ESTIMATES OF HURRICANE WIND SPEEDS BY "PEAKS OVER THRESHOLD" METHOD

By N. A. Heckert, E. Simiu, Fellow, ASCE, and T. Whalen

ABSTRACT: We report results that lend support to the hypothesis that extreme hurricane wind speeds are described predominantly by reverse Weibull distributions, which have limited upper tails. This is consistent with the result reported recently in the atmospheric sciences literature that, on physical grounds, hurricanes have finite maximum intensity. The sampling errors typical of our results are controlled by the relatively short length of the historical record (about one hundred years, i.e., a few tens of hurricane events for any one location) and are, in practice, independent of whether the size of the simulated data samples for any specified location is of the order of 1,000 or larger. According to our estimates, mean recurrence intervals of wind speeds corresponding to nominal ultimate wind loads specified in the ASCE standard 7-95 for extratropical storm regions on the one hand, and hurricane-prone regions on the other, are mutually inconsistent with respect to risk.

INTRODUCTION

A fundamental theorem in extreme value theory states that sufficiently large values of independent and identically distributed variates are described by one of three extreme value distributions: the Fréchet distribution (with an infinite upper tail), the Gumbel distribution (whose upper tail is also infinite, but shorter than the Fréchet distribution's), and the reverse (negative) Weibull distribution, whose upper tail is finite. For a detailed exposition, see Castillo (1988).

The consideration of reverse Weibull distributions in wind engineering is the result of recent developments in extreme value theory, notably, the use of the "peaks over threshold" approach. Results reported by Simiu and Heckert (1995) lend support to the hypothesis that extreme wind speeds in extratropical storm regions are best fitted by reverse Weibull distributions. The purpose of this paper is to analyze hurricane wind speed data with a view to answering the question of whether this is also true of extreme hurricane wind speeds. This question is of interest in a structural reliability context, and in relation to the finding recently reported in the atmospheric sciences literature that, on physical grounds, hurricanes have finite maximum intensities (Emanuel 1988).

In the following sections, we describe the data and the peaks over threshold method of analysis used in this work, present results obtained by that method, comment on those results, and present comparisons with results based on simulations and analyses by other authors. The last section presents our conclusions.

HURRICANE WIND SPEED DATA

The wind speed data analyzed in this paper were obtained by simulation (Batts et al. 1980) to yield estimates of hurricane wind speeds that were used to develop the wind speed map included in the ASCE standard A7-93 and are currently used in an ASCE project on the development of load factors for combined wind and flood (Mehta, personal communication, 1997). They are available on tape (Accession No. PB821)

'Comp. Specialist, Computational and Appl. Math. Lab., Nat. Inst. of Standards and Technol, Gaithersburg, MD 20899.

²NIST Fellow, Building and Fire Res. Lab., Nat. Inst. of Standards and Technol., Gaithersburg, MD

³NIST-NRC Postdoctoral Res. Assoc., Building and Fire Res. Lab. Nat. Inst. of Standards and Technol., Gaithersburg, MD.

Note. Associate Editor: Ahsan Kareem. Discussion open until September 1, 1998. To extend the closing date one month, a written request must be filed with the ASCE Manager of Journals. The manuscript for this paper was submitted for review and possible publication on December 13, 1996. This paper is part of the *Journal of Structural Engineering*, Vol. 124, No. 4, April, 1998. ©ASCE, ISSN 0733-9445/98/0004-0445-0449/\$4.00 + \$.50 per page. Paper No. 12249.

32259, National Technical Information Service, Springfield, VA) and in anonymous files accessible to the reader as indicated in Appendix I. For each of more than 50 mileposts located at 50 nautical mile intervals on the Gulf and Atlantic coasts (mileposts are shown in Fig. 1 for multiples of 100 nautical miles), and for each of 999 simulated hurricane events, data are available as maximum wind speeds within each of the sixteen equally spaced azimuths: N, NNE, NE, ENE, E, ..., W, WNW, NW, NNW. The data represent nominal fastest one-minute hurricane speeds at 10 m above ground over open terrain at the coastline, in knots, and were originally obtained as the product of estimated fastest ten-minute speeds by a nominal factor of 1.18. In addition, the estimated annual rates of occurrence of hurricane events are available for each location.

In this study, we analyze data sets in which each of the data points is the maximum wind speed in a hurricane event, regardless of direction. To obtain the fastest ten-minute speeds in m/s, the wind speed data described above must be multiplied by the factors 1.151 mph/knot and 0.447 (m/s)/mph, then divided by the nominal factor 1.18 mentioned earlier. Finally, we obtain nominal mean hourly speeds by dividing the result by a factor of 1.05, representing the approximate ratio of fastest ten-minute speed to mean hourly speed.

ANALYSES AND RESULTS

Estimation of Tail Length Parameter of Generalized Pareto Distribution

The Generalized Pareto Distribution (GPD) is an asymptotic distribution whose use in extreme value theory rests on the fact that exceedances y of a sufficiently high threshold u are rare events to which the Poisson distribution applies. The expression for the GPD is

$$G(y) = \text{Prob}[Y \le y] = 1 - \{ [1 + (cy/a)]^{-1k} \}$$

$$a > 0, \quad (1 + (cy/a)) > 0$$
(1)

where a and c = location and the tail length parameter, respectively. Eq. 1 can be used to represent the conditional cumulative distribution of excess Y = X - u of the variate X over the threshold u, given X > u for a sufficiently large u (Pickands 1975). The cases c > 0, c = 0, and c < 0 correspond respectively to Fréchet (Type II Extreme Value), Gumbel (Type I Extreme Value), and reverse Weibull (Type III Extreme Largest Values) domains of attraction. For c = 0, the expression betweeen braces is understood in a limiting sense as the exponential $\exp(-y/a)$ (Castillo 1988).

For mileposts 150 through 2850 (successive mileposts are separated by 50 nautical miles), Simiu et al. (1996) reported

JOURNAL OF STRUCTURAL ENGINEERING / APRIL 1998 / 445



Fig. 1. Locator Map with Coastal Distance Intervals Marked (Nautical Miles; 1 Nautical Mile ~ 1.9 km) (Ho et al. 1987)

estimates of the tail length parameter c and 95 percent confidence bounds, as well as estimates of mean hourly speeds X_R at 10 m elevation over open terrain at the coastline for mean recurrence intervals R = 25, 50, 100, 1,000, and 2,000 years. The estimates are based on analyses of sets of data exceeding various thresholds u. They were obtained by using the de Haan procedure [de Haan 1994; see Simiu and Heckert (1995) for a review and details]. A few examples of estimates are shown in Fig. 2. Note that the smaller the threshold is, the larger the sample size (for example, for milepost 150, the sample size for a 38 m/s threshold is 26, whereas for a 37 m/s threshold, the sample size is 36), and the smaller the sampling errors are (i.e., the narrower the confidence bands). On the other hand, the smaller the threshold is, the larger the deviation from the assumption inherent in extreme value theory that the data are asymptotically large. If this deviation is too large, the extreme value model is no longer appropriate (Castillo 1988). In view of the dependence of the estimates upon threshold, the estimation is performed subjectively on the basis of plots such as those of Fig. 2, as discussed, for example, in Simiu and Heckert (1995) and references quoted therein. Given the errors associated with high thresholds on the one hand and low thresholds on the other, an optimum threshold in principle exists near which the graph is approximately horizontal [see Simiu and Heckert (1995) for additional details]. When choosing a reasonable value for the estimated value of tail length parameter c on the basis of our inspection of the graphs, it should be recalled that a larger estimate implies a longer tail and is therefore conservative from a structural engineering viewpoint.

We note that the confidence bounds of Fig. 2 are associated with the sampling errors due to the limited number of simulated wind speed data being analyzed. In addition to these sampling errors, the estimates are affected by climatological sampling errors, that is, sampling errors due to the limited number of climatological parameter data (pressure defect, radius of maximum wind speeds, translation velocity), on the basis of which the simulation of the wind speed data was carried out. Based on the data of Batts et al. (1980) used in this

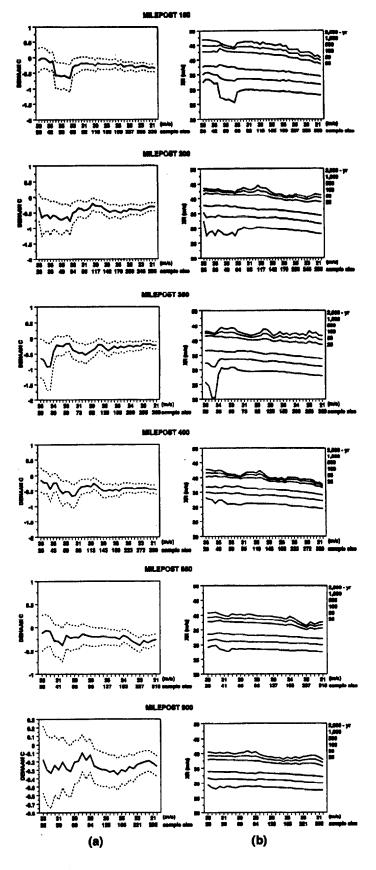
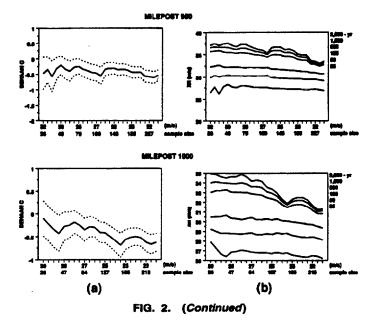


FIG. 2. (a) Estimated Tail Length Parameter c and 95% Confidence Bounds versus Threshold and Number of Threshold Exceedances; (b) Estimated Hourly Mean Speeds at 10 m above Ground over Open Terrain for Mean Recurrence Intervals of 25 Years (Lowest Curve), 50 Years, 100 Years, 500 Years, 1,000 Years (Successive Curves in Ascending Order) and 2,000 Years (Top Curve)



paper, for estimates of the 50-yr speeds, the estimated coefficient of variation of climatological sampling errors is about 10%, whereas the estimated coefficient of variation of sampling errors associated with a sample of 1,000 simulated hurricanes is about 2% (Batts et al. 1980; see also Delaunay 1988). The coefficient of variation of the total error is then $[0.1^2 + 0.02^2]^{1/2} = 0.102$. Augmenting the sample size of the simulated hurricanes from 1,000 to 10,000, for example, decreases the coefficient of variation of the total error from 0.102 to $[0.1^2 + (0.02)^2/10]^{1/2} = 0.1002$, an improvement of about 2%. An approximate estimate, based on the use of a method of moments formula [see, e.g., Simiu and Scanlan (1996)] yields for the 10,000-yr wind a coefficient of variation of the total error of about $[0.25^2 + 0.05^2] = 0.255$ if the sample size of the simulated hurricanes is 1,000 and about $[0.25^2 + 0.05^2]$ 10] = 0.250 if the sample size is 10,000, i.e., a comparable relative improvement. The reason for the marginal improvement brought about by the use of a larger number of simulated hurricanes is that the number of historical hurricane events from which the probabilistic models of the climatological parameters are constructed is very small (typically a few tens of hurricane events at any one station) in relation to the size of the simulated data sample (which, for the data used in this paper, is about 1,000). To summarize, by increasing the number of simulated hurricanes by a factor of ten, the precision of the estimates, conditional on a zero sampling error associated with the historical hurricane occurrences, is significantly improved. However, if that sampling error is taken into account, the improvement of the precision is insignificant. We conclude that increasing that sample size beyond about 1,000 achieves no useful purpose, except perhaps for certain types of directional wind speed analyses, which are not dealt with in this paper.

We now consider, as an example, the graph for the estimate of the tail length parameter c, for milepost 850 (Fig. 2). It is reasonable to infer from this graph that the estimated value of c is about -0.2. As a second example, consider the graph for milepost 900. In this case, a reasonable choice for the estimated value of c would be, for instance, -0.25. Our results show that the estimated values of the tail length parameter c are predominantly negative [see Simiu et al. (1996) for details]. This predominance is an indication that the reverse Weibull distribution is a better model of the hurricane wind speeds than the Gumbel distribution.

Estimation of Wind Speeds with Various Mean Recurrence Intervals

The mean recurrence interval R of a given wind speed, in years, is defined as the inverse of the probability that the wind speed will be exceeded in any one year. (This definition is valid for integer values $R \ge 1$, which are of interest in this work.) In this section, we give expressions that allow the estimation from the GPD of the value of the variate corresponding to any percentage point $1 - 1/[\lambda(u)R]$, where $\lambda(u)$ is the mean crossing rate of the threshold u per year. Note that, for any given location, $\lambda(u) = \mu_o n(u)/999$, where $\mu_o =$ annual rate of occurrence of hurricane events at that location; n(u) = number of wind speed data in excess of the threshold u; and 999 = number of wind speed data in the set with the lowest possible threshold (i.e., the number of data obtained by simulation for each location). Set

$$Prob[Y(u) < y] = 1 - 1/[\lambda(u)R]$$
 (2)

Using (1)

$$1 - [1 + c(u)y(u)/a(u)]^{-1/c(u)} = 1 - 1/[\lambda(u)R]$$
 (3)

therefore

$$y(u) = -a(u)\{1 - [\lambda(u)R]^{c(u)}\}/c(u)$$
 (4)

Recalling the definition of y(u), we get the value of the variate X being sought as

$$X_R(u) = y(u) + u \tag{5}$$

[for further details see Davison and Smith (1990)]. Consider for example the graph for milepost 850 (Fig. 2). Note a similarity between the dependence on sample size of the estimate of parameter c on the one hand and estimates of speeds with large mean recurrence intervals (e.g., R=2000 years and R=1000 years) on the other. The similarity is less pronounced for speeds with smaller mean recurrence intervals, for example, R=50 years. With relatively small error, it may be inferred conservatively from the graph that the mean hourly wind speeds are $X_{25} \sim 29$ m/s, $X_{50} \sim 32$ m/s, $X_{100} \sim 34$ m/s, $X_{500} \sim 38$ m/s, $X_{1000} \sim 40$ m/s, and $X_{2000} \sim 41$ m/s. For milepost 950, the choices for the 100-yr and 2000-yr mean hourly speeds are about 32 m/s and 38 m/s.

Table 1 shows estimated hourly mean hurricane wind speeds with 50-yr, 100-yr, and 2,000-yr mean recurrence intervals at 10 m above ground over open terrain near the coastline. Also shown in Table 1 are hourly mean speeds based on values estimated by Batts et al. (1980), Georgiou et al. (1983), and Vickery and Twisdale (1995). The values of Table 1 corresponding to estimates by various authors are mutually consistent from the point of view of averaging time, as pointed out in the discussion of Simiu (1997).

ANALYSIS OF RESULTS

For wind speeds over the coastline, the most significant difference between physical models used by the various authors listed in Table 1 involves the representation of the hurricane boundary layer. Unlike the other sets of estimates of Table 1, which used identical or similar empirical boundary layer models, the estimates by Vickery and Twisdale (1995) are based in part on the Shapiro modeling approach to simulating the boundary layer flow in a translating hurricane. As noted by Shapiro (1983), "the simple slab model with constant depth used in the present analysis cannot describe the detailed structure of the boundary layer, especially near the convective eye wall." Additional research would help to estimate the degree to which a simple slab model improves upon empirical models used in simulations of hurricane wind speeds.

TABLE 1. Estimated Hourly Mean Hurricane Wind Speeds with 50-yr, 100-yr, and 2,000-yr Mean Recurrence Intervals at 10 m above Ground over Open Terrain near Coastline, in m/s

Mean recurrence intervals (1)	Coastal distance* (2)	2 (3)	3 (4)	4 (5)	5 (6)	6 (7)	7 (8)	8 (9)			1	12 (13)	_		1							•	•					1
50-ут	Batts et al. (1980) This paper Georgiou et al. (1983) Vickery and Twisdale (1995)	34 35 33 30	33 34 31 29	32 35 32 31	31 33	31 33	32 33 33 34	32 32 35 34	31 32 34 32	29 29 31 32	30 34	33 32 35	36 34 36 32	36 36 39 35	35 37 38 36		32 31 31	28	33 33			34 33 33 32		27 29 29 30	30 30 30	33 34 32 30	33 33 31	28
100-ут	Batts et al. (1980) This paper Georgiou et al. (1983) Vickery and Twisdale (1995)	37 38	36 36 35 35	35 37 36	34 34 36	35 35 37	35 35 37 38	35 35 39	34 34 37	31 31 34	32 31 36	36 35 39	38 37 40	39 39 43	38 39 43	37 38 40	34 32	33 31 36	36 35 35	36 36 35	37 34 36	36 37 36	33 34	32 32 31	34 33 32 32	36 37 35	36 35	33 35 31 33
2,000-ут	Batts et al. (1980) This paper Georgiou et al. (1983) Vickery and Twisdale (1995)	47 46 51 45	46 42 47	45 43 48	45 44	44 44 49	44 41 50	45 44 51	44 41 50 45	40 35 46	43 35 48	46 46 51 44	46 43 55	47 46 57	47 47 56	45	40 39 45	41 38	47 44 49	46 43 48	46 42 49	45 47 48	44 43 45 45	43 41 43	45 43 47	47 46 50 45	48	46 45

Note: The estimates of this paper are based on the reverse Weibull distribution, which has a limited upper tail. All other estimates are based on other distributions. In hundreds of nautical miles (see Fig. 1).

Modeling and computational errors notwithstanding, the various sets of estimates of 50-yr winds listed in Table 1 are by and large comparable. This is understandable in view of the informal calibration of the models effected in investigations with a view to obtaining results that "make sense." We note, however, that there are differences of about ±10% between the estimates of this paper and those of Georgiou et al. for mileposts 200, 300, 400, 800, 1,100, 1,200, 1,400, and 1,800. The 50-yr wind speed estimates based on Vickery and Twisdale (1995) are in most cases comparable to the estimates of this paper, with differences in excess of 10% occurring for only very few locations. Given the many uncertainties that affect each set of estimates, it is difficult in our opinion to argue that one set of 50-yr wind speed estimates is much better than another.

The highest estimated hourly 2,000-yr speed at 10 m over water near the coastline, based on the results of this paper, is about 47 m/s \times 1.2 = 56.4 m/s (126 mph). (The coefficient 1.2 accounts for the difference between speeds at the standard 10 m elevation over water on one hand and over open terrain on the other (Simiu and Scanlan 1996). The corresponding fastest-minute speed and 3-sec speed are 56.4 \times 1.32 = 74.7 m/s (167 mph) and 56.4 \times 1.67 = 94 m/s (211 mph) (based on Krayer and Marshall 1992; ASCE 7-95).

It can be verified from Fig. 2 and similar plots (Simiu et al. 1996) that, for the estimates of this paper, the wind load factor $\phi_w = 1.3$ (adjusted to account for the hurricane importance factor of 1.05) specified in ASCE standard 7-95 and earlier versions thereof would, in most cases, correspond for wind-sensitive structures to nominal ultimate wind loads with mean recurrence intervals of, roughly, 500 years or less. (If epistemic uncertainties were taken into account, those mean recurrence intervals would be smaller than 500 years.)

Let us now consider extratropical storm wind speeds and assume that the coefficient of variation of the extreme value population is $V_v = 0.125$ —a typical value for such speeds. Mean recurrence intervals of wind speeds inducing nominal ultimate wind loads can be easily calculated for any given coefficient of variation of the extreme wind speed population and any tail length parameter of their reverse Weibull distribution (Simiu and Scanlan 1996). For example, for $V_v = 0.125$ and c = -0.2, the mean recurrence intervals of extreme wind speeds inducing nominal ultimate loads specified by ASCE 7-95 are, on average, more than two orders of magnitude larger than 500 years (again, under the assumption that epistemic errors are negligible). This strongly suggests that wind load factors specified in ASCE 7-95 for extratropical storm regions

on one hand and hurricane-prone regions on the other are mutually inconsistent with respect to risk, and therefore that, at least for structures whose design is governed by wind loads, the issue of specifying realistic wind load factors warrants further investigation.

It follows from the results of Table 1 that, on average, wind load factors would be closer to the value $\phi_w - 1.3$ currently specified in ASCE A7-93 if based on speeds modeled by the reverse Weibull distribution rather than existing models used by other researchers. To see this, note that the average estimated ratios of 2,000-yr speeds to 50-yr speeds are about 1.3, 1.4, 1.45, and 1.5 for the sets based on this report, Batts et al. (1980), Vickery and Twisdale (1995), and Georgiou et al. (1983), respectively, so that the squares of these values are about 1.7, 1.95, 2.1, and 2.25, respectively. A similar ordering would be obtained if speeds with other large mean recurrence intervals were considered instead of the 2,000-yr speeds.

CONCLUSIONS

The main conclusions of this work are:

- The results of our analyses are consistent with the assumption that reverse Weibull distributions are an appropriate probabilistic description of simulated extreme hurricane speeds along the Gulf and Atlantic coasts. A similar conclusion was reached by Simiu and Heckert (1995) with regard to wind speeds in regions not affected by hurricanes.
- 2. For any specified reasonably long mean recurrence interval, e.g., 2,000 years, simulated hurricane wind speeds described by reverse Weibull distributions tend to be lower than simulated speeds estimated by other procedures reported in the literature. In our opinion, it would be of interest to apply in the future the peaks over threshold approach to simulated hurricane wind speed data developed by other authors, should such data become available for public use.
- 3. Nominal ultimate wind loads obtained through multiplication of the 50-year loads by the load factor φ_w = 1.3 specified in the ASCE standard 7-95 appear to have significantly shorter estimated mean recurrence intervals than corresponding loads for extratropical storm regions. This apparent inconsistency with respect to risk warrants an investigation of the basis for wind load factor specification. Such an investigation is currently being conducted by the writers.

ACKNOWLEDGMENTS

We acknowledge with thanks partial support by the National Science Foundation (Grant No. CMS-9411642 to the Department of Civil Engineering, The Johns Hopkins University, subsequently transferred to the University of Colorado at Boulder), and partial support by the NATO Office of Scientific Affairs for collaborative work with a team headed by Dr. S. Coles of the Department of Mathematics and Statistics, University of Lancaster, U.K. Useful interactions with Dr. Coles and with Prof. R. B. Corotis of the University of Colorado are also acknowledged with thanks.

APPENDIX I. INSTRUCTIONS FOR ACCESSING DATA AND COMPUTER PROGRAMS

To access data and programs type first:ftp ftp.nist.gov; \u00fcuser anonymous; enter password \u00edyour e-mail address; \u00fcd /pub/ bfrl/emil. This places you in the main directory. Datasets and programs are stored in three subdirectories named maxyear, hurricane and directional. Each subdirectory has a readme file.

For example, to access the readme file for the hurricane directory, type from the main directory: cd hurricane; get readme. To get back to the main directory, type cd. / To access hurricane datasets or programs, from the main directory type: cd hurricane/datasets (to access datasets) or cd hurricane/programs (to access programs). Then type prompt off; dir; mget * (this copies all the data files). Once you are in the subdirectory hurricane/datasets, if you wish to get a specific file, type: get (nist.name) (local name) (example: get file35.dat file35.dat). To finish the session type:)quit

APPENDIX II. REFERENCES

- Batts, M. E., Cordes, M. R., Russell, L. R., Shaver, J. R., and Simiu, E. (1980). "Hurricane wind speeds in the United States." NBS Building Sci. Series 124, Nat. Bur. Stand., Washington, D.C.
- Batts, M.E., Cordes, M. R., and Simiu, E. (1980). "Sampling errors in estimation of extreme hurricane winds." J. Struct. Div., ASCE, 106(11), 2109-2115.

- Castillo, E. (1988). Extreme Value Theory in Engineering. Academic Press, Inc., New York, N.Y.
- Davison, A. C., and Smith, R. L. (1990). "Models of exceedances over high thresholds." J. Royal Stat. Soc., B-52, 339-442.
- de Haan, L. (1994). "Extreme value statistics." Extreme value theory and applications, J. Galambos, J. Lechner, and E. Simiu, eds., Vol. 1, Kluwer Academic Publishers, Boston, Mass.
- Delaunay, D. (1988). Vents extrêmes dus aux cyclones tropicaux dans les DOM-TOM. Centre Scient. et Techn. du Bâtim., Paris, France (in French).
- Durst, C. (1960). "Wind speeds over short periods of time." *Met. Mag.*, 89, 181-186.
- Emanuel, K. A. (1988). "The maximum intensity of hurricanes." J. Atmos. Sci., 45, 1143-1155.
- Georgiou, P. N., Davenport, A. G., and Vickery, B. J. (1983). "Design wind loads in regions dominated by tropical cyclones." J. Wind Eng. Ind. Aerodyn., 13, 139-152.
- Ho, F., Su, J., Hanevich, K., Smith, R., and Richards, F. (1987). "Hurricane climatology for the Atlantic and Gulf Coasts of the United States." NWS 38, National Oceanic and Atmospheric Admin., Silver Spring, Md.
- Krayer, W. R., and Marshall, R. D. (1992). "Gust factors applied to hurricane winds." Bull. Am. Met. Soc., 73, 613-617.
- Pickands, J. (1975). "Statistical inference using order statistics." Ann. Statistics, 3, 119-131.
- Shapiro, L. J. (1983). "The asymmetric boundary layer flow under a translating hurricane." J. Atmos. Sci., 40, 1984-1998.
- Simiu, E., and Scanlan, R. H. (1996). Wind effects on structures, 3rd Ed., Wiley-Interscience, New York, N.Y.
- Simiu, E., and Heckert, N. A. (1995). "Extreme wind distribution tails: a 'peaks over threshold' approach." NIST Building Sci. Series 174, NIST, Gaithersburg, Md.
- Simiu, E., Heckert, N. A., and Whalen, T. (1996). "Estimates of hurricane wind speeds by the 'peaks over threshold' method." NIST Tech. Note 1416, NIST, Gaithersburg, Md.
- Simiu, E. (1997). "Discussion of 'Prediction of hurricane wind speeds in the United States,' by Peter J. Vickery and Lawrence A. Twisdale." J. Struct. Engrg., ASCE, 123(11), 1549-1550.
- Vickery, P. J. and Twisdale, L. A. (1995). "Prediction of hurricane wind-speeds in the U.S." J. Struct. Engrg., ASCE, 121(11), 1691-1699.