Reliability analysis for wind turbines with incomplete failure data collected

Haitao Guo[*], Simon Watson[*], Peter Tavner[+], Jiangping Xiang[*]

[*]Centre for Renewable Energy Systems Technology, Loughborough University, LE11 3TU ,UK
[+]School of Engineering, New & Renewable Energy, Durham University

Abstract

Reliability has an impact on wind energy project costs and benefits. Both life test data and field failure data can be used for reliability analysis. In wind energy industry, wind farm operators have greater interest in recording wind turbine operating data. However, field failure data may be tainted or incomplete, and therefore it needs a more general mathematical model and algorithms to solve the model. The aim of this paper is to provide a solution to this problem. A three parameter Weibull model is discussed and the parameters are estimated by Maximum Likelihood and Least Squares. The average failure numbers of the wind turbines in Denmark and Germany are used for this study. The traditional Weibull model is also employed for comparison. Analysis shows that the three parameter Weibull model can obtain more accuracy on reliability growth of turbine lifetime. The proposed three parameter Weibull model is also applicable to the life test of the components in use to shorten testing time. This work will be helpful in the understanding of the mechanical behaviour, durability and management aspects of wind energy systems.

Key words: reliability; failure rate; three parameter Weibull model; maximum likelihood; least squares; wind turbines

1. Introduction

Wind turbines are a renewable source of energy and will play an increasingly important role in providing electricity, because wind turbines capacities and the number of grid-connected wind turbines are increasing. Wind turbine reliability is a significant factor in ensuring the success of a wind power project. Walford outlined the issues relevant to wind turbine reliability for wind turbine power generation projects and the relationship between wind turbine reliability and operation & maintenance costs[1]. Reliability and condition monitoring apparently benefit the maintenance management of wind power systems [2] by reducing the O&M costs by giving advance warning of failures. Some authors have combined wind power generation and wind speed models to analyze power production reliability[3][4]. Others have applied Probabilistic Safety Assessment (PSA) to wind turbines to assess system reliability qualitatively and quantitatively, based upon component failures[5][6]. Another authors used the monthly variation of energy production to weight the shutdown time, which included both maintenance and fault hours. This showed a large difference exists between the original downtime and weighted downtime[7].

After a wide review, Herbert etc. concluded that fewer authors have worked on reliability evaluation of wind turbine systems[8]. Valuable information for wind turbine reliability analysis can be derived from failure data by statistical analysis [9][10]. Climate change can also be taken into account using statistical data [11].

In wind energy industry, there are plenty of field data available for reliability analysis, but the field failure data is usually tainted or incomplete. In most of the cases, data is not collected from the time of the installation of wind turbines and the population of the investigated wind turbines changes. To study the reliability characteristics using the available field data, a new model considering the above problems is presented in this paper. The new model is a three parameter Weibull model that uses a third parameter to look into the past running time. The three parameters are estimated by two techniques, Maximum Likelihood and Least Squares. Two wind turbine populations are analyzed using the presented model and traditional Weibull model. The data is extracted from Windstats Newsletters[17]. The results of both models are compared.

2. Windstats Data

Windstats Newsletter is a quarterly international publication which provides various information about the wind energy converted in wind turbines in various countries in the world. The data analyzed in this paper has been extracted from Windstats Newsletters from wind turbines in Demark and Germany. The data collects the numbers of wind turbine subassembly failures in a fixed interval, one month for Danish turbines and one quarter for German turbines. To simplify the problem and concentrate on the methods demonstrated in this paper, it is assumed that any subassembly failure will lead to a wind turbine failure. By that assumption, the wind turbine failures in an interval is equal to the sum of the subassembly failures.

Danish data starts from Oct. 1994 to Dec. 2003 with population varying from highest 2345 turbines to lowest 851 turbines; German data starts from Dec. 1995 to Sep. 2004 with population varying from highest 4285 turbines to lowest 1578. Danish data shows a decreasing number of installed turbines, while German wind turbines increase rapidly[9][10]. Since the population changes, it's necessary to eliminate the population difference by normalizing the wind turbine reliabilities. Furthermore, individual wind turbines have similar subassemblies and architecture. Therefore, the number of failures in an interval is divided by corresponding number of turbines to get the average failure number of that interval, which means the number of failures per interval per turbine and is suitable for modelling the reliability of the wind turbines.

The interval of Danish data is not converted into a quarter (the interval of the German data) by synthesizing data for 3 months, because this paper is aimed at providing reliability analysis methods for wind turbines using incomplete failure data recorded in different intervals.

Windstats Newsletter also provides additional information, besides failure numbers, about wind turbines such as production, capacity factor, which can be used to analyze other aspects of wind turbines[11].

3. Weibull Model

The Weibull process is widely used to model the non-homogeneous Poisson process. Its intensity function (or failure rate) is

$$\lambda(t) = \phi \cdot t^{\mu} \tag{1}$$

With $\mu<0$, $\mu=0$, $\mu>0$, the Weibull model can depict infant mortality stage, normal stage and wear-out stage of the so-called bathtub curve respectively. Equation (1) can also be written as the popular equation (2) , which simplifies calculating the integral of intensity function,

$$\lambda(t) = \rho\beta t^{\beta-1} \tag{2}$$

where $\beta=\mu+1$ and $\rho=\phi/\beta$. In some literatures, ρ is called the scale parameter and β the shape parameter.

For a Poisson process, the probability of N events occurring over period $(a, b]$ is [12]

$$P\{n(a,b]=N\} = \frac{\left[\int_a^b \lambda(x)dx\right]^{N_i} e^{-\int_a^b \lambda(x)dx}}{N!} \tag{3}$$

The failure data of wind turbines collected by Windstats is grouped data, monthly for Danish wind turbines and quarterly for German. It is assumed that the distribution of wind turbine failures is of Weibull and individual groups are independent to each other. Thus, a joint probability distribution function PDF of k grouped data is

$$P\{n(t_0,t_1]=N_1, n(t_1,t_2]=N_2, \cdots, n(t_{k-1},t_k]=N_k\}$$
$$= \prod_{i=1}^k P(n(t_{i-1},t_i]=N_i) \tag{4}$$

Combined with equation (2) and (3), the following is derived

$$P\{n(t_0,t_1]=N_1, n(t_1,t_2]=N_2, \cdots, n(t_{k-1},t_k]=N_k\}$$
$$= \prod_{i=1}^k \frac{\left[\rho(t_i^{\beta}-t_{i-1}^{\beta})\right]^{N_i} e^{-\lambda(t_i^{\beta}-t_{i-1}^{\beta})}}{N_i!} \tag{5}$$

Next, calculate the maximum likelihood estimates of β and ρ.

$$\text{Ln} P\{n(t_0,t_1]=N_1, n(t_1,t_2]=N_2, \cdots, n(t_{k-1},t_k]=N_k\}$$
$$= \sum_{i=1}^k N_i \text{Ln} \rho + \sum_{i=1}^k N_i \text{Ln}(t_i^{\beta}-t_{i-1}^{\beta}) - \sum_{i=1}^k \rho(t_i^{\beta}-t_{i-1}^{\beta}) - \sum_{i=1}^k N_i! \tag{6}$$

Then

$$\frac{\partial \text{Ln} P}{\partial \rho} = \sum_{i=1}^k \frac{N_i}{\rho} - \sum_{i=1}^k (t_i^{\beta}-t_{i-1}^{\beta}) \tag{7}$$

By Letting $\dfrac{\partial \text{Ln} P}{\partial \rho} = 0$ and if $t_0=0$, the estimate of ρ is

$$\hat{\rho} = \frac{\sum_{i=1}^k N_i}{t_k^{\beta}} \tag{8}$$

Substitute ρ in equation (6) by $\hat{\rho}$ of equation (8)

$$\text{Ln} P\{n(t_0,t_1]=N_1, n(t_1,t_2]=N_2, \cdots, n(t_{k-1},t_k]=N_k\}$$
$$= \sum_{i=1}^k N_i \text{Ln}(\sum_{i=1}^k N_i) - \sum_{i=1}^k N_i \beta \text{Ln}(t_k) + \sum_{i=1}^k N_i \text{Ln}(t_i^{\beta}-t_{i-1}^{\beta}) - \sum_{i=1}^k N_i - \sum_{i=1}^k N_i!$$

Then

$$\frac{\partial \mathrm{Ln}\,P}{\partial \beta} = -\sum_{i=1}^{k} N_i \,\mathrm{Ln}(t_k) + \sum_{i=1}^{k} N_i \frac{t_i^{\beta}\,\mathrm{Ln}\,t_i - t_{i-1}^{\beta}\,\mathrm{Ln}\,t_{i-1}}{t_i^{\beta} - t_{i-1}^{\beta}}$$

$$= \sum_{i=1}^{k} N_i \left[\frac{t_i^{\beta}\,\mathrm{Ln}\,t_i - t_{i-1}^{\beta}\,\mathrm{Ln}\,t_{i-1}}{t_i^{\beta} - t_{i-1}^{\beta}} - \mathrm{Ln}\,t_k \right] \tag{9}$$

Again, by Letting $\dfrac{\partial \mathrm{Ln}\,P}{\partial \beta} = 0$ and $t_0 \mathrm{Ln} t_0 = 0$, we can get $\hat{\beta}$ by solving the following

formula

$$\sum_{i=1}^{k} N_i \left[\frac{t_i^{\beta}\,\mathrm{Ln}\,t_i - t_{i-1}^{\beta}\,\mathrm{Ln}\,t_{i-1}}{t_i^{\beta} - t_{i-1}^{\beta}} - \mathrm{Ln}\,t_k \right] = 0 \tag{10}$$

The above result is the same as that in [13]. In this paper, time interval T is fixed, which means $t_i = iT$, then the above formula is changed to

$$\sum_{i=1}^{k} N_i \left[\frac{i^{\beta}\,\mathrm{Ln}(iT) - (i-1)^{\beta}\,\mathrm{Ln}[(i-1)T]}{i^{\beta} - (i-1)^{\beta}} - \mathrm{Ln}(kT) \right] = 0 \tag{11}$$

Then, the estimates of ϕ and μ can be derived from ρ and β.

4.   A Three parameters Weibull Model

The Weibull process is a Non-homogeneous Poisson Process with the intensity function of equation (1). According to the statements above, the number of failures in fixed intervals are known, so the parameters ϕ and μ should be estimated using those measurements. After rewriting equation (1) in term of equation (2), the number of failures, Λ, over period $(0, t)$ is calculated as below:

$$\Lambda(t) = \int_0^t \lambda(x)dx = \rho t^{\beta} \tag{12}$$

The above equation is suitable for the situations where the data is recorded from the wind turbine installation. However, the data collected in Windstats is not necessarily from the date of wind turbine's installation as it may start from some years after later. Another parameter α called time factor in this paper is reasonably introduced into the Weibull model in order to describe the past running time. The α shifts the intervals along the time axis from 0, T, …, kT to αT, $\alpha T + T$, ..., $\alpha T + kT$. T for Danish turbines is the number of hours of a month and a quarter for German turbines. Therefore, the task has changed to estimate the parameters ϕ, α, μ ☐ from the average failure numbers Λ_1, Λ_2, …, Λ_k.

4.1. Maximum Likelihood Estimates

After α is introduced, equation (1) becomes

$$\lambda(t) = \phi(t + \alpha T)^{\mu} \tag{13}$$

Correspondingly, equation (2) changes into

$$\lambda(t) = \rho\beta(t + \alpha T)^{\beta-1} \tag{14}$$

Following the same procedures in the last section, the estimate of ρ can be derived as

$$\hat{\rho} = \frac{\sum\limits_{i=1}^{k} N_i}{[(k+\alpha)T]^{\beta} - (\alpha T)^{\beta}} \tag{15}$$

and the estimates of α and β are the solutions of the equations below which are solved by trust-region dogleg method[14].

$$\sum_{i=1}^{k} N_i \left[\frac{(i+\alpha)^{\beta} \operatorname{Ln}[(i+\alpha)T] - (i-1+\alpha)^{\beta} \operatorname{Ln}[(i-1+\alpha)T]}{(i+\alpha)^{\beta} - (i-1+\alpha)^{\beta}} - \frac{(k+\alpha)^{\beta} \operatorname{Ln}[(k+\alpha)T] - \alpha^{\beta} \operatorname{Ln}(\alpha T)}{(k+\alpha)^{\beta} - \alpha^{\beta}} \right] = 0 \tag{16}$$

$$\frac{\beta}{T} \sum_{i=1}^{k} N_i \left[\frac{(i+\alpha)^{\beta-1} - (i-1+\alpha)^{\beta-1}}{(i+\alpha)^{\beta} - (i-1+\alpha)^{\beta}} - \frac{(k+\alpha)^{\beta-1} - \alpha^{\beta-1}}{(k+\alpha)^{\beta} - \alpha^{\beta}} \right] = 0 \tag{17}$$

Then, the estimates of ϕ and μ can be derived from ρ and β.

4.2. Least squares Estimates

From Equation (12), the number of failures in each interval is given by

$$\Lambda_1 = \rho t^{\beta} \Big|_{\alpha T}^{(1+\alpha)T} = \rho((1+\alpha)T)^{\beta} - \rho(\alpha T)^{\beta}$$

$$\Lambda_2 = \rho t^{\beta} \Big|_{(1+\alpha)T}^{(2+\alpha)T} = \rho((2+\alpha)T)^{\beta} - \rho((1+\alpha)T)^{\beta}$$

$$\vdots$$

$$\Lambda_k = \rho t^{\beta} \Big|_{(k-1+\alpha)T}^{(k+\alpha)T} = \rho((k+\alpha)T)^{\beta} - \rho((k-1+\alpha)T)^{\beta}$$

Thus, for the interval i, the average number of failures is

$$\Lambda_i = \rho((i+\alpha)T)^{\beta} - \rho((i-1+\alpha)T)^{\beta} \qquad i=1,....,k \tag{18}$$

Define

$$\xi = \rho T^{\beta} \tag{19}$$

then Equation (18) becomes

$$\Lambda_i = \left\{ (i+\alpha)^{\beta} - (i-1+\alpha)^{\beta} \right\} \xi, i=1,...,k \tag{20}$$

Equation (20) can then be written in matrix form as:

$$\mathbf{\Lambda} = \mathbf{C}\xi \tag{21}$$

Where

$$\mathbf{C} = \begin{bmatrix} (1+\alpha)^{\beta} - (0+\alpha)^{\beta} \\ (2+\alpha)^{\beta} - (1+\alpha)^{\beta} \\ \cdot \\ \cdot \\ \cdot \\ (k+\alpha)^{\beta} - (k-1+\alpha)^{\beta} \end{bmatrix} \tag{22}$$

And

$$\mathbf{\Lambda} = \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \\ . \\ . \\ . \\ \Lambda_k \end{bmatrix} \tag{23}$$

Next, the least squares estimate of ξ can be derived as following

$$\hat{\xi} = (\mathbf{C}^T\mathbf{C})^{-1}\mathbf{C}^T\mathbf{\Lambda} \tag{24}$$

which is a function of α and β. Inserting Equation (24) into Equation (20) gives the estimate of Λ_i

$$\hat{\Lambda}_i = [(i+\alpha)^\beta - (i-1+\alpha)^\beta]\hat{\xi}(\beta,\alpha) \tag{25}$$

Now the problem becomes

$$\min_{\beta,\alpha} Q(\beta,\alpha) \tag{26}$$

Where

$$Q(\beta,\alpha) = \sum_{i=1}^{k}(\Lambda_i - \hat{\Lambda}_i)^2 \tag{27}$$

Since Q is highly non-linear, the above minimum is solved by a large-scale algorithm which is a subspace trust region method and is based on the interior-reflective Newton method[15][16]. With the values α and β, by equations (19) and (24) an estimate for the parameter ρ can be obtained. Then, the estimates of ϕ and μ can be derived from ρ and β.

5. **Modelling Results Analysis**

The two algorithms presented above are applied to Danish and German populations of Windstats data, analyzing the reliability growth. Table 1 and 2 give estimated parameters from the two algorithms, but Danish data is estimated in monthly intervals and German data is estimated in quarterly intervals. The average relative errors are calculated as following

$$Error_{avg\_rel} = \frac{1}{k}\sum_{i=1}^{k}\frac{\left|\Lambda_i - \hat{\Lambda}_i\right|}{\hat{\Lambda}_i} \tag{28}$$

Danish population has a larger time factor(α) than German, which implies that Danish wind turbines are put into use earlier than German wind turbines. Windstats data (2003) confirms Danish first turbine installations was in 1987 and German first turbine installations was in 1990. Therefore, combined with the starting points of data of two populations (Danish is Oct. 1994 and German is Dec. 1995), those α estimated by ML in the tables show their consistence. Those α achieved by LS have a little bit larger deviation. The ML estimates coincide with the LS estimates when the noise is zero-mean Gaussian distributed. The noise in Windstats data is not of that characteristic[11]. ML and LS give similar intensity distribution functions, which explains that similar curves (in Fig. 5 and Fig. 6) are obtained by ML and LS.

6

Table 1. Estimated parameters of the 3-parameter Weibull Model for monthly Danish data

α		ϕ		μ		Average Relative Error	
LS	ML	LS	ML	LS	ML	LS	ML
112.32 (9.36 yrs)	93.7 (7.80 yrs)	9.93	7.33	-0.999999	-0.985097	0.29	0.29

Table 2. Estimated parameters of the 3-parameter Weibull Model for quarterly German data

α		ϕ		μ		Average Relative Error	
LS	ML	LS	ML	LS	ML	LS	ML
26.45 (6.61 yrs)	21.42 (5.36 yrs)	16.94	11.32	-0.999985	-0.976861	0.24	0.24

Failure numbers and failure rates of Danish wind turbines are shown in figure 1 and figure 2, while failure numbers and Failure rates of German wind turbines are shown in figure 3 and figure 4. Results of traditional Weibull model and 3-parameter Weibull model(Least squares estimates) are also illustrated in the figures. In the figures of failure numbers, Windstats data is also given(asterisks). It should be noted that failure number is the numbers within an interval, a month for Danish turbines and a quarter for Greman turbines. The starting times(αT) are marked out. Figure 5 and figure 6 show the comparison of failure number and failure rate between Maximum Likelihood estimates and Least squares estimates with Danish data analyzed. In figure 5, a copy of Windstats data is marked as circles for ML result. Besides, it can be seen that:

- The curves of failure number vs. time has a similar shapes to those of failure rate vs. time. This applies to both populations and both techniques.
- The Danish and German curves have similar shapes in accordance with the estimated values of μ which are close between the two populations.
- Failure numbers and failure rates of 3-parameter model are rather high close to time zero, whereas those of 2-parameter model are rather high close to time αT(years). That is a shortcoming of Weibull modelling. The monthly failure number by the first year per turbine is about 0.73.
- The 3-parameter model provides more information about the period before failure data was collected, but 2-parameter model can only be used to predict reliability performance for the subsequent period.
- The wind turbines are all shown to be in the stage of infant mortality, because all the values of μ are less than zero.
- Similar curves are obtained by applying ML and LS techniques to 3-parameter models.
- LS gives a larger α than ML, but the time factor differences between two populations using two techniques are close to each other. From table 1, it can be calculated that the time factor difference is 2.75 years for LS and 2.44 years for ML.
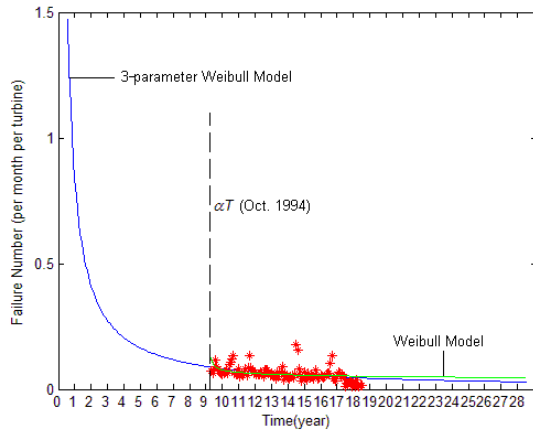
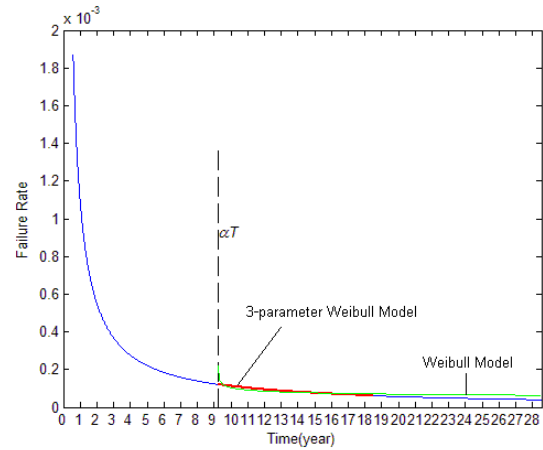Figure 1. Failure number of Danish wind turbines

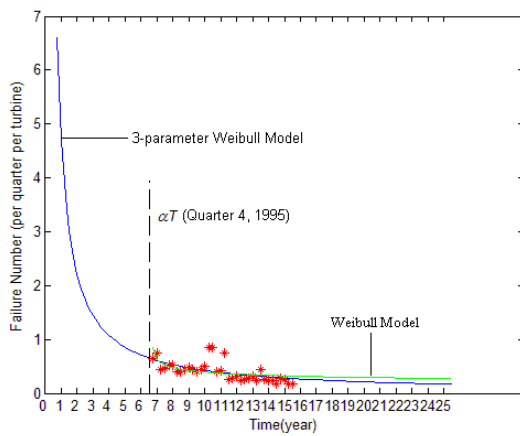Figure 2. Failure rate of Danish wind turbines

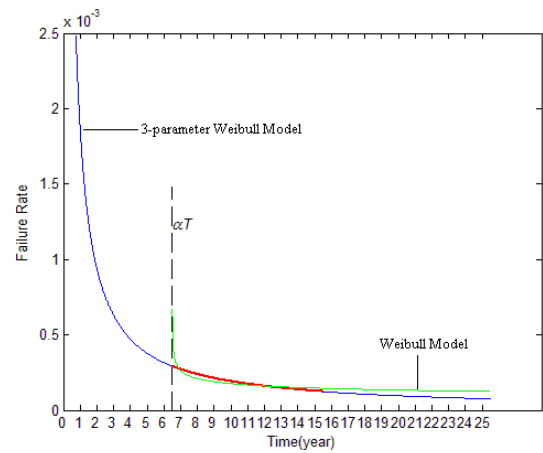Figure 3. Failure number of German wind turbines

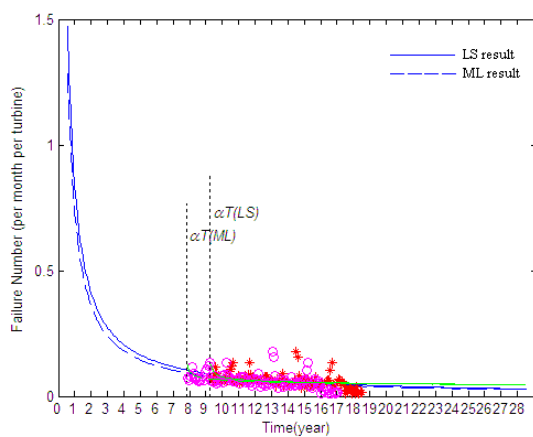Figure 4. Failure rate of German wind turbines

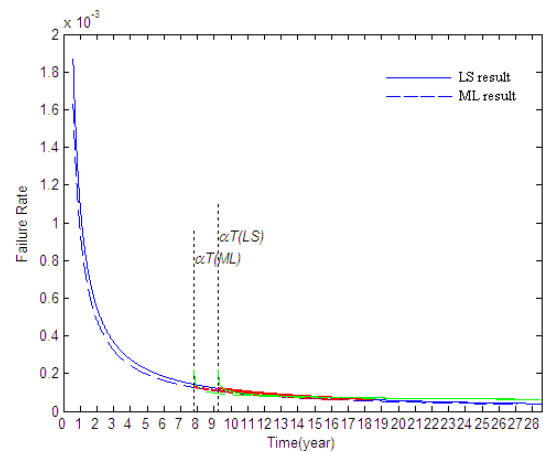Figure 5. Failure number comparison of ML and LS estimate results (Danish population)

Figure 6. Failure rate comparison of ML and LS estimate results (Danish population)

6. Conclusions

Windstats Newsletters provide failure data of wind turbines, but it is incomplete. In order to take such incompletion into account and obtain a more accurate reliability growth of wind turbines, a 3-parameter Weibull model is presented in this paper and its parameters are estimated by two techniques, Maximum Likelihood and Least squares. Similar results have been achieved by the two techniques.

Three parameters Weibull model presented in this paper has advantages over traditional Weibull in dealing with incomplete data. However, three parameters Weibull model shrinks to of the traditional Weibull model on the condition that α is set to 0. That is in accordance with the fact that three parameters Weibull model is promoted from traditional Weibull model by introducing α into it. Therefore, the proposed three parameter Weibull model is a general model that is applicable to both complete data, like life test data, and incomplete data, like field failure data in Windstats.

Because three parameters Weibull model provides an extra earlier part of reliability curve, it is helpful in planning a better maintenance schedule for wind energy systems. In other words, the remaining life time of a wind turbine can be estimated as a reference of the maintenance schedule, if a period of data is available for reliability analysis.

Reference:

[1] Walford C. Wind Turbine Reliability: Understanding and Minimizing Wind Turbine Operation and Maintenance Costs. SANDIA REPORT, SAND2006-1100, 2006.
[2] Nilsson J, Bertling L. Maintenance Management of Wind Power Systems Using Condition Monitoring Systems—Life Cycle Cost Analysis for Two Case Studies. IEEE TRANSACTIONS ON ENERGY CONVERSION, 2007, 22(1):223-229
[3] Giorsetto P, Utsurogi K. DEVELOPMENT OF A NEW PROCEDURE FOR RELIABILITY MODELING OF WIND TURBINE GENERATORS.IEEE Transactions on Power Apparatus and Systems, 1983, PAS-102(1):134-143
[4] Karki R, Hu P, Billinton R. A Simplified Wind Power Generation Model for Reliability Evaluation. IEEE TRANSACTIONS ON ENERGY CONVERSION, 2006, 21(2):533-540

[5]   Seebregts A, Rademakers L, Van Den Horn B. RELIABILITY ANALYSIS IN WIND TURBINE ENGINEERING. Microelectron. Reliab., Vol. 35, Nos 9-10, pp. 1285-1307, 1995

[6]   Michos D, Dialynas E, Vionis P. Reliability and Safety Assessment of Wind Turbines Control and Protection Systems. WIND ENGINEERING, 2006, 26(6):359–369

[7]   Abderrazzaq M, Hahn B. Analysis of the turbine standstill for a grid connected wind farm (case study).Renewable Energy, 2006, 31:89-104

[8]   Herberta J, Iniyanb S, Sreevalsanc E, etc. A review of wind energy technologies. Renewable and Sustainable Energy Reviews, 2007, 11:1117–1145

[9]   Tavner P, Xiang J, Spinato F. IMPROVING THE RELIABILITY OF WIND TURBINE GENERATION AND ITS IMPACT ON OVERALL DISTRIBUTION NETWORK RELIABILITY. 18th International Conference on Electricity Distribution, Turin, 2005.

[10]   Tavner P, Xiang J, Spinato F. Reliability Analysis for Wind Turbines. Wind Energy, 2007; 10:1–18

[11]   Xiang, J., Hor, C. and Watson, S.J., "Climate Change Impacts on Future Reliability and Capacity Factors of Wind Turbines", ENERGYX 2006, SINTEF, Elsevier, Norway, June 2006, ISBN 008045108X.

[12]   Rigdon S, Basu A. Statistical Methods for the Reliability of Repairable Systems, Wiley, New York, 2000, p43.

[13]   Reliability Growth Management (Military Handbook 189). Department of Defense: Washington, DC, 1981

[14]   Moré J, Garbow B, and Hillstrom K. User Guide for MINPACK 1, Argonne National Laboratory, Rept. ANL-80-74, 1980

[15]   Coleman, T.F. and Y. Li, "An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds," SIAM Journal on Optimization, Vol. 6, pp. 418-445, 1996

[16]   Coleman, T.F. and Y. Li, "On the Convergence of Reflective Newton Methods for Large-Scale Nonlinear Minimization Subject to Bounds," Mathematical Programming, Vol. 67, Number 2, pp. 189-224, 1994

[17]   Windstats, www.windstats.com