Extreme wind load estimates based on Gumbel distribution of dynamic pressures: an assessment

by

E. Simiu
Building and Fire Research Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899 USA

N.A. Heckert and J.J. Filliben
Information Technology Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899 USA

and

S.K. Johnson
Building and Fire Research Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899 USA

Reprinted from Structural Safety, Vol. 23, No. 3, 221-229, 2001.

NIST CENTENNIAL 1901-2001

NIST
National Institute of Standards and Technology
Technology Administration, U.S. Department of Commerce

ELSEVIER

Structural Safety 23 (2001) 221–229

STRUCTURAL SAFETY

www.elsevier.com/locate/strusafe

Extreme wind load estimates based on the Gumbel distribution of dynamic pressures: an assessment

E. Simiu[a,*,1], N.A. Heckert[b,2], J.J. Filliben[b,3], S.K. Johnson[a,4]

[a]*Building and Fire Research Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA*
[b]*Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD 20899, USA*

Received 20 September 1998; received in revised form 24 January 2000; accepted 22 November 2000

Abstract

We present a contribution to the current debate on whether it is more appropriate to fit a Gumbel distribution to the time series of the extreme dynamic pressures (i.e. of the squares of the extreme wind speeds) than to fit an extreme value distribution to the time series of the extreme wind speeds themselves. It has been shown that the use of time series of the extreme dynamic pressures would be justified if the time series of the wind speed data taken at small intervals (e.g. 1 h) were, at least approximately, Rayleigh-distributed. We show that, according to sets of data we believe are typical, this is not the case. In addition, we show results of probability plot correlation coefficient (PPCC) analyses of 100 records of sample size 23 to 54, according to which the fit of reverse Weibull distributions to largest yearly wind speeds is considerably better than the fit of Gumbel distributions to the corresponding largest yearly dynamic pressures. We interpret the data and results presented in the paper as indicating that there is no convincing support to date for the hypothesis that the Gumbel distribution should be used as a model of extreme dynamic pressures.

Keywords: Building technology; Extreme value statistics; Structural engineering; Wind engineering; Wind forces

Introduction

Unless they are associated with resonant amplification or aeroelastic effects, wind loads may in general be assumed to be proportional to the squares of the wind speeds. Given a time series of

* Corresponding author. Tel.: + 1-301-975-6076; fax: + 1-301-869-6275.
E-mail address: emil.simiu@nist.gov (E. Simiu).
[1] NIST Fellow.
[2] Computer specialist.
[3] Mathematical statistician.
[4] Computer specialist.

0167-4730/01/$ - see front matter
PII: S0167-4730(01)00016-9

the extreme wind speeds, two methods have been proposed in the literature for estimating wind loads corresponding to various mean recurrence intervals. A first method uses time series of extreme wind speeds to fit an Extreme Value distribution to the data and estimate percentage points of the wind speeds. The estimated percentage points of the wind forces are then proportional to the squares of the estimated percentage points of the extreme wind speeds. A second method is based on the time series of the squares of the extreme wind speeds. (By definition, these are proportional to the dynamic pressures.) From these time series an Extreme Value Type I distribution is fitted to the squares of the wind speeds, and estimates are made of the percentage points of the squares of the extreme wind speeds. It has been suggested in the literature and in standards committees that estimates of wind forces based on the second method may be closer to the "true" percentage points than estimates based on the first method—see Refs. [1,2].

Differences between load estimates based on the two methods are significant. Monte Carlo simulations were reported by Simiu et al. [3] for 1000 sets of size 50 taken from an extreme wind speed population with Gumbel distribution and the reasonably typical expectation 30 m/s and standard deviation 4.5 m/s. The simulations showed that, under Cook's assumption [1] that both the extreme dynamic pressures and the corresponding extreme wind speeds have Gumbel distributions, the ratios of wind loads calculated by the second method to those calculated by the first method are 0.93, 0.85, and 0.78 for loads with 100-, 1000-, and 10,000-y mean recurrence intervals, respectively. These results are consistent with estimates by Cook [1].

In this paper we offer a contribution to the debate concerning the relative merits of the two methods just described. In Section 2 we review and assess the fundamental assumptions on the basis of which it has been stated that the second method is superior to the first: that it is possible to identify a parent population of the extreme wind speeds, and that this population is best fitted by a distribution that is, at least approximately, of the Rayleigh type [1,2]. In Section 3 we present results of PPCC analyses concerning the relative goodness of fit of the reverse Weibull distribution to sets of maximum yearly speeds on the one hand, and of the Gumbel distribution to the corresponding sets of dynamic pressures on the other. Section 4 contains our conclusions.

2. Hourly time series of wind speeds and the assumption of Rayleigh-distributed parent populations

The argument used by Cook [1] for using the dynamic pressure rather than the wind speeds when fitting the Gumbel distribution to extreme value time series rests on the assumption that the parent population from which the extreme speeds are extracted is fitted by a distribution that is, approximately, of the Rayleigh type. Cook based this assumption on analyses of sets of wind speeds measured at 1-h intervals, that is, on sets of $8760 \times n$ wind speed data, where n denotes number of years [1, p. 297]. If the Rayleigh distribution were correct, then the rate of convergence to the asymptotic Gumbel distribution of epochal maxima would be faster if the maxima consisted of dynamic pressures than if they consisted of wind speeds. Hence, an analysis in which a Gumbel distribution were fitted to a time series of extreme dynamic pressures would yield more realistic results than one in which a Gumbel distribution were fitted to a time series of extreme wind speeds. The assumption that the Rayleigh distribution models reasonably well the parent population of the extreme wind speeds is also used by Naess [2, pp. 254 and 256], who notes that since an annual record yields a total of 8760 h of data per year, there would be a reasonably high level of confidence attached to the Weibull—or Rayleigh—distribution of the parent population fitted to such a record, in spite of the existence of correlations among such data.

To assess Cook's and Naess's assumption we created histograms of wind speeds measured at one-hour intervals provided by the National Climatic Center (NCC) for the years 1996–1997 for the following stations: Bismarck, N. Dakota, Valentine, Nebraska, Harrisburg, Pennsylvania, Reno, Nevada, Boise, Idaho, Tucson, Arizona, and Dayton, Ohio. The histograms of the first four of these stations are shown in Figs. 1–4. (The data provided by NCC represent 2-min speeds in knots. To transform them into m/s the data should be multiplied by the factor 0.447 m/s/mph$\times$1.15 knots/mph$=$0.514 m/s/knot.) Note in Figs. 1–4 that at Bismarck and Valentine the mode of the wind speeds is about 7 knots$\times$0.514 m/s/knot $\approx$ 3.6 m/s; at Reno, the mode is considerably lower; and at Harrisburg the histogram is multimodal.

We now comment on the use of data such as those of Figs. 1–4 for inferences on the distribution of the extremes. In our opinion, the preponderance of very weak speeds casts doubt on the validity of such inferences. Weak speeds are mostly distinct meteorologically from the extreme wind speeds associated with powerful storms. To resort to a well-known comparison, inferences on powerful winds based on predominantly weak winds—morning breezes and so forth—are as unwarranted as inferences on the height of adults based on the heights of children in a kindergarten class. Inferences are a fortiori unwarranted if the distribution is multimodal, pointing even more strikingly to the existence of distinct types of winds, that is, of winds belonging to different classes of meteorological phenomena. Given the meteorological inhomogeneity of the data we believe that inferences on the probability distribution of a putative parent population from which the extremes are taken cannot be made with confidence from time series of 8760 hourly data per

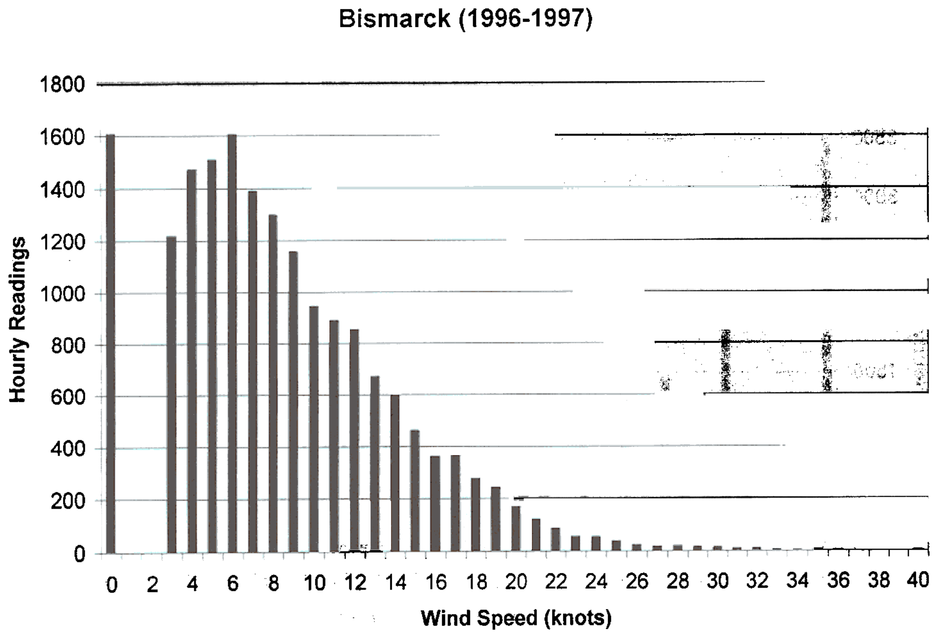

Fig. Histogram of wind speeds measured at 1-h intervals, Bismarck, ND.

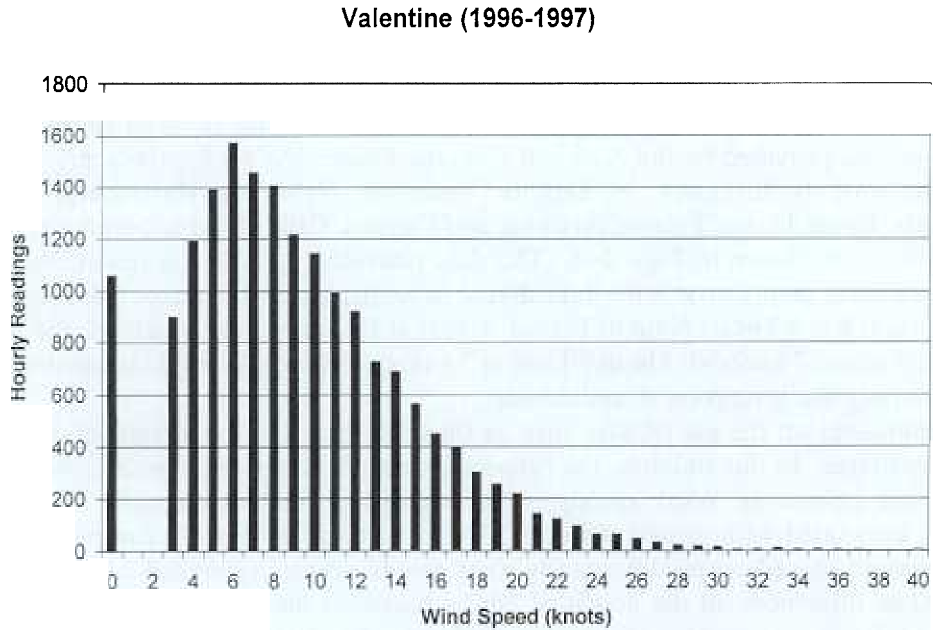

Fig. 2. Histogram of wind speeds measured at 1-h intervals, Valentine, NE.

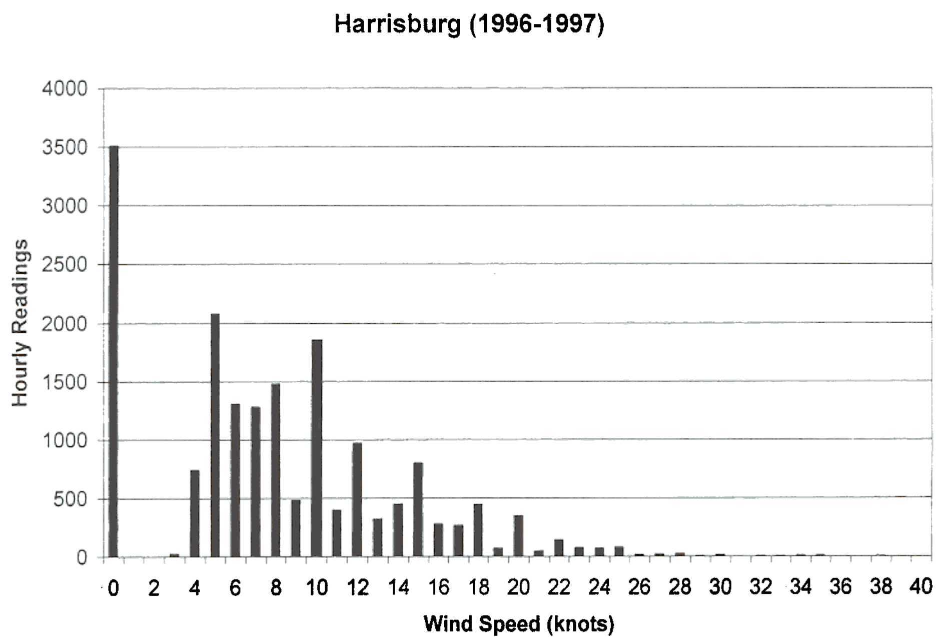

Fig. 3. Histogram of wind speeds measured at 1-h intervals, Harrisburg, PA.

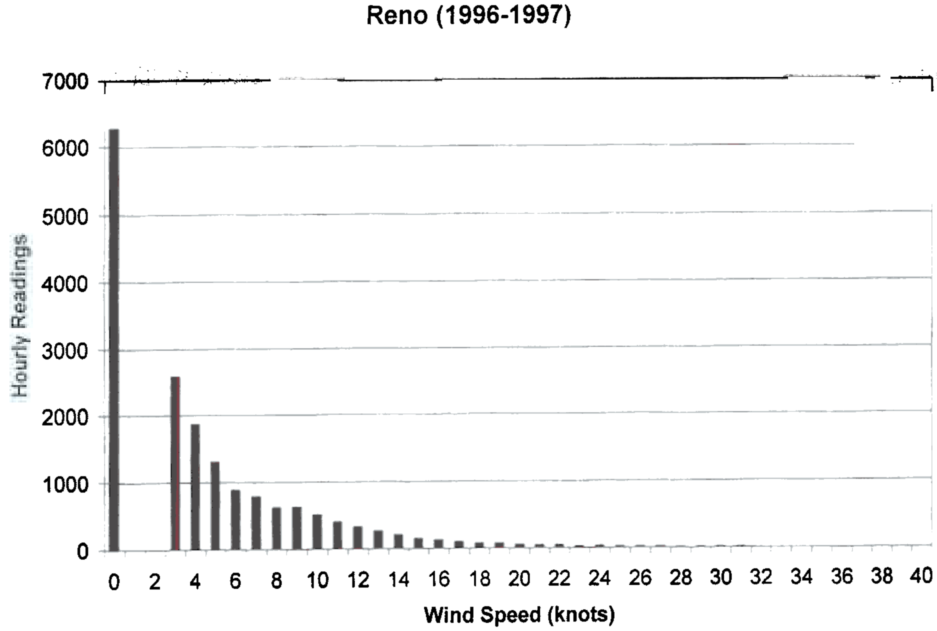

Fig. 4. Histogram of wind speeds measured at 1-h intervals, Reno, NV.

year, in spite of the small sampling errors that, ideally (i.e. in the absence of correlations among the data), would be inherent in such a large sample. We believe our conclusion would be warranted even if it were true that the best fitting distribution of the hourly data were Rayleigh. However, probability plot correlation coefficients (PPCC) goodness of fit tests indicated that this does not appear to be the case. The analyses consisted of PPCC calculations under the assumption that the data are fitted by a Weibull, a power lognormal, a lognormal, a reverse Weibull, a Gumbel, a Fréchet, a power normal, a normal, a Pareto, and a Rayleigh distribution. For none of the seven stations being analyzed was the PPCC largest for the Rayleigh distribution, that is, in all cases it was found that the Rayleigh distribution was not the best fitting distribution—by far—among the set of distributions just listed.

Rather than analyzing the entire set of hourly data, one may analyze the hourly data that exceed a sufficiently high threshold. This type of analysis, referred to in an Extreme Value context as a "peaks over threshold" approach, would be reliable if the data exceeding the threshold were, for practical purposes, statistically independent. Rather than applying a "peaks over threshold" approach to hourly data, it is more appropriate to use such an approach for sets of uncorrelated data extracted from relatively long sets of maximum daily data. Such use is consistent with the theory underlying the "peaks over threshold" approach. "Peaks over threshold" analyses of uncorrelated wind speed data have been performed by, among others, Simiu and Heckert [4]. According to their results the reverse Weibull distribution is an appropriate model of the extreme wind speeds. Since the reverse Weibull distribution is a tail-limited distribution, the distribution of the square of the wind speeds would also be tail-limited, and therefore it would not be a Gumbel distribution. Thus, whether the entire sample of 8760 wind speed data per year, or just those data exceeding a sufficiently high threshold, were used in the analysis, there appears to be

Table 1
Probability plot correlation coefficients (PPCCs) for 100 records of 23-y to 54-y length. For each station, following the years of record (in parentheses), the first and second number are the PPCC for reverse Weibull distribution of extreme wind speeds and Gumbel distribution of squares of extreme wind speeds, respectively. The closer the PPCC is to unity, the better is the fit of the distribution to the data.

Birmingham, AL (1944–1977):	0.98981, 0.98920
Montgomery, AL (1950–1983):	0.96001, 0.92078
Tucson, AZ (1948–1987):	0.97203, 0.96020
Yuma, AZ (1949–1987):	0.99129, 0.98962
Fort Smith, AZ (1952–1982):	0.97079, 0.96857
Little Rock, AK (1943–1981):	0.99056, 0.97900
Fresno, CA (1939–1975):	0.99377, 0.99346
Red Bluff, CA (1945–1986):	0.98424, 0.97722
Sacramento, CA (1949–1987):	0.98148, 0.97218
San Diego, CA (1940–1987):	0.95433, 0.91381
Denver, CO (1951–1983):	0.99027, 0.98949
Grand Junction, CO (1947–1979):	0.98009, 0.97517
Pueblo, CO (1941–1983):	0.99085, 0.98958
Washington, DC (1945–1984):	0.98629, 0.98935
Atlanta, GA (1935–1976):	0.99530, 0.98584
Macon, GA (1950–1982):	0.99499, 0.99425
Boise, ID (1940–1987):	0.99100, 0.98997
Pocatello, ID (1939–1987):	0.97735, 0.97108
Chicago Midway, IL (1943–1979):	0.99578, 0.99438
Moline, IL (1944–1987):	0.98855, 0.98246
Peoria, IL (1943–1984):	0.99266, 0.98744
Springfield, IL (1948–1979):	0.98090, 0.97633
Evansville, IN (1941–1984):	0.98777, 0.97947
Fort Wayne, IN (1942–1987):	0.99034, 0.98960
Indianapolis, IN (1944–1979):	0.96430, 0.93478
Burlington, IA (1942–1964):	0.97539, 0.96788
Des Moines, IA (1951–1987):	0.98460, 0.98665
Sioux City, IA (1942–1987):	0.98525, 0.97108
Dodge City, KS (1943–1983):	0.99437, 0.98348
Topeka, KS (1950–1983):	0.98295, 0.97328
Wichita, KS (1941–1981):	0.98329, 0.96684
Louisville, KY (1946–1984):	0.99119, 0.99156
Portland, ME (1941–1983):	0.97946, 0.96357
Detroit, MI (1934–1979):	0.99185, 0.99068
Grand Rapids, MI (1951–1979):	0.97941, 0.97690
Lansing, MI (1949–1986):	0.98407, 0.98267
Sault Ste Marie, MI (1941–1987):	0.99333, 0.99188
Duluth, MN (1950–1985):	0.99352, 0.99218
Minneapolis, MN (1938–1979):	0.96500, 0.93869
Jackson, MS (1948–1976):	0.98611, 0.98309
Columbia, MO (1950–1985):	0.99128, 0.97394
Kansas City, MO (1934–1984):	0.99321, 0.99136
Springfield, MO (1941–1984):	0.98520, 0.97639

(continued on next page)

Table 1 (*continued*)

Billings, MT (1939–1987):	0.99525, 0.99134
Great Falls, MT (1944–1987):	0.99121, 0.98798
Havre, MT (1961–1987):	0.98001, 0.97429
Helena, MT: (1940–1987):	0.98162, 0.97896
Missoula, MT (1945–1987):	0.96309, 0.94657
North Platte, NE (1949–1979):	0.99143, 0.98038
Omaha, NE (1936–1986):	0.93909, 0.97864
Valentine, NE (1956–1982):	0.99337, 0.97324
Ely, NV (1939–1987):	0.99607, 0.99465
Reno, NV (1942–1987):	0.99276, 0.99133
Winnemucca, NV (1950–1987):	0.98080, 0.97425
Concord, NH (1941–1986):	0.99146, 0.97834
Albuquerque, NM (1933–1984):	0.98747, 0.97347
Roswell, NM (1947–1982):	0.99039, 0.98373
Albany, NY (1938–1983):	0.96499, 0.94537
Binghamton, NY (1951–1985):	0.99433, 0.99337
Buffalo, NY (1944–1987):	0.96468, 0.94320
Rochester, NY (1941–1985):	0.98825, 0.98369
Syracuse, NY: (1941–1985):	0.98296, 0.98122
Charlotte, NC (1951–1979):	0.97860, 0.97610
Greensboro, NC (1930–1979):	0.98062, 0.95779
Bismarck, ND (1940–1979):	0.98662, 0.97678
Fargo, ND (1942–1986):	0.93881, 0.97282
Cleveland, OH (1942–1976):	0.99067, 0.98845
Columbus, OH (1952–1981):	0.98787, 0.97750
Dayton, OH (1943–1983):	0.99363, 0.99112
Toledo, OH (1943–1987):	0.98329, 0.95904
Oklahoma City, OK (1952–1981):	0.99354, 0.98980
Tulsa, OK (1943–1977):	0.97734, 0.97182
Portland, OR (1950–1987):	0.97037, 0.93645
Harrisburg, PA (1939–1976):	0.98926, 0.98887
Scranton, PA (1955–1987):	0.99217, 0.99259
Greenville, SC (1942–1984):	0.98864, 0.98652
Huron, SD (1939–1987):	0.99462, 0.99130
Rapid City, SD (1942–1984):	0.98579, 0.92885
Chattanooga, TN (1941–1975):	0.98643, 0.97540
Knoxville, TN (1942–1974):	0.98795, 0.98777
Nashville, TN (1942–1975):	0.98480, 0.96923
Abilene, TX (1944–1979):	0.92649, 0.94375
Amarillo, TX (1941–1974):	0.98574, 0.98263
Austin, TX (1943–1979):	0.98258, 0.98696
Dallas, TX (1941–1972):	0.99048, 0.98633
El Paso, TX (1943–1974):	0.97992, 0.98013
San Antonio, TX: (1941–1976):	0.96119, 0.92443
Salt Lake City UT (1942–1987):	0.99374, 0.99246
Burlington, VT (1944–1983):	0.98785, 0.97866
Lynchburg, VA (1944–1987):	0.98249, 0.97653
Richmond, VA (1951-1983):	0.98013, 0.97405

(*continued on next page*)

Table 1 (*continued*)

North Head, WA (1912–1952):	0.83733, 0.88610
Spokane, WA (1941–1987):	0.99011, 0.98738
Tatoosh Island, WA (1912–1965):	0.98857, 0.98863
Green Bay, WI (1949–1984):	0.94201, 0.95575
Madison, WI (1947–1987):	0.98563, 0.98189
Milwaukee, WI (1941–1982):	0.99322, 0.99092
Cheyenne, WY (1936–1981):	0.99226, 0.97843
Lander, WY (1946–1987):	0.99127, 0.98850
Sheridan, WY (1941–1984):	0.98236, 0.97781.

no support for the belief that the time series of the extreme dynamic pressures has a Gumbel distribution.

3. Results based on sets of maximum yearly wind speeds

We show in Table 1 the PPCCs calculated for 100 full sets of maximum yearly speeds under the assumption that the sets are best fitted by reverse Weibull distributions, and the PPCCs calculated for the corresponding sets of dynamic pressures under the assumption that those sets are best fitted by the Gumbel distribution. A comparison between the respective PPCCs shows that for 88% of the stations the fit of the reverse Weibull distribution to the wind speeds is better than the fit of the Gumbel distribution to the dynamic pressures. In our opinion this suggests there is no support for the belief that fitting a Gumbel distribution to extreme dynamic pressures yields better estimates of extreme wind loads than fitting a reverse Weibull distribution to the corresponding extreme wind speeds.

4. Conclusions

- For physical reasons—the unrepresentativeness of low wind speeds from the point of view of extreme wind speed estimation—probability distributions fitted to time series of wind speeds recorded at small intervals (e.g. 1 h) are in our opinion unlikely to provide a useful basis for inferences on extreme wind speeds.
- Seven sets of data recorded at one-hour intervals over two years at stations chosen randomly from stations not subjected to hurricane winds were subjected to probability plot correlation coefficient analyses. The distributions that were tested were: Weibull, power lognormal, lognormal, reverse Weibull, Gumbel, Fréchet, power normal, normal, Pareto, and Rayleigh. In all cases the Rayleigh distribution was—by far—not the best fitting distribution.
- Probability plot correlation coefficient (PPCC) analyses of sets of maximum yearly speeds and of the corresponding sets of squares of the wind speeds showed that, for 88 of the 100 stations for which analyses were performed, the fit of the reverse Weibull distribution to the sets of maximum yearly wind speeds is better than the fit of the Gumbel distribution to the

corresponding sets of squares of the wind speeds. Note, however, that in many instances the difference between the respective values of the PPCC was very small, and that additional analyses of this type would therefore be desirable.

- The results of the analyses presented in this paper are consistent with results published by Walshaw [5], Simiu and Heckert [4], and Holmes and Moriarty [6], according to which the reverse Weibull distribution is an appropriate probabilistic model of the extreme wind speeds. Calculations reported by Minciarelli et al. (2001) [7] show that the use of the reverse Weibull distribution in structural reliability estimates for structures subjected to wind loads result in nominal safety levels comparable to those of structures subjected to gravity loads, whereas the use of the Gumbel distribution results in far lower nominal safety levels (Ellingwood et al., 1980) [8]. This is another possible indication that the reverse Weibull distribution is a reasonable model of extreme wind speeds. Nevertheless, we do not advocate the use of the reverse Weibull distribution for codification purposes at this time. Rather, we believe that further investigations are desirable with a view to establishing definitively, if possible, the probabilistic model most appropriate for practical use.
- In our opinion, the results presented in this paper do not support the use for engineering calculations of estimates based on the Gumbel distribution of extreme dynamic pressures, as was advocated by Cook [1] and Naess [2].

Acknowledgements

We thank M.E. Changery, A. Chen and S. McCown of the National Climatic Data Center, National Oceanic and Atmospheric Administration, Asheville, NC, for their help in the acquisition of the hourly wind speed data.

References

[1] Cook NJ. Towards better estimates of extreme winds. International Journal of Wind Engineering and Industrial Aerodynamics 1982;9:295–323.
[2] Naess A. Estimation of long return period design values for wind speeds. Journal of Engineering Mechanics 1998; 124:252–9.
[3] Simiu E, Heckert NA, Filliben JJ. Comparisons of wind loading estimates based on extreme speeds and on square of extreme speeds. Proceedings, 13th ASCE Eng. Mechs. Div. Conference, Baltimore, 1999.
[4] Simiu E, Heckert NA. Extreme wind distribution tails: a "peaks over threshold" approach. NIST Building Science Series 174, and Journal of Structural Engineeering 1995, 1996:122;539–47.
[5] Walshaw D. Getting the most out of your extreme wind data. Journal of Research of the National Institute of Standards and Technology 1994;99:399–411.
[6] Holmes JD, Moriarty WW. Application of the generalized pareto distribution to extreme value analysis in wind engineering. Journal of Wind Engineering and Industrial Aerodynamics 1999;83:1–10.
[7] Minciarelli F, Gioofre M, Grigoriu M, Simiu E. Estimates of extreme wind effects and wind load factors: influence of knowledge uncertainties. Probabilistic Engineering Mechanics 2001;16:331–340.
[8] Ellingwood BR, Galambos TV, MacGregor JG, Cornell CA. Development of a probability based load criterion for American National Standard A58, NIST Special Publication 577. Washington, DC: National Bureau of Standards, 1980.