Methods for Estimating Wind Speed Frequency Distributions

C. G. JUSTUS, W. R. HARGRAVES, AMIR MIKHAIL AND DENISE GRABER

School of Aerospace Engineering, Georgia Institute of Technology, Atlanta 30332

(Manuscript received 28 June 1977, in final form 3 November 1977)

ABSTRACT

The Weibull function is discussed for representation of the wind speed frequency distribution. Methods are presented for estimating the two Weibull parameters (scale factor c and shape factor k) from simple wind statistics. Comparison is made with a recently proposed method based on the "square-root-normal" distribution with mean wind speed and fastest mile data as input statistics. The Weibull distribution is shown to give smaller root-mean-square errors than the square-root-normal distribution when fitting actual distributions of observed wind speed. Another advantage of the Weibull distribution is the available methodology for projecting to another height the observed Weibull distribution parameters at anemometer height.

1. Introduction

For several years the Weibull distribution has been used to represent wind speed distributions for application in wind loads studies (Davenport, 1963). Recently, the Weibull distribution has also been found useful and appropriate for wind energy applications (Justus *et al.*, 1976a,b; Hennessey, 1977). The purpose of this report is to point out some of the advantages of the use of the Weibull distribution for these purposes, and to describe some simple methods for estimating the two Weibull parameters from the average wind speed and other simple statistics.

The Weibull distribution for wind speed V is expressed by the probability density function (wind speed frequency curve)

$$p(V)dV=(k/c)(V/c)^{k-1}\exp[-(V/c)^k]dV, \quad (1)$$

where c is the scale factor (units of speed) and k the shape factor (dimensionless). The equivalent cumulative probability function (wind speed duration curve) is

$$p(V\leqslant V_x)=\int_0^{V_x} p(V)dV=1-\exp[-(V_x/c)^k]. \quad (2)$$

Advantages of the Weibull distribution are that 1) it is a two-parameter distribution, depending only on c and k [hence more general than the Rayleigh distribution, which has $k=2$, and easier to work with than the more general bi-variate normal distribution, which requires five parameters], 2) in a wide number of cases (Justus *et al.*, 1976a; Hennessey, 1977) the Weibull seems to give a reasonable fit to observed distributions, and 3) with Weibull c and k values known at one height, a consistent methodology (Justus and Mikhail, 1976) can be used to adjust these parameters to another desired height. The bivariate normal distribution assumes normal distributions for each wind component and requires five parameters for representation: $\bar{u}$, σ_u, $\bar{v}$, σ_v and ρ_{uv}, where σ_u, σ_v and ρ_{uv} are, respectively, the standard deviations in u and v and the cross correlation between u and v. If the simplifying assumptions $\bar{u}=\bar{v}=\rho_{uv}=0$ and $\sigma_u=\sigma_v=\sigma$ are made, then the bivariate normal wind speed distribution reduces to the Rayleigh distribution, which is Eq. (1) with $k=2$. Allowance for k values other than $k=2$ is an empirical means of accounting for situations between these two extremes.

2. Methods for Weibull parameter estimation

There are several methods which can be used to estimate the Weibull parameters c and k, depending on which wind statistics are available and what level of sophistication in data analysis one wishes to employ.

a. *Least-squares fit to observed distribution*—METHOD 1

If the observed wind speeds are divided into n speed interval $0-V_1$, V_1-V_2, ..., $V_{n-1}-V_n$, having frequencies of occurrence $f_1, f_2, \ldots, f_n$ and cumulative frequencies $p_1=f_1$, $p_2=f_1+f_2$, ..., $p_n=p_{n-1}+f_n$, then Eq. (2) transforms to the linear form $y=a+bx$ by the relations

$$x_i=\ln V_i, \quad (3)$$

$$y_i=\ln[-\ln(1-p_i)]. \quad (4)$$

Best-fit linear coefficient values a and b can be found by either an unweighted or frequency-of-occurrence-weighted least-squares process. The Weibull parameters c and k are related to the linear coefficients a and b by

$$c=\exp(-a/b), \quad (5)$$

$$k=b. \quad (6)$$

0021-8952/78/0350-0353$05.00

If the observed wind distribution is at the desired height, there is little reason to prefer the Weibull distribution fit over the observed distribution. However, if the wind distribution is desired at some height other than the anemometer level, the advantage of the use of the Weibull distribution is that c and k values c_a and k_a determined at anemometer height z_a can be adjusted to any desired height z by the relation (Justus and Mikhail, 1976)

$$c(z)=c_a(z/z_a)^n, \qquad (7)$$

$$k(z)=k_a[1-0.088\ln(z_a/10)]/[1-0.088\ln(z/10)], \qquad (8)$$

where z and z_a are in meters and the power law exponent n is given by

$$n=[0.37-0.088\ln c_a]/[1-0.088\ln(z_a/10)]. \qquad (9)$$

b. *Median and quartile wind speeds*—METHOD 2

If the complete observed wind speed distribution is not available but the median speed V_m (50% probable) and quartile speeds $V_{0.25}$ and $V_{0.75}$ [$p(V\leqslant V_{0.25})=0.25$, $p(V\leqslant V_{0.75})=0.75$] are available, then c and k can be computed by the relations

$$k=\ln[\ln(0.25)/\ln(0.75)]/\ln(V_{0.75}/V_{0.25})$$
$$=1.573/\ln(V_{0.75}/V_{0.25}), \qquad (10)$$

$$c=V_m/(\ln 2)^{1/k}. \qquad (11)$$

Ljungstrom (1976) has used a wind speed distribution formula

$$p(V\geqslant V_x)=2^{-(V_x/V_m)^\beta} \qquad (12)$$

which is equivalent to the Weibull with $\beta=k$.

c. *Mean wind speed and standard deviation*—METHOD 3

If only the mean wind speed $\bar{V}$ and standard deviation σ are available [where $\sigma^2=\langle(V-\bar{V})^2\rangle$, and the angle brackets denote an average], then c and k can be estimated from these statistics, since c and k are related to $\bar{V}$ and σ by

$$\bar{V}=c\Gamma(1+1/k), \qquad (13)$$

$$(\sigma/\bar{V})^2=[\Gamma(1+2/k)/\Gamma^2(1+1/k)]-1, \qquad (14)$$

where Γ is the usual gamma function and $\sigma/\bar{V}$ is the coefficient of variation. The c and k values can best be found by using the approximate relation for (14), i.e.,

$$k=(\sigma/\bar{V})^{-1.086} \qquad (15)$$

and the inverse of (13)

$$c=\bar{V}/\Gamma(1+1/k). \qquad (16)$$

TABLE 1. Observed mean wind speed, fastest mile and speed distribution (%) for Concord, N. H.

No.	Month	Year	$\bar{V}$ (mph)	V_{max} (mph)	Wind speed levels (mph) $\leqslant 5$	$\leqslant 10$	$\leqslant 15$	$\leqslant 20$	$\leqslant 25$	$\leqslant 30$	$\leqslant 35$	$\leqslant 40$
1	Oct	1971	5.7	25	44.8	87.5	96.4	99.2	100.0			
2	Nov	1971	7.6	30	35.4	74.2	87.1	98.3	99.6	100.0		
3	Dec	1971	8.1	39	35.9	68.1	85.9	95.2	99.2	100.0		
4	Jan	1972	7.8	44	40.3	69.8	87.1	94.8	99.2	100.0		
5	Feb	1972	9.9	34	22.4	57.3	81.5	92.2	98.3	99.1	99.6	100.0
6	Mar	1972	9.7	26	16.5	63.3	86.7	100.0				
7	Apr	1972	8.9	28	25.4	67.9	87.9	97.5	99.6	100.0		
8	Aug	1972	5.3	25	48.4	84.7	97.2	100.0				
9	Oct	1972	5.7	23	48.8	81.5	95.6	99.6	100.0			
10	Nov	1972	6.0	32	50.0	78.8	93.3	98.8	100.0			
11	Dec	1972	6.4	40	44.8	80.2	91.5	97.8	99.2	100.0		
12	Jan	1973	8.2	32	35.9	64.9	83.9	94.4	99.2	100.0		
13	Feb	1973	7.9	23	33.9	66.5	88.8	99.1	100.0			
14	Apr	1973	8.9	32	28.8	63.8	87.9	95.8	99.6	99.6	100.0	
15	Aug	1973	3.4	20	71.0	96.4	100.0					
16	Oct	1973	4.6	30	58.5	88.7	98.0	99.2	100.0			
17	Nov	1973	6.5	26	45.4	72.9	91.3	98.8	100.0			
18	Dec	1973	6.5	28	44.4	71.8	91.1	99.2	100.0			
19	Feb	1974	7.8	40	37.5	73.7	85.7	94.6	98.2	100.0		
20	Mar	1974	10.2	49	22.2	54.0	79.0	93.1	99.2	100.0		
21	Aug	1974	5.0	37	58.1	89.1	98.8	99.6	100.0			
22	Oct	1974	5.6	32	50.4	81.0	96.0	100.0				
23	Nov	1974	5.6	29	55.8	80.4	94.6	99.2	100.0			
24	Dec	1974	5.5	30	53.2	85.1	92.7	98.4	99.8	100.0		
25	Jan	1975	6.0	40	50.8	84.3	94.0	97.6	99.6	100.0		
26	Feb	1975	6.3	35	49.1	80.8	93.3	98.2	100.0			
27	Mar	1975	9.0	38	32.3	61.7	79.8	92.7	100.0			
28	Apr	1975	9.0	31	29.2	59.6	84.6	97.5	99.6	99.6	100.0	
29	Aug	1975	5.6	24	50.4	87.5	98.8	100.0				
30	Oct	1975	6.6	28	43.5	79.4	96.4	98.0	100.0			

TABLE 2. RMS errors of distribution fit (%) for Concord, N. H. data (Table 1).

Method description	Wind speed levels (mph) $\leqslant 5$	$\leqslant 10$	$\leqslant 15$	$\leqslant 20$	$\leqslant 25$	$\leqslant 30$	$\leqslant 35$	$\leqslant 40$
Weibull least squares	1.40	2.63	1.61	0.47	0.35	0.14	0.09	0.07
Square-root normal	11.57	5.20	4.29	1.83	0.54	0.33	0.14	
$\bar{V}$ and V_{max} Weibull	5.07	4.46	3.88	1.67	0.55	0.34	0.14	0.07
k vs $\bar{V}$ Weibull	3.88	3.44	3.00	1.75	0.47	0.18	0.11	

d. Mean wind speed and fastest mile—METHOD 4

Since data listings such as *Local Climatological Data* give monthly mean wind speed but not the standard deviation, Eqs. (15) and (16) cannot be used unless some estimate of the standard deviation or equivalent parameter is obtained. Since the monthly fastest mile V_{max} [average speed (mph) associated with most rapid 1 mi run of wind] is routinely published in *Local Climatological Data* it can be used as an estimator for the Weibull k value. The cumulative probability associated with V_{max} is

$$p(V \geqslant V_{max}) = 1/(24\, V_{max} d) = \exp[-(V_{max}/c)^k], \quad (17)$$

where d is the number of days per month. The left side of (17) comes from the definition V_{max} of the fastest mile, i.e.,

$$V_{max}(\text{mph}) = 1(\text{mi})/t(\text{h}) \quad \text{or} \quad t = 1(\text{mi})/V_{max}]$$

and the fact that the probability $p(V \geqslant V_{max})$ is $t/(24\, d)$. Inversion of (17) to solve for V_{max}, and division by Eq. (13) yields the relation

$$V_{max}/\bar{V} = [\ln(24\, V_{max} d)]^{1/k}/\Gamma(1+1/k), \quad (18)$$

which, for observed $\bar{V}$ and V_{max}, can be solved iteratively for the appropriate k value. With k thus determined, the Weibull c parameter can be found from (16). Since $\Gamma(1+1/k) \approx 0.9$ over the usual range of k values, approximate solutions to (18) can be found from

$$k = \ln[\ln(24\, V_{max} d)]/\ln(0.9 V_{max}/\bar{V}). \quad (19)$$

Widger (1977) recently proposed a similar use of fastest mile data, but he suggested that a "square-root-normal" distribution be used, i.e., one in which the square root of the wind speed has a Gaussian distribution.

e. Trend of k vs $\bar{V}$—METHOD 5

Earlier studies (Justus *et al.*, 1976b) have determined that a general trend exists between Weibull k values (or variance of the wind distribution) and the mean wind speed. Mathematically these results can be expressed for average, high (90 percentile), and low (10 percentile) variability (i.e., $\sigma/\bar{V}$) sites by

$$k = \begin{vmatrix} 1.05\bar{V}^{\frac{1}{2}} & \text{(low)} \\ 0.94\bar{V}^{\frac{1}{2}} & \text{(average)} \\ 0.83\bar{V}^{\frac{1}{2}} & \text{(high)} \end{vmatrix} \quad (20)$$

for $\bar{V}$ in meters per second, or

$$k = \begin{vmatrix} 0.70\bar{V}^{\frac{1}{2}} & \text{(low)} \\ 0.63\bar{V}^{\frac{1}{2}} & \text{(average)} \\ 0.49\bar{V}^{\frac{1}{2}} & \text{(high)} \end{vmatrix} \quad (21)$$

for $\bar{V}$ in miles per hour. In (20) or (21) $\sigma/\bar{V}$ corresponds to high k and vice versa. Thus, with some qualitative estimate of the degree of wind variability of the site, and only the mean wind speed, k can be estimated from (20) or (21) and c can be computed from (16).

Another method, potentially more accurate but somewhat more difficult to apply, is the maximum likelihood iteration technique of Takle and Brown (1977). These authors also introduce a third parameter F_0 to account for a finite number of absolute calm winds [i.e., $p'(V) = F_0\delta(V) + (1-F_0)p(V)$, where δ is the Dirac delta function, $p(V)$ the Weibull distribution, and $p'(V)$ the adjusted distribution with extra calms].

3. Comparison of the methods

Table 1 (from Widger, 1977) gives observed monthly speed distribution for 30 months from five years of data at Concord, N. H. Using the least-squares-fit method (method 1) the best Weibull distributions were found and rms errors are given in Table 2, which also gives the results for the square-root-normal distribution which Widger fit to the data on Table 1 using his mean speed and fastest mile technique. As Widger pointed out, the fit of his model is not especially good for the lowest speed interval. Table 2 also shows approximate Weibull distribution fits using the simple methodologies based on mean speed and fastest mile (method 4) and the trend of k versus mean speed (method 5). In application of method 5, it was observed (based on examination of a few selected months) that the variability for Concord fell between the average and 90 percentile levels. Based on these observations, the relation

$$k = 0.58\bar{V}^{\frac{1}{2}} \quad (V \text{ in miles per hour}) \quad (22)$$

was used, where the intermediate value 0.58 provided the best qualitative agreement with the selected sample month data. Comparison of the data from Table 2 shows that either of the simple-method Weibull estimates produces a lower rms error from the observed distribution than does the square-root normal.

The methods presented here offer options for estimating Weibull distribution parameters, based on which wind statistics are available and the rigor required in the data analysis. The simplest methods (4 and 5) provide reasonably accurate representation of the actual observed distribution.

Acknowledgments. Work reported here was sponsored in part under U. S. Energy Research and Development Administration Contract EY-76-S-06-2439. By acceptance of this article, the publisher and/or recipient acknowledges the U. S. Government's right to retain a nonexclusive, royalty-free license in and to any copyright covering this paper.

REFERENCES

Davenport, A. G., 1963: The relationship of wind structure to wind loading. *Prof. Conf. Wind Effects on Structures*, Nat. Phys. Lab., London, England, 19–82.

Hennessey, J. P., Jr., 1977: Some aspects of wind power statistics. *J. Appl. Meteor.*, **16**, 119–128.

Justus, C. G., and Amir Mikhail, 1976: Height variation of wind speed and wind distribution statistics. *Geophys. Res. Lett.*, **3**, 261–264.

——, W. R. Hargraves and Ali Yalcin, 1976a: Nationwide assessment of potential output from wind powered generators. *J. Appl. Meteor.*, **15**, 673–678.

——, —— and Amir Mikhail, 1976b: Reference wind speed distributions and height profiles for wind turbine design and performance evaluation applications, ERDA ORO/5108-76/4 [Available from NTIS].

Ljunstrom, O., 1976: Large scale wind energy conversion system (WECS) design and installation as affected by site wind energy characteristics, grouping arrangement and social acceptance. *Proc. Int. Symp. Wind Energy Systems*, British Hydro. Res. Assoc., Cambridge, England, A1.1–A1.23.

Takle, E. S., and J. M. Brown: Note on the use of Weibull statistics to characterize wind speed data. Journal paper No. J-8941, Iowa Agriculture and Home Economics Experiment Station, Ames Iowa, Project 2116.

Widger, W. K., Jr., 1977: Estimations of wind speed frequency distributions using only the monthly average and fastest mile data. *J. Appl. Meteor.*, **16**, 244–247.