APPLICATION OF THE GENERALIZED PARETO DISTRIBUTION TO WIND ENGINEERING

J.D. Holmes[+], W.W. Moriarty[*]

[+] CSIRO, Division of Building, Construction and Engineering, P.O. Box 56, Highett, Victoria, 3190
[*] Consultant Meteorologist, 143 Reservoir Road, Sunbury, Victoria, 3429

INTRODUCTION

The Generalized Pareto Distribution (G.P.D.) is of fairly recent origin [1], but clearly is of considerable significance for geophysical phenomena such as floods or extreme windstorms due to its close relationship with the Extreme Value Distributions [2,3]. In this paper, the application of the Distribution to the analysis of wind gusts generated by the passage of downbursts is described, and predicted extreme wind speeds by this method are compared with those derived from conventional extreme value analysis of annual maxima. The G.P.D. approach has been used for wind speeds previously by Lechner et al [4]; it was applied to data from 100 U.S. stations, but to annual maxima not separated by storm type.

GENERALIZED EXTREME VALUE DISTRIBUTION

Although the tradition in wind engineering has been to fit the Type I Extreme Value Distribution to extreme wind speeds following the work of Gumbel [5], in fact this Distribution is a special case of the generalized Extreme Value Distribution (G.E.V.):

$$F_v(V) = \exp\{-[1 - k\,(V-u)/a]^{1/k}\} \qquad (1)$$

In (1), k is a shape factor. When $k < 0$, the G.E.V. is known as the Type II Extreme Value (or Frechet) Distribution; when $k > 0$, it becomes a Type III Extreme Value Distribution (a form of the Weibull Distribution). As k tends to 0, Equation (1) becomes Equation (2) in the limit. Equation (2) is the well-known Type I Extreme Value Distribution, or Gumbel Distribution.

$$F_v(V) = \exp\{-\exp[-(V-u)/a]\} \qquad (2)$$

The G.E.V. with k equal to -0.2, 0 and 0.2 are plotted in Figure 2, in a form that the Type I appears as a straight line. As can be seen the Type III ($k = +0.2$) curves in a way to approach a limiting value at high return periods. Thus the Type III Distribution is appropriate for phenomena that are limited in magnitude for geophysical reasons. The Type I and Type II predict unlimited values as the return period increases; for this reason they would appear to be inappropriate for wind engineering applications. Thus, although it has been customary in extreme wind studies to assume that the Type I (or Gumbel) distribution applies, there seems to be no reason (apart from a desire to be conservative) why the Type III should not be preferred.

EXCESSES OVER HIGH THRESHOLDS - THE GENERALIZED PARETO DISTRIBUTION

The Generalized Pareto Distribution for a variate, y, takes the form :

$$F_y(y) = 1 - [1 - (k\,y / \sigma)]^{1/k} \qquad (3)$$

where σ is a scale factor, and k is a shape factor. The case $k<0$ is the usual Pareto Distribution.

The significance of the G.P.D. is that it is the appropriate distribution for the *excesses*, or differences, of independent observations above a defined threshold, given that the limiting parent distribution of the observations is one of the extreme value distributions [2,3]. It also turns out that the value of the shape factor, k, in the G.P.D. is the same shape factor appropriate to the underlying G.E.V. (Equation (1)).

The G.P.D. also has a 'threshold stability' property, i.e. if Y has a Generalized Pareto Distribution and u>0, then the conditional distribution of Y-u given Y>u is also Generalized Pareto. This leads to a useful fitting method from which k and σ can be determined [3]. Under certain conditions it can be shown [3] that :

$$E(Y-u \mid Y>u) = [\sigma - k(u-u_o)]/(1+k) \qquad (4)$$

where u_o is the lowest threshold chosen (>0). E( ) is the expectation operation.

Thus, if the mean observed excess over u is plotted against ($u-u_o$), then the plot should follow a straight line with slope $-k/(1+k)$ and intercept $\sigma/(1+k)$, if the Generalized Pareto Distribution is applicable. From the slope and intercept, the vaues of k and σ can be determined. A negative slope indicates that the data is 'in the domain of attraction' of the Type III Extreme Value Distribution. In reference [4], this approach is described as CME (conditional mean exceedance).

An estimate of the R-year return period value, V_R, of the underlying variable, V, can be obtained as follows. If the exceedance rate of the level u_o, is λ (per year), then the mean crossing rate of the level V_R is given by : $\lambda\,[1- F_y\,(V_R - u_o)] = \lambda\,[1- k\,(V_R - u_o)/\sigma]^{1/k}$, from Equation (3).

Setting this equal to 1/R, we obtain [3]:

$$V_R = u_o + \sigma\,[1 - (\lambda R)^{-k}] / k \qquad (5)$$

Note that, for k>0, as $R \rightarrow 0$, $V_R \rightarrow u_o + (\sigma/k)$, i.e. at high return periods, the predicted extreme wind speeds tend to a limiting value. The application of the method is explained in the following section.

EXAMPLE - DOWNBURST WINDS AT MOREE, N.S.W.

A survey of the Dines Anemometer records obtained by the Bureau of Meteorology for Moree, N.S.W., between 1965 and 1992, revealed 56 gusts of 40 knots (20.6 m/s) or greater, whose 'signature' indicated that they were produced by downbursts. An analysis of the average excess over the thresholds u = 20.3 (u_o), 22, 24,....34 (m/s), gave the values shown plotted against ($u-u_o$) in Figure 2. A straight line was fitted, from which values for σ of 5.507 (m/s), and for k of 0.190, were obtained. The value of λ is 2.0 (56 events in 28 years). Then applying Equation (5) for return periods between 10 and 10000 years, we obtain values for V_R given in Table I. These value show that an upper limit of about 45 m/s is being approached at high return periods.

There are only four years between 1965 and 1992 in which no downburst gust of 40 knots or greater, was recorded. Assuming that in these years, the maximum downburst gust was 39 knots, a full set of annual maxima becomes available. Then these values were ordered and plotted against the unbiassed reduced variate appropriate to the Type I Extreme Value Distribution suggested by Gringorten [6] :

$$x = -\ln\{-\ln[1-(i-0.44)/(N+1-0.88)]\} \qquad (6)$$

where i is the order and N is the total number of values (28 in this case)

A straight line fitted by the least squares method was used to produce the predicted values of V_R shown in Table I.

As indicated earlier, a value of k greater than 0 obtained from the Generalized Pareto Plot (Figure 2) indicates a Type III underlying Extreme Value Distribution. Since the Type III is a three parameter distribution, the value of the shape factor, k, must be specified when deriving an equivalent straight line fit. The equivalent to Equation (6) for the Type III, is the following :

$$x = a\{1 - \exp(-k\,x')\}/k \qquad \text{with } x' = -\ln\{-\ln[1-(i-0.30)/(N+1-0.60)]\} \tag{7}$$

In this case the Benard plotting position parameter has been used, as recommended by Cunnane[7] for the Type III Extreme Value Distribution. The value of k of 0.190 obtained from the Generalized Pareto Plot was used in Equation (7) for the Moree data The values of V_R for various values of return period, R, from the Type III fit to the annual maxima are also shown in Table I. Perhaps not surprisingly, as the same value of shape factor, k, applies, the values obtained were quite similar to those obtained via the Generalized Pareto Analysis via Equation (5). However, both these estimates are *significantly lower* at high return periods (and approaching an upper limit), than those obtained by fitting the Type I Distribution to the annual maxima. The Type I Extreme Value, or Gumbel, approach to prediction of extreme wind speeds for design appears to be a conservative one.

CONCLUSIONS

This paper has discussed the Generalized Pareto Distribution and its application to the statistical analysis of extreme wind speeds. Its main advantage is that it makes use of all relevant data on the high wind gusts produced by the storms of interest, not just the annual maxima, and it is not necessary to have a value for every, or nearly every, year to carry out the analysis. The Type I Extreme Value (Gumbel) Distribution appears to be a conservative one for geophysical phenomena, because it is unlimited as the return period increases, a property which contradicts intuitive expectations of the physical world. The Type III Extreme Value Distribution (k>0) is a more appropriate model for extreme wind speeds. Fitting the G.P.D. to the excesses over thresholds is a convenient method of determining the appropriate value of the shape factor, k.

ACKNOWLEDGEMENT

The work described in this paper was carried out as part of a project "Design of Transmission Line Structures under Severe Thunderstorm Winds" sponsored by the Australian Electricity Supply Industry Research Board, Powerlink (Queensland) and TransGrid (N.S.W.). The support of these bodies is gratefully acknowledged by the authors.

REFERENCES

1. Pickands, J. Statistical inference using extreme order statistics. Annals of Statistics, Vol.3, pp 119-131, 1975.

2. Hosking, J.R.M. and Wallis, J.R. Parameter and quantile estimation for the generalized Pareto Distribution. Technometrics, Vol. 29, pp 339-349, 1987.

3. Davison, A.C. and Smith, R.L. Models for exceedances over high thresholds. J.R.Statist. Soc. B, 1990.

4. Lechner, J.A., Leigh, S.D. and Simiu, E. Recent approaches to extreme value estimation with application to wind speeds. Part 1: The Pickands method. J.W.E. & Ind. Aerodyn., Vol.41, pp509-519, 1992.

5. Gumbel, E.J. Statistics of extremes. Columbia University Press, New York, 1958.

6. Gringorten, I.I. A plotting rule for extreme probability paper. J. Geophys. Res., Vol. 68, pp 813-814, 1963.

7. Cunnane, C. Unbiased plotting positions - a review. J. Hydrology, Vol. 37, pp 205-222, 1978.

TABLE I. Predictions of extreme winds (m/s) due to downbursts at Moree, N.S.W.

Return Period (years)	Generalized Pareto Method (excesses over threshold)	Extreme Value Type I (annual maxima)	Extreme Value Type III (k=0.190) (annual maxima)
10	32.9	31.8	31.8
20	34.9	34.4	33.7
50	37.2	37.7	35.8
100	38.7	40.2	37.2
200	40.0	42.8	38.4
500	41.5	46.1	39.7
1000	42.4	48.6	40.6
2000	43.3	51.1	41.3
5000	44.2	54.4	42.2
10000	44.9	56.9	42.8

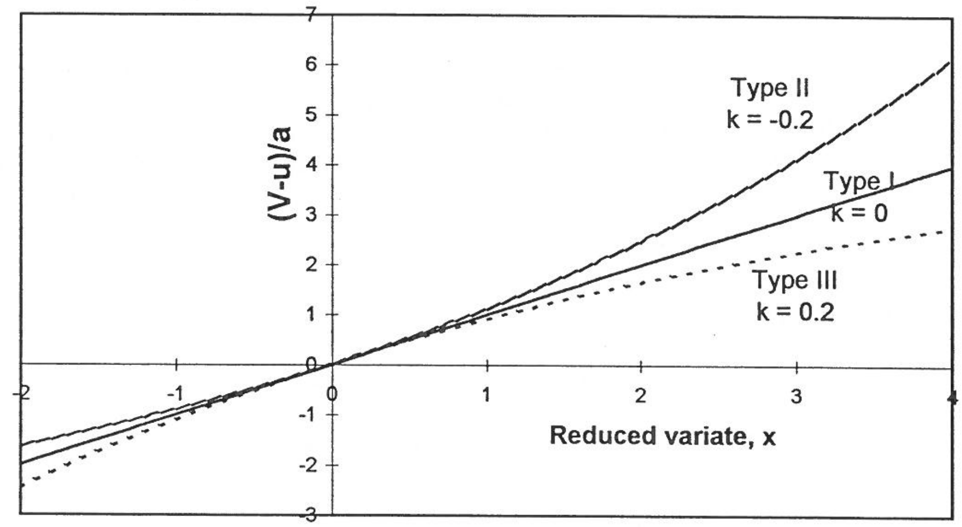

Figure 1. Generalized Extreme Value Distribution with k = -0.2, 0.0 and +0.2

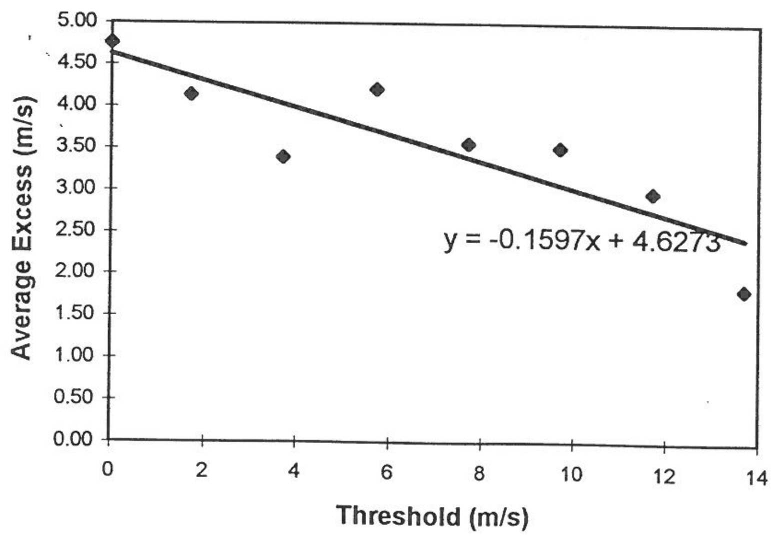

Figure 2. Generalized Pareto Plot for Moree downburst gusts