Wind and Structures, *Vol. 19, No. 3 (2014) 271-294*
DOI: http://dx.doi.org/10.12989/was.2014.19.3.271

Detecting artefacts in analyses of extreme wind speeds

Nicholas J. Cook *

Consultant, RWDI, Unit 4, Lawrence Way, Dunstable, Bedfordshire, LU6 1BD, UK

(Received April 19, 2014, Revised May 10, 2014, Accepted June 25, 2014)

Abstract. The impact of artefacts in archived wind observations on the design wind speed obtained by extreme value analysis is demonstrated using case studies. A signpost protocol for detecting candidate artefacts is described and its performance assessed by comparing results against previously validated data. The protocol targets artefacts by exploiting the serial correlation between observations. Additional "sieve" algorithms are proposed to identify types of correctable artefact from their "signature" in the data. In extreme value analysis, artefacts displace valid observations only when they are larger, hence always increase the design wind speed. Care must be taken not identify large valid values as artefacts, since their removal will tend to underestimate the design wind speed.

Keywords: wind observations; validation; bias; artefacts; extreme values; design wind speed

1. Introduction

Wind speed observations provide an essential starting point for many civil and structural engineering projects, ranging from assessing the service performance of structures using the "parent" wind speeds to design against wind storms using the "extreme" wind speeds. A major issue is that anemograph networks are designed and managed for other purposes, principally aviation meteorology, and are not intended to provide data of the accuracy and precision needed for engineering design. It is therefore necessary to make the best possible use of the observational data records that exist, in particular to locate and eliminate errors that may distort analyses of the data, especially in analyses of extreme wind speeds. It is in the nature of extreme value analysis that artefacts replace valid observations only when they are greater, always increasing projected design values, so the issue is therefore more one of economy of design than of safety. Nevertheless, Gatey and Millar (2007) remark on the absence of a clearly defined method for identifying outliers and suggest that methods for statistically identifying outliers should be explored.

For many locations around the world, the hourly surface synoptic observations (SYNOP) or the aviation meteorological reports (METAR) are the only readily available sources. Depending on the national observing protocols, these give between two-minute to ten-minute mean wind speed and direction values, reported on the hour (UTC) and, sometimes, at 30 or 20 minute intervals in between. Commercial imperatives for instant access and minimal cost drive structural engineers to rely mostly on the US National Climatic Data Centre (NCDC) archive, because this supplies data

*Corresponding author, Dr., E-mail: njc.wind@ntlworld.com

http://www.techno-press.org/?journal=was&subpage=8 ISSN: 1226-6116 (Print), 1598-6225 (Online)

from around the World rapidly and free of charge. The primary issue with these "raw" observations is their quality.

There are two principal forms of error in wind data: bias errors – errors that apply consistently in the data set and are mostly amenable to correction; and artefacts – erroneous values that occur intermittently in the data set and are only sometimes correctable. In reviewing errors in wind observations, Cook (2014b) showed that the most common types of bias error are:

1. Unit bias – where the archived units are not the originally acquired units and the values have been re-rounded to integers. Many countries* acquire wind speeds as integer values of knots (kn), but NCDC archives these data as integer values of miles-per-hour (MPH) in the Surface Hourly Abbreviated Format (SHAF) or as integer values of tenths of a metre per second (m/s ×10) in the newer Integrated Surface Hourly Data (ISHD) TD3505 format. In both cases, the conversion factor is greater than unity and the original values of the observations are directly recoverable.

2. Odd-Even bias – In manually-acquired data, a weak psychological bias occurs in the selection of values read by an observer, such that decadal values: 10, 20, 30 ... are preferred, then half-decadal values: 5, 15, 25 ..., then even values. A much stronger bias is due to the optical illusion caused by the standard anemograph chart which has scale lines at 2 kn intervals. Observations which lie in the space between lines are 12% more likely to be assigned to the central odd value than to the even value of the scale (Cook 2014b). One might expect that automated acquisition procedures would eliminate observational bias but, unless they are well designed, odd-even bias may still occur due to mismatches between the acquisition and archiving precisions. Individual observations cannot be corrected for odd-even bias, but frequency tables or histograms may be corrected by offsetting the bins by ½ a unit and placing half the counts from each of the two adjacent bins into the new offset bin. For the cumulative distribution function (CDF) this re-binning is directly equivalent to taking the median, or central order statistic, for each integer wind speed value.

3. "Calm bias", or "Form 6910 bias" – is a side-effect of reporting "calms". A common protocol for observers is that whenever the mean speed is 0 or 1 and the wind vane shows gusty variations, the speed is reported as 2 and the gust as missing. A mean speed of 0 is only reported if the speed trace indicates calm over the complete 1 hour period and the vane is unmoving. Thus flat calms are reported as 0, values in the range 0 to 2 are reported as 2, while there should be no values of 1 at all. Information is lost by this protocol, so it is impossible to correct individual values. However, as the protocol preserves the total count of values from 0 to 2, the CDF for 3 and above is unaffected.

4. Anemometer torque hysteresis – Many anemometers are the generator type, using a multi-pole permanent magnet that requires a torque equivalent to a wind speed of around 5 to 6 kn to start rotating from rest. Once rotating, the rotor is given sufficient impetus from the previous pole to overcome the next one and will keep rotating if the wind speed drops below 5 kn. This non-linear behaviour creates a hysteresis bias, with too many counts in the range 0 − 2 kn and too few in the range 2 − 5 kn. Again, individual observations cannot be corrected for this effect. When fitting the CDF to a model distribution, the values for 5 kn and below should be plotted to preserve the correct plotting positions, but should be discounted from the fit.

5. Changes in calibration – Unintended changes are often not detected until large enough to be obvious, in which case the observational record may be erroneous for some considerable time

* The procedure to determine whether the observations were originally in integer knots is to make the conversion to knots, then to confirm that all values fall within ±0.1 kn of integer values (Cook 2014b).

previously. When there is another anemometer close to the site, calibration changes may be revealed by comparing the long-term average wind speeds, often called "buddy checking". Intended changes, e.g., adjusting the height of the anemometer above ground level, may be reversed by applying revised calibration factors to the corresponding periods in the data record, provided that the date of each change is known and that the new calibration factors are also known, i.e., provided that adequate metadata records have been kept for the site and are available to the user.

6. Changes in exposure – While every effort is made to locate each anemometer at 10m above ground in the standard open-country exposure, this is almost impossible to achieve for all wind directions. It is therefore likely that a significant calibration factor for exposure will need to be applied by direction. Exposure may change during the observation period, requiring recalibration for specific periods of the observations. There are standard methodologies for assessing the effect of the surface roughness and orography around a site on the mean and gust wind speeds (Cook 1985) and for assessing the shelter caused by large obstructions, such as buildings (Taylor and Salmon 1993). Observations are corrected by applying these gain factors to the respective periods of record.

7. Seasonal bias - Except for investigations of seasonal dependence, e.g., as in Cook (1983), wind speed analyses are generally conducted irrespective of season. At locations with a strong seasonal dependence, short data records that do not comprise whole numbers of years, or have significant periods missing, may contain different numbers of observations in each season, leading to a bias to the more frequently observed season. Seasonal dependence may be subsumed into the analysis by the use of continuous whole-year records. Alternatively, where there are significant periods missing, by ensuring the data contains equal numbers of observations for each calendar month. When analysing annual maxima in the northern hemisphere, it is better to replace calendar years by "wind-years", starting on 1 July and finishing on 30 June, so that each wind-year contains a complete windy season. This removes any possibility that a major storm spanning the change of year contributes to both year's maxima.

Artefacts in the upper tail of the data are a particular issue when analysing extremes, e.g., annual maxima. There may typically be only around 20 annual maxima available for a given site and, if just a few of these values are erroneous, the predicted once-in-50-year wind speed used for engineering design may be significantly in error. Cook (2014b) showed that the most common types of artefact include:

1. Human errors – Also described by Le Blanc (2010), these are typically:
    - Transpositions – The SYNOP or METAR code contains the right digits, but in the wrong order. A common transposition is of the leading zero of a value less than 10, e.g. $07 \rightarrow 70$.
    - Substitutions – Where other digits within the SYNOP or METAR code are substituted for the correct digits, often imposing a pattern on the data that is not really there.
    - Wrong digit – Where one digit is confused with another, similar looking digit; or where the wrong key is pressed, typically a key immediately adjacent to the correct key on the keyboard ("fat-finger" syndrome).
    - Duplications – where an obviously erroneous or missing observation has been replaced by a copy of the preceding, or succeeding observation.
2. Processing and archiving errors – Decoding and archiving the observational reports provides opportunities to introduce additional artefacts of the types listed above. Additional examples of archiving errors are assignment of the wrong date/time or the wrong location to the observations.

3. Interference by an external agency – It is well known that cup anemometers tend mostly to over-speed when not correctly aligned. Within the meteorological folk-lore, there are apocryphal accounts of bent or twisted cup-arms when anemometers have been used as targets for shooting practice and of jet-wash or prop-wash from aircraft taxiing or queuing by the anemometer location.

2. Extent of the problem – case study: Heraklion, Crete

2.1 Location and observations

Heraklion Airport lies near the middle of the northern coast of the island of Crete, which consists of a long east-west orientated mountain ridge with a narrow coastal plain. A ten year record of observations in SHAF[†] format was obtained from the NCDC archive and the procedures to remove bias errors prescribed in Cook (2014b) were applied. The mean wind speeds were converted back from the archived units of miles-per-hour (MPH) to the originally observed units of knots (kn). While implemented to remove unit bias, this procedure was also essential for identifying most of the artefact types caused by human error.

2.2 Annual maximum mean wind speeds

In locations for which there is no adequate code of practice, or for structures requiring special treatment, the established construction industry standard is to base structural design studies on an analysis of extreme winds at the nearest anemograph station. Typically, this is done using the observed record of annual maxima using the conventional Gumbel (1958) methodology, but nowadays using updated plotting positions (Gringorten 1963) to reduce analysis bias and "bootstrapped" confidence limits (Efron and Tibshirani 1993) to assess the accuracy of the result.

Fig. 1 shows the classical Gumbel analysis, including confidence limits and outlier rejection, applied to the ten reported annual maxima from 1996 to 2005 inclusive. The maximum recorded value (98 kn) lies well outside the confidence limits, so qualifies under the classical Gumbel method as an "outlier" to be discounted from the analysis and considered separately. This leaves 9 surviving maxima (assuming here that only the annual maxima are available so that the outlier cannot be replaced). These 9 values, in their revised plotting positions, lie within the confidence limits, except for the lowest value, and are a reasonable fit to the Gumbel model. Job well done, one would think.

[†] After this work was completed, NCDC replaced the database in SHAF format by a new database in ISH (Integrated Surface Hourly) format, where the wind speeds are now reported in tenths of metres per second. This change has introduced a new type of artefact, not described in Cook (2014b), that is confined to SYNOP (FM-12) reports issued on the hour (UTC) in which the wind speed is erroneously reported in units of tenths of a knot. This artefact always follows ten minutes after a valid METAR (FM-15) report and is detectable because the reported SYNOP speed is always exactly ten times the integer knot value of the METAR speed. For example, a METAR speed reported as 144 corresponds to 28 kn and the following SYNOP reports this as 280. As the calibration factor from m/s to kn is 1.944, this artefact approximately doubles the wind speed.

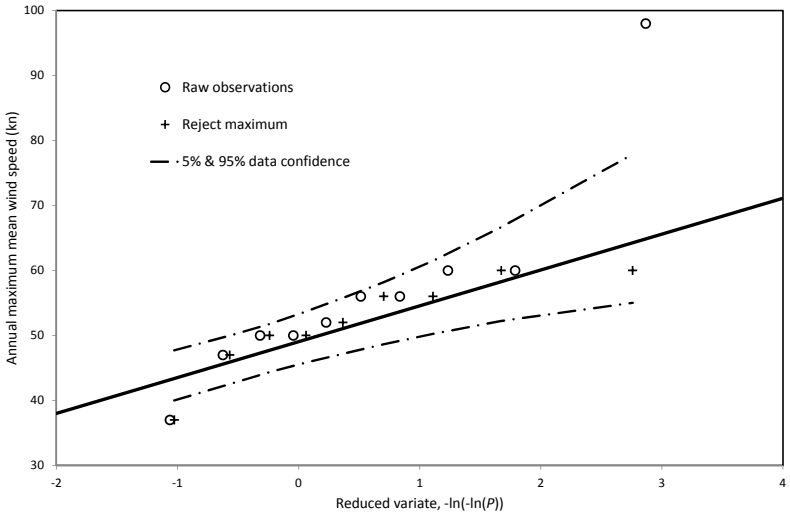

Fig. 1 Gumbel plot of reported annual maximum mean wind speeds at Heraklion, Crete, 1996-2005

2.3 Count-back procedure to cull artefacts

With access to the original wind record, the annual maximum for each year may be examined in the context of the trend established by the previous and subsequent observations. If deemed to be an artefact, a rejected value can be replaced by the next-highest annual value. The process starts with the highest annual maximum and works downwards in value until all the surviving, or replacement, annual maxima are valid.

Starting with the highest annual maximum and proceeding in descending order of value, Table 1 shows the two observations preceding and succeeding each reported or replacement annual maximum mean wind speed. Each annual maximum is compared with the trend established by the preceding and succeeding values and is replaced by the next-highest annual maximum if it appears to be an isolated outlier compatible with one of the artefact types listed above. The reasons for rejection, which are given in the "Attribution" column, include "fat-finger syndrome" where a key adjacent to the intended key has also been struck, or a key has been struck twice, and transpositions of the leading zero. Where the artefact is correctable, the correct value is given in brackets. The 46 value in 2003 does not correspond to any of the artefact types, but is nevertheless rejected the grounds of trend alone. Two directions are shown in bold type where the values are not a multiple of 10°: one is attributed to "column shift"[‡] in decoding, i.e., the 0 had been replaced by first digit of the speed and the speed digits shifted by one column; the other is un-attributable. It is acknowledged that in some cases these attributions are speculative and so are indicated by the "?" symbol, but in other cases, such as transpositions, the attribution is more certain. It is quite possible that some are valid observations caused by short-duration thermal events (e.g., thunderstorm downburst) which ought to be considered separately anyway.

[‡] As the NCDC databases are in fixed column format, another indicator of "column shift" during decoding is that the parameters in the record to the right of the shift are shown as invalid ("999").

Table 1 Count-back procedure for annual maxima at Heraklion, Crete, 1996 – 2005

Year	Month	Day	Hour	Min	Dirn	Speed	Attribution
2002	8	23	21	0	280	9	
2002	8	23	22	20	290	8	
2002	8	23	22	50	300	98	Fat finger
2002	8	23	23	20	300	98	Fat finger
2002	8	24	0	0	300	9	(300, 08)
2002	8	24	0	20	290	11	
1996	6	7	21	0	300	12	
1996	6	7	21	50	300	12	
1996	6	7	22	50	200	60	Decadal
1996	6	8	0	0	300	11	
1996	6	8	3	0	300	9	
2003	5	24	7	50	10	6	
2003	5	24	8	20	10	6	
2003	5	24	8	50	40	56	Fat finger
2003	5	24	9	50	180	5	(040, 06)
2003	5	24	9	0	220	8	
2005	4	14	0	0	190	7	
2005	4	14	0	50	210	8	
2005	4	14	0	20	200	56	Fat finger
2005	4	14	1	50	200	5	(200, 06)
2005	4	14	1	20	200	7	
2003	9	20	11	20	330	13	
2003	9	20	11	50	310	15	
2003	9	20	12	0	110	54	Column shift?
2003	9	20	12	20	320	16	(310, 15)?
2003	9	20	12	50	320	17	
1997	12	7	3	0	160	11	

1997	12	7	3	50	170	19	Missing first digit?
1997	12	7	5	0	70	52	Transposed speed?
1997	12	7	6	0	190	12	(170, 25)?
1997	12	7	6	50	180	13	
2005	5	19	19	50	160	8	
2005	5	19	20	20	170	12	
2005	5	19	20	50	300	52	Fat finger
2005	5	19	21	50	290	4	(300, 05)
2005	5	19	21	0	260	6	
1998	6	3	22	50	330	15	
1998	6	4	0	0	330	14	
1998	6	4	0	50	360	50	Missing first digit?
1998	6	4	1	50	330	12	(360, 15)?
1998	6	4	3	0	330	12	
1999	12	22	20	20	200		
1999	12	22	20	50	180	6	
1999	12	22	21	0	100	50	Transposed
1999	12	22	21	50	240	6	(100, 05)
1999	12	22	22	50	210	7	
2000	6	30	20	50	180	5	
2000	6	30	21	50	180	5	
2000	6	30	21	0	180	50	Transposed
2000	6	30	22	50	180	5	(180, 05)
2000	7	1	0	0	180	5	
2002	11	5	9	0	180	19	
2002	11	5	9	20	180	20	
2002	11	5	9	50	180	50	Wrong digit
2002	11	5	10	20	180	15	(180, 20)
2002	11	5	10	50	170	17	

1997	9	11	12	0	330	10	
1997	9	11	13	20	320	10	
1997	9	11	14	20	**117**	48	Invalid direction ?
1997	9	11	15	0	320	9	
1997	9	11	16	20	350	6	
2004	7	30	4	50	160	7	
2004	7	30	5	20	150	7	
2004	7	30	5	50	150	47	Fat finger
2004	7	30	6	0	140	3	(150, 07)
2004	7	30	6	20	90	3	
2003	12	25	21	20	170	6	
2003	12	25	21	0	170	8	
2003	12	25	21	50	200	46	?
2003	12	25	22	20	220	6	
2003	12	25	22	50	220	7	
2003	9	13	3	50	220	5	
2003	9	13	4	20	240	5	
2003	9	13	4	50	240	45	Fat finger
2003	9	13	5	20	220	4	(240, 05)
2003	9	13	5	50	250	6	
2004	7	10	4	20	280	14	
2004	7	10	5	20	280	15	
2004	7	10	5	50	**291**	45	Column shift
2004	7	10	6	0	290	14	(290, 14)
2004	7	10	6	20	290	17	
2004	12	22	15	20	220	4	
2004	12	22	15	50	190	4	
2004	12	22	16	20	100	44	Double key?
2004	12	22	17	20	999	0	(100, 04)
2004	12	22	17	50	140	3	

2002	1	6	5	50	340	23	
2002	1	6	6	50	330	23	
2002	1	6	7	0	350	43	Wrong first digit?
2002	1	6	7	50	360	22	(350, 23)
2002	1	6	8	50	340	21	
2005	2	15	12	20	170	38	
2005	2	15	12	50	170	38	
2005	**2**	**15**	**13**	**0**	**160**	**43**	**Valid observation**
2005	2	15	13	50	150	29	
2005	2	15	13	20	160	35	

This count-back process is complete when every year in the record contributes a valid annual maximum value. Note also that some second-, third- and even fourth-highest annual maxima are themselves artefacts, requiring repeated replacement until the largest valid value is found. For brevity, Table 1 is truncated at the first valid annual maximum – which happens to be the only valid value of the original ten reported annual maxima.

2.4 Re-analysis of validated annual maxima

Fig. 2 shows the Gumbel plot for the culled observations, compared with the original observations. Of the ten original annual maxima, only one is valid – the outlier in the lower tail in Fig. 1 – and this now lies near the middle of the range of the culled observations. The projected design once-in-50-year mean wind speed falls from V_{50} = 70.6 kn (36.3 m/s) to V_{50} = 46.3 kn (24.0 m/s), more than halving design loads which are proportional to the square of wind speed.

Some post-justification of this count-back procedure is provided by two tests. First is the histogram of integer knot counts for the original and culled data in Fig. 3. The Weibull distribution model has been fitted to the culled data and indicates, by where it crosses the 1 count axis, that valid values greater than 43 kn are unlikely to occur naturally in the record. (Note that the fit is poor below 5 kn, as expected, due to the effects of calm bias and anemometer torque hysteresis.) The second is provided by the average rate of artefacts exceeding the valid annual maximum, i.e., the annual frequency of culls, λ = 1.8. Fig. 4 investigates whether this value of λ corresponds a purely random process by comparing the distribution of the number of artefacts exceeding the valid annual maximum in each year with the expectation for the Poisson recurrence model. Given there are only 18 culled values in the ten years, the match for this small sample is remarkably good.

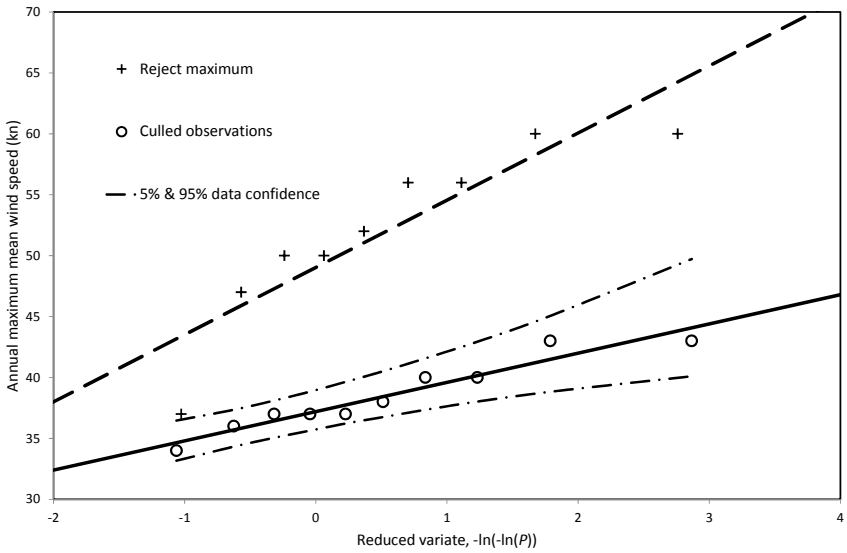

Fig. 2 Gumbel plot of reported and validated annual maximum mean wind speeds at Heraklion, 1996-2005

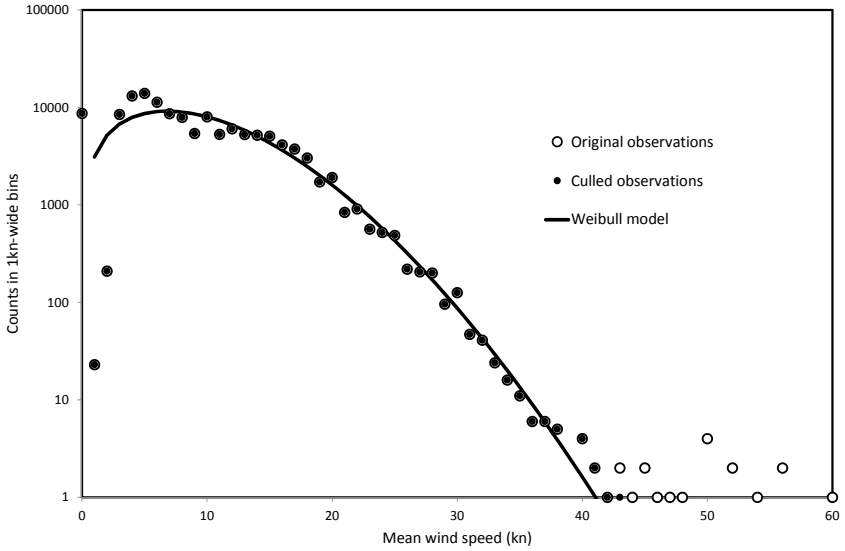

Fig. 3 Histogram of integer knot wind speed counts for Heraklion, Crete, 1996 – 2005

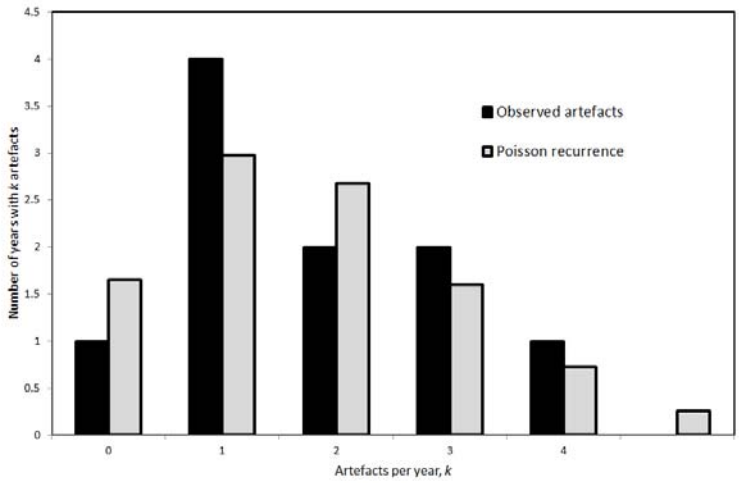

Fig. 4 Distribution of the number artefacts exceeding the valid annual maximum in each year

3. Signpost protocol for sequential observations

3.1 Requirements for a signpost protocol

An ideal signpost protocol is one which, using only the sequential observations of wind speed, detects all artefacts, missing none and without any false positives. This is, of course, impossible and so it is necessary to aim lower. A practical protocol should balance the conflicting requirements of minimising both the number of missed artefacts and the number of false positives.

In the count-back procedure of 2.3, above, the criterion for an artefact is based on the size of its deviation from the temporal trend. In effect, it exploits the strong serial correlation between successive values. Hence, a candidate signpost protocol can be derived by considering the difference between successive values, rather than the values themselves. This form of approach is not novel. During the late 1960s – early 1970s, Jones (1973) analysed atmospheric turbulence in terms of sequential incremental changes, which led to the development of the statistical discrete gust approach to predict aircraft response to turbulence (Jones 1989). Here, the incremental change is defined as the increment between the current observation and the value for previous hour – hence

$$\Delta_V(h) = V(h) - V(h-1) \tag{1}$$

for wind speed values, V, in sequential hours, h.

3.2 Data for development of the signpost protocol

Development and testing the efficiency of any signpost protocol is greatly helped by knowing in advance where the artefacts are in a set of observations. Cook (2014b) had already detected artefacts in the NCDC archive for SYNOP observations at Marham airfield, UK, from 1973 to 2000 and this validated set of artefacts was used to develop and evaluate the signpost protocol.

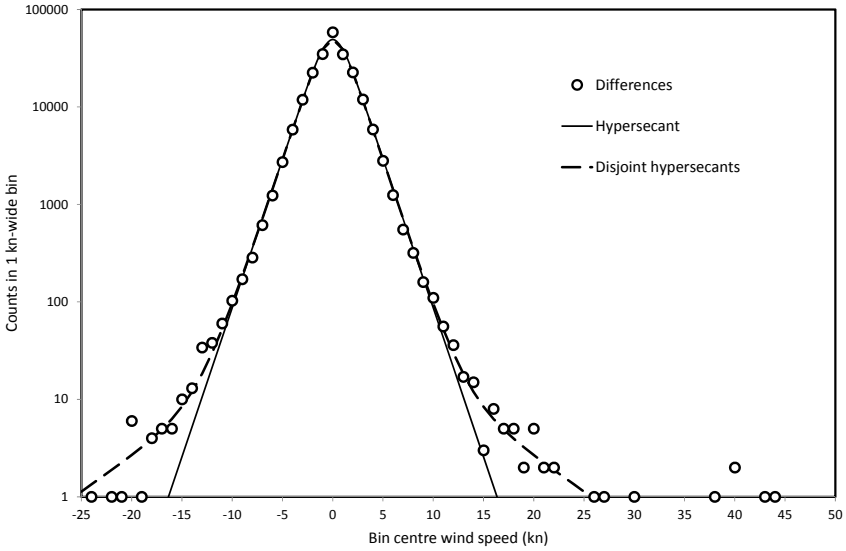

Fig. 5 Frequency of hourly increments of mean wind speed in NCDC data for Marham, UK, 1973 – 2000

Fig. 5 shows the frequency distribution of the hourly increments for the mean wind speeds at Marham. This appears symmetrical with exponential tails, and is well modelled by the hyperbolic secant (HS) distribution over three orders of magnitude

$$p(x) = \frac{1}{2}\operatorname{sech}\left(\frac{\pi x}{2\sigma(x)}\right) = \frac{1}{\exp\left(\frac{\pi x}{2\sigma(x)}\right) + \exp\left(\frac{-\pi x}{2\sigma(x)}\right)} \tag{2}$$

Here $\sigma(x)$ is the standard deviation of x. Note that the hyperbolic secant decomposes into a pair of exponential terms, each representing one of the two tails. The observations appear to deviate systematically from the HS model in the far tails, indicating the possibility of a rare mechanism giving large increments, e.g., squalls. The dashed curve shows the fit for two disjoint HS distributions, indicating that this second mechanism, if it exists, occurs for only 0.3% of the time.

The fitted HS distribution in Fig. 5 crosses the 1 count axis at $|\Delta_V|$ = 17 kn for the single HS and at 26 kn for the disjoint HS. Values of $|\Delta_V|$ less than 17 kn are indicative of valid observations, values greater than 26 kn almost certainly indicate artefacts, while intermediate values may be artefacts or the supposed rare events. However, Cook (2014b) notes that many "wrong first digit" artefacts correspond to an error of 10 kn, and would not be indicated by a threshold larger than this value.

Cook's validated artefact error values for a threshold of 10 kn, ε_V, are compared with the corresponding hourly increments of wind speed, Δ_V, as the open circles in the Q-Q plot, Fig. 6, showing almost one-to-one correspondence. Accordingly, the increments, Δ_V, are a good proxy for the artefact errors, ε_V, and so can be used as the basis of the signpost protocol.

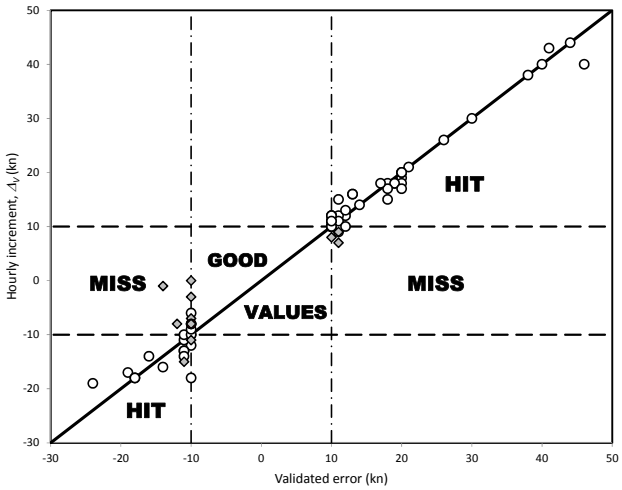

Fig. 6 Regression between validated error and hourly speed increment: open circles – artefacts identified by protocol; diamonds – artefacts missed by protocol

3.3 Evolution of the signpost protocol

Applying a simple threshold value of 10 kn produces 564 exceedances in the Marham data, but a significant number of these appear as pairs in consecutive hours. Consider the case of the transposition artefact in the sequence of wind speeds: V = 4, 5, 50, 6, 5. In the hour, h, of the artefact $\Delta_V(h) = 50 - 5 = 45$, but in the following hour, $\Delta_V(h+1) = 6 - 50 = -44$, giving a second indication of the same artefact. Taking this behaviour into consideration results in 493 unique exceedances which, when compared with the candidate artefacts, corresponds to 87 matches, 7 misses and 399 false positives. This simple threshold process succeeds in reducing the subsequent burden of checking for artefacts from a possible 222661 observations down to only 493 likely candidates, but it remains a tedious task to validate these manually. To reduce this burden further, the protocol needs to be more subtle.

Cook (2014b) showed that most artefacts are "singletons" in that the values in the preceding and succeeding hours are valid, as in the above example. Pairs of consecutive artefacts occur but are much rarer, while three or more consecutive artefacts are unlikely to occur, except as precursors to a period of missing data which indicates an equipment failure. Hence, in order to identify singletons and pairs the protocol requires an increment in hour h exceeding the threshold to be reversed by another increment exceeding the threshold but of the opposite sign in hour h+1 or h+2. This eliminates valid large, but sustained, increments in speed caused by synoptic changes, such as the passage of fronts. However, strong rising/falling trends in wind speed add complications. Consider an artefact which occurs during a strong rising trend of 5 kn per hour: where the validated error is ε_V = 13 kn, but $\Delta_V(h)$ = 16 kn and $\Delta_V(h+1) = -5$ kn, so the threshold is exceeded in the first, but not the second hour. The protocol compensates by adding the hourly trend to the threshold values: in this example the trend is +5 kn, so the thresholds become ≥+15 kn for hour h and ≤−5 kn for hour h+1, and thus the artefact is detected.

Table 2 Performance metrics of protocol

Real artefacts found by protocol	71
Real artefacts missed by protocol	7
Miss rate	9.0%
False positives	175
Efficiency	29%
Artefacts as percentage of observations	0.04%

Table 2 shows the performance metrics for the signpost protocol operating on the Marham observations. For a threshold of ±10 kn, with trend compensation, the miss rate is 9%. The detection efficiency is 29%, defined as the proportion of detected candidates that are actually artefacts. The detected candidates represent 0.04% of the total population of observations, which greatly reduces the burden of manual checking of the data. Lowering the threshold to 9 kn halves both the miss rate and detection efficiency, doubling the manual checking burden. Raising the threshold to 11 kn is disastrous as it fails to identify all of the "missing first digit" and most of the "wrong first digit" artefacts.

This protocol can be deemed successful if the missed artefacts are not significant. The missed artefacts for the 10 kn threshold are indicated in Fig. 6 by the diamond symbols. These are clustered against the 10 kn threshold for the validated error, ε_V, and indicate where either $\Delta_V(h)$, $\Delta_V(h+1)$ or $\Delta_V(h+2)$ only just failed to exceed the threshold, despite the trend compensation. Fig. 7 shows the distribution of $|\Delta_V|$ for the false positives which are also clustered against the 10 kn threshold, such that 90% of the false positives lie within $3\sigma_\varepsilon$ of the threshold.

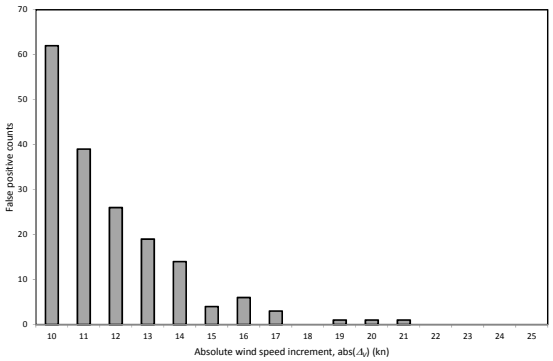

Fig. 7 Distribution of false positives from signpost protocol

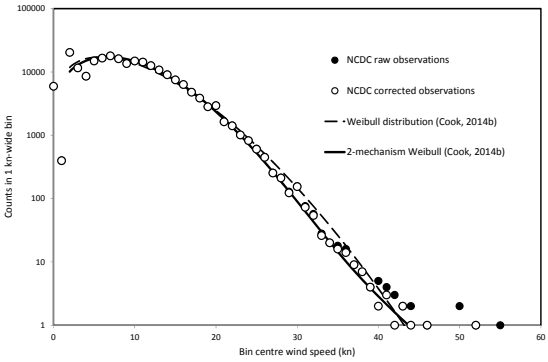

Fig. 8 Frequency of mean wind speeds in NCDC data for Marham, UK, 1973 – 2000: raw and corrected observations using 10 kn threshold with trend compensation

3.4 Correcting the wind speeds

Having identified the artefacts in the wind speeds, the options for correction are to discard these observations or, especially when the application for the data requires sequential observations, to interpolate replacement values. In the first case, the body of frequency distributions will hardly be affected by the loss of around 0.04% of the observations, while the upper tail will have had the largest and most reliably detected artefacts removed. In the second case, interpolation for false positives should only make small differences to the values, because of the strong serial correlation between adjacent hours.

Hence, instead of validating the individual candidate artefacts by manual inspection, depending on the use to which the data will be put, it may be satisfactory to run an interpolation procedure on all the candidates to automate the process fully. The effect on the wind speed observations is summarised in Fig. 8 which shows the frequency distribution for each integer value of wind speed in the raw and the corrected observations. (Note, again, that the fit is poor below 5 kn, as expected.) Elimination of the artefacts gives better resolution of the upper tail, so that the corrected observations lie on the fits for the datum UK Met Office data from Cook (2014b) for the single and the two-mechanism disjoint Weibull model – the two-mechanism model clearly giving the better fit. The Weibull parameters for these fits are presented in Table 3 and indicate that the suggested rare mechanism occurs ~1% of the time, but dominates in the far upper tail due to its lower shape parameter.

Table 3 Weibull distribution model parameters fitted to MetO data set for Marham

Weibull parameter		Single model	Disjoint 2-mechanism model	
Frequency	$f =$	1.00	0.100	0.900
Shape	$w =$	1.67	1.28	1.87
Scale (kn)	$C =$	10.0	8.09	10.4

3.5 Adapting the protocol for wind direction artefacts

Two issues have to be addressed before the signpost protocol can be adapted to detect artefacts in wind direction.

1. Owing to the cyclical nature of direction, it is not possible be certain whether hourly increments are positive (veering) or negative (backing). Applying the principle of Occam's razor, it is assumed that the smaller possibility always applies, so that increments of wind direction lie within the range ±180°, setting limits to the largest value.

2. Large increments of wind direction occur much more frequently in lighter winds and, without compensation for this, the protocol would be strongly biased to low wind speeds.

Both these issues may be accommodated by transforming the increment in wind direction, Δ_D, into the incremental arc wind speed, Δ_A, defined by

$$\Delta_A = V \times \mathrm{rad}(\Delta_D) \tag{3}$$

This has dimensions of wind speed and is unlimited in both tails because V is unlimited.

Fig. 9 shows the frequency distribution of the arc increments, Δ_A, in the same format as Fig. 5. It is obvious that a single HS distribution (the dashed curve) is not a good model because the deviations in the tails are more prominent than in Fig. 5 and, also, neither the body nor the tails are symmetric. The deviations can again be accommodated by the disjoint sum of two distributions, but accommodating the asymmetry requires an asymmetry parameter, α, in (2) to scale the slopes of each tail, thus

$$p(x) = \frac{\left[\frac{\alpha}{2} + \frac{1}{2\alpha}\right] \sin\left(\frac{\pi}{1+\alpha^2}\right)}{\exp\left(\frac{\alpha \pi x}{2\sigma(x)}\right) + \exp\left(\frac{-\pi x}{2\alpha\sigma(x)}\right)} \tag{4}$$

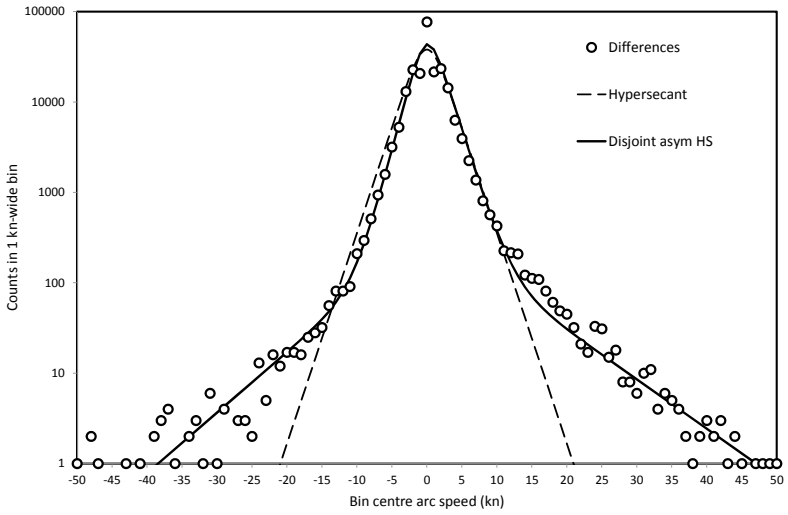

Fig. 9 Frequency of hourly increments of arc wind speed in NCDC data for Marham, UK, 1973 – 2000

Eq. (4) is a generalisation of the hyperbolic secant distribution and reverts to (2) when $\alpha = 1$, and the solid curve in Fig. 9 is the fit to this. The asymmetry parameter for the principal mechanism is $\alpha = 0.99$ (nearly symmetrical), $\sigma(\Delta_A) = 2.5$ kn, and it operates for 98% of the time. The asymmetry parameter for the rare mechanism is $\alpha = 0.91$ (distinctly asymmetrical), $\sigma(\Delta_A) = 11.3$ kn, and it operates for 2% of the time.

It follows from Fig. 9 that the signpost protocol will indicate candidate directional artefacts by applying a threshold somewhere between 20 and 40 kn. Using 30 kn, all the validated artefacts from Cook (2014b), comprising 16 singletons and one pair, were found by the protocol (no misses) together with 41 false positives (efficiency 29%). It also follows, that values of $|\Delta A| \leq 20$ kn will correspond to a mixture of the principal and the rare mechanisms, but values in the range $20 \leq |\Delta A| \leq 30$ kn should contain only valid events from the rare mechanism, indicating 63 backing events and 91 veering events per year.

3. Design application of the protocol – Ta'if Regional Airport, Saudi Arabia

The city of Ta'if, at an altitude of around 2000m on a plateau southeast of Mecca, is a popular summer resort and is undergoing extensive development, including new high-rise structures requiring sophisticated engineering design calculations. Such calculations are only as good as the meteorological data from Ta'if Regional Airport that supports them. At the time of this design study, the available meteorological record from NCDC was from 1983 to 2008, comprising a mixture of SYNOP and METAR reports. Fig. 10 shows the frequency histogram of wind speeds in integer knot bins for the raw observations (open circles) and after re-binning (filled circles) to remove the very strong bias to even values evident in the data – 86% are even and 14% are odd. An excess of the decadal values 40, 50 and 60 kn is very evident, and an excess of 10, 20 and 30 kn values is also indicated within the range of valid values. There are no decadal values higher than 60 kn, suggesting that, if these had existed, they may have been culled from the observations as being obvious artefacts.

The very strong even bias in the wind speeds is also evident in the frequency distribution of wind speed increments, Fig. 11. Differences between consecutive even speeds and between consecutive odd speeds always produce even differences, while odd differences can only occur between and odd and even speeds. The differences bias of 76% even and 24% odd exactly matches the expectation from the observation bias.

Applying the signpost protocol to the Ta'if observations with a threshold of 20 kn produced 110 candidates. This value was high enough to limit the number of false positives, but was too high to detect 10 kn decadal artefacts, which were judged to be sufficiently diluted by the very large population of valid observations around this value. Instead of inspecting the candidates manually, an algorithm was run to sift out decadal values where adjacent values were less than 10, which indicate transpositions of leading zeros, e.g., 05→50, and artefacts where the wind direction was not a multiple of 10°, which indicate other transpositions within the SYNOP or METAR codes. The 36 unattributed candidates left over were sorted by direction and plotted as Fig. 12. This shows around half are concentrated in the narrow range of direction between 10°T and 30°T, implying they may be due to thermal events occurring regularly from that bearing, or interference by an external agency from that direction.

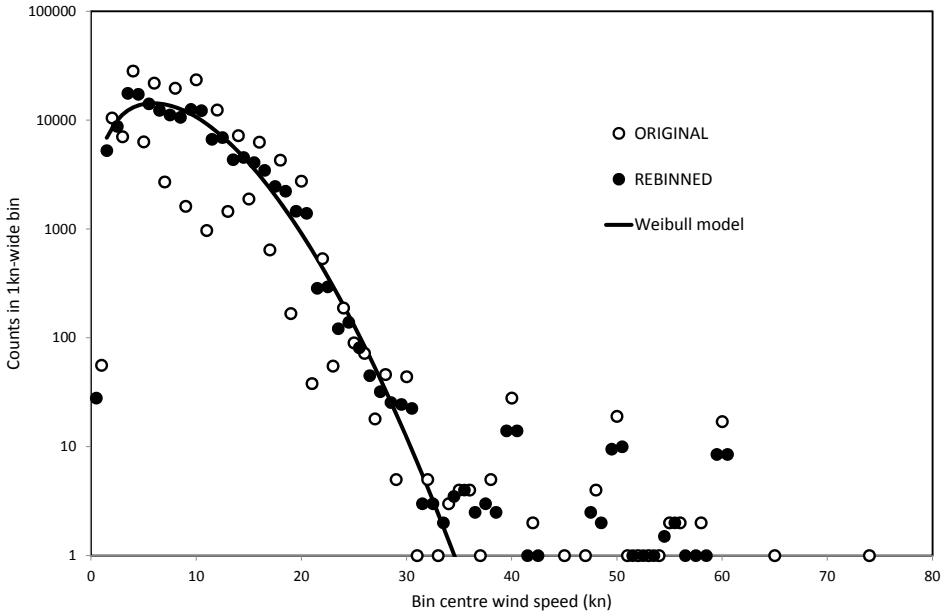

Fig. 10 Frequency of mean wind speeds at Ta'if Regional Airport, Saudi Arabia, 1983 – 2008

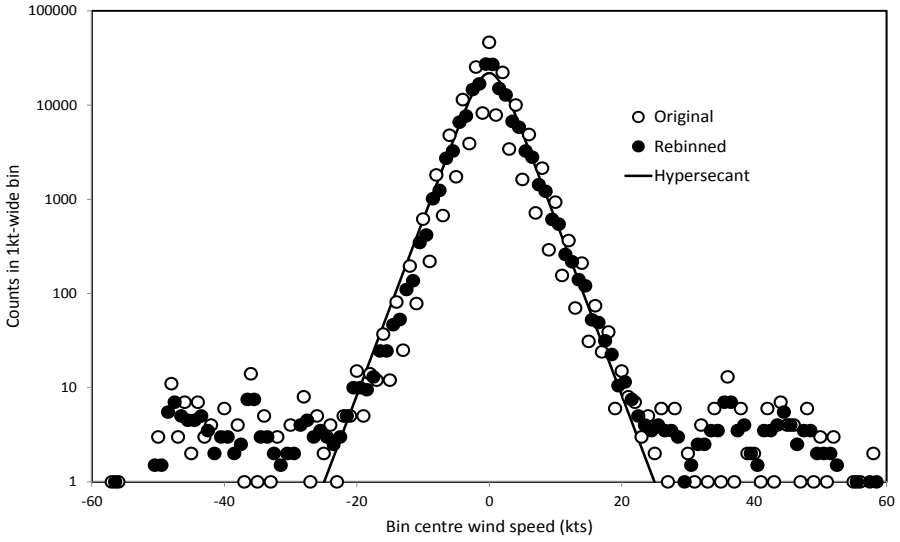

Fig. 11 Frequency of hourly increments of mean wind speed for Ta'if, Saudi Arabia, 1983 – 2008

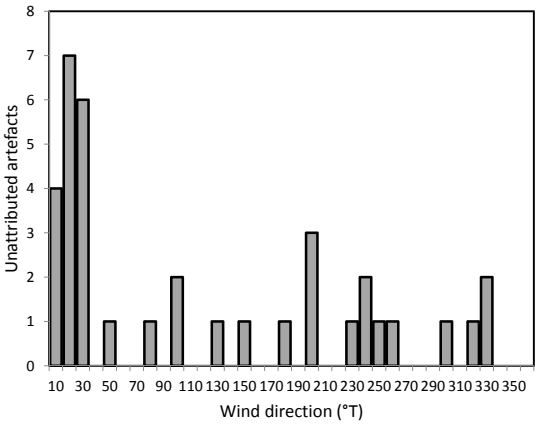

Fig. 12 Frequency by direction of unattributed artefacts for Ta'if, Saudi Arabia, 1983 – 2008

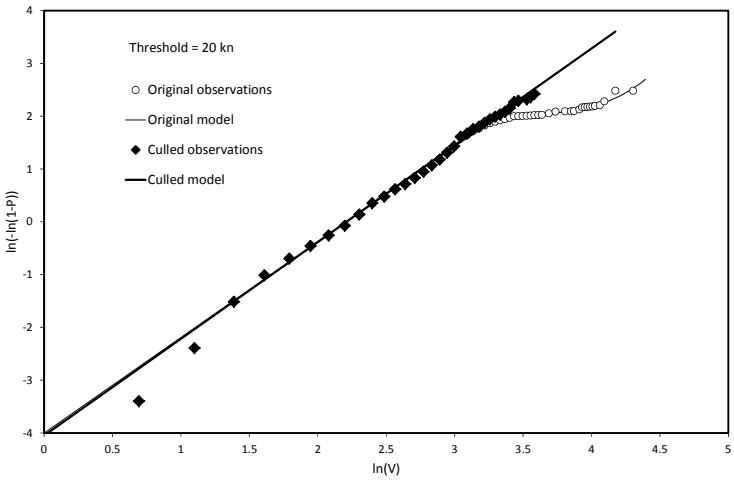

Fig. 13 CDF on Weibull axes for mean wind speeds at Ta'if, Saudi Arabia, 1983 – 2008

The parent CDF of the wind speeds is plotted on Weibull axes in Fig. 13, showing the original and the culled wind speed values. The deviation from the Weibull model in the upper tail of the original values is almost entirely due to the excess of decadal values from transpositions of speeds below 10 kn and this disappears after culling of these artefacts. Note that torque hysteresis bias is evident in the values below 5 kn (for $\ln(V) < 1.7$). Having eliminated artefacts from the full data record, it is more appropriate to use an extreme value analysis method that utilises sub-annual extremes in order to minimise the statistical uncertainty. This author favours the XIMIS method (Harris 2009) for hourly data and the NIST methodology (Lombardo *et al*. 2009) when the data is discontinuous – as recently reviewed in Cook (2014a).

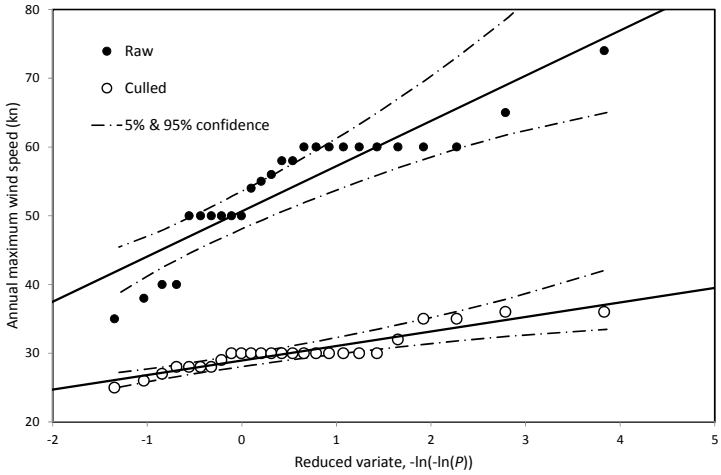

Fig. 14 Extreme value plot on Gumbel axes for annual maximum mean wind speeds at Ta'if, Saudi Arabia, 1983 – 2008

Nevertheless, for compatibility with the earlier Heraklion analysis in Fig. 2, the annual maxima are plotted on Gumbel axes in Fig. 14, which shows that only the smallest of the original annual maxima (35 kn) is a valid observation and corresponds to the 3rd or 4th highest annual maximum. All the other raw annual maxima correspond to transposition or "fat finger" artefacts. Removing the artefacts reduces the design wind speed from V_{50} = 76 kn to 30 kn, a reduction by a factor of four in pressures and loads. Note that the thirteen 30 kn values surviving in the annual maxima are valid values because the values in the preceding and succeeding hours exceed 20 kn: the dominance of these thirteen 30 kn values appears to be a case of observer preference to decadal values at the expense of 29, 31, 32 and 33.

5. Discussion

5.1 "Decadal" artefacts

"Decadal" artefacts crop up regularly in observations around the world. Most are due to transpositions of the leading zero in the SYNOP or METAR code when the values are in the range 02 to 09. However an excess of decadal values also occurs in wind storms, and this may be due to difficulty in reading the meter or chart when the wind is gusty, so that the nearest "round number" is chosen.

5.2 Setting the signpost protocol threshold

The threshold level is a compromise between the value indicated by the observed distributions of differences, e.g., Figs. 5 and 11, and the need to detect the commonest artefact types. Whether

or not to err on the side of a low threshold to catch more artefacts at the expense of more false positives depends on whether the number of resulting candidates is small enough to validate manually. With extremes, it is more important not to exclude a large valid value than it is to exclude a large artefact. Artefacts always act to increase the design wind speed because they must be larger than the valid extreme before they affect the result, but the loss of valid large extremes is unsafe as this decreases the design value. Hence, in the case of extremes, the need for a safe result requires false positives to be minimised by always inspecting the high-value candidates manually, using the count-back procedure. As the differences between valid observations increase as the wind speed increases, there is a case for modulating the threshold by some measure of the "windiness" in addition to the trend compensation.

5.3 Further development of the protocol

Some of the artefact types reviewed by Cook (2014b) are more readily detected by their "signature" in the data than others. It is relatively simple to sift out the "decadal" transpositions that convert values below 10 into large-valued artefacts: *viz:* IF(MOD(V[h],10)=0) AND (V[h-1]+ V[h+1])<=20) THEN (artefact); and the errors where wind direction is not a multiple of 10: *viz:* IF(MOD(D[h],10)<>0 THEN (artefact). The intention is to develop specific "sieves" for specific artefact types: "fat-finger syndrome", "column shift" and "repeated observation" types are obvious next candidates, but more should be possible. This is analogous with the "signatures" used by anti-virus software to detect specific computer viruses.

5.4 The tails of the distribution of increments

The tails of both wind speed and arc (directional) increments in the Marham data suggest two disjoint mechanisms: a principal mechanism operating 99% & 98% of the time, respectively; and a rare mechanism operating 1% & 2% of the time, respectively. There is no indication of any link between the two rare mechanisms: the frequencies are different; the wind speed is symmetrical and the arc speed is not; but, more importantly, there is no significant correlation between the values or increments corresponding to each tail.

The rare mechanism proposed to account for the deviation in the upper tail of the wind speeds at Marham, Fig. 8, may be due to squalls, thunderstorm downbursts, or may even be a manifestation of Browning's "sting in the tail" (Browning 2004): the so-called "sting-jet" phenomenon.

The rare mechanism proposed to account for the large arc (directional) increments, is asymmetric in the sense that veering is ten times more likely than backing. European synoptic charts at 00:00 UTC, from 1999 to date, are available for download from the on-line weather archive www.wetterzentrale.de. The signpost protocol was re-run for the period 2000-2011 at Marham and the large arc increments detected, as before, for comparison. All the strong veering events appear associated with the passage of a cold front where the isobars are strongly bent and, when the event occurs close to 00:00 UTC, the corresponding front lies directly over Marham. On the other hand, all the strong backing events appear to be associated with the passage of localised troughs that are not directly associated with fronts. This suggests that the asymmetry in these directional increments is due to Coriolis effects, so would be expected to be reversed in the southern hemisphere.

5.5 The possibility of jet-wash affecting the anemometer

The vulnerability of airport-based anemometers to jet-wash or prop-wash is not something usually admitted, although the suggestion crops up from time to time in wind-engineering folk-lore. There is no strong evidence that jet-wash from aircraft on the ground has a real and significant effect on observations, but it is a known hazard to people, vehicles and buildings at airports. The article "Engine Thrust Hazards in the Airport Environment" by the Boeing Corporation (http://www.boeing.com/commercial/aeromagazine/aero_06/textonly/s02txt.html) states that the engine efflux velocity of large commercial airliners lies in the range 300 kn to 500 kn and indicates that "wind" speeds in excess of 130 kn occur at 60 m "downwind" of aircraft on the ground. Gaussian jet theory indicates that the maximum wind speed in the plume decays in inverse proportion to the distance "downwind", hence to above 80 kn at 100 m, 40 kn at 200 m, 16 kn at 500 m, etc. It follows that it is quite possible that some observational artefacts that cannot be attributed to the common types are due to jet-wash from aircraft on the ground.

5.6 Short-duration thermal events

It must be admitted that some of the remaining unattributed candidate artefacts may be valid observations caused by short-duration thermal events. Inland dinghy sailors are well aware of wind shifts in "light and variable" conditions that sometimes rotate the wind direction through 360° in less than a minute. Some small thermally-driven events will last long enough to have a significant influence on the two-minute to ten-minute means in METAR and SYNOP reports. Larger events, such as squalls and thunderstorm downbursts, are indistinguishable from singleton artefacts without recourse to additional data. Fortunately, the optional section of SYNOP and METAR reports include a "present weather" (PW) code that identifies these events and forms the foundation for another specific "sieve". However, the PW code is optional and is often missing.

In the Ta'if observations, a current thunderstorm or squall occurs for sixteen directional candidates, but only one speed candidate – a culled value of 40 kn which, now reinstated, corresponds to the maximum valid observation. As this lies outside the confidence limits, it qualifies under the Gumbel methodology as an outlier which needs to be accounted for separately. Indeed, separation of the wind events from different mechanisms, first highlighted by Gomes & Vickery (1978) is now regarded as an essential first step in the analysis of extreme wind speeds in mixed climates (Cook *et al*. 2003, Lombardo *et al*. 2009), and the signpost protocol also assists in this process.

6. Conclusions

- Extreme-value analysis of wind observations to derive design wind speeds is particularly susceptible to the presence of artefacts in the data.
- Most artefacts in wind data are the result of human errors in acquiring, encoding and decoding the values, and this leads to recognisable forms of artefact such as the transposition of digits.
- Annual maxima may be validated by manual inspection and comparison with the synoptic trend indicated by the observations in the preceding and succeeding hours. Artefacts should be replaced by the next-highest valid observation for each year.

- A signpost protocol is demonstrated which exploits the serial correlation between adjacent hours in the observational record to detect candidate artefacts automatically.
- Additional "sieve" algorithms are proposed to detect specific artefact types from their "signature" in the data.
- Valid, short-duration thermally-driven events are included in the candidate artefacts. Such events would normally be assessed separately, so their inclusion is not necessarily an issue and they may be recognised by reference to the "present weather" code.
- The removal of artefacts from the observations always reduces design wind speeds because artefacts displace valid annual maxima only when they are larger.
- Care must be taken not to exclude the largest valid observations by mistaking them for artefacts as this will lead to underestimation of the design wind speed.

Acknowledgements

The author gratefully acknowledges the assistance given by Mr R I Harris in generalizing the hyperbolic secant distribution, (2), into the asymmetric form, (4).

References

Browning, K.A. (2004), "The sting at the end of the tail: Damaging winds associated with extra-tropical cyclones", *Q. J. Roy. Meteor. Soc.*, **130**(597), 375-399.

Cook, N.J. (1983), "Note on directional and seasonal assessment of extreme winds for design", *J. Wind Eng. Ind. Aerod.*, **12**(3), 365-372.

Cook, N.J. (1985), *The Designer's Guide to Wind Loading of Building Structures*, Part 1, Butterworths, London, UK..

Cook, N.J. (2014a), "Consolidation of analysis methods for sub-annual extreme wind speeds", *Meteorol. Appl.*, **21**(2), 403-444.

Cook, N.J. (2014b), "Review of errors in archived wind data", *Weather*, **69**(3), 72-81.

Cook, N.J., Harris, R.I. and Whiting, R.J. (2003), "Extreme wind speeds in mixed climates revisited", *J. Wind Eng. Ind. Aerod.*, **91**(3), 403-422.

Efron, B. and Tibshirani, R.J. (1993), *An introduction to the bootstrap*, Chapman and Hall, New York, USA.

Gatey, D.A. and Miller, C.A. (2007), "An investigation into 50-year return period wind speed differences for Europe", *J. Wind Eng. Ind. Aerod.*, **95**(9-11), 1040-1052.

Gomes, L. and Vickery, B.J. (1978), "Extreme wind speeds in mixed climates", *J. Wind Eng. Ind. Aerod.*, **2**(4), 331-344.

Gringorten, I.I. (1963), "A plotting rule for extreme probability paper", *J. Geophys. Res.*, **68**(3), 813-814.

Gumbel, E.J. (1958), *Statistics of Extremes*, Columbia University Press, New York, USA.

Harris, R.I. (2009), "XIMIS – a penultimate extreme value method suitable for all types of wind climate", *J. Wind Eng. Ind. Aerod.*, **97**(5-6), 271-286.

Jones, J.G. (1973), *Statistical discrete gust theory for aircraft loads: a progress report*, *RAE Technical Report* 73167, Royal Aircraft Establishment, Farnborough, UK.

Jones, J.G. (1989), "Statistical discrete gust method for predicting aircraft loads and dynamic response", *J. Aircraft* , **26**(4), 382-392.

Le Blanc, F. (2010), "Rescuing old meteorological data", *Weather*, **65**, 277-280.

Lombardo, F.T., Main, J.A. and Simiu, E. (2009), "Automated extraction and classification of thunderstorm and non-thunderstorm wind data for extreme-value analysis", *J. Wind Eng. Ind. Aerod*., **97**(3-4), 120-131.

Taylor, P.A. and Salmon, J.R. (1993), "A model for the correction of surface wind data for sheltering by upwind obstacles", *J. Appl. Meteor.*, **32**(11), 1683-1694.

JH