CORRESPONDENCE

A Simple Estimator of the Shape Factor of the Two-Parameter Weibull Distribution

ROLAN D. CHRISTOFFERSON

Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado/NOAA, Boulder, CO 80309

DALE A. GILLETTE

Geophysical Monitoring for Climatic Change, Air Resources Laboratory, NOAA, Boulder, CO 80303

19 April 1986 and 25 October 1986

ABSTRACT

A method is given to estimate the shape factor (K) of the Weibull distribution directly from wind observations, without iteration, plotting, or sorting of data. The estimate is of comparable accuracy as that given by the maximum-likelihood estimate but is obtained more economically in computer time. Estimation of the scale factor (C) is done by maximum likelihood.

1. Introduction

Wind energy researchers, e.g., Corotis et al. (1978) and Hennessey (1977), have shown that the Weibull distribution and the special case of the Weibull distribution, the $\chi - 2$ (Rayleigh distribution), fit probability densities for wind data quite well. Probability densities are essential for computing expectations of wind-related quantities, e.g., wind energy and wind erosion. Therefore, the estimation of the parameters of the distributions when only unfitted wind data are available is of interest.

Of the two distributions, the $\chi - 2$ is simpler to use, since it can be calibrated purely from the mean of the velocity. Corotis et al. (1978) state, however, that Weibull-based distributions give somewhat closer fits than the $\chi - 2$ distributions. Since these better fits may translate into significantly better estimates for expectations of higher moments of wind speed, it is worthwhile to efficiently estimate the parameters of the Weibull distribution K (the shape factor) and C (the scale factor).

The form of the Weibull distribution, $P(U)$, is

$$P(U)dU = \left(\frac{K}{C}\right)\left(\frac{U}{C}\right)^{K-1} \exp\left[-\left(\frac{U}{C}\right)^{K}\right]dU. \qquad (1)$$

The parameters K and C must be estimated iteratively, graphically or by a linear least-squares fit to the data or by third moments (Takle and Brown, 1978; Justus et al., 1976) or by use of mean, standard deviation, and the "fastest mile" (Justus et al., 1978). Because the iterative maximum-likelihood method is somewhat tedious and expensive in computer time, and the other methods provide estimates considerably less accurate than those of the maximum-likelihood method, we sought a simple nongraphical method, other than linear least squares, that estimates K directly from the wind data and that gives results comparable with maximum likelihood. We limited our search for a simple method to estimate K since the maximum-likelihood method for estimating C is efficient.

2. Development of a statistic for evaluting K of the Weibull distribution

One method often used to estimate the parameters K and C of the Weibull distribution from wind data is the maximum-likelihood estimate given by Johnson and Kotz (1970). This method conveniently expresses the parameter C in terms of the parameter K and equally spaced wind observations as

$$C = \left(\frac{1}{N}\sum_{i=1}^{N} U_i{}^K\right)^{1/K} \qquad (2)$$

where

N the number of wind observations
U_i the observed wind speed for observation i.

In this method, the parameter K is evaluated by solving the equation

$$K = \left(\frac{\sum_{i=1}^{N} U_i^K \ln U_i}{\sum_{i=1}^{N} U_i^K} - \frac{1}{N}\sum_{i=1}^{N} \ln U_i\right)^{-1}. \qquad (3)$$

Because K appears on both sides of the equation, the equation must be solved iteratively, and to find a convergent value for K, several iterations are required. Therefore, we sought an alternative method for evaluating K without the need of iteration.

Our method utilizes the use of the statistic β:

$$\beta = \frac{1}{(N-1)}\left[\frac{1}{N}(\sum_{i=1}^{N} U_i^m)(\sum_{i=1}^{N} U_i^{-m}) - 1\right] \quad (4)$$

where m is an arbitrary positive number. The expectation of β is

$$E(\beta) = \frac{\pi m}{K}\csc\left(\frac{\pi m}{K}\right). \quad (5)$$

That is, the expectation of β is a function only of K (the desired Weibull parameter) and m (the arbitrary power of the wind speed used in the statistic). If we now specify that

$$\beta = \frac{\pi m}{[K]}\csc\left(\frac{\pi m}{[K]}\right) \quad (6)$$

where $[K]$ is the estimated value of the Weibull parameter K, then the $[K]$ can be obtained by substituting (5) in (6). By making the substitution

$$\theta = \frac{\pi m}{[K]}, \quad (7)$$

Eq. (6) becomes

$$\frac{\sin\theta}{\theta} = \frac{1}{\beta}. \quad (8)$$

The solution to Eq. (8) may be found iteratively for arbitrary m by using Newtonian iteration. We found that standard deviations of the estimates for K became smaller as the value for m approached zero. Therefore, we used the approximation

$$\frac{\sin\theta}{\theta} \approx 1 - \frac{\theta^2}{6} \quad (9)$$

so that

$$\theta = [6(1 - 1/\beta)]^{0.5}. \quad (10)$$

Thus,

$$[K] = \pi m[6(1 - 1/\beta)]^{-0.5}. \quad (11)$$

Now, if we let m approach zero and apply L'Hospital's rule twice,

$$[K] = \frac{\pi}{\sqrt{6}}\left[\frac{N(N-1)}{N(\sum_{i=1}^{N} \ln^2 U_i) - (\sum_{i=1}^{N} \ln U_i)^2}\right]^{0.5}. \quad (12)$$

Using (12), we solved for the estimated values of K for 1000 sets of size 100 each of Weibull-distributed random numbers having known values of C and K. Six values of C times six values of K were chosen to give a total of 36 000 sets of testing numbers. The same distributions were used by the maximum-likelihood method, and the resulting K estimates were compared. Table 1 gives the mean K and C values estimated by the maximum-likelihood method along with their standard deviations. Table 2 gives the mean and standard deviations of the estimated K and C values using our method. The two tables show that both methods give about the same accuracy with our estimator often being slightly more accurate. The maximum-likelihood method required about 20 to 40 iterations to converge to the estimate whereas the method using (12) required no iteration for a value of roughly comparable accuracy and precision.

3. Conclusion

An alternative method to maximum likelihood was developed to estimate the parameter K of the Weibull distribution without iteration. The estimate of K thus produced was of approximately comparable accuracy and precision to that produced by the maximum-likelihood method which required 20–40 iterations. The estimates produced by our alternative method were economical in computer time, since they did not require sorting of the data, or any iteration.

TABLE 1. Mean and standard deviations (in parentheses) of estimates of K (upper) and C (lower) by maximum likelihood from sample populations developed from random numbers having a Weibull distribution with parameters K and C shown on margins of table.

	C					
K	0.5	1.0	1.5	2.0	2.5	3.0
0.5	0.5158 (0.0669) 0.5053 (0.1056)	0.5122 (0.0640) 0.9952 (0.2002)	0.5155 (0.0659) 1.5092 (0.3096)	0.5130 (0.0634) 2.0186 (0.4463)	0.5138 (0.0674) 2.5353 (0.6013)	0.5170 (0.0668) 3.0028 (0.7252)
1.0	1.0301 (0.1313) 0.5057 (0.0639)	1.0299 (0.1312) 0.9973 (0.1016)	1.0256 (0.1343) 1.4896 (0.1670)	1.0321 (0.1329) 1.9919 (0.2604)	1.0239 (0.1319) 2.4956 (0.4183)	1.0200 (0.1351) 3.0173 (0.5400)
1.5	1.5412 (0.2088) 0.5039 (0.0585)	1.5400 (0.2031) 1.0022 (0.0673)	1.5414 (0.1410) 1.4888 (0.1236)	1.5427 (0.1923) 1.9923 (0.2159)	1.5438 (0.2051) 2.4901 (0.3445)	1.5382 (0.1939) 2.9959 (0.4805)
2.0	2.0453 (0.2632) 0.5031 (0.0506)	2.0466 (0.2615) 1.0011 (0.0505)	2.0541 (0.2668) 1.4913 (0.1094)	2.0648 (0.2745) 1.9874 (0.2123)	2.0667 (0.2628) 2.4813 (0.3324)	2.0532 (0.2612) 2.9896 (0.4644)
2.5	2.5652 (0.3351) 0.5055 (0.0488)	2.5645 (0.3304) 0.9987 (0.0399)	2.5703 (0.3164) 1.4989 (0.0952)	2.5694 (0.3126) 1.9899 (0.1887)	2.5574 (0.3346) 2.4967 (0.3170)	2.5834 (0.3332) 2.9705 (0.4379)
3.0	3.0897 (0.4119) 0.5056 (0.0484)	3.0649 (0.3975) 1.0001 (0.0335)	3.1070 (0.3799) 1.4856 (0.0890)	3.0156 (0.3943) 1.9932 (0.1871)	3.0831 (0.4036) 2.4948 (0.3200)	3.0811 (0.4042) 3.0001 (0.4548)

TABLE 2. Mean and standard deviations (in parentheses) of estimates of K (upper) and C (lower) by our method from sample populations developed from random numbers having a Weibull distribution with parameters K and C shown on margin of table.

	C					
K	0.5	1.0	1.5	2.0	2.5	3.0
0.5	0.5114 (0.0524) 0.5056 (0.1093)	0.5058 (0.0498) 1.0021 (0.2215)	0.5080 (0.0517) 1.5334 (0.3382)	0.5075 (0.0510) 2.0487 (0.4593)	0.5068 (0.0526) 2.5674 (0.5767)	0.5108 (0.0521) 3.0750 (0.6396)
1.0	1.0179 (0.1035) 0.5014 (0.0522)	1.0189 (0.1012) 1.0025 (0.1109)	1.0190 (0.1041) 1.5006 (0.1616)	1.0202 (0.1056) 2.0187 (0.2153)	1.0114 (0.1026) 2.4938 (0.2621)	1.0121 (0.1025) 3.0014 (0.3209)
1.5	1.5327 (0.1600) 0.5005 (0.0378)	1.5263 (0.1626) 1.0050 (0.0791)	1.5261 (0.1520) 1.4987 (0.1077)	1.5186 (0.1536) 2.0055 (0.1391)	1.5180 (0.1601) 2.5025 (0.1863)	1.5205 (0.1532) 3.0019 (0.2083)
2.0	2.0345 (0.2092) 0.5001 (0.0269)	2.0303 (0.1979) 1.0027 (0.0535)	2.0327 (0.2102) 1.4983 (0.0856)	2.0445 (0.2161) 2.0045 (0.1117)	2.0358 (0.2112) 2.5076 (0.1347)	2.0354 (0.2025) 2.9996 (0.1642)
2.5	2.5296 (0.2581) 0.5007 (0.0218)	2.5396 (0.2540) 1.0001 (0.0439)	2.5287 (0.2555) 1.5061 (0.0655)	2.5439 (0.2537) 2.0036 (0.0874)	2.5223 (0.2608) 2.4934 (0.1090)	2.5525 (0.2663) 3.0021 (0.1309)
3.0	3.0510 (0.3181) 0.4999 (0.0176)	3.0257 (0.3162) 1.0002 (0.0371)	3.0663 (0.3053) 1.4988 (0.0553)	3.0475 (0.3078) 2.0004 (0.0742)	3.0456 (0.3117) 2.5023 (0.0907)	3.0441 (0.3125) 3.0032 (0.1082)

Acknowledgments. This work has been supported by NOAA as part of the National Acid Precipitation Assessment Program.

REFERENCES

Corotis, R. B., A. B. Sigl and J. Klein, 1978: Probability models of wind velocity magnitude and persistence. *Sol. Energy,* **20,** 483–493.

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

Johnson, N. L., and S. Kotz, 1970: *Continuous Univariate Distributions, Vol. 2.* Houghton Mifflin, 306 pp.

Justus, C. G., W. R. Hargraves and A. Mikhail, 1976: Reference wind speed distributions and height profiles for wind turbine design and performance evaluation applications. ORO/5108-76/4 UC60, 52 pp. (Available from NTIS).

——, ——, —— and D. Graber, 1978: Methods for estimating wind speed frequency distributions. *J. Appl. Meteor.,* **17,** 350–353.

Takle, E. S., and J. M. Brown, 1978: Note on the use of Weibull statistics to characterize wind-speed data. *J. Appl. Meteor.,* **17,** 556–559.