Wind and Structures, *Vol. 19, No. 2 (2014) 169-184*
DOI: http://dx.doi.org/10.12989/was.2014.19.2.169

An alternative method for estimation of annual extreme wind speeds

Yi Hui[*1], Qingshan Yang[2] and Zhengnong Li[1]

[1]*College of civil engineering, Hunan University, Yuelushan, Changsha, Hunan, 410082, China*
[2]*Beijing's Key Laboratory of Structural Wind Engineering and Urban Wind Environment, School of Civil Engineering, Beijing Jiaotong University, Beijing, 100044, China*

(Received February 18, 2014, Revised June 10, 2014, Accepted June 16, 2014)

Abstract. This paper presents a method of estimation of extreme wind. Assuming the extreme wind follows the Gumbel distribution, it is modeled through fitting an exponential function to the numbers of storms over different thresholds. The comparison between the estimated results with the Improved Method of Independent Storms (IMIS) shows that the proposed method gives reliable estimation of extreme wind. The proposed method also shows its advantage on the insensitiveness of estimated results to the precision of the data. The volume of extreme storms used in the estimation leads to more than 5% differences in the estimated wind speed with 50-year return period. The annual rate of independent storms is not a significant factor to the estimation.

Keywords: extreme wind speed; extreme value estimation; annual maximum; independent storms; weighted least-squares method

1. Introduction

A good estimation of extreme wind speed is essential for estimating the wind load effects on structures. It has been a long time since extreme value theory was introduced into the field of wind engineering. The extreme value theory states that sufficiently large values of independent and identically distributed variates can be described by one of extreme value distributions.

In order to make good use of the extreme wind speed, researchers tried to analyze independent sub-annual maximum wind speed data for decades. Extreme wind speed estimation based on the independent maximum values was firstly described by Jensen and Franck (1970). The essence of this method is to obtain the maximum daily wind speed and then to inspect the maxima and select only those which apply to independent storms, so that the extremes are from meteorologically uncorrelated events. Cook (1982) derived the Method of Independent Storms (MIS) for this purpose; it was then improved (Harris 1999) by introducing the theoretically derived plotting position and the corresponding variance for each ranking position. This method was further improved in subsequent studies (Cook and Harris 2003, Harris 2009, 2014). Besides the use of maximum wind speeds to estimate extreme wind speed distribution, the peaks over threshold (POT)

*Corresponding author, Assistant professor, E-mail: alihui@hnu.edu.cn

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

method is another approach that leads to the Generalized Pareto Distribution (GPD) (Lechner *et al*. 1993, Simu and Heckert 1996, Holmes and Moriaty 1999). An and Pandy (2007) provided another alternative method for the extreme wind speed estimation which is called the r-Largest Order Statistics(r-LOS) method. Karpa and Naess (2013) introduced the average conditional exceedence rate (ACER) method to estimate the extreme wind speed. Hong *et al.* (2013) discussed generalized least-squares method for the annual maximum wind speed estimation.

There have been discussions for a long time on which type of extreme distribution, Fisher-Tippett distribution type I, II or III, the extreme wind speed follows (Cheng and Yeung 2002, Harris 2004, 2005, An and Pandey 2005). It is a crucial question for clarification of the underlying mechanism of wind. However, the estimated results from the two kinds of models do not have significant differences for the 50- or 100- years return period wind speed from an engineering point of view.

In this paper, an alternative estimation method based on the Poisson Process model for extreme wind speed is introduced and applied to analyze approximately 15 years consecutive 10-min mean wind records from 5 stations in North America. The extreme wind speed is assumed to follow a Gumbel distribution. This method models the extreme wind speed by counting the number of storms over several thresholds. One thing that should be noted is that, although these places are not located in hurricane-prone regions, some other kinds of wind storms might occur at the stations, such as thunderstorms. Because of lack of detailed climate data, such storms cannot be specified and removed from the data. Thus, the estimated results might possibly be distorted in some level due to such unknown mixed climate. The idea of events over threshold has been used to estimate extreme rainfall (Buishand 1984) showing great advantage for the extreme value estimation of correlated sequence data. The feasibility of the method will be discussed below with analysis on the wind speed data.

The paper is organized as follows. The theoretical background and the methodology are presented in Section 2. Estimated results from data recorded at 5 stations are compared with results from the IMIS and POT methods for validation in Section 3. The effects of threshold are discussed in Section 4. Section 5 discusses the data precision effects on the estimated results in this method. The effects of the annual rate of storms on the estimation results are checked in Section 6. Section 7 summarizes the finding of this paper.

2. Method introduction

2.1 Basic theory

For a general random process with $F(x)$ as the distribution function, "n" (x_1, x_2, x_3 ,…,x_n) values are obtained independently from this random process. The probability of the maxima X_n of all the n values smaller than a value z is expressed as

$$P(X_n < z) = P(x_1 < z, x_2 < z, ..., x_n < z) = F^n(z) \tag{1}$$

and if $F(x)$ is close to unity, the following approximation exists as

$$\ln F(x) \approx -[1 - F(x)] \tag{2}$$

Therefore

$$F^n(x) \approx e^{-n[1-F(x)]} \tag{3}$$

where $n[1\text{-}F(x)]$ is the expected number of x_i (i=1, 2,…, n) that exceeds the threshold x. Let

$$a(x) = n[1 - F(x)] \tag{4}$$

and if the function $F(\cdot)$ is assumed to be an exponential distribution as

$$F(z) = 1 - e^{-z} \tag{5}$$

where z can be a function of a variable x, $z=h(x)$, then two parameters related to x are the location parameter μ and scale parameter α. Let z be defined as

$$z = \alpha(x - \mu) \tag{6}$$

Parameter μ is called the "characteristic largest value" of z, expressed as $\breve{z}$ which satisfies

$$n[1 - F(\breve{z})] = 1 \tag{7}$$

This equation indicates that for a sample of data x with volume n, there is one value of x which is expected to be greater than μ, basing on Eq. (3). So that

$$\mu = \ln(n) \tag{8}$$

The dimensions of α are the inverse of those of x, and α can be solved according to Harris(1982)

$$\alpha = f(z)/[1 - F(z)] = 1 \tag{9}$$

Therefore, z can be expressed as

$$z = x - \ln(n) \tag{10}$$

and the following equation can be obtained for large n as

$$F^n(x) = (1 - e^{-x-\ln(n)})^n = (1 - \frac{e^{-x}}{n})^n = \exp(-\exp(-x)) \tag{11}$$

This equation means that if $F(x)$ follows an exponential type distribution in the tail with a large enough n, $F^n(x)$ follows the Gumbel distribution. For a Gumbel distribution, the non-dimensional variable x on the right hand side of Eq. (11) can be replaced by a function $\alpha_v(V - \mu_v)$, where μ_v is the location parameter and α_v is the scale parameter related to a specific variable V.

2.2 Application to extreme wind speed

For extreme value analysis of long-term consecutive 10 min or 1 hour mean wind speed data, the formulas discussed in previous session 2.1 cannot be directly applied. This is because the consecutive wind speed data are highly correlated with each other. This means that data within a very long period are not independent with each other. Only storms that are separated for a time duration longer than a specific period can be treated as independent storms. $a(u)$ is defined as the number of independent storms whose wind speeds exceed a certain value u, which is expressed as

$$a(u) = r \times Q(D) \tag{12}$$

where r is the number of total independent wind storms within a certain time, normally taken as one year in wind engineering. Therefore, r is also called as the annual rate of storms. $Q(D)$ is the probability of event D occurring, where D indicates storms whose maximum wind speed are higher than u--expressed as

$$D = \{U_i \ge u, U_{i+1} < u, ..., U_{i+m} < u\} \tag{13}$$

and m is the least interval to separate independent storms. If any U_{i+n} occurs with $n<m$, the U_i and U_{i+n} are assumed to be from the same storm and are therefore not independent.

If r is large, the probability of the maximum wind speed $\hat{U} < u$ during a period, e.g., one year, is

$$\mathrm{P}(\hat{U} < u) \approx e^{-rQ(D)} = e^{-a(u)} \tag{14}$$

by assuming the wind speed distribution follows an exponential distribution at the upper tail region. Considering Eq. (11)

$$a(u) = e^{-\alpha_a(u-\mu_a)} \tag{15}$$

where α_a and μ_a are the scale parameter and location parameter of the Gumbel distribution. $a(u)$ can be obtained directly from the wind speed data by counting the number of storms D for different threshold u_i (i=1, 2,…, N) with u_1, the lowest value of threshold and u_N, the highest value of threshold. If the observation lasts for M years, and the total number of storms D during the M years are $b(u_i)$, $a(u_i)$ can be calculated by

$$a(u_i) = b(u_i)/M \tag{16}$$

Consequently, parameters α_a and μ_a can be estimated through the least mean squares method.

2.3 Variance of number of storms over threshold

Theoretically, the exact probability density for the number of storms D is given by

$$\phi(u) = \binom{r}{k} F(u)^{r-k}[1-F(u)]^k \tag{17}$$

where r is the annual rate of storms. The expected value of occurrence time $a(u)$ of storm D is

$$E(a(u)) = r[1-F(u)] \tag{18}$$

and the variance of occurrence time $a(u)$ of storm D is

$$\sigma(a(u))^2 = r \cdot F(u) \cdot [1-F(u)] \tag{19}$$

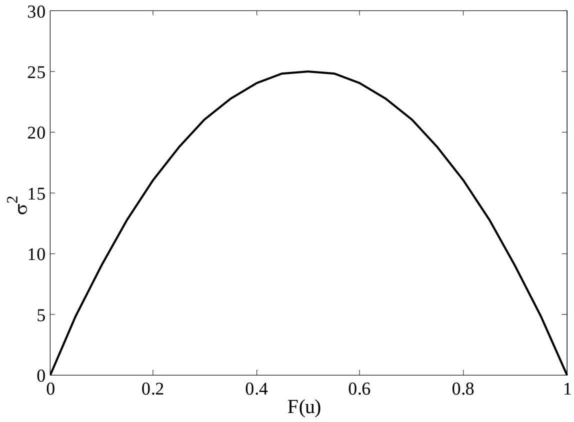


Fig. 1 Variance changes with probability

From the equations above, it can be seen that the variance σ^2 is a function of $F(u)$ and r. The relation of σ^2 and $F(u)$ is shown in Fig. 1 when r equals 100. This number is generally agreed to be a typical annual rate of independent wind storms.

If the threshold u is large, it will have $F(u)$ close to unity. It can be seen from Fig. 1 that while $F(u)$ approaches unity, the variance decreases sharply accordingly. This indicates that a lower threshold u is associated with larger variability of $a(u)$.

2.4 Fitting process

The two parameters α_a and μ_a can be directly obtained according to Eq. (15) by fitting the function

$$\ln a(u) = -\alpha_a (u - \mu_a) \tag{21}$$

with a straight line (Buishand 1984). However, this curve fitting ignores the variance $a(u_i)$ associated with different thresholds. Therefore, the idea of weighted least-squares technique introduced by Harris (1996) is adopted in this process. Weights w_i (i=1, 2,…, N) are introduced to compensate for the systematic change of variance of $a(u_i)$ corresponding to each threshold u_i. The problem then becomes that of finding values α_a and μ_a which minimize the value S; the smallest value of S indicates the best fit of the curve to the observed data.

$$S = \sum_{i=1}^{N} w_i [a(u_i) - e^{-\alpha_a (u_i - \mu_a)}]^2 \tag{22}$$

and the weights need to satisfy

$$\sum_{i=1}^{N} w_i = 1.0 \tag{23}$$

The choice of these weights is to make the variance of $w_i[a(u_i) - e^{-\alpha_a(u_i-\mu_a)}]^2$ constant and independent of threshold u_i. Therefore, the weight w_i is expressed as

$$w_i = \frac{1/\sigma_i^2}{\sum_{i=1}^{N} 1/\sigma_i^2} \tag{24}$$

According to Eq. (11), the probability distribution of the independent wind speed in the tail part can be expressed as

$$F(u) = 1 - \frac{e^{-\alpha_a(u-\mu_a)}}{r} \tag{25}$$

Substitute $F(u)$ in Eq. (19) with Eq. (25), we have

$$\sigma(u)^2 = \left(1 - \frac{e^{-\alpha_a(u-\mu_a)}}{r}\right) \cdot e^{-\alpha_a(u-\mu_a)} \tag{26}$$

The variance is noted to decrease with an increase in $F(u)$. In other words, it will decrease while the wind speed u is increasing. Therefore, it is easy to know that the weight corresponds to the lower threshold is smaller. This can be explained further that a larger variance indicates a greater probability of errors. If a smaller weight is associated with this term, such errors will be reduced.

It can also be seen that the variance is also a function of both parameters α_a and μ_a in Eq. (26). Therefore, they cannot be estimated directly in explicit expression by solving $\frac{\partial S}{\partial \alpha_a} = 0$ and $\frac{\partial S}{\partial \mu_a} = 0$ of Eq. (22). A curve-fitting method is therefore performed with the help of the MATLAB toolbox with an iterative fitting process being adopted as follows

1) Assuming all the weights are identical, $w_{i1} = 1/N$.

2) Fit the exponential curve and get the initial estimates of α_{a1} and μ_{a1}.

3) Calculate w_{i2} based on α_{a1} and μ_{a1}.

4) Repeat Steps 1 to 3 for an accurate estimation until

$\alpha_{ai} - \alpha_{a(i-1)} = 0.01\alpha_{a(i-1)}$, and $\mu_{ai} - \mu_{a(i-1)} = 0.01\mu_{a(i-1)}$

The estimations of the two parameters then become $\hat{\alpha}_a = \alpha_{ai}$ and $\hat{\mu}_a = \mu_{ai}$.

3. Results and discussion

Harris(1999) proposed the well known Improved Method of Independent Storms(IMIS) for estimating extreme wind speeds. This method was improved by Harris (2009) to reduce the error introduced by the asymptotic convergence; it is still considered as a benchmark for the extreme wind speed estimation method. Many studies have been conducted on this method. The top U largest storms out of N independent wind storms in R years observations are selected for estimation. The storm maxima(v) of these selected storms are pre-processed to v^2 first. The reason for adopting v^2 in estimation is that the probability distribution of square of the wind speed is assumed to follow more closely the exponential distribution in the upper tail region than wind speed itself. However, in this paper, for the sake of convenience, the pre-processing of wind speed is avoided. The scale and location parameters (α_a and μ_a) are estimated based on the following equations

$$\hat{\alpha}_a = \frac{\sum_{v=1}^{U} w_v \bar{y}_v v - (\sum_{v=1}^{U} w_v \bar{y}_v)(\sum_{v=1}^{U} w_v v)}{\sum_{v=1}^{U} w_v v^2 - (\sum_{v=1}^{U} w_v v)^2} \tag{27}$$

$$\hat{\mu}_a = \sum_{v=1}^{U} w_v v - (\sum_{v=1}^{U} w_v \bar{y}_v) / \hat{\alpha}_a, \tag{28}$$

where w_v is the weight, v is wind speed, and y_v is the plotting position which is derived and introduced into the estimation procedure.

Holmes and Moriarty (1999) applied the POT method to estimate the annual extreme wind speed distribution and the GPD was used for the modeling. The estimation function for the scale parameter α_a and shape parameter k are related basing on the following equation

$$E(Y-u|Y>u) = [\frac{1}{\alpha_a} - k(u-u_0)] \Big/ (1+k) \tag{29}$$

where u_0 is the lowest threshold, and $E(\cdot)$ is the expectation operation. This equation indicates that if the mean of observed excess value over u is plotted against (u-u_0), the plot should follow a straight line with slope $k/(1+k)$ and an intercept $1/[\alpha_a \cdot (1+k)]$ from which α_a and k can be determined. The parameters of GEV distribution can then be obtained. However, we just assume the annual extreme wind speed follows the Type I (Gumbel) distribution in this study as discussed above, so the shape parameter k is set to be zero. The R-year return period wind speed V_R can then be obtained as

$$V_R = u_0 + \sigma \ln(\lambda R) \tag{30}$$

where λ is the annual exceedence rate of the threshold u_0. Consequently, the location parameter μ_a can be obtained as

$$\hat{\mu}_a = \mu_0 + \sigma \ln \lambda \tag{31}$$

The results estimated by the proposed method were compared to the results from the IMIS method and POT methods. Approximately 15 years consecutive 10 min mean wind speeds recorded at 5 sites in the United States, in areas not subjected to mature hurricane winds, are used for the analysis. These data were published by the National Oceanic and Atmospheric Administration's National Data Buoy Center. T able 1 also shows the estimated scale and location parameters based on the three methods for the station at Sheboygan, Wisconsin. From the estimated 50-year return period wind speed V_{50} shown in the last column of Table 1, it can be seen that the estimated results of IMIS method and the proposed method are the same, but the POT method gives somewhat different results. There are, in total, 1006 independent storm maxima counted at this station and the top 26 of them are used. Table 2 gives details of the estimation at the Sheboygan city by IMIS method. The plotting position and weight was obtained by adopting

Cook's bootstrapping method (Cook 2004). For the POT method, the lowest threshold u_0 was set to 19.0 m/s, and the interval of each threshold was 0.5 m/s. The value λ was 2.29 (32 events in 14 years). Fig. 2 shows the observed average storm numbers above each threshold per year as well as the fitted line of the proposed method. The distance between each threshold is also 0.5 m/s.

Table 1 Estimated results at Sheboygan, Wisconsin (1996-1999 and 2001-2010)

Method	α_a(s/m)	μ_a(m/s)	V_{50}(m/s)
IMIS	0.54	20.2	27.4
POT	0.64	20.3	26.4
Proposed	0.55	20.2	27.4

Note: U is 26 for IMIS method; The lowest threshold is 19 m/s for POT and Proposed method

Table 2 Details of analysis at Sheboygan by IMIS method

Rank	Wind speed (m/s)	Plotting position	Weight
1	25.4	3.275	0.0018
2	24.6	2.278	0.0045
3	23.8	1.782	0.0073
4	22.8	1.444	0.0103
5	22.5	1.193	0.0133
6	22.4	0.994	0.0165
7	21.6	0.827	0.0196
8	21.2	0.686	0.0223
9	20.8	0.560	0.0255
10	20.6	0.450	0.0283
11	20.6	0.350	0.0311
12	20.5	0.258	0.0338
13	20.3	0.174	0.0367
14	20.1	0.096	0.0396
15	20.0	0.025	0.0421
16	19.8	-0.042	0.0451
17	19.8	-0.105	0.0480
18	19.7	-0.164	0.0516
19	19.6	-0.220	0.0550
20	19.5	-0.272	0.0578
21	19.5	-0.322	0.0610
22	19.5	-0.370	0.0642
23	19.4	-0.417	0.0669
24	19.4	-0.461	0.0694
25	19.3	-0.504	0.0725
26	19.3	-0.545	0.0756

Tables 3 to 6 give the estimated V_{50} and model parameters based on records from the other four stations. Comparing the two parameters extreme wind speed models for the five stations, the location parameter μ is relatively insensitive to the methods. The estimated μ_a from all the three methods have similar values, but the scale parameter α_a has a relatively larger variation. It can also be seen from the tables that the proposed method has much closer results to the IMIS method, but the results from the POT method have larger differences. The largest difference between results from the proposed method and the IMIS method is: 6.8% for α_a at Dunkirk station, 0.5% for μ_a at Sheboygan station, and 1.8% for V_{50} at Dunkirk station. The largest difference between results from the POT method and the IMIS method is: 51.6% for α_a at South Bass Island, 1.7% for μ_a at South Bass Island, and 8.9% for V_{50} at Dunkirk station. Therefore, it is concluded that the proposed method is a suitable one to estimate the extreme wind speed based on the Gumbel model.

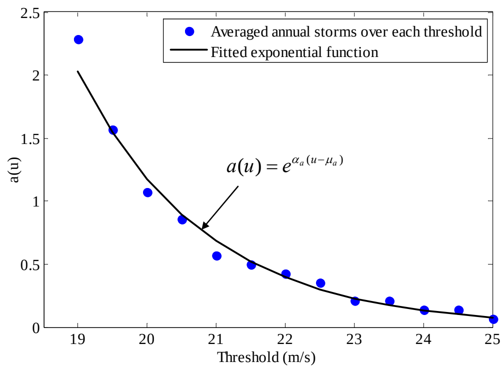


Fig. 2 Storms above each threshold and the fitted exponential function

Table 3 Estimated results at Passage Island, MI (1997-2012)

Method	α_a(s/m)	μ_a(m/s)	V_{50}(m/s)
IMIS	0.61	23.7	30.1
POT	0.64	23.8	29.9
Proposed	0.56	23.5	30.5

Note: U is 26 for IMIS method; The lowest threshold is 23m/s for POT and Proposed method

Table 4 Estimated results at South Bass Island, OH (1996-2012)

Method	α_a(s/m)	μ_a(m/s)	V_{50}(m/s)
IMIS	0.60	23.4	29.9
POT	0.91	23.0	27.3
Proposed	0.59	23.4	30.0

Note: U is 26 for IMIS method; The lowest threshold is 22.5m/s for POT and Proposed method

Table 5 Estimated results at Dunkirk, New York (1996-2012)

Method	α_a(s/m)	μ_a(m/s)	V_{50}(m/s)
IMIS	0.44	24.7	33.6
POT	0.65	24.5	30.5
Proposed	0.47	24.7	33.0

Note: U is 26 for IMIS method; The lowest threshold is 23.5m/s for POT and Proposed method

Table 6 Estimated results at Potato Point, Alaska (1996-2012)

Method	α_a(s/m)	μ_a(m/s)	V_{50}(m/s)
IMIS	0.55	22.8	29.9
POT	0.64	22.6	28.7
Proposed	0.56	22.8	29.8

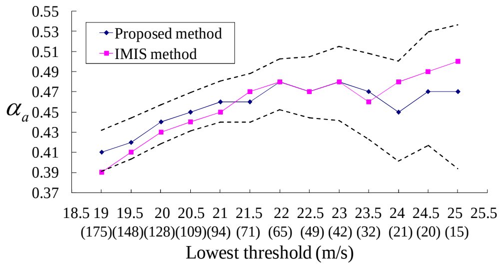
Note: U is 26 for IMIS method; The lowest threshold is 22m/s for POT and Proposed method

4. Effects of lowest threshold to the estimation results

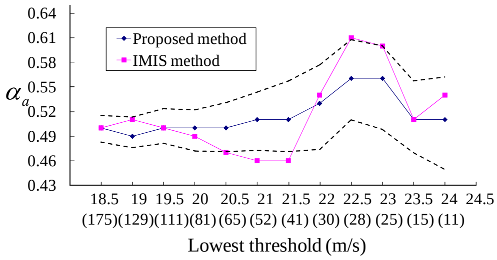
There is always a contradiction between the limited observation data and the desire of researchers to have as much as possible data for the estimation. Many methods have been developed to improve the traditional Gumbel method using one maximum from each sample for the estimation for this purpose, like MIS, IMIS, POT, *et al*. However, all these methods share a common question, that is, how much data should be selected from a fixed volume of parent data for the estimation? A small data size cannot provide a solid base for estimation. However, if too much data are selected from the limited observations, the estimated model might be unsuitable as a representation of the real extreme value distribution, as the selected data are not really “extreme values”. In this section, the effects of the lowest threshold are discussed and the effect of the amount of data is also checked.

The average number of storms selected for estimation varies from less than 1 storm in a year to around 10 storms/year. The variation of estimated results with different average number of storms per year was studied. Fig. 3 shows that the estimated parameter α_a changes with the lowest threshold for each station. The values in bracket under each lowest threshold wind speed are the number of storms with maximum speed greater than that lowest threshold value. Therefore, the variation of α_a with the number of storms selected for estimation can also be checked in Fig. 3. The maximum wind speed of each of the selected storms has also been applied to the IMIS method for estimation in this paper. It can be seen that the change of estimated α_a at each station from both methods have similar trend and most of the corresponded values are quite close to each other. The results, shown in Fig. 3, also show that the variation of α_a estimated by the proposed method are relatively weak compared to the α_a estimated by the IMIS method. This is due to the fact that in the proposed method, a lower threshold will be associated with a smaller weight. Therefore, a lower threshold does not have a strong effect on the estimated results. The 95% confidence limits of estimated α_a by the proposed method are also shown in Fig. 3 (plotted in black dashed lines). It can be seen that estimated results by IMIS method are almost all located within the range of 95% of the proposed method. It means that the results estimated by the two methods are compatible. It can also be observed that width of confidence limits of all cases gradually become narrower when the number of storms used by the estimation process become larger. However, when the average annual storms (number of storms divided by the observation years) used in the estimation process are greater than about 4-5, the width of the confidence limits become relatively constant. The estimated μ_a changes with the lowest threshold of both methods as shown in Fig. 4. It can be seen that the estimated μ_a are relatively more stable than α_a. The estimated results from both methods are also quite close to each other for most of the cases. It can be concluded that the results obtained by IMIS and the proposed method have similar stability. Confidence limits of the estimated μ_a are also drawn in Fig. 4, and indicated by dashed lines. It can be checked that the estimated results by the two methods are also quite compatible with each other. What is different from the results of α_a is that the width of confidence limits for μ_a stay relatively constant with the change of the lowest threshold.

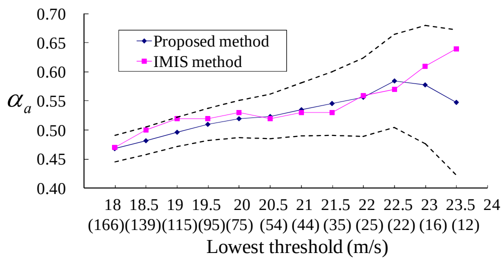
Table 7 gives the maxima and minima of estimated 50 year return period wind speeds of each station, based on the two parameters estimated by the proposed method, with different lowest threshold values. Differences up to 7% are noted, but for most of the cases the differences are less than 5%. These results indicate that the proposed method is stable with the change of amount of data. Also based on the maxima and minima of the estimated values shown in Table 7, the range of the variation of the 50 year return period wind speed can be approximately estimated. The mean values of all estimated speeds at each station are also given in Table 7. If the 95% confidence limits of the estimated extreme wind speeds are of interest, the results shown in Figs. 3 and 4 can be used for this purpose.



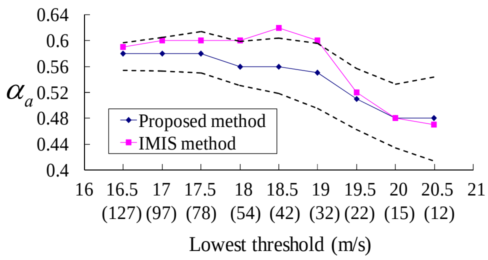
(a) Dunkirk, New York



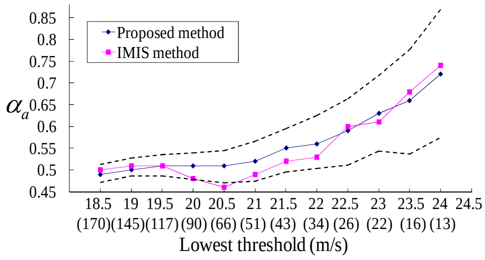
(b) Passage Island, Michigan



(c) Potato Point, Alaska

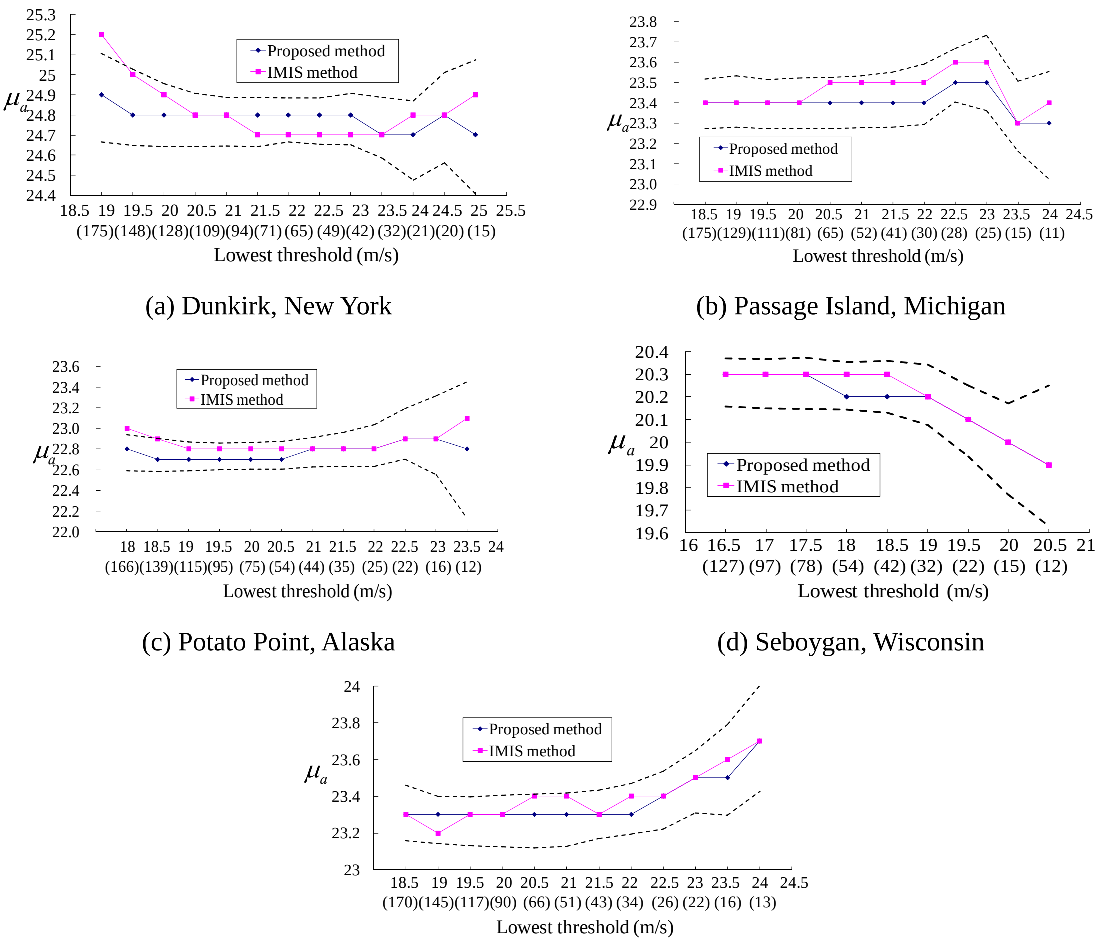


(d) Seboygan, Wisconsin



(e) South Bass Island, Ohio

Fig. 3 Parameter α_a changes with the lowest threshold value



(a) Dunkirk, New York

(b) Passage Island, Michigan

(c) Potato Point, Alaska

(d) Seboygan, Wisconsin

(e) South Bass Island, Ohio

Fig. 4 Parameter μ_a changes with the lowest threshold value

Table 7 The maxima and minima of the estimated 50 year return period wind

Station	Dunkirk	Passage Island	Potato Point	Sheboygan	South Bass Island
Max V_{50} (m/s)	34.4	31.2	31.4	28.1	31.3
Min V_{50} (m/s)	33.0	30.5	29.1	27.0	29.1
Difference	4.1%	2.2%	5.4%	3.9%	7%
Mean (m/s)	33.4	31.0	30.3	27.4	30.4

5. Sensitivity to precision of collected data

As discussed above, the major difference of the proposed method from the other methods is that it does not use the wind speed values directly in the estimation process. Only the number of storms whose maximum wind speeds are greater than each specified threshold are counted. Therefore, the sensitivity of this proposed method to the precision of the recorded data has been checked, in order to investigate the robustness of this method

To do this the estimated results based on the original recorded data and the data that are rounded to the nearest integer (less precise) are compared (shown in Table 8). For reference, the results by IMIS method are also provided (Table 9). The results shown in these two tables are based on the same data set for each station, for which approximately 4-5 storms/year are used. From 500-year return period extreme wind speeds shown in Table 8, it can be seen that the results from the original and rounded data are quite close — the largest difference is 1.0%. By the IMIS method the difference of the results from two kinds of data are relatively larger — the largest difference is about 3.2%, although it is also tolerable from engineering point of view. This is because α_a obtained by the proposed method are less sensitive to the precision of data than that by IMIS method. However, μ_a obtained by the proposed method is more sensitive to the precision of data. Since for the estimation of extreme values, especially long return period extreme values, the scale parameter α_a is more important, the results estimated by the proposed method do not vary a lot.

Table 8 Estimated results based on original and rounded data by proposed method

	Station	Dunkirk	Passage Island	Potato Point	Sheboygan	South Bass Island
α_a (s/m)	Original	0.46	0.50	0.52	0.58	0.51
	Rounded	0.46	0.52	0.53	0.58	0.53
	Difference	0%	4%	1.9%	0%	3.9%
μ_a (m/s)	Original	24.8	23.4	22.7	20.3	23.3
	Rounded	25.1	23.7	23.1	20.6	23.5
	Difference	1.2%	1.3%	1.8%	1.5%	0.9%
V_{500} (m/s)	Original	38.3	35.8	34.6	31.0	35.5
	Rounded	38.6	35.6	34.8	31.3	35.2
	Difference	0.8%	0.5%	0.5%	1.0%	0.7%

Table 9 Estimated results based on original and rounded data by IMIS method

	Station	Dunkirk	Passage Island	Potato Point	Sheboygan	South Bass Island
α_a (s/m)	Original	0.47	0.47	0.52	0.60	0.46
	Rounded	0.44	0.48	0.51	0.61	0.50
	Difference	6.4%	2.1%	1.9%	1.7%	8.7%
μ_a (m/s)	Original	24.7	23.5	22.8	20.3	23.4
	Rounded	25.0	23.6	23.0	20.4	23.4
	Difference	1.2%	0.4%	0.9%	0.4%	0%
V_{500} (m/s)	Original	37.9	36.7	34.7	30.7	36.9
	Rounded	39.1	36.5	35.2	30.6	35.8
	Difference	3.2%	0.5%	1.2%	0.2%	3.0%

6. Sensitivity to annual rate of independent storms

Although there is general agreement that the annual rate of independent storms for many locations in temperate climates is around 100 (Cook 1982, Gumley and Wood 1982), because of the four day peak appearing at the macro-meteorological wind speed spectrum. There is no well established method or algorithm to calculate the exact number of annual independent storms. Some research (Harris 2014) has been made to calculate the annual rate for a better estimation of extreme wind speed, as such a rate is also one of the factors which directly affect the number of storms being counted in the estimation procedures. Thus, this section will discuss its effect on the estimation results. Cook and Harris (2004) have also discussed theoretically that IMIS method shows the insensitivity of that method. The sensitivity to annual rate of storms based on the actual data is described in this paper.

Table 10 gives the estimated results corresponding to each averaged annual rate of independent storms of the Dunkirk station. The smallest interval time (the value “m” discussed in Section 2) is the period of time to obtain the corresponding number of independent storms set in the algorithm of estimation. Such a period is to make sure that each selected storm is separated sufficiently from the next one. It is known that the criterion of independence mainly affects the counting of storms with relatively smaller maximum speed. Two lowest thresholds were adopted for the proposed method in this case, and the effects of annual rate were checked respectively.

It can be seen from Table 10 that the differences of annual rate has a lesser effect on the estimated results than the change of lowest threshold. While the annual rate changes from 60 to 200, the parameter α_a only changes from 0.40 to 0.42, if a lowest threshold of 19 m/s is used. The results also indicate that with an increase in the annual rate, the rate of increment of α_a and V_{50} tend to reduce. This means that if the annual rate stays between 100 and 200, the estimated results just change a little. It can also be seen that if the lowest threshold is set to be larger at 22 m/s, the effects of the annual rate becomes less. While the annual rate changes from 60 to 200, parameter α_a just changes from 0.474 to 0.480. When the annual rate stays between 100 and 200, the estimated results stay relatively constant. The parameter μ_a is more or less constant with the change of annual rate in both cases.

Table 10 Results for the Dunkirk station - changes with annual rate of independent storms

Annual rate		60	80	100	120	140	160	180	200
Least interval (10 min)		400	290	220	160	140	120	100	90
Lowest threshold 19 m/s	α_a	0.401	0.406	0.411	0.414	0.415	0. 418	0.421	0.422
	μ_a	24.9	24.9	24.9	24.9	24.9	24.9	24.9	24.9
	V_{50}(m/s)	34.65	34.53	34.41	34.34	34.32	34.25	34.19	34.17
Lowest threshold 22 m/s	α_a	0.474	0.476	0.478	0.480	0.480	0.480	0.480	0.480
	μ_a	24.8	24.8	24.8	24.8	24.8	24.8	24.8	24.8
	V_{50}(m/s)	33.05	33.01	32.98	32.96	32.96	32.96	32.96	32.96

7. Conclusions

An alternative method, for estimation of annual extreme wind speeds, is introduced in this paper. By comparisons of results from the IMIS and POT methods, the proposed method is shown to be capable to obtained estimated results that are close to those from the IMIS method. This method also appears to have better reliability for extreme wind speed estimation than the POT method for the Fisher-Tippet I (Gumbel) model. The above conclusions are based on analysis of approximately 15 years wind speeds records at 5 locations in North America.

By checking the effects of the lowest threshold to the estimation, the scale parameter is shown relatively sensitive to the selection of lowest threshold, but the location parameter is not. These two parameters estimated by the proposed method are all close to those from the IMIS method, which shows that this method can be one of the choices for extreme wind speed estimation.

The results show that the proposed method is relatively more robust to the precision of the recorded data than the existing methods, especially for the estimation of long return period extreme wind speed.

The results also show that the annual rate of independent storms does not have a strong effect on the estimated results. With an increase in the annual rate, its effect on the estimated parameters becomes weaker. When the annual rate varies between 100 and 200, the estimated extreme wind speed is almost the same.

Acknowledgments

This study is supported by the young researcher cultivating program (531107040741) of Hunan University. It is also supported by the 111 project (B13002) and the National Natural Science Foundation of China (91215302). The authors give special thanks to Prof. Xinzhong Chen and Prof. S.S. Law for their great help.

References

An, Y. and Pandey, M.D. (2005), "A comparison of methods of extreme wind speed estimation", *J. Wind Eng. Ind. Aerod*., **93**(7), 535-545.

An, Y. and Pandey, M.D. (2007), "The *r* largest order statistics model for extreme wind speed estimation", *J. Wind Eng. Ind. Aerod.*, **95**(3), 165-182.

Buishand, T.A. (1985), "The effect of Seasonal variation and serial correlation on extreme value distribution of rainfall data", *J. Clim. Appl. Meteorol.*, **24**, 154-160.

Chenga, E. and Yeung, C. (2002), "Generalized extreme gust wind speeds distributions", *J. Wind Eng. Ind. Aerod.*, **90**(12-15), 1657-1669.

Cook, N.J. (1982), "Towards better estimation of extreme winds", *J. Wind Eng. Ind. Aerod.*, **9**(3), 295-323.

Cook, N.J. (2004), "Confidence limits for extreme wind speeds in mixed climates", *J. Wind Eng. Ind. Aerod*., **92**(1), 41-51.

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

Cook, N.J. and Harris, R.I. (2004), "Exact and general FT1 penultimate distributions of extreme wind speeds drawn from tail-equivalent Weibull parents", *Struct. Saf.*, **26**(4), 391-420.

Gumley, S.J. and Wood C.J. (1982), "A discussion of extreme wind-loading probabilities", *J. Wind Eng. Ind. Aerod*., **10**(1), 31-45.
Harris, R.I. (1982), "An improved method for the prediction of extreme values of extreme values of wind effects on simple buildings and structures", *J. Wind Eng. Ind. Aerod.*, **9**(3), 343-379.
Harris, R.I. (1996), "Gumbel re-visited – a new look at extreme value statistics applied to wind speeds", *J. Wind Eng. Ind. Aerod.*, **59**(1), 1-22.
Harris, R.I. (1999), "Improvements to the method of independent storms", *J. Wind Eng. Ind. Aerod.*, **80**(1-2), 1-30.
Harris, R.I. (2004), "Extreme value analysis of epoch maxima—convergence and choice of asymptote", *J. Wind Eng. Ind. Aerod.*, **92**(11), 897-918.
Harris, R.I. (2005), "Generalised Pareto methods for wind extremes. Useful tool or mathematical mirage?", *J. Wind Eng. Ind. Aerod.*, **93**(5), 341-360.
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), 1-30.
Harris, R.I. (2014), "A simulation method for the macro-meteorological wind speed and the implications for extreme value analysis", *J. Wind Eng. Ind. Aerod.*, **125,** 146-155.
Holmes, J.D. and Moriarty, W.M. (1999), "Application of the generalized pareto distribution to wind engineering", *J. Wind Eng. Ind. Aerod.*, **83**(1-3), 1-10.
Hong, H.P., Li, S.H. and Mara, T.G. (2013), "Performance of the generalized least-squares method for the Gumbel distribution and its application to annual maximum wind speeds", *J. Wind Eng. Ind. Aerod.*, **119,** 121-132.
Jensen, M. and Franck, N. (1970), *The climate of strong winds in Denmark*, Danish Technical Press, Copenhagen.
Karpa, O. and Naess, A. (2013), "Extreme value statistics of wind speed data by the ACER method", *J. Wind Eng. Ind. Aerod.*, **112**, 1-10.
Lechner, J.A., Simiu, E. and Heckert, N.A. (1993), "Assessment of 'peaks over threshold' methods for estimating extreme value distribution tails", *Struct. Saf.*, **12**(4), 305-314.
Simiu, E. and Heckert, N.A. (1996), "Extreme wind distribution tails: a "peaks over threshold" approach", *J. Struct. Eng. - ASCE*, **122**(5), 539-547.

JH