The statistically independent parent distribution of wind speeds derived by run-of-wind resampling

Nicholas Cook*, Consultant, RWDI*

Wind speeds sampled on an hourly, or 3-hourly basis are strongly correlated between adjacent observations. Harris [1] shows that this serial correlation has an integral time scale of T = 22.1 h at Boscombe Down, UK. It is often assumed that observations separated by more than 3T are statistically independent. This is the approach used by Simiu & Heckert [2] who used 4-day extremes to deduce the distribution of annual maxima. The 1982 Method of Independent Storms (MIS) [3] identifies individual synoptic events, or "storms", and extracts the maximum wind speed from each. When the events are statistically independent, the distribution of annual maxima, Φ, is related to the distribution of storm maxima, P_S, by the fundamental statistical relationship: $\Phi = P_S^{\ r}$, where r is the annual rate of events. It is often claimed that the distribution of annual maxima can be derived from the distribution of the correlated hourly parent, P, above some suitable threshold using the annual rate of *independent* events, r_i, which is less than the rate of all events. But r_i will be a constant only when the degree of correlation is the same for all wind speeds, otherwise r_i will be **wind-speed dependent**.

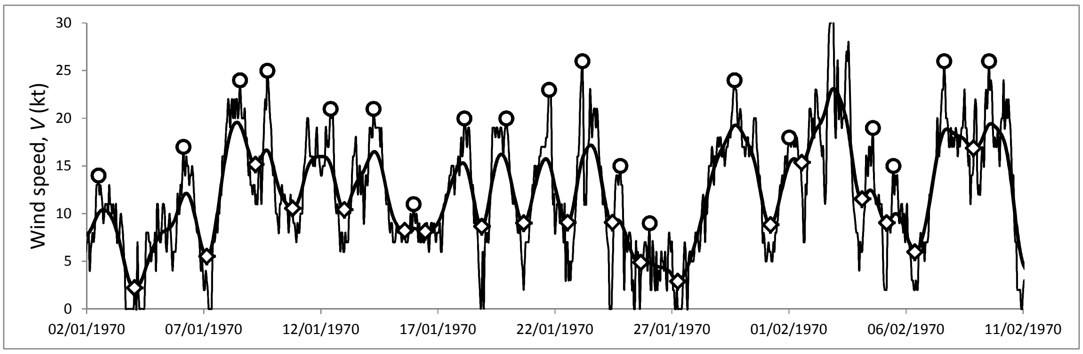


Figure 1. Method of independent storms operating with 22 h constant time scale filter

Figure 1 shows eight days of mean wind speeds from Boscombe Down using T = 22.1 h, the value for the integral time scale found by Harris [1]. An autoregressive digital filter of the form: $y_n = (1-a)\, x_n + a\, y_{n-1}$ is used. This is analogous to a single-pole R-C filter for $a = \exp(-2\pi\Delta t/T)$, where Δt is the time between samples (1 h) and T is the time constant of the filter. Applying this filter six times, alternatively *forwards* and *backwards* though the data, gives a sharp low pass filter *with no phase lags*. The thin line is the hourly mean data and the thicker curve is the filtered data. The diamond symbols mark the minima of the filtered data which the method uses to denote the start of a new "storm". The circle symbols show the "storm" maxima. Note that here, and elsewhere in the paper, $Q = 1 - P$, is the probability of exceedence.

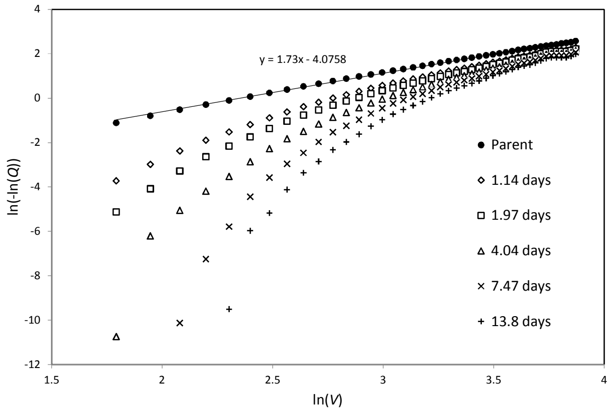


Figure 2. CDF of *n*-day maxima on Weibull axes.

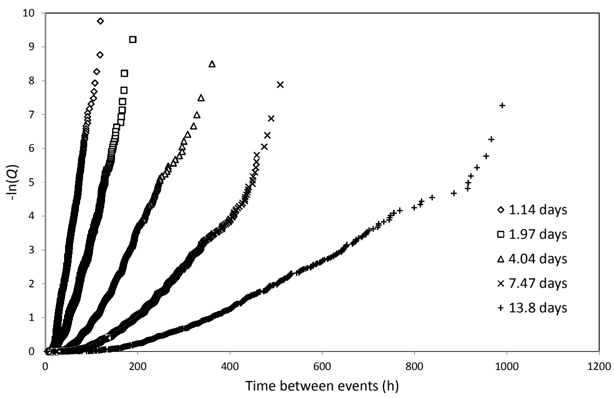


Figure 3. CDF of time between storm maxima

When MIS is used to replicate the Simiu and Heckert [2] n-day maxima, we find that $n \approx 2T$ because there must be a minimum between each maximum. Figure 2 shows the CDFs for the hourly parent and the n-day maxima plotted on Weibull axes – the hourly parent is a good fit to shape factor w = 1.73, and the n-day maxima diverge progressively towards the FT1 asymptote. Figure 3 shows the corresponding CDFs of time between storm maxima plotted on axes that give a straight line with slope T for a Poisson process. The distributions are curved, indicating the time scale is not constant and n-day maxima of wind speeds do not conform to a Poisson process.

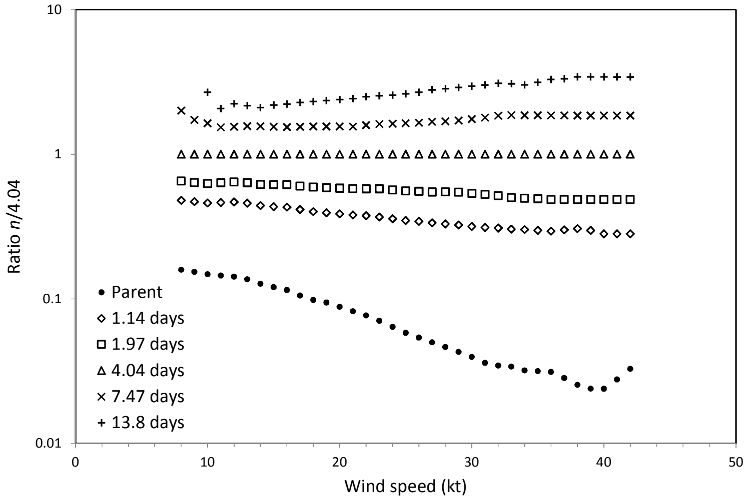


Figure 4. Ratio $r = n/4.04$ for each wind speed

When n-day maxima are independent, the CDFs for different n are related by : $P_2 = P_1^{\,r}$ where $r = n_2/n_1$, i.e. by the ratio of their epochs. Using the 4-day maxima as datum, Figure 4 shows the values of r for each integer wind speed in the observed range at Boscombe Down. It is apparent that r is not constant for the hourly parent and that the serial correlation between hourly observations is stronger at lower wind speeds. For statistical independence, the values for each of the n-day maxima should be constant and be a factor of ~2 apart. Instead, the values converge together at lower wind speeds, indicating a residual statistical dependence remains in the n-day maxima.

An important finding from Harris' study [1] is that the Macrometeorological spectrum of wind speeds exhibits a −5/3 power law decay range in the same manner as the Micrometeorological spectrum and rough-wall boundary layers, which are length-scale dependent. This implies that the Macrometeorological spectrum is also length scale dependent and that the observed wind speed dependence of r is caused by the same correlated "patch" of wind being sampled more frequently when it is advected past the observer at low wind speeds than at high wind speeds. It follows that serial correlation may be equalised through the wind speed range by resampling the observations at equal intervals of run-of-wind. A fixed time constant T gives a fixed coefficient a in the recursive filter – but varying the value of T in inverse proportion to the wind speed gives a value a corresponding to a fixed run-of-wind. Figure 5 shows the same eight days of wind speeds as Figure 1, but filtered at a fixed run-of-wind that corresponds to a 22h period at the overall mean wind speed. The constant run-of-wind results in filtered data that have flatter, U-shaped minima and sharper maxima. This also gives storm maxima which are 4 days apart, on average, but is seen to produce fewer maxima at low wind speeds and more maxima at higher wind speeds.

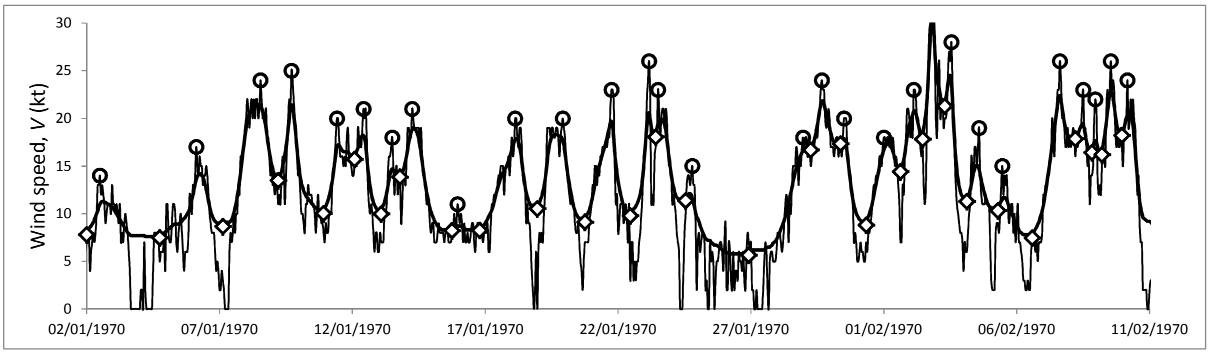
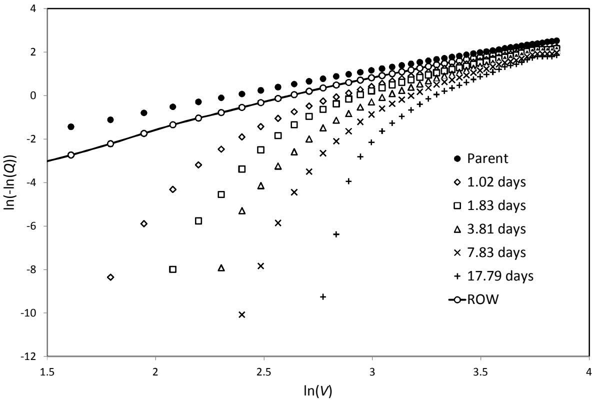


Figure 5. Method of independent storms operating with 22-h average constant run-of-wind filter



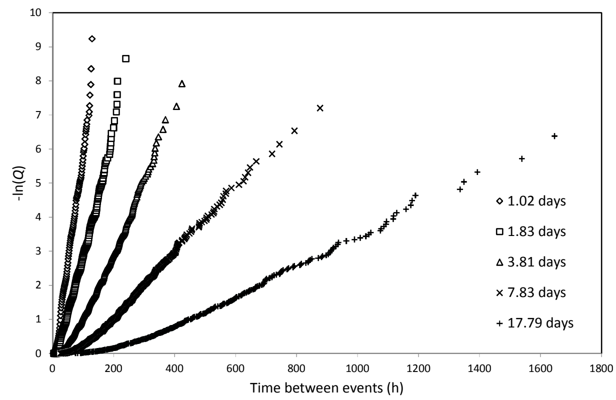


Figure 6. CDF of ROW maxima on Weibull axes.

Figure 7. CDF of time between ROW maxima

Figure 6 shows the CDFs of the hourly parent and *n*-day run-of-wind maxima on Weibull axes, corresponding to Figure 2. By reducing the serial correlation at low wind speeds, the ROW resampled *n*-day maxima diverge towards the FT1 asymptote more quickly. The additional curve, marked “ROW” is the corresponding ROW parent CDF of wind speeds, derived later. Figure 7 shows the CDF of time between ROW maxima on Poisson axes, corresponding to Figure 3 – the distributions are straightened, indicating that a consistent time scale has been obtained.

Figure 8. Ratio *r* = *n*/3.81 for each wind speed, run-of-wind resampled

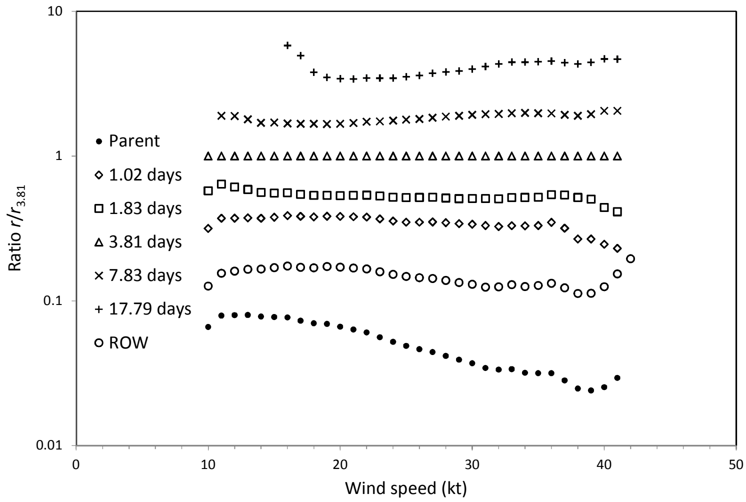


Figure 8 shows value of *r* for each integer wind speed, corresponding to Figure 4 after run-of-wind resampling. There are small differences in the equivalent values of *n*, here based on the overall mean wind speed. For the *n*-day maxima, *r* is much more consistent within the expected experimental variation, and no longer converges at low wind speeds. Values for the hourly parent are the same as Figure 4, except normalised by the new datum *n* = 3.81. The additional set of data, marked “ROW”, are the values for the ROW parent mean wind speeds, described later. Note that these values lie horizontal and parallel to the values for the *n*-day maxima, as is required if statistically independent.

Serial correlation in the hourly parent contributes too many observations at lower wind speeds, compared with higher wind speeds, in the frequency table for the integer wind speed ranges. This is corrected by factoring the count in each wind speed bin by V/L, where V is the wind speed of the bin and L is the desired run-of-wind. For example: if the wind speed is V = 10 kt and the run of wind is L = 100 NM, the same run of wind is sampled 10 times by hourly data, so the count must be factored by 10/100 = 0.1 for each run-of-wind to contribute just once. The **parent distribution of statistically independent wind speeds** is therefore obtained by posing the question “what run-of-wind is required to relate the CDF of this parent to the CDF of the datum *n*-day maxima using a single constant value of *r*?” A good starting point for L is given by the product of the overall mean wind seed and Harris’ value for the integral time scale: for Boscombe Down $\overline{V}$ = 9.3 kt and T = 22.1 h, giving L = 205 NM. The optimum value of L was obtained by minimising the error in ln (-ln(Q)), i.e. in the fit on Weibull axes, using the Excel “Solver” non-linear optimiser. For Boscombe Down, this yields L = 143.8 NM, corresponding to an annual rate of independent events r_i = 572. The “ROW” data in Figures 6 and 8 are the results of this optimisation.

Figure 9. The parent distribution of statistically dependent wind speeds for Boscombe Down, UK, on Weibull axes

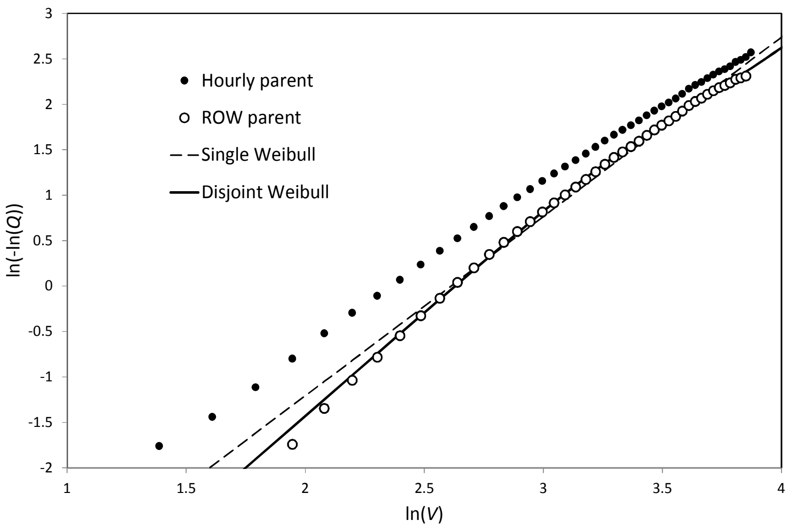


Figure 9 shows the CFD of this “ROW” parent plotted on standard Weibull axes, with the hourly parent for comparison. The ROW data exhibit more of a curve than does the hourly parent. The corresponding straight-line fit gives Weibull parameters w = 1.97 and C = 13.6 kt.

Cook and Harris [4] previously found that the observations for Boscombe Down fitted a disjoint Weibull model when the wind speeds has been separated synoptically into cyclonic and anti-cyclonic sets using the Jenkinson-Lamb index of weather types. Here the disjoint fit was made without prior separation and is shown by the solid line curve in Figure 9, corresponding to the parameters in Table 1, below. Mechanism 1 is dominant in the upper tail and the values correspond well with the previous analysis of the cyclonic component [4]. Mechanism 2 is dominant in the lower tail, provides the curve in the body of the distribution, and acts as a guarding degree of freedom in the fit for Mechanism 1.

Table 1. Weibull parameters for disjoint fit to statistically independent parent wind speeds at Boscombe Down

Parameter	Mechanism 1 (Cyclonic)	Mechanism 2 (Anti-cyclonic)
Disjoint frequency, *f*	0.229	0.771
Shape parameter, *w*	2.00	2.35
Scale parameter, *C*	15.59 kt	13.46 kt
Annual rate, *r*	131.1	440.8

The relevance of the parent distribution of statistically independent wind speeds is that it permits the direct estimation of extremes, including the distribution of annual maxima, **without recourse to any extreme value theory**, because the annual rate of each mechanism is known. The fundamental relationship $\Phi = P^r$ can be used directly, or else the Poisson model is appropriate, $\Phi = \exp(-rQ)$, because the rate parameter is large. Figure 10 shows the CDF of the annual maximum mean wind speed for Boscombe Down, comparing the observations with the two Weibull fits.

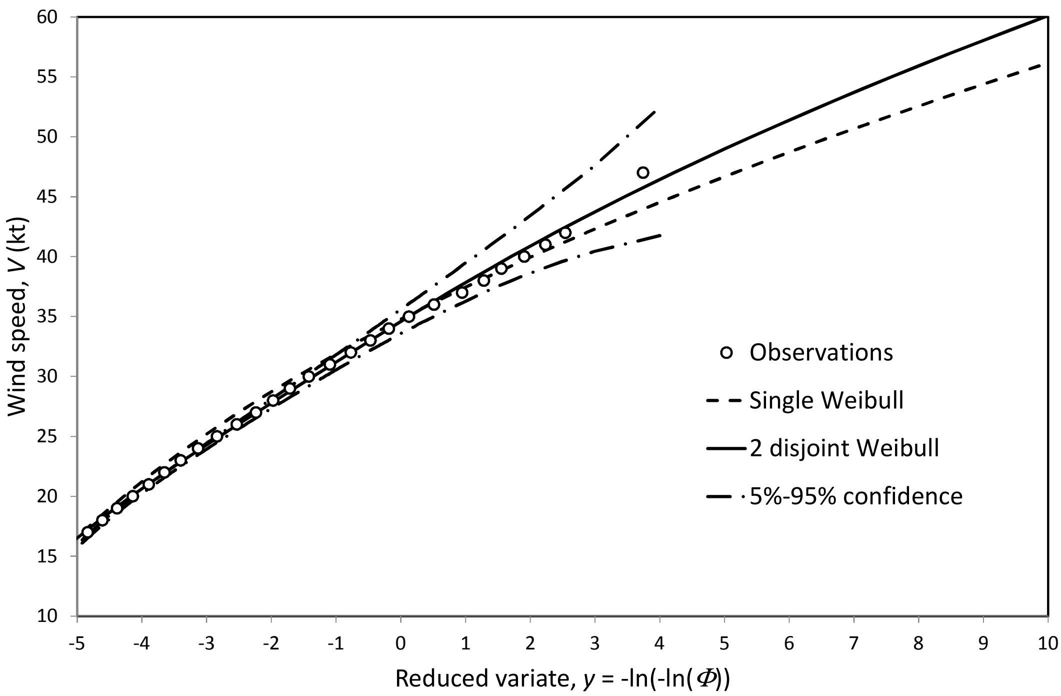


Figure 10. CDF of annual maximum wind speed for Boscombe Down directly from the statistically independent parent

References

1. Harris R.I.: The Macrometeorological spectrum – A preliminary study, JWEIA, (2008) **96**, 2294-2307.
2. Simiu E, Heckert NA: Extreme wind distribution tails: a "peaks over threshold" approach, JASCE, (1996) 539 – 547.
3. Cook NJ: Towards better estimation of extreme winds, JWEIA, (1982) **9**, 295 – 323.
4. Cook NJ, Harris RI: Postscript to "Exact and general FT1 penultimate distributions of wind speeds drawn from tail-equivalent Weibull parents". Structural Safety, 30 (2008), 1-10.