Statistics of Wind Resource Assessment using Measure-Correlate-Predict Methodology

by

Niyazi Gur

A Thesis
Submitted to
Heriot-Watt University
In Partial Fulfillment of the Requirements for the
Master of Science Degree
In Renewable Energy Engineering

Heriot-Watt University
School of Engineering and Physical Sciences

August, 2011

Academic Supervisor: Dr Wolf-Gerrit Früh
Industrial Supervisor: Dr Peter J. M. Clive

August 26, 2011

DECLARATION

I hereby declare that the work entitled *"Statistics of Wind Resource Assessment using Measure-Correlate-Predict Methodology"* submitted to the Heriot-Watt University under the requirement of masters dissertation project, is a record of an original work expressed in my own words. Any references made to the work done by other authors in any form (e. g. ideas, figures, text, tables) are acknowledged at their point of use.

Matriculation number: H00020405,

Yours sincerely,

Niyazi Gur

Abstract

Six data sets with widely varying characteristics wind speed and direction data are analyzed within the developed Measure-Correlate-Predict (MCP) methodology and so as to investigate the predictive capability of the method, comparisons are made by performing the Linear Regression Method. "Sector Pairing" procedure, which is another method that is developed for the purpose of this project, applied to each datasets in each geographical group and in order to deal with meaningful intrinsic direction sectors 15 month of short-term concurrent data are used to perform the procedure. Measure-Correlate-Predict method allows establishing a relationship between wind speeds measured at target and reference site during the short-term wind speed measurement campaign. This relationship then can be used to predict the long-term wind regime at a potential wind farm development site using the long-term data from a nearby meteorological station. Analysis had shown that increase in the length of data used when predicting the long-term wind speed regime at a site, regardless the direction, tends to increase the predictive capability of the developed MCP methodology and it has been investigated that the correlation coefficient "r" between the reference and the target sites over the short-term tends to increase as the distance between and elevation differences between the sites decreases, however this has been uncertain for some group predictions. In the analysis for each individual group predictions are made by performing the developed method, this method named as "Modified Weibull Scaling Method" and had showed better agreement in predicting the long-term wind speeds at the target sites compared to that of the Linear Regression method, where the linear regression has been used to make comparisons to see that how well the predictions are. Linear regression method has the potential to generate robust predictions of wind speed at a target site. Of the evaluated measure-correlate-predict methodologies, the "Modified Weibull Scaling Method" is recommended for general usage regardless of the correlation coefficient value or the availability of historical data from a nearby station. However, on the success of the Modified Weibull Scaling Method more analysis are necessary to disaggregate the influence of correlation coefficient, length of historical data over the long-term and the length of concurrent data over the short-term. The correlation coefficient is highest between sites with uniform terrain such as coastal locations and lower

between sites with complex topography. In addition to this, predictions performed by performing the "Modified Weibull Scaling Method" showed better agreement of the predictions with the measured wind speeds for the sites located in coastal areas and experience from Atlantic wind speed patterns. Though both MCP methods produced highly uncertain results for sites with lower correlation coefficients. Though for sites with higher correlation, special consideration should be paid to the uncertainty of estimated metrics introduced by the Modified Weibull Scaling Method and an increase in sector pairs such as taking arbitrary 9 pairs of 40-degree sectors rather than 3 pairs would enhance the predictive capability of the Modified Weibull Scaling Method. However in case of an increase in sector pairing, the use of longer period of short-term concurrent data will be necessary so as to deal with meaningful amount of data in each paired dataset, meaning that there would not be enough data to represent a better prediction if the use of longer concurrent data is disregarded. The Linear Regression Method is not recommended due to high prediction error observations from the analysis.

Acnowledgements

The Author of this masters dissertation project would like to thank to his both academic and company supervisors Dr Wolf-Gerrit Früh and Dr Peter J. M. Clive, respectively, for their valuable guidance and advice, and the Engineering and Physical Science Department of Heriot-Watt University for facilitating the development of this project.

I also would like to extend my special thanks to Ginny Spencer, from the Postgraduate Office of Engineering and Physical Science Department, for her invaluable support whenever required throughout the academic year.

Table of Contents

LIST OF FIGURES

Figure **Page**

LIST OF TABLES

LIST OF ACRONYMS

MCP	=	Measure-Correlate-Predict
LRM	=	Linear Regression Method
MWSM	=	Modified Weibull Scaling Method
TWE	=	Total weighted error
cdf	=	Cumulative distribution function
Dist	=	Distance
Elev	=	Elevation
Long	=	Longitude
Lat	=	Latitude
Q-Q Plot	=	Quantile-Quantile Plot
SWE	=	Sector-vise variability of the prediction error
NN	=	Northerly Northerly Sector
NE	=	Northerly Easterly Sector
NW	=	Northerly Westerly Sector
EN	=	Easterly Northerly Sector
EE	=	Easterly Easterly Sector
EW	=	Easterly Westerly Sector
WN	=	Westerly Northerly Sector
WE	=	Westerly Easterly Sector
WW	=	Westerly Westerly Sector

LIST OF NOTATION

r^2	=	Correlation coefficient
N	=	North
W	=	West
M	=	Meters
km	=	Kilometer
Z_{hub}	=	Hub height of the turbine
Z_{mast}	=	Distance between ground level and the wind speed data at mast height
Z_0	=	Roughness length
U_{mast}	=	Wind speed at mast level
V	=	Wind speed at height z above ground level.
V_{ref}	=	Reference wind speed.
s'^2	=	Predicted variance of wind speed
s^2	=	Measured variance of wind speed
k	=	Shape parameters
c	=	Scale parameters
u	=	Random variable or wind speed
v	=	The mean speed
s	=	Standard deviation
y	=	Target site wind speed
x	=	*Reference site* wind speed
c	=	Offset
k_x	=	Shape parameter at the reference site
k_y	=	Shape parameter at the target site
λ_x	=	Scale parameter at the reference site
λ_y	=	Scale parameter at the target site
m/s	=	Meters per second
km^2	=	Kilometer square
agl	=	Above ground level

INTRODUCTION

Climate change is one of the biggest threats for the planet earth that mankind faces today and renewable energy generation is one of the primary options in tackling this issue. The use of wind energy convertion systems as an alternavive to conventional energy generation can reduce demand on fossil fuel, enhance air and water quality, and reduce greenhouse gas emission [27]. Wind power generation especially in countries like Scotland, where there are best mean wind speed variations in the continent of Europe [1] [26], may play a big role in achieving carbon reduction targets that are set by the European Commision for each individual country in the European Union or other renewable energy generation targets elsewhere in the world.

Deployment of large-scale wind farms requires huge capital investments [24]. Therefore, comprehensive wind resourse assessment of long-term wind regime is required in order to estimate relaibility of a site chosen for a wind farm development since this is esential for wind farm developers so as to estimate the pay back period and profitability of the invesment. This is one of the biggest concern for the developeres from the economics standpoint.

Meanwhile, on-shore wind farm development is one of the cheapest options to generate decarbonized electricity with a capital cost of 1300-1600£/kW, Table 15, page 66, and this figure is predicted to be gradualy decrease in the near future due to being one of the fastest growing industry in the globe alongside the other factors, such as fuel cost, opetration and maintanence costs, govenment interventions, etc. [19] [20] [21].

Due to the variable nature of wind speed, which may show changes in hourly, daily and annually basis, at least one year of concurrent data is recomended to be used in order to make reliable energy yield prediction of a wind power development project [5]. Nevertheless, more than 12 months wind speed monitoring campaign would be very expensive and time consuming to make. In addition to this only one year monitoring campaign may not represent annually variations in mean wind speed and direction [5] [31].

Measure-Correlate-Predict Methodology (MCP) is one of the most popular and widely used wind resource assessment techniques in predicting the long-term with resources at a potential wind farm development site. This method involves establishing a relationship between the measured concurrent data from reference and target site over the short-term period, then this relationship correlated along with the long-term wind speed data from a reference site with a historical record so as to predict the long-term wind speed variations at a candidate site [18]. This study investigates the predictive capability of two MCP methods, as well as analyzes the associated errors in the predictions of the Modified Weibull Scaling Method (MWSM), which is developed for the purpose of this study and the Linear Regression Method (LRM) so as to compare the predictive capability of the developed method. These methods performed by applying the *sector pairing procedure*, which is another technique developed in order to perform the analysis.

Group A wind speed and direction data that is used in this study provided from the Sgurr Energy, which is a Renewable Energy Consultancy based in Glasgow, UK and sponsor for this project. The rest of data (Group B to F) extracted directly from the British Atmospheric Data Centre (BADC) website with an access to Met Office - MIDAS Land Surface restricted datasets (Table 2, page 20). All data required further processing, however, to minimize the discretization problems associated with the use of integer values in recording, converting from *knots* to *m/s* and storing the data. Data sets were chosen for their availability, as well as their geographical and topographical characteristics and location that have potential for wind energy generation in the United Kingdom.

1. LITERATURE REVIEW

1.1. Wind Resource Assessment

Measure-Correlate-Predict (MCP) Method is a tool in the centre of wind resource assessment techniques in predicting the long-term with regime at a potential wind farm development site [18]. Due to the variable nature of wind speed, which may show different characteristics from location to location and from time to time, an appropriate estimation of the long-term wind speed data of a period of at least 20 years in length is required [29] [28], which is a typical lifetime of a wind turbine [2], and is recommended to use in making long-term wind resource predictions. This is required to examine the suitability of a wind farm at a site.

Location of a wind farm development site is often identified through consultation of wind resource variability maps, visiting field, and data from nearby meteorological stations. Installing a mast on-site to monitor the wind speed follows the process above. A mast usually comprises of anemometers for wind speed, wind vanes for direction and sometimes temperature and pressure sensors. Data are usually logged in the system in hourly basis and seasonal and diurnal variations can be derived from the hourly logged data [3].

1.2. Physics of Wind

Wind is the flow of gasses on large scales and is the consequence of temperature and pressure gradients in the earth's atmosphere, which are caused by the uneven heating of the Earth's surface by the sun. In accordance to this wind cycles prone to the variations in both speed and direction, therefore, the use of short-term wind speed measurement campaigns in predicting the long-term availability of wind speeds at a site cannot be representative in the nature of the statement specified above. Over long-term periods, the mean speed reaches to a characteristic value [25]. However, can still be prone to changes due to influence of the long-term climatic events such as El Niño (ENSO), the North Atlantic Oscillation and the Arctic Oscillation [4] [32].

Wind speed varies considerably with height above the ground level so called as *wind sheer*. In theory a hub height of, for example 50m will experience far stronger wind speed variations due to lack of obstacles at higher hub heights, near-to-ground boundary layer up to 100m. At height of ground (z) = 0 the air speed is always zero [22]. Within the height of local obstructions wind speed increases erratically, and violent directional fluctuations can occur in strong winds. Above this erratic region, the height/wind speed profile is given by the equation 1 below [6];

The Log Law to extrapolate wind speed is shown below;

= *h* = *h* / 0 / 0) *h* 0= *h*

h (1)

Where;

Z_{hub}	=	Hub height of the turbine
Z_{mast}	=	Distance between ground level and the wind speed data at mast height
Z_0	=	Roughness length (Table 1)
U_{mast}	=	Wind speed at mast level

= * (2)

Where;

V	=	Wind speed above ground level.
V_{ref}	=	Reference wind speed.

Various methodologies are used to assess the mean wind speed at hub height. Among them, the Log Law is often used method in extrapolating wind speed to a different height, from a reference hub height with a known wind speed [13].

It is important to mention that using this method may give inaccurate predictions of extrapolated height mean wind speeds. It has been found that for hilly terrain with no trees, the difference between the predicted hub height (50m) and the measured wind speed values are ranged from about 1% to 13% [13]. It should be noted that wind shear changes with a number of variables. In estimating the hub height wind speeds with Log Law method the table below is used:

Type of terrain	Roughness length (m)
Cities, forests	0.7
Suburbs, wooded countryside	0.3
Villages, countryside with trees and hedges	0.1
Open farmland, few trees and buildings	0.03
Flat grassy plains, rough pasture	0.01
Lawn grass	0.008
Blown sea	0.0005

Table 1: Typical surface Roughness Lengths [8, 13].

1.3. Wind Speed Distribution

The Weibull distribution is a commonly used probability density function in order to evaluate the wind speed distribution for a potential wind farm development site with knowing two parameters, "*k*", which is the shape factor and "*c*", which is the scale parameter. In the case of The Rayleigh distribution, which is a special case of the Weibull distribution, the shape parameter *k* assumed to be equal to 2, this simplification applied only when the wind speed at a site is known.

The Weibull distribution function is [6]:

$$(\ )= \qquad -1\exp- \tag{3}$$

Where *k* and *c* are the Weibull "shape" and "scale" parameters, respectively, *u* is the value of the random variable; *u* is the wind speed in the case of wind resource assessment.

In the case of Weibull distribution, higher value of shape parameter "*k*" indicates that the wind speed variability at a site is low and smaller value of shape parameter "k" indicates that there is greater variability in the mean wind speed variations. Therefore, when the shape parameter is small it can be assumed that there will be high probability occur in low wind speeds.

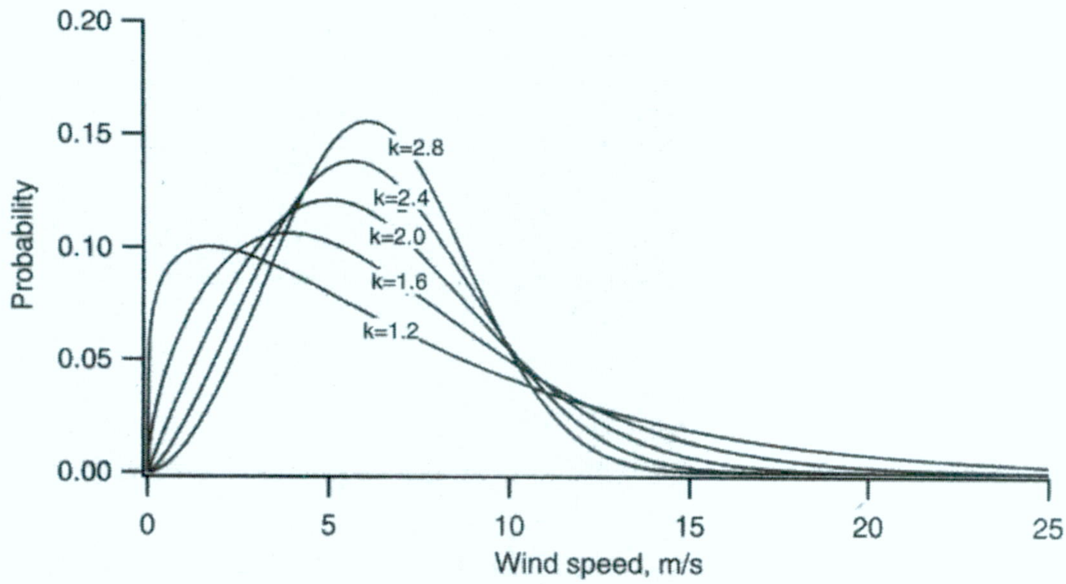


Figure 1: Influence of shape parameter "k", where the mean wind speed taken to be constant at 8 m/s to show the Weibull distribution of different shape parameters [13].

1.4. Methods for Estimating the Weibull Distribution Parameters

There is variety of different methods in estimating the shape and scale parameters of the Weibull distribution. Two of them are used to perform the analysis in this study, which are the Method of Moments and Quantile-Quantile plot (Q-Q Plot) and reviewed in details.

1.4.1. Method of Moments

The advantage of fitting the data by performing the Method of Moments that it is less biased estimator of the shape and scale parameters of the Weibull distribution compared to that of the Q-Q Plots. It is also easy to implement except there is an iterative aspect that there is not in the use of Q-Q Plot.

The first moment and second corrected moments of the two-parameter Weibull distribution are as follows [9]:

$$(\ )= \quad 1+1 \qquad (4)$$

$$2= 2 \quad 1+2 \quad - \quad 21+1 \qquad (5)$$

Therefore, given sample estimators of the mean speed *v*, and its standard deviation *s*, the following can be solved iteratively for *k*, where is the gamma function [9], equation (6);

$$2 \quad 2= \quad 1+2 \quad - \quad 21+1 \quad 21+1$$

The scale parameter *c* can then be estimated from, equation (7);

$$= \quad 1+1$$

Essentially the Method of Moment requires iteratively a sort of a change in shape parameter until it is found that it gives a theoretical Weibull distribution, where there are moments of the same as the moments of measured distribution.

1.4.2. Quanlile Quantile Probability Plots

A Quantile versus Quantile probability plot (Q-Q plot) is a probability plot, which allows to the user to compare the probability distributions visually by plotting their quantities against one to another. It can be used to compare the shape and scale parameters of the Weibull distribution, which provides a graphical view to show that how the parameters of the Weibull distribution similar or different.

Q-Q plots have extensively been used in order to perform the analysis in this study. The feature of displaying data on a Q-Q plot is that the slope and the intercept of a line fitted to the points on the plot, which in turn provide estimations of the scale and shape parameters of a Weibull distribution. Q-Q plot is an efficient and unbiased estimator of the shape and scale parameters. Another advantage of the Q-Q plot is that the every single data, i.e. hourly wind speed, are fitted against one to another by equating their squared correlations, r^2, rather than replacing x_i with $y_i = j + kx_i$. Consequently there will be changes in shape and scale parameters due to slope and intercept resulting linear end which will in turn reflect the changes in both shape and scale parameters [10].

The cumulative distribution function (cdf) f(x) of a Weibull distribution is [10];

$$f(x) = 1 - exp(-(x/\lambda)^k) \tag{8}$$

The cumulative distribution function f(x) value can also be calculated from the rank of the data. For example, the data with a rank, which is one fifth of the total number of data, is the value below which 1/5 of the data lie and above which 4/5 of the data lie. Therefore, the cumulative distribution function is also given by:

$$f(x) = Rank(x)/(Count + 1) \tag{9}$$

Subtract 1 from everything;

$$1 - Rank(x)/(Count + 1) = exp(-(x/\lambda)^k) \tag{10}$$

Logs of both sides taken;

$$Ln[1 - Rank(x)/(Count + 1)] = -(x/\lambda)^k \tag{11}$$

Negate (logs of the both side of the equation are taken):

$$-Ln[1 - Rank(x)/(Count + 1)] = (x/\lambda)^k \tag{12}$$

$$Ln\{1/[1 - Rank(x)/(Count + 1)]\} = (x/\lambda)^k \tag{13}$$

Logs taken again;

$$Ln(Ln\{1/[1 - Rank(x)/(Count + 1)]\}) = Ln\{(x/\lambda)^k\} \tag{14}$$

$$Ln(Ln\{1/[1 - Rank(x)/(Count + 1)]\}) = k\,Ln(x) - k\,Ln(\lambda) \tag{15}$$

So if the wind speed is called = x and call the number that has derived from its rank $Ln\{1/[1 - Rank(x)/(Count + 1)]\} = y$, which is a log-log plot of a straight line whose slope is the shape parameter k, and whose offset then gives the scale parameter λ. the advantage of using Q-Q plot is that it can all be done in a straightforward way in Microsoft Excel without any iterative procedures.

Alternatively, swapping the axes to plot $ln(x)$ against $ln(y)$, which is perhaps easier to handle gives;

$$ln(x) = ln(y)/k + ln(\lambda) \tag{16}$$

When fitting the data by performing Q-Q plot method, the data are ranked from 1 to N and the measured data sorted from smallest to largest, the value of the cumulative distribution function is empirical for the 1 value is $1/N$. The cumulative distribution function is also calculated on the basis of the value itself in terms of the shape and scale parameters. Therefore, having empirical value for the cumulative distribution function and the theoretical value in terms of shape and scale parameters, a straight line can be generated by plotting them one to another.

The problem with the Q-Q Plot is that it is not unbiased estimator of the shape and scale parameters of the Weibull distributions because there is a need of rearrangement in order to get in the form of straight line and where we end up with an expression of x and y which are prone to uncertainties and introduces bias in data fitting of the equation of the theoretical and empirical cumulative distribution function values.

The use of a straight line is required because the point is to equate the theoretical and empirical cumulative distribution functions against to each other, which allows for the rearrangement of the terms to get an equation of a straight line.

The benefit of using Q-Q Plot is that it is very easy to implement and allows user to deal with the data that is not uniformly Weibull distributed, which in turn allows user to exclude calms and extreme wind speeds.

1.5. Measure-Correlate-Predict Methodology

The measure-correlate-predict (MCP) methodology uses the statistical relationship between the data measured at a station nearby to a wind power development site to make a prediction of wind speed at a target site, which based on a long-term wind speed data at a reference site. MCP methods can be explored by using various correlation techniques in the literature depending on the site location features and topography, available length of the data or quality of the data and other factors [5] [30] [33]. Establishing a relationship between sites is very complicated by stochastic variations in wind speed and direction, which is subject to changes over a period time and also the distance, the variability of wind speed would also be obstructed by surrounding nature such as; instability of atmospheric environment, terrain effects on the flow of wind speed and direction, weather conditions. Therefore choosing an appropriate modeling technique is very important to avoid issues that may be arisen in the exploration of the MCP method. MCP method, along with fluid dynamic modeling schemes, is a commonly used method in the wind power generation industry to allow investors to estimate the long-term power production of a wind farm based on a short- term period of monitoring campaign [28], this is done so as to investigate the reliability of the site chosen for a wind farm development from the economics standpoint.

The term of *target site* is used extensively throughout this report and it refers to the site where a wind farm development may take place and *candidate site* used interchangeably with this term. A close by station with a long historical record, referred to as the *reference site*, where the long-term of the wind speed variability is known.

An overall illustration of an MCP method can be seen in Figure 2, 3 and 4 below. From an available reference site data over the long-term and concurrent data of the target and reference site over the short-term data has been shown in Figure 2 and 3 below, the length of the data over the short term usually taken to be 12 months. Figure

3 illustrates the target wind speeds over the short-term wind speed measurement campaign at a potential wind farm development site. A relationship would then be established using the reference wind speed data over the long-term with the concurrent data over the short term so as to predict the wind speed variability at the target site, an example of this can be seen in Figure 4 below.

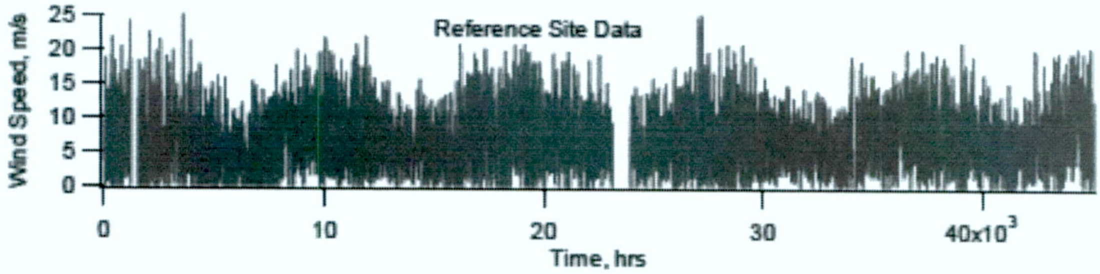


Figure 2: Available reference site data over the long-term [11]

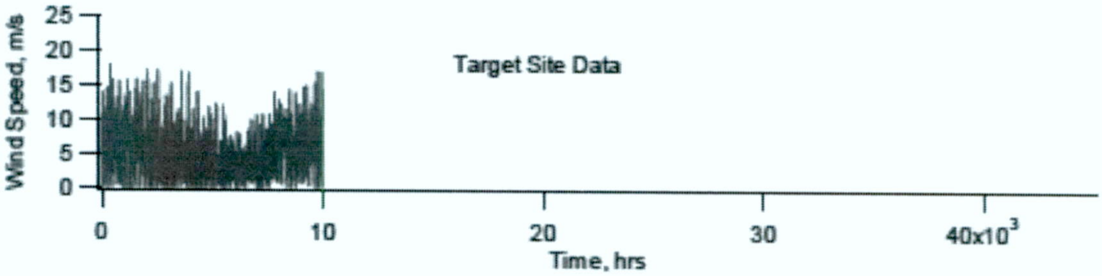


Figure 3: Available target site data over the short-term [11]

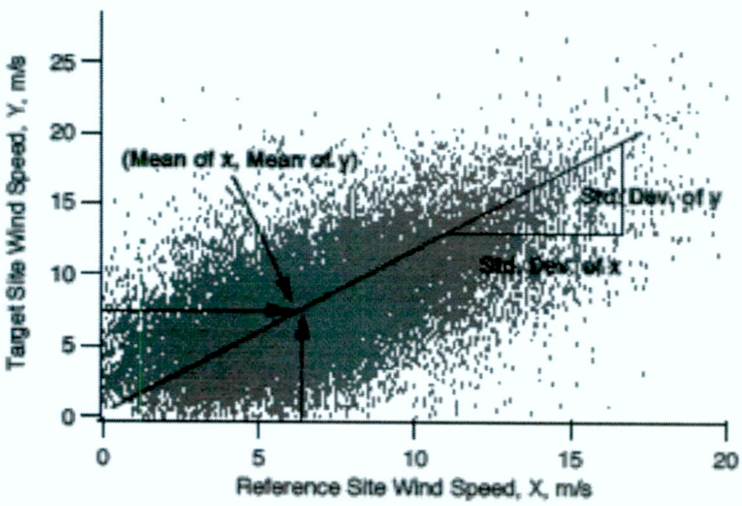


Figure 4: Estimation of the target wind speed over the long-term by plotting a straight line between the datasets [11]

The energy yield of a wind farm is well related with the availability of wind resources at a candidate site for a wind farm development; the more the assumptions appropriate means is the more accurate the energy yield estimation are. Therefore estimation of the long-term wind regime is a crucial part in implementation of wind conversion technologies, availability of the resources at a candidate site is predicted by modeling the relationship (often wind speed) between the short-term concurrent wind data from both target and the reference sites, this is in general taken to be up to 12 months time period, though in some journal papers it has been concluded that even 8 month of concurrent data would adequately establish a good relationship with some certain wind resource assessment techniques [5]. The relationship over the short-term data then can be used in predicting the long-term wind regime at a target site with forming a relationship with the reference site, where the historical wind regime is known and this is usually 10-20 years of wind speed data [14]. The general methodology of MCP process is illustrated in Figure 5 below.

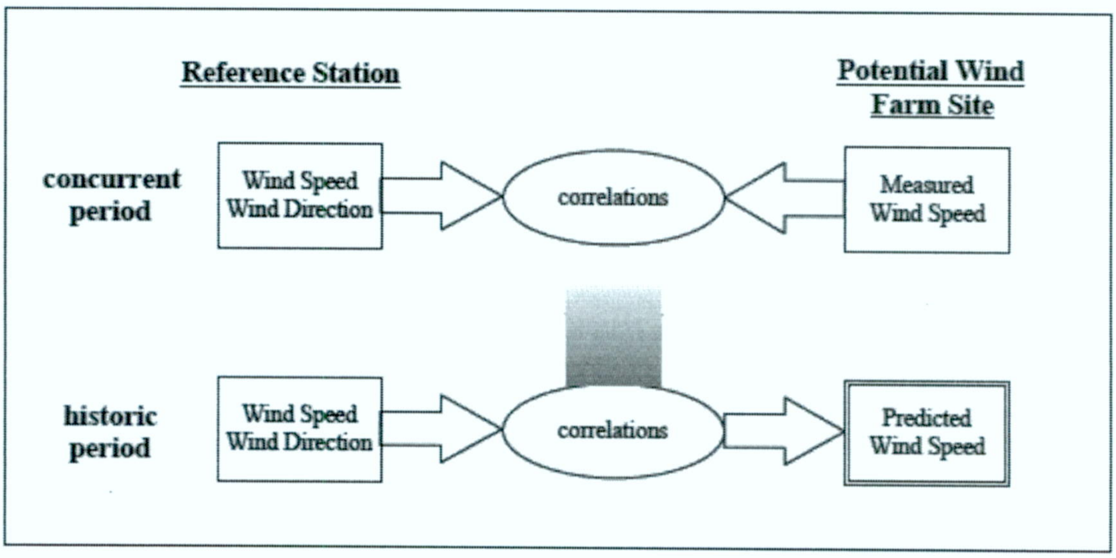


Figure 5: Methodology of Measure- Correlate-Predict Techniques [12]

Variety of different MCP techniques have been proposed in the last 15 years time [12], these all are addressing the consequences of uncertainties within the conventional methodologies based on analysis performed and all varies in modeling the definitions, directional sectors and adequacy of available data.

Establishing a relationship between the concurrent data over the short-term from the reference and target sites are used to make prediction. The data could be grouped or binned into different portions such as directional sectors, however, in this case directional behavior of wind speed is binned independent from the wind speed in the current literature [14] [15]. MCP methods use different algorithms in fitting the parameters. This fitting than would be applied into the historical long-term data from the reference site in order to make prediction for the long-term target site wind speeds.

There are different aspects of choosing the required period of concurrent wind speed data over the short-term with respect to MCP methodologies; this is done to ensure that the length of the available data over the short-term will be reliable enough to establish a good relationship over the long-term between the target and reference site. This would be depending on the specific site features such as topography of the site. However, in a journal paper it has been concluded that 9 months or longer concurrent data is required so as to decrease the uncertainties arisen from the MCP methodologies [15].

In this literature review two different MCP methods are reviewed. These methods are *Linear Regression Method* and *Weibull Scaling Method*, respectively;

- Linear Regression Method
- Weibull Scaling Method

1.5.1. Linear Regression Method

The regression MCP methods use traditional linear regression MCP analysis, polynomials of other orders are used in establishing a relationship (i.e. 2^{nd}, 3^{rd} or 4^{th} order), and however higher-order polynomials do not provide much improvement in fitting the data, and moreover are prone to instabilities in the fit [14]. In addition to the linear fitting, in some certain cases a better relationship would be drawn by forcing the line passing through the origin (0,0), this is done to avoid unrepresentative results in low wind speeds, such as calms, for example when the wind speed is 3 m/s at the reference site and a wind speed 0 m/s at the target site, which may be unrealistic to establish a relationship, for example if the proximity between the sites is short, meaning that considering such a relationship might be wrong since the wind speeds at both site may be assumed to show very high correlation . However, also forcing the regression line passing through the origin is typically represents poorer fit to the data [16].

Thøgersen et al. (2007) concluded that analysis performed by using WindPRO program by implementing linear regression method to the available data, as much as 10 per cent of the energy would dramatically be lost in the long-term correction of the data due to not including the random errors (residuals). Even if that there is a good equation in reality it is not going to necessarily conform to the straight line [17]. Therefore there is a need of a stochastic random variable so as to fix this issue. This is due to the fact that the error variance also includes stochastic differences between the sites along with their measurement errors [17].

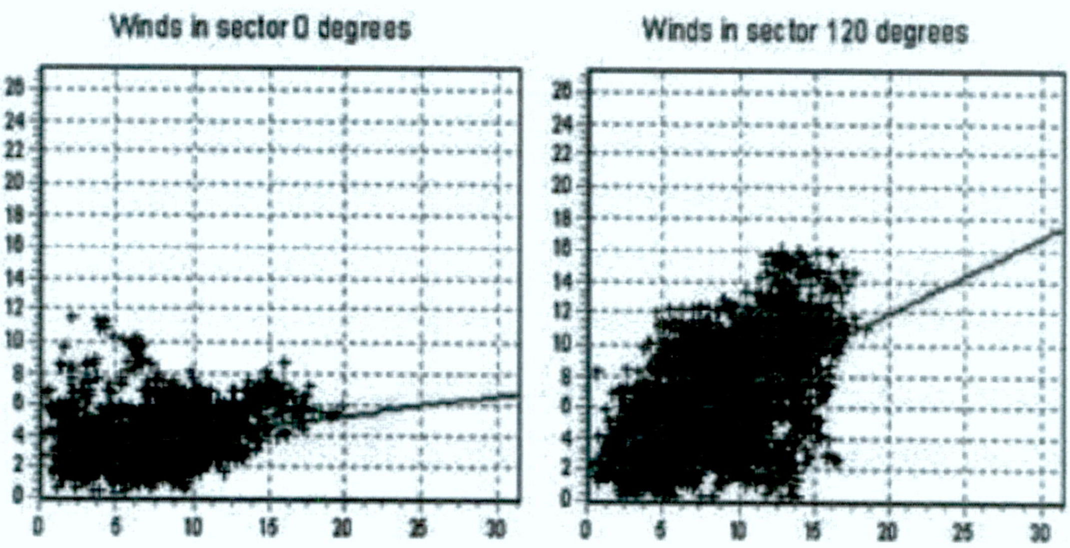


Figure 6: Linearly fitted wind speeds (x axis indicates the wind speed at the reference site and y axis indicates the wind speed at the target site) [17]

Regression model

Regression models are used so as to establish a relationship between the sites, which assumed to model the wind speeds appropriately enough [17];

$$y = f(x) + c \qquad (17)$$

where y is the wind speed at a candidate site, whose long-term wind regime is required to be predicted, x is the wind speed at the reference site, whose historical wind regime is available, c is the offset (residual) and $f(x)$ represents the regression model, which are determined from linear regression and often assumed to be mx, where x is the wind speed at the reference site and m is the slope [17].

The regression model would be represented by polynomials of any order. However, a linear method in general, assumed that this model gives a reasonable result in estimating the wind energy production. Clive (2008) found that both 2^{nd} and 5^{th} polynomial order fitted to the data establish a good relationship with the 3-parameter power law, however, at low wind speeds 5^{th} order polynomial displays a considerable wobble, whereas 2^{nd} order polynomial fit provides a close approximation to the power law, see the Figure 7 on the next page. McKenzie et al. (2008) noted that if linear fit provides a good relationship, 2^{nd} order polynomial fit would also provide a good approximation of the prediction.

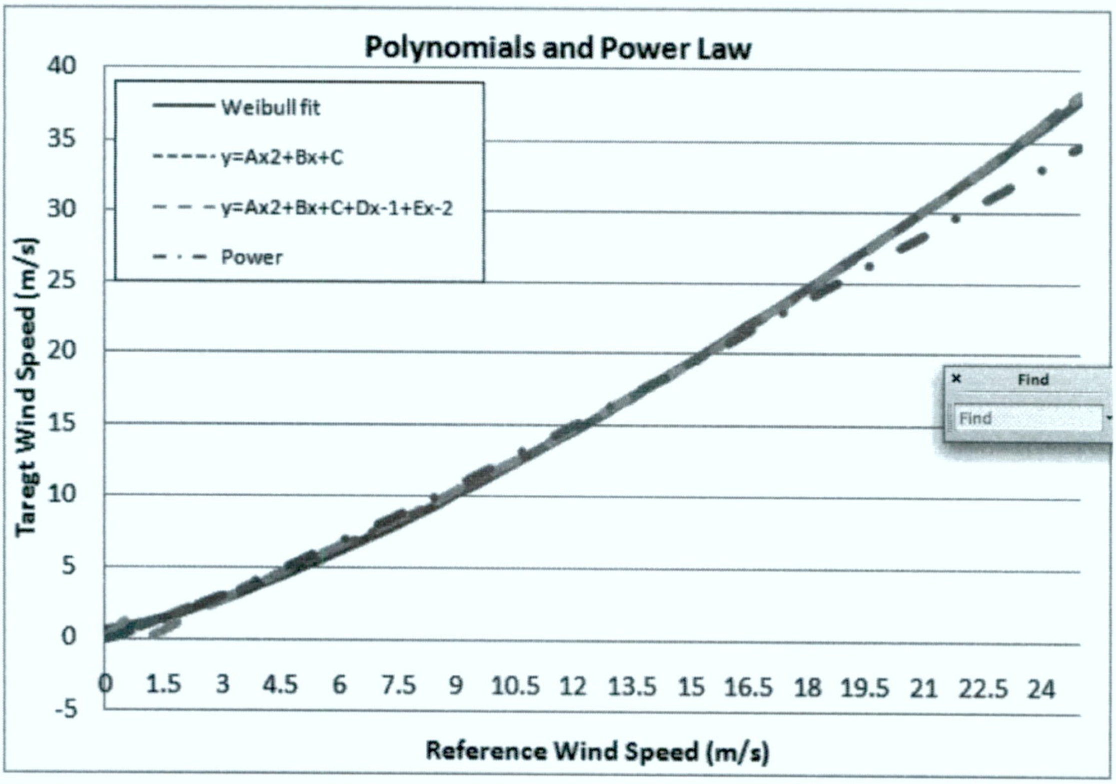


Figure 7: Comparison of polynomial and power law fits [14]

In the case of using "Linear Regression method", the predicted mean wind speed at the candidate site will be closer to the value of the measured mean wind speed [5]. However the predicted variance at the candidate site will be less than the measured variance in accordance to the expression below [7]:

$$s'^2 = s^2 r^2 \tag{18}$$

where s'^2 and s^2 are the predicted and measured variance of wind speed respectively, r^2 is the correlation coefficient of the regression fit. This is because of the non-linear relationship between wind speed and wind power, this can result in significant underestimations of the available wind resource, in particular where the sites that are less correlated [7].

1.5.2. Weibull Scaling Method

The Weibull scaling method is a conventional empirical method, which establishes a linear fitting directly onto the Weibull distributions of both target and reference site wind speeds. Weibull distribution is assumed to be that it will establish a reasonable relationship in most places, however, in reality the relationship between Weibull distributions is actually not linear, where this relationship assumed to be linear in performing the analysis by using "Weibull Scaling Method" in the current literature [16];

$$F_x(x)= F_y(y) \Leftrightarrow 1\text{-}e^{-(x/ \lambda x)kx}= 1\text{-}e^{-(x/ \lambda y)ky} \Leftrightarrow (y/\lambda_y)^{ky}=(x/\lambda_x)^{kx} \tag{19}$$

Hence:

$$y=\lambda_{\,y}(x/\lambda_x)^{kx/ky} \tag{20}$$

where k_x, k_y and λ_x, λ_y are the shape and scale parameters of the reference and target site and y and x represent target and reference wind speeds, respectively. The non-linear relationship between k_x and k_y can be seen from the equation 19, 20 and 21, however, in the literature the relationship between k_x and k_y assumed to be unity [16];

$$y=\lambda_{\,y}(x/\lambda_x) \tag{21}$$

It can be seen that the shape parameters of the both locations are assumed to be equal in equation 21 above, and effects of this associated with the long-term wind speed predictions is well underestimated. Appropriate correlations are necessary in the nature of this method so as to result in more sensible and accurate calculations.

It is clear that an inappropriate use of Weibull Scaling Method has been adopted in establishing a relationship by assuming that target and reference shape parameters (k_x, k_y) are equal to each other, which are not unless that the both sites are identical and have the same shape parameters. Therefore, an error may occur in making predictions if the shape parameters are differing at the reference and target sites.

2. METHODS

The analyses are split into four parts. The first part titled as "Data Summary", presents relevant descriptive statistics of the datasets that are used in order to perform the analysis of the measure-correlate-predict (MCP) methods. The second, titled "Estimation of Weibull Distribution Parameters", investigates the predictive capability of shape and scale parameters of Weibull distributions by comparing three different approaches, these are Method of Moments, Quantile-Quantile Plot (including calms and extreme wind speeds) and Quantile-Quantile Plot (excluding calms and extreme wind speeds). The Third, titled "Sector pairing", describes sector-pairing procedure, which is the technique developed in order to identify the data subsets by pairing direction sectors. The forth, titled "Modified Weibull Scaling Method", which also has been developed for the purpose of this project, in this analysis "Linear Regression Method" and "Modified Weibull Scaling Method" are used to analyze the available datasets in each geographical group pair and makes comparison of their predictive capability between these two method as well as the associated error of the methods.

Overall methodology for parts two, three, and four as described above are identified and explained under this chapter. For each set of paired data, the length of the concurrent data is taken to be 15 months and predictions of the wind speed over the long-term are based on this data in order to perform the correlation methods. The resulting values are used to make a prediction at the target site(s), based on the historical data at the reference site(s). The resulting prediction then compared to the measured wind speed data from the target site over the long-term. The correlate and predict procedures are repeatedly applied on each geographical data pair, and wind speed availability at the target sites in each group are calculated to compare the predicted versus the actual values.

2.1. Data Summary

Wind speed and direction and other optional atmospheric observations (e.g., barometric pressure, temperature, gust, etc.), are obtained and paired in two stations. Six groups of data sets consisting of 12 sites in total were acquired (Table 2, page 20 4, page 29, Figure 9, page 28). The sites in each group greatly vary in distance between them, where the closest sites in a group, in Group A, Boscombe Down and Larkhill are just under 6km apart and the furthest sites in a group, in Group E, Stornoway Airport in Western Isles are just over 190 km apart. The data in each group have 20 years long-term available data outside the exemption of the Group A data, where the analysis performed by using 10 years long-term data for Group A predictions.

The wind speed data gives hourly variations of the wind speed and direction and is used an airflow model to estimate the effect of topography on wind speed. The wind speed measurement model by MIDAS uses a one km^2 resolution at 10m above ground level (agl) and does not take account of topography on small scale, of local surface roughness, such as crops, stone walls or tress, which in reality may have effect on the wind speed and direction [34]. List of stations with period, location, elevation and distance between each pair can be seen in the Table 2 on the next page.

Group	Geographic area	Station name	Start of Period	End of Period	Lat. (N)	Long (W)	Elev. (m MSL)	Dist. (km)
A	Argyll North	Machrihanish	2000	2010	55.4408	-5.6957	10	
	Lanarkshire	Salsburgh	2000	2010	55.8330	-3.8660	240	122.8
B	Wiltshire	Boscombe Down	1990	2010	51.1613	-1.7532	126	
	Wiltshire	Larkhill	1990	2010	51.2012	-1.8044	132	5.697
C	Cambridgeshire	Wittering	1990	2010	52.6110	-0.4596	73	
	Bedfordshire	Bedford	1990	2010	52.2265	-0.4638	85	42.76
D	Shetland	Fair Isle	1990	2010	59.5265	-1.6278	57	
	Shetland	Sumburgh	1990	2010	59.8793	-1.2974	7	43.39
E	Argyll	Tiree	1990	2010	56.4998	-6.8796	9	
	Western Isles	Stornoway A.	1990	2010	58.2138	-6.3177	15	193.5
F	Orkney	Kirkwal	1990	2010	58.9539	-2.8999	26	
	Caithness	Wick airport	1990	2010	58.4541	-3.0884	36	56.63

Table 2: List of stations with period, location, elevation and distance between each pair. The first region was used as the target station; the second is the reference station. All data periods ended on 31 December 2010.

2.2. Estimation of Weibull Distribution Parameters

The wind speeds for Machrihanish in Group A (Table 2) over the short-term has been analyzed by performing three different approaches in estimating the shape and scale parameters of the Weibull distribution;

i. Quantile-Quantile Plot (Q-Q Plot) including calms and extreme wind speeds.
ii. Quantile-Quantile Plot (Q-Q Plot) excluding calms and extreme wind speeds.
iii. Method of Moments

For bullet points “i“ and “ii” above; calms indicate the wind speeds under 3 m/s and extreme wind speeds indicate the wind speeds over 25 m/s. where in these wind speed most of the turbines stops generating electricity, this is because either insufficient wind speed variability to rotate the rotor of the turbine or safety purposes [23].

Quantile-Quantile Plot that is fitted for the Machrihanish wind speed over the short-term data including calms and extreme wind speeds has given the best prediction of the shape and scale parameters of the Weibull distribution, however with an higher uncertainty of 0.73% in shape parameter and 0.21% in scale parameter than that of the analysis performed by excluding the calms and extremes, where these were found to be 0.22% and 0.02% for those shape and scale parameters, respectively, when the calms and extreme wind speeds excluded in fitting the data.

The reason that the higher uncertainty experienced when the calms and extreme wind speeds included in fitting is mainly extend to which the data is actually Weibull distributed so the higher the uncertainty is the less Weibull distributed the data are. Therefore, the less the calms and extreme wind speeds included, the prediction get closer to the actual measurement but the uncertainty increases, see Table 3 below.

	Q-Q Plot calms and extreme wind speeds excluded	**Q-Q Plot calms and extreme wind speeds included**
r^2	0.00%	0.00%
k	0.22%	0.73%
λ	0.02%	0.21%
Mean	0.02%	0.23%

Table 3: Uncertainty arisen from Q-Q Plot analysis when fitting the short-term Machrihanish data. r^2 is the correlation coefficient, k is the shape parameter, λ is the scale parameter and mean is the mean wind speed

The estimations that are gained by performing the Method of Moments has given the least realistic results. This maybe because this approach does not allow excluding the calms and extreme wind speeds, meaning that the actual equation between the measured and the theoretical moments of the Weibull distribution function assumed to be equating the theoretical moments of the Weibull distribution to the moments of actual dataset. Even though the Method of Moments is mathematically the most correct approach, the uncertainty found to be higher than that of the Q-Q Plot estimations. This balance between theoretical correctness and practical implementation is not well correlated when using the Q-Q Plot. However, better estimations of the shape and scale parameters of the Weibull distribution may be gained if the data can be fitted partially Weibull distributed.

2.3. Sector Pairing

Current implementation of MCP methodologies seem to be unrepresentative in most places in wind resource assessment process by underestimating a number of parameter, this would be explained due to establishing a separate relationship between the reference and target sites in terms of the direction sector pairing, which in turn results in a separate wind speed and directional sectors, in reality these are very much correlated in calculating the actual energy yield of a wind farm.

An alternative MCP methodology would accommodate the use of significant intrinsic direction sectors, which are related to the physical exposure at the reference and the target site at the same time and in the same direction when performing the analysis rather than putting a limit on directional sectors, which they are usually 8-12 sectors in the literature, nevertheless this model also allows establishing a relationship for bimodal distributions, meaning that separation of bimodal distribution is possible. In accordance to this approach the data for both the reference and target site has to be identified by pairing the directional sectors, this would then allow us to establish a better and more realistic relationship between the datasets. An MCP method that enables the available resource to be better represented would have the following implementations;

i. Dealing with meaningful intrinsic direction sectors, which are coupled with using physical exposures at the site rather than, for example; arbitrary 12 of 30-degree sectors. This allows the separation of bimodal distributions into unimodal distributions from physically significant direction sectors, meaning that more appropriate separation of the data interns of the direction sectors would be made with the available data.

ii. This approach could also be adopted segment the target and reference site data in daily or yearly basis and separation of these into diurnally and seasonally variability.

iii. The reference and target site data has to be identified by pairing the data into directional sectors, such as three 120-degree sectors and other segments such as on daily or yearly basis from the concurrent data over the short-term and target and reference data over the long-term.

iv. The relationship between the target and the reference site should be analyzed by using the "Modified Weibull Scaling Method", meaning that also the shape parameters of both target and reference site will be taken into account. An alternative to this empirical relationship derived by associating wind speeds of equal rank from the target and reference data subsets would also be used.

The correctness of the assumption of that the linear relationship between two Weibull distributions is invalid, which in reality characterized by an exponent α (equation 22 and 23 further below under the Modified Weibull Scaling Method headline), where this assumes that the target and reference site Weibull shape parameters are equal to each other. Though, in reality it is not and an error is incurred in this method based on linear regression if the shape parameters of the Weibull distributions at the target and reference site differ. In general it may not be establish an adequate relationship between the sites. Therefore, Modified Weibull Scaling Method or an empirical relationship based on associating similarity ranked target site speeds can be adopted so as to establish more reliable predictions of long-term wind speeds.

The approach mentioned in bullet point "iii" establishes the veer correction itself. For example if there are three 120-degree sectors, there will be nine possible datasets, the diagonal datasets would then be dataset 1, 5 and 9 and taking the rest of the datasets into account accommodates the veer correction. Therefore there is no need of an iterative process in terms of the veer correction in accordance to approach in item "iii" above.

Meanwhile, not putting adequate enough data into a single sector may be arisen issues of ending up insufficient amount of information; however, by reducing the number sector pairs and increasing the size of target and reference direction sectors might help to degrade this uncertainty. This would be done to avoid bimodal wind speed distributions.

The wind speed distributions at the target and reference site in each sector pair might not be described adequately well by a Weibull distribution, in alternative to this consequence, with an empirical approach the relationship between target and reference site would ultimately be described by observing that ranked wind speeds will coincide for well correlated data.

The Q-Q Plot method has been used as a correlation technique in fitting the data for each group of datasets, short-term data taken to be 15 months period so as to avoid that the length of data would not be sufficiently enough in each dataset to establish a reliable prediction of long-term wind speeds at each target site in every Group of data pair. In general, MCP methods are used in only terrestrial applications where the relationships between wind speeds can be more complicated with a low correlation coefficient. To assess the predictive capabilities of Linear Regression and Modified Weibull Scaling Method analysis and Q-Q Plot as a correlation technique are conducted using six data sets (Table 2). For each geographic group of target and reference site, sector pairing procedure has been applied for both over the short and long-term data, number of sectors taken as three 120-degree equally distributed sectors and Linear Regression and Modified Weibull Scaling Method has been applied to each nine datasets of the wind speeds on the basis of time stamps of the data to assess the predictive capability of these two MCP method, including MWSM. Long-term historical data accounts for 20 year of period for Groups B to F and 10 years for Group A. The stations in the first rows of each group in Table 2 taken to be the target site and predictions are made on these sites. Hourly wind speed and direction data are used for each site, there were no iterative action in sorting the data except there were a need of converting the wind speed from knots to m/s since the

MIDAS data logged as knots rather than m/s.

To sum up, Predictive capabilities performed by two MCP methods have been assessed in details. The Weibull Scaling Method Method is modified from the literature by making arrangements in Weibull Scaling parameters.

2.4. Modified Weibull Scaling Method

Uncertainties with respect to Weibull Scaling Method arisen from conventional approach within the general use of the method that has been explained in Literature Review chapter under Weibull Scaling Method title. It is clear that the relationship between reference and target site is actually non-linear, an modifications in this method would be possible by including target and reference Weibull distribution shape parameters as shown in the equations 22 and 23 below [16];

$$k_x / k_y = \alpha = k_x / k_y \Leftrightarrow k_\eta = (k_\xi\, k_y)/k_x \tag{22}$$

Hence:

$$\lambda_\eta = (\lambda_\xi\, \lambda_y)^\alpha / \lambda_x{}^\alpha \tag{23}$$

Where k_x, k_y and λ_y λ_x are the shape and scale parameters of the reference and target site respectively whilst the short term campaign, and k_ξ, k_η *and* λ_ξ, λ_η are the shape and scale parameters of the reference and target site respectively whilst the long-term wind measurement campaign. The key consideration here is not to equate the shape parameters of the reference and target site as it has been reviewed under the Weibull Scale Method previously.

Weibull scaling similar to the method described above has previously been discussed, though the possibility of the same target and reference Weibull distribution is only exists where the sites are identical, unless specified otherwise, the effect of Weibull parameters in resulting analysis are underestimated in long-term wind speed predictions in the current literature by assuming that the relationship between the shape parameters is equal to unity [3].

2.4.1. True Relationship Between Weibull Distributions

In order the perform that the true relationship is not actually linear, from $y=\lambda_y(x/\lambda_x)^{kx/ky}$, the following equation can be drawn [17]:

$$y = mx^{\alpha} + c \tag{24}$$

In this case α, which is the ratio of the shape parameters and can be defined as follows [17];

$$\alpha = k_x/k_y \tag{25}$$

and

$$m = \lambda_y \lambda_x^{-\alpha} \tag{26}$$

This implies that the relationship between the Weibull distributions will be linear if the shape parameters are the same at both the target and reference sites.

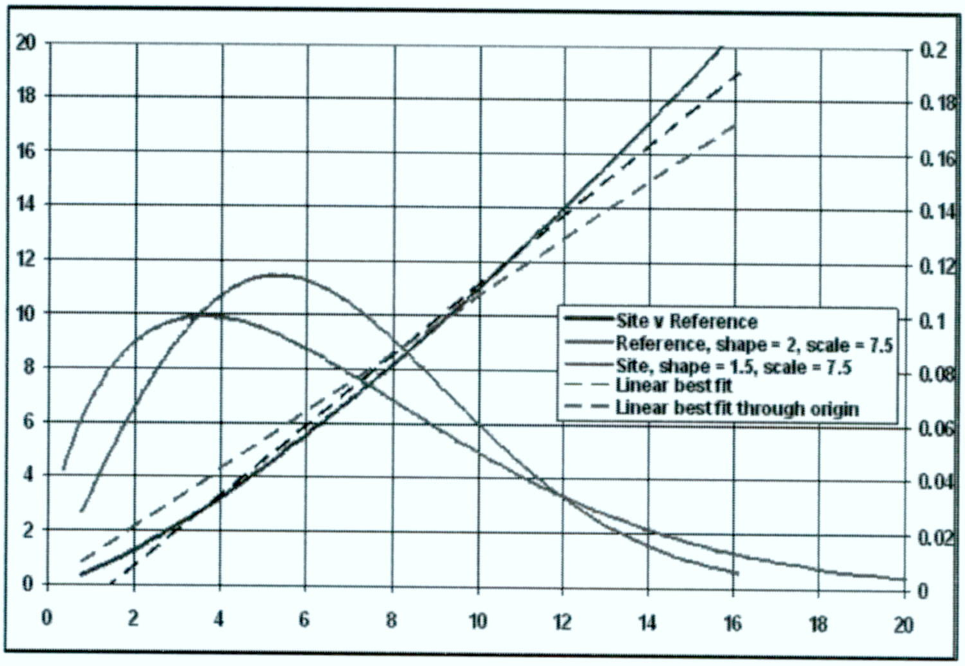


Figure 8: Establishing a relationship between 2 Weibull distributions using with power law with a straight line, the numbers in the *x-axis* indicated the scale parameters, the numbers in *y-axis* indicated the shape parameters and the numbers from 0.02 to 0.2 probability [16].

In the Figure 8 on the previous page, reference site Weibull distribution shown as red in colour and the target is in green, In this case the shape parameter of the reference site is greater than the target one. The true relationship between two distributions has been shown in blue line whereas the linear best fit has been shown as a dashed black line. Therefore, it can be seen that the true relationship is clearly non-linear. The red dashed forced to pass from the origin so as to represent the best fit through the origin. It could be concluded that the fitting is prone to a systematic error if the shape parameters of the both site are not identical [16].

3. RESULTS

The results of the analysis are presented in this chapter and explanation, interpretation, and analysis of the results are reserved for the Discussion chapter.

3.1. Data Summary

Geographical locations and the data pairs used in the analysis are presented in Figure 9, page 28 and Table 2, page 20 and Table 4, page 29. Results of Weibull parameter estimation capability of Q-Q Plot and Method of Moment are represented in Figure 10, 11 and 12. Tables showing the predictive capability of the Linear Regression and Modified Weibull Scaling method and their associated total weighted errors (TWE) for Group A, B, C and D are represented in Table 8 to 11, page 24 and 25. Scatter plots over the short-term are also presented (Figure 26 to 29, page 56 and 57). Other tables and figures for the remaining groups, Group E and F, are represented in Appendix A, B and C.

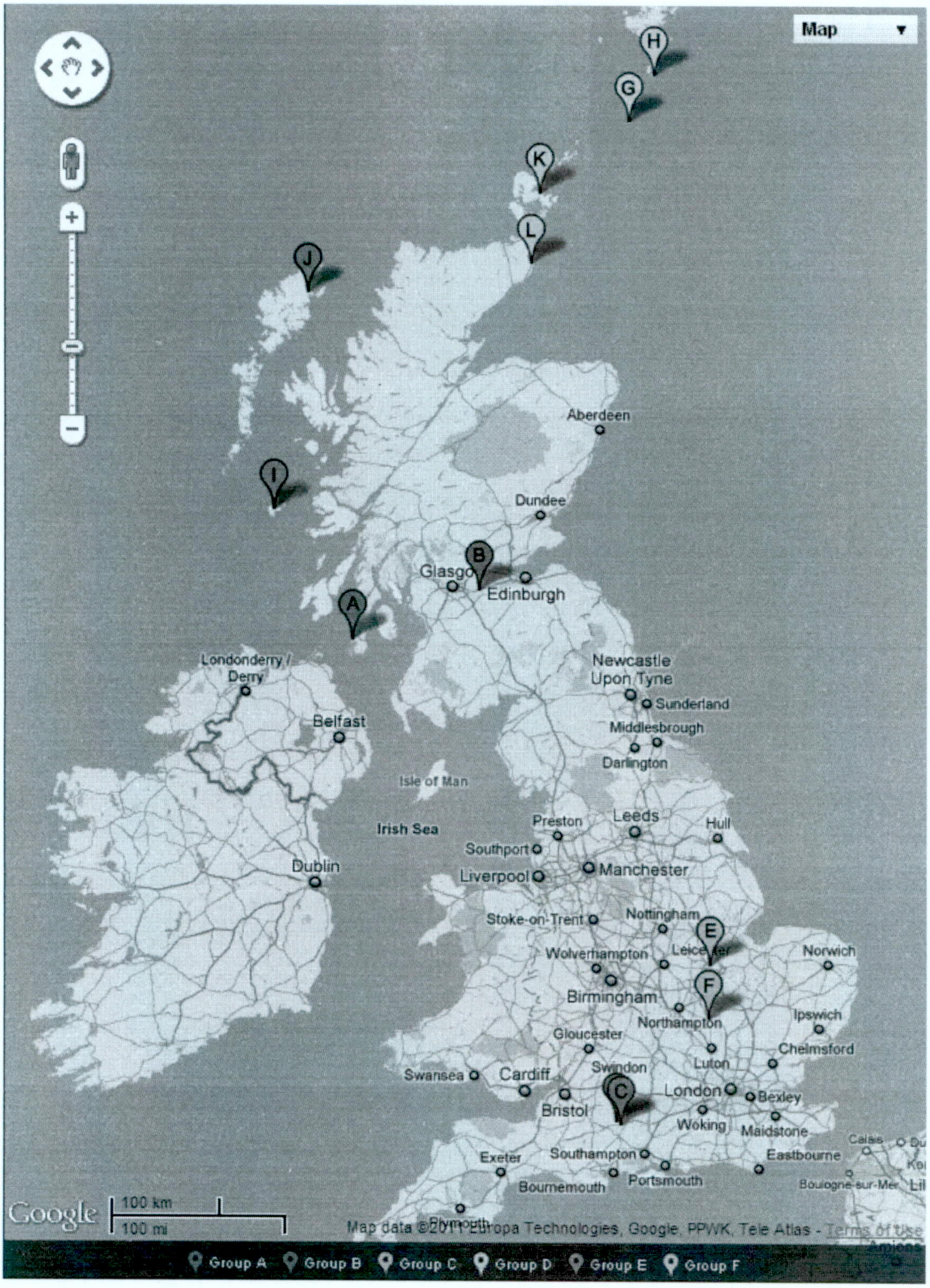


Figure 9: Map of the UK showing the 6 pairs of weather stations (Station D locates right behind the Station C). Coloured pinpoints represent individual pairs.

Dataset pairs	Map pinpoints	Geographic area	Station name
Group A	A	ARGYLL	MACHRIHANISH
	B	NORTH LANARKSHIRE	SALSBURGH
Group B	C	WILTSHIRE	BOSCOMBE DOWN
	D	WILTSHIRE	LARKHILL
Group C	E	CAMBRIDGESHIRE	WITTERING
	F	BEDFORDSHIRE	BEDFORD
Group D	G	SHETLAND	FAIR ISLE
	H	SHETLAND	SUMBURGH
Group E	I	ARGYLL	TIREE
	J	WESTERN ISLES	STORNOWAY AIRPORT
Group F	K	ORKNEY	KIRKWAL
	L	CAITHNESS	WICK AIRPORT

Table 4: Name of the stations that are showed with letters on the map

3.2. Estimation of Weibull Distribution Parameters Analysis

In the analysis it has been observed that the predictions have given the best accuracy for both shape and scale parameters of the Weibull distribution when the data fitted by performing Q-Q Plots with including the calms and extreme wind speeds. Method of Moments performed better than the Q-Q Plot predictions when the calms and extremes excluded, see Figure 10, 11 and 12, page 30 and 31.

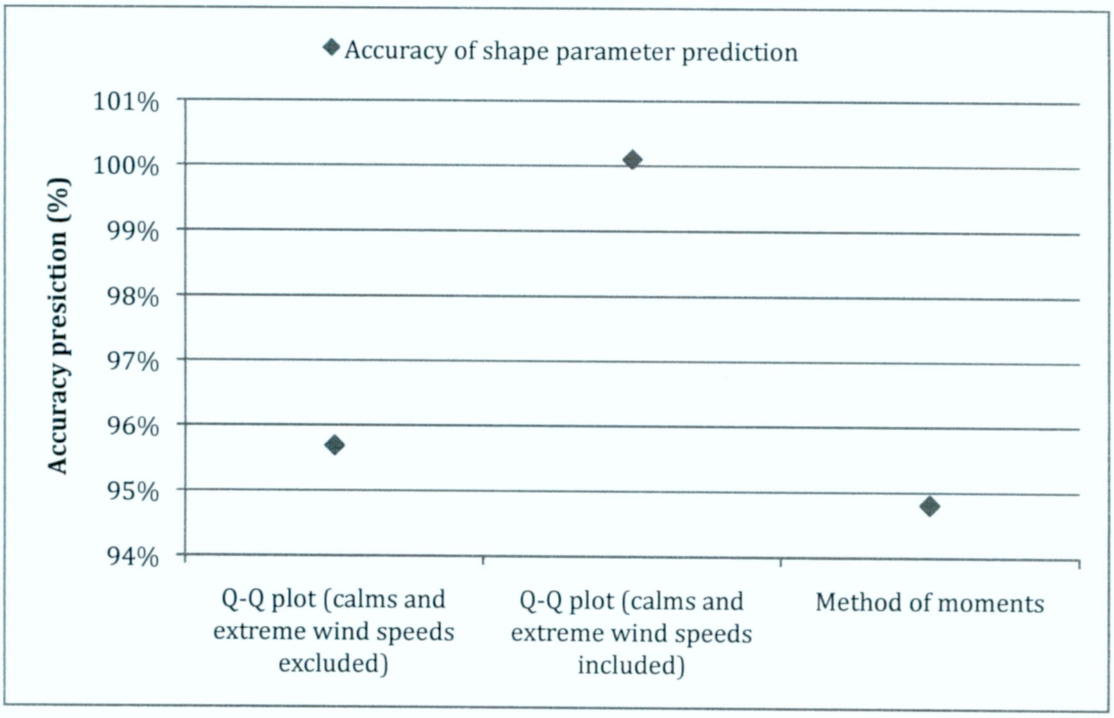


Figure 10: Shape parameter accuracy analysis of three different methods for Machrihanish wind speeds

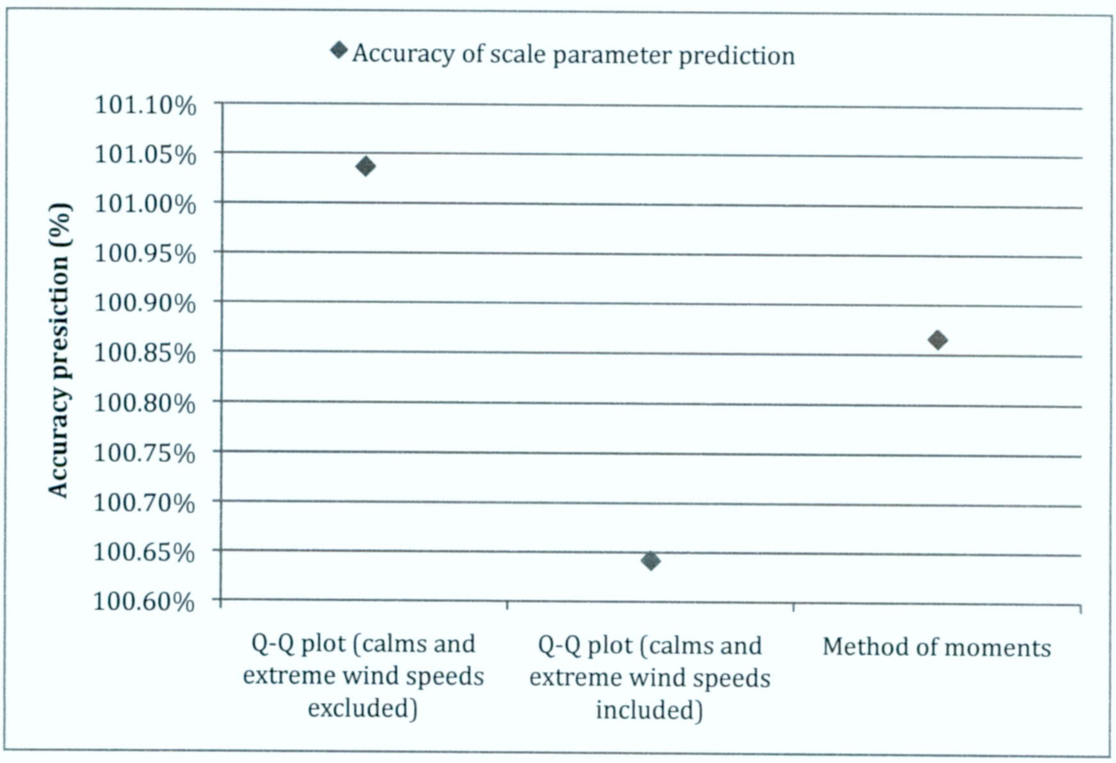


Figure 11: Scale parameter accuracy analysis of three different methods for Machrihanish wind speeds

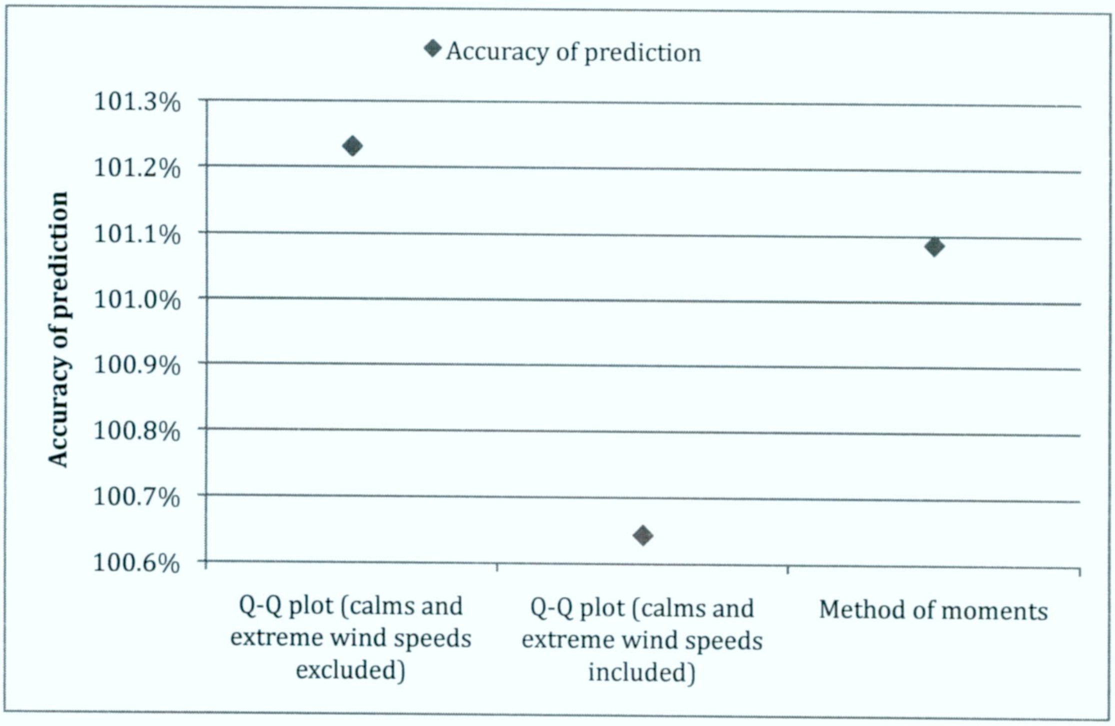


Figure 12: Accuracy analysis of three different methods for Machrihanish wind speeds

3.3. Sector Pairing and Modified Weibull Scaling Method Analysis

The MCP methodology is evaluated on all paired data sets using both Linear Regression and Modified Weibull Scaling Method, which are used as the correlation techniques in this study and Quantile versus Quantile Plots used in estimation of the shape and scale parameters of the Weibull distribution. Their performance is quantified by calculating the accuracy of prediction and prediction errors, which are calculated for each individual group datasets so as to calculate the total weighted error (TWE) and sector wise variability of the prediction error, calculated as the weighted standard deviation of the sector prediction error from its weighted mean for both Linear Regression and Modified Weibull Scaling Method and comparisons made between them. This process followed the same for each geographical Group A to F. Though only Group A to D results are represented under the Results chapter in Table 8 to 11, page 34 and 34 and Figure 13 to 20, page 46 to 40 and comparisons between all groups represented in a single figure (Figure 21, page 40) and remaining results for Group E and F are represented in Appendix B, however their Weibull distribution parameter estimation analysis by using Linear Regression and Modified Weibull Scaling Method are also included under the Results chapter, Figure 23, 24 and 25, page 42, 43 and 44. Scatter graphs showing the correlation coefficients and equation of the straight line for all group pair (Group A to F) by plotting a straight line between the reference target site wind speed over the short-term represented and can be find in Appendix A, page 56.

Number of direction sectors are taken to be three equally distributed sectors, therefore there will be nine possible datasets including three target and three reference sector, where each of them covers a 120-degree space. This allowed us to sort the data into 9 distinct datasets, meaning that we have concurrent target and reference data for each dataset; this is followed by establishing Linear Regression and Modified Weibull Scaling relationship from the reference site dataset over the long-term, which in turn nine different predictions correspond to each dataset are analyzed; for example dataset one (Northerly-Northerly sector) is the prediction of the long-term target wind speeds when the wind blows at the same time and in the same direction at both target 1 (Northerly sector) sector and reference one (Northerly sector) sector (Table 5, 6 and 7, page 32 and 33), the corresponding datasets tabulated in Table 5, page 32, the numbers in the cells indicates wind speeds.

Sector pairing tool has been developed via using Microsoft Excel, for example 10 times plus sector one from the reference sector and the target sector gives a 2 digit number, a unique value that put on the look up table to determine corresponding datasets (Table 6 and 7).

	Reference site sector 1	**Reference site sector 2**	**Reference site sector 3**
Target site sector 1	$4.893_{\text{Dataset 1}}$	$4.505_{\text{Dataset 2}}$	$8.977_{\text{Dataset 3}}$
Target site sector 2	$3.482_{\text{Dataset 4}}$	$6.136_{\text{Dataset 5}}$	$7.763_{\text{Dataset 6}}$
Target site sector 3	$4.159_{\text{Dataset 7}}$	$6.056_{\text{Dataset 8}}$	$6.324_{\text{Dataset 9}}$

Table 5: Sector pairing results for each individual dataset.

Sector	Minimum	Maximum
1 (Northerly sector)	300°	60°
2 (Easterly sector)	60°	180°
3 (Westerly sector)	180°	300°

Table 6: Showing the range of each sector that has been adopted for the analysis.

Target Sector	Reference Sector	Dataset
1 (North)	1 (North)	D1 (NN)
1 (North)	2 (East)	D2 (NE)
1 (North)	3 (West)	D3 (NW)
2 (East)	1 (North)	D4 (EN)
2 (East)	2 (East)	D5 (EE)
2 (East)	3 (West)	D6 (EW)
3 (West)	1 (North)	D7 (WN)
3 (West)	2 (East)	D8 (WE)
3 (West)	3 (West)	D9 (WW)

Table 7: Corresponding datasets of sectors.

Each dataset populated within times (Dataset 1 to 9) shows that how frequent the data is, therefore, counts in the diagonal cells (Dataset 1, 5 and 9 in Table 5, page 32) are higher than the data counts in the off diagonal cells and represents the wind speeds at the same time, direction and sector. The counts in the off diagonal cells are important because they systematically establish the veer correction, if there were no counts in the off diagonal cells there would be no veer correction, therefore. This method accommodates the need for the veer correction itself. The idea of sector pairing is to reduce the uncertainty in the sense of trying to exclude the data that are not correlated. Therefore, not dividing the data based on sector pairing, meaning that including the concurrent data examples of different relationships, would accommodate scatter in the data. Sector pairing method is trying to separate these two different relationships and in theory more accurate prediction of the long-term wind speeds should be experienced, practical implementation of this approach has confirmed the theoretical assumption (Table 8 to 13, page 34, 35 and 59).

Datasets	wind speed (m/s)	LRM predictions (m/s)	MWSM predictions (m/s)	Accuracy prediction for LRM (%)	Accuracy prediction for MWSM (%)	Prediction error for LRM (%)	Prediction error for MWSM (%)	Count of data
D1	5.67	6.68	4.89	117.82%	86.28%	-17.82%	13.72%	1585
D2	5.32	6.39	4.51	120.14%	84.69%	-20.14%	15.31%	4183
D3	6.28	7.22	8.98	115.07%	142.99%	-15.07%	-42.99%	7629
D4	4.94	2.62	3.48	53.06%	70.44%	46.94%	29.56%	2867
D5	5.99	8.61	6.14	143.76%	102.43%	-43.76%	-2.43%	7591
D6	5.43	6.48	7.76	119.21%	142.89%	-19.21%	-42.89%	14674
D7	6.01	5.22	4.16	86.87%	69.17%	13.13%	30.83%	2673
D8	6.11	4.52	6.06	74.06%	99.20%	25.94%	0.80%	7008
D9	7.03	9.45	6.32	134.39%	89.98%	-34.39%	10.02%	13342
					TWE	-15.42%	-9.48%	

Table 8: Analysis fo Group A dataset by using both Modified Weibull Scaling Method and Linear Regression Method. Total weighted error shows how well the methods perform.

Datasets	Measured wind speed (m/s)	LRM predictions (m/s)	MWSM predictions (m/s)	Accuracy prediction for LRM (%)	Accuracy prediction for MWSM (%)	Prediction error for LRM (%)	Prediction error for MWSM (%)	Count of data
D1	3.92	4.46	3.73	113.74%	95.15%	-13.74%	4.85%	44180
D2	3.26	5.05	3.52	155.03%	108.08%	-55.03%	-8.08%	4681
D3	2.87	1.72	1.93	59.80%	67.16%	40.20%	32.84%	1887
D4	1.62	2.85	2.68	175.37%	164.90%	-75.37%	-64.90%	2035
D5	3.59	4.88	3.30	136.08%	92.13%	-36.08%	7.87%	34318
D6	2.79	2.86	2.22	102.56%	79.64%	-2.56%	20.36%	2116
D7	2.52	1.51	2.62	60.11%	104.17%	39.89%	-4.17%	8635
D8	2.65	2.11	2.02	79.68%	76.20%	20.32%	23.80%	2851
D9	4.46	3.49	3.74	78.24%	83.88%	21.76%	16.12%	80867
					TWE	-0.13%	9.67%	

Table 9: Analysis fo Group B dataset by using both Modified Weibull Scaling Method and Linear Regression Method. Total weighted error shows how well the methods perform.

Datasets	Measured wind speed (m/s)	LRM predictions (m/s)	MWSM predictions (m/s)	Accuracy prediction for LRM (%)	Accuracy prediction MWSM (%)	Prediction error for LRM (%)	Prediction error for MWSM (%)	Count of data
D1	4.72	4.60	4.64	97.39%	98.29%	2.61%	1.71%	38880
D2	3.91	2.39	3.07	61.18%	78.44%	38.82%	21.56%	3463
D3	3.93	4.39	3.45	111.70%	87.75%	-11.70%	12.25%	3674
D4	3.09	2.91	3.45	94.15%	111.51%	5.85%	-11.51%	3742
D5	4.23	3.78	4.39	89.19%	103.59%	10.81%	-3.59%	21970
D6	3.16	2.95	1.86	93.28%	58.78%	6.72%	41.22%	3135
D7	3.94	3.97	3.86	100.91%	97.96%	-0.91%	2.04%	6472
D8	4.19	3.44	4.19	82.09%	100.00%	17.91%	0.00%	2887
D9	5.13	4.98	4.85	97.11%	94.48%	2.89%	5.52%	77115
					TWE	4.60%	3.92%	

Table 10: Analysis fo Group C dataset by using both Modified Weibull Scaling Method and Linear Regression Method. Total weighted error shows how well the methods perform.

Datasets	Measured wind speed (m/s)	LR predictions (m/s)	MWSM predictions (m/s)	Accuracy prediction for LRM (%)	Accuracy prediction MWSM (%)	Prediction error for LRM (%)	Prediction error for MWSM (%)	Count of data
D1	7.40	5.50	7.96	74.36%	107.62%	25.64%	-7.62%	28699
D2	5.94	5.78	5.44	97.24%	91.49%	2.76%	8.51%	1292
D3	5.80	5.67	5.72	97.82%	98.71%	2.18%	1.29%	1486
D4	4.99	1.68	3.97	33.58%	79.46%	66.42%	20.54%	4094
D5	7.78	7.43	7.78	95.51%	100.01%	4.49%	-0.01%	28820
D6	7.04	7.91	7.32	112.39%	103.97%	-12.39%	-3.97%	4118
D7	6.77	2.86	6.97	42.26%	103.08%	57.74%	-3.08%	5773
D8	6.48	6.00	6.33	92.65%	97.69%	7.35%	2.31%	1967
D9	8.35	7.83	8.85	93.79%	106.01%	6.21%	-6.01%	38560
					TWE	14.63%	-3.34%	

Table 11: Analysis fo Group D dataset by using both Modified Weibull Scaling Method and Linear Regression Method. Total weighted error shows how well the methods perform.

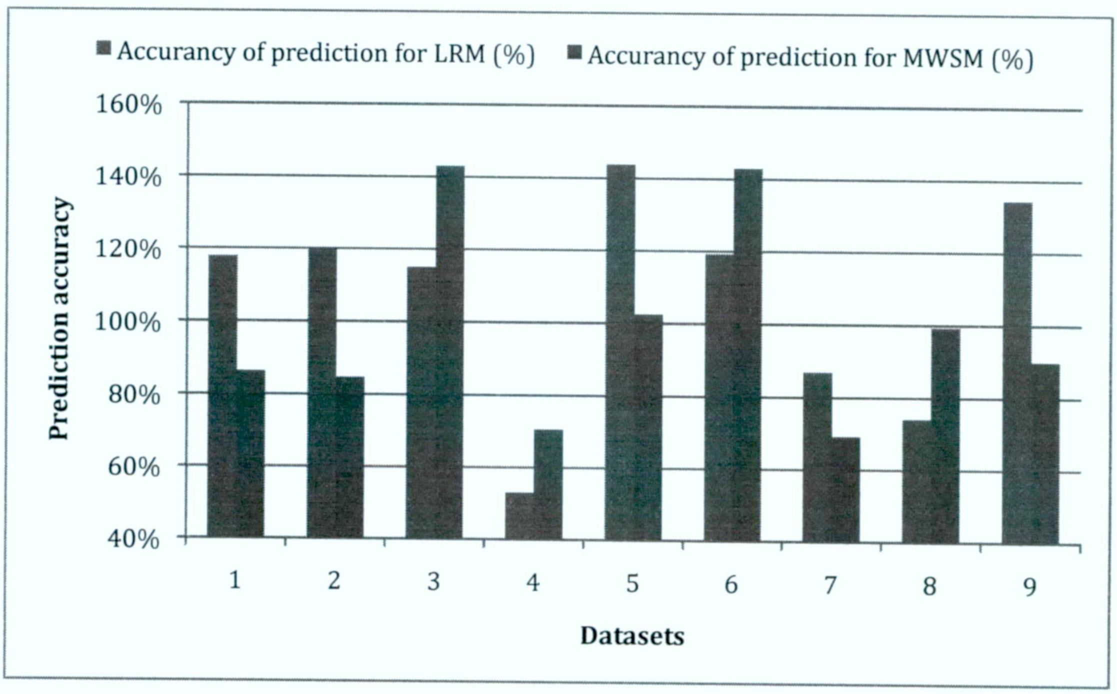


Figure 13: Group A prediction accuracy comparison betweern Linear Regression and Modified Weibull Scaling Method results

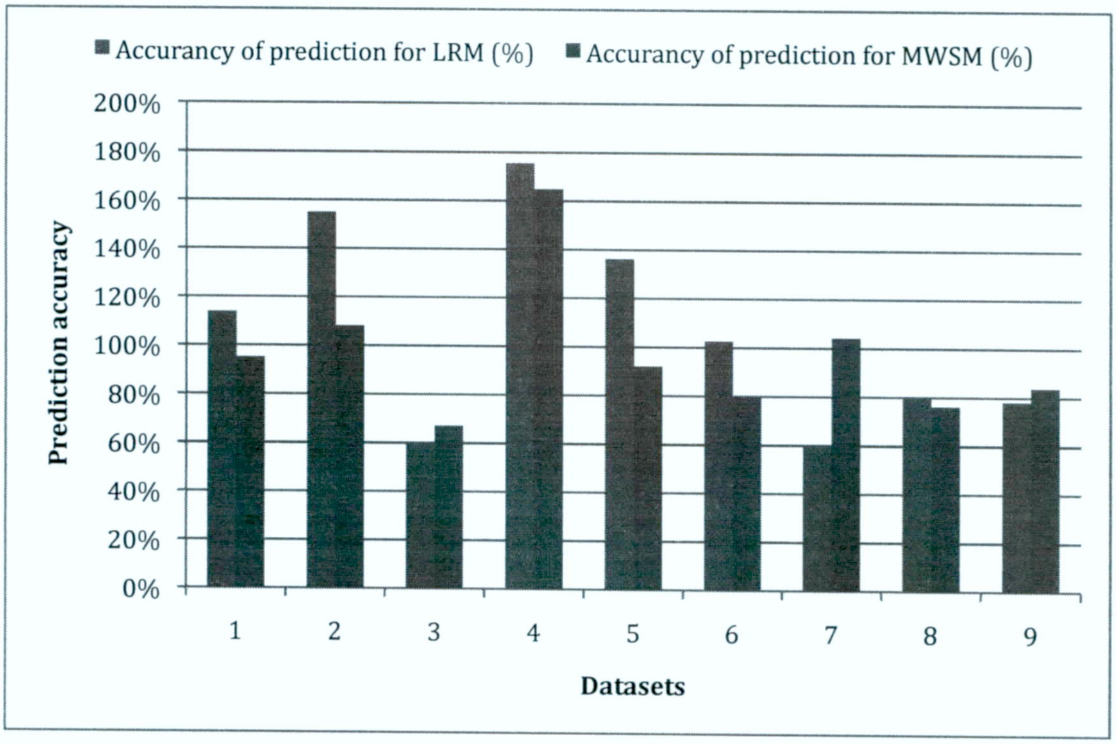


Figure 14: Group B prediction accuracy comparison betweern Linear Regression and Modified Weibull Scaling Method results

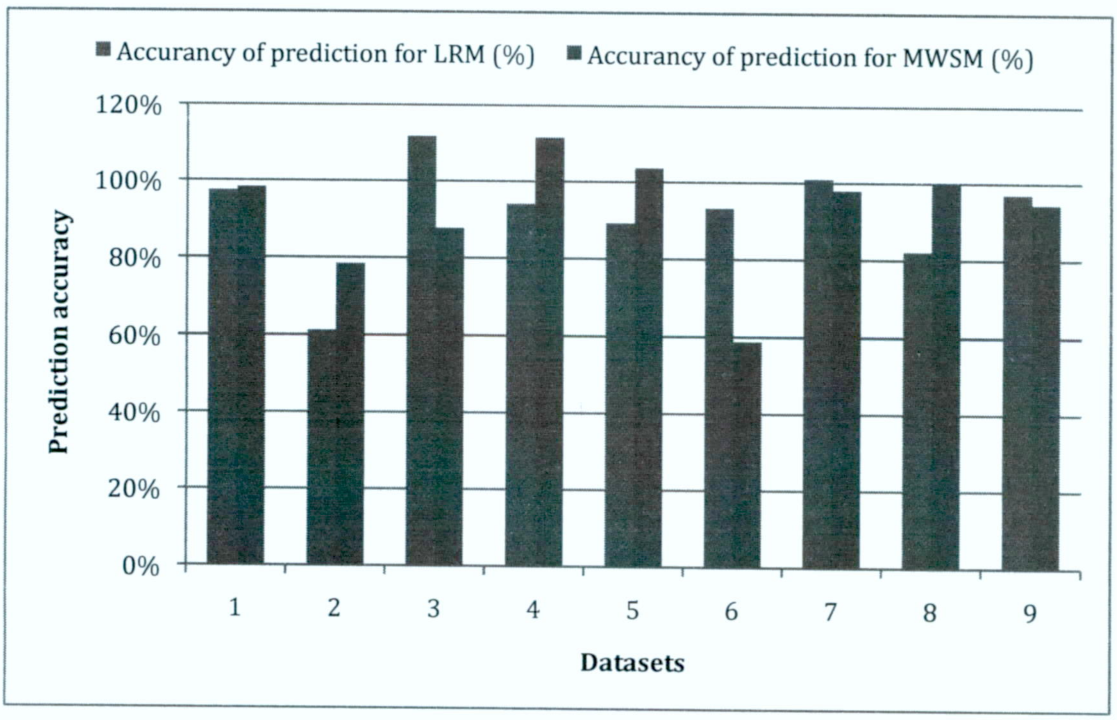


Figure 15: Group C prediction accuracy comparison betweern Linear Regression and Modified Weibull Scaling Method results

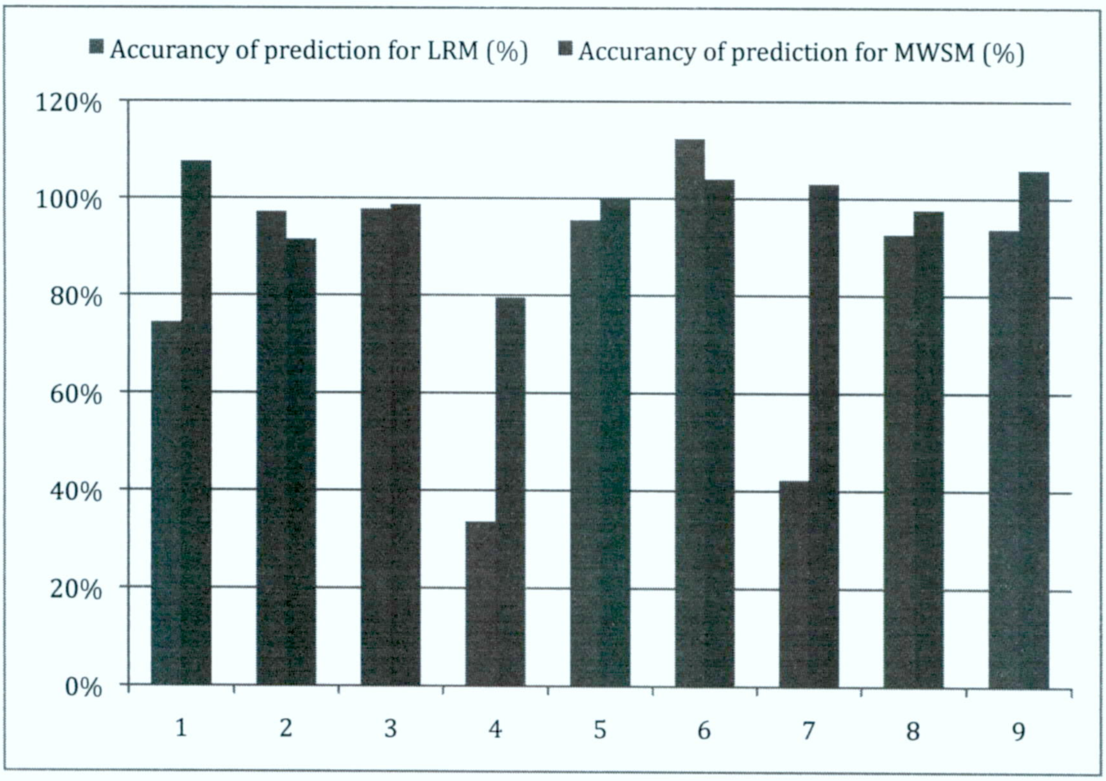


Figure 16: Group D prediction accuracy comparison betweern Linear Regression and Modified Weibull Scaling Method results

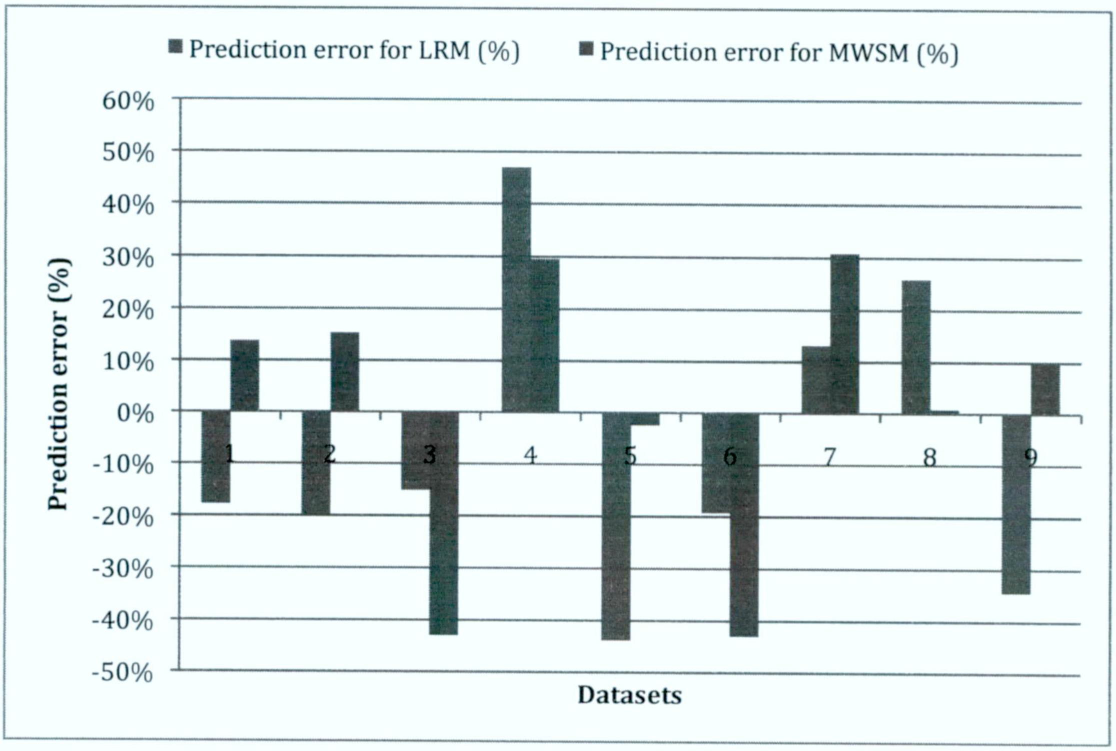


Figure 17: Group A prediction error comparison betweern Linear Regression and Modified Weibull Scaling Method results

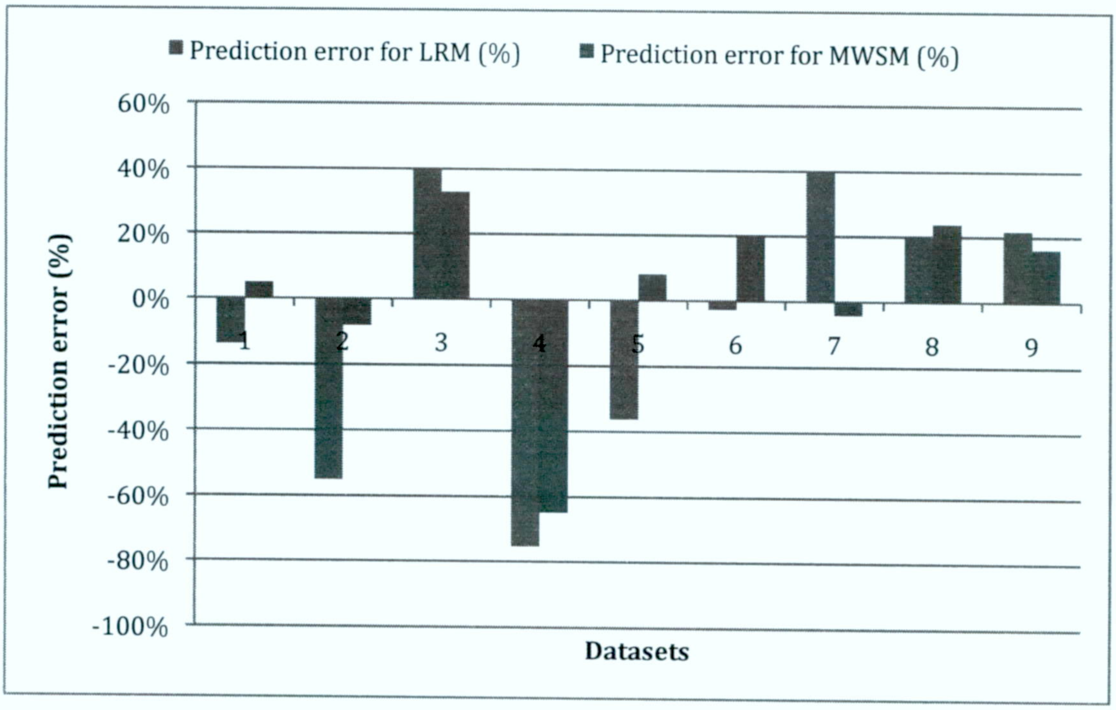


Figure 18: Group B prediction error comparison betweern Linear Regression and Modified Weibull Scaling Method results

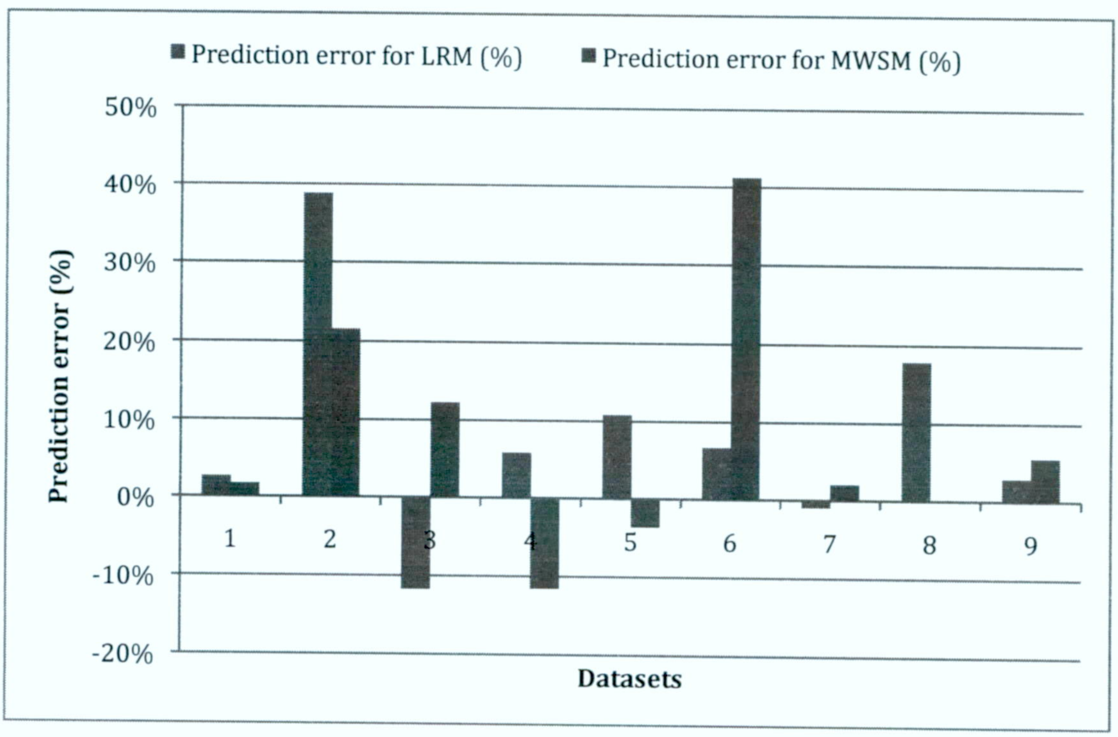


Figure 19: Group C prediction error comparison betweern Linear Regression and Modified Weibull Scaling Method results

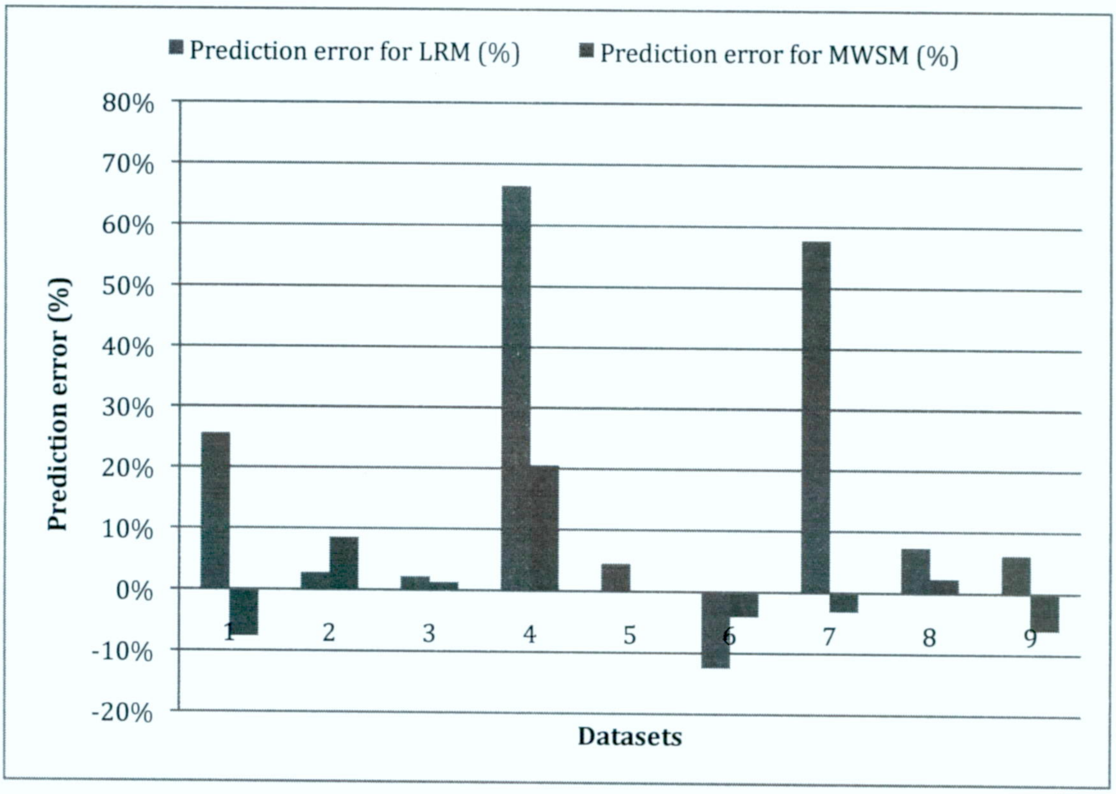


Figure 20: Group D prediction error comparison betweern Linear Regression and Modified Weibull Scaling Method results

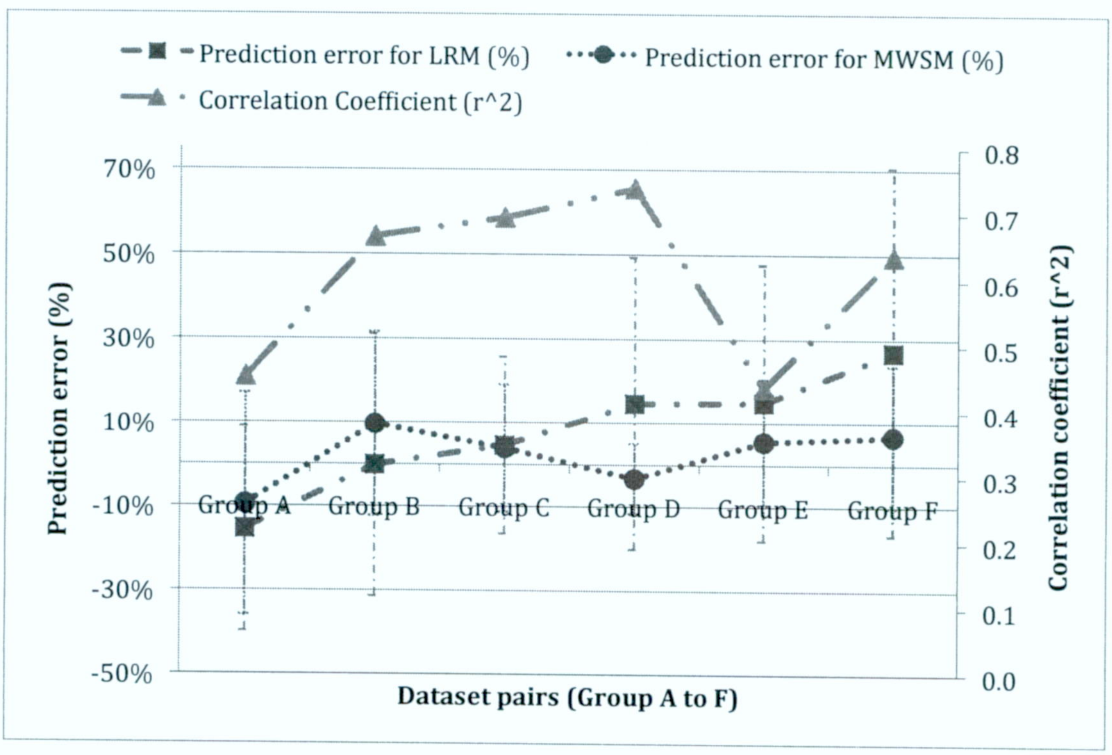
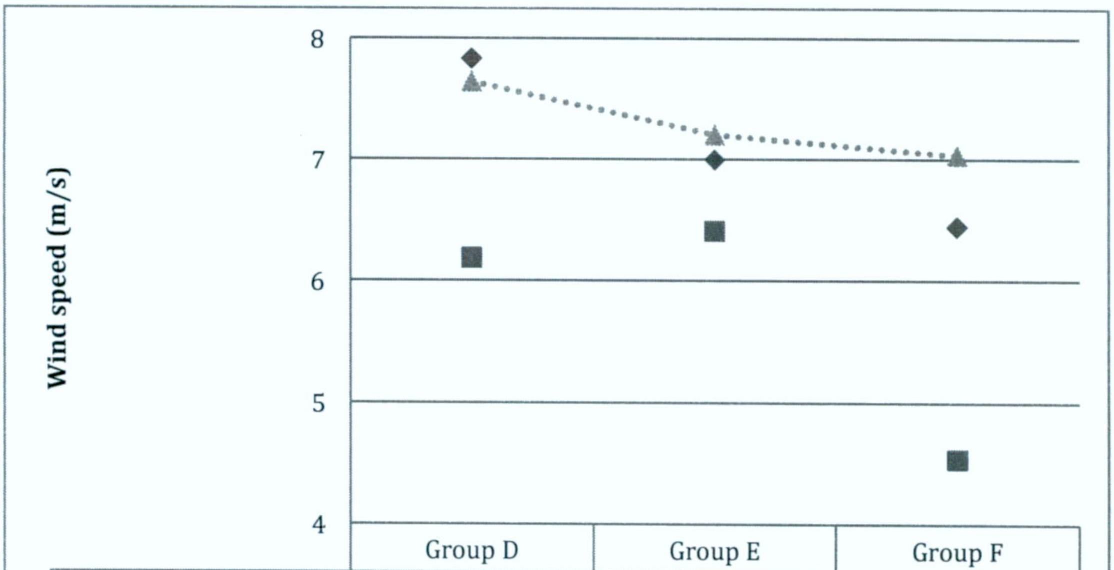
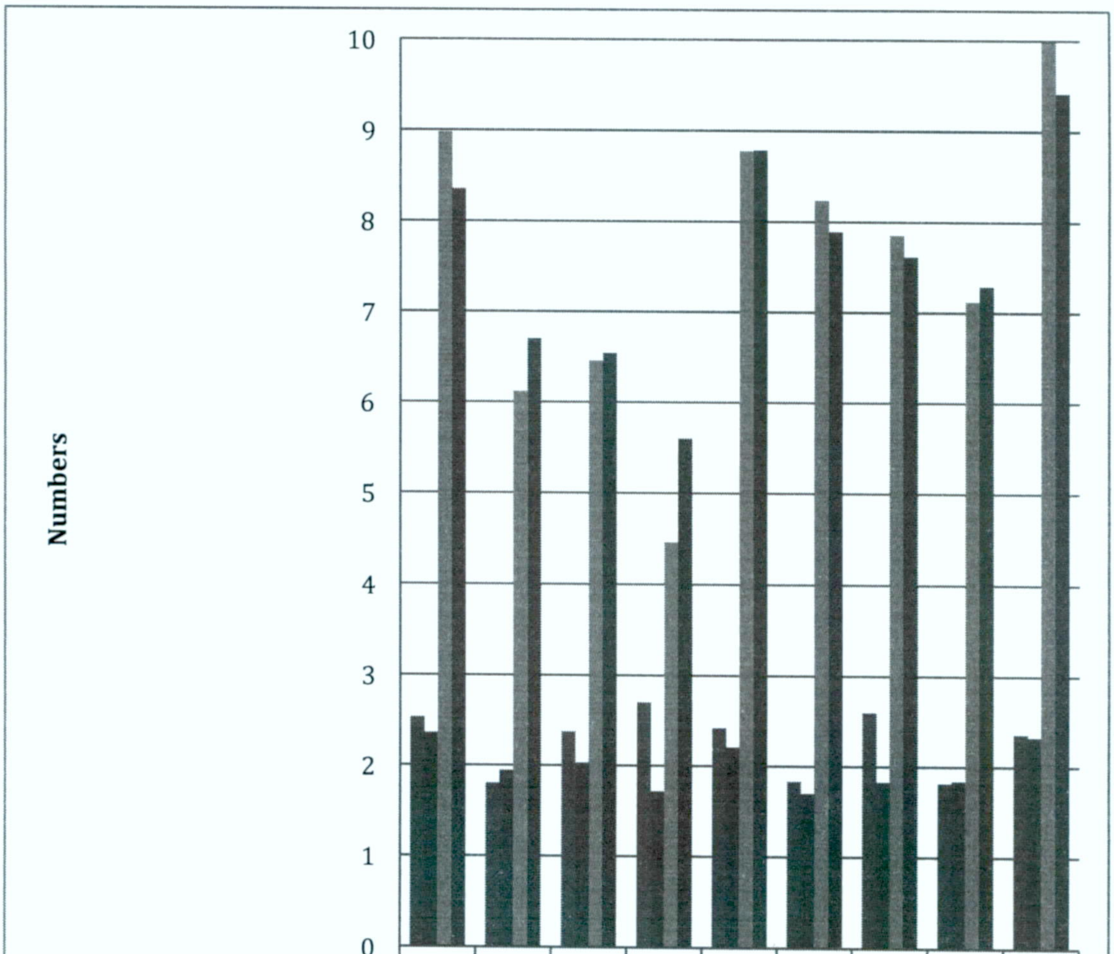


Figure 21: Comparison between Group`s correlation coefficient and predictive capability of Linear Regression and Modified Weibull Scaling Method



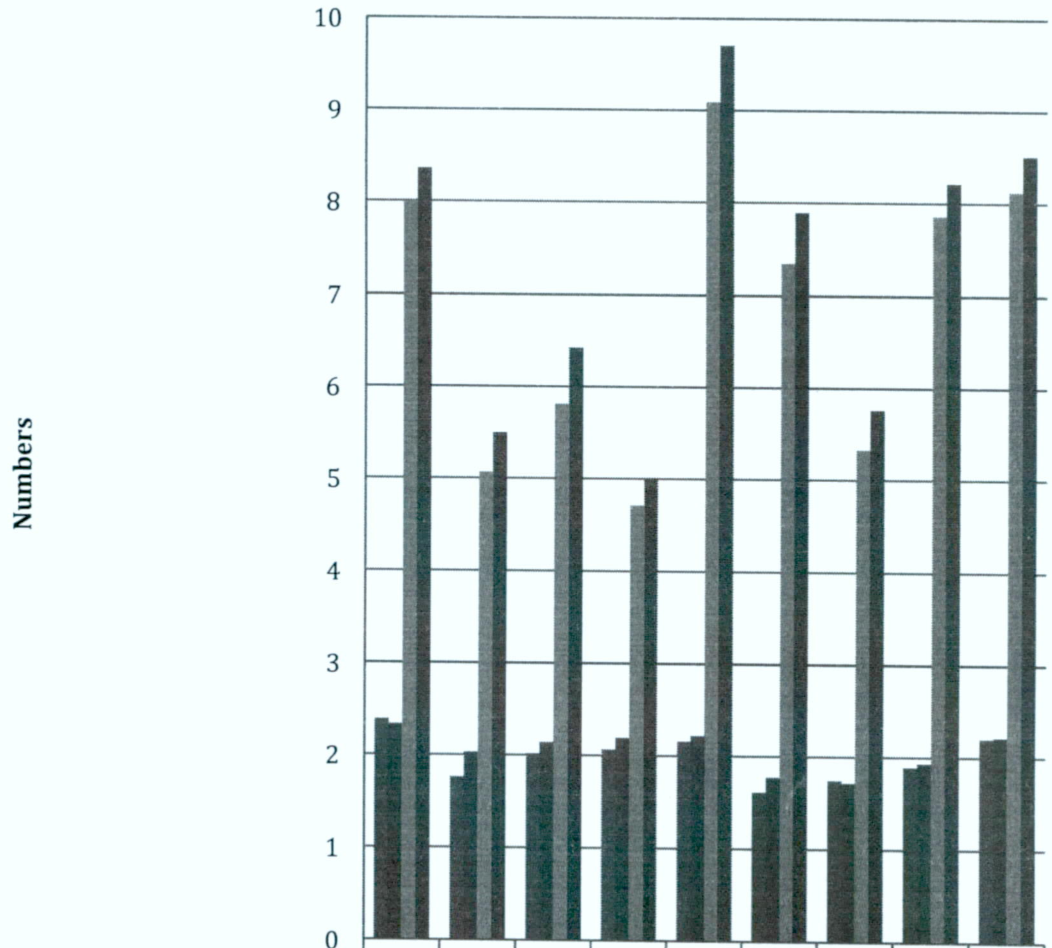
	Group D	Group E	Group F
LR prediction	6.187	6.408	4.536
MWSM prediction	7.829	6.997	6.443
Measured	7.638	7.203	7.033

Figure 22: Comparison between Linear Regression and Modified Weibull Scaling Method without sector pairing



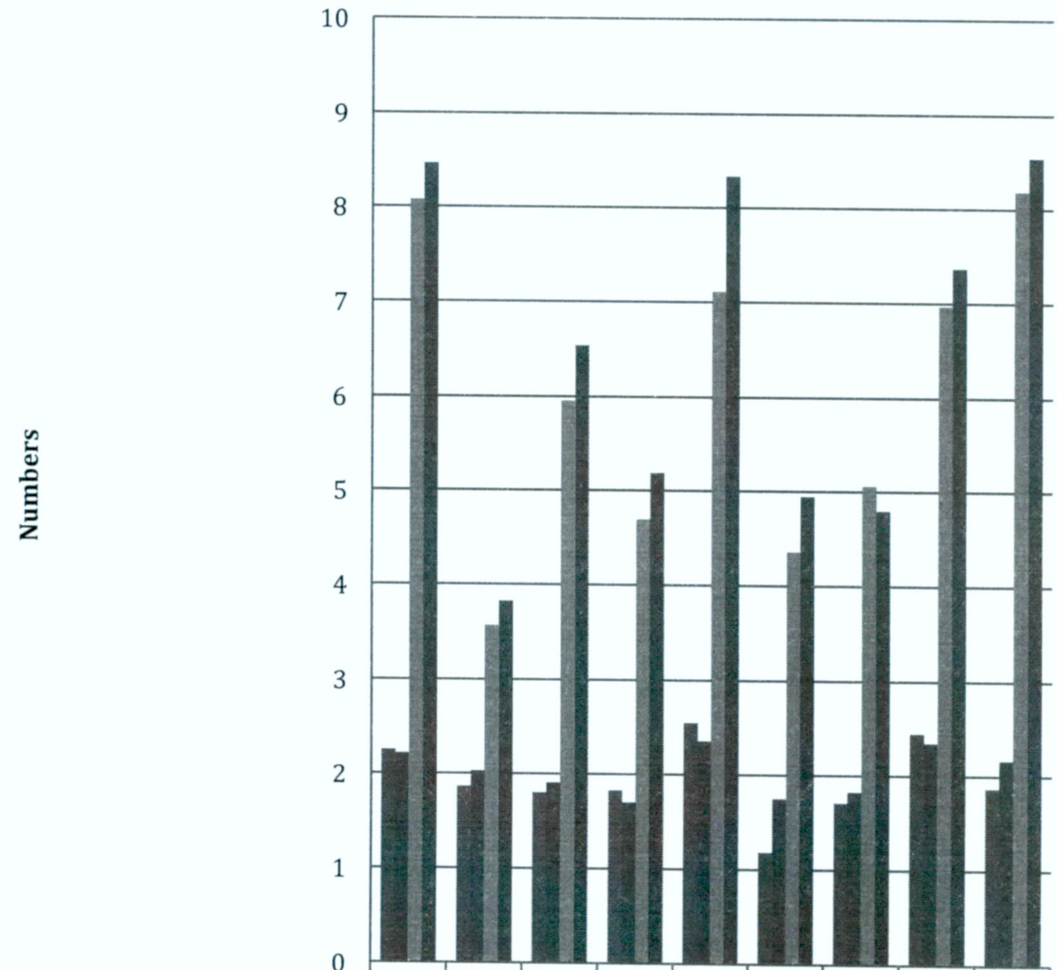
	D1	D2	D3	D4	D5	D6	D7	D8	D9
Predicted shape parameter	2.531	1.805	2.377	2.703	2.423	1.832	2.597	1.814	2.358
Measured shape parameter	2.361	1.943	2.031	1.715	2.207	1.698	1.832	1.843	2.326
Predicted scale parameter	8.973	6.115	6.456	4.462	8.774	8.237	7.853	7.118	9.983
Measured scale parameter	8.349	6.701	6.542	5.599	8.783	7.888	7.615	7.291	9.419

Figure 23: Shape and scale parameter predictive capability of Modified Weibull Scaling Method for Group D dataset pairs



	D1	D2	D3	D4	D5	D6	D7	D8	D9
Predicted shape parameter	2.383	1.758	2.014	2.064	2.153	1.606	1.739	1.887	2.199
Measured shape parameter	2.332	2.028	2.135	2.187	2.217	1.773	1.712	1.928	2.212
Predicted scale parameter	7.999	5.064	5.808	4.708	9.090	7.350	5.327	7.860	8.123
Measured scale parameter	8.358	5.494	6.429	5.009	9.709	7.892	5.761	8.211	8.509

Figure 24: Shape and scale parameter predictive capability of Modified Weibull Scaling Method for Group E dataset pairs



	D1	D2	D3	D4	D5	D6	D7	D8	D9
Predicted shape parameter	2.250	1.858	1.796	1.824	2.543	1.176	1.704	2.446	1.866
Measured shape parameter	2.212	2.028	1.904	1.698	2.353	1.747	1.822	2.347	2.163
Predicted scale parameter	8.069	3.572	5.953	4.695	7.110	4.359	5.062	6.967	8.184
Measured scale parameter	8.459	3.830	6.540	5.187	8.331	4.945	4.796	7.367	8.541

Figure 25: Shape and scale parameter predictive capability of Modified Weibull Scaling Method for Group F dataset pairs

4. DISCUSSION

4.1. Data Summary

The meteorological monitoring sites chosen for this analysis represents a wide range of wind characteristics in both wind speed and direction and different topographical characteristics. Group A and E station pairs are furthest apart from each other with a distance of over 100km and located in entirely different terrain types and both have very different physical exposures of wind speeds. Outside the exemption of Group B and C target mean wind speeds, which are 3.08m/s and 4.03 m/s, respectively, all of the target site wind speeds have good mean wind speed variations and the highest mean wind speed experienced at Group D target site with a magnitude of 6.73 m/s. These wind speeds represents the mean wind speed variability at the target sites over the short-term. Majority of these high wind speeds can be attributed at least in some part to the height of the monitoring equipment. Nevertheless, the variety of wind speed magnitude in the data sets and their quality of record reflect the true nature of the Measure-Correlate-Predict methodology, meaning that, data measurement campaigns prone to instabilities and have mechanical aspect that may fail in certain cases such as freeze and control circuit failures.

There are also instabilities in the available data with missing data series. Ideally, neither the reference nor target sites would have gaps in their period of record, however, in the light of including of such a deficiency in this analysis also serves to reflect the true nature of long-term monitoring stations. Outages frequently would occur, especially at sites close to the shore. If the period of record is sufficiently long, however, the presence of gaps should not have a large impact in the use of the MCP method, unless the gaps disproportionately favor certain seasons of the year, where the effect of this in resulting can be seen in Group A wind speed predictions and in sanity check in figure 39 Appendix D. In particular in Group A data sets, where around infrequent of two and half years of data is missing at both reference and target site, this has been resulted in around 3 years displaced data after applying the sector pairing procedure. This might be one of the reasons why that the Modified Weibull Scaling method did not perform very well from the point of the theoretical assumption, which implies there has to be persistent relationship between the target and reference site wind speed data when the directional sectors paired and the correlation method is applied. Another reason would be locating entirely different terrains, where Machrihanish wind speeds characterized by the Atlantic wind speed patterns and Salsburgh located in such a dedicated area between Edinburgh and Glasgow, UK. In Group A wind speeds there is 1.5 year data in target site and 1 year data in reference site missing and these occurs in different time periods, even though 10 years data used in Group A analysis, there were around 7 years data remained after sector pairing procedure, outside the exemption of this the theoretical assumption conform to the practical implementation perfectly for Groups B to F. For Group A analysis, the data counts in diagonal and off-diagonal cells showed inconsistent results, for example for dataset 1 (NN Sector) shown a count of 1585 where as the counts in dataset 9 accounted for 13342, which is the most populated dataset, the counts of each paired dataset for Group A geographical pair can be seen in Table 8. The inconsistency experienced in dataset 1 would be due to missing data for a particular sector, the same inconsistency is also observed for the dataset 5. Dataset 1, 5 and 9 takes account the wind speed at the same time and direction for wind speeds at both target and reference site. Apart from the Group A data, the remaining geographical group analysis has proven the that there is a persistent relationship between the target and the reference long-term wind speeds when the directional sectors paired.

4.2. Estimation of Weibull Distribution Parameters

It can be summarized that the use of Q-Q Plots including calms and extreme wind speeds would be most appropriate to use when fitting the data since it provides the best estimation of the shape and scale parameters of the Weibull distributions of the data. Despite the Method of Moments is mathematically the most correct approach, Q-Q Plot analysis has shown the most proximate results to the measured values of the shape and scale parameters. Results and comparisons between for three different approaches as in reviewed under the Literature Review can be seen in Appendix C, page 63, Figure 36, 37 and 38, page 63 and 64.

4.3. Sector Pairing and Modified Weibull Scaling Method Analysis

As shown in Group A to F analysis (Table 14, page 60, Appendix B, page 59), Modified Weibull Scaling Method produce the least weighted error outside the exemption of Group B analysis, where for this geographical group pair, Linear Regression has shown better prediction of the mean wind speed over the long-term, however MWSM is produced less prediction error for individual dataset predictions, where the total weighted error counted higher than that of the Linear Regression predictions when the data counts in each dataset taken into account. This would be due to influence of the data counts in each dataset. Though, sector-wise variability error for MWSM has shown better agreement of the prediction with a value of 21.93%, whereas this is valued at 31.26% for LRM (Table 14, page 60, Appendix B, page 59). Group B predictions require further analysis so as to come up with a logical explanation of the factors that have influenced the resulting total weighed error. This process would include seasonal variations and predictions based on this and going on an increase in the sector pairing procedure.

Diagonal dataset, which are 1, 5 and 9, which corresponds to wind speeds at the same time and in the same direction at both sites, this implies that these datasets should be more populated than that of the off diagonal cells, consequently it has been observed that at some points diagonal cells counted less than that of the off-diagonal ones. For example Dataset 5 in group A data found to be count half of the dataset 6, which in theory it should be other way around. In addition to the possible factors explained previously, this might because, for example; the wind direction when in north-east direction at target site and when in north-west direction at the reference site which both datasets are in the northerly sector. This implies that there would be the necessity of more sector pairs and longer period of data in order to avoid high uncertainty when performing the methods, for example, short-term data may need to be taken longer than 15 months in order to end up with a meaningful amount of data for each dataset.

A sanity check, time against direction by plotting the target and reference data has been performed for the group A short-term concurrent data to see how the diagonal and off-diagonal interacts each other, for example reference sector 1 and target sector 1 and so on. It has been observed that the data in group A is quite unpaired; this might be another reason why that the Modified Weibull Scaling method did not perform well from the theoretical standpoint.

The same predictions have also been analyzed by performing the Linear Regression method to see how the both methods agree with the actual measurements. The procedure that is followed was exactly the same when performing the analysis by using the Modified Weibull Scaling Method. This procedure is applied to each data set (dataset 1 to 9) repeatedly and it has been found that the fundamental assumption, which implies that there is a persistent relationship between the wind speeds over the long-term predictions found to be valid when performing the Modified Weibull Scaling method.

5. CONCLUSIONS AND FURTHER RESEARCH

Six data sets with widely varying meteorological characteristics were assembled and summarized. A method for estimating long-term wind regime at a site was developed and evaluated. Sector pairing procedure as explained under the Methods chapter was applied to each geographical group (Group A to F) datasets. The predictive capabilities of two measure-correlate-predict methods (Regression and the Modified Weibull Scaling Method) and associated errors were explored using these data sets. Date sets D, E and F and results gained with using MWSM from these datasets were examined in greater details. The ability to estimate shape and scale parameters of Weibull distribution parameters for Machrihanish wind speed data in Group A was evaluated using the Method of Moments and Q-Q plot so as to make comparisons between two method. The impact of site characteristics on the performance of the Modified Weibull Scaling Method was examined.

The following conclusions and recommendations were reached:

- In estimating the shape and scale parameters of the Weibull distribution, using the Q-Q Plots (calms and extreme wind speeds included) performed significantly better than that of the Q-Q Plot (calms and extreme wind speeds excluded) and the Method of moments.

- A decrease or increase in the value of the correlation coefficient (r^2) between neighboring stations seen to be not affecting the performance of the MWSM. The same conclusion can be made for the performance of the LRM. These outcomes are valid only when the sector pairing procedure is applied.

- Among the evaluated measure-correlate-predict methodologies, only the Modified Weibull Scaling Method is recommended for general usage. It can summarized that the predictive capability of the MWSM is increasing proportional to the correlation coefficient (r^2) with the exemption of Group A and E results.

- Linear Regression Method has the potential to produce robust predictions of wind speed at a target site.

- Special consideration should be paid to the Modified Weibull Scaling Method. Such as an increase in sector pairs, for example taking more sectors than three equally paired sectors might greatly increase the predictive capability of this method. However in this case a longer period of short-term concurrent data will be necessary so as to reach a meaningful amount of counts in each dataset.

The following analyses remain for further research:

- Long-term data (20 years or more) for Group A datasets are available on MIDAS website with an access to restricted datasets. Group A analysis should be performed again to overcome with the uncertainty in the results.

- The effects of the use of concurrent data in making predictions should reanalyzed in order to assess the sensitivity of using the concurrent data by taking shorter or longer than 15 months. This should be done along with making sequential increases in sector pairing procedure, i.e. 4 or 5 or more equal sectors.

- Method compassions can also be made with an example turbine to assess the expected energy production. This approach is important to evaluate the success of the MCP techniques since they provide measures of how effective the applied technique could be from the economics standpoint.

- In order to make reliable comparisons the predictive capability of the developed Modified Weibull Scaling Method should be compared with the other methods in the literature, such as, Weibull Scale, Matrix, Wind Index methods.

- A model in deciding the period of short-term concurrent data can be used in order to make sure that the short-term data does not represent the unusual wind speed and direction variations, for example in that year when the wind measurement campaign took place, wind speeds may have been calm or extreme compared to historical wind speed record. Assuming that there is no long-term wind speed records available for a target site, correlation of the data measured over the wind speed monitoring campaign can be made by using the correlation made over the wind speeds at the reference site. Therefore, A model that estimates the yearly average wind speed amongst the 20 year or more historical data at the target site and implement this relationship onto measured wind speeds may be adopted to tackle with any issue that may rise due to inconsistent wind speed patterns in a year. This method implies that an artificial concurrent data so as to make more sensible long-term wind speed predictions can be created depending on the annual average mean speed variations in long-term wind speed variability at the target site, meaning that rather than taking the precise period of short-term concurrent in establishing a relationship. Accommodation of such a model will be necessary in artificial generation of a concurrent data specified above.

- It can be summed up that making comparisons of the correlation coefficients (r^2) with the wind speed predictions performed by different MCP methods would be unrepresentative. Correlation coefficients of the short-term wind speeds at both target and reference site along with the correlation coefficient of the long-term reference wind speeds can be used in establishing a sort of different relationship of the predicted wind speeds.

BIBLIOGRAPHY

[1] Sinden, G. (2011). *Wind Power and the UK Wind Resource* [online]. Oxford: Graham. Available from http://www.eci.ox.ac.uk/publications/downloads/ sinden05-dtiwindreport.pdf [Accessed: 20 June 2011].

[2] Früh, W. G. (2010). Wind Energy Introduction and Basic Resource. [Lecture notes]. Renewable Energy Technologies Module. Heriot-Watt University, Earl Mountbatten Building, Room EM244, 1st Semester.

[3] Bailey, B. H. & S. L. McDonald. (1997). *Wind Resource Assessment Handbook*. National Renewable Energy Laboratory, 1617 Cole Boulevard, Golden, CO.

[4] Gipe, P. (2004). *Wind Power*. White River Junction, VT: Chelsea Green Publishing Company.

[5] Rogers, A.L., Rogers, J.W. & Manwell, J.F., (2005). Comparison of the performance of four measure-correlate-predict algorithms. *Jounal of Wind Engineering and Industrial Aerodynamics*. 93 (3) p.243–264.

[6] Twidell, J. & Weir, T. (2006). *Renewable Energy Resources*. [2nd Ed]. New York: Taylor & Francais Group.

[7] Draper, N. R. & H. Smith. (1966). *Applied Regression Analysis*. New York: Wiley.

[8] Burton T., Sharpe D., Jenkins N., & Bossanyi E. (2001). *Wind Energy Handbook*. New York USA: John Wiley & Sons.

[9] Dorvlo, A.S.S.. (2002). Estimating wind speed distribution. *Energy Conversion and Management*. 43 (3) p.2311-2312.

[10] F. F. Gan, Kenneth J. K. & John C. T. (1991). Probability Plots and Distribution Curves for Assessing the Fit of Probability Models. *The American*

Statistician. 45 (1) p.14.

[11] Rogers, A. L., Rogers, J. W. & Manwell, J. F. (2005). Uncertainties in Results of Measure-Correlate-Predict Analyses. *American Wind Energy Association*, Denver, CO.

[12] Anderson, M. & Bass, J. (2004). "*A Review of MCP Techniques*", Renewable Energy Systems Ltd, Report No. 01327R00022, Issue No. 3.

[13] Manwell, J. F., McGowan, J. G. & Rogers A. L. (2009). *Wind Energy Explained*. [2nd Ed]. USA: John Wiley and Sons Ltd.

[14] McKenzie, J. H., Clive, P. J. M., Chindurza, I. & Bulté, H. (2008). Considering the Correlation in Measure-Correlate-Predict Techniques. *Proceedings of Wind Power Asia*, Beijing.

[15] Sreevalsan, E., Das, S. S., Sasikumar, R. & Ramesh, M. P. (2007). Wind Farm Site Assessment Using Measure-Correlate-Predict (MCP) analysis, *Wind Engineering*. 31 (2) p.111-116.

[16] Clive P. J. M. (2008). Non-linearity in MCP with Weibull Distributed Wind Speeds. *Wind Engineering*. 32 (3) p. 319-324.

[17] Thøgersen, M.L., Motta, M., Sørensen, T. & Nielsen, P. (2007). Measure-Correlate-Predict Methods: Case Studies and Software Implementation, *Proceedings of the EWEA*, Milan.

[18] Clive P. J. M. (2011). *Extracting information from data: MCP in real world*. Unpublished.

[19] Dale, L., Milborrow, D., Slark, R., Strbac, G. (2004). "Total Cost Estimates for Large-Scale Wind Scenarios in UK". *Energy Policy*, (32), p.1949-1956.

[20] Department of Trade and Industry (2001), *Wind energy Fact Sheet 3*: The economics of onshore wind energy [online]. Available at:

http://webarchive.nationalarchives.gov.uk/+/http://www.berr.gov.uk/files/file17776.pdf [Accessed: 27 June 2011].

[21] British Wind Energy Association [BWEA]. (2010). *"Wind Energy Generation costs, Fact Sheet 04"*, [online]. Available at: http://www.bwea.com/pdf/briefings/Wind-Energy-Generation-Costs.pdf [Accessed: 27 July 2011].

[22] White, F. M. (2003). *Fluid Mechanics*. McGraw Hill, Boston, MA, [5th Ed].

[23] Environmental Change Institute. (2011). *Wind Power and the UK Wind Resource*. [online] Available at: http://www.eci.ox.ac.uk/publications/downloads/sinden05-dtiwindreport.pdf [Accessed: 28 July 2011].

[24] Dale, L., Milborrow, D., Slark, R. & Strbac, G. (2004). Total cost estimates for large-scale wind scenarios in UK. *Energy Policy for a Sustainable Energy Future.* 32 (17) p.19-52.

[25] Pidwirny, M., Lemke, K. & Faulkner, D. (2006). *Understanding Physical Geography*, [1st Ed]. Michael Pidwirny.

[26] Renewable energy potential. (2008). [online] Available at: http://www.scotland.gov.uk/News/Releases/2008/02/14122107 [Accessed: 2 March 2011].

[27] Scottish Government, *Climate Change Delivery Plan.* (2009). [online] www.scotland.gov.uk/Publications/2009/06/18103720/0 [Accessed: 1 July 2011].

[28] Landberg, L., Myllerup, L., Rathmann, O., Petersen, E. L., Jørgensen, B. H., Badger, J. & Mortensen, N. G. (2003). Review Wind Resource Estimation - An Overview. *Wind Engineering*. 6 p.261-271.

[29] Salmon, J. R. & Walmsley, J. L. (1999). A two-site correlation model for wind speed, direction and energy estimates. *Journal of Wind Engineering and*

Industrial Aerodynamics. (79) p.233-268.

[30] Derrick, A. (1992). Development of the Measure-Correlate-Predict Strategy for Wind Site Assessment. *British Wind Energy Association*. National Wind Turbine Center, NEL, UK.

[31] Alexiadis, M., Dokopoulos, P., Sahsamanoglou, H., & Manousaridis, I. (1998) Short-term forecasting of wind speed and related electrical power. *Solar Energy*. 63 (1) p.61-68.

[32] IPCC (2007). IPCC Fourth Assessment Report: Climate Change 2007 (AR4). Cambridge, United Kingdom and New York, NY, USA.: Cambridge University Press.

[33] Bechrakis, D., Deane, J., & McKeogh, E. (2004). Wind Resource Assessment of an Area Using Short-Term Data Correlated To a Long-Term Data Set. *Solar Energy*. (76) p.725-732.

[34] The British Atmospheric Data Centre. (2011). *Met Office Surface Data Users Guide*. [online]. Available from: http://badc.nerc.ac.uk/data/ukmo-midas/ukmo_guide.html#2.6 [Accessed: 2 June 2011].

Appendix A

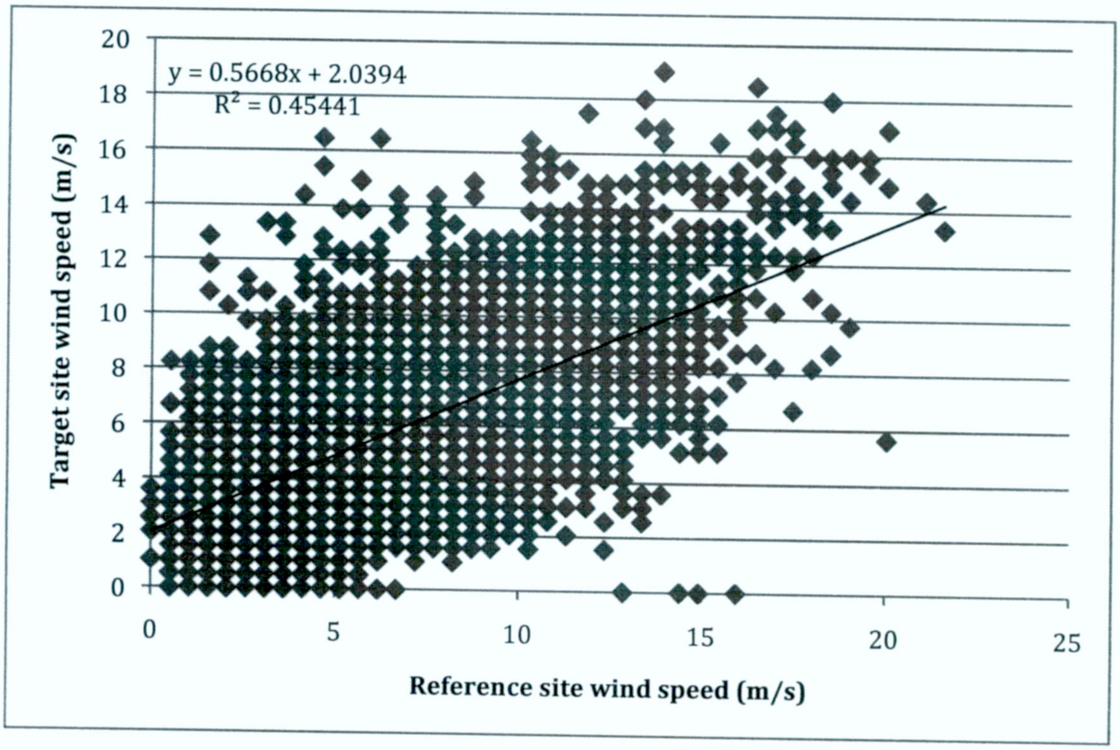


Figure 26: The wind speed scatter for Group A wind speeds over the short-term

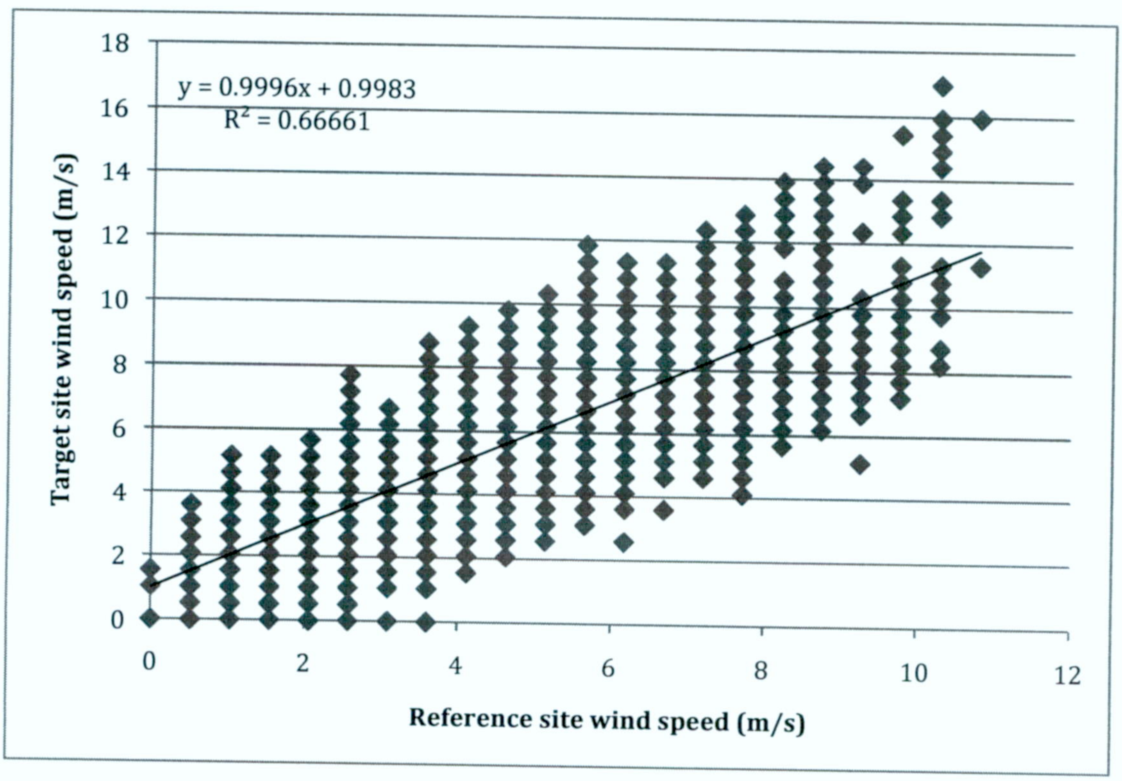


Figure 27: The wind speed scatter for Group B wind speeds over the short-term.

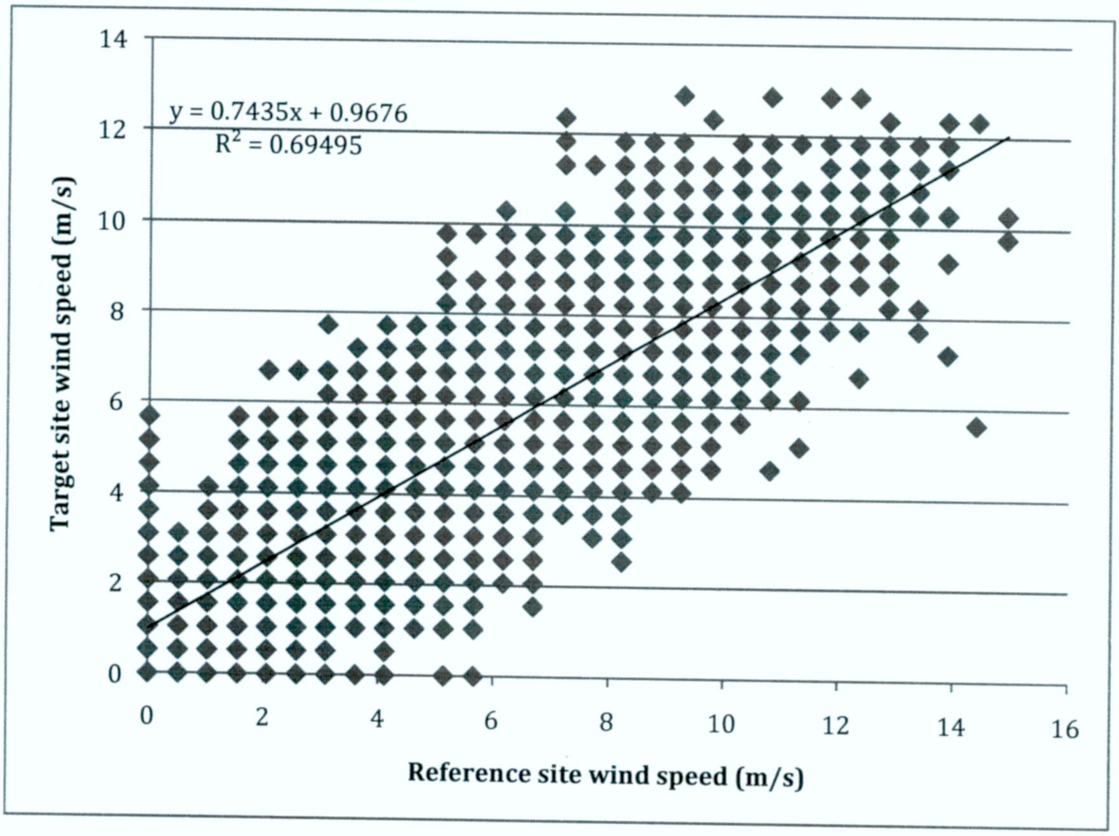


Figure 28: The wind speed scatter for Group C wind speeds over the short-term.

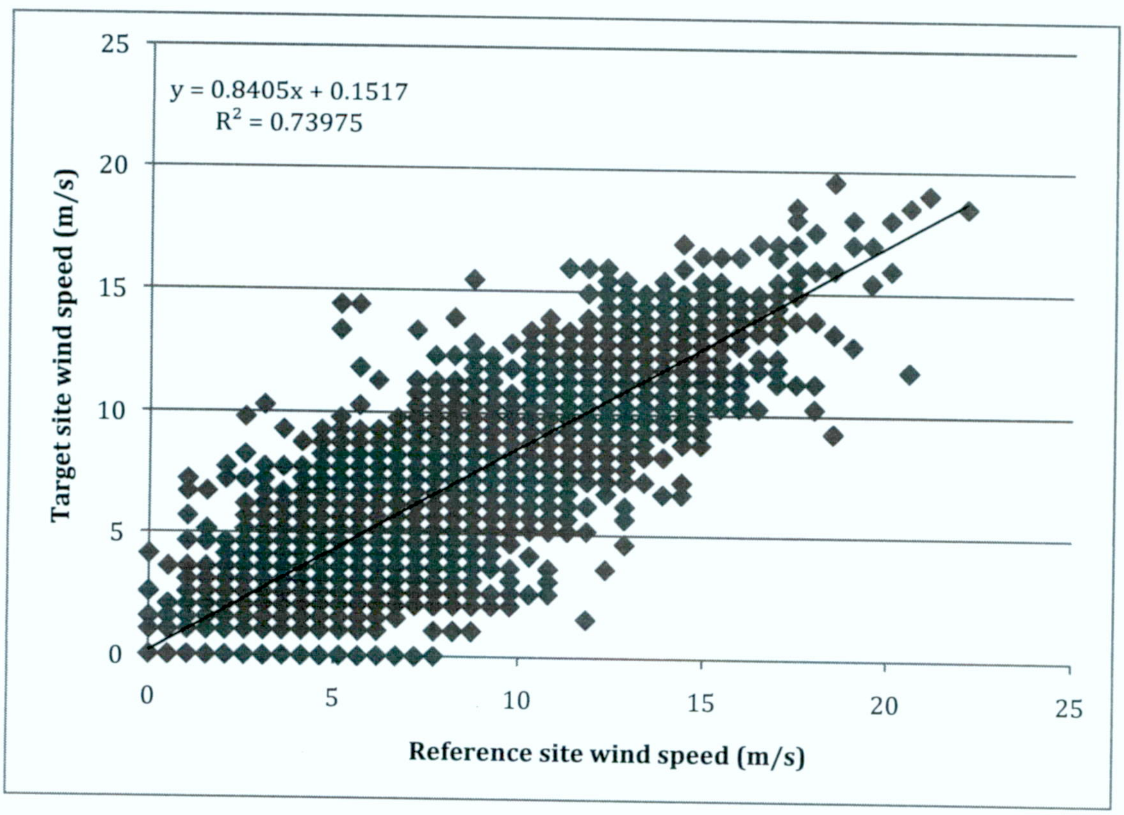


Figure 29: The wind speed scatter for Group D wind speeds over the short-term.

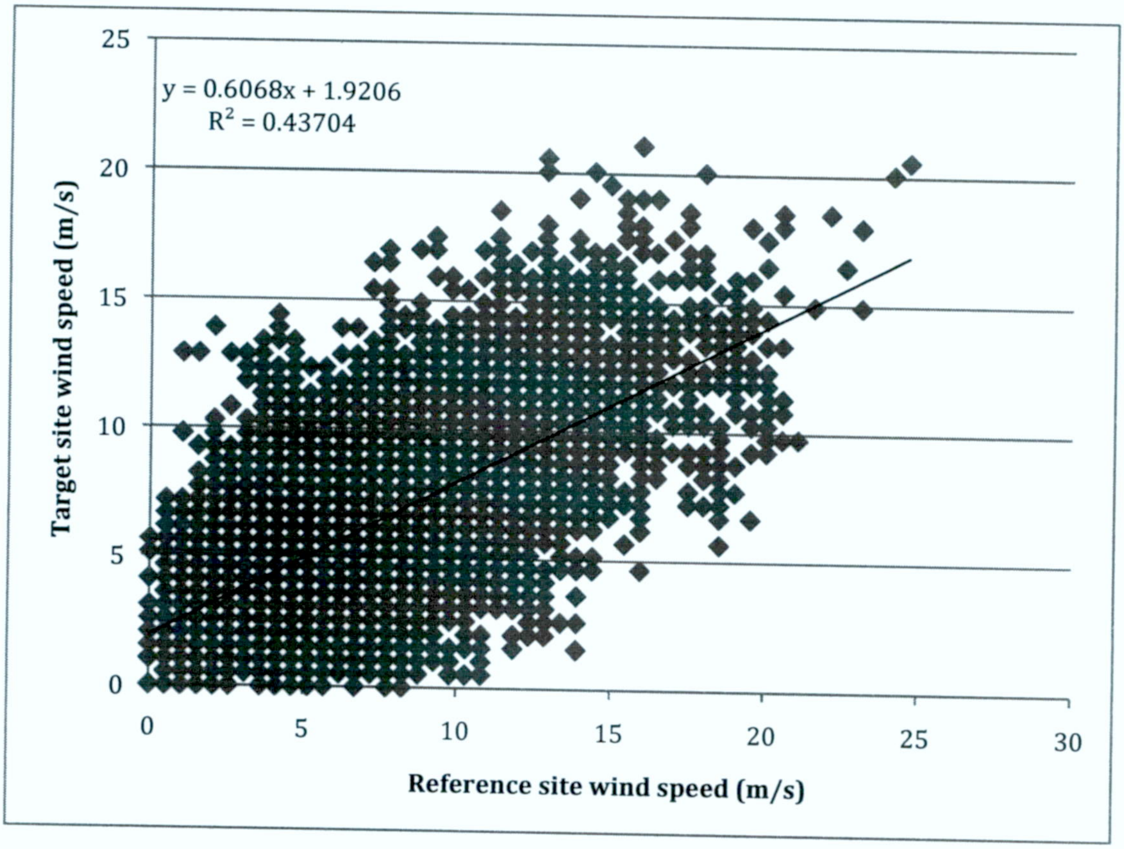


Figure 30: The wind speed scatter for Group E wind speeds over the short-term.

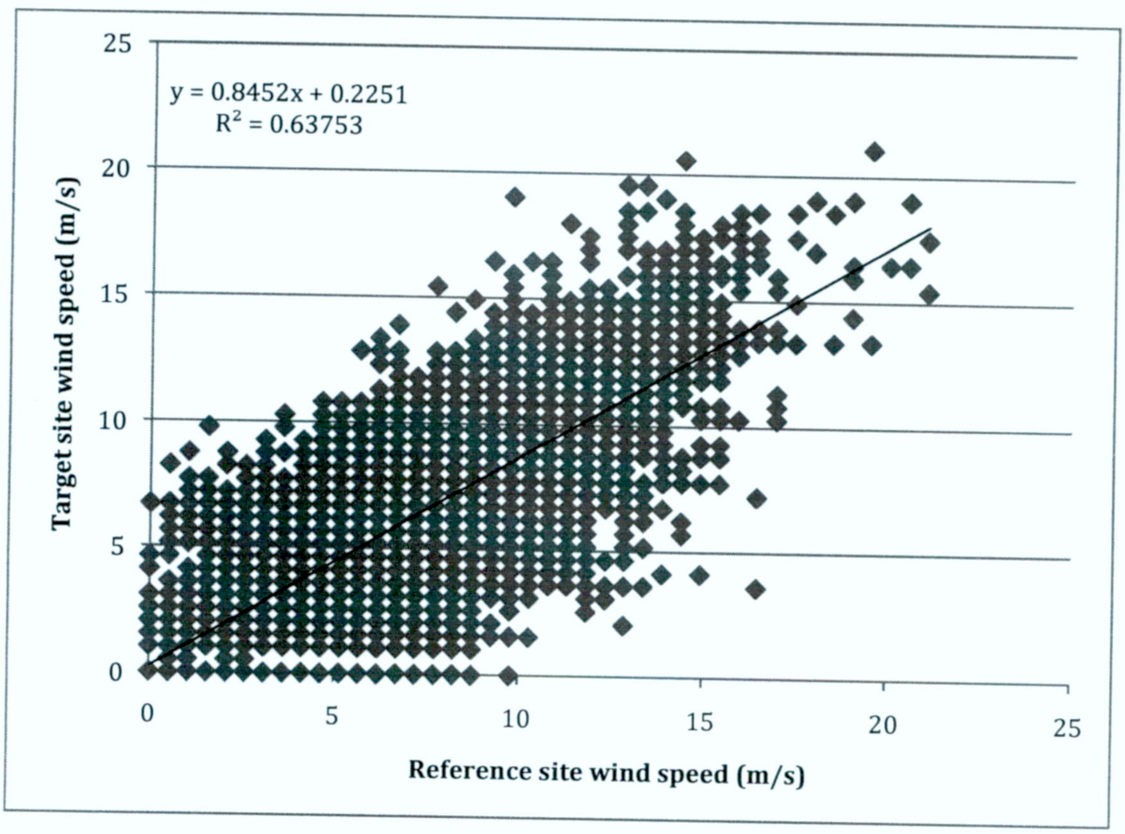


Figure 31: The wind speed scatter for Group F wind speeds over the short-term.

Appendix B

Datasets	Measured wind speed (m/s)	LR predictions (m/s)	MWSM predictions (m/s)	Accuracy prediction for LRM (%)	Accuracy prediction MWSM (%)	Prediction error for LRM (%)	Prediction error for MWSM (%)	Count of data
D1	7.41	5.41	7.09	73.06%	95.74%	26.94%	4.26%	26280
D2	4.87	5.65	4.51	116.12%	92.63%	-16.12%	7.37%	2976
D3	5.69	4.05	5.15	71.18%	90.40%	28.82%	9.60%	10985
D4	4.44	2.91	4.17	65.52%	94.03%	34.48%	5.97%	7299
D5	8.60	7.24	8.05	84.25%	93.61%	15.75%	6.39%	21946
D6	7.02	4.79	6.59	68.26%	93.79%	31.74%	6.21%	8626
D7	5.14	4.36	4.75	84.77%	92.36%	15.23%	7.64%	3356
D8	7.28	7.35	6.98	100.98%	95.79%	-0.98%	4.21%	6390
D9	7.54	7.29	7.19	96.77%	95.46%	3.23%	4.54%	50758
					TWE	14.83%	5.48%	

Table 12: Analysis fo Group E dataset by using both Modified Weibull Scaling Method and Linear Regression Method. Total weighted error shows how well the methods perform.

Datasets	Measured wind speed (m/s)	LR predictions (m/s)	predictions (m/s)	Accuracy prediction for LRM (%)	Accuracy prediction for MWSM (%)	Prediction error for LRM (%)	Prediction error for MWSM (%)	Count of data
D1	7.49	5.77	7.15	77.01%	95.40%	22.99%	4.60%	37791
D2	3.39	4.55	3.17	134.00%	93.49%	-34.00%	6.51%	2240
D3	5.80	3.79	5.29	65.32%	91.23%	34.68%	8.77%	9921
D4	4.63	2.70	4.17	58.40%	90.14%	41.60%	9.86%	5735
D5	7.38	5.29	6.31	71.59%	85.48%	28.41%	14.52%	48863
D6	4.41	3.48	4.12	79.04%	93.59%	20.96%	6.41%	4013
D7	4.26	1.65	4.52	38.79%	105.92%	61.21%	-5.92%	8417
D8	6.53	4.14	6.18	63.39%	94.64%	36.61%	5.36%	14844
D9	7.56	5.90	7.27	77.97%	96.06%	22.03%	3.94%	74576
					TWE	26.89%	6.74%	

Table 13: Analysis fo Group F dataset by using both Modified Weibull Scaling Method and Linear Regression Method. Total weighted error shows how well the methods perform.

	Group A	Group B	Group C	Group D	Group E	Group F
TWE error for LRM (%)	-15.42%	-0.13%	4.60%	14.63%	14.83%	26.89%
TWE error for MWSM (%)	-9.48%	9.67%	3.92%	-3.34%	5.48%	6.74%
Correlation Coefficient (r^2)	0.454	0.667	0.695	0.740	0.437	0.638
SWE for LRM	24.36%	31.26%	21.07%	34.68%	32.83%	43.66%
SWE for MWSM	26.51%	21.93%	15.20%	8.38%	15.04%	16.98%

Table 14: A table showing the total weighted errors for both Linear Regression (LRM) and Modified Weibull Scaling Method (MWSM) and associated sector-wise variablity of the prediction (SWE) error calculated as the weighted standard deviation of the sector prediction error from its weighted error for each Geographical Group wind speed predictions.

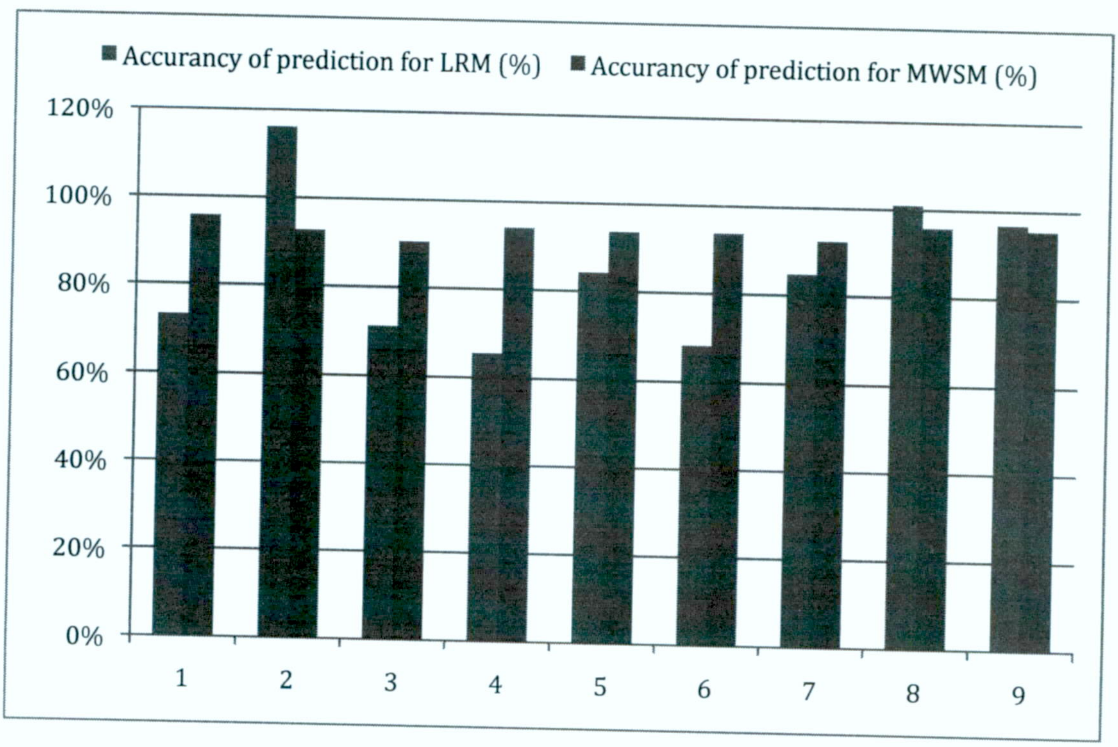


Figure 32: Group E prediction accuracy comparison betweern Linear Regression and Modified Weibull Scaling Method results

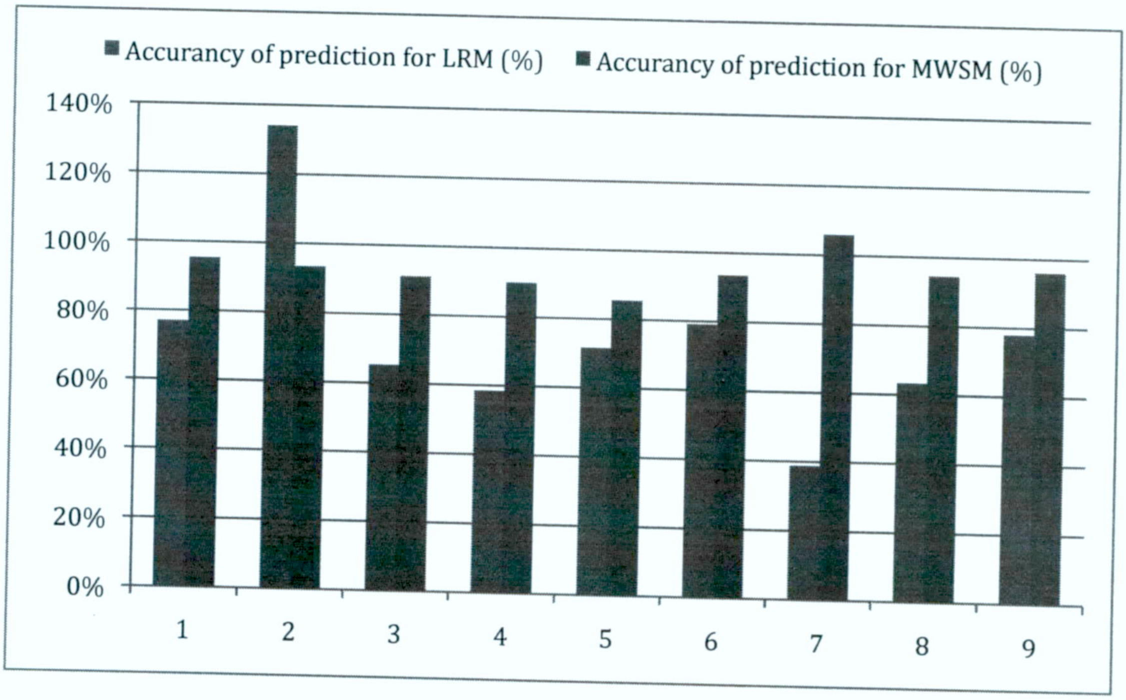


Figure 33: Group F prediction accuracy comparison betweern Linear Regression and Modified Weibull Scaling Method results

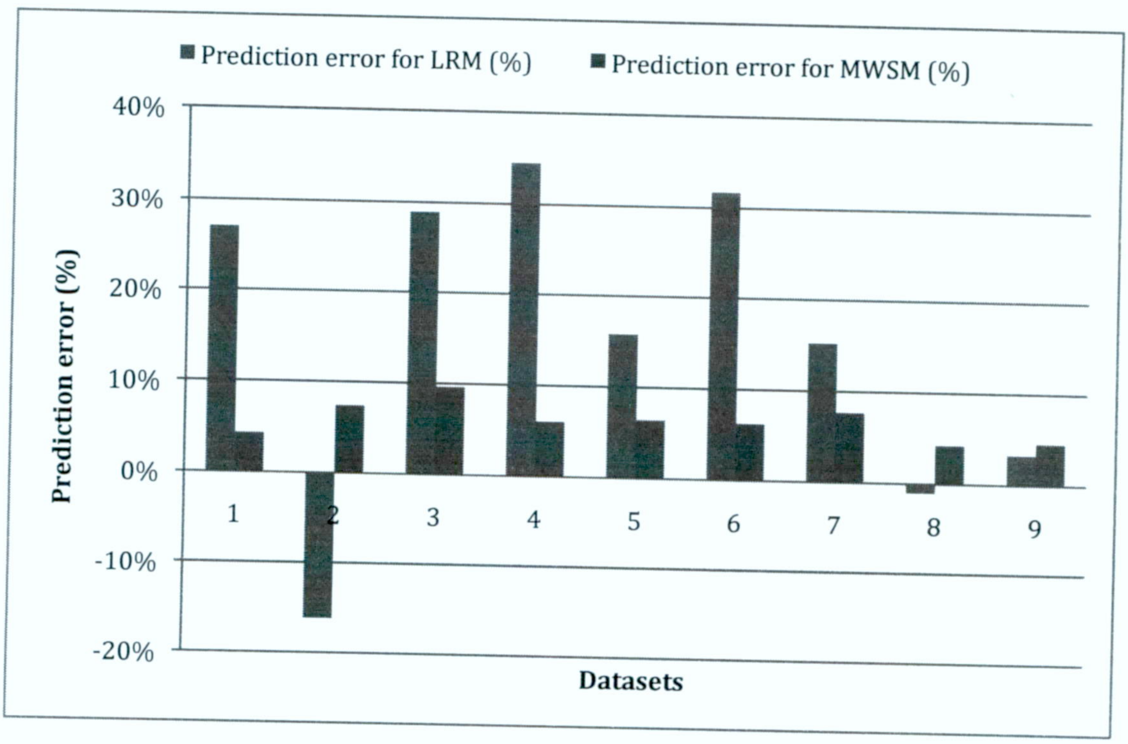


Figure 34: Group E prediction error comparison betweern Linear Regression and Modified Weibull Scaling Method results

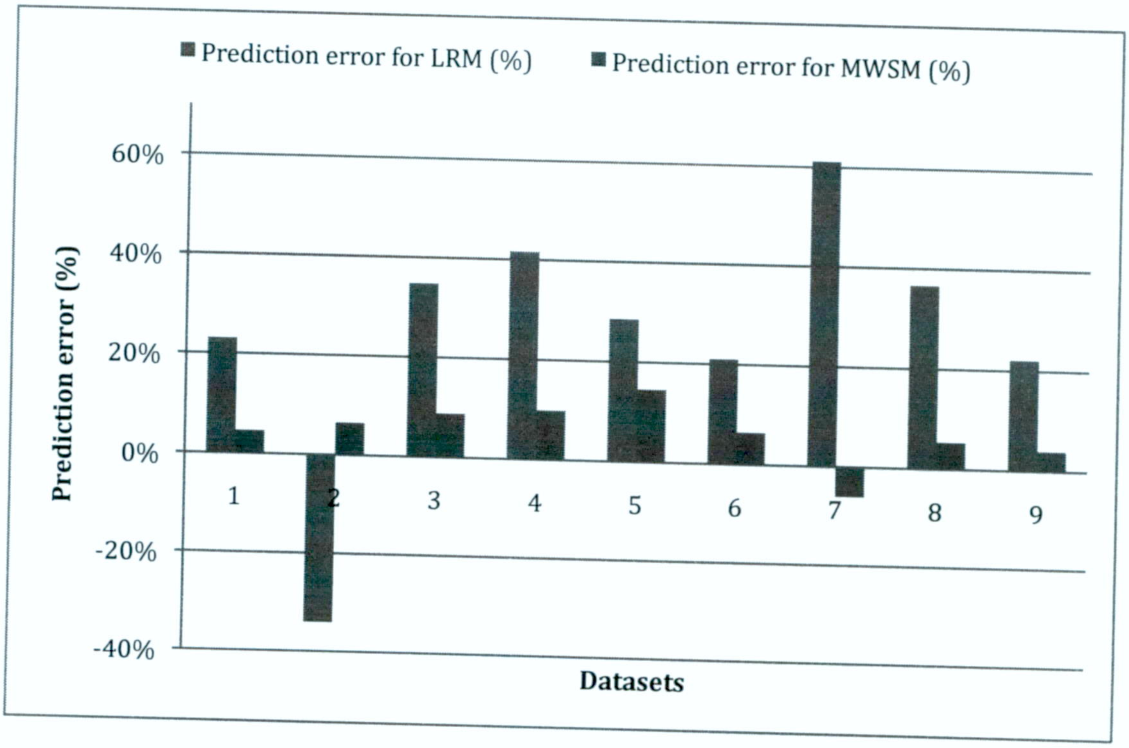


Figure 35: Group F prediction error comparison betweern Linear Regression and Modified Weibull Scaling Method results

Appendix C

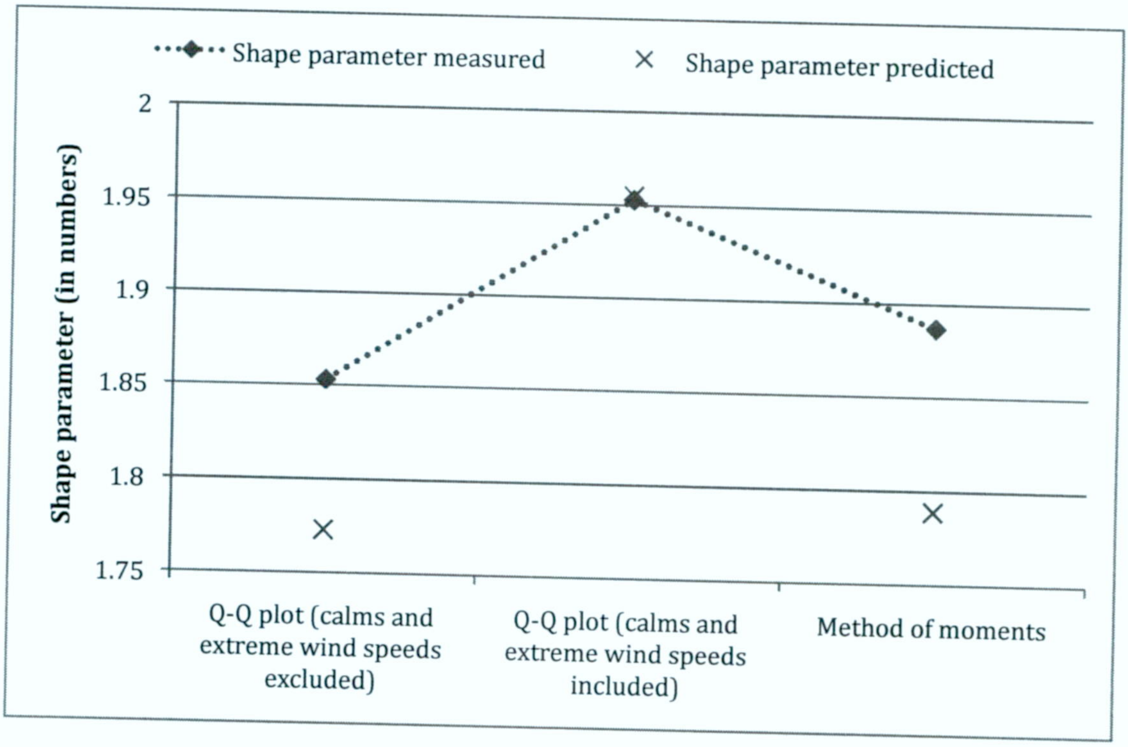


Figure 36: Shape parameter predictive capability of three different methods for Machrihanish wind speeds

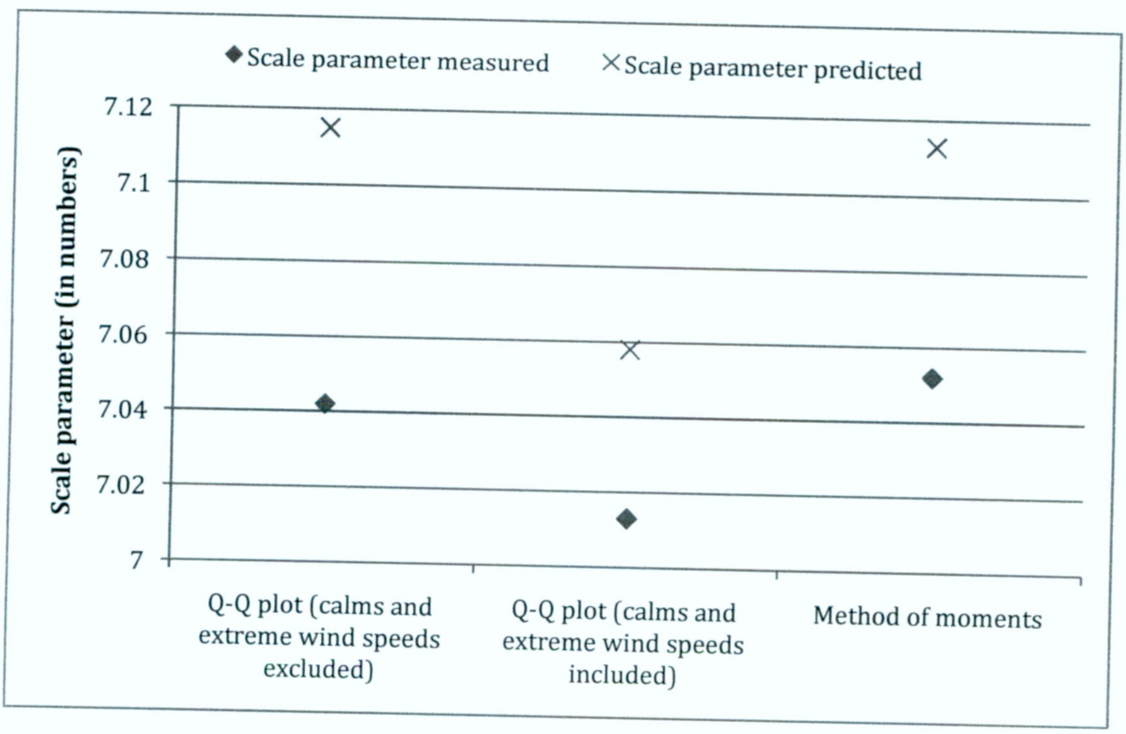


Figure 37: Scale parameter predictive capability of three different methods for Machrihanish wind speeds

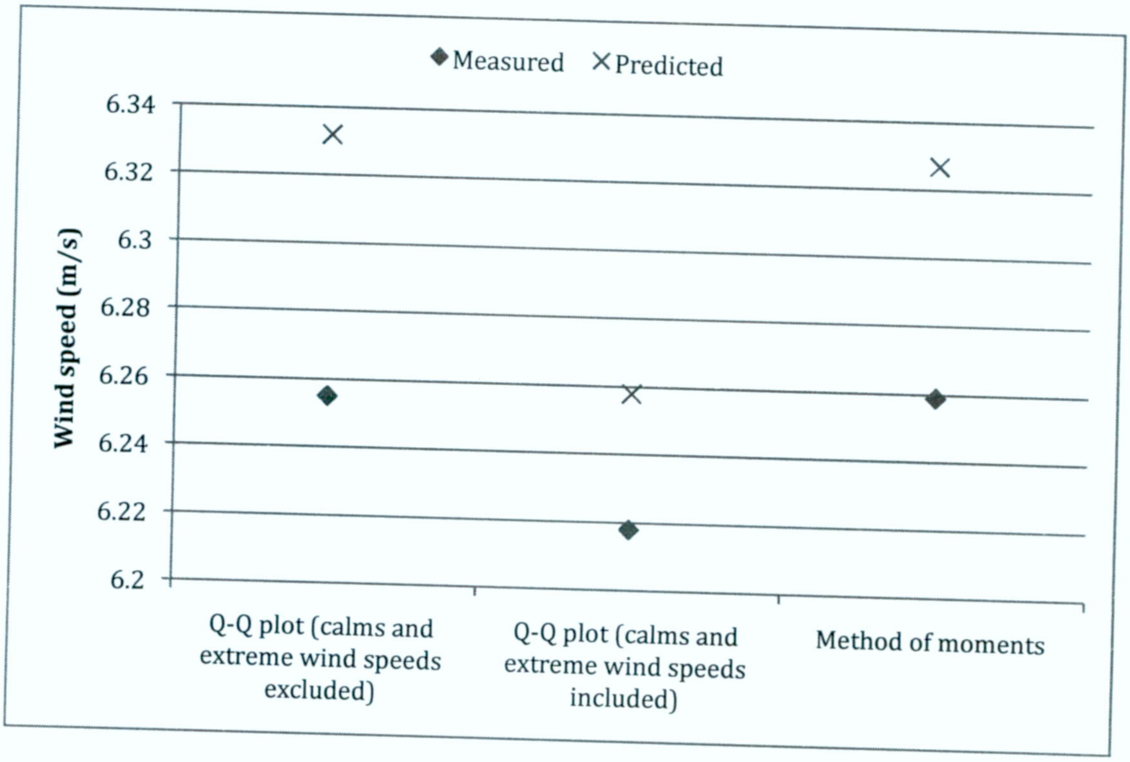


Figure 38: Predictive capability of three different methods for Machrihanish wind speeds

Appendix D

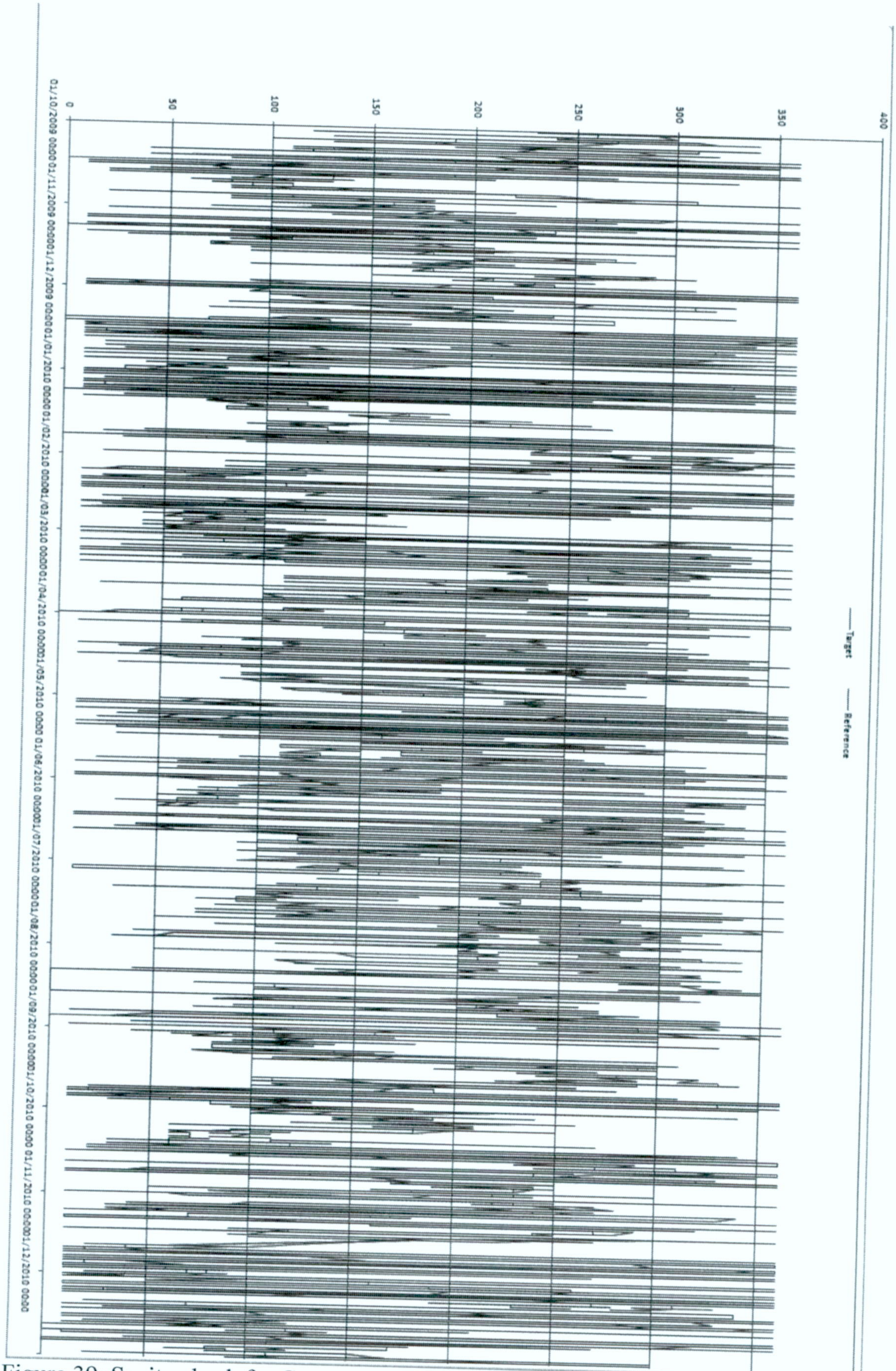


Figure 39: Sanity check for Group A wind speeds over the short-term

Appendix E

Technology	Capital cost, £/kW	O&M cost, £/kW/yr	O&M cost, £/MWh	Fuel cost, £/MWh
Gas	650		4	14
Coal	1650		7	7.5
Nuclear	2000-3000		9-11	6.5
On-shore wind	1300-1600	54		Free
Offshore wind	2500	79		Free

Table 15: Electricity generating costs of different technologies [21].