A Statistical Analysis of Cape Town Wind Profile

Mbuyamba Tshibanda
Faculty of Engineering and the Built Environment
University of Johannesburg, South Africa
mtshibanda@uj.ac.za

Jan-Harm C Pretorius
Faculty of Engineering and the Built Environment
University of Johannesburg, South Africa
jhcpretorius@uj.ac.za

***Abstract*—The increased integration of wind power into electric power systems presents new challenges for effective planning and operation of these systems. The Weibull distribution is a widely used distribution, especially for modelling the random variable of wind speed. In this respect, the authors present a comparative analysis of a number of methods used for estimating Weibull parameters. Results for a real-world database are presented in a case study format. The techniques require historical wind speed data, collected over a particular time interval, to establish the parameters of wind speed distribution for a specific location, namely Cape Town, South Africa.**

Keywords—Wind speed data; wind energy; probability density function; Weibull distribution.

I. Introduction

The utilization of renewable energy in electric power systems is expanding rapidly due to environmental concerns as well as those related to the depletion of conventional power generation sources. Wind power is a fast growing electric generation technology. Wind generation offers enormous benefits to power systems, such as an economical form of energy (in comparison to thermal generation), emission reduction, availability to large areas and relatively easy implementation of wind power farms [1]. However, wind generation presents a series of difficulties to traditional power systems, namely uncontrolled level of power generation which is solely dependant on wind availability that may result in poor predictability of wind generation (irregular fluctuation and intermittence of power generation) [2].

The most significant properties of wind generation include the frequency and magnitude of wind speed. The power output obtained from wind is directly proportional to the cube of wind speed. One of the primary characteristics of wind is that it is highly variable and its properties vary from one location to another. Wind speed changes continuously and the statistical approach is considered a viable method to estimate its speed and frequency values. In this respect, wind speed probability density function plays an important role in electric power generation applications of wind turbines. The most important requirements for effective wind power planning and operation in power systems is an accurate estimation of wind speed distribution. Therefore, investigation of wind power generation should be carefully performed in accordance with wind speed probabilistic characters.

This paper presents a statistical analysis of the main characteristics of wind speed in the region around Cape Town, South Africa. Wind speed data is measured as average hourly values, being statistically analyzed over a one year period. The probability density distribution is derived from the obtained values and their distributional parameters are evaluated in accordance with acceptable statistical methods.

In order to evaluate wind speed distribution, the key factors of wind speed probability density function, height dependence and wind direction are presented in section two. The main statistical methods used to evaluate the parameters of probability density functions are reported in section three. In addition, two statistical assessments used to establish the accuracy of the applied methods, are reported. In section four, a numerical analysis is developed with the purpose of evaluating the statistical parameters, based on the hourly wind speed database collected from the Cape Town region. Finally, conclusions are presented in section five.

II. Wind Speed Distribution

Available wind energy depends on wind speed, which could be considered as a random variable. Values representing wind-speed occurrence over an extended period of time could be described in terms of probability distribution function.

A. Wind Speed Probability Density Function

A number of studies have been published in scientific literature related to wind energy, which propose to use a variety of probability density functions (e.g. normal, lognormal, gamma, Rayleigh, Weibull) to describe wind speed distributions [3,4,5]. The common conclusion of these studies is that the two-parameter Weibull distribution may be successfully utilized to describe the principle wind speed variation. The Weibull probability density is expressed by the following:

$$f_W(v) = \frac{\beta}{\alpha}\left(\frac{v}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{v}{\alpha}\right)^{\beta}\right]. \tag{1}$$

The corresponding cumulative probability function of the Weibull distribution is:

$$F_W(v) = 1 - \exp\left[-\left(\frac{v}{\alpha}\right)^{\beta}\right], \tag{2}$$

where α (m/s) is the scale parameter and β (dimensionless) is the shape parameter of the Weibull distribution.

The Weibull distribution is one of the most widely used distributions in a number of technical fields. This distribution has a particular property in that it lacks a characteristic shape and assumes the attributes of other distributions, based on various values of shape parameter, as demonstrated in Figure. 1below:

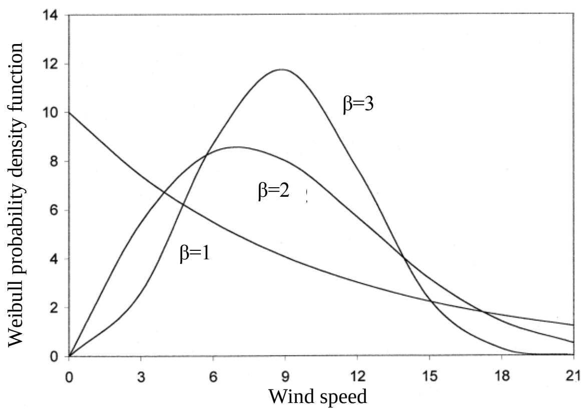

Fig. 1. Weibull probability distribution function with different values of shape parameter

The Weibull distribution becomes hyper-exponential when shape parameter is less than uniform and it is a well-known exponential distribution when the shape parameter is equal to one. On the other hand, it becomes a Rayleigh distribution when shape parameter is equal to two and is a normal distribution when β=3.4, or an approximate normal distribution when β approaches a value of 4, respectively.

In general, scale parameter provides information about the average wind speed profile, whilst shape parameter describes the deviation of wind speed values around the mean, along with the feature of probability density function. It is evident that Weibull distribution becomes relatively narrower and higher as shape parameter increases. In addition, the peak of density function moves in the direction of higher wind speeds as shape parameter increases. Shape and scale parameters are also interconnected through analytical expressions of the mean and variance of the Weibull probability density function.

B. *Variation of Wind Speed with Height*

Wind tends to blow faster at higher altitudes due to the influence of lower ground surfaces and decreased density of air. The most common expression for wind speed variation with height, applies the law of wind profile power (based on the ground friction coefficient), described by the following equation:

$$v(z)/v(z_r) = (z/z_r)^k \tag{3}$$

In the above equation, v(z) and v(z_r) pertain to wind speeds at desired z and registered z_r heights, whilst k is the friction coefficient, which depends on surface roughness and atmospheric stability [6]. Numerically, the friction coefficient ranges between 0.05 for smooth terrains and 0.5 for rough terrains, with the most frequently adopted value occurring at around 0.14.

The following relationships exist between parameters of the Weibull distribution for varying heights:

$$\alpha(z) = \alpha(z_r)\cdot(z/z_r)^k \,;\, \beta(z) = \beta(z_r) \tag{4}$$

This relationship is based on the expression of the Weibull distribution that is amended in accordance with [7], in respect of the wind speed and height relationship. As per equation (4) above, it is clear that shape parameter is a fixed property of wind profile, whereas scale parameter may be modified in narrow range, by adjustment of the desired height.

C. *Wind Direction Distribution*

Wind typically blows from several directions, which could be depicted by a wind rose diagram. The wind rose is a chart which offers a view of the manner in which wind speeds and directions are distributed at a particular location over a specific period of time. It is a useful representation as it allows for a large quantity of data to be encapsulated within a single plot.

If terrain roughness is similar in all directions, it can be said that little disparity occurs in wind speeds above or at the hub of the wind turbine, since the turbine yaw system invariably causes the rotor to follow the directionality in which wind blows. Therefore, notwithstanding exceptional cases, wind directionality is precluded and all wind is reasonably assumed to emanate from the same direction.

Conversely, if terrain around the turbine is significantly variable in terms of roughness (or obstacles), values for wind speeds at the hub of the wind turbine would be dissimilar, depending on the distribution of wind direction, thus necessitating the utilisation of detailed calculations.

III. Methods for Parameter Estimation

Wind speed distribution is determined in its entirety when its parameters are numerically established. Weibull distribution parameters may be determined through the use of various estimation methods, classified as either graphical or analytical [8,9,10]. The frequently utilized analytical methods include the Maximum Likelihood Estimator, along with the Methods of Moments and Least Squares. Each analytical method, as it applies in the current context, has specific criteria which yield estimates considered most suitable to particular situations.

Weibull parameters play a fundamental role in developing an electric power wind generator model, therefore it is imperative for diverse estimation methods to be compared in order to ensure that the parameters of the Weibull distribution correspond to a wind speed database.

As a result, this paper attempts to determine which method would produce the most suitable Weibull parameter estimation. The performance of the selected methods are analyzed in terms of the same wind speed database where the Relative Mean Bias Error (RMBE) and Relative Root Mean Square Error (RRMSE) are applied in statistical evaluation of the performance of Weibull parameters.

A. *Maximum Likelihood Estimator*

The Maximum Likelihood Estimator (MLE) is a widely used analytical method that is frequently applied in engineering and mathematical problems. In accordance with MLE theory and in the current context of the Weibull distribution of wind speed, the likelihood function is built as joint density of n number of random variables and is a function of the following unknown two parameters:

$$L(\alpha,\beta)=\prod_{i=1}^{n} f(v_i)=\prod_{i=1}^{n}\frac{\beta}{\alpha}\left(\frac{v_i}{\alpha}\right)^{\beta-1}\exp\left[-\left(\frac{v_i}{\alpha}\right)^{\beta}\right] \quad (5)$$

where α and β values can be achieved by using the iterative or limits method. The final method of parameter evaluation involves taking the partial derivatives of the likelihood function as they apply to the parameters and setting the resulting equations equal to zero as follows:

$$\frac{\partial \ln(L)}{\partial \beta}=\frac{n}{\beta}+\sum_{i=1}^{n}\ln v_i-\frac{1}{\alpha}\sum_{i=1}^{n}v_i^{\beta}\cdot\ln v_i=0 \text{ and}$$
$$\frac{\partial \ln(L)}{\partial \alpha}=-\frac{n}{\alpha}-\frac{1}{\alpha^2}\sum_{i=1}^{n}v_i^{\beta}=0 \quad (6)$$

The values of α and β are derived from simultaneous solving of both equations.

B. *The Method of Moments*

Distribution parameters could also be established by an analytical method referred to as the Method of Moments (MOM). For a known set of wind data, the moments of unknown parameters (that depend on the two-parameter Weibull distribution), are equalized with empirical moments.

The analytical expression of the mean and variance of Weibull distributions can be directly calculated from the equations below:

$$M(v)=\alpha\cdot\Gamma(1+1/\beta) \text{ and}$$
$$D^2(v)=\alpha^2\cdot\left[\Gamma(1+2/\beta)-\left(\Gamma(1+1/\beta)\right)^2\right] \quad (7)$$

where Γ( ) represents the gamma function, whilst empirical moments are calculated from the following equations:

$$\bar{v}=\sum_{i=1}^{n}\frac{v_i}{n} \quad \text{and} \quad \sigma^2=\frac{1}{n}\sum_{i=1}^{n}\left(v_i-\bar{v}\right)^2 \quad (8)$$

The β parameter can be obtained from the coefficient of variation (by dividing the variance on the square mean), and the α parameter can subsequently be established from the expression depicted as (7) above.

C. *Least Squares Method*

The Least Squares Method (LSM) is extensively used in engineering problems for the estimation of Weibull parameters. The method provides a linear relation between the two parameters, having as start point twice logarithms of the Weibull cumulative distribution function, as follows:

$$\ln\ln\left[1-\frac{1}{F_w(v)}\right]=\beta\ln(v)-\beta\ln(\alpha) \quad (9)$$

This relationship represents a straight line, expressed as: $Y=a\cdot X+b$, where:

$$Y=\ln\ln\left[1-\frac{1}{F_w(v)}\right],\ X=\ln(v), \text{ and}$$
$$a=\beta\ ,\ b=-\beta\ln(\alpha) \quad (10)$$

Performing rank regression on Y requires that a straight line be mathematically fitted to a set of data points such that the sum of the squares of deviations from the data points to the line, is minimized. This in essence, is the same methodology applied to probability plotting, however the principle of least squares is used to determine the line through data points. With the use of simple linear regression, α and β parameters are derived from the coefficient of polynomial linear fitting.

D. *Statistical Test Analysis*

Relative Mean Bias Error (RMBE) and Relative Root Mean Square Error (RRMSE) have been used in statistical evaluation of Weibull distribution performance. These statistical tests are based on the following expressions:

$$RMBE(\%)=\frac{1}{N}\sum_{i=1}^{N}\left(v_i-v_i^*\right)\bigg/\frac{1}{N}\sum_{i=1}^{N}\left(v_i^*\right)\cdot 100 \quad (11)$$

$$RRMSE(\%)=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(v_i-v_i^*\right)^2}\bigg/\frac{1}{N}\sum_{i=1}^{N}\left(v_i^*\right)\cdot 100 \quad (12)$$

where v_i is the ith actual data, v_i^* the ith predicted data with Weibull distribution and N the number of observations. The RMBE and RRMSE assessments provide information about the model's performance with lower values being desirable. Therefore, the preferred distribution function may be selected according to the lowest values of RMBE and RRMSE.

IV. Case Study of Cape Town, South Africa

In the current study, wind potential of the Cape Town region was statistically analysed based on hourly measurements extending over a one-year period. A wind speed database was recorded from prior studies in the South-East area of South Africa, in particular, Cape Town. The region of interest is specifically located on the South Atlantic seaboard, at 33°55' South latitude, 18°25' East longitude and 25 m above sea level. The recorded data is applicable to the year 2012. The measurements available in the initial database are characterized by one hour acquisition intervals, with the average hourly value being recorded.

Wind speed values collected at anemometer height (10 m above ground) were involved in adjustments to wind turbine height. As a result, the expression referred to in (3), was used to evaluate wind speed values at the hub of wind turbine height (100 m), assuming the same terrain roughness occurred around the wind turbine and that 0.14 was the friction coefficient. Figure 2 demonstrates the applicable wind speed values available at hub wind turbine height (100 m).

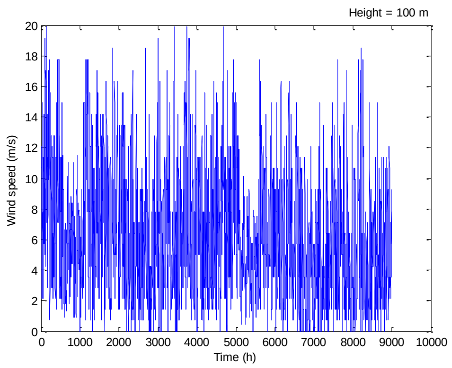

Fig. 2. Hourly Wind Speed Database Collected for Cape Town, South Africa

Similarly, the variation of wind speed to direction has been evaluated in prior studies. In this respect, wind direction and speed occurring over a period of time at a specific location may be graphically presented as a wind rose plot. To create a wind rose, average wind direction and wind speed values are arranged according to wind direction to allow the percentage of time to be determined, of wind blowing from each direction. Figure 3 shows the wind rose plotted for the hourly wind speed database of Cape Town:

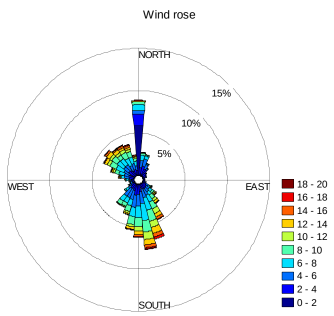

Fig. 3. Wind Rose Applicable to Wind Database Depicted in Fig 2.

Wind direction data is typically sorted into twelve equal arc segments, comprising 30° per segment. This is done in preparation for plotting a circular graph, in which the radius of each of the twelve segments represents the percentage of time that wind emanates from each direction (or each segment). As depicted in Figure 3 above, the primary wind direction follows a North-South route. Assuming the same roughness of terrain occurs around the wind turbine and taking into account that the rotor follows wind directionality, it may be concluded that wind speed values are not affected by wind direction.

On the basis of the above measurements, Weibull distribution parameters which approximate the existing database of wind speed frequency, were estimated. In order to compare the stipulated methods of parameter estimation, a Matlab® program was developed in accordance with previously described methods and the existing wind speed database.

In order to evaluate the performance of these methods, the Relative Mean Bias Error (RMBE) and Relative Root Mean Square Error (RRMSE) were applied to establish the accuracy of the estimated probability density function in relation to the actual distribution. RMBE and RRMSE are statistical tests used widely to evaluate the distinction between values provided by an estimated probability density function and established values of database distribution. Table I presents the Weibull parameters of the analyzed database, with scale and shape parameters being determined in accordance with methods reported in section three.

TABLE I. WEIBULL PARAMETERS FOR THE PERIOD 2012

Parameters / statistical tests	**Method**		
	MLE	**MOM**	**LSM**
Scale parameter α	7.3067	7.3277	6.9457
Shape parameter β	1.7833	1.8086	1.6594
RMBE	0.5869	0.6941	-1.5563
RRMSE	1.9135	2.0532	3.1003

The evaluated Weibull distribution parameters appear to be very close, demonstrating that the scale parameter lies between 6.9457 and 7.3277 m/s, while the shape parameter lies between 1.6594 and 1.8086. Similarly, the RMBE and RRMSE values for each of the applied methods are also presented. MLE appears to be superior in accuracy with a smaller margin of error in comparison to MOM and LMS. In terms of the current analysis, it appears that LMS can be considered the least accurate method.

In order to establish if any significant difference exists between the diurnal variations of wind speed, the wind speed database was divided into diurnal sets of values, namely daytime and nighttime, with each assessed in isolation. In this instance and for the purpose of simplicity, daytime is defined as the period between 7 AM and 7 PM whilst nighttime occurs from 7 PM to 7 AM.

A typical diurnal representation of wind speed is shown in Figure. 4 below:

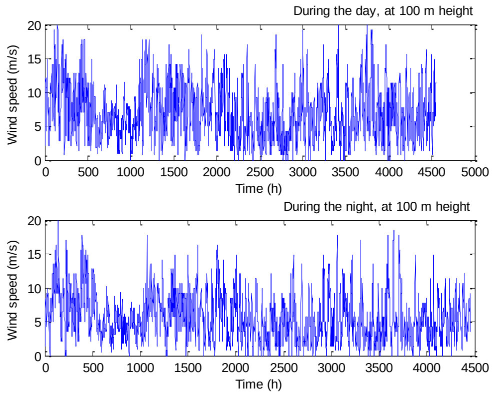

Fig. 4. Daytime and Nighttime Wind Speed Database

A similar analysis of diurnal wind directions was conducted. Wind rose diagrams for both sets of wind speed, over daytime and nighttime periods, are depicted in Figure. 5:

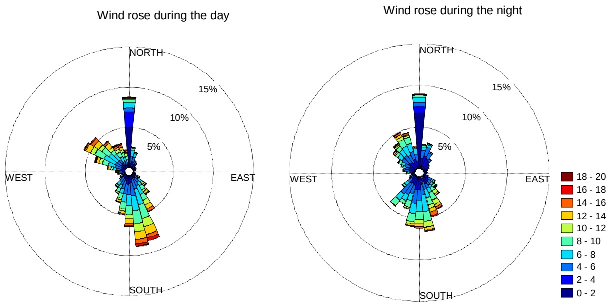

Fig. 5. Wind Rose Applicable to the Diurnal Wind Database

The Weibull parameters derived from analytical assessment of the available data (in accordance with the methods described in section 3) are presented in Table II. It is apparent that the scale factor varies between 7.8021 and 8.1882 m/s, the shape factor ranges from 1.7664 to 1.9442 and the scale factor varies between 6.0779 and 6.4376 m/s for wind speeds during the day, while the shape factor ranges from 1.6181 to 1.7529 during the night.

TABLE II. WEIBULL PARAMETERS FOR A DIURNAL WIND SPEED DATABASE

	Parameters / Statistical Tests	Method		
		MLE	MOM	LSM
Day time	Scale parameter α	8.1699	8.1882	7.8021
	Shape parameter β	1.9222	1.9442	1.7664
	RMBE	0.6140	0.7268	-1.7390
	RRMSE	2.1600	2.2660	3.6603
Night time	Scale parameter α	6.4376	6.4603	6.0779
	Shape parameter β	1.7258	1.7529	1.6181
	RMBE	0.5629	0.6665	-1.4718
	RRMSE	1.8682	1.9900	2.9136

RMBE and RRMSE values were also calculated. The scale and shape parameters derived from diurnal values and the wind speed database as a whole, revealed the best estimation when MLE was used. Therefore, an inference can be drawn that MLE appears to be the best method for estimating the parameters of the two-parameter Weibull distribution, taking into consideration RMBE and RRMSE as measurements of comparison.

V. Conclusions

In practice, it is imperative to describe wind speed variations for the optimal design of wind generation systems. The wind variation for a typical site is commonly described in terms of the Weibull distribution. Therefore it is vital to understand the best method of parameter evaluation that presents a minimal margin of error.

In this respect, the current study was developed to compare the results of three methods of parameter estimation, for the same database. Hourly wind speed data of the Cape Town region were statistically analyzed. The probability density distributions were derived from this database and distributional parameters were evaluated. The computational results reveal that the method with the lowest values of statistical testing was MLE, for a whole year database and for diurnal values, respectively. However, in terms of accuracy, MOM and LSM also appeared to be suitable methods of assessing Weibull function since these parameters seemed to approximate those of MLE. Moreover, the associated Matlab package contained functions and tools that estimated the parameters and confidence intervals of Weibull data.

References

[1] Alroza Khaligh, Omar G. Onar, Energy Harvesting: Solar, Wind and Ocean Energy Conversion Systems, Taylor and Francis Group, LLC, New York (2010), pp. 101-109.

[2] Gary L. Johnson, Wind Energy Systems, KS, Manhattan (2006), pp. 100-105.

[3] Carta J.A., Ramirez P., "A review of wind speed probability distributions used in wind energy analysis. Case studies in the Canary Islands", Renewable and Sustainable Energy Reviews, vol. 13, issue 5, June 2009, pp. 933-955.

[4] Villanueva D., Feijoo A., "Wind power distributions: A review of their applications", Renewable and Sustainable Energy Reviews 14, 2010, pp. 1490–1495.

[5] A.N. Celik "A statistical analysis of wind power density based on the Weibull and Rayleigh models at the southern region of Turkey," Renew. Energy 29, 2003, pp.593–604.

[6] Ahmad Mahir Razali, Ali A. Salih and Asaad A. Mahdi Estimation Accuracy of Weibull Distribution Parameters. Journal of Applied Sciences Research 5, 2009, pp. 790-795.

[7] C. Nemes, M. Istrate, "Effects of wind profile in wind energy systems performance," 6th International Workshop on Deregulated Electricity Market Issues in S-E Europe, Bled, Slovenia, 2011, pp. 95-100.

[8] Bain L., Engelhardt M., Introduction to Probability and Mathematical Statistics. Duxbury Press, California (1992).

[9] Lange P. M., "On the uncertainty of wind power predictions – analysis of the forecast accuracy and statistical distributions of errors", Journal of Solar Energy Engineering, vol. 127, 2005, pp. 177-184.

[10] Isaac Y. F. Lun, Joseph C. Lam, "A study of Weibull parameters using long-term wind observations". Renewable Energy Journal, Vol 20, Issue 2, June 2000, pages 145-153.