{ m/s}, 0.5 \text{ m/s})$
 Wind veer: $(\mu_{\Delta \theta}, \sigma_{\Delta \theta}) = (5.0 \text{ deg}, 7.0 \text{ deg})$
 Correlation $\rho_{\Delta u \Delta \theta} = 0.1$ (left figure), $\rho_{\Delta u \Delta \theta} = 0.9$ (right figure)

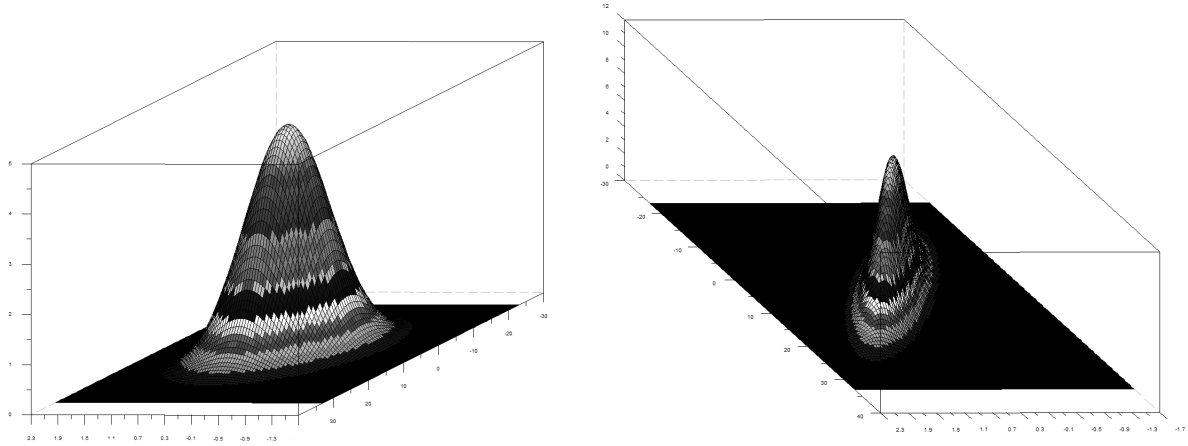


Figure 4: Bivariate Gaussian Distribution: Example of the Joint Speed-Up and Wind Veer Distribution. The parametric model for the veer and speed-up distribution is typically only used in the cases where only few samples are available in the bin. In that case, the polynomials mentioned in Section 4.2.2 are used to find the distribution parameters. Then a realization of the wind veer and the wind speed-up are made using the appropriate techniques.

3.3 Calculating the Long Term Corrected Data

As in the case of regression MCP (see Chapter 3), the long term corrected meteorological data are calculated using Bootstrap and Monte-Carlo simulation techniques, i.e. probabilistic methods enabling generation of the long term corrected wind distribution through an 'artificial' time series. For details on Monte-Carlo simulation, see [15]. For an introduction to the Bootstrap, see [12].

An outline of the Matrix MCP calculation and simulation procedure is as follows:

1. In the concurrent period, the wind veer and wind speedup are calculated as the site data minus reference data.
2. All sample sets (ΔU , $\Delta\theta$) are sorted/grouped into the appropriate bins.
3. From the binned sample distributions, the statistical moments are now calculated (mean, standard deviation and correlation).
4. The polynomial surfaces used to describe the statistical moments are fitted.
5. Given the long term reference wind distribution a random sample of wind direction and wind speed, (W_{ref} , θ_{ref}) is simulated.
6. Knowing the wind speed and wind direction on the reference site, the appropriate bin is now found. Using the data connected to this bin, a realization of the wind veer and wind distribution is now simulated from either the real sample data (by a bootstrapping technique) or using the modelled polynomials and the joint Gaussian model (by a Monte Carlo simulation technique).
7. Repeat 5-6 until the number of samples is 'sufficiently' large. This is typically the case when the number of 'artificial' samples has reached the number of samples expected in the long term reference period. If a time series is available for the entire reference period it is preferable to pick the actual measurements sequentially from the reference in (5). The artificial time series is then complete when all reference measurements have been transformed.
8. From the sample distribution generate table data and fit Weibull table data.

The simulation of samples from the long term reference distribution (step 3) may be done using data from either the fitted Weibull data, the distribution table or from a long term reference time series. It is recommended to use the data from the distribution table or time series, as the Weibull data are subjected to modelling bias (as the Weibull distribution fit used is an 'Energy' fit).

When modelling from Weibull or table data, then no correlation structure in the generated time series is preserved. Thus only the resulting distribution tables should be compared to the actual measurements and used in further analysis.

If however the reference data input is a time series a correlation between measured site data and predicted site data is a measure of the success of the prediction.

4. Wind Index MCP

4.1 Introduction

The wind energy index of a particular period is defined as ratio between the energy in the considered period (such as a particular month) and the average (monthly) energy taken over a longer reference period. For instance: a monthly wind index in January 2010 of 115 means that this particular month has a 15% higher wind energy than the average month in the reference period – which is typically taken as 20 or 30 years. The formulation used in WindPRO for the Wind Index is:

$$I = \left[\frac{\text{Average power output in some period, } T}{\text{Average power output in reference period, } T_{ref}} \right] \cdot 100$$



Brokilde Wind Farm

The wind energy index could be derived from wind data (as in the WindPRO MCP module) or as production indexes (derived from databases on wind turbine performance (such as the Danish Wind Index available from www.vindstat.dk).

Thus, the index correlation method is a method for making the MCP analysis by using averages of the energy yield (typically monthly averages), thus disregarding the directional distribution of the winds. Even though this method may seem rather crude when comparing to other more advanced MCP methods, which takes the wind veer into account; this method has its advantages in stability and performance as it may even succeed in the cases where other MCP methods seem to fail.

This is due to the fact, that the wind indexes are related directly to WTG energy yield. The Wind Index MCP method in WindPRO offers the opportunity to calculate the wind indexes using real power curves from the wind turbines included in the wind turbine catalogue in WindPRO. Also a generic power curve based on a truncated squared wind speed approach may be chosen. When the wind indexes has been calculated, the MCP correction is done on the estimated WTG energy yield, i.e. by multiplying the production estimated with a correction factor based on the difference in the wind index from the short term site data to the long term site data estimate.

4.2 The Power Curve used in the Wind Index Calculation

The energy level in the wind is proportional to the third power of the wind speed. However, since the power curve of a WTG is a non-linear function of the wind speed, then the wind index is typically calculated as either

1. Through a generic power curve, e.g. the square of the wind speed for wind below the stall onset and a constant above the stall onset:

$$P(u) = \begin{cases} u^2 & \text{for } u < u_{stall} \\ u_{stall}^2 & \text{for } u \geq u_{stall} \end{cases}$$

2. From a real power curve, $P(u)$, see an example from a Vestas V44 in the Figure 1.

The WindPRO Wind Index MCP tool offers a possibility to calculate the wind index using both types of power curve methods.

4. Wind Index MCP

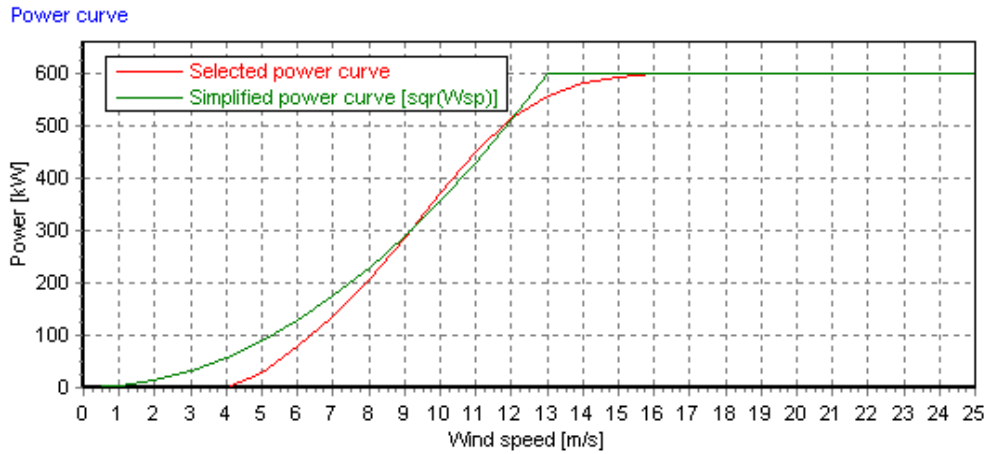


Figure 1: Power Curve from Vestas V44 and Simple Power Curve (scaled to fit the maximum power of the Vestas WTG).

4.3 Wind Speeds in the Wind Index Calculation

In order for the power output - calculated for local site and reference site - to be comparable; they must be based on a similar mean wind speed. This is done by assuming a sector uniform shear. This shear is then applied so that both concurrent mean when speeds are set to a fixed user-inferred wind speed, typically the expected mean wind speed at hub height.

The individual wind speed measurements are thus multiplied with the relevant factor. Both full time series wind speeds will be adjusted with the same ratio as was applied to the respective concurrent time series. The argument for this operation is that the variations in wind speed will only be interpreted correctly in terms of wind energy if a comparable section of the power curve is considered.

4.3 Calculating the Wind Index

The average power output in some period, W_T , is calculated using the modelled or measured power curve, $P(u)$:

$$W_T = \frac{\sum_{i=1}^N P(u_i)}{N}$$

Where N is the number of measured wind speeds within the period considered
 u_i is the i -th wind speed measurement (typically 10 minute mean wind speeds)

In WindPRO Wind Index MCP four different average power outputs are calculated or modelled. These are:

- W_{Rf} Average power output the full reference period of the reference wind dataset
- W_{Rc} Average power output for the concurrent part of the reference wind data set
- W_{Sf} Average power output for the full reference period of the on-site wind dataset
- W_{Sc} Average power output for the concurrent part of the on-site wind dataset

4. Wind Index MCP

The power output for the full period on the reference site, W_{Rf} is set to index =100, unless specifically stated otherwise, and the ratio in power output between W_{Rf} and W_{Rc} then gives the index of the concurrent period so that

$$I_{Rf} = 100$$
$$I_{Rc} = \frac{W_{Rc} \cdot 100}{W_{Rf}}$$

Where I is the wind index of the relevant period.

The assumption is now that the index for the concurrent period at the reference site is the same as the index for the concurrent period at the on-site, thus

$$I_{Sc} = I_{Rc}$$

That this is a valid assumption needs to be validated through correlation, see next section on Wind Index Correlation. Now knowing the index of the concurrent period on the site means that it possible to find the expected average power output for the entire site measurement period. This is done simply by multiplying the average power output of the concurrent period for the on-site measurements with the wind index. The relation is:

$$I_{Rc} = \frac{W_{Sc} \cdot 100}{W_{Sf}}$$
$$W_{Sf} = (W_{Sc} \cdot 100) / I_{Rc}$$

This is then the wind index for the measurement period. Please note that in this method the entire measured site data set is retained with original frequency and period length.

4.4 Wind Index Correction

When the expected long-term average power for measured site data has been found, W_{sf} it may be converted to a correction factor that may be used on other calculations, such as the energy yield based on a wind statistic from the local meteorological mast.

$$W_{sf} = C_{sf} \cdot W_{sc}$$
$$C_{sf} = \frac{W_{sf}}{W_{sc}} = \frac{100}{I_{Rc}}$$

Where C_{sf} is the correction factor for the full site data set.

This correction factor needs to be applied to the final result of the energy calculation. If a wind statistic is created in relation to the long term correction calculation the correction factor is automatically embedded in the resulting wind statistic and will be applied whenever the wind statistic is used. The term used for the correction factor in WindPRO is the *Regional Correction Factor, RCF*.

4.5 An Example

On a particular site, the following average power outputs have been measured (or calculated via WindPRO):

4. Wind Index MCP

Reference site: $W_{Rfs} = 450 \text{ kW}$; $W_{Rc} = 350 \text{ kW}$
 On-site: $W_{Sc} = 200 \text{ kW}$

The wind index for the concurrent period on the reference site is calculated below to 77.8. Thus, this period of consideration holds only 77.8% of the energy as the full long term reference period.

$$I_{Rc} = \frac{350 \cdot 100}{450} = 77.8$$

Now the Wind Index Correction factor can be calculated 1.285, i.e. the expected long term yield is 28.5% higher than the local measurements show:

$$C_{sf} = \frac{100}{77.8} = 1.285$$

4.6 Validating the Wind Index Correlation

In order to make the crucial validation that the wind index for the concurrent period of the reference station is identical to the index of the concurrent period of the site; then it is necessary to establish whether there is a fair correlation between the two data sets. This can be established by establishing a set of monthly wind index using the same method as described above.

Wind indexes are calculated for each month during the concurrent period comparing the monthly average power output to that of the entire concurrent period. This is done for both reference and site data.

When the monthly wind indexes are plotted against each other like shown in Figure 2 and Figure 3, then the course of the indexes must be similar. If they are divergent it is a sign that the climates at the two locations are different and the assumption of similar index for the same period grows weak. Similarly does the assumption that the reference data are representative for long term conditions at the site. The correlation is measured in either correlation value of the monthly indexes or as a standard error of the difference between the index graphs.

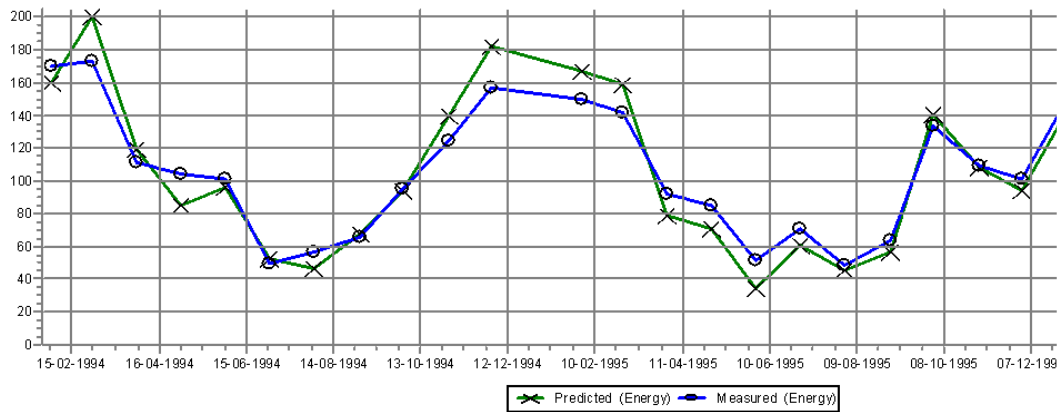


Figure 2: Monthly wind index for reference (green) and concurrent data sets (blue). The correlation is a test of how well the two graphs fit together.

4. Wind Index MCP

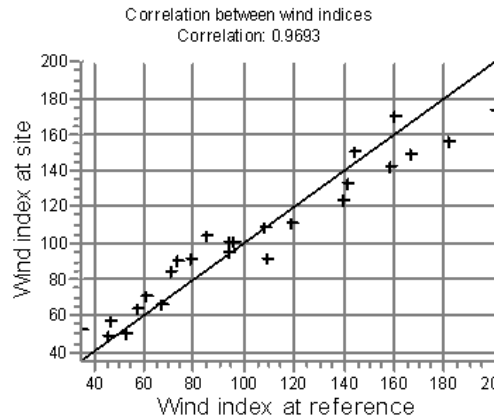


Figure 3: Correlation between wind indexes on the site and the reference station.

4.7 Calculating a LTC Wind Statistic based on a Wind Energy Index

All MCP methods in WindPRO have the opportunity to output a wind statistic where the wind speeds are corrected into the proper long term corrected (LTC) reference level. This MCP corrected wind statistic is primarily used for energy yield estimations such as with a WindPRO/PARK calculation.

While the calculation of a long term corrected wind statistic is fairly straightforward for MCP-methods based on modelling of the wind speeds; then this is not the case when a wind energy index is involved. Since the Wind Energy Index relates to the power yield from an actual turbine, then this power curve must also be taken into consideration when using the wind index with the wind statistic. The correction method as implemented in WindPRO is based on a numerical optimization scheme, where the Weibull form and scale parameters are fitted to match the increase in energy as calculated from the Wind Index MCP method. An outline of the procedure is given in the following pseudo-code:

```
For each pair of A and k in the Wind Statistic do
  1. Calculate original AEP based on original A and k as well as turbine power curve
  2. Calculate LTC-AEP based on original AEP multiplied with Wind Energy Index
  3. Find the LTC Weibull parameters by numerical optimization to match LTC-AEP in (2)
  4. Save the LTC Weibull parameters in new LTC corrected wind statistic
End For
```

5. Weibull Scale MCP

5.1 Introduction

The Weibull Scale method is a very simple empirical method, which does its linear manipulation directly on the Weibull form and scale parameters (A, k) as well as adjustments on the frequency distribution.

The Weibull method has the advantage, that it will match the nature of the wind at most places, but beware that application of this method should be done with caution on locations with significant non-Weibull distributions.

Also, the scaling which is linear, is a quite simple and radical assumption. A very good directional distribution correlation is needed for the calculation to make sense. The method works best when only small corrections are needed.



Enercon E33 WTGs at Fehmarn, Germany.

5.2 The Weibull Scale Method

The Weibull Scale method presumes that the relationship between the Weibull distribution parameters and the frequency follow the general relation:

$$\lambda_{site}^{long} = \left[\lambda_{site}^{short} / \lambda_{reference}^{short} \right] \cdot \lambda_{reference}^{long}$$

where λ is the distribution parameter under consideration (Weibull A, k)

In the case of considering the frequencies, the modified long term frequency distribution must be normalized to 100%, i.e. for the N sectors under consideration:

$$f_{site,i}^{long} = \left[\left[\frac{f_{site,i}^{short}}{f_{reference,i}^{short}} \right] \cdot f_{reference,i}^{long} \right] / \left[\sum_{i=1}^N \left[\frac{f_{site,i}^{short}}{f_{reference,i}^{short}} \right] \cdot f_{reference,i}^{long} \right]$$

where f is the frequency
 N is the number of sectors (typically 12)
 i is the sector under consideration

5.3 Doing the Weibull Scale MCP

Weibull Scale MCP primarily requires look-up in the appropriate Weibull distributions, calculating the correction table and finally doing the calculation of the long term distribution. An example of the application of a Weibull Scale Method is shown in the Figure 1 below.

5. Weibull Scale MCP

Reference full: Alstedgårde 3yr

Height: 10.00														
Sector		0-N	1-NNE	2-ENE	3-E	4-ESE	5-SSE	6-S	7-SSW	8-WSW	9-W	10-WNW	11-NNW	Mean
A- parameter	[m/s]	4.12	4.57	4.86	5.61	5.86	4.44	4.59	5.54	5.49	5.95	4.56	4.11	5.14
Mean wind speed	[m/s]	3.69	4.06	4.31	4.97	5.19	3.93	4.07	4.90	4.87	5.27	4.05	3.66	4.55
k- parameter		1.655	1.821	2.001	2.016	2.276	1.949	1.919	2.176	2.415	2.428	1.916	1.715	2.013
Frequency	[%]	7.10	3.38	4.82	9.40	8.61	5.12	8.30	13.27	12.57	11.67	9.56	6.21	100.00

Reference Concurrent: Alstedgårde 3yr

Height: 10.00														
Sector		0-N	1-NNE	2-ENE	3-E	4-ESE	5-SSE	6-S	7-SSW	8-WSW	9-W	10-WNW	11-NNW	Mean
A- parameter	[m/s]	4.15	5.40	5.55	6.26	6.40	4.32	4.30	4.74	4.84	5.15	4.42	3.74	5.02
Mean wind speed	[m/s]	3.70	4.78	4.91	5.55	5.68	3.83	3.82	4.21	4.28	4.57	3.91	3.33	4.45
k- parameter		1.738	1.985	2.025	2.360	2.603	2.073	1.877	1.837	2.172	2.241	2.109	1.757	1.971
Frequency	[%]	8.75	4.45	6.39	12.13	9.98	4.71	8.21	11.21	9.96	9.29	8.55	6.36	100.00

Local Site Concurrent: Risø mast96-97(1y)

Height: 44.00														
Sector		0-N	1-NNE	2-ENE	3-E	4-ESE	5-SSE	6-S	7-SSW	8-WSW	9-W	10-WNW	11-NNW	Mean
A- parameter	[m/s]	6.50	5.37	6.17	6.96	7.91	6.33	5.16	6.11	6.80	7.12	7.17	5.74	6.65
Mean wind speed	[m/s]	5.76	4.77	5.47	6.19	7.05	5.61	4.57	5.42	6.03	6.31	6.35	5.22	5.89
k- parameter		2.109	1.930	2.184	2.769	2.900	2.454	2.015	1.940	2.151	2.087	2.137	1.433	2.125
Frequency	[%]	7.35	5.24	7.60	12.06	11.45	7.51	5.61	7.34	9.62	10.03	9.75	6.44	100.00

Correction Factors

Sector		0-N	1-NNE	2-ENE	3-E	4-ESE	5-SSE	6-S	7-SSW	8-WSW	9-W	10-WNW	11-NNW
A- parameter		0.993	0.848	0.877	0.896	0.916	1.027	1.067	1.167	1.136	1.154	1.032	1.098
k- parameter		0.952	0.917	0.988	0.854	0.874	0.940	1.022	1.185	1.112	1.083	0.909	0.976
Frequency		0.811	0.760	0.754	0.775	0.863	1.086	1.010	1.184	1.262	1.256	1.117	0.977

Local MCP-Corrected Data: Risø mast96-97(1y)

Height: 44.00														
Sector		0-N	1-NNE	2-ENE	3-E	4-ESE	5-SSE	6-S	7-SSW	8-WSW	9-W	10-WNW	11-NNW	Mean
A- parameter	[m/s]	6.46	4.55	5.41	6.24	7.25	6.50	5.50	7.13	7.73	8.21	7.40	6.30	6.87
Mean wind speed	[m/s]	5.72	4.05	4.79	5.53	6.44	5.75	4.87	6.32	6.85	7.28	6.56	5.75	6.09
k- parameter		2.008	1.770	2.158	2.366	2.536	2.307	2.060	2.299	2.392	2.261	1.942	1.399	2.035
Frequency	[%]	6.00	4.01	5.77	9.41	9.95	8.21	5.71	8.74	12.23	12.68	10.97	6.33	100.00

Figure 1: Sample Calculation with the Weibull Scale Method.

Note that required input to the Weibull Scale Method is distributions from three different sources:

- long term reference data (named 'Reference Full' above)
- the overlapping parts of the reference data (named 'Reference Concurrent'), and finally
- the site data (named 'Local Site Concurrent' in the Figure 1)

The calculation of the concurrent datasets is made directly in WindPRO MCP using the data from the correlation table. These data are then fitted to a distribution table and finally the Weibull parameters are fitted from the distribution table.

Please note that as this method does not involve the transformation of a time series it is not possible to compare a measured site time series with a predicted concurrent time series.

6. Measure & Correlate: Evaluating the Data for the MCP-Analysis

6.1 Introduction

Before doing the actual MCP analysis, some pre-processing of the measured data must be made in order to assure that the data are not contaminated with erroneous measurements. Furthermore, the pre-processing should aim in validating that the long term reference data is really representative for the local site conditions.

This chapter outlines a procedure for doing MCP source data evaluation. Typically, most of the evaluations could be done by using the Meteo-Object or the Meteo-Analyzer in WindPRO.



NEG Micon 750 kW wind turbine

6.2 First Step

First step is to evaluate as well the long-term reference as the short-term measurements. If the data do not correlate reasonable well or there are erroneous data, no method will work properly. A good example of poor data is a long-term reference that continuously decreases during the years due to urbanization or growth around the measure mast. Another reason for poor correlation can be that the wind climate simply differs too much from measure site to reference site. But here it is important to keep track on if it really is the wind climate or “just” differences in local terrain conditions.

6.3 Suggested Evaluations of the Long-Term Reference Data

- Is the reference mast sited in same wind climate as the measure mast (especially if one mast is in costal region or mountain region there might be larger wind climate changes)
- Were the same equipment used during all years
- Were the same measurement height used during all years
- Has the equipment regularly been calibrated?
- Did the surroundings change during the years (urbanization, new forest or windbreaks within 5 km radius), local obstacles within 1 km (buildings, trees etc.) or other changes, that might influence the measured wind speeds?

6.4 Evaluation of the Correlation

The suggested approach in evaluating the correlation of the reference data versus the local data is as follows:

- Graphical inspection of parallel data – disable obviously erroneous data. Time shift if necessary.
- Sector match graph or table (identify possible direction off set – possibility to correct). Low wind speed disabling while low wind speeds often has no meaningful direction (how to handle this has to be considered at a later stage, while moving low wind speeds from one sector to another can change prediction quite much).
- Draw and inspect figures of sector wise correlations

Furthermore, when one or more MCP calculations have been performed, then the final tab named ‘Predict’ in the MCP tool gives an overview of the calculations as shown below in Figure 1. For each calculation some key data are shown together with several diagnostic statistical quantities (‘r’ and ‘s’ values) indicating the goodness of each MCP prediction. If a particular MCP prediction was ended with an additional ‘StatGen’ calculation also two energy level indicators “Wind energy’ and ‘WTG energy’ are shown for the wind statistic.

6. Evaluating the Data

The ‘r’ values are relative measures of goodness of fit of the prediction, identical to the correlation coefficient in case of a linear model (see section 2.2.4 on Coefficient of Determination). ‘r’ can range between 0 and 1, with 1 being a perfect prediction.

Method	Description	Measure height a.g.l. [m]	Mean wind in measure height [m/s]	Key height a.g.l. [m] *	Mean wind in key height [m/s]	Wind energy	WTG energy	Wind statistic saved	Very long term wind [m/s]	Very long term wind [%]	r - wind speed	s - wind speed [m/s]	r - wind index	s - wind index [%]	Time of calculation
Regression MCP	Wishek met mast 7/01 to 6/02.40,0m - /NC...	40.0	8.15	50.0 ...	8.49	140.1	143.7	No			0.5368	2.8303	0.8540	5.1594	07-02-200...
Matrix MCP	Wishek met mast 7/01 to 6/02.40,0m - /NC...	40.0	8.08	50.0 ...	8.42	137.7	141.2	No			0.5207	2.8577	0.7969	6.2075	07-02-200...
Weibull Scale M...	Wishek met mast 7/01 to 6/02.40,0m - /NC...	40.0	8.14	50.0 ...	8.48	139.6	142.6	No							07-02-200...
Wind Index MCP	Wishek met mast 7/01 to 6/02.40,0m - /NC...	40.0		50.0	8.51	138.3	143.4	No					0.9054	8.4167	07-02-200...

Figure 1: The ‘Predict’ tab in a WindPRO MCP calculation.

The ‘s’ values are absolute measures of fit as is also seen from their units. In statistical terms the descriptive name is ‘root mean squared residual’ as it indicates the typical residual (i.e. error or deviation) of the predicted time series. ‘s’ is an unbounded positive quantity. The equation for “s” is:

$$s = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n - 2}}$$

where s is the root mean squared residual
 Y_i is the sample considered
 n is the number of samples
 \hat{Y}_i is the response value for the sample i – calculated using the model of consideration

For the Regression and Matrix methods there is an ‘r’ and ‘s’ pair based on the predicted wind speed time series and a pair based on monthly energy indices. For the Index method there is only an ‘r’ and ‘s’ pair for monthly energy indices as no wind speed time series is predicted. For the Weibull scale method there are no time series and, thus, no ‘r’ and ‘s’ values at all.

6.4.1 A Small Note on Comparing Different MCP-Methods

The estimates of correlation (r) and standard error (s) on wind speed for the linear regression method and the matrix method are based on a comparison of local on-site wind speed measurements and the predicted wind speed time series by transferring data from the reference series to the local site. The reported value is for individual measurements, though in the graphical presentation in WindPRO - it is possible to obtain values with different averaging periods.

The estimates of correlation (r) and standard error (s) on wind index for the linear regression method and the matrix method are based on a comparison of normalized production time series for local data and transferred (predicted) data. The production time series are obtained by applying a power curve to the wind speed time series (selected in the Setup tab). The normalization is done against the average production for the concurrent period. The reported value is using monthly averages (in effect the monthly windiness index), though other averaging periods are available in the graphical presentation in WindPRO.

The estimates of correlation (r) and standard error (s) for wind index method is based on a comparison of normalized production time series from the local and the reference data sets. No transfer function is included as the method does not predict a new time series for the local site, but rather apply a general correction factor to the local data set.

The fundamental difference on the calculation methods is the reason why the correlation and standard error values cannot be compared directly between Linear regression and Matrix method on the one hand and the Index method on the other hand.

6.5 Considerations Concerning Sample Averaging Period

Standard in meteorological measurements is to take the last 10 minutes as the typically available reference data; 1-hour, 3-hour or 6-hour values. If other method is used, this might lead to that the data set should be manipulated.

7. References & Literature

Below is a list of selected literature related to the Measure-Correlate-Predict methods and also the WindPRO implementation of the various MCP-methods. Additional literature and references are found at the end of the individual chapters which holds the detailed information of the implemented MCP-methods.



Enercon E40 and two Tackle Turbines at Schwarzenborn Knull, Germany

- [1] Woods, J.C. & S.J. Watson: *A new matrix method of predicting long-term wind roses with MCP*, Journal of Wind Engineering and Industrial Aerodynamics 66 (1997), p. 85-94.
- [2] Joensen, A., L. Landberg & H. Madsen: *A new measure-correlate-predict approach for resource assessment*, Proceedings of the 1999 European Wind Energy Conference, pp. 1157-1160.
- [3] Dale Addison, J.F, A. Hunter, J. Bass, M. Rebbeck: *A neural network version of the measure correlate predict algorithm for estimating wind energy yield*,
- [4] Gerdes, G. & M.Strack: *Long-term Correlation of Wind Measurement Data*, DEWI Magazin Nr. 15, August, 1999.
- [5] Derrick, A.: *Investigations into the Measure-Correlate-Predict Strategy for Site Wind Assessment*, ETSU W/11/00236/REP.
- [6] Nason, G.P. & T. Sapatinas: *Wavelet Packet Transfer Function Modelling of Nonstationary Time Series*
- [7] Frandsen, S. et al: *Power Performance Assessment*, Risø National Laboratory, Final Report, June 1999, pp. 137.
- [8] Riedel, Strack & Waldl: *Robust Approximation of Functional Relationship Between Meteorological data: Alternative Measure-Correlate-Predict Algorithms*, EWEA 2001.
- [9] Zaphitopoulos, Dellaportas, Morfiadakis & Glinou: *Prediction of Wind Speed and Direction at a Potential Site*, Wind Engineering no 3, 1999
- [10] Addison, Hunter, Bass & Rebbeck: *A neural network version of the measure correlate predict algorithm for estimating wind energy yield*
- [11] Ross, S.M: *Introduction to Probability and Statistics for Engineers and Scientists*, Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 1987
- [12] Efron, B & R.J. Tibshirani: *An Introduction to the Bootstrap*, Monographs on Statistics and Applied Probability, Chapman & Hall.
- [13] Press, W.H; B.P. Flannery; S.A.Teukolsky, W.T.Vetterling, *Numerical Recipes in Pascal*, The Art of Scientific Computing, Cambridge University Press, 1989.
- [14] Thøgersen M.L et al: *WindPRO Online: Remote Sensing Data in WindPRO* (WindPRO Reference manual on online services). EMD International A/S, 2010 (for WindPRO users: available from <http://help.emd.dk/knowledgebase>)
- [15] Sørensen, J.D. & I. Enevoldsen: *Lecture Notes in Structural Reliability Theory*, Aalborg University, Spring 1996.