INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

UNITED NATIONS INDUSTRIAL DEVELOPMENT ORGANIZATION

INTERNATIONAL CENTRE FOR SCIENCE AND HIGH TECHNOLOGY
c/o INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS 34100 TRIESTE (ITALY) VIA GRIGNANO, 9 (ADRIATICO PALACE) P.O. BOX 586 TELEPHONE 040-224572 TELEFAX 040-224575 TELEX 460449 APH I

SMR/760-41

**"College on Atmospheric Boundary Layer
and Air Pollution Modelling"
16 May - 3 June 1994**

"Statistical Analysis of Extreme Wind Speeds"

G. SOLARI
Institute of Construction Science
University of Genoa
Genoa, Italy

Please note: These notes are intended for internal distribution only.

Statistical analysis of extreme wind speeds

G. Solari

Institute of Construction Science, Genoa University, Italy

1 Introduction

The statistical analysis of extreme wind speeds represents the fundamental preliminary step of all the studies concerned with structural safety under wind loads.

This chapter provides a systematic framework of the methods mostly used at present for determining the probability distribution of the maximum wind speed over a fixed period of time. Within this ambit, it deals initially with the most reliable procedures for well-behaved climates characterized by the occurence of ordinary winds only. It later provides a synthetic presentation of the methods available to generalize the above mentioned analysis to extraordinary aeolic phenomena such as hurricanes and tornadoes.

2 Ordinary aeolic phenomena

Consider an aeolic data base characterized by the representativeness, correctness and homogeneity properties discussed in paragraph 2 of the previous chapter ..... The data base consists of M values of wind speed V. Extract from this N samples, formed by $n = M/N$ values and corresponding to a

This text is based on the lecture given at the *College on Atmospheric Boundary Layer Physics*, 2nd Workshop on Modelling of the Atmospheric Flow Fields, 21 May – 1 June 1990.

Figure 1: Up-crossings of the v threshold.

unitary time period T, following a principle equivalent to casuality. The maxima of the N samples, provided independent, are defined as extremes. Within the ambit of engineering applications, it is usual to assume $T = 1$ year; in this case the extremes are identified with the annual maxima. Different approaches to this method are discussed by Lagomarsino et al [1990].

The probability distribution of the extreme values is usually assigned by means of the distribution function $F_M(v)$ of the maximum v being the state variable of V. As an alternative, it can be given by expressing v as a function of its mean return period $\overline{R}(v)$:

$$\overline{R}(v) = \frac{1}{1 - F_M(v)} \qquad (1)$$

defined as the number of time intervals T (normally years) which run, on average, between two successive aeolic occurrences exceeding the v threshold.

With the current extent of knowledge, the evaluation of these quantities in the presence of ordinary winds only is usually carried out on the basis of three alternative procedures defined as process analysis (paragraph 2.1), population analysis (paragraph 2.2) and asymptotic analysis (paragraph 2.3). A synthetic comparison between the results obtained when applying these procedures is given in paragraph 2.4.

2.1 Process analysis

The process analysis considers wind speed as a stochastic and stationary process [Gomes and Vickery, 1977]. The mean number $\overline{N}$ of the up-crossings of the v threshold (fig. 1) in the unit of time T is provided by the relationship [Rice, 1944, 1945]:

$$\overline{N}(v) = \int_{o}^{\infty} \dot{v} f_{V\dot{V}}(v, \dot{v}) d\dot{v} \qquad (2)$$

where $\dot{V}$ is the first derivative of the process V, and $f_{V\dot{V}}(v, \dot{v})$ is the joint density function of V and $\dot{V}$. Assuming V and $\dot{V}$ to be statistically independent, $f_{V\dot{V}}(v, \dot{v}) = f_V(v) f_{\dot{V}}(\dot{v})$, where $f_V(v)$ the density function of V (see Chapter ...) and $f_{\dot{V}}(\dot{v})$ is the density function of $\dot{V}$. In this case eq. (2) becomes [Lagomarsino et al., 1990]:

$$\overline{N}(v) = \lambda f_V(v) \qquad (3)$$

where:

$$\lambda = \int_{o}^{\infty} \dot{v} f_{\dot{V}}(\dot{v})\, d\dot{v} \qquad (4)$$

Assuming that the v threshold is sufficiently high, up-crossings may be considered as rare and independent events; as such they constitute a Poissonian process. The extreme distribution function is given therefore by the expression:

$$F_M(v) = e^{-\overline{N}(v)} \qquad (5)$$

Note that eq. (5) gives the distribution of maxima without necessarily knowing them. Making use of the population density, it becomes applicable to data bases containing a limited amount of data. Unlike analogous equations discussed in the following paragraphs, it does not allow to express v as an explicit function of the mean return period $\overline{R}$.

There are four techniques for estimating parameter λ. The first, based on its own definition (eq. 4), requires the knowledge of $f_{\dot{V}}(\dot{v})$, a problem at this stage. The second, proposed by Gomes and Vickery [1977], expresses λ as a function of the expected frequency of the wind speed process, analyzing the harmonic content of its power spectral density; this technique appears to be very burdensome and rather delicate from the computational point of view. The third thecnique, again by Gomes and Vickery [1977], enumerates the N up-crossings of the v threshold in time T, expressing λ according to the formula:

$$\lambda = E\left[\frac{\overline{N}(v)}{f_V(v)}\right] \qquad (6)$$

where $E[\ ]$ is the mean of the observed value. The fourth and last technique [Lagomarsino et al, 1990] estimates the λ parameter applying the order statistics by directly regressing eqs. (3) and (5) on the basis of the extremes:

$$\lambda = E\left[\frac{-lnF_M(v)}{f_V(v)}\right] \qquad (7)$$

Using a "Gumbel probability paper" [Benjamin and Cornell, 1970], fig. 2 shows some extreme distributions corresponding to estimates of λ obtained by means of eqs. (6) and (7). These diagrams also indicate the observed maxima. It is to be noted that eq. 6 fits the extremes only in the domain of the highest $\overline{R}$ values.

2.2 Population analysis

On the assumption that the N samples extracted from the data base are each composed of n statistically independent values, the distribution function of the maxima assumes the expression [Benjamin and Cornell, 1970]:

$$F_M(v) = F_V^n(v) \qquad (8)$$

$F_V(v)$ being the distribution function of V (see Chapter ...). In reality, by virtue of the process continuity, the hypothesis of statistical independence is quite unacceptable. This incongruity can be eliminated, while preserving the formal simplicity of eq. (8), by expressing the extreme distribution function by means of the formula [Davenport, 1968; Cook, 1982]:

$$F_M(v) = F_V^{n'}(v) \quad , \quad n' \leq n \qquad (9)$$

where n' is the amount of independent data during time T. Similarly to eq. (8), eq. (9) also provides the distribution of the maxima without necessarily having knowledge of them. This expression is also applicable to limited data bases. Substituting eq. (19) or eq. (21) of Chapter... in eq. (9), it is possible to render v explicit as a function of $\overline{R}$. The same result cannot be reached expressing $F_V(v)$ by eq. (25) in Chapter....

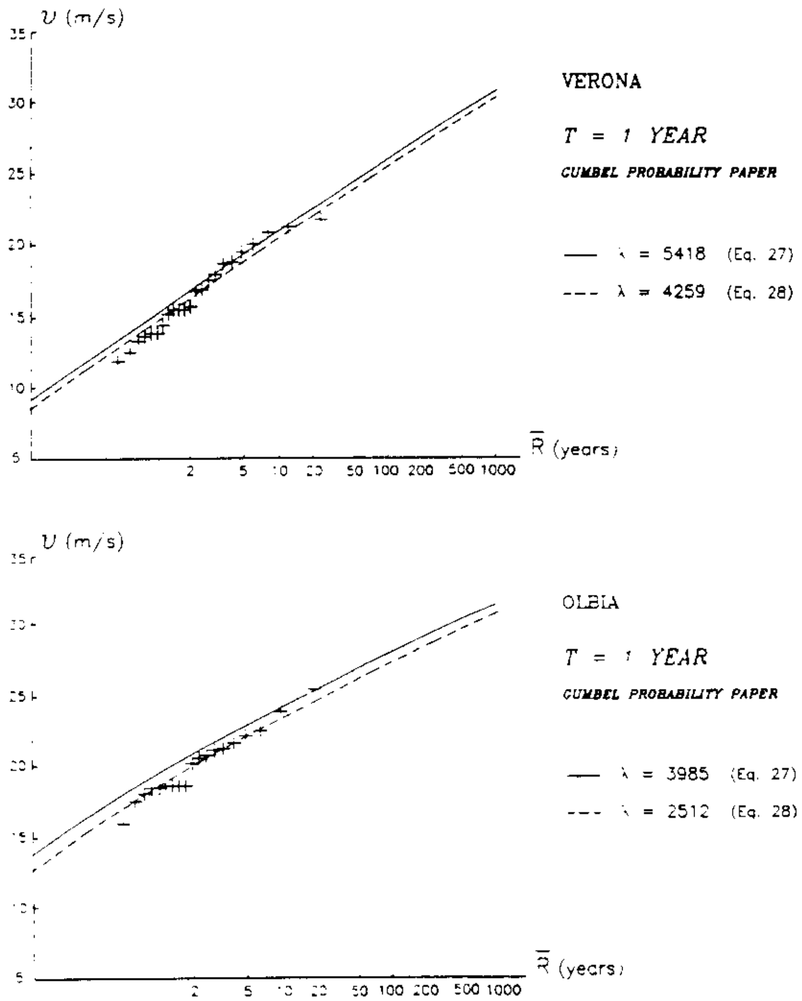

Figure 2: Examples of extreme wind speed statistics computed by process analysis.

The classic procedures proposed for estimating n' can be divided into three distinct categories. The first evaluates n' by means of the autocovariance function, according to a range of alternative methods discussed by Essenwanger [1986]. The second associates n' to the mean number of independent storms which occur during the period T [Cook, 1982]. The third expresses n' as a function of the expected frequency of the aeolic process, analyzing the harmonic content of its power spectral density [Davenport, 1968; Gomes and Vickery, 1977]. However, weak objective procedures are involved, sometime theoretically inconsistent and computationally burdensome. Nevertheless Lagomarsino et al [1990] have recently proposed two alternative methods. The first evaluates n' on the basis of the observed extremes $v_i(i = 1, ...N)$ through the non-linear least square regression:

$$\sum_{i=1}^{N} ln\{F_V(v_i)[F_V^{2n'}(v_i) - F_M(v_i)F_V^{n'}(v_i)]\} = 0 \qquad (10)$$

The second expresses $F_V(v)$ through eq. (19) in Chapter... and evaluates n', k and c simultaneously, k and c being the Weibull parameter; substituting eq. (18) of Chapter ... in eq. (9), the latter is traced in the "probability paper" governed by the expression:

$$ln\{-ln[1 - F_M^{1/n'}(v)]\} = klnv - klnc \qquad (11)$$

Applying the order statistics and least squares methods iteratively, n', k and c are estimated by minimizing the error function represented by the standard deviation of the distances from the observed extremes. This procedure has the intrinsic limit of precluding the expression of $F_V(v)$ by means of eq. (21) or eq. (25) in Chapter ... On the other hand, it possesses the considerable merit of allowing an estimate of k and c and, therefore, an allotment of $f_V(v)$ and $F_V(v)$, only the extremes being known.

Fig. 3 shows some results provided by the two above methods, assuming $T = 1$ year. The comparison between fig. 3 and fig. 5 in Chapter indicates that eq. (11) offers valuations of k and c reasonably close those obtained regressing the population by the censored technique. Moreover, utilizing three free parameters, it considers the maxima observed in a clearer manner than that possible by eq. (10). This does not mean that eq. (11) is to be preferred to eq. (10): although the latter estimates n', k and c instead of n' only, the former, being regressed over the entire data population, implicitly contains decidedly richer information.

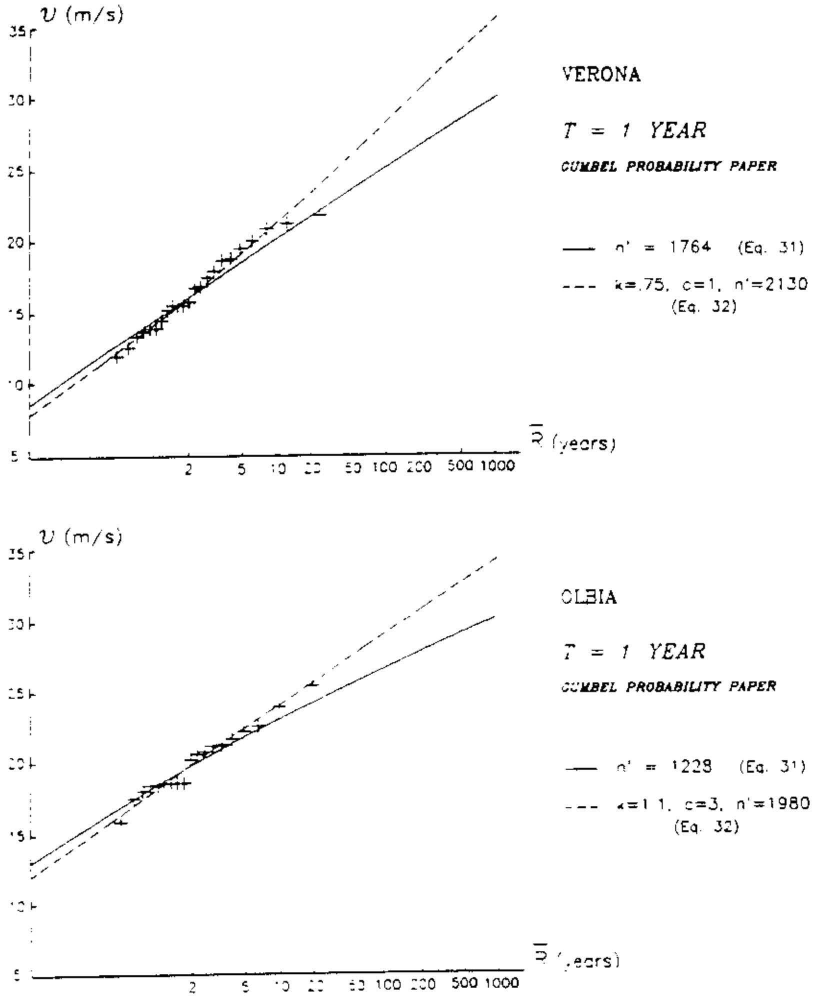

Figure 3: Examples of extreme wind speed statistics computed by population analysis.

2.3 Asymptotic analysis

It has been demonstrated [Fisher and Tippett, 1928] that, for n or n' tending towards the infinite, the distributions of the extreme values provided by eq. (8) and eq. (9) tend, according to rules depending on the tail of the population distribution [Gumbel, 1958], towards one of the three following asymptotic distributions denominated respectively of the I, II and III type:

$$F_M(v) = exp\{-exp[-a(v-u)]\} \quad , \quad -\infty < v < +\infty \qquad (12)$$

$$F_M(v) = exp\{-\left(\frac{\omega}{v}\right)^{\alpha}\} \quad , \quad v \geq 0 \qquad (13)$$

$$F_M(v) = exp\{-\left(\frac{y-v}{y-x}\right)^{\beta}\} \quad , \quad v \leq y \qquad (14)$$

where $a, u, w, \alpha, \beta, x, y$ are the parameters. Eqs. (12), (13) and (14) have the merit of modelling the extremes without requiring knowledge of the data population.

Literature on this subject, especially the less recent, gives preference, with different reasoning and according to the problem dealt with, to type I [Court, 1953; Shellard, 1963; Davenport, 1968; Simiu et al, 1978], type II [Bernstein, 1967; Thom, 1968] or type III distribution [Simiu and Filliben, 1980]. The more modern tendencies, however, agree on judging the type I asymptotic distribution as optimal. The reasons for this choice are many: assuming that the population distribution is represented by the Weibull model, it falls within the exponential tail function class where it is demonstrated that the distribution of the maximum tends towards eq. (12); the estimation of u and a is simpler than the estimation of the parameters of the type II and III distributions; once u and a are known, v may be expressed explicitly as a function of $\overline{R}$. The applicative studies indicate that the eq. (12) is the asymptotic form best fitting the extremes observed. It is to be noted, however, that this expression, being inferiorly unlimited, is not sufficiently adequate from the conceptual viewpoint for representing wind speeds. To this, it must be added that, if it is substantially acceptable to assume n or n' as tendentially infinite in the domain of ordinary return period values ($\overline{R} = 20 - 100$ years), this is found so much less correct the greater becomes this quantity [Cook, 1982; Lagomarsino et al., 1992].

The estimation of the parameters u and a is generally made on the basis of three classes of procedure: the moments method, the least square method and the BLUE method. The moments method has been found to be decidedly inadequate due to the non-symmetry of the type I distribution. The least square method has the limitation of being "biased". Among the BLUE methods, the technique proposed by Lieblein [1974] is currently the most accredited though it does not always provide completely reliable results. Some preliminary analyses indicate the use of robust and resistant methods as being the most promising [Hoaglin et al., 1983].

Fig. 4 compares some type I extreme distributions obtained by regressing the parameters u and a with different methods. It confirms the negative opinion expressed on the moments method; in the cases shown, substantial diversities do not appear between the use of the least squares method and the Lieblein method.

2.4 Comparison between different procedures

The practical comparison between the process, the population and the asymptotic analysis reveals that, when correctly applied, they furnish almost coincident results in the range of the return periods not greatly exceeding the number of years for which measurements are available. On the other hand, they often lead to increasing divergencies when considering return periods well above this range (fig. 5).

Lagomarsino et al. [1992] have demonstrated that these tendencies strictly depend on the k parameter of wind speed distribution (eqs. 18 - 21 in Chapter ...). If $k > 1$, the asymptotic analysis furnishes results on the safe side, while the process analysis is, with respect to the population analysis, more prudent. On the other hand, if $k < 1$, the asymptotic analysis becomes unconservative, while the population analysis is, at the same time, the most prudent and reliable. In the same paper, a general framework is also given of the most suitable approximate relationships aimed at estimating the parameters of a given model, assuming the parameters of another as known.

3 Extreme wind speed statistics - extraordinary aeolic phenomena

A sub-data base corresponding to winds of extraordinary type such as hurricanes (paragraph 3.1) or tornadoes (paragraph 3.2) contains, in principle, too small a number of elements (in most cases it does not contain any at all) to allow application of the statistical analysis techniques described in rela-

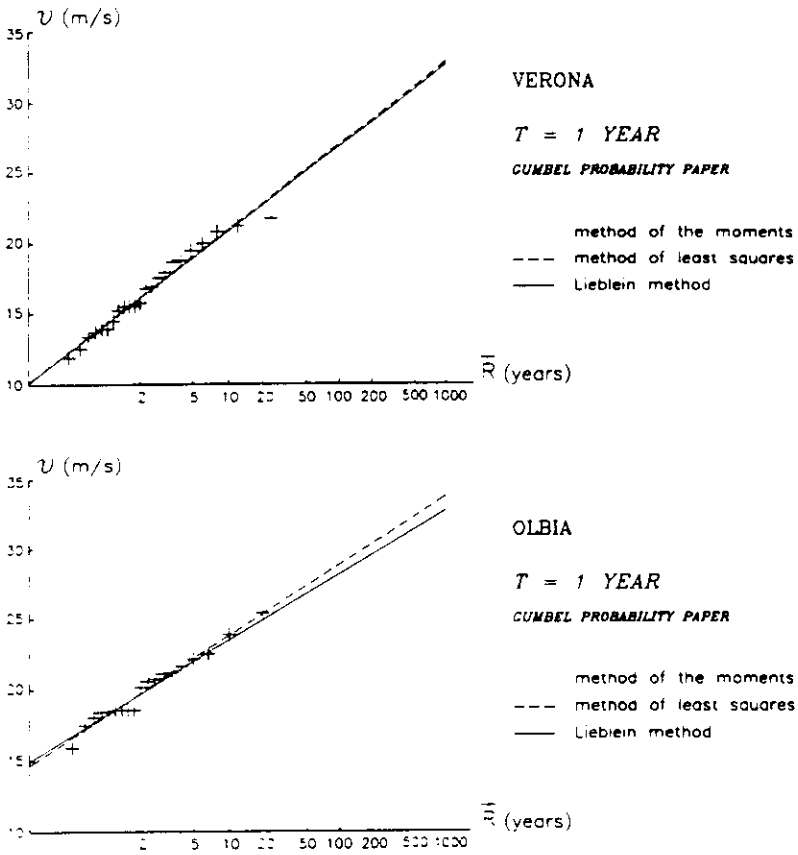

Figure 4: Examples of extreme wind speed statistics computed by asymptotic analysis.

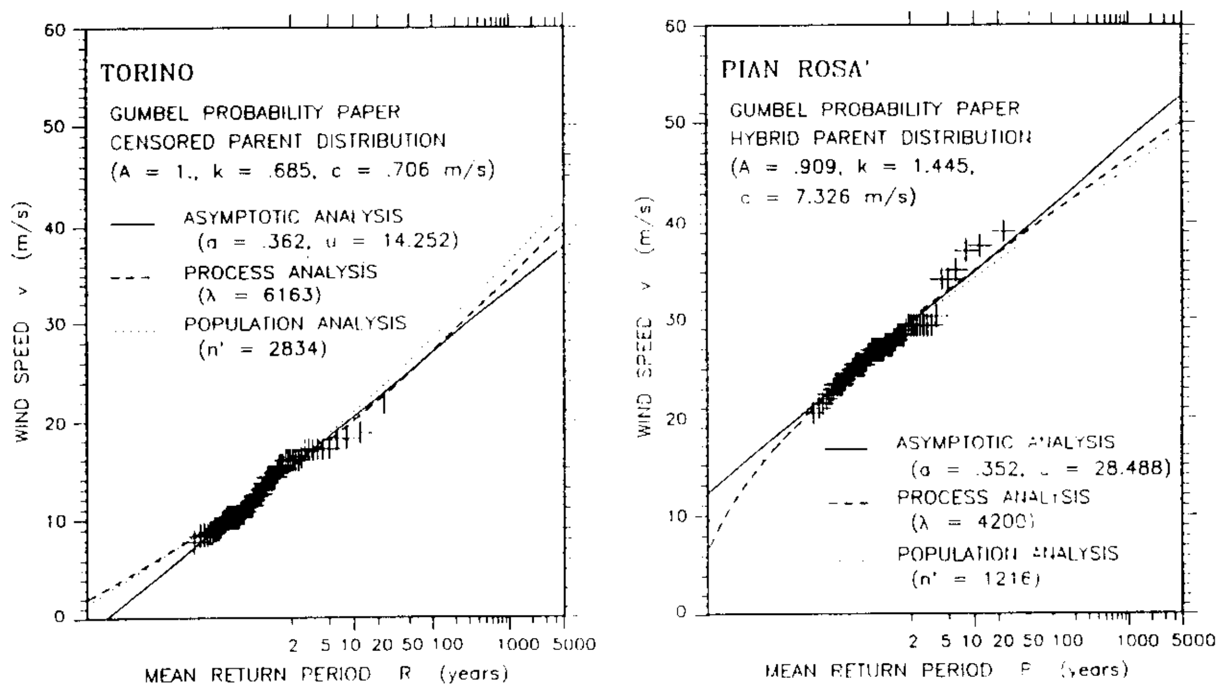

Figure 5: Comparison between the applications of different extreme wind speed statistics.

tion to ordinary winds. This makes recourse to alternative methodologies indispensable.

3.1 Hurricanes

Statistical analysis of wind speeds associated with hurricanes has long since been carried out utilizing the procedure formulated by Russell [1971] and summarized below:

1. the geographical areas in which hurricanes develop capable of striking the station considered are identified;

2. a mathematical model based on the physical parameters governing the behaviour of the hurricane is set up. The model parameters are characterized through probability distributions gained from the data gathered, over the course of years, in the areas under examination (fig. 6);

3. applying the Monte Carlo method, a historic series of hurricanes is generated such as to produce sufficient data at the station considered;

4. the artificial data thus acquired are submitted to the statistical analysis techniques previously described, thus determining the distribution function $F_M(v)$ of the maxima due to hurricanes for each station.

Detailed explanations and considerable developments of this procedure have been carried out by Tryggvason et al. [1976], Batts et al. [1980], Georgiou et al. [1983], Delaunay [1987], Sanchez-Sesma et al. [1988].

3.2 Tornadoes

Statistical analysis of wind speeds associated with tornadoes may, in principle, be carried out paraphrasing the criterion illustrated in relation to hurricanes. This results de facto of little advantage, since the area struck by an individual tornado is so limited as to render the generation time interval tendentiously infinite such as to produce sufficient data.

With the current state of knowledge the analysis is generally carried out following the formulation given to the problem by Thom [1963]:

1. the territory under consideration is divided into M regions characterized by homogeneous type tornadic occurrence;

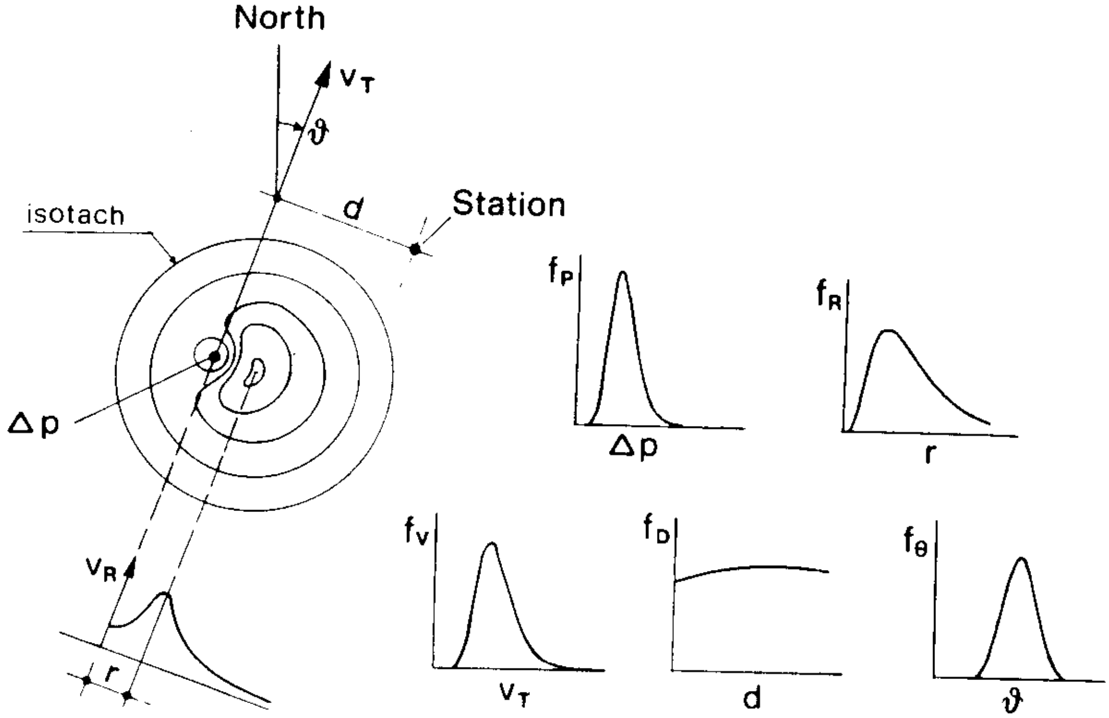

Figure 6: Hurricane model.

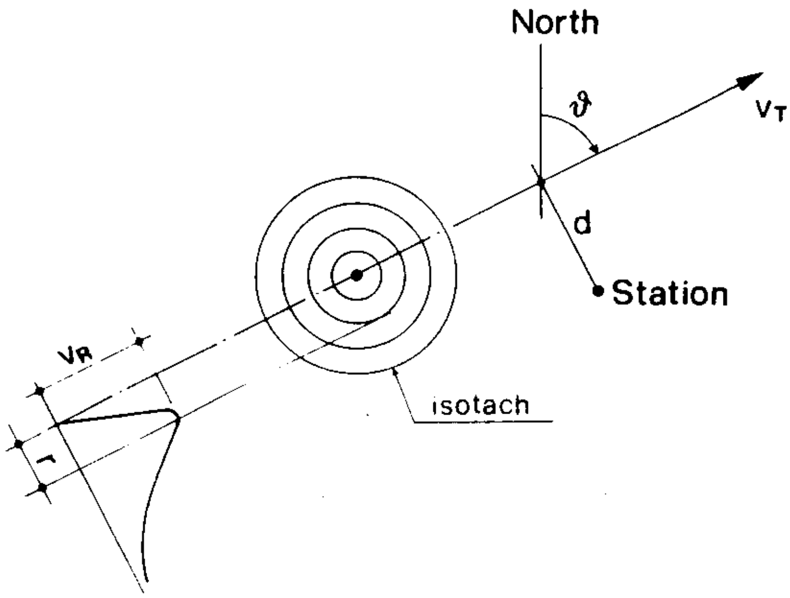

Figure 7: Tornado model.

2. treating the tornado as a Poissonian event, and utilizing the information gathered, the probability $F^*_M(v)$ is estimated that the maximum speed does not exceed the v threshold at a given point of the region containing the station examined;

3. the cumulative distribution of the maximum $F_M(v)$ is expressed as the product of $F^*_M(v)$ by the probability that the tornado may strike the station considered (fig. 7).

Substantial developments of this method have been proposed by Wen and Chu [1973], Garson et al [1975], Twisdale [1978], Twisdale and Dunn [1982].

Independently of the mathematical aspects of these procedures, it should be mentioned, however, that to the knowledge of the author, there are no current single anemometric recordings able to prove the occurrence of these phenomena (the extremely few instruments which have experienced a tornado have duly ended up out of order or destroyed); all values currently available have been gained by scales of intensity, following reconnaissance carried out in the areas hit [Metha et al., 1976; Wolde-Tinsae et al., 1985].

3.3 Mixed wind statistics

With knowledge of the distribution functions $F^{(k)}_M(v)$ $(k = 1, ..., L)$ of the annual speed maxima associated with the L aeolic phenomena considered, the global distribution function $F_M(v)$ may be expressed in an elementary form by introducing the hypothesis that said phenomena are statistically independent from each other [Gomes and Vickery, 1977/1978]. In this case it results:

$$F_M(v) = \prod_{k=1}^{L} F^{(k)}_M(v) \qquad (15)$$

Fig. 8 exemplifies this result taking advantage of a choice of coordinated axes such as to render the functions $v^{(k)}(\overline{R})(k = 1, ...L)$ linear; this is possible only in the case in which all the L distribution functions $F^{(k)}_M(v)$ are expressed through the same model. In figure 8, diagram (a) illustrates the case in which ordinary winds only are present. Diagram (b) shows what happens when, alongside ordinary winds (dotted line), exceptional type aeolic phenomena occur (dot-dash lines); the more these are rare and intense, the more the diagrams representing the same increase in their gradient and translate towards the greater values of the return period.

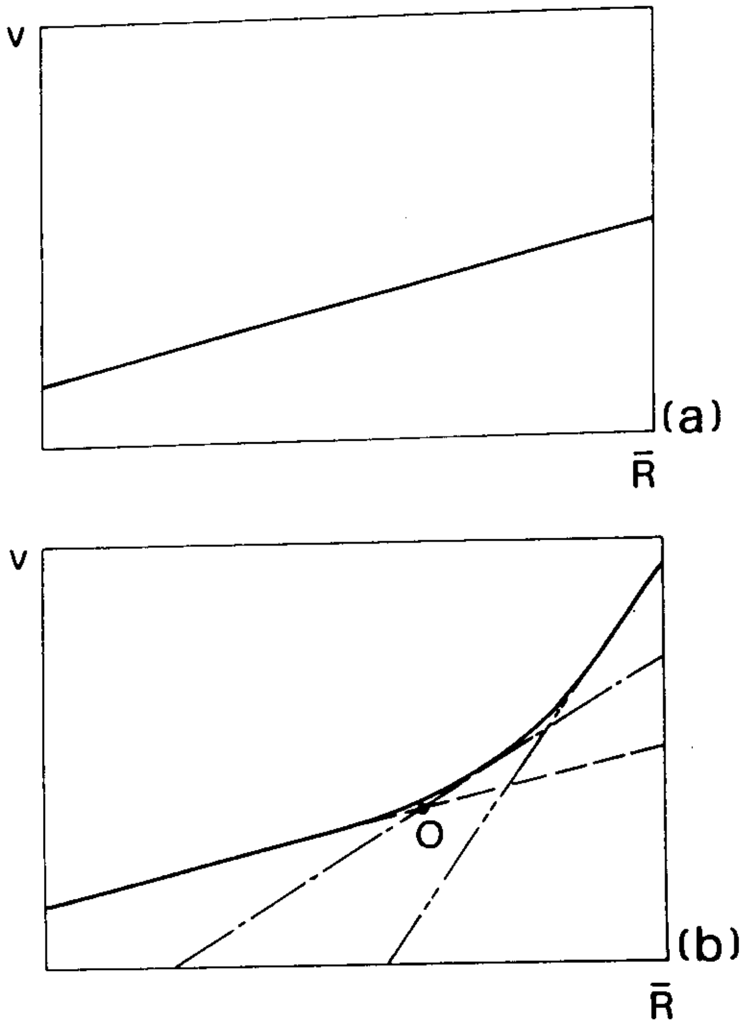

Figure 8: Comparson between extreme wind speed statistics referred to ordinary winds and mixed situations.

Experience shows that, in most cases, the first point of intersection 0 of the sheaf of curves in diagram (b) is set to the right of the number of years for which measurements are available. Significative exceptions to this principle are represented, for example, by the U.S. east coast and by the north Australian coast where tropical cyclones possess extremely high frequency and intensity.

4 Conclusions and perspectives

A synthetic outline has been furnished of the methods at present most accredited for analyzing extreme wind speed statistics. In addition to these methods some recent results are also described of a series of researches carried out by the author together with his co-workers prof. ing. Giulio Ballio, dott. ing. Sergio Lagomarsino and dott. ing. Giuseppe Piccardo.

The writer maintains that the quality of results obtained when applying the above techniques can be substantially improved by re-formulating the same procedures, taking into consideration the effects associated with wind direction. With the aim of developing this subject, a series of studies are at present being carried out, the preliminary results of which strongly confirm this concept.

Acknowledgements

The present paper is a summary and a compendium of a number of works concerning an extensive research program being carried out by the author in cooperation with prof. ing. Giulio Ballio of the Department of Structural Engineering of Milan Polytechnic, and with dott. ing. Sergio Lagormarsino and dott. ing. Giuseppe Piccardo of the Institute of Construction Science of Genoa University.

References

1. Batts, M.E., Russell, L.R., Simiu, E., 1980: "Hurricane wind speeds in the United States", *J. Struct. Div., ASCE*, **106**, 2001 - 2016.

2. Benjamin, J.R., Cornell, C.A., 1970: "Probability, statistics and decision for civil engineers", Mc-Graw Hill, New York.

3. Bernstein, M.F., 1968: "Theoretical bases for the method adopted in A.G.S.R. for the dynamic design of tall slender structures for wind effects", Proc. Int. Conf. on Wind Effects on Buildings and Structures, Ottawa, Canada, 2, 1 – 38.

4. Cook, N.J., 1982: "Towards better estimation of extreme winds", *J. Wind Engng. Ind. Aerod.*, **9**, 295 - 323.

5. Court, A., 1953: "Wind extremes as design factors", *J. Franklin Inst.*, **256**, 39 - 55.

6. Davenport, A.G., 1968: "The dependence of wind loads on meteorological parameters", Proc. Int. Conf. on Wind Effects on Buildings and Structures, Ottawa, Canada, 1, 19 - 82.

7. Delaunay, D., 1987: "Extreme wind speed distributions for tropical cyclones", *J. Wind Engng. Ind. Aerod.*, **28**, 61 - 68.

8. Essenwanger, O.M., 1986: "Elements of statistical analysis, in: General Circulation", 1B, Elsevier, Amsterdam.

9. Fisher, R.A., Tippett, L.H.C., 1928: "Limiting forms of the frequency distribution of the largest or smallest member of a sample", Proc. Cambridge Phil. Soc., 24, 180 - 190.

10. Garson, R.C., Catalan, J.M., Cornell, C.A., 1975: "Tornado design winds based on risk", *J. Struct. Div., ASCE*, **101**, 1883 - 1897.

11. Georgiou, P.N., Davenport, A.G., Vickery, B.J., 1983: "Design wind speeds in regions dominated by tropical cyclones", *J. Wind Engng. Ind. Aerod.*, **13**, 139 - 152.

12. Gomes, L., Vickery, B.J., 1977: "On the prediction of extreme wind speeds from the parent distribution", *J. Ind. Aerod.*, **2**, 21 - 36.

13. Gomes, L., Vickery, B.J., 1977/1978: "Extreme wind speeds in mixed wind climates", *J. Ind. Aerod.*, **2**, 331 - 344.

14. Gumbel, E.J., 1958: **Statistics of extreme**, Columbia University, New York.

15. Hoaglin, D.C., Mostaller, F., Tukey, J.W., 1983: **Understanding robust and exploratory data analysis**, John Wiley, New York.

16. Lagomarsino, S., Piccardo, G., Solari, G., 1990: "Criteri di modellazione probabilistica di basi dati coliche", Atti Primo Convegno ANIV, Firenze, Italy, 209 - 224.

17. Lagomarsino, S., Piccardo, G., Solari, G., 1992: "Statistical analysis of high return period wind speed", *J. Wind Engng. Ind. Aerod.*, **41**, 485 - 496.

18. Lieblein, J., 1974: "Efficient methods of extreme-value methodology", NBSIR 74 - 602, National Bureau of Standards, Washington, DC.

19. Metha, K.C., Minor, J.E., Mc Donald, J.R., 1976: "Wind speed analysis of April 3-4, 1974 tornadoes", *J. Struct. Div., ASCE*, **102**, 1709 - 1724.

20. Rice, S.O., 1944: "Mathematical analysis of random noise", *Bell. System Techn. J.*, **23**, 282 - 332.

21. Rice, S.O., 1945: "Mathematical analysis of random noise", *Bell. System Techn. J.*, **24**, 46 - 156.

22. Russell, L.R., Probability distributions for hurricane effects", *J. Water. Harb. Coast. Engng. Div., ASCE*, **97**, 139 - 154.

23. Sanchez-Sesma, J., Aguirre, J., Sen., M., 1988: "Simple modeling procedure for estimation of cyclonic wind speeds", *J. Struct. Engng., ASCE*, **114**, 352 - 370.

24. Shellard, H.C., 1963: "The estimation of design wind speeds", Proc. Symp. on Wind Effects on Buildings and Structures, Teddington, U.K., 1, 29 - 52.

25. Simiu, E., Bietry, J., Filliben, J.J., 1978: "Sampling errors in estimation of extreme winds", *J. Struct. Div., ASCE*, **104**, 491 - 501.

26. Simiu, E., Filliben, J.J., 1980: "Weilbull distributions and extreme wind speeds", *J. Struct. Div., ASCE*, **106**, 2365 - 2374.

27. Thom, H.C.S., 1963: "Tornado probabilities", *Mon. Wea. Rev.*, **91**, 730 - 736.

28. Thom, H.C.S., 1968: "New distributions of extreme winds in the United States", *J. Struct. Div., ASCE*, **94**, 1787 - 1801.

29. Tryggvason, B.V., Surry, D., Davenport, A.G., 1976: "Predicting wind-induced response in hurricane zones", *J. Struct. Div., ASCE*, **102**, 2333 - 2350.

30. Twisdale, L.A., 1978: "Tornado data characterization and wind speed risk", *J. Struct. Div., ASCE*, **104**, 1611 - 1630.

31. Twisdale, L.A., Dunn, W.L., 1982: "Probabilistic analysis of tornado wind risks", *J. Struct. Engng., ASCE*, **109**, 468 - 488.

32. Wen, Y., Chu, S., 1973: "Tornado risks and design wind speed", *J. Struct. Div., ASCE*, **99**, 2409 - 2421.

33. Wolde-Tinsae, A.M., Porter, M.L., Mc Keown, D.I., 1985: "Wind speed analysis of tornadoes based on structural damage", *J. Clim. Appl. Meteorol.*, **24**, 699 - 710.