DOI: 10.14621/ce.20180103

Wind Energy Modelling and Investigation using Statistical Methods: A Case Study of El-Sherouk City in Egypt

Bahy Abdel-Mesih[1]*, Mohamed Osman[2]

[1)]Appalachian State University, Department of Sustainable Technology and the Built Environment
Boone, North Carolina. USA; abdelmesihbs@appstate.edu

[2)]The British University in Egypt, Faculty of Engineering
Suez Desert Road, El-Sherouk City, 11837 Cairo, Egypt

Abstract

Wind energy is a well-established and reliable source of clean non-polluting energy. The power available in wind is proportional to the cube of wind speed. Therefore wind speed is the most critical data that needs to be analysed in order to assess the energy potential at a certain site. In this work, wind data is collected via the weather station installed at The British University in Egypt (BUE) in El-Sherouk City; northeast of Cairo, Egypt. As a new and large urban community, El-Sherouk City's wind energy potential has not been investigated yet. Therefore the objective of this work is to mathematically model and investigate wind energy potential using accurate wind data published for the first time. Monthly mean wind speed, and standard deviation for a period of two years are presented. The annual and monthly Weibull parameters are determined. For this task, four statistical methods are compared: 1 - Moment Method (MOM), 2 - Least Linear Square Method (LLSM), 3 - Statistical Maximum Likelihood Estimator Method (SMLEM) and 4 - Energy Pattern Factor Method (EPFM). The accuracy of the results of each method is investigated using 1 - Mean Square Error (MSE), 2 - R Square Method (R^2) and 3 - Chi Square Method (χ^2). Finally the paper presents the estimated annual and monthly power densities at different heights.

Keywords:	Wind energy modelling; Wind data statistical analysis; Weibull distribution parameters; Wind power characteristics

Article history:	Received:	27 March 2017
	Revised:	04 November 2017
	Accepted:	19 January 2018

1. Introduction

With high population increase rate, Egypt is struggling to meet its growing energy needs. Egypt's demand for electricity increases rapidly between 1.5 and 2 GW per year due to urbanization and economic growth [1]. Recently Egypt start to focus on increasing its share in renewable energy. According to Egypt 's New and Renewable Energy Authority (NREA) the renewable energy share of local power consumption will reach 20 % on 2022 [2]. Moreover, in order to meet the increasing demand on power consumption for the new industrial projects an additional 13 GW must supplied by 2020. Due to the introduction of the feed-in tariff (FIT) introduced in 2014, Egypt will procure 4.3 GW of renewable power production by 2017. Accordingly, the solar photovoltaic market and wind market are projected to grow cumulatively to approximately 2 to 3 GW each by 2020 [3]. Therefore, with the increasing demand on electricity, general attitude for using renewable energy as main source for electricity generation, the need for more accurate investigation for wind power potential for urban communities arises. One of most developing urban communities in Egypt is El-shorouk city which located on the northeast board of Cairo.

In order to mathematically model the wind speed for those two areas it was decided to use Weibull distribution, which depends on two parameters namely; the shape parameter (K) and the scale parameter (C), the calculation of those two parameters will be done using four methods 1 - Moment Method (MOM), 2 - Least Linear Square Method (LLSM), 3 - Statistical Maximum Likelihood Estimator Method (SMLEM) and 4 - Energy Pattern Factor Method (EPFM), then the accuracy of each method will be investigated using 1 - Root Mean Square Error (RMSE), 2 - R Square R and 3 -

Chi Square χ^2 . Finally, the most accurate method will be incorporated in the site's actual power potential calculation. The power potential will be investigated using the assumption of ideal wind power $P_w = 1/3\ \rho V^3$ (W/m^2).

2. Weibull distribution

A number of studies have focused on assessing the wind energy potential of specific sites around the world using statistical modelling techniques [4, 5]. The most widely used technique is the Weibull distribution function that aims at establishing a good fit approximation for the gathered experimental data. The shape of Weibull distribution curve is determined by two parameters, namely the shape parameter (K) and the scale parameter (C) [6, 7, 8]. Where C is the most probable measured wind speed and is related to the mean wind speed via the relation:

$$v_{avg} = c\left[0.568 + \frac{0.434}{K}\right]^{1/K} \quad (1)$$

The Weibull probability distribution function f(v) is:

$$f(v) = \frac{K}{C}(\frac{v}{C})^{K-1} exp^{(-\left(\frac{v}{C}\right)^K)} \quad (2)$$

3. Weibull parameters estimation methods

3.1. Statistical Maximum Likelihood Estimator Method

Maximum Likelihood Estimator Method is a popular method that is used to determine the Weibull parameters K and C. MLEM is based on the mathematical likelihood function expression as [9, 10]

$$\mathrm{L} = \prod_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{f_i}(\mathrm{x_i}\,, \theta) \quad (3)$$

f_i (X_i, θ) is the probability function for a set of data with sample size n like X_1, X_2, ……… X_n and θ is the unknown parameter. L is the likelihood of this sample data; is the probability of having this sample data given the probability function f_i (X_i, θ). Therefore, the value of θ, which maximizes the (L), is called the Maximum Likelihood Estimator. The solution is by applying $\frac{dln(L)}{d\theta} = 0$ to equation (3), so we get:

$$L(x_1 \dots x_n, k, c) = \prod_{i=1}^{n}(\frac{k}{c})(\frac{x_i}{c})^{k-1}\, exp^{-(\frac{x_i}{c})^k} \quad (4)$$

By taking the partial derivative once with respect to K and second time with respect to C and then eliminating C from both equations and simplifying, we obtain a relation for K:

$$K = \left[\frac{\sum_{i=1}^{n} x_i^k \ln(x_i)}{\sum_{i=1}^{n} x_i^k} - \frac{1}{n}\sum_{i=1}^{n}\ln(x_i)\right]^{-1} \quad (5)$$

Solving for scale parameter C we get:

$$c = (\frac{\sum(v^k)}{n})^{\frac{1}{k}} \quad (6)$$

Equation (5) shows that the only way to determine K is to use iterative method. As a result, Christofferson and Gillette [11] introduced a modification on the method where a statistical method used to solve equation (5) from the collected data without wasting time in iterations. This method uses statistics to solve expectation beta E(β) which is a function of K only and then applying L' Hospital's rule twice we have:

$$K = \frac{\pi}{\sqrt{6}} * \sqrt{\frac{n(n-1)}{n(\sum ln^2(v)) - (\sum \ln(v))^2}} \quad (7)$$

3.2. Least Square Method

The LLSM is a direct numerical method used to determine the linear or nonlinear equation for the best fit line for a set of data without the need for drawing it. This is done by minimizing the error between each point and the best fit line. Therefore it was decided to use this method instead of the Graphical Method. In order to avoid the sign difference between the points under and above the best fit line the error is squared. Weibull parameters are computed based on the X and Y coordinates of any given data as follows:

$$K = \frac{n\sum_i X_i Y_i - \sum_i X_i Y_i}{n\sum_i (X_i)^2 - (\sum_i X_i)^2} \quad (8)$$

and

$$C = -(exp^{\frac{\sum_i (X_i)^2\ \sum_i Y_i - \sum_i X_i \sum_i X_i Y_i}{n\sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}} \quad (9)$$

3.3. Moment Method

C. G. Justus [12] introduced an empirical method considered as special case from moment method which gives an acceptable approximate solution for K. K is a function of average speed and standard deviation σ:

$$K = (\frac{\sigma}{\bar{v}})^{-1.086} \quad (10)$$

This approximation is valid in case that K is ranges from one to ten. Once K is determined C could be calculated as well:

$$C = \frac{\bar{v}}{\Gamma(1+\frac{1}{K})} \quad (11)$$

3.4. Energy Pattern Factor Method

The EPFM is a numerical method that determines K using an equation that depends on the average energy of the wind:

$$K = 1 + \frac{3.69}{(E_{pf})^2} \quad (12)$$

where E_{pf} is the energy pattern factor,

$$E_{pf} = \frac{\frac{1}{n}\sum_{i=1}^{n} v^3}{(\frac{1}{n}\sum_{i=1}^{n} v)^3} \quad (13)$$

C is computed from equation (6).

4. Evaluation criteria

4.1. Root Mean Square Error Method

The RMSE determines the fitting of the model to the original data, this is done by calculating the average error *e* between each point of model and original data [13].

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e^2} \quad (14)$$

4.2. R^2 Method

At this method, a number called R^2 is calculated; where this number represents how much the best fit line actually coincides with the gathered data. The R^2 value ranges from 0 to 1 and the more the value the better the model matches the experimental data.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_d - Y_c)^2}{\sum_{i=1}^{n}(Y_d - \overline{Y_d})^2} \quad (15)$$

Where Y_d is the original data value and Y_c are the calculated value.

4.3. χ2 Method

The Chi – Square (χ^2) is another statistical method that aims at assessing the goodness of a data fitting. The more the value approaches zero, the less significant is the level of difference between the original and calculated data [14]. Y_d is the real data cumulative probability and Y_c is the calculated cumulative probability.

$$X^2 = \sum_{i=1}^{n}\frac{(Y_d - Y_c)^2}{Y_c} \quad (16)$$

5. Original data analysis

The data collected from the weather station at the British University in Egypt (BUE), El-Sherouk City, Cairo, is 207 meter above mean sea level ($30^0$07'05.72''N and $31^0$36'32.54''E). The anemometer is placed at a height 10-meter with average wind speed is collected every five minutes. The graph on Figure 1 shows the measured mean wind speed over a period of two consecutive years. This chart shows that the year can be divided to two periods the first from January till July with mean wind speed 3.4 m/s and the second period from August till December with mean wind speed 3 m/s. The annual mean speed is 3.25 m/s.

Table 1 shows the wind speed standard deviation over the two-year period. Since the standard deviation of two months January and February is relatively higher than the other months, it is expected to have wider range of wind speed variation from the mean during these two months. In addition, the annual mean standard deviation shows a wide range of variation in wind speed from the mean.

Figure 1. Monthly average wind speed at 10 m height over two years

Table 1: Monthly wind speed Standard Deviation over a period of two years at height 10 m

Mean Wind Speed	Month												Annual SD (m/s)
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	
V_{10}	2.13	2.20	1.67	1.64	1.65	1.58	1.40	1.32	1.46	1.43	1.63	1.69	1.67

Table 2: Energy pattern Factor

Parameter	Month												Annual mean
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	
K	1.50	1.60	2.16	2.20	2.19	2.37	2.80	2.44	2.27	2.16	1.87	1.80	2.06
C	3.50	3.84	3.90	3.90	3.83	3.90	3.78	3.38	3.50	3.44	3.30	3.32	3.67

Table 3: Moment Method

Parameter	Month												Annual mean
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	
K	1.58	1.72	2.22	2.26	2.30	2.45	2.85	2.43	2.30	2.15	1.95	1.92	2.14
C	3.50	3.86	3.90	3.91	3.87	3.90	3.78	3.38	3.50	3.44	3.30	3.30	3.67

Table 4: Statistical Maximum Likelihood Estimator Method

Parameter	Month												Annual mean
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	
K	2.03	2.10	2.37	2.40	3.57	2.58	2.68	2.26	2.27	2.00	2.10	2.23	2.27
C	3.79	4.10	3.96	3.97	3.99	3.95	3.74	3.34	3.50	3.40	3.37	3.46	3.73

Table 5: Least Linear Square Method

Parameter	Month												Annual mean
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	
K	1.91	1.71	2.40	2.50	2.38	2.38	2.67	2.39	2.29	2.27	1.97	2.05	1.87
C	3.73	4.24	3.82	3.80	4.14	4.04	3.34	2.90	3.37	3.00	3.36	3.50	3.96

6. Weibull probability distribution analysis

At this section, the K and C for each mentioned set of data is calculated using the four previous mentioned methods. Tables 2–5 show the results of Weibull parameters identification.

Previous tables show that both the MOM and the EPFM give the same value for C and very close value for K. Therefore both the MOM and the EPFM will give nearly the same probability distribution with one curve little narrower and higher than the other.

Figure 2 shows the Weibull probability distribution retrieved from each method and the original data. The

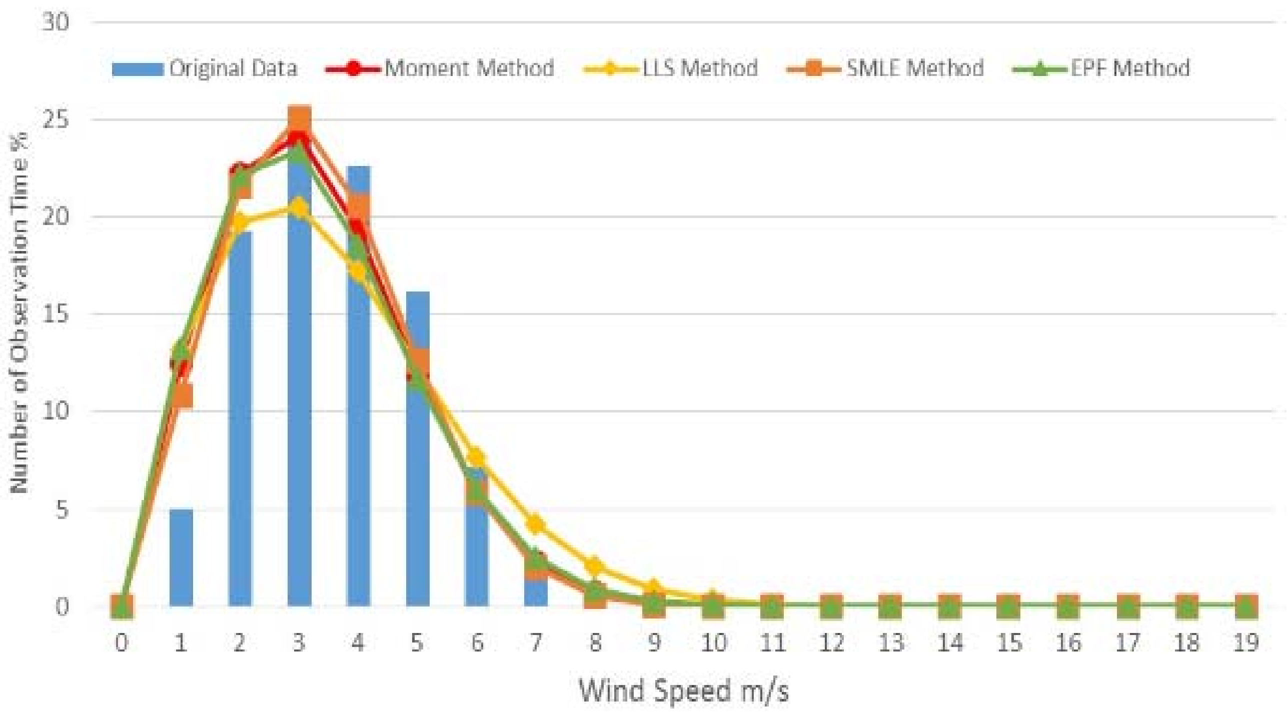

Figure 2. Annual Weibull Probability Distribution using the four modelling methods

Table 6: Weibull Probability Distribution Accuracy Estimation

Method / Accuracy	EPFM	MOM	SMLEM	LLSM
RMSE	0.0113	0.0094	0.0142	0.0345
R^2	0.9998	0.9999	0.9997	0.9983
X^2	0.0071	0.004	0.009	0.0399

EPFM gives the best fit for the original data. LLSM shows the worst fitting for the data. In addition, the MOM shows the second best fit and the SMLEM shows the third best fit. In order to better evaluate the previous methods, Table 6 shows the results for the three evaluation methods discussed in section 4.

Table 6 shows that all the methods give a very good result with very slightly difference between them not exceeding 0.05. However, the RMSE and R^2 show that the EPFM and the MOM have a better accuracy than other two methods. SMLEM follows by a small margin of accuracy. In addition, it shows that the LLSM have less accuracy than the other methods. Finally, the Chi-square test shows that there is very little significant difference between the modelling techniques and the original data.

7. Power density

The parameters K and C obtained from the EPFM are used to calculate the power density defined as:

$$P_{dw} = \int_{V_{min}}^{V_{max}} PD(V)\, f(v)\, dv \qquad (17)$$

Where V is the wind speed, P_{dw} is the output power density at this wind speed and f(V) is the Weibull statistical probability distribution [15]. Equation (17) is compared to the power density, equation (18) with results furnished in Table 7 and Figure 3.

$$P_d = \frac{1}{2}\rho v_w^3 \qquad (18)$$

The chart on Figure 3 shows very slightly difference between the two previously mentioned methods results, this mean that the Weibull distribution using Energy Patter Factor Method give highly accurate results with mean error 1 % positive or negative. In addition, it shows that power density ranged from 26 W/m^2 till 60 W/m^2, with lowest potential in August, and peak at February. The yearly mean power density is 38.7 W/m^2.

Table 7: Monthly wind power density at 10 m high

Wind power density (W/m^2)	Month												Annual mean
	Jan.	Feb.	Mar.	Apr.	May	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	(W/m^2)
P	51.12	59.7	43.6	43	41.3	42.5	36.6	26.2	30.5	28.7	29.2	31.8	38.7
P_{dw}	52.5	62	45	44.3	42.1	41.7	34.3	26.6	31.2	30.9	31.7	33.9	39.2
Bias error	-0.38	-2.3	-1.4	-1.3	-0.8	0.8	2.3	-0.4	-0.7	-2.2	-2.5	-2.1	-0.5

Figure 3. Variation of monthly wind power computed at 10 m height by using two different methods

8. Conclusion

The original wind data analysis shows small variation in wind speed during the year with annual mean 3.25 m/s at 10-meter height. In addition, the annual stander deviation was equal to 1.67 m/s which indicated small variation in wind speed from the mean. Finally, the wind rose chart shows that most of the year the wind direction change gradually from North - West to East with mean wind speed 3.5 m/s.

The EPFM showed best data fitting with less than 0.012 % error with no significance difference between it and the original data. This is followed by (MOM) and (SMLEM) which show second best fitting, whereas the (LLSM) shows the worst fitting for original data. Finally, based on this outcome, the Weibull parameters are to be calculated using (EPFM) for the statistical analysis of power density.

The power density calculation shows maximum power density 60 W/m^2 in February, lowest power density 26 W/m^2 in August, with annual mean power density 38.7 W/m^2. In addition, the result retrieved from this method depends on Weibull probability distribution showing an average ± 1 % bias error compared to the results retrieved using original data mean and variance. Finally, it was concluded that the site in El-Sherouk lacks the potential for large scale wind power generation and low potential for small scale wind power generation using low starting torque vertical wind turbines.

The results obtained for this site in El-Sherouk City indicates poor potential for large utility-scale wind turbine application. However, the site is still suitable for small utility scale like running home appliances using low starting torque vertical axis wind turbines [16]. This work is important for any potential investor or decision makers investigating available options in renewable energy sector for the development of northeast region of Cairo.

References

[1] Croker, Angela, Renewable energy in Egypt: hydro, solar and wind, *Norton Rose Fulbright* (2013), www:nortonrosefulbright.com/knowledge/publications/74735/renewable-energy-in-egypthydro-solar-and-wind

[2] Burger, Andrew, Egypt's renewable energy drive gains steam, *Renewable Energy World* (2015). www:renewableenergyworld:com/articles/2015/06/egypt-s-renewable-energy-drive-gainssteam.html

[3] Abdelhamid, Aisha, Egypt fast becoming a hot renewable energy market, *Clean Technica* (2015). http://cleantechnica:com/2015/03/22/egypt-fast-becoming-a-hot-renewable-energy-market/

[4] Kidmo, D. , Danwe, R., Doka,S. , Djongyang, N., Statistical analysis of wind speed distribution based on six Weibull methods for wind power evaluation in Garoua, Cameroon, *Revue des Energies Renouvelables*, *Vol. 18*, (2015), 1, pp. 105-125.

[5] Rocha, P. A. C., de Sousa ,R. C., de Andrade , C. F., da Silva , M. E. V., Comparison of seven numerical methods for determining Weibull parameters for wind energy generation in the northeast region of Brazil, *Applied Energy*, *Vol. 89*, (2012), 1, pp. 395-400.

[6] Sarkar, A., Singh, S., Mitra, D., Wind climate modelling using Weibull and extreme value distribution, *International Journal of Engineering, Science and Technology, Vol. 3*,(2011),5, pp. 100-106.

[7] Caretto, Larry, Probability distribution functions for wind/California State University Northridge, www.csun.edu/~lcaretto/me483/probability.doc

[8] Bhattacharjee, Paritosh, A study on Weibull distribution for estimating the parameters, *Journal of Applied Quantitative Methods, Vol. 5*,(2010),2, pp. 234-241.

[9] Engineering Statistics Handbook / Maximum likelihood estimation, www.itl.nist.gov/div898/handbook/apr/section4/apr412.htm

[10] Shata, Ahmed, Theoretical investigation and mathematical modelling of a wind energy system case study for Mediterranean and red sea, Ph.D. thesis, Berlin University, Berlin, Germany, 2008.

[11] Christofferson, Rolan, Gillette, Dale, A simple estimator of the shape factor of the two-parameter Weibull distribution, *Journal of Applied Meteorology and Climatology, Vol. 26*, (1987), 2, pp.323–325, DOI: http://dx.doi.org/10.1175/1520-0450(1987)026<0323:ASEOTS>2.0.CO;2

[12] Justus, Carl. *Winds and wind system performance (Solar energy series)*, Franklin Institute Press, Philadelphia, USA, 1978.

[13] Chai ,T., Draxler ,R. R., Root mean square error (RMSE) or mean absolute error (MAE)?- Arguments against avoiding RMSE in the literature, *Geoscientific Model Development*, *Vol. 7*, (2014), 3, pp. 1247-1250, DOI: 10.5194/gmd-7-1247-2014

[14] Chang, Tian, Wind speed and power density analyses based on mixture Weibull and maximum entropy distributions, *International Journal of Applied Science and Engineering, Vol. 8*, (2010), 1, pp. 39-46.

[15] Shata, Ahmed, Potential wind power generation in south Egypt, *Renewable and Sustainable Energy Reviews, Vol. 16*, (2012), 3, pp. 1528-1536.

[16] Dixon, Maureen, Hall, Cesare, *Fluid Mechanics and Thermodynamics of Turbomachinery*, Seventh edition, Elsevier Inc., Massachusetts, USA, 2010.