NOTES

Note on the Use of the Inverse Gaussian Distribution for Wind Energy Applications

W. E. BARDSLEY[1]

Physics Department, University of Otago, New Zealand

26 September 1979 and 4 May 1980

ABSTRACT

The inverse Gaussian distribution is suggested as an alternative to the three-parameter Weibull distribution for the description of wind speed data with low frequencies of low speeds. A comparison of the two distributions indicates a region of strong similarity, corresponding reasonably well to three-parameter Weibull distributions which have been fitted to wind data.

Maximum likelihood estimation of the inverse Gaussian parameters is much simpler than the iterative technique required for the three-parameter Weibull distribution. In addition, the inverse Gaussian distribution features the mean wind speed as a parameter, a desirable property for wind energy investigations. A summation-reproductive property of the distribution permits estimation of the mean wind energy flux from a sequence of speed averages.

1. Introduction

The Weibull distribution has been widely employed to describe frequency distributions of wind speeds for wind energy evaluation. Hennessey (1977) and others have applied the two-parameter form of the distribution, but Stewart and Essenwanger (1978) found that the three-parameter form produced an improvement of fit. As noted by Stewart and Essenwanger (1978), the addition of the third parameter introduces estimation difficulties, the maximum likelihood procedure being particularly awkward. Furthermore, a positive value of the Weibull location parameter leads to the unrealistic condition of zero probability of wind speeds less than the parameter value.

The purpose of this note is to suggest the use of the inverse Gaussian distribution as a possible alternative model in situations where the data appear to indicate the use of the three-parameter Weibull distribution with a positive location parameter. Such data could arise from potential wind energy sites possessing low probabilities of low wind speeds. Alternatively, the speeds may have been obtained as averages over long time intervals, effectively eliminating the low speeds. When applicable, the inverse Gaussian distribution permits a much greater simplicity of parameter estimation and possesses some other desirable properties with respect to wind energy evaluation.

[1] Present affiliation: Earth Sciences Department, University of Waikato, Hamilton, New Zealand.

0021-8952/80/091126-05$05.25

2. Distribution comparison

The probability density function of the inverse Gaussian distribution can be written

$$f_X(x) = [\mu\phi/(2\pi x^3)]^{1/2} \exp(-\tfrac{1}{2}\phi x \mu^{-1} + \phi - \tfrac{1}{2}\mu\phi x^{-1}), \quad x > 0, \quad \mu > 0, \quad \phi > 0, \quad (1)$$

where μ is the distribution mean and ϕ a shape parameter. As ϕ increases, the distribution tends toward the normal. A description of moments and other properties is given by Tweedie (1957).

The inverse Gaussian distributions are a positively skewed unimodal family, rather similar in appearance to the unimodal forms of the two-parameter Weibull distribution. However, unlike the latter distribution, the main body of the inverse Gaussian is capable of considerable "displacement" from zero, while at the same time maintaining a significant amount of skewness (Fig. 1). The different behavior of the two distributions is illustrated in Fig. 2, showing a plot of skewness against T, a dimensionless ratio obtained by dividing the 0.0001 quantile by the mean of the distribution concerned. The skewness values in Fig. 2 represent the usual moment measure $\mu_3/(\mu_2)^{3/2}$, where μ_i is the ith central moment.

It is apparent from Fig. 1 that for larger ϕ values, the inverse Gaussian distribution exhibits a very abrupt increase in probability density, corresponding approximately to the 0.0001 quantile. This feature, together with the degree of skewness, suggests that the three-parameter Weibull distribution could be mimicked by inverse Gaussian distributions with the 0.0001 quantile corresponding to the Weibull location parameter. Of course, the similarity would apply only to a restricted set of three-parameter Weibull distributions, since the shape of the inverse Gaussian distribution is correlated to the displacement of the distribution from zero.

In order to measure the degree of similarity between the inverse Gaussian distribution and the three-parameter Weibull, a comparison was made of their respective distribution functions. Following Schuster (1968), the inverse Gaussian distribution function can be written

$$F_X(x) = \phi[(\phi\mu/x)^{1/2}(x\mu^{-1} - 1)] + \exp(2\phi)\phi[-(\phi\mu/x)^{1/2}(1 + x\mu^{-1})], \quad (2)$$

where ϕ represents the standard normal integral. Using the symbolism of Stewart and Essenwanger (1978), the distribution function of the three-parameter Weibull distribution is given by

$$G_W(w) = 1 - \exp\{-[(w - \gamma)/\theta]^\beta\}, \quad w \geq \gamma, \quad (3)$$

where γ, θ and β are location, scale and shape parameters, respectively.

The measure of similarity is given by M, defined as

$$M = \max|F_X(x) - G_W(x)| \quad (4)$$

for $F_X(\gamma) = 0.0001$ and μ set equal to the three-parameter Weibull mean. An M value of 0 indicates perfect coincidence of the distributions, while an M of 1 indicates no overlap. M is independent of the X, W units of measurement, so all distribution comparisons can be carried out with $\mu = 1$. It will be noted that a change in measurement scale in the three-parameter Weibull distribution implies changes in both γ and θ, resulting in an unaltered value of the ratio

$$R = \gamma/[\gamma + \theta\Gamma(1 + \beta^{-1})], \quad (5)$$

where the term in the square brackets in (5) represents the Weibull mean and Γ denotes the gamma function.

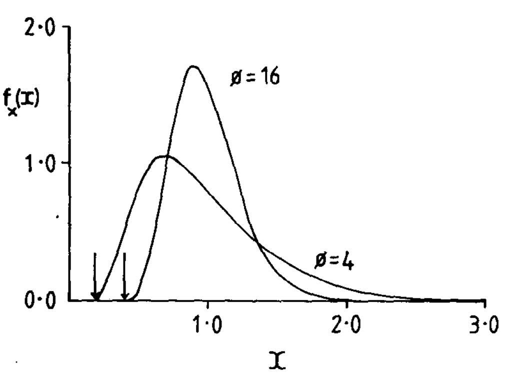

FIG. 1. Probability density functions of two inverse Gaussian distributions, standardized at $\mu = 1$. Arrows indicate the location of the 0.0001 quantiles.

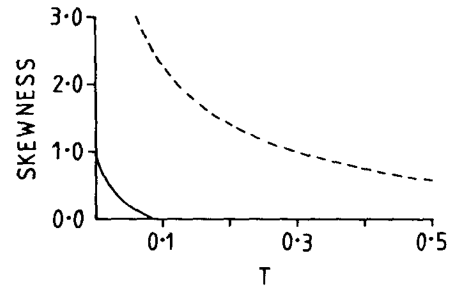

FIG. 2. Comparison of skewness of the inverse Gaussian distribution (broken) and the two-parameter Weibull distribution (solid) plotted as a function of the standardized 0.0001 quantile. See text for symbol definitions.

It is apparent from (2) that for given μ, any specified quantile value uniquely determines the inverse Gaussian-shape parameter ϕ. With ϕ obtained, the value of M can be found by evaluating the distribution functions at closely spaced values of x. Since μ and the Weibull mean are held constant at unity, each γ value will generate a particular inverse Gaussian distribution, regardless of the value of β. It follows that similarity of the three-parameter Weibull and the inverse Gaussian distribution can only be expected for a restricted range of β values for a given γ. Because of the scale-invariant nature of the comparison, the same M values will result for constant values of β and R.

Contours of M are shown in Fig. 3 expressed as a function of β and R. As might be expected, the region of greatest similarity occurs in the vicinity of equal variances. An approximate correspondence between the two distributions can be expected in the region bounded by the 0.1 contour, and the similarity is quite marked within the 0.05 contour (Fig. 4).

If the inverse Gaussian distribution is to be a practical substitute for the three-parameter Weibull, it is desirable that the regions of similarity should coincide with observed distributions of wind speed data. In order for this to be evaluated, points were superimposed on Fig. 3, using the positive-γ Weibull parameter estimates listed in Table 6 of Stewart and Essenwanger (1978). Points with R values ≤ 0.05 were excluded since the two-parameter Weibull distribution should be applicable in such cases. It is apparent from Fig. 3 that although a considerable degree of scatter is evident, there is still a sufficient number of points within the 0.1 contour to indicate the potential of the inverse Gaussian distribution for describing wind speed data with a dearth of low values. In particular cases, of course, it may happen that the inverse Gaussian gives a better fit than the three-parameter Weibull.

The above section has been concerned with the

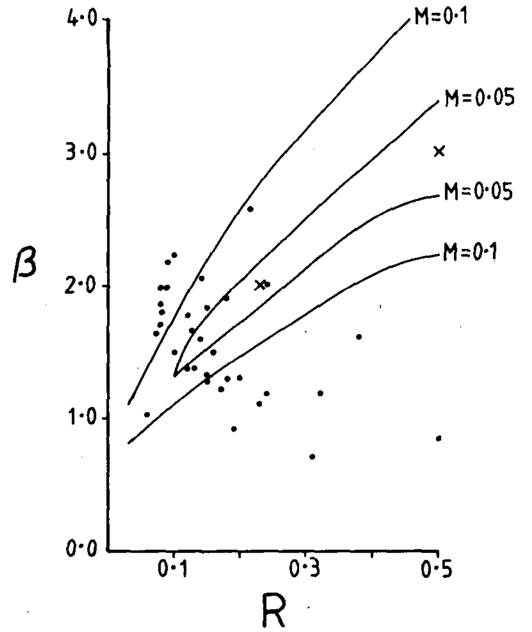

FIG. 3. Similarity contours of the inverse Gaussian and three-parameter Weibull distributions. Points correspond to β, R values taken from Table 6 of Stewart and Essenwanger (1978). Crosses correspond to the β, R values of Fig. 4. See text for symbol definitions.

use of relatively high-ϕ inverse Gaussian distributions as possible substitutes for three-parameter Weibull distributions. For low values of ϕ, the inverse Gaussian distribution is useful for describing frequency distributions with high peaks near zero and long right tails. Although such wind speed distributions are probably uncommon, the inverse Gaussian should be kept in mind as a possible alternative to the two-parameter Weibull if this situation arises.

3. Some further features of the inverse Gaussian distribution

This section describes some useful statistical properties of the inverse Gaussian distribution, as related to the description of wind speed data in general. It is assumed that a sample of N nonzero wind speed measurements has been obtained, and these measurements represent independent observations drawn from an inverse Gaussian distribution with unknown parameters.

The maximum likelihood estimators of μ and ϕ are given by

$$\hat{\mu} = \bar{x}, \tag{6}$$

$$\hat{\phi} = (\bar{x}\bar{y} - 1)^{-1}, \tag{7}$$

where $y_i = x_i^{-1}$. If required, exact confidence limits can be constructed around $\hat{\mu}$ (Chhikara and Folks, 1976). If N is large, the simplicity of (6) and (7) is clearly advantageous for calculation purposes, particularly when compared to the iterative procedure required for the maximum likelihood estimation of Weibull parameters. The estimators (6) and (7), of course, can be applied to any number set, $\hat{\mu}$ and $\hat{\phi}$ representing constants acting as a form of shorthand description of the data set concerned.

The presence of the population mean as a distribution parameter is a useful feature, since μ^3 multiplied by the appropriate constant represents a lower bound to the mean wind energy flux. A conservative estimate of this lower bound could be obtained through using a lower confidence limit for μ, given some specified level of probability.

An estimate of the mean wind energy flux is obtained by substituting $\hat{\mu}$ and $\hat{\phi}$ into the expression for the third moment of X about zero:

$$\mu_z = \tfrac{1}{2}\rho E(X^3) = \tfrac{1}{2}\rho\mu^3(1 + 3\phi^{-1} + 3\phi^{-2}), \tag{8}$$

where E denotes expectation, μ_z mean wind energy flux, and ρ is air density. The wind-energy pattern factor is given by

$$E(X^3)/\mu^3 = 1 + 3\phi^{-1} + 3\phi^{-2} \tag{9}$$

and is estimated by substituting $\hat{\phi}$ into (9). It is apparent that (9) decreases rapidly toward unity as

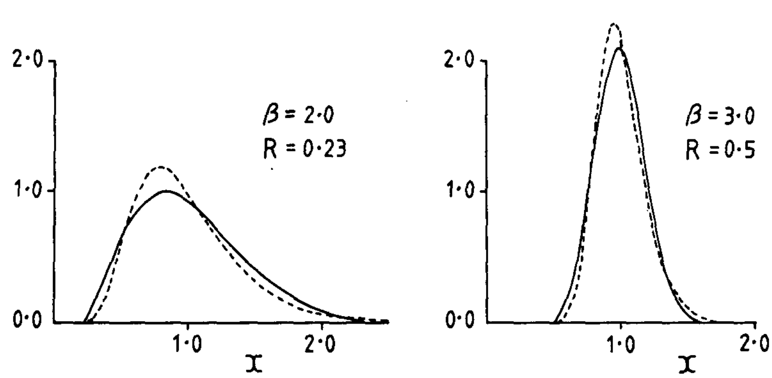

FIG. 4. Comparison of standardized probability density functions of the inverse Gaussian distribution (broken) and the three-parameter Weibull distribution (solid). All means are set to unity. See text for symbol definitions.

the inverse Gaussian distribution tends toward the normal with increasing ϕ.

Unlike the log-normal and two-parameter Weibull distributions, the inverse Gaussian distribution is not stable with respect to power transformations of X. The energy flux distribution is therefore not inverse Gaussian and is given by

$$h_Z(z) = \tfrac{1}{3}a^{-1}(\mu\phi/2\pi)^{1/2}(z/a)^{-7/6} \times \exp[-(z/a)^{1/3}\phi/2\mu + \phi - (z/a)^{-1/3}\mu\phi/2] \quad (10)$$

obtained by the change of variable $Z = aX^3$, where $a = \tfrac{1}{2}\rho$ and Z is wind energy flux. It is shown in the Appendix that $h_Z(z) \to 0$ as $z \to 0$, and there is a single mode at

$$z_m = a\{[(49 + 4\phi^2)^{1/2} - 7]\mu/2\phi\}^3. \quad (11)$$

The distribution function of Z is obtained through (2) as

$$H_Z(z) = F_X[(z/a)^{1/3}], \quad (12)$$

which can be used to construct power-duration curves for a given inverse Gaussian distribution of wind speed.

The above results apply to wind speeds distributed as inverse Gaussian random variables over the entire speed range. In wind energy surveys, it may only be necessary to describe the speed distribution within a specified interval, corresponding to the functional range of particular aerogenerators. In this situation the doubly truncated inverse Gaussian distribution could prove useful. As might be expected, the estimation procedures and moment properties are more complex than in the nontruncated case. Relevant details are given by Johnson and Kotz (1970, p. 146).

The inverse Gaussian distribution possesses a reproductive property under summation, a potentially useful feature for wind energy studies. It sometimes happens that the only data available from a site is in the form of speed values averaged over some relatively long time interval ΔT. For the purpose of estimating the mean wind energy flux, it is desirable to obtain some estimate of the speed distribution which would have been observed, had the speeds been averaged over some smaller time interval defined by

$$\Delta t = \Delta T/K, \quad (13)$$

where K is an integer. If the speeds averaged over Δt can be assumed to represent random values from an inverse Gaussian distribution, then the observed speeds averaged over ΔT also represent inverse Gaussian random variables. The parameters of the two distributions are related by

$$\mu_t = \mu_T, \quad (14)$$

$$\phi_t = \phi_T\Delta t/\Delta T, \quad (15)$$

and it follows from (8) that an estimate of the mean wind energy flux can be obtained by substituting the relevant parameter estimates into the expression

$$\mu_z = \tfrac{1}{2}\rho\mu_T{}^3[1 + 3(\Delta T/\Delta t)\phi_T{}^{-1} + 3(\Delta T/\Delta t)^2\phi_T{}^{-2}]. \quad (16)$$

Although (16) could prove useful in estimating mean energy flux from long-term speed averages, some care is required in selecting the size of Δt. It is evident that if K is large, ϕ_t will be small and the inferred speed distribution will have an unrealistically long right tail, resulting in very high mean energy flux estimates. Furthermore, in practice, a small Δt is likely to produce serial correlation, violating the random variable requirement. It would be of interest to test the usefulness of (16) by applying it to high-resolution wind data, using various combinations of ΔT and Δt.

At potential aerogenerator sites, it may be necessary to use wind speed measurements from some reference level L_0 to estimate the speed distribution at some higher level L_1. If the wind speed increases with height in terms of the logarithmic or power law models, then any speed distribution at L_1 must be the same as that at L_0 except for a scale change. In the case of the inverse Gaussian distribution (1), the scale change is achieved by a change in μ. Therefore, μ_1 can be expressed as a function of height and μ_0. For the power law model, this relation is simply

$$\mu_1 = (L_1/L_0)^\alpha\mu_0, \quad (17)$$

where α is the power law exponent.

4. Conclusion

The inverse Gaussian distribution represents a useful alternative to the three-parameter Weibull for the description of wind data containing low frequencies of low speeds. Although a good fit will not always be obtained, it would be of value to check the suitability of the inverse Gaussian before fitting the three-parameter Weibull distribution. Such preliminary investigations would require very little calculation time because of the simplicity of estimating μ and ϕ.

The inverse Gaussian distribution possesses a number of useful features with respect to wind energy evaluation. It is not suggested, of course, that any one property is unique to the distribution. Wise (1966) has noted a similarity of form between the log-normal and inverse Gaussian distributions, and the gamma distribution also possesses an additive reproductive property. However, when taken together, the various useful properties of the inverse Gaussian would indicate that the distribution could play a more important role in wind energy studies.

Acknowledgment. This work was carried out with the aid of funding provided by the New Zealand Energy Research and Development Committee.

APPENDIX

Form of the Energy Flux Distribution

1. Limiting value of Eq. (10) as $z \to 0$

As $z \to 0$, Eq. (10) simplifies to

$$(+\text{ constant})[z^{7/6}\exp(cz^{-1/3})]^{-1}, \qquad \text{(A1)}$$

where $c = a^{1/3}\mu\phi/2$. After expansion of the exponential term, the expression within the square brackets becomes

$$z^{7/6} + cz^{5/6} + c^2z^{3/6}/2! + c^3z^{1/6}/3! + c^4z^{-1/6}/4! + \cdots. \qquad \text{(A2)}$$

It is evident that $z \to 0$ implies (A2) $\to \infty$; therefore (10) $\to 0$ as $z \to 0$.

2. Derivation of z_m

Substituting $q = (z/a)^{1/3}$ into (10) and taking the derivative with respect to q gives

$$(+\text{ term})(-\phi\mu^{-1}q^2 - 7q + \mu\phi). \qquad \text{(A3)}$$

For positive q, there exists only a single solution to the quadratic in (A3)

$$q_m = [(49 + 4\phi^2)^{1/2} - 7]/2\phi\mu^{-1}. \qquad \text{(A4)}$$

Therefore, the value of z corresponding to q_m is given by

$$z_m = a\{[(49 + 4\phi^2)^{1/2} - 7]/2\phi\mu^{-1}\}^3. \qquad \text{(A5)}$$

REFERENCES

Chhikara, R. S., and J. L. Folks, 1976: Optimum test procedures for the mean of first passage time distribution in Brownian motion with positive drift (inverse Gaussian distribution). *Technometrics*, **18,** 189–193.

Johnson, N. L., and S. Kotz, 1970: *Continuous Univariate Distributions*, Vol. 1. Houghton Mifflin, 300 pp.

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

Schuster, J., 1968: On the inverse Gaussian distribution function. *J. Amer. Statist. Assoc.*, **63,** 1514–1516.

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

Tweedie, M. C. K., 1956: Statistical properties of inverse Gaussian distributions, I. *Ann. Math. Stat.*, **28,** 362–377.

Wise, M. E., 1966: Tracer dilution curves in cardiology and random walk and lognormal distributions. *Acta Physiolog. Pharmacol. Neerlandica*, **14,** 175–204.