The Characteristics of Wind Velocity that Favor the Fitting of a Weibull Distribution in Wind Speed Analysis

STANTON E. TULLER[1] AND ARTHUR C. BRETT[2]

University of Victoria, Victoria, British Columbia, Canada V8W2Y2

(Manuscript received 21 July 1982, in final form 4 June 1983)

ABSTRACT

The derivation of the Weibull distribution from the bivariate normal distribution provides theoretical justification for its use in wind speed analysis if four conditions are met. These conditions are that the orthogonal components of horizontal wind velocity transformed by raising them to the power $k/2$ are normally distributed, have zero means, have equal variances, and are uncorrelated. These four conditions specify a circular normal distribution for the transformed wind velocity components. Real world wind velocity patterns, however, are seldom, if ever, circular normal. Instead, the effects of topography and frontal systems produce a different wind speed distribution from each direction. This helps explain why the Weibull distribution gives only an approximate fit to the observed wind speed frequency distribution. The fit, at seven British Columbia coastal stations, was best at those stations with the most nearly circular wind velocity patterns and the lowest proportions of calms. It is suggested that these two characteristics, readily available in published wind normals, might provide a useful preliminary indication of the utility of attempting to fit a Weibull distribution.

1. Introduction

A number of studies in recent years have investigated the fitting of specific distributions to wind speed data for use in such practical applications as air pollution modelling, estimation of wind loads on buildings and wind power analysis. Several distributions have been tested, but the Weibull distribution has received the most recent attention (e.g., Justus *et al.*, 1976; Hennessey, 1977; Justus *et al.*, 1978; Stewart and Essenwanger, 1978; Takle and Brown, 1978; Van der Auwera *et al.*, 1980).

Most studies published in the widely available periodical literature report that the Weibull distribution gives an approximate but generally good fit to the observed wind speed distribution. The selection of the Weibull distribution is often attributed to its flexibility, which permits a good fit to the observed data to be obtained, and its ease of use with the necessity to estimate only two or three parameters. It can be shown, however, that if the observed winds meet certain conditions then the Weibull distribution is the mathematical function to fit to wind speed frequency data. Stauffer (1979), for example, presents a derivation of the Weibull distribution that provides theoretical justification for its use in fitting tree diameter and species abundance frequency data in forestry.

The present paper highlights the derivation of the Weibull distribution from the bivariate normal distribution. Although a Weibull wind speed frequency distribution may occur in other specific situations or arise by chance, it is felt that the situation presented here is the most useful to consider when dealing with natural winds. The purposes of the present paper are to: 1) highlight the precise relationship between the Weibull and the bivariate normal distribution, 2) provide some theoretical justification for the use of the Weibull distribution (and its variants) in wind speed analysis, 3) show that satisfactory fits need not be regarded as fortuitous events explainable only in terms of the inherent flexibility of the distribution function, 4) emphasize those conditions of wind velocity that favor the fitting of the Weibull distribution and 5) show why the Weibull distribution might provide a good fit in some cases but a poor fit in others.

It is hoped that the information presented here will provide a more complete appreciation of the Weibull distribution and allow its users to better evaluate its applications in wind speed analysis.

2. The Weibull distribution

The ordinary (two-parameter) Weibull distribution is a special case of the generalized three-parameter gamma distribution (Stacy, 1962). The Weibull probability density function may be expressed as follows:

$$P_w(V) = (k/c)(V/c)^{k-1} \exp[-(V/c)^k],$$
$$V \geq 0, \quad c > 0, \quad k > 0, \qquad (1)$$

[1] Department of Geography.
[2] Academic Systems.

where V is the wind speed (m s^{-1} in this study), c is called the scale parameter (in the same units as V), and k is a dimensionless quantity called the shape parameter.

Integration of (1) from 0 to some wind speed value $x > 0$ yields the ordinary Weibull cumulative distribution function

$$F_{w0}(x) = 1 - \exp[-(x/c)^k], \quad x \geqslant 0. \tag{2}$$

Regarding the Weibull distribution simply as a special case of the gamma distribution, however, does not seem to offer much practical insight when dealing with wind speed distributions.

A more meaningful approach might be to initially consider wind velocity. Wind velocity may be viewed as a bivariate distribution with the two orthogonal components (commonly north–south and east–west) as the two variables. It is well known that, under certain conditions, the Rayleigh distribution (a special case of the Weibull distribution with $k = 2.0$) may be derived from the bivariate normal distribution (see, for example, Justus *et al.*, 1978 and Lindgren, 1968). Although Stewart and Essenwanger (1979) observed that "under the customary postulation of a Gaussian distribution and a bivariate model for the wind speed components" the Weibull distribution is an approximation, the precise relationship between the Weibull and the bivariate normal distributions has not been reported in the meteorological periodical literature. Thus, it is useful at this time to show that, given certain conditions, the Weibull distribution may be derived from the bivariate normal distribution and to explore how closely the observed wind velocity approaches a bivariate normal distribution.

The Weibull distribution can be derived from the bivariate normal distribution under the following conditions:

a) If u and v are orthogonal components of the wind velocity (say, the east–west and north–south components, respectively), then the transformed variables $y = \mathrm{sgn}(u) \cdot |u|^{k/2}$ and $z = \mathrm{sgn}(v) \cdot |v|^{k/2}$, for some constant $k > 0$, are normally distributed, where |arg| denotes absolute value, and sgn(arg) = 1.0 if the argument is positive or zero, and -1.0 if the argument is negative;

b) y and z have equal variances $c^k/2$, with $c > 0$;

c) the means of y and z both are zero, and thus, the mean wind velocity is zero; and

d) y and z are independent.

Condition d) can be relaxed to the weaker condition that the correlation of y and z is zero.

Condition a) results in a bivariate normal probability density function in the transformed variables y and z. Application of conditions b), c), and d) yields a circular normal density function about the origin of the y, z axes. Transformation from the cartesian y, z axis system to polar coordinates, followed by transformation of the resulting radial distance variable from units of (m $s^{-1})^{k/2}$ back to m s^{-1} yields the Weibull probability density function for wind speed, (1). In the special case the constant $k = 2$, and thus the untransformed wind velocity components satisfy the four conditions, this derivation produces the Rayleigh distribution.

Thus, if the distribution of the transformed wind velocity is circular normal then the Weibull distribution should give a perfect fit to the wind speed distribution. If the original wind velocity distribution is circular normal the Rayleigh distribution should give a perfect fit. This suggests that the closer the wind velocity distribution is to circular normal the better able we are to say *a priori* that the Weibull distribution is a good distribution to fit in wind speed analysis. If the Weibull distribution is not found to give a good fit to observed data then the four conditions describing the circular normal distribution indicate the types of transformations that might be attempted to improve the fit.

As x approaches zero in (2), $F_{w0}(x)$ approaches zero; however, wind speeds of zero are recorded often. Takle and Brown (1978) proposed a hybrid Weibull distribution which included a non-zero probability of calms. Their hybrid Weibull cumulative distribution function appears as

$$F_{wH}(x) = F_0 + (1 - F_0)F_{w0}(x), \quad x \geqslant 0, \tag{3}$$

where $F_{w0}(x)$ is given by (2) and F_0 is estimated from the relative frequency of calms. The net effect of this procedure is to fit the Weibull distribution to the nonzero wind speeds and to adjust the resulting distribution, together with the moments computed therefrom for the occurrence of zero wind speeds. The hybrid, like the ordinary Weibull distribution, may be derived from the bivariate normal distribution if the four conditions noted above are met.

Stewart and Essenwanger (1978) discussed a three-parameter variant of the Weibull distribution appearing as

$$F_{w3}(x) = 1 - \exp\{-[(x - s)/c]^k\}, \quad x \geqslant s, \tag{4}$$

where the new parameter s (in the same units as x) is called the location parameter. All other symbols in (4) are as previously defined.

The location parameter (s) can be accommodated in a derivation of the three-parameter Weibull from a bivariate normal distribution by applying the transformation $u^* = \mathrm{sgn}(u)\max(0,|u| - s)$ to each of the orthogonal components of wind velocity (u and v). The operation $\max(0, |u| - s)$ returns the greater of the two values 0 and $|u| - s$. Thus, a positive location parameter acts as a threshold, censoring all velocities smaller than the s value and reducing the magnitude of the uncensored velocities. A negative location parameter has the effect of increasing velocity magnitudes by an amount equal to the value of s. Most of the location parameters for cases examined in the present study and reported in Stewart and Essenwanger (1978) are negative. If the transformed variables $y^* = (u^*)^{k/2}$ and $z^* = (v^*)^{k/2}$ satisfy the four conditions specified above for y and z, then $x - s$ where $(x - s)^k = (u^*)^k + (v^*)^k$ is Weibull distributed with distribution function (4).

Stewart and Essenwanger (1978) have found the three-parameter distribution to be generally superior to the ordinary Weibull for estimating certain probability thresholds and Van der Auwera *et al.* (1980) found another three-parameter variant of the Weibull distribution gave a better fit to observed wind speeds than did a number of other distribution functions.

3. Stations

Hourly wind speed and direction data from seven stations on the British Columbia coast were utilized in the present study (Fig. 1, Table 1). Twenty-three years of record (1953–1975) were used for six of the stations (Cape St. James, Comox Airport, Spring Island, Vancouver International Airport, Victoria Gonzales Heights and Victoria International Airport). Sixteen years of data (1960–1975) were available for Tofino Airport. The number of hourly observations ranged from 140 197 at Tofino Airport to 201 568 at Vancouver International Airport.

The wind pattern at all stations was influenced to some degree by topography. Local relief, land and sea distribution and local factors such as nearby forest all had an effect (Table 1). In addition, the pattern of winds associated with storm or frontal passages influenced the average wind speed–direction pattern. Strong southeasterly winds frequently occur in advance of a front at stations on the British Columbia coast. The wind shifts to relatively strong southwesterlies, westerlies, or northwesterlies after the front has passed. The directions of the strongest winds at each station result from a combination of the topographic and synoptic factors. The overall pattern is different at each station but the strongest mean annual winds are from the southeast at all stations but Vancouver International Airport where they come from the northwest.

4. Methods

The ordinary, hybrid and three-parameter Weibull distributions were fitted to the cumulative distribution function of observed wind speeds by the method of least squares. Functions (2), (3) and (4) can be transformed into an equation of the form:

$$\ln\{-\ln[1 - F_w(x)]\} = k \ln x - k \ln c, \qquad (5)$$

where $F_w(x)$ is the appropriate Weibull distribution. If we set

$$Y_i = \ln\{-\ln[1 - F(x_i)]\}, \qquad (6)$$

where $F(x_i)$ is now the observed distribution, and

$$t_i = \ln x_i, \qquad (7)$$

we can write a linear equation in which Y_i is the independent variable and t_i is the dependent variable:

$$Y_i = at_i + b + e_i, \qquad (8)$$

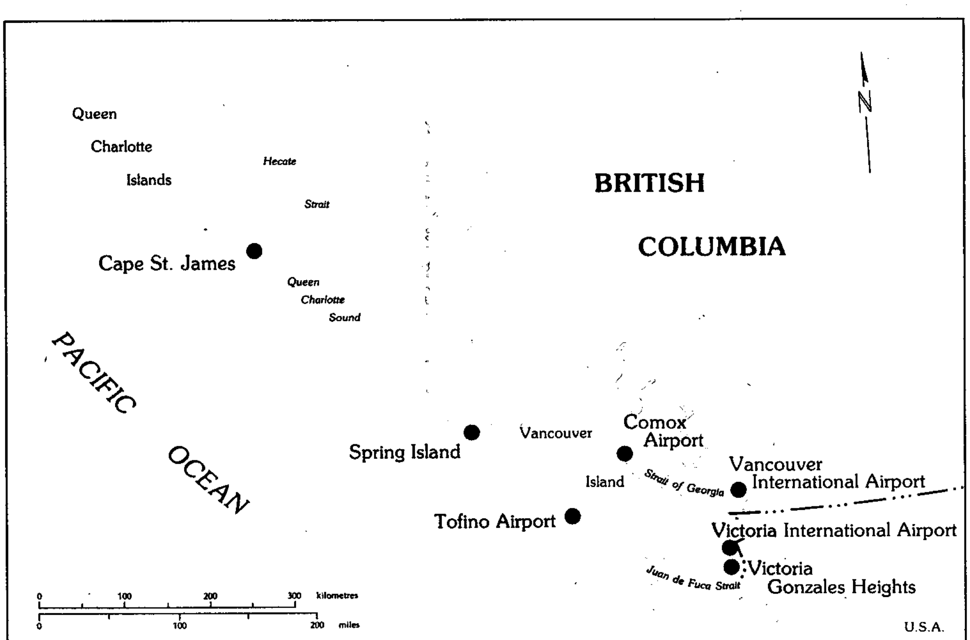

FIG. 1. Stations.

TABLE 1. Stations.

Station	Latitude	Longitude	Number of hourly observations	Topographic situation
Cape St. James	51°56′	131°01′	174 903	The station sits on a cone shaped hill about 100 m high on the southern extremity of the Queen Charlotte Islands. Higher hills lie to the NNW. Turbulence is created by local topography. SE or E gales are frequent in winter. May be influenced by the outflow of air from the continental valleys in winter.
Comox Airport	49°43′	124°54′	201 543	Located on the east coast of central Vancouver Island. Mountains trending NW–SE lie about 32 km to the west and low hills with the same trend are within 3 km to the west. The sea lies about 1.5 km to the east and 3 km to the SE. The station is most open to the sea from N through E to SE.
Spring Island	50°00′	127°25′	201 503	The station is on top of a low hill on a small island in the Kyuquot Sound area of Vancouver Island's west coast. Hills to 60 m lie close to the north and higher hills to 600 m trending NW–SE are 7 km to the NE. Exposure to the anemometer is nonetheless considered to be excellent in all directions.
Tofino Airport	49°05′	124°46′	140 197	Open to the sea from SE through W. Scattered islands and irregular relief provide partial obstruction from W through NE and continuous hills lie to the NE, E and SE. The local area is heavily wooded and the anemometer has been placed at about 19 m to improve the exposure.
Vancouver International Airport	49°11′	123°10′	201 568	Located on the flat Sea Island in the Fraser River delta. The sea lies to the west and mountains with peaks over 1000 m are about 20 km to the N. Low hills to about 90 m lie in the residential areas of Vancouver about 5 km N. Residential areas and flat farm land lie to the E and S. Strong NW winds channelled by the Strait of Georgia follow passage of a front.
Victoria Gonzales Heights	48°25′	123°19′	179 656	Located on top of a hill 70 m above Juan de Fuca Strait to the S. Winds from the SW through ESE reach the station after passage over water. Exposure is excellent in all directions. The hills of southern Vancouver Island divert flow into and out of Juan de Fuca Strait to SW and NE winds at Gonzales Heights. Well developed fronts are preceded by strong SE winds and followed by SW winds.
Victoria International Airport	48°39′	123°26′	201 565	A 300 m hill lies within 3 km to the SSW and lower hills (to 150 m) are located about 4 km N. Trees are found to the W and NW. The best exposure is in the NE–SE quadrant.

where e_i is the estimate of the error for the i-th observation and a and b are constants. a is the estimator of the Weibull shape parameter k, and c can be determined via

$$c = \exp(-b/a). \tag{9}$$

The values of a and b were determined by minimizing the sum of squared errors ($\Sigma\ e_i^2$) over all observed wind speeds included in the analysis.

Equation (7) becomes

$$t_i = \ln(x_i - s) \tag{10}$$

when fitting the three-parameter Weibull distribution. The linear equation (8) can still be used to estimate the parameters c and k but the problem of estimating s is nonlinear. Therefore, a nested linear–nonlinear procedure was used whereby (8) and (9) were used to estimate c and k for specific values of s. These computations were nested within a golden section algorithm which varied the value of s until the minimum sum of squared errors was achieved (Kowalik and Osborne, 1968). The golden section method is a conventional technique of handling a mixed linear–nonlinear problem and software for the procedure is available. It is a simpler method than that employed by Stewart and Essenwanger (1978) or the maximum likelihood technique which involves the simultaneous solution of three equations through iteration. It is a straightforward and economic solution and its use might be considered by those interested in practical application of the three-parameter Weibull distribution.

Stewart and Essenwanger (1978) reported that the method of moments responds better to the higher values in the distribution, and thus provided a better fit for estimating specific probability thresholds than did the least squares method, although the latter may provide a better overall fit across all values in the observed distribution. Generally the least squares method gave better fits to the observed data in the present study than did the method of moments. Only the results of the least squares method are presented here.

Many investigators report that the maximum likelihood method is preferred on the basis of accuracy, although it is more complicated and time consuming to use (e.g., Van der Auwera *et al.*, 1980; Takle and Brown, 1978; Stewart and Essenwanger, 1978).

The goodness of fit of the computed distributions to the observed wind speed distributions was measured by two statistics. The rms error was determined via

$$R_{\mathrm{MSE}} = [\sum_{i=1}^{N} e_i^2/(N - Z)]^{1/2}, \tag{11}$$

where e_i is the error in estimating the cumulative frequency of the i-th wind speed value, N is the total number of wind speed values and Z is either 2 or 3 depending on the number of degrees of freedom for the particular distribution being tested.

The maximum absolute difference, which forms the basis of the Kolmogorov–Smirnov goodness of fit test, is simply the greatest absolute difference between the computed and observed cumulative distribution functions over the entire range of wind speed values.

5. Results

a. Goodness of fit

The goodness of fit of the ordinary Weibull cumulative distribution function fitted by method of least squares was measured by both the rms error and the maximum absolute difference (Table 2). The Kolmogorov–Smirnov goodness of fit test, based on the maximum absolute difference, showed that with the extremely large sample sizes available in this study we must reject the null hypothesis that there is no difference between the ordinary Weibull and the actual wind speed distribution at at least the 0.01 level at all stations.

The 0.01 critical values used in this test were derived using Lilliefors' (1967) asymptotic formula for the one sample Kolmogorov–Smirnov statistic. Although Lilliefors' formula was derived for testing a normal distribution the values of the observed maximum absolute difference (Table 2) were one to two orders of magnitude greater than the critical values which ranged from 0.00275 at Tofino Airport to 0.00230 at Vancouver International Airport (see Table 7). Littell *et al.* (1979) have tabulated the critical values for testing a Weibull distribution for sample sizes between 10 and 40. The Littell *et al.* values are smaller than the Lilliefors values for sample sizes greater than ten at the 0.01 and 0.05 levels of significance. If this trend continues at larger sample sizes then the results using the Lilliefors asymptotic formula are likely to err on the conservative side when rejecting the null hypothesis. Therefore, we conclude that the fit of the ordinary Weibull distribution was not statistically significant.

The fit was poorest at those stations where the wind speed with direction frequency pattern was most asymmetric, that is, the wind velocity distribution was least circular (Tofino Airport, Comox Airport and Spring Island, Fig. 2); and at those with a high proportion of calms (Tofino Airport and Comox Airport, Table 3).

The problem of calms is treated by both the hybrid distribution where they are removed before fitting the Weibull distribution, and by the three-parameter distribution in which the shift of the coordinates effected by a negative location parameter or the censoring produced by a positive location parameter help accommodate some frequency of calms.

Some improvement in fit is seen with the three-parameter Weibull distribution, especially at those stations with a high proportion of calms (Table 2). The hybrid Weibull distribution produced some improvement in the maximum absolute difference at Comox Airport and Tofino Airport, the two stations with a high proportion of calms. In general, however, the hybrid distribution did not seem to offer much advantage over the ordinary Weibull distribution and in

TABLE 2. Measures of goodness of fit of the ordinary, three-parameter and hybrid Weibull distributions fitted by method of least squares to the cumulative distribution function of observed wind speeds.

Station	Maximum absolute difference			Root mean squared error		
	Ordinary	3-parameter	Hybrid	Ordinary	3-parameter	Hybrid
Cape St. James	0.034	0.024	0.029	0.010	0.009	0.011
Comox Airport	0.209	0.056	0.142	0.037	0.015	0.046
Spring Island	0.087	0.091	0.120	0.021	0.021	0.029
Tofino Airport	0.221	0.101	0.187	0.048	0.020	0.068
Vancouver International Airport	0.080	0.073	0.088	0.020	0.019	0.031
Victoria Gonzales Heights	0.030	0.035	0.088	0.008	0.009	0.027
Victoria International Airport	0.055	0.054	0.105	0.015	0.015	0.032

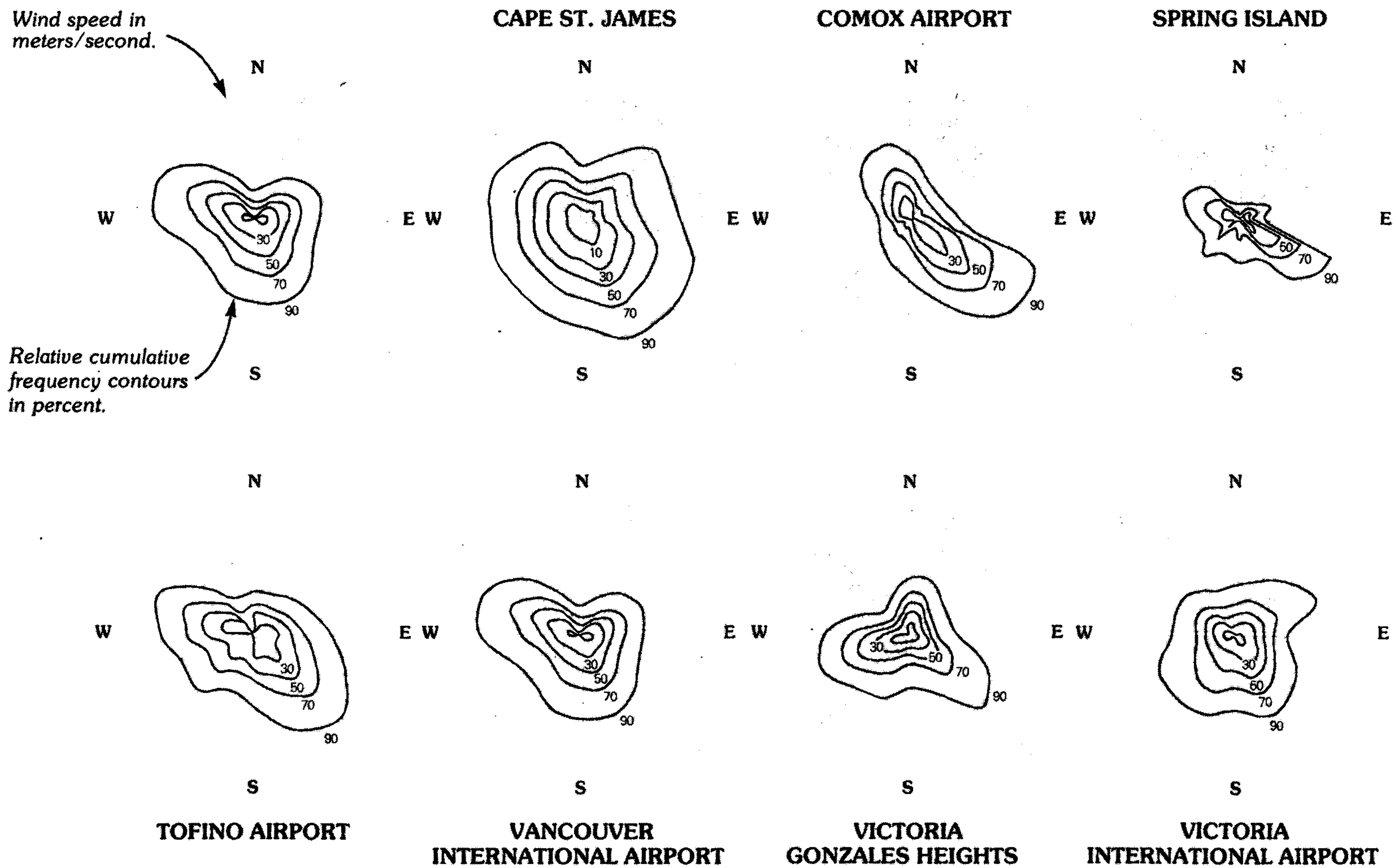

FIG. 2. Relative cumulative frequency contours of hourly wind speed.

TABLE 3. Proportion of calms.

Station	Proportion of calms
Cape St. James	0.008
Comox Airport	0.209
Spring Island	0.087
Tofino Airport	0.221
Vancouver International Airport	0.080
Victoria Gonzales Heights	0.027
Victoria International Airport	0.056

many cases actually produced a worse fit to the observed data (Table 2).

An exact fit of any mathematical function to real world data is not expected and even a statistically significant fit is extremely difficult to achieve with the sample sizes available in the present study. All that is usually required is that the applied distribution allow an adequate estimation of a particular parameter. The definition of "adequate" depends on the situation.

The Weibull distribution has received wide attention because it easily allows the estimation of the third raw moment of the wind speed ($\overline{V^3}$) for use in wind power analysis. The three-parameter Weibull distribution fitted by method of least squares was able to predict the third raw moment of the observed wind speed to within six percent at five of the seven British Columbia coastal stations (Table 4). Both of the exceptions (Comox Airport and Spring Island) had a very non-circular wind velocity pattern (Fig. 2). Comox Airport had a high proportion of calms (Table 3). The magnitude of the relative error at these two stations suggests that caution is still advised when using the Weibull distribution to estimate the third raw moment of the wind speed.

The goodness of fit of the Weibull distribution at the seven stations examined here confirms what has been found in most other studies. The Weibull distribution gives an approximate but often usable fit to the observed wind speed distribution.

TABLE 4. Comparison of the third raw moment of the wind speed ($\overline{V^3}$) with that estimated by the three-parameter Weibull distribution fitted by method of least squares.

Station	$\overline{V^3}$ observed ($m^3 s^{-3}$)	$\overline{V^3}$ estimated ($m^3 s^{-3}$)	Difference ($m^3 s^{-3}$)	Difference (%)
Cape St. James	1,874.7	1,936.2	+61.5	+3.3
Comox Airport	168.4	149.3	−19.1	−11.4
Spring Island	424.3	343.4	−80.9	−19.1
Tofino Airport	153.0	146.8	−6.2	−4.1
Vancouver International Airport	111.9	114.3	+2.4	+2.2
Victoria Gonzales Heights	321.4	303.7	−17.7	−5.5
Victoria International Airport	93.9	90.3	−3.6	−3.8

b. *Testing the four conditions*

According to classical logic, if a consequent proposition is false; then the antecedent proposition(s) must be false. Although some of the fits of the Weibull distribution to the observed wind speeds at the seven British Columbia coastal stations might be considered good or adequate for many purposes they were, nonetheless, not statistically significant and in some cases were poor. Therefore, we must conclude that the transformed wind velocity distribution was not circular normal at any of the stations and one or more of the four antecedent conditions must have been violated.

Although the circular normal distribution has long been used to describe wind distributions (e.g., Brooks *et al.,* 1946), it has also been widely appreciated for some time that many wind velocity distributions are not strictly circular normal. Brooks *et al.* (1946) stated that this distribution would not be expected in samples that include more than one prevailing wind system or cover the whole year. Brooks and Carruthers (1953) also noted the asymmetrical patterns that can occur in the wind speed–direction frequency distributions in samples that cover contrasting seasons, or that are from locations between two prevailing wind systems. They also stated that, "Near the ground, distributions of wind frequencies may be very irregular for they are generally much influenced by local topography; it is doubtful whether any simple empirical distribution can often be found that is truly representative." Crutcher (1957) noted that many seasonal upper air wind distributions were more elliptical than circular and commented on situations that may produce non-circular distributions including the boundary between continental and marine areas and surface winds. Essenwanger (1976) also gives examples of the non-circularity of upper level wind velocity distributions and discusses the application of bivariate elliptical normal and bivariate logistic distributions.

With the increasing use of the Weibull distribution in wind speed analysis it is instructive to re-emphasize this point by presenting further graphical examples of the nature of the wind velocity pattern at stations in topographically diverse areas and to examine in greater detail the validity of the four conditions necessary for a bivariate circular normal distribution of the transformed wind velocity. This should help researchers and practitioners to better evaluate the potential benefits of using the Weibull distribution in a particular situation or select possible methods of attempting an improvement in fit through various data transformations.

Local topography and the effects of frontal passages created a wind velocity pattern at all seven British Columbia coastal stations that did not satisfy the four conditions inherent in the circular normal distribution. The transformed orthogonal components of wind velocity were not normally distributed [condition a)]; they had neither b) equal variances, nor c) zero means; and d) they were not uncorrelated.

The effects of topography and the different wind speeds and directions associated with frontal systems mean that any station will receive stronger winds from some directions than it will from others. Although exponentiation by $k/2$ has the effect of reducing the asymmetry of the distribution when k is less than two, it is not able to transform a non-circular into a circular distribution. Therefore, the transformed mean wind velocity will seldom, if ever, be zero; the variances of the transformed orthogonal components will not be equal; and there will not be zero correlation between the transformed orthogonal components. The wind patterns found at the seven British Columbia coastal stations can be illustrated by the relative cumulative wind speed frequency contours (rcf contours) (Fig. 2). These contours indicate the proportion of wind observations from a particular direction that are below a given speed. They easily show the unequal distribution of wind speed with direction observed at these stations. The weighted central point of the rcf contours is the mean wind velocity. At none of the seven British Columbia coastal stations was this point at zero, the origin in the polar coordinate system (Fig. 2). The mean wind velocity and the means of each transformed orthogonal component were not zero at any of the stations (Table 5).

The effects of topography and frontal passages elongated the rcf contours along a preferred axis (northwest–southeast at most stations) and shifted the mean wind velocity to the southeast or southwest. The pattern of the rcf contours at Comox Airport, Tofino Airport and Spring Island corresponded to the trend of nearby hills and mountains and is repeated in the pattern of wind direction frequency.

The effect of strong southeasterlies preceding a front is seen at all stations. It was the dominant factor in the southeasterly lobes of the rcf contours at Cape St. James and Victoria Gonzales Heights where topography would not be a major factor and southeasterly winds have a very low annual frequency.

The relatively high speed of northwesterly winds at Vancouver International Airport is a result of both synoptic and topographic factors. Strong northwesterlies frequently blow behind a cold front aligned in a northeast–southwest direction, especially if it is followed by a strong high pressure ridge (Kendrew and Kerr, 1955). The channeling of the wind by the Strait of Georgia helps to funnel the northwesterly and increase its mean velocity.

Although the pattern of the rcf contours will be different at each observation station, it is probably safe to say that no location is immune to the influence of local topography and frontal passage. The topographic effects are doubtless greater at coastal stations than at those on the prairies or plains, but no station is without any topographic effects. Therefore, a mean of zero for each of the transformed orthogonal velocity components and for the overall wind velocity should be the exception rather than the rule.

The noncircular nature of the rcf contours indicates that the variances of the transformed orthogonal components of wind velocity [the basis of condition b)] were not equal (Table 5).

If the rcf contours were circular, the condition of zero correlation between the transformed orthogonal wind velocity components [condition d)], would also be met. This condition was not met at any of the seven British Columbia coastal stations with the highly asymmetrical pattern in their rcf contours. The correlation coefficients between the east–west and north–south components were significantly different from zero with the extremely large sample sizes available for this study (Table 6).

The fourth condition, normality of the transformed orthogonal components was tested by use of the Kolmogorov–Smirnov goodness of fit test. The maximum absolute differences between the east–west and north–south components of the transformed wind velocity and a normal distribution were computed. k used in the transformation was determined by a least squares fit of an ordinary Weibull distribution. The critical

TABLE 5. Means and variances of transformed east–west and north–south components of wind velocity.

Station	Mean $[(\mathrm{m\ s^{-1}})^{k/2}]$		Variance $[(\mathrm{m^2\ s^{-2}})^{k/2}]$	
	East–west	North–south	East–west	North–south
Cape St. James	1.36	0.42	20.25	24.80
Comox Airport	−0.13	0.13	1.88	2.07
Spring Island	−0.49	0.35	2.69	2.04
Tofino Airport	−0.12	0.26	2.82	1.85
Vancouver International Airport	−0.40	0.19	3.69	1.42
Victoria Gonzales Heights	0.74	0.37	6.45	4.67
Victoria International Airport	0.19	0.38	3.57	2.02

TABLE 6. Correlation of transformed east–west and north–south components of wind velocity.

Station	Correlation coefficient
Cape St. James	−0.32
Comox Airport	−0.64
Spring Island	−0.50
Tofino Airport	−0.26
Vancouver International Airport	−0.15
Victoria Gonzales Heights	+0.33
Victoria International Airport	−0.22

values for testing the null hypothesis that the distributions of the orthogonal components were normal were computed using the asymptotic formulas of Lilliefors (1967). In all cases the maximum absolute differences exceeded the critical values at both the 0.05 and 0.01 levels of significance (Table 7). Therefore, we must reject the null hypothesis that there is no difference between the transformed and normal distributions with the very large sample sizes used in this study.

None of the conditions necessary for the derivation of the Weibull distribution from the bivariate normal distribution were met by the winds observed at the stations examined here. The net result of the asymmetrical wind velocity pattern is a somewhat irregular wind speed frequency distribution. The associated cumulative frequency distribution is not a smooth, regular curve that can be perfectly reproduced by a single mathematical function with a reasonable number of free parameters. There is no *a priori* reason to believe that the Weibull distribution will give a statistically significant fit to the observed data and it or any other distribution can only be an approximation of the observed wind speed distribution. A good fit could be achieved by chance, however, or a fit may be deemed quite usable even when a very large sample size means it is not significant. Much of the error occurs at the low wind speeds and at zero even with the hybrid and three-parameter distributions. Therefore, it is suggested that if a station does have a nearly circular pattern of rcf contours we would expect the Weibull (or Rayleigh) distribution to give a good fit on the basis of its derivation from the bivariate normal distribution. If such a pattern is not found, however, caution and due regard for the acceptable degree of error is advised.

It is instructive to look further at the four conditions since many of these can be overcome by transformations of the data. The choice of east–west and north–south axes is arbitrary and the requirement of zero correlation may be overcome by rotation of the cartesian coordinate system. The rotation necessary to achieve zero correlation at the seven stations examined here is presented in Table 8. It is apparent from the rcf contours (Fig. 2), however, that wind speed is not independent of direction and hence the orthogonal components of wind velocity are not independent although it is possible to choose a cartesian coordinate system in which the orthogonal wind velocity components are uncorrelated. Selection of an appropriate coordinate system will not overcome the lack of zero means, equal variances and normality of the orthogonal wind velocity components (Tables 8 and 9).

Zelen and Severo (1966) present a transformation to achieve zero means, unit variances and zero correlation of the orthogonal components. The problem of non-normality remains, however, (Table 9) as does the recovery of the actual wind speed distribution from the highly transformed wind velocity distribution. Although this study has emphasized the concept of non-circularity because it is easy to present visually, it appears as though non-normality of the orthogonal wind velocity components presents the more difficult problem.

Essenwanger (1959) reported that the wind speed distribution can be made approximately normal via a square root or cube root transformation. Raising wind velocity to the power $k/2$ has the effect of a square

TABLE 7. Kolmogorov–Smirnov maximum absolute differences for normality tests of the distributions of the transformed orthogonal components of wind velocity.

	Lilliefors asymptotic critical values* Significance level		Maximum absolute difference	
Station	0.01	0.05	East–west	North–south
Cape St. James	0.00246	0.00212	0.0690	0.1198
Comox Airport	0.00230	0.00197	0.1267	0.1860
Spring Island	0.00230	0.00197	0.1127	0.1981
Tofino Airport	0.00275	0.00237	0.1916	0.2449
Vancouver International Airport	0.00230	0.00197	0.1127	0.2149
Victoria Gonzales Heights	0.00243	0.00209	0.1341	0.0788
Victoria International Airport	0.00230	0.00197	0.0909	0.1786

* The 0.01 and 0.05 critical values were calculated using the Lilliefors (1967) asymptotic formulas $1.031/\sqrt{n}$ and $0.886/\sqrt{n}$, respectively, where n is the sample size.

TABLE 8. Means and variances of transformed cartesian components of wind velocity estimated following rotation of axis system to achieve zero correlation of the components.

Station	Rotation angle, θ (deg)*	Mean $[(\mathrm{m\ s^{-1}})^{k/2}]$ E/W + θ	Mean $[(\mathrm{m\ s^{-1}})^{k/2}]$ N/S + θ	Variance $[(\mathrm{m^2\ s^{-2}})^{k/2}]$ E/W + θ	Variance $[(\mathrm{m^2\ s^{-2}})^{k/2}]$ N/S + θ
Cape St. James	36.9	1.49	−0.57	19.94	40.68
Comox Airport	43.5	0.00	0.37	3.10	12.09
Spring Island	−39.8	−0.83	−0.07	6.86	2.16
Tofino Airport	−26.0	−0.43	0.31	11.16	6.25
Vancouver International Airport	−12.4	−0.47	0.09	4.41	1.86
Victoria Gonzales Heights	34.6	0.90	−0.03	9.11	3.93
Victoria International Airport	−21.3	0.00	0.44	4.36	2.20

* Positive sign indicates counterclockwise rotation when viewing coordinate system from above.

root transformation when $k = 1.0$ and a cube root transformation when $k = 0.67$. Although k values reported in the literature vary widely, a number of cases have values relatively close to 1.0.

6. Summary and conclusions

The derivation of the Weibull distribution from a bivariate normal distribution establishes that the Weibull is a logical distribution to use to represent the wind speed distribution if four conditions are met. These four conditions are: a) the orthogonal components of the wind velocity transformed by raising to the power $k/2$ are normally distributed; the transformed orthogonal components have b) equal variances and c) zero means; and d) they are uncorrelated.

It is unlikely, however, that these conditions will be fully satisfied by the actual winds observed in nature. The effects of local topography and the pattern of winds that accompany storm systems usually produce an irregular and asymmetric pattern of wind speed with direction. This is one reason why the Weibull distribution provides only an approximate representation of the actual wind speed distribution. The closeness of the fit will vary from station to station. It is felt that, with the increasing use of the Weibull distribution in wind speed analysis, greater publicity for these conditions that favor a good fit is useful at this time. This will help workers to better plan and interpret the results of any project where the Weibull distribution is employed.

The results of the present study reconfirm the fact that wind velocity distributions are not circular normal. The results, however, indicate that the fit of the Weibull distribution was best at those stations with the most nearly circular pattern of relative cumulative wind speed frequency contours and a low frequency of calms. It is suggested that a knowledge of the average wind speed from each direction can give an approximate idea of how circular the pattern of the rcf contours will be at a particular station. This, along with the percentage of calms, should provide a preliminary, qualitative indication of the potential of the Weibull distribution to provide a useful fit to the actual wind speed distribution at a particular location based on widely published, readily available wind summary data. The plotting of the cumulative frequency distribution on Weibull probability paper might be the next step to specifically test the utility of the Weibull distribution and determine the parameters (Berrettoni, 1964; Essenwanger, 1976; Mage, 1979; Takle and Brown, 1978).

Other patterns of the wind velocity distribution such

TABLE 9. Kolmogorov–Smirnov maximum absolute differences for normality tests of the distributions of orthogonal components transformed to achieve zero correlation, zero means and unit variances.

Station	Lilliefors asymptotic critical values Significance level 0.01	Lilliefors asymptotic critical values Significance level 0.05	Maximum absolute difference Component 1	Maximum absolute difference Component 2
Cape St. James	0.00246	0.00212	0.1032	0.0660
Comox Airport	0.00230	0.00197	0.1228	0.1791
Spring Island	0.00230	0.00197	0.1393	0.1409
Tofino Airport	0.00275	0.00237	0.1307	0.1022
Vancouver International Airport	0.00230	0.00197	0.0665	0.0853
Victoria Gonzales Heights	0.00243	0.00209	0.0700	0.0902
Victoria International Airport	0.00230	0.00197	0.0888	0.0821

as four-lobe symmetry of the rcf contours could also form the basis for the derivation of the Weibull distribution. It is felt that the circular distribution is the most likely to be approached by winds observed in nature and the easiest to evaluate in practical application, however. In any event the rcf contours at the seven British Columbia coastal stations did not display any pattern that would suggest a perfect fit of the Weibull or Rayleigh distributions.

The goodness of fit of the Weibull distribution was not statistically significant at any of the seven coastal stations analyzed here. The three-parameter distribution, however, was able to estimate the third raw moment of the wind speed to within about six percent at five of the stations.

The patterns of the rcf contours, the values of the Weibull parameters and the goodness of fit varied from station to station. It is unlikely, therefore, that Weibull parameters determined for one location will be adequate for use elsewhere in a topographically diverse area. The patterns are different enough that no single distribution could be expected to give good results at all stations.

The patterns of the rcf contours were so asymmetrical that it would be very difficult to devise a function with a manageable number of free parameters that would give a significant fit to the observed data. If one were found it would be valid for only that one location.

This paper has highlighted some of the theoretical reasons why the Weibull distribution might be the distribution to fit to a wind speed distribution if certain assumptions as to the nature of the wind velocity pattern from which the wind speed distribution is derived are fulfilled. The use of the Weibull or other distributions to approximate the distribution of actual wind speeds at a particular location can have a number of practical benefits. Their application, however, should be based on firm foundation and must be tempered with caution. It is hoped that the present paper will help contribute to this end.

Acknowledgments. The Atmospheric Environment Service wind data used in this study were supplied by the Climatology Unit of the British Columbia Ministry of Environment. The analysis was aided by computing and programming grants from the University of Victoria.

REFERENCES

Berrettoni, J. N., 1964: Practical application of the Weibull distribution. *Ind. Qual. Control,* **21,** 71–79.

Brooks, C. E. P., and N. Carruthers, 1953: *Handbook of Statistical Methods in Meteorology.* Her Majesty's Stationary Office, 412 pp.

——, C. S. Durst and N. Carruthers, 1946: Upper winds over the world. Part I. The frequency distribution of winds at a point in the free air. *Quart. J. Roy. Meteor. Soc.,* **72,** 55–73.

Crutcher, H. L., 1957: On the standard vector-deviation wind rose. *J. Meteor.,* **14,** 28–33.

Essenwanger, O. M., 1959: Probleme der Windstatistik. *Meteor. Rundsch.,* **12,** 37–47.

——, 1976: *Applied Statistics in Atmospheric Science. Part A. Frequencies and Curve Fitting.* Elsevier, 424 pp.

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

Justus, C. G., W. R. Hargraves and A. Yalcin, 1976: Nationwide assessment of potential output from wind-powered generators. *J. Appl. Meteor.,* **15,** 673–678.

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

Kendrew, W. G., and D. Kerr, 1955: *Climate of British Columbia and the Yukon Territory.* Queen's Printer, 222 pp.

Kowalik, J., and M. R. Osborne, 1968: *Methods for Unconstrained Optimization Problems.* Elsevier, 148 pp.

Lilliefors, H. W., 1967: On the Kolmogorov–Smirnov test for normality with mean and variance unknown. *J. Am. Stat. Assoc.,* **62,** 399–402.

Lindgren, B. W., 1968: *Statistical Theory.* Macmillan, 521 pp.

Littell, R. C., J. T. McClave and W. W. Offen, 1979: Goodness-of-fit tests for the two-parameter Weibull distribution. *Commun. Stat.-Simul. Comput.,* **B8,** 257–269.

Mage, D. T., 1979: Comments on "Frequency distribution of wind speed near the surface". *J. Appl. Meteor.,* **18,** 1089–1091.

Stacy, E. W., 1962: A generalization of the gamma distribution. *Ann. Math. Stat.,* **33,** 1187–1192.

Stauffer, H. B., 1979: A derivation for the Weibull distribution. *J. Theor. Biol.,* **81,** 55–63.

Stewart, D. A., and O. M. Essenwanger, 1978: Frequency distribution of wind speed near the surface. *J. Appl. Meteor.,* **17,** 1633–1642.

——, and ——, 1979: Reply (to Mage). *J. Appl. Meteor.,* **18,** 1091–1092.

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.

Van der Auwera, L., F. de Meyer and L. M. Malet, 1980: The use of the Weibull three-parameter model for estimating mean wind power densities. *J. Appl. Meteor.,* **19,** 819–825.

Zelen, M., and N. C. Severo, 1966: Probability functions. *Handbook of Mathematical Functions,* M. Abramowitz and I. A. Stegun, Eds., U.S. Dept. of Commerce, 925–996.