# PROBABILISTIC ANALYSIS

The probabilistic analysis of the mean wind velocity consists in four steps:

1. Analysis of the wind velocity data base;
2. Probabilistic analysis of the current wind velocity, in order to determine its parent distribution;
3. Probabilistic analysis of the maximum wind velocity, usually the yearly maximum, in order to determine its extreme distribution;
4. Transformation of the parent and extreme distributions to a set of points in a given area and wind climate analysis.

## 1. Analysis of the wind velocity data base

In order to ensure that a wind data base can be correctly submitted to a probabilistic analysis, this shall be representative, reliable and homogeneous. A wind data base is representative if it is acquired over a sufficiently long period of time by an adequately located anemometric station. A wind data base is reliable if it is error-free. A wind data base is homogeneous if it is composed of values recorded in uniform conditions.

### 1.1. Representativeness of a wind data base

The representativeness of a wind data base is of the relative type. It depends on the duration of the recording period and on the location of the anemometer examined.

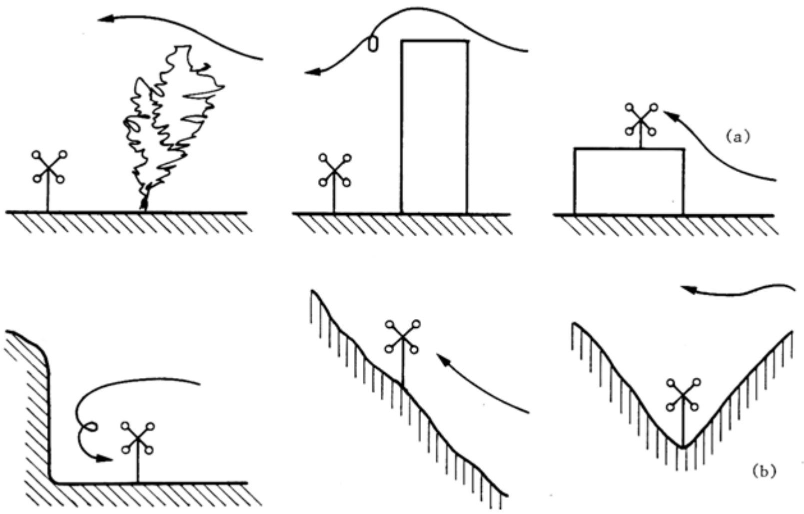
The duration of the recording period that makes a wind data base representative depends on:

1. The problem under examination. For instance, the safety analysis of a structure with a long nominal life requires a long-term data base;

2. The quality of the procedure applied for carrying out its probabilistic analysis. For instance, making recourse to refined probabilistic techniques allows one to analyze shorter data bases;

3. The accuracy sought. For instance, in order to obtain highly reliable probabilistic analyses longer data bases are necessary.

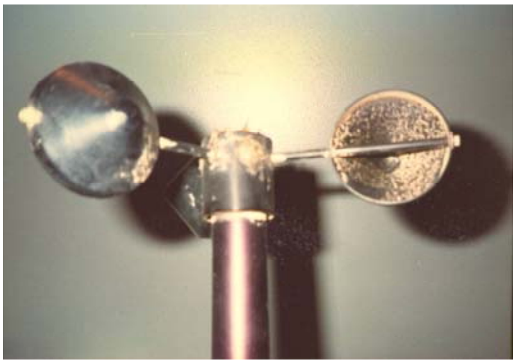
The location of the anemometer is a crucial point, provided that the data recorded are used to derive statistics in other positions. In such case, the anemometer should be put in an open terrain, at least at 10 m height; the vicinity of perturbing obstacles is acceptable to the degree in which transformation procedures successively employed are effective.




## 1.2 Reliability of a wind data base

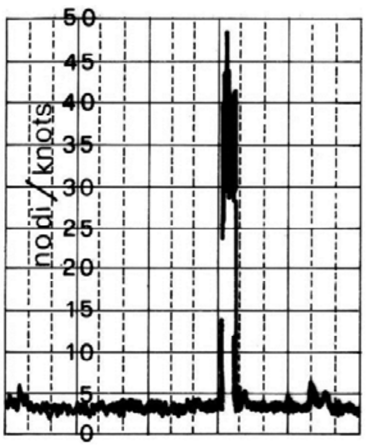
A wind data base may contain four main classes of errors:

1. Measurements made with poor-functioning instruments due, for instance, to lack of maintenance of wear;

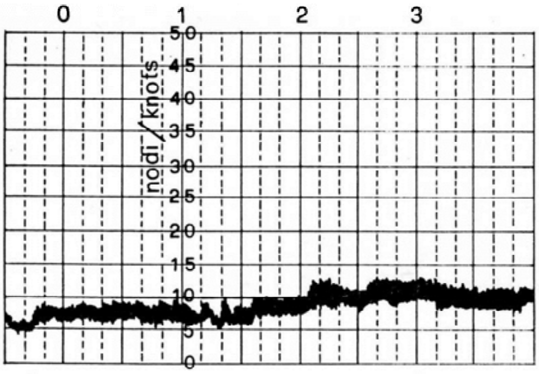


2. Measurements made outside the reading interval of the anemometer ($v_1$, $v_2$), such as during wind calms ($0 < v < v_1 \sim 0.5\text{-}2.0$ m/s) or wind phenomena of exceptional intensity ($v > v_2 \sim 50$ m/s);

3. Recording of events extraneous to the problem examined, e.g. airport measures concurrent with aircraft landing or take-off;



4. Human errors arisen in the transfer of data from graphs to cards or to digital files, as it was typical of the first periods in which data were recorded (around 1950-1975).



(a)

| hh | dd | ff |
|---|---|---|
| 00 | 24 | 06 |
| 03 | 23 | 42 |
| 06 | 25 | 04 |
| 09 | 00 | 00 |
| 12 | 00 | 00 |
| 15 | 00 | 00 |
| 18 | 00 | 00 |
| 21 | 00 | 00 |

(b)

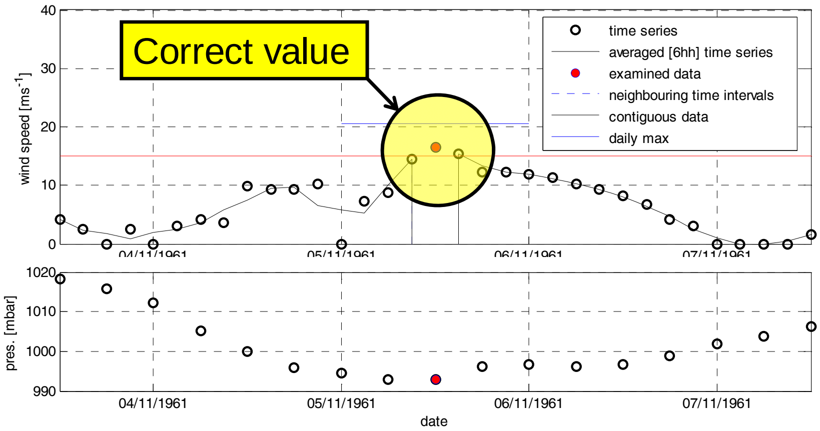
Besides, a wind data base may involve missing data due to instrument shut-down caused by natural events (e.g. extreme wind speeds or lightening), by operators (e.g. maintenance, replacement or transfer of anemometers), or by questionable acquisition procedures (e.g. mean wind velocity averaged over 10 minutes recorded every 3 hours, or stations operative for given daily periods).

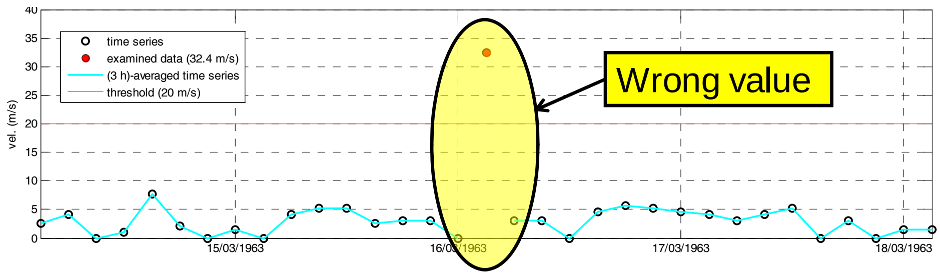
In all these cases, corrections procedures shall be applied. Such procedures are based upon the following criteria:

1. Correction of errors in the lower velocity range (e.g. wind calms) is fundamental for determining the parent distribution;

2. Correction of errors in the middle velocity range is usually not necessary since, statistically, such errors compensate with each other;

3. Correction of errors in the high velocity range is essential to determine the extreme distribution.

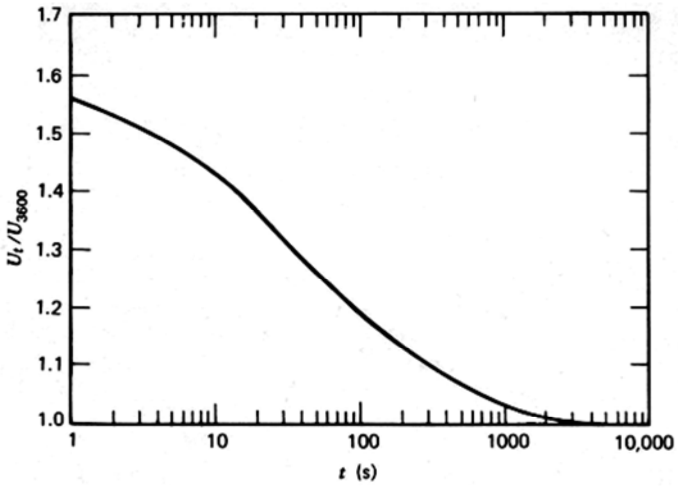


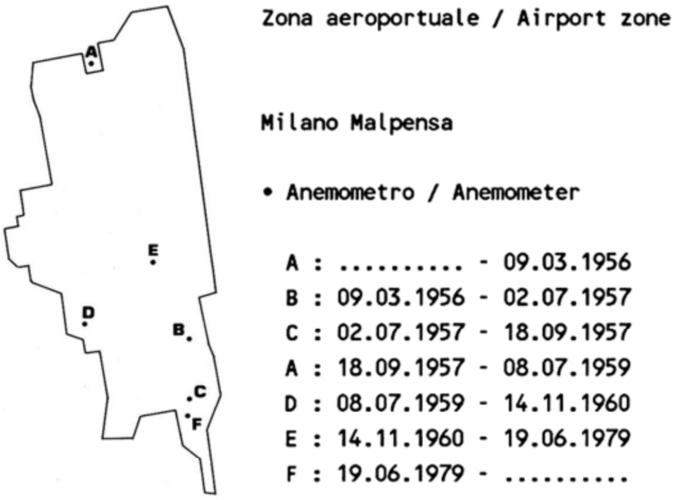
## 1.3 Homogeneity of a wind data base

In general terms, a wind data base may contain 3 main classes of heterogeneity:

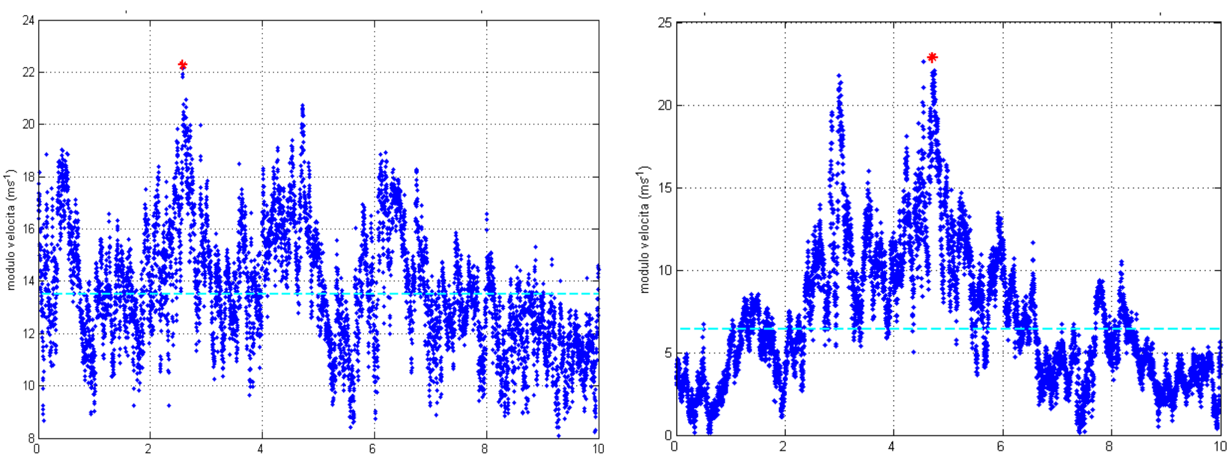
1. Values averaged on different time periods;




2. Shifting of the anemometer or evolution of the terrain surrounding the instrument;




3. Data associated with different wind phenomena, e.g. extra-tropical cyclones, gust fronts or tornadoes.




R

# 2. Parent distribution

## 2.1 Rayleigh and Weibull models

Studies carried out from 1946 have shown that, under suitable simplifying hypotheses, the parent population can be represented, in probabilistic terms, by the Rayleigh model:

$$F_V(v) = 1 - \exp\left(-\frac{v^2}{2\sigma^2}\right) \quad v \ge 0$$

$$f_V(v) = \frac{v}{\sigma^2}\exp\left(-\frac{v^2}{2\sigma^2}\right) \quad v \ge 0$$

where $\sigma$ is the only model parameter. This fact makes the Raileigh model rather rigid in order to represent a wide range of wind data base.

The choice of representing the parent population by the Weibull model originates from mere empiricism, and assumes the form:

$$F_V(v) = 1 - \exp\left[-\left(\frac{v}{c}\right)^k\right] \quad v \ge 0$$

$$f_V(v) = \frac{k}{c}\left(\frac{v}{c}\right)^{k-1}\exp\left[-\left(\frac{v}{c}\right)^k\right] \quad v \ge 0$$

The mean and the variance are given by:

$$\mu_V = c\,\Gamma\left(\frac{1}{k}+1\right)$$

$$\sigma_V^2 = c^2\left[\Gamma\left(\frac{2}{k}+1\right) - \Gamma^2\left(\frac{1}{k}+1\right)\right]$$

where $\Gamma(x)$ is the gamma function:

$$\Gamma(x) = \int_0^\infty \xi^{x-1} e^{-\xi} d\xi \quad (x > 0)$$

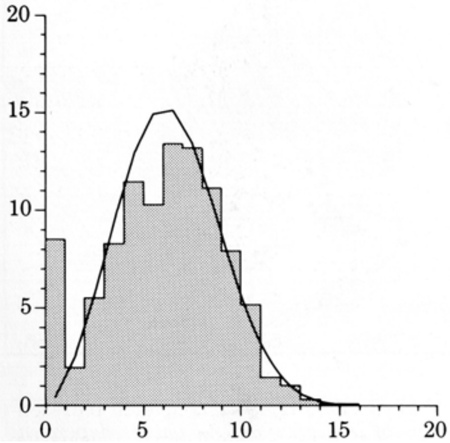
For $k > 1.3,\ \mu_X \simeq 0.89c,\ \sigma_X / \mu_X \simeq k^{0.92}$.

The Weibull distribution involves two parameters, k and c, which replace the single parameter $\sigma$ of the Rayleigh model. This ensures a greater flexibility for the data regression, without foregoing its formal simplicity. Moreover, the Weibull model coincides with the Rayleigh model for:

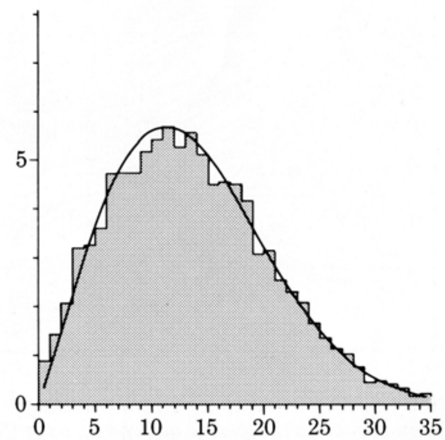
$$k = 2, \quad c = \sigma\sqrt{2}$$

The following figures show the accuracy with which a Weibull model with suitable k and c parameters fits the data recorder in several parts of the world.

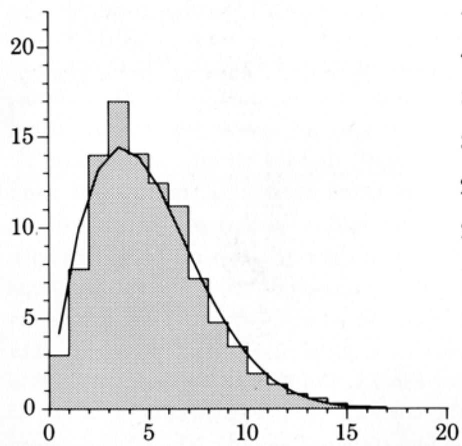
Fuerteventura Canary Islands, Spain
c = 7.2 ms⁻¹, $k$ = 2.78



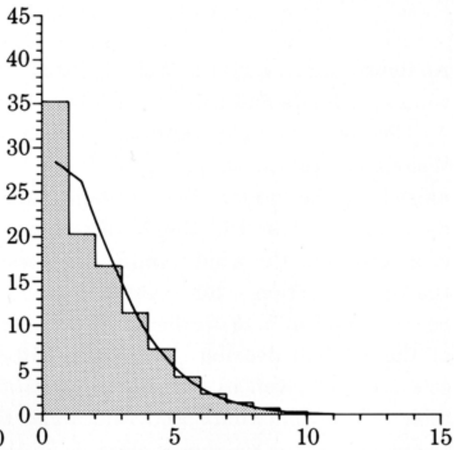
Snaefell, UK
c = 15.4 ms⁻¹, $k$ = 2.08

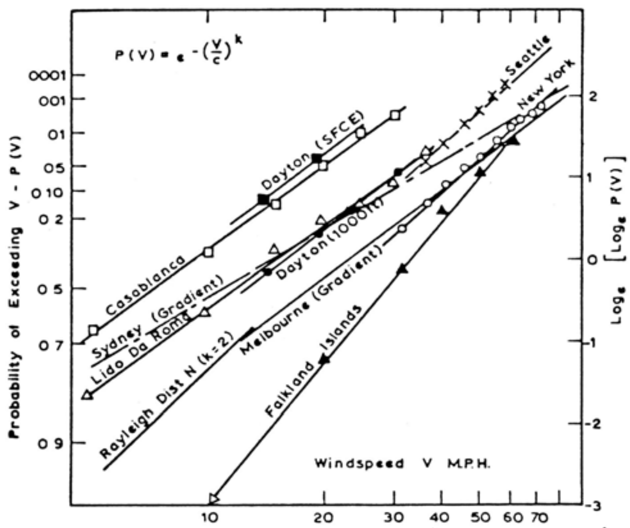


Schiphol, The Netherlands
$A$ = 5.6 ms⁻¹, $k$ = 1.83



Mont de Marsan, France
$A$ = 2.4 ms⁻¹, $k$ = 1.24





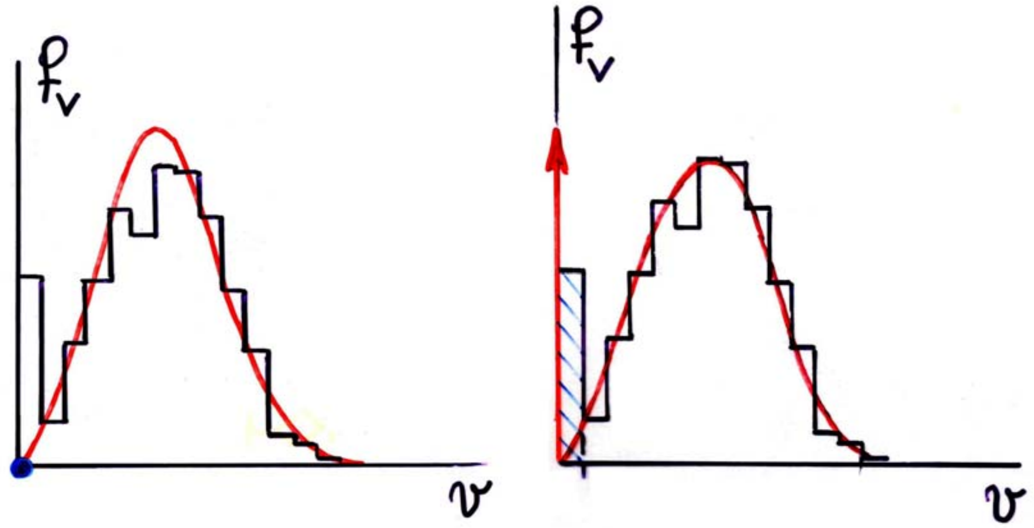

| | c(mph) | k |
|---|---|---|
| Dayton (Ohio) Surface | 8.8 | 1.7 |
| Dayton (Ohio) 1000 ft. | 15.9 | 1.7 |
| Casablanca | 9.3 | 1.5 |
| Falkland Islands | 33.0 | 2.5 |
| Lido da Roma | 15.2 | 1.5 |
| New York (F.J.K. airport – 500M) | 25.5 | 1.87 |
| Seattle airport (300M) | 19.0 | 1.9 |
| Sydney (gradient) | 17.4 | 1.39 |
| Melbourne (gradient) | 23.4 | 1.64 |

## 2.2. Wind calms

The Weibull model involves the condition:

$$F_V(0) = f_V(0) = 0$$

The anemometric records, on the contrary, report more or less long periods of wind calms, this pointing out that V = 0 has a non-nil probability of occurrence. Thus, V is a mixed random variable and shall be modelled by generalized distribution functions.




This circumstance may be taken into account replacing the classical Weibull model by the following hybrid Weibull model:

$$F_V(v) = F_0 + (1 - F_0)\left\{1 - \exp\left[-\left(\frac{v}{c}\right)^k\right]\right\} \quad v \geq 0$$

$$f_V(v) = F_0\delta(v) + (1 - F_0)\frac{k}{c}\left(\frac{v}{c}\right)^{k-1}\exp\left[-\left(\frac{v}{c}\right)^k\right] \quad v \geq 0$$

where $F_0$ is the probability that V = 0 and δ is the Dirac operator. Using this model, the parameters k and c shall be regressed solely on the basis of the values greater than zero. $F_0$ may be estimated as the ratio between the number of the nil values and the total number of the data recorded.

## 2.3 Compound distributions

A suitable way to estimate refined distributions of the parent population is subdividing the full data base into a set of sub-data bases characterized by homogeneous values, estimating the distribution of each sub-data base by a Weibull model, and assembling such models within a so called compound Weibull model. This method is frequently applied by using sub-data bases each one corresponding to a range of wind directions.

Using this method, the probability distribution of the non-nil mean wind velocities coming from the i-th sector (i = 1,2,..m) can be expressed by:

$$F_{V_i}(v) = 1 - \exp\left[-\left(\frac{v}{c_i}\right)^{k_i}\right] \quad v \geq 0;\ i = 1, 2, ..m$$

$$f_{V_i}(v) = \frac{k_i}{c_i}\left(\frac{v}{c_i}\right)^{k_i - 1} \exp\left[-\left(\frac{v}{c_i}\right)^{k_i}\right] \quad v \geq 0;\ i = 1, 2, ..m$$

Accordingly, the compound probability distribution is given by:

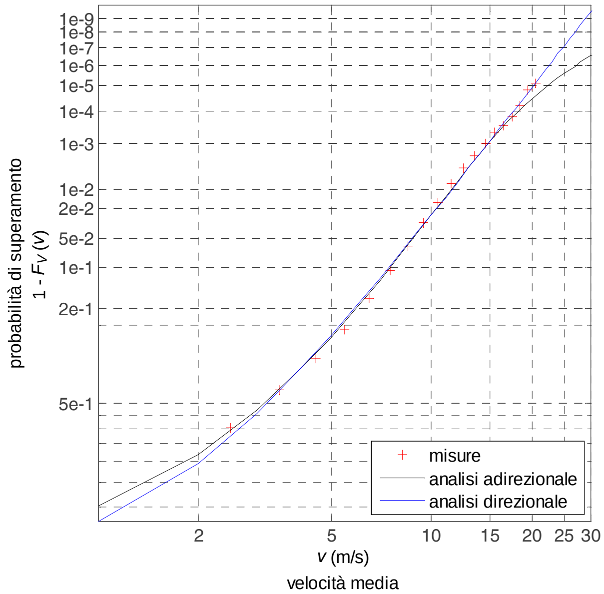
$$F_V(v) = 1 - (1 - F_0)\sum_{i=1}^{m} A_i \exp\left[-\left(\frac{v}{c_i}\right)^{k_i}\right] \quad v \geq 0$$

$$f_{V_{ij}}(v) = F_0\delta(v) + (1 - F_0)\sum_{i=1}^{m} A_i \frac{k_i}{c_i}\left(\frac{v}{c_i}\right)^{k_i - 1} \exp\left[-\left(\frac{v}{c_i}\right)^{k_i}\right] \quad v \geq 0$$

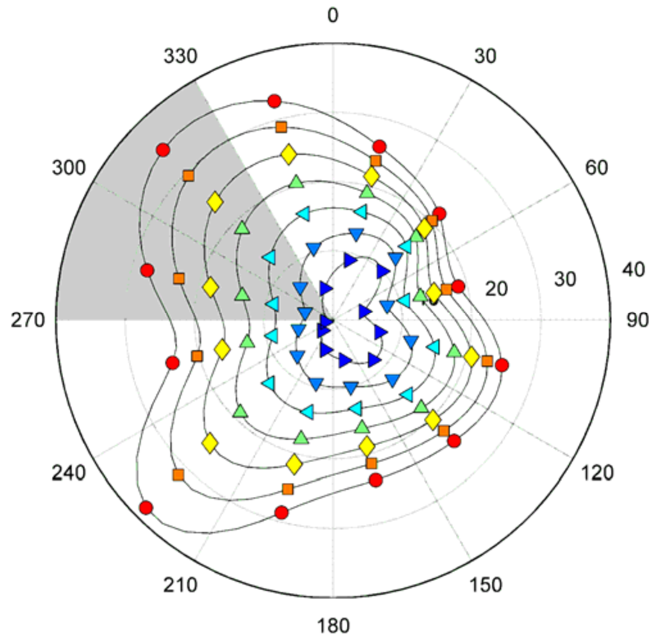
$A_i$ being the probability, conditional to V > 0, that the wind comes from the i-th sector:

$$\sum_{i=1}^{m} A_i = 1$$

The following figure compares the probability of exceedance of v (at the meteorological station of Genova) corresponding to the classical Weibull model and to the compound one.

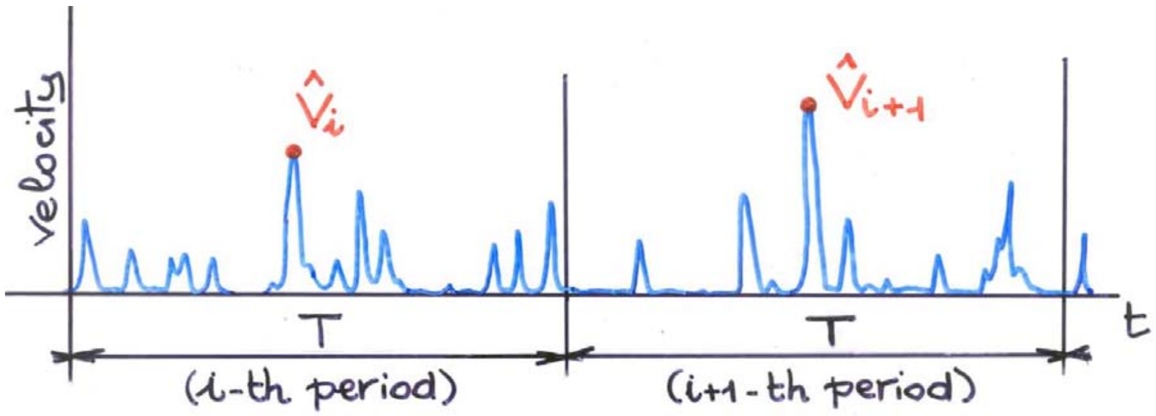



The following figure shows the joint probability of occurrence of the mean wind velocity and its direction. The points of the inner diagram correspond to mean wind velocities coming from sectors 30° wide, with exceedance probability 1%; the outer diagrams correspond to exceedance probabilities 0.1%, 0.01%, 0.001% and 0.0001%.




## 3. Extreme distribution

The following figure depicts a typical variation of the mean wind velocity V in a long time period, and its maximum values $\hat{V}$ in an interval T.




Usually the distribution of the maximum value of V is referred to an interval T = 1 year, and the distribution of the maximum value of V in a period T = N years is given by:

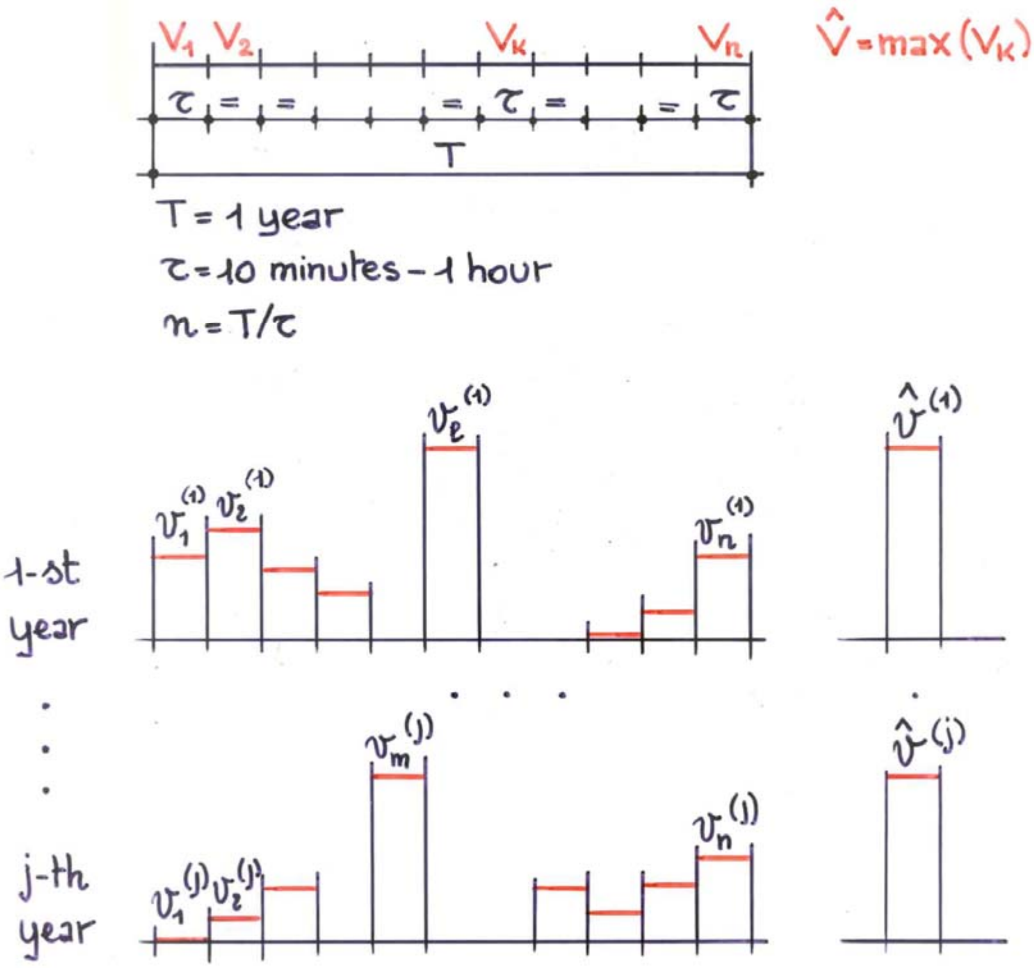
$$F_{\hat{V},N}(v) = \left[F_{\hat{V},1}(v)\right]^N$$

where $F_{\hat{V},N}$ and $F_{\hat{V},1} = F_{\hat{V}}$ are, respectively, the distribution functions of the maximum value of V over N years and over 1 year. The above equation presumes that the yearly maxima are statistically independent.

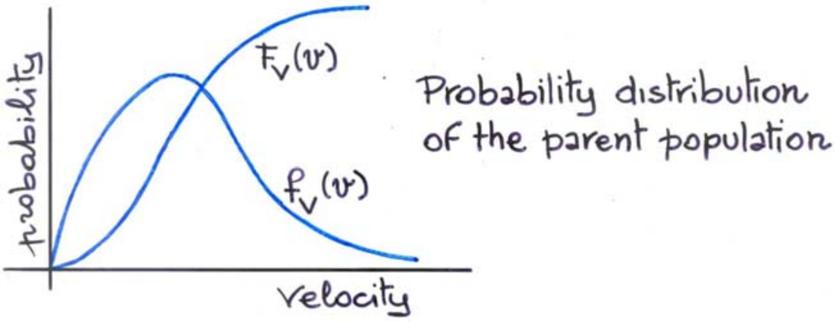
## 3.1 Yearly maximum distribution

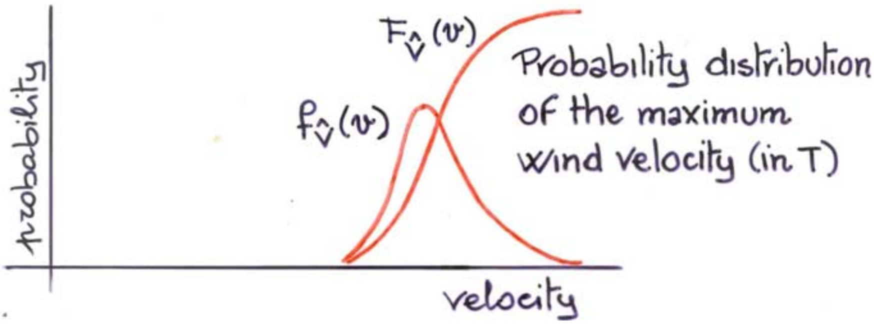
Let us consider n random variables $V_1$, $V_2$, ... $V_n$. $V_k$ corresponds to the mean wind velocity, averaged over $\tau$ = 10 minutes, in the k-th 10 minutes period in the course of one year; n is the number of the 10 minutes periods in one year (n = 365×24×6). $\hat{V}$ = max{$V_1$, $V_2$, ... $V_n$} is the yearly maximum value of the mean wind velocity.

In order to interpret the physical meaning of the random variables $V_1$, $V_2$, ... $V_n$, let us consider the series of the values $v_1^{(j)}, v_2^{(j)}, \ldots v_n^{(j)}$ assumed by $V_1$, $V_2$, ... $V_n$ in the j-th year. $\hat{v}^{(j)} = \max\left\{v_1^{(j)}, v_2^{(j)}, \ldots v_n^{(j)}\right\}$ (j = 1,2,...n) is the series of the maximum values assumed by $\hat{V}$ in the course of the years. The following figure clarifies these concepts.




As the random variable $\hat{V}$ collects the maximum values of the random variables V, its distribution tends to translate towards the largest values of the mean wind velocity.




The evaluation of the distribution of $\hat{V}$ is very easy and expressive, provided that the random variables $V_1$, $V_2$, ... $V_n$ are statistically independent and identically distributed.

With utmost generality, the distribution function of $\hat{V}$ is given by the relationship:

$$F_{\hat{V}}(v) = P\left[\hat{V} \le v\right] = P\left[V_1 \le v, V_2 \le v, \ldots V_n \le v\right]$$

Assuming that $V_1$, $V_2$, ... $V_n$ are statistically independent, it follows that:

$$F_{\hat{V}}(v) = P\left[\hat{V} \le v\right] = P\left[V_1 \le v\right] \cdot P\left[V_2 \le v\right] \cdot \ldots \cdot P\left[V_n \le v\right]$$

Besides, assuming that $V_1$, $V_2$, ... $V_n$ are identically distributed ($F_V = F_{V_i}\ \forall i = 1, 2, \ldots n$), it follows that:

$$F_{\hat{V}}(v) = \left[F_V(v)\right]^n \tag{1}$$

Thus, the probability density function of $\hat{V}$ is given by:

$$f_{\hat{V}}(v) = \frac{dF_{\hat{V}}(v)}{dv} \Rightarrow$$

$$f_{\hat{V}}(v) = n\left[F_V(v)\right]^{n-1} f_V(v) \tag{2}$$

Unfortunately, the random variables $V_1$, $V_2$, ... $V_n$ are neither statistically independent nor identically distributed. They are not statistically independent because the process of the mean wind velocity has a correlation structure that excludes this case. They are not identically distributed because, for instance, the parent distribution of the mean wind velocity changes in the course of the year due to the alternation of the seasons.

These shortcomings may be overcome making recourse to two distinct approaches.

The first consists in expressing the distribution function of the yearly maximum mean wind velocity by the relationship:

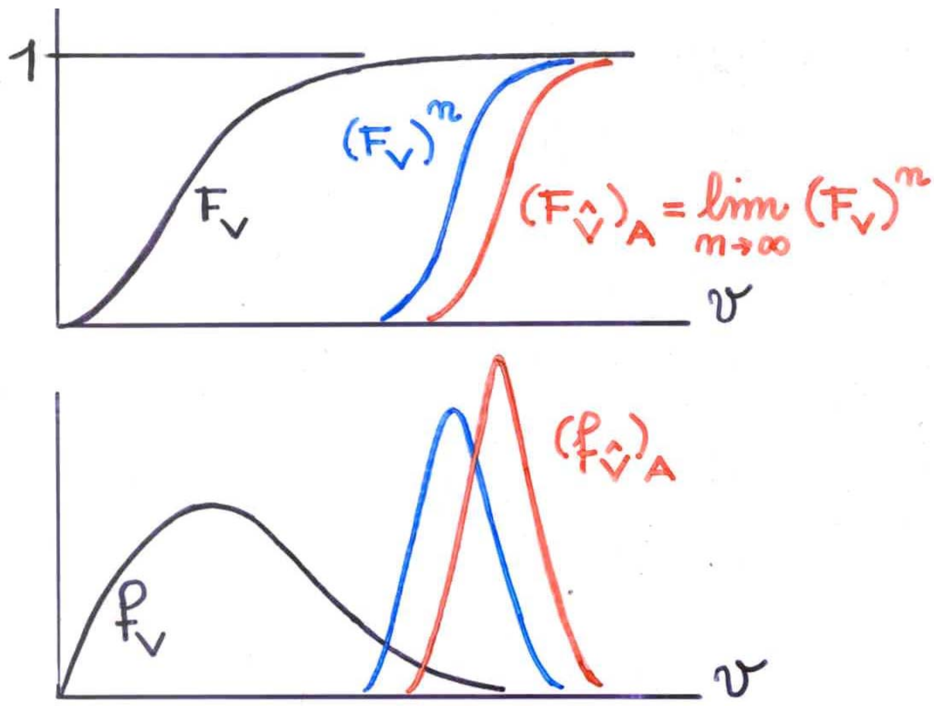
$$F_{\hat{V}}(v) = \left[F_V(v)\right]^{n'}$$

where n’ < n is the number of independent storms in the period T = 1 year. This quantity can be derived, for instance, from the auto-correlation function of the mean wind velocity process.

Due to the uncertainties of this derivation, the second approach consists in the use of the asymptotic distribution functions assuming that n and n’ tends to infinity.

### 3.2 Asymptotic extreme distributions

Fisher and Tippett (1928) demonstrated that, for n tending to infinite, Eqs. (1) and (2) tend to limit distributions, so called asymptotic distributions. The analytical expression of the asymptotic distributions of $\hat{V}$ depends on the shape of the tail of the underlying distribution of V.




Gumbel (1958) demonstrated the existence of three main asymptotic distributions, referred to as the type I, II e III extreme distributions.

Extreme distribution of the type I (Gumbel distribution)

$\hat{V}$ has a type I distribution provided that V has the following two properties:

1) $f_V(v)$ is unlimited in both the directions ($-\infty < v < +\infty$);

2) the tail of the distribution of V towards the maximum values has the form:

$$F_V(v) = 1 - e^{-\rho(v)}$$

where ρ(v) is a growing monotonic function of v. It follows that:

$$F_{\hat{V}}(v) = \exp\left\{-\exp\left[-a(v-u)\right]\right\} \quad (-\infty < v < \infty)$$

$$f_{\hat{V}}(v) = a\exp\left\{-a(v-u)\right\}\exp\left\{-\exp\left[-a(v-u)\right]\right\} \quad (-\infty < v < \infty)$$

In this case the mean value and the variance of $\hat{V}$ are given by:

$$\mu_{\hat{V}} = u + \frac{\gamma}{a} = u + \frac{0,5772}{a}$$

$$\sigma^2_{\hat{V}} = \frac{\pi^2}{6a^2} = \frac{1,6449}{a^2}$$

<u>Extreme distribution of the type II</u> (Fréchet distribution)

$\hat{V}$ has a type II distribution provided that V has the following two properties:

1) $f_V(v)$ is unlimited towards the maximum values ($0 < v < +\infty$);
2) the tail of the distribution of V towards the maximum values has the form:

$$F_V(v) = 1 - \beta\left(\frac{1}{v}\right)^k \quad (v > 0; k > 0)$$

It follows that:

$$F_{\hat{V}}(v) = \exp\left\{-\left(\frac{\mu}{v}\right)^k\right\} \quad (v > 0)$$

$$f_{\hat{V}}(v) = \frac{k}{u}\left(\frac{u}{v}\right)^{k+1}\exp\left\{-\left(\frac{u}{v}\right)^k\right\} \quad (v > 0)$$

In this case the mean value and the variance of $\hat{V}$ are given by:

$$\mu_{\hat{V}} = \mu\Gamma\left(1 - \frac{1}{k}\right) \quad (k > 1)$$

$$\sigma^2_{\hat{V}} = \mu^2\left[\Gamma\left(1 - \frac{2}{k}\right) - \Gamma^2\left(1 - \frac{1}{k}\right)\right] \quad (k > 2)$$

where Γ(x) is the gamma function:

<u>Extreme distribution of the type III</u> (reversed Weibull distribution)

$\hat{V}$ has a type III distribution provided that V has the following two properties:

1) $f_V(v)$ is limited towards the maximum ($v < w$);
2) in proximity of w the distribution of V assumes the form:

$$F_V(v) = 1 - c(w-v)^{\gamma} \quad (v < w, \gamma > 0)$$

It follows that:

$$F_{\hat{V}}(v) = \exp\left\{-\left(\frac{w-v}{w-\eta}\right)^{\gamma}\right\} \quad (v < w)$$

$$p_{\hat{V}}(v) = \frac{\gamma}{w-\eta}\left(\frac{w-v}{w-\eta}\right)^{\gamma-1} \exp\left\{-\left(\frac{w-v}{w-\eta}\right)^{\gamma}\right\} \quad (v < w)$$

In this case the mean value and the variance of $\hat{V}$ are given by:

$$\mu_V = w - (w-\eta)\Gamma\left(1 - \frac{1}{\gamma}\right)$$

$$\sigma_{\hat{V}}^2 = (w-\eta)^2\left[\Gamma\left(1+\frac{2}{\gamma}\right) - \Gamma^2\left(1-\frac{1}{\gamma}\right)\right]$$

<u>Mean return period</u>

The (mean) return period R of v is the mean number of years that elapses between two events with a mean wind velocity greater or equal to v. This quantity is given by the relationships:

$$R(v) = \frac{1}{p(v)} = \frac{1}{1 - F_{\hat{V}}(v)}$$

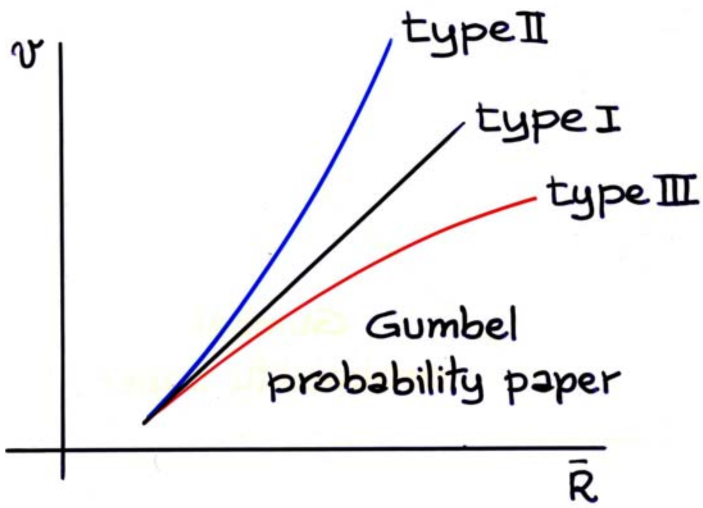
where p is the probability that the yearly maximum value of the mean wind velocity is greater than v. Adopting this property, the type I, II and III distributions of the maximum can be rewritten as:

$$v = u - \frac{1}{a}\ln\left[-\ln\left(1 - \frac{1}{R}\right)\right]$$

$$v = \mu\left[-\ln\left(1 - \frac{1}{R}\right)\right]^{-1/k}$$

$$v = w - (w-\eta)\left[\ln\left(1 - \frac{1}{R}\right)\right]^{1/\gamma}$$

Drawing a diagram v(R) whose coordinates are such as that v is a linear function of R under the condition that $\hat{V}$ has a distribution of the type I, then v(R) has a concavity upwards or downwards, under the condition that $\hat{V}$ has a distribution of the type II or III, respectively.




A wide scientific and technical debate has been developed over the years in order to identify the best asymptotic distribution that represents the yearly maximum value of the mean wind velocity.

Assuming that the parent distribution of the mean wind velocity V is represented by the Weibull model, no couple of the conditions for which $\hat{V}$ has a distribution of the type I, II and III is satisfied. However, V satisfies at least the second condition for which $\hat{V}$ has a distribution of the type I. As this distribution is also the easiest to use, especially in the technical and codification sectors wind engineering usually adopts this model. In spite of this, several authors recommended the use of the distribution of the type III. Few are in favor of the distribution of the type II.

### 3.3 Generalized extreme distribution

The three distributions of the type I, II and III may be expressed by means of a unique model provided by the relationship (Jenkinson, 1955):

$$F_{\hat{V}}(v) = \exp\left\{-\left[1-\beta\alpha(v-\eta)\right]^{1/\beta}\right\} \qquad (\beta \neq 0)$$

$$F_{\hat{V}}(v) = \exp\left\{-\exp\left[\alpha(v-\eta)\right]\right\} \qquad (\beta = 0)$$

In this case the mean value and the variance of $\hat{V}$ are given by:

$$\mu_{\hat{V}} = \eta + \frac{1}{\alpha\beta}\left[1-\Gamma(1+\beta)\right], \quad \beta > -1$$

$$\sigma^2_{\hat{V}} = \left(\frac{1}{\alpha\beta}\right)^2\left[\Gamma(1+2\beta)-\Gamma^2(1+\beta)\right], \quad \beta > -1/2$$

Jenkinson demonstrated that:

- for $\beta$ tending to 0, the generalized extreme distribution tends to the distribution of the type I;

- for β < 0, the generalized extreme distribution tends to the distribution of the type II, provided that v > (η + 1/αβ);
- for β > 0, the generalized extreme distribution tends to the distribution of the type III, provided that v < (η + 1/αβ).

### 3.4 Penultimate distribution

The asymptotic distributions are often applied in spite of the fact that n' is too low to meet the basic assumption that it tends to infinity. The type I penultimate distribution (Cook and Harris 2004, 2008) overcomes this limitation by taking into account the actual finite value of n'. It assumes the following easy and elegant shape when the extremes are drawn from a tail-equivalent Weibull parent distribution:

$$F_{\hat{V}}(v) = \exp\left\{-\exp\left[\alpha^k\left(v^k - \eta^k\right)\right]\right\}$$

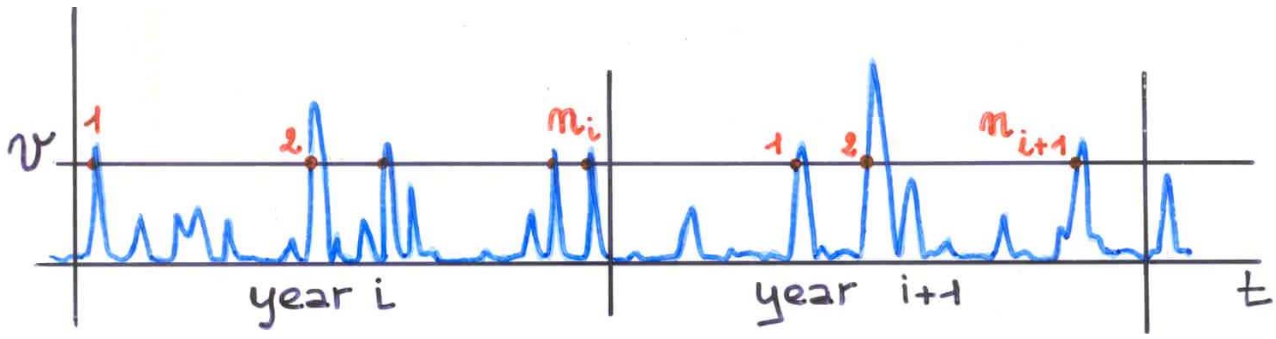
where α = 1/c; k and c are the parameters of the Weibull distribution. If the Weibull distribution of the population is known, only η should be estimated from extremes.

### 3.5 Process analysis

Process analysis deals with the mean wind velocity as a random stationary process. The mean number of the up-crossings of the v threshold in the unit of time (in this case one year) is provided by the relationship (Rice 1944, 1945):

$$\bar{N}(v) = \int_0^\infty \dot{v} f_{V\dot{V}}(v,\dot{v})\, d\dot{v}$$

where $\dot{V}(t)$ is the first derivative of the process $V(t)$; $f_{V\dot{V}}(v,\dot{v})$ is the joint density function of $V(t)$ and $\dot{V}(t)$.




Assuming that $V(t)$ and $\dot{V}(t)$ are statistically independent, then:

$$f_{V\dot{V}}(v,\dot{v}) = f_V(v)\cdot f_{\dot{V}}(\dot{v}) \Rightarrow$$

$$\bar{N}(v) = \int_0^\infty \dot{v} f_{V\dot{V}}(v,\dot{v})\, d\dot{v} = \int_0^\infty \dot{v} f_V(v) f_{\dot{V}}(\dot{v})\, d\dot{v} = f_V(v)\int_0^\infty \dot{v} f_{\dot{V}}(\dot{v})\, d\dot{v} \Rightarrow$$

$$\bar{N}(v) = \lambda f_V(v)$$

where:

$$\lambda = \int_0^\infty \dot{v} f_{\dot{V}}(\dot{v}) d\dot{v}$$

It is worth noting that $V(t)$ and $\dot{V}(t)$ are uncorrelated random processes independently of their distribution; as such, they are statistically independent provided that $V(t)$ is also Gaussian. Unfortunately, $V(t)$ is a Weibull process and $V(t)$ and $\dot{V}(t)$ cannot be independent. Thus, the above treatment is approximated; experience shows that such approximation is very good.

Assuming that the v threshold is sufficiently high, its up-crossings may be considered as rare and independent events; as such they constitute a Poissonian process. Thus, the probability that during one year the threshold v is up-crossed n times results:

$$P(n,v) = e^{-\bar{N}(v)} \frac{\left[\bar{N}(v)\right]^n}{n!}$$

It follows that the probability that during one year v is up-crossed zero times results:

$$P(0,v) = e^{-\bar{N}(v)}$$

Such probability identifies with the distribution function of the maximum yearly mean wind velocity:

$$F_{\hat{V}}(v) = P(0,v) = e^{-\bar{N}(v)} \Rightarrow$$

$$F_{\hat{V}}(v) = e^{-\lambda f_V(v)}$$

This expression establishes a link between the extreme and the parent distribution.

### 3.6 <u>Peak over threshold (POT) method</u>

Let us consider a variable V with a parent distribution $F_V(v)$ such that, for large n', the distribution of the largest value in the period T = 1 year converges towards an asymptotic distribution. Pickands (1975) demonstrated that the excess of such variable over the threshold u, X = V – u, tends to the Generalized Pareto Distribution (GPD):

$$F_X(x) = 1 - \left[1 - \beta \frac{x}{\delta}\right]^{1/\beta} \quad (\beta \neq 0)$$

$$F_X(x) = 1 - \exp\left(-\frac{x}{\delta}\right) \quad (\beta = 0)$$

provided that u is arbitrarily large; β is the shape factor previously introduced for the GEV distribution.

If the threshold u is large enough, its crossings can be assumed to be independent, and the number m of values over u in the period T = 1 year has a Poisson distribution with a rate parameter $\lambda_u$. It follows that the distribution of the largest value of V is given by:

$$F_{\hat{V}}(v) = \exp\left\{ -\lambda_u \left[ 1 - \beta \frac{v-u}{\delta} \right]^{1/\beta} \right\} \qquad (\beta \neq 0)$$

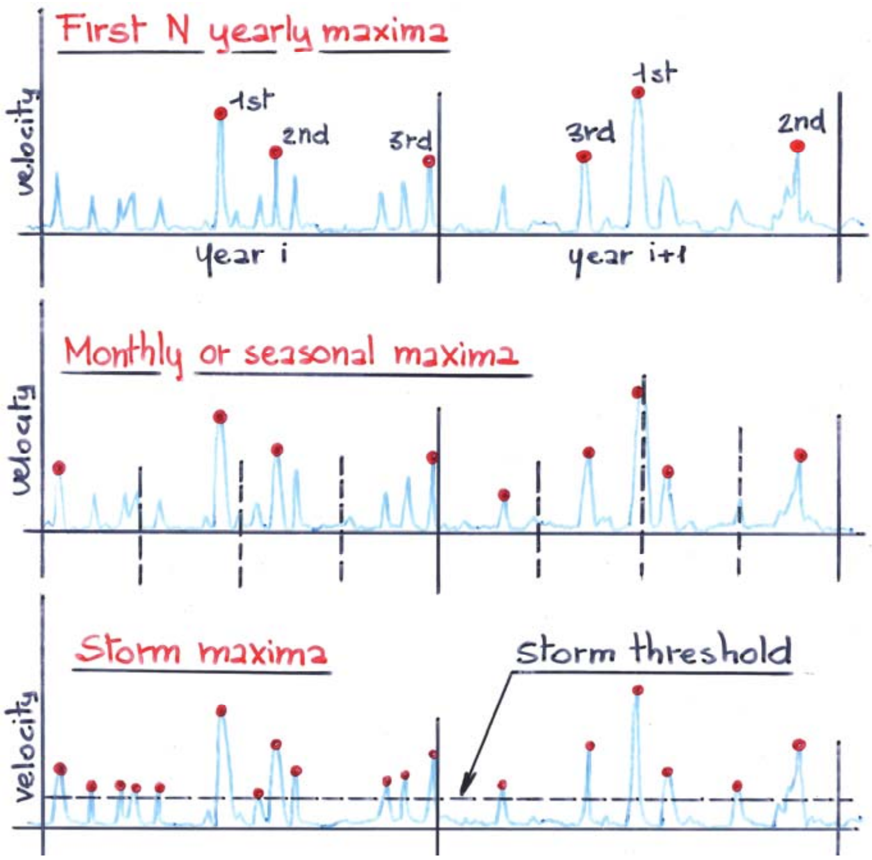
$$F_{\hat{V}}(v) = \exp\left\{ -\lambda_u \exp\left[ -\frac{v-u}{\delta} \right] \right\} \qquad (\beta = 0)$$

A reliable estimator of $\lambda_u$ is the total number m of exceedance of u in T = 1 year.

### 3.7 Enlarging the dataset of yearly maxima

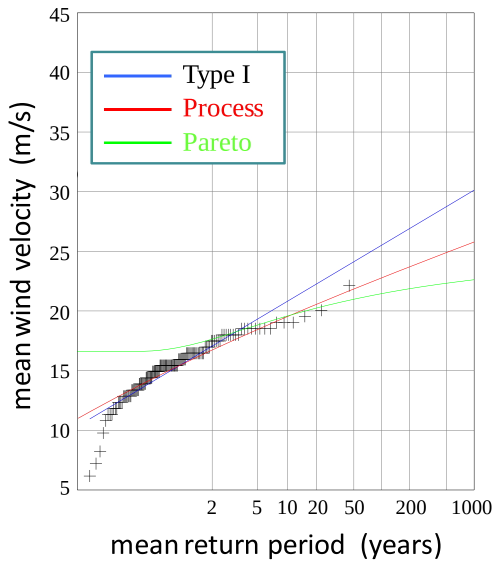
Using asymptotic distributions it is usual to regress the extreme values by considering only the yearly maxima. This approach has the shortcoming that only a limited number of information is used and extremes higher than yearly maxima are often neglected.

Several methods have been proposed to overcome this shortcoming. The r-Los method selects the r largest order statistics in T = 1 year. The method of Independent Storms (MIS) (Cook, 1982) increases the number of the data available for the regression of the extreme distribution considering the highest value of each independent storm; the collection of independent events is based on the assumption of suitable thresholds to identify the storm and given breaks for the event separation, according to the climatic features and to the typology of data available. This method has been subsequently improved by Harris (1999), becoming known as the IMIS.

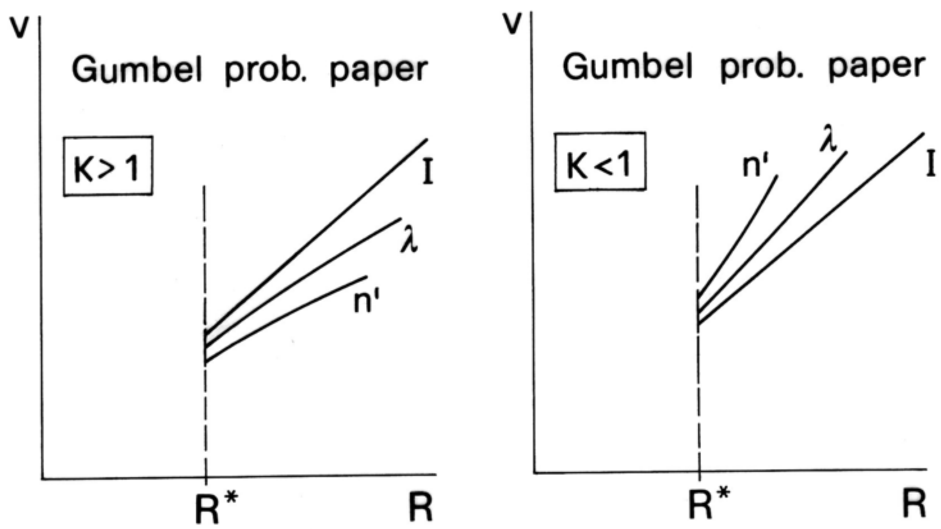


### 3.8 Comparison between different distributions

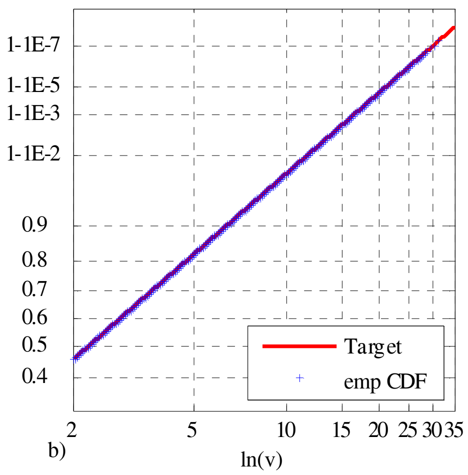
The comparison between different models of the extreme distribution of the mean wind velocity points out a relevant aspect. When the return period R is less than the number of the years for which measured data is available, different models provide averagely coherent results. On the other hand, when R is greater than the number of years for which measured data is available, different models tend to diverge. In this case, the lacking of enough data precludes understanding what the best model is; it follows that such understanding is possible only by means of theoretical evaluations with reference to which there is yet no agreement.

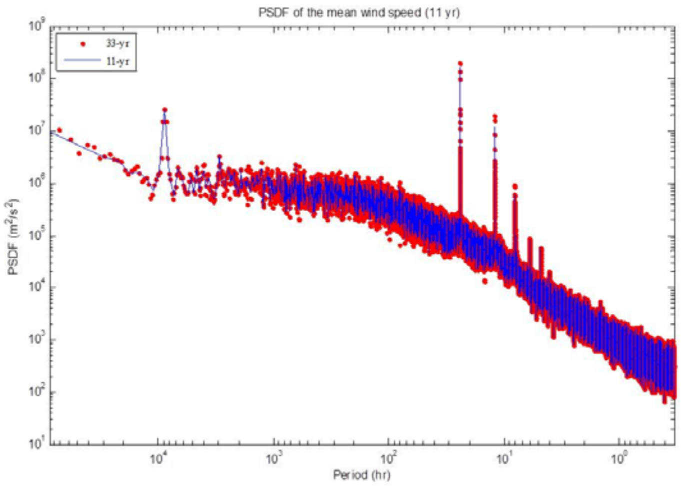


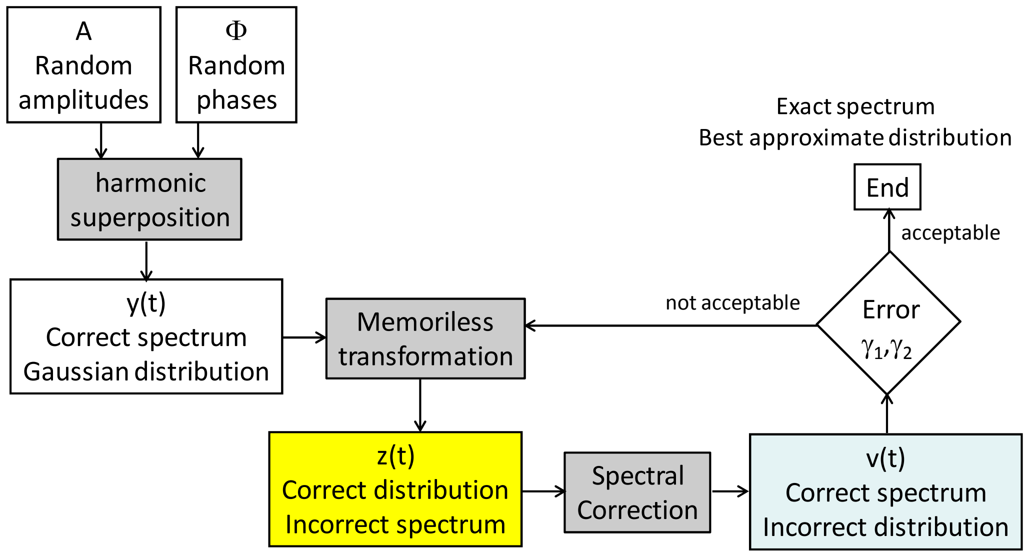

In spite of this, some general considerations are shown by the following figures.

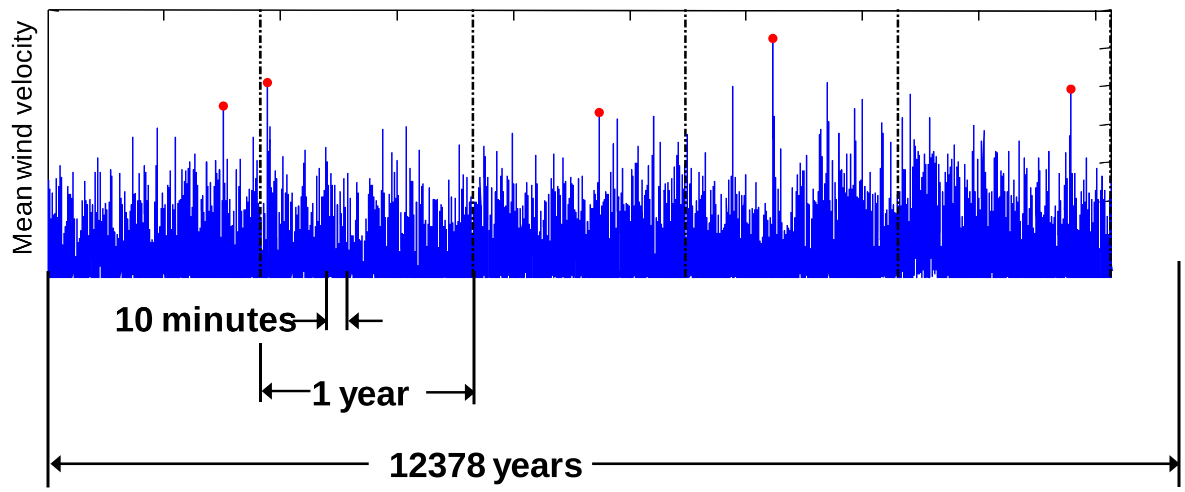



A promising technique to clarify the above doubts is represented by the Monte Carlo simulation of long term mean wind velocity histories (Torrielli et al 2011, 2012). The following figures show, in their order, the target density function and power spectral of the mean wind velocity to simulate, the iterative algorithm used for the simulation, a simulated mean wind velocity history lasting for over 10.000 years.

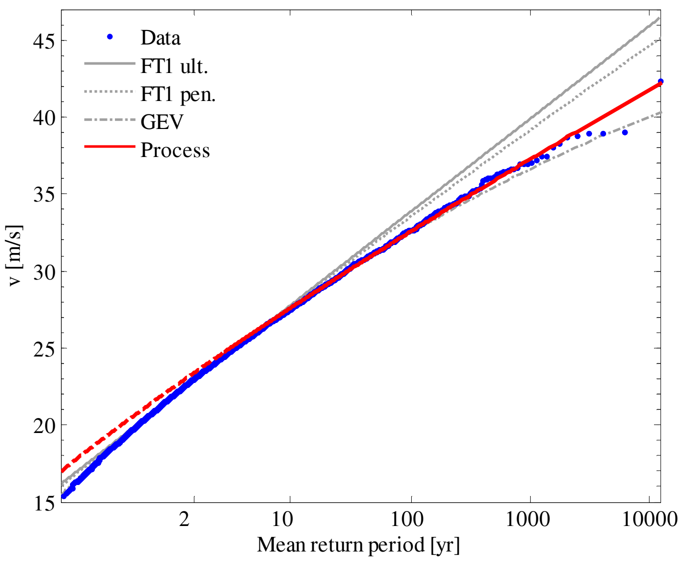



The following figure shows a comparison between different extreme distributions, reporting also the simulated yearly maximum values. The comparison, extended up to R = 10.000 years, points out that the process analysis and the penultimate distribution provide the best representation.

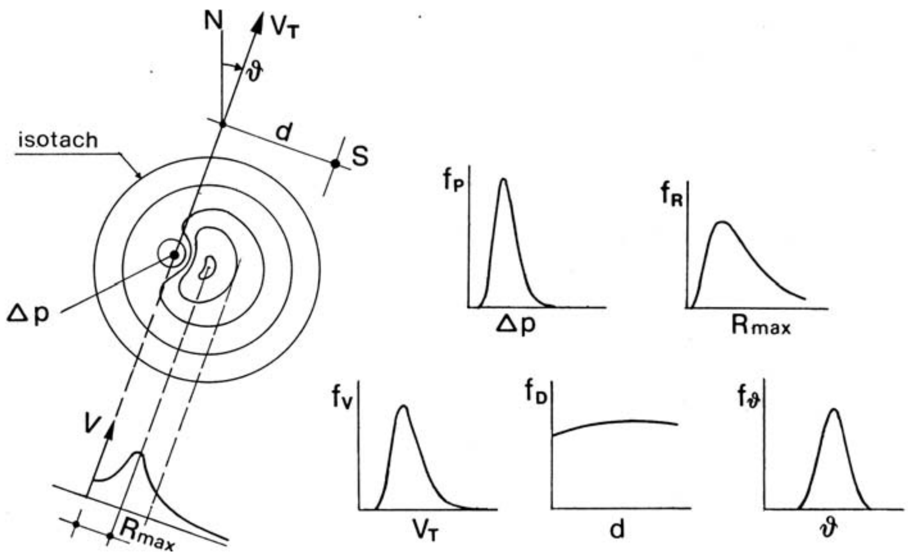



## 3.9 Extreme distributions for hurricanes

Hurricanes rarely strike a meteorological station. Thus, the available data concerning hurricanes are not enough to apply the above models. This shortcoming is usually overcome by applying the following 4-steps approach:

1. The geographical areas into which hurricanes develop capable of striking the site considered are identified;

2. A mathematical model based on physical parameters governing hurricanes is set up. The model parameters are characterized by probability distributions gained from the data gathered in the areas under examination;




3. Applying the Monte Carlo technique, a historic series of hurricanes is simulated such as to produce enough data at the considered site;

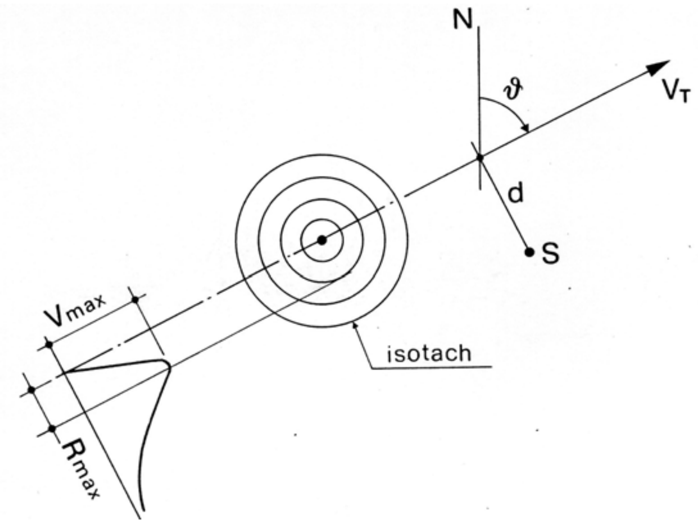
4. The simulated data are analyzed statistically in order to determine the distribution function of the maximum wind velocity at the site due to hurricanes.

## 3.10 Extreme distributions for tornadoes

The probabilistic analysis of the extreme wind velocity due to tornadoes may be carried out, in principle, paraphrasing the method illustrated for hurricanes. This results de facto of little advantage, since the area struck by an individual tornado is so limited as to make the generation time period tendentiously infinite as to produce enough data. The following 3-steps procedure is consequently usually applied:

1. The territory under consideration is subdivided into M regions characterized by a homogeneous occurrence of tornadoes;

2. Treating the tornado as a Poissonian event, and utilizing the information gathered, the probability $F^*(v)$ is estimated that the maximum velocity due to tornadoes does not exceed the v threshold at a given point of the region containing the examined site;




3. The distribution function of the maximum wind velocity due to tornadoes is given by the product of $F^*$ by the probability that a tornado may strike the considered site.

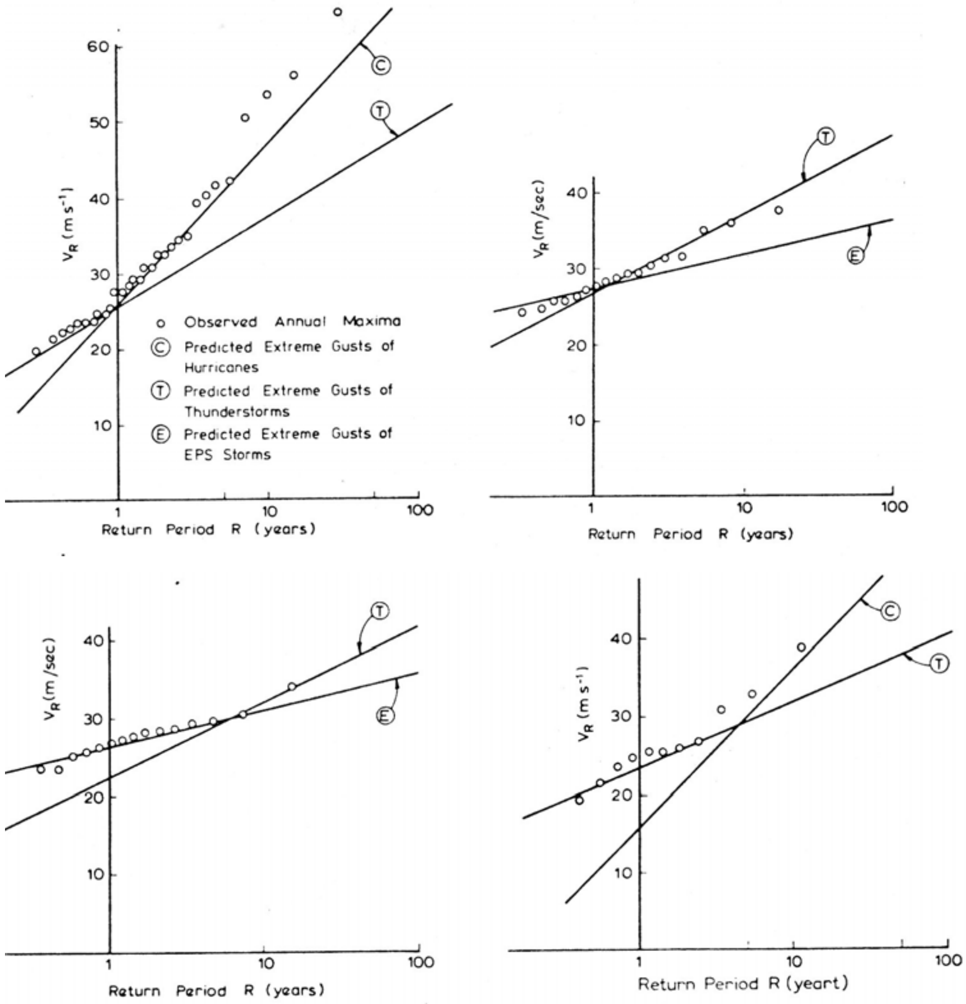
## 3.11 Mixed extreme distributions

The coexistence of different wind events in the same area, e.g. extra-tropical cyclones, downbursts and tornadoes, requires a 2-steps approach. First, each event should be analyzed separately with the aim of obtaining its own extreme distribution. Then, these distributions should be composed in a unique comprehensive model, referred to as a mixed extreme distribution.

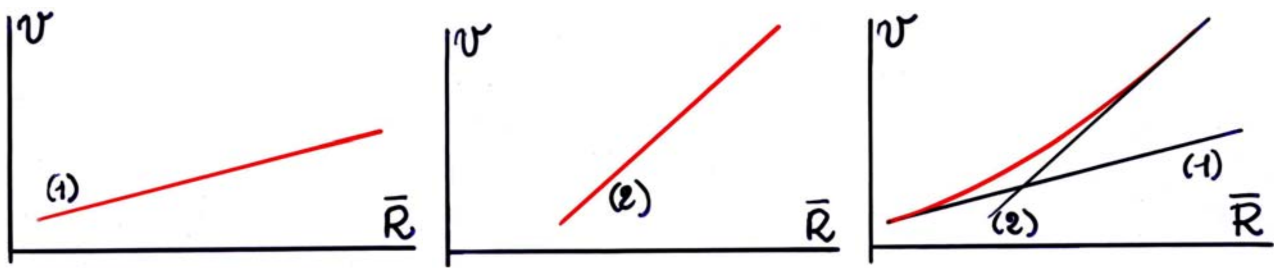
Let us consider the distribution functions $F_{\hat{V}}^{(k)}(v)$ (k = 1,2,..L) of the yearly maximum wind velocity associated with L different Aeolian phenomena. Assuming that these phenomena are statistically independent of each other, the distribution function of the yearly maximum mean wind velocity is given by:

$$F_{\hat{V}}(v) = \prod_{k=1}^{L} F_{\hat{V}}^{(k)}(v)$$

The following figure shows some examples of actual situations referred to different parts of the world.




The following figure shows the result of applying a mixed extreme distribution, using coordinated axes such as to render linear the functions $v^{(k)}(R)$ (k = 1, 2, ... L); this is possible provided that all the distribution functions $F_{\hat{V}}^{(k)}(v)$ are expressed by the same probabilistic model.




The first and the second scheme show the case in which L = 1. The third scheme shows what it happens for L = 2. The more an event is rare and intense, the more its diagram increases its gradient and translates towards the higher values of the mean return period.

# 4. Wind climate

The probability distributions obtained by means of the methods described above are strictly representative of the site where the anemometer is placed or the simulations are carried out. In order to extend such representativeness to other sites, the recorded data base shall be transformed into simulated data bases corresponding to the sites of interest, inside a suitable neighbourhood of the anemometer itself. Such transformation can be carried out through an intermediate transformation of the mean wind velocity at the gradient height. The probability distributions of the simulated data bases are then determined using the same methods described above.

The result of this analysis may be represented by the parent and extreme distributions of the mean wind velocity at a reference site (e.g. the mean wind velocity recorded by an ideal anemometer put at 10 m height in a flat, open and homogeneous terrain), in another specific point (e.g. at the site of a structure to analyze), or in a set of points (such as to realize a so called wind map).

These evaluations are usually referred to as wind climate analyses.