ENVIRONMETRICS
Environmetrics 2003; **14**: 27–43 (DOI: 10.1002/env.555)

Bootstrap simulations for evaluating the uncertainty associated with peaks-over-threshold estimates of extreme wind velocity

M. D. Pandey[1], P. H. A. J. M. Van Gelder[2*,†] and J. K. Vrijling[2]

[1]*University of Waterloo, Waterloo, Canada*
[2]*TU Delft, Civil Engineering, Hydraulic and Offshore Engineering Section, P.O. Box 5048, NL-2600 GA Delft, The Netherlands*

SUMMARY

In the peaks-over-threshold (POT) method of extreme quantile estimation, the selection of a suitable threshold is critical to estimation accuracy. In practical applications, however, the threshold selection is not so obvious due to erratic variation of quantile estimates with minor changes in threshold. To address this issue, the article investigates the variation of quantile uncertainty (bias and variance) as a function of threshold using a semi-parametric bootstrap algorithm. Furthermore, the article compares the performance of L-moment and de Haan methods that are used for fitting the Pareto distribution to peak data.

The analysis of simulated and actual U.S. wind speed data illustrates that the L-moment method can lead to almost unbiased quantile estimates for certain thresholds. A threshold corresponding to minimum standard error appears to provide reasonable estimates of wind speed extremes. It is concluded that the quantification of uncertainty associated with a quantile estimate is necessary for selecting a suitable threshold and estimating the design wind speed. For this purpose, semi-parametric bootstrap method has proved to be a simple, practical and effective tool. Copyright © 2003 John Wiley & Sons, Ltd.

KEY WORDS: peaks-over-threshold method; extremes; bootstrap; simulations; uncertainty; L-moments; wind speed; return period; Pareto distribution; quantiles

1. INTRODUCTION

Extreme quantile estimates of wind velocity corresponding to return periods of the order of 50 to 1000 years are commonly utilized in the assessment of wind load-effects on civil engineering infrastructures. The peaks-over-threshold (POT) method is a widely used approach for extreme value estimation that includes several of the largest order statistics exceeding a sufficiently high threshold in the available data. A greater amount of data incorporated in this way is aimed at reducing the sampling uncertainty. The POT approach has received considerable attention since it has been shown that the Pareto distribution (GPD) arises as the limiting distribution of peaks of a random variable (Pickands, 1975).

Practical applications of the POT method are confounded by the problem of selection of a suitable threshold, which is not known a priori. This problem becomes particularly serious when quantile

*Correspondence to: P. H. A. J. M. Van Gelder, TU Delft, Civil Engineering, Hydraulic and Offshore Engineering Section, P. O. Box 5048, NL-2600 GA Delft, The Netherlands.
†E-mail: P.vangelder@ct.tudelft.nl

Received 27 March 2001
Accepted 28 November 2001

estimates exhibit large variability with respect to minor variations in threshold. A case in point is a recent study by Simiu and Heckert (1996) that illustrates an erratic variation of extreme wind speed estimated by a method proposed by de Haan (1994), which is widely used in mathematical statistics.

The article is primarily concerned with the selection of an optimal threshold considering the variation of quantile uncertainty (bias and variance) as a function of threshold. The bootstrap simulation method is commonly used to assess the variance of a statistical estimate obtained from a single random sample. However, the original non-parametric version of the bootstrap method is rarely applicable to the extreme quantile estimation problem, because the available random sample may not contain any observations in the region of tail extrapolation. To overcome this difficulty, a semi-parametric bootstrap algorithm, developed by Caers and Maes (1998), is applied in this study.

As an alternative to de Haan's method for the GPD parameter estimation, the article also examines the effectiveness of the method of L-moments, which are linear combinations of expectations of order statistics (Hosking, 1990). The use of L-moments has become quite popular in hydrology, since they exhibit much smaller sampling variability than estimates of ordinary moments. A recent study has shown that the L-moment method results in more stable quantile estimates than those obtained using the de Haan method (Pandey *et al.*, 2001a).

The main objectives of the article are: (i) to evaluate the performance of a bootstrap algorithm for quantile uncertainty estimation; and (ii) to compare statistical uncertainty associated with quantiles estimated by de Haan and L-moment methods, which has not been reported in the literature. For this purpose, the article utilizes simulated as well actual wind velocity data collected at various stations in the United States.

The article is organized as follows. The following section reviews basic concepts of the POT approach and the de Haan estimation method. The L-moment method is briefly outlined in Section 3. Section 4 describes a bootstrap method for extreme value estimation problem, and numerical results are presented in Section 5. The main conclusions and a list of references are presented in Sections 6 and 7, respectively.

2. THEORETICAL BACKGROUND

2.1. Peaks-over-threshold approach

Let $X_1, X_2, \ldots, X_n$ be a series of independent random observations of a random variable X with the distribution function (DF) $F(x)$. To model the upper tail of $F(x)$, consider k exceedances of X over a threshold u and let $Y_1, Y_2, \ldots, Y_k$ denote the excesses (or peaks), i.e. $Y_i = X_i - u$. Pickands (1975) showed that, in some asymptotic sense, the conditional distribution of excesses follows the generalized Pareto distribution. Thus the DF of $Y_i = [(X_i - u)\,|\,X_i > u]$, $i = 1, 2, \ldots, k$, is given as

$$G(y) = 1 - \left(1 + \frac{c(y-h)}{a}\right)^{-1/c} \tag{1}$$

where h, a and c denote the location, scale and shape parameters, respectively. Generally, the location parameter is taken as zero. The distribution is unbounded, i.e. $0 < y < \infty$ if $c \geq 0$ and bounded as $0 < y < a/c$ if $c < 0$. The exponential DF is a special case of (1) when $c = 0$. The GPD exhibits a unique threshold-stability property, i.e. if X follows GPD then the conditional distribution of excesses, i.e. $G(y)$, also follows GPD with the same shape parameter as that of X. It can also be shown that the distribution of maximum excesses, i.e. $Z = \max(Y_1, Y_2, \ldots, Y_k)$ follows the generalized extreme value distribution provided that exceedances over the threshold are generated from a Poisson process (Davison and Smith, 1990). These elegant mathematical properties have motivated the use of the GPD model in extreme quantile estimation.

The calculation of an R-year a quantile value, x_R, is based on the quantile of peaks, Y, corresponding to a return period of λR, where λ is the mean exceedance (or crossing) rate per year of X over u. Thus,

$$x_R = G^{-1}\left(1 - \frac{1}{\lambda R}\right) + u \tag{2}$$

where $G^{-1}(p)$ denotes the Pareto quantile function. If n denotes the number of samples collected over m years and k is the number of exceedances, then the mean crossing rate is estimated as $\lambda = k/m$. As the threshold is lowered to include more data in the inference, the crossing rate λ increases due to increasing values of k. An interesting trade-off is thus revealed from (2), i.e. the more the data considered, the farther in the tail region one has to go for quantile estimation due to increasing values of the effective return period λR. Conversely, by inclusion of additional data the accuracy of tail modelling must increase at a higher rate than the movement farther in the tail region. Otherwise, the accuracy of POT estimates would deteriorate (Pandey *et al.*, 2001a).

2.2. *Parameter estimation by de Haan method*

De Haan (1994) proposed the estimation of scale and shape parameters using the order statistics of exceedances, $\{X_{n-k,n}, \ldots, X_{n,n}\}$, where $X_{n-k,n}$ is the smallest data point to exceed a given threshold. The shape parameter of the Pareto distribution was derived as

$$c = c_1 + c_2$$

where

$$c_1 = M_n^{(1)}$$

and

$$c_2 = 1 - \frac{1}{2}\left(1 - \frac{\left(M_n^{(1)}\right)^2}{M_n^{(2)}}\right)^{-1} \tag{3}$$

in terms of moments of peaks of log-transformed exceedance data:

$$M_n^{(r)} = \frac{1}{k}\sum_{i=1}^{k}[\log(X_{n-i+1,n}) - \log(X_{n-k,n})]^r \tag{4}$$

The scale parameter was derived as

$$a = u\frac{M_n^{(1)}}{\rho} \tag{5}$$

where $\rho = 1$ if $c \geq 0$, and $\rho = 1/(1-c)$ if $c < 0$.

3. METHOD OF L-MOMENTS

3.1. Order statistics and L-moments

Using the density function of an rth order statistic, $X_{r:n}$ (Kendall and Stuart, 1977), along with a transformation $u = F(x)$, its expectation can be expressed in terms of the quantile function, $x(u)$, as

$$\mathrm{E}[X_{r:n}] = r\binom{n}{r}\int_0^1 x(u)u^{r-1}(1-u)^{n-r}\,\mathrm{d}u \tag{6}$$

Expectations of the maximum and minimum of a sample of size n can be easily obtained from (6) by setting $r = n$ and $r = 1$, respectively:

$$\mathrm{E}[X_{n:n}] = n\int_0^1 x(u)u^{n-1}\,\mathrm{d}u$$

and

$$\mathrm{E}[X_{1:n}] = n\int_0^1 x(u)(1-u)^{n-1}\,\mathrm{d}u \tag{7}$$

The probability-weighted moment (PWM) of a random variable was formally defined by Greenwood *et al.* (1979) as

$$M_{i,j,k} = \mathrm{E}[X^i u^j (1-u)^k] = \int_0^1 x(u)^i u^j (1-u)^k\,\mathrm{d}u \tag{8}$$

Two special forms of PWMs, $\alpha_k = M_{1,0,k}$ and $\beta_k = M_{1,k,0}(k = 0, 1, \ldots, n)$ are particularly simple and useful. From (7) and (8), it can be shown that α_k and β_k are directly related with expectations of the minimum and maximum, respectively, in a sample of size k, as

$$\alpha_k = \frac{1}{k}\mathrm{E}[X_{1:k}], \qquad \beta_k = \frac{1}{k}\mathrm{E}[X_{k:k}], \qquad k \geq 1 \tag{9}$$

Certain linear combinations of PWMs, referred to as L-moments, are shown to be analogous to ordinary moments in a sense that they also provide measures of location, dispersion, skewness, kurtosis, and other aspects of the shape of probability distributions or data samples (Hosking, 1990). An rth order L-moment is mathematically defined as

$$\lambda_r = \sum_{k=1}^{r} p^*_{r-1,\,k-1}\beta_k \tag{10}$$

where $p^*_{r,k}$ represents the coefficients of shifted Legendre polynomials (Hosking, 1990). From an ordered random sample of size n, unbiased estimates b_k and a_k of β_k and α_k, respectively, can be obtained as (Hosking and Wallis, 1997)

$$b_k = \frac{1}{n}\sum_{i=1}^{n}\binom{i-1}{k}X_i \Big/ \binom{n-1}{k} \quad \text{and} \quad a_k = \frac{1}{n}\sum_{i=1}^{n}\binom{n-i}{k}X_i \Big/ \binom{n-1}{k} \tag{11}$$

Normalized forms of higher order L-moments, $\tau_r = \lambda_r/\lambda_2, r = 3, 4, \ldots$, are convenient to work with due to their bounded variation, i.e. $|\tau_r| < 1$. An extensive simulation-based comparison with an information-theoretic measure, namely, the divergence, confirms that L-moments are effective in summarizing distribution properties and quantile estimation (Pandey *et al.*, 2001b).

3.2. Parameter estimation

Hosking and Wallis (1997) illustrated that L-moments are efficient in estimating parameters of a wide range of distributions. In general, the bias of small sample estimates of higher-order L-moments is fairly small as compared to traditional moment estimates. Furthermore, the computation of L-moments is very simple, requiring limited arithmetic operations only. The first three L-moments of GPD are given as

$$\lambda_1 = h + \frac{a}{(1-c)}, \qquad \lambda_2 = \frac{a}{(1-c)(2-c)} \quad \text{and} \quad \tau_3 = \frac{(1+c)}{(3-c)} \tag{12}$$

Therefore, using three sample L-moments, the location (h), scale (a) and shape (c) parameters can be estimated as

$$c = \frac{(3\tau_3 - 1)}{(\tau_3 + 1)}, \qquad a = (1-c)(2-c)\lambda_2 \quad \text{and} \quad h = \lambda_1 - (2-c)\lambda_2 \tag{13}$$

Here, c is a function of skewness only, whereas scale and location parameters depend on the mean and dispersion measures. In case the location parameter is known (typically $h = 0$), the first two L-moments can be used to estimate c and a as

$$c = 2 - \frac{(\lambda_1 - h)}{\lambda_2} \quad \text{and} \quad a = (1-c)(\lambda_1 - h) \tag{14}$$

In contrast with (12), here c depends on mean and dispersion, and ignores the sample skewness. Since a previous study has shown that the use of three L-moments results in superior performance, the two L-moments method (14) is not used in this article.

4. BOOTSTRAP SIMULATION METHOD

4.1. General

Consider that a random sample of observation, $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$, is used to obtain a sample estimate θ_s of a parameter of interest θ, which can be a quantile or some other statistic. The purpose of bootstrap simulations is to estimate the uncertainty (bias and variance) associated with the sample estimate θ_s. In the standard version of bootstrap (Efron and Tibshirani, 1993), a random sample of size n is drawn with replacement from the ordered sample $\{X_{1:n}, X_{2:n}, \ldots, X_{n:n}\}$ as

$$X_j^* = F_E^{-1}(p) = X_{[np]+1} \qquad \text{for } j = 1, n \tag{15}$$

where $F_E^{-1}(p)$ denotes the empirical (sample) quantile function, p is a uniform *rv* (0–1) and $[np]$ denotes the integer floor function. This method is also known as non-parametric bootstrap, as no parametric DF is used for simulating bootstrap samples.
Using a kth bootstrap sample, denoted by

$$\mathbf{X}^*(k) = \{X_1^*, X_2^*, \ldots, X_n^*\}, \qquad k = 1, 2, \ldots, b \tag{16}$$

a new bootstrap estimate θ_k^* of θ_s can be obtained. Here, b denotes the number of bootstrap simulations. The set of estimates, $\boldsymbol{\theta}^* = \{\theta_1^*, \theta_2^*, \ldots, \theta_b^*\}$, constitutes the sampling distribution of θ_s. The bootstrap estimate of the true bias $(\mathrm{E}[\theta_s] - \theta)$ is given as

$$\text{Bias} = (\theta_m^* - \theta_s) \tag{17}$$

where θ_m^* is the average of all bootstrap estimates $\boldsymbol{\theta}^*$. The positive bias would imply the overestimation of the sample estimate. Van Noortwijk and Van Gelder (1998) proposed to penalize overestimation (i.e. positive bias) differently from underestimation (or negative bias), depending on the consequences of a wrong estimate (for instance failure of the structure). The concept of (asymmetric) loss functions will, however, not be the subject of this article.
The variance of θ_s is estimated as

$$\text{variance} = (\text{standard error})^2 = \frac{1}{b-1} \sum_{k=1}^{b} (\theta_k^* - \theta_m^*)^2 \tag{18}$$

4.2. Semi-parametric bootstrap method

The extreme quantile estimation is essentially a statistical extrapolation beyond the existing data set. In such cases the standard non-parametric bootstrap method is not considered suitable. The reason is that a given random sample may not contain any extreme observations that are in the proximity of region of tail-extrapolation. In other words, the empirical distribution from which bootstrap samples are generated cannot provide a good approximation of the distribution tail.

In the context of the POT method, alternative procedures have been investigated that primarily rely on double (nested) bootstrap simulations based on some asymptotic distribution tail models (Caers *et al.*, 1998; Draisma *et al.*, 1999). In this method, a POT sample, chosen above a certain threshold (u_1), is further divided into two sub-samples by choosing another high threshold (u_2). For a given set of u_1 and $u_2 (u_2 > u_1)$, bootstrap estimates of first and second order tail indices are calculated. Threshold u_1 and u_2 are varied to minimize mean square error in an asymptotic sense. Mathematical complexities of such formulations and uncertainty about the selection of u_1 and u_2 make them less amenable to practical engineering applications.

To overcome these difficulties, Caers and Maes (1998) proposed a simple and effective algorithm in which a semi-parametric distribution model is used to generate bootstrap samples. Samples above a threshold (u) are generated from a smooth parametric distribution, whereas samples below u are simulated from the empirical distribution. The bootstrap distribution is thus modelled as

$$\begin{aligned} F_{\mathrm{B}}(x \mid u) &= (1 - F_{\mathrm{E}}(u))\, F_{\mathrm{P}}(x) + F_{\mathrm{E}}(u) \qquad && \text{for } x > u \\ &= F_{\mathrm{E}}(x) && \text{for } x \leq u \end{aligned} \tag{19}$$

where $F\mathrm{P}(x)$ is the Pareto distribution fitted to POT data by de Haan or L-moment method. A similar approach is also proposed by Hutson (2000). Using results of direct Monte Carlo simulation, Caers and Maes (1998) illustrated the applicability of this approach (19) to distributions with bounded tail. Here, the semi-parametric model is applied to distributions with unbounded upper tail, e.g. lognormal and Gumbel, and its validity is examined in the following section.

5. RESULTS

5.1. Simulations

The study compares quantile uncertainty estimated from semi-parametric bootstrap method against benchmark values calculated from direct Monte Carlo simulations. The comparison consists of two steps. Firstly, random samples, each of size n, were simulated from a known parent distribution. The simulated data series is assumed to represent independent maximum values collected over a period of T years. POT samples were constructed from upper k order statistics, such that the mean exceedance rate becomes $\lambda = k/T$. Using 10 000 simulated samples, quantile average, bias and standard error (SE) were evaluated for both de Haan (DHN) and L-moment (LMM) estimates. In the second step, 500 bootstrap simulations were performed on a random sample of size n generated from the same parent DF. From the bootstrap distribution of quantile, its bias, SE and average values were calculated. The bootstrap results were averaged over 250 different simulated samples of size n. These simulations were repeated for several k values.

5.1.1. Example 1. Here, data series ($n = 900$, $T = 30$ years) were simulated from a Gumbel distribution ($\mu = 30$ mph, $\sigma = 6.5$), similar to that used by Gross *et al.* (1994). The POT sample size (k) was varied from 30 to 300 (i.e. $\lambda = 1$–30), and quantile average, bias and SE were estimated for 100 and 1000-year return periods (R).

The average quantile values estimated from bootstrap (BT) and Monte Carlo (MC) simulations for varying threshold values are compared in Figures 1 and 2 for $R = 100$ and 1000 year, respectively. For the L-moment method, both bootstrap and MC estimates of 100-year quantile are close to the exact value (67.25 mph). In a sense, quantile estimates are almost unbiased for a wide range of threshold values (Figure 1). MC results suggest that the DHN underestimates the quantile value, and bootstrap estimates depart further from MCS results. In Figure 2, results for 1000-year LMM estimates reveal an interesting feature. For high threshold ($k < 150$), LMM overestimates the quantile value, which gradually decreases with lowering the threshold and results in underestimation for $k > 200$. In contrast, the DHN method consistently underestimates the quantile value.

Comparisons of standard error (SE) estimated from BT and MCS are presented in Figures 3 and 4 for $R = 100$ and 1000 year, respectively. It is seen from Figure 3 ($R = 100$) that bootstrap slightly

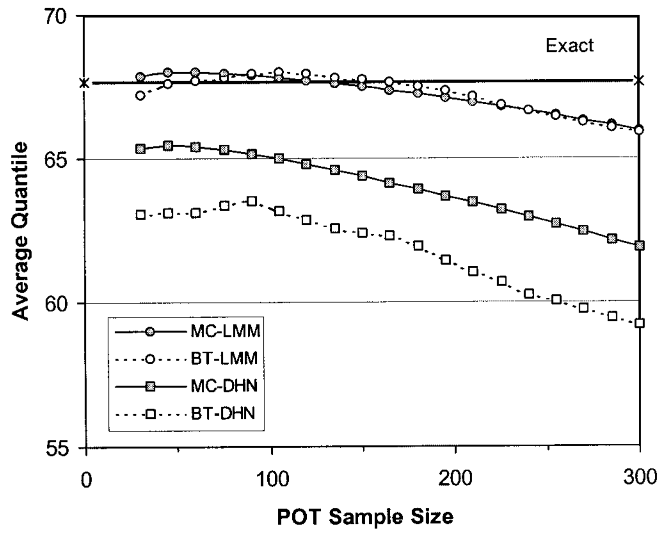

Figure 1. Comparison of Monte Carlo (MC) and bootstrap (BT) estimates of Gumbel quantiles ($R = 100$ year)

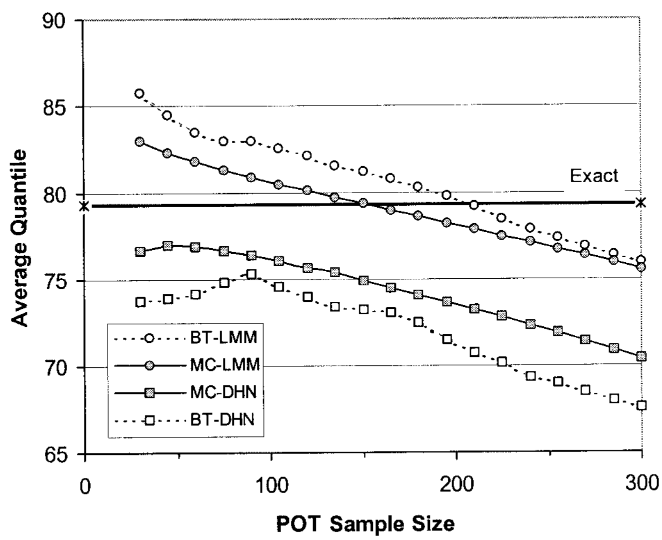

Figure 2. Comparison of Monte Carlo and bootstrap estimates of Gumbel quantiles ($R = 1000$ year)

underestimates the DHN-SE, which decreases significantly for 1000-year quantile (Figure 4). In the case of LMM with $k < 200$, SE is slightly overestimated by bootstrap, whereas an underestimation is seen for $k > 200$ (Figure 3). In the case of 1000-year LMM estimates obtained with $k < 100$, bootstrap SE is higher than that of MCS. For lower thresholds, however, they are in fairly close agreement.

5.1.2. Example 2. This example is similar to the previous one, except that data series ($n = 900$, $T = 30$ years) were simulated from a lognormal distribution ($\mu = 1$ mph, $\sigma = 0.3$). Results shown in Figures 5–8 are also qualitatively similar to the previous case. The LMM bootstrap quantile average

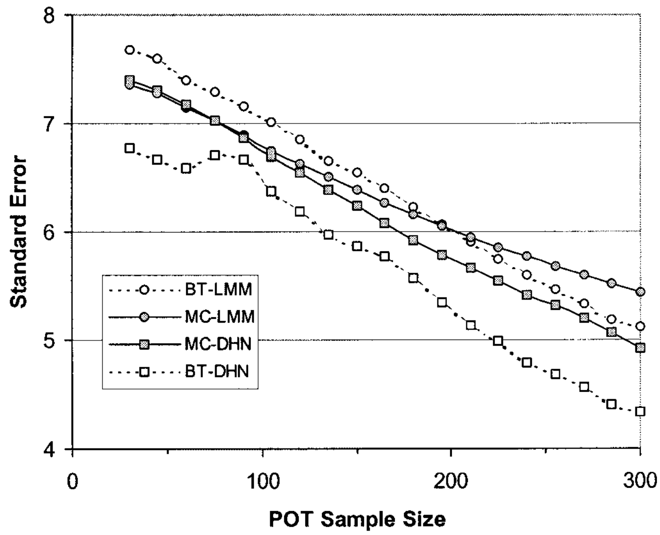

Figure 3. Comparison of standard error of Gumbel quantiles ($R = 100$ year)

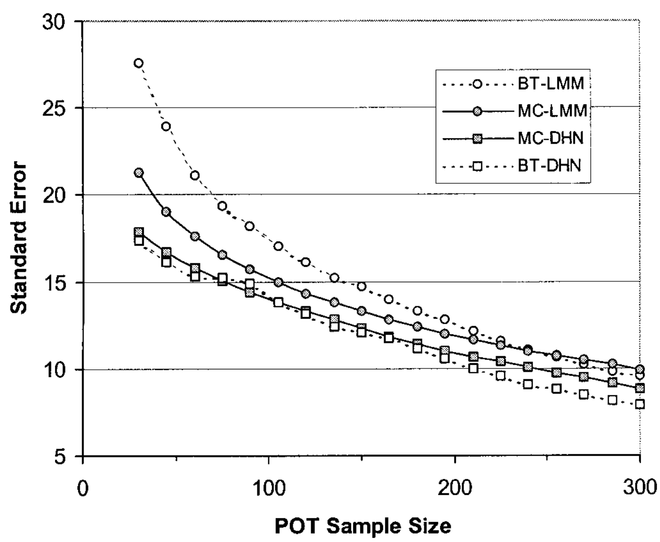

Figure 4. Comparison of standard error of Gumbel quantiles ($R = 1000$ year)

is generally close to MC results for wide ranging k. An exception to this observation is seen for $k < 100$ and $R = 1000$ year (Figure 6), where bootstrap estimates are higher than MC values. Bootstrap method tends to underestimate SE, though its magnitude is fairly small. In the case of the LMM method, a transition is seen about $k \approx 100$ from an over to underestimation of SE.

5.1.3. Remarks. DHN method consistently underestimates the quantile value with SE smaller than LMM. A transition of the bias of LMM estimates from positive to negative values is interesting in the sense that a k value resulting in an almost unbiased estimate can be determined. The bias of LMM

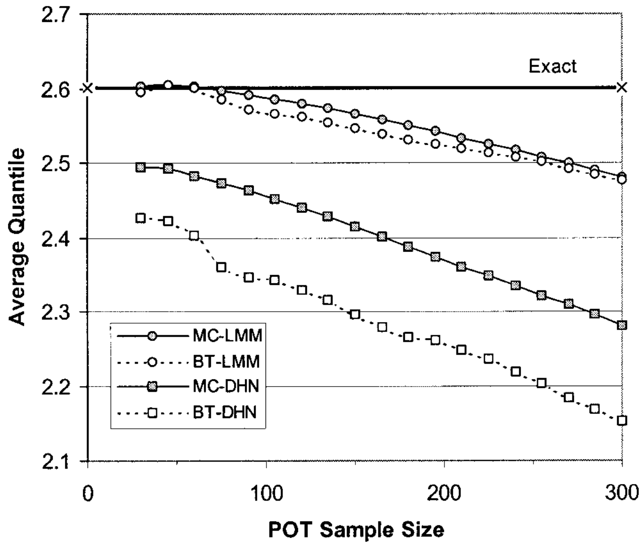

Figure 5. Comparison of Monte Carlo (MC) and bootstrap (BT) estimates of lognormal quantiles ($R = 100$ year)

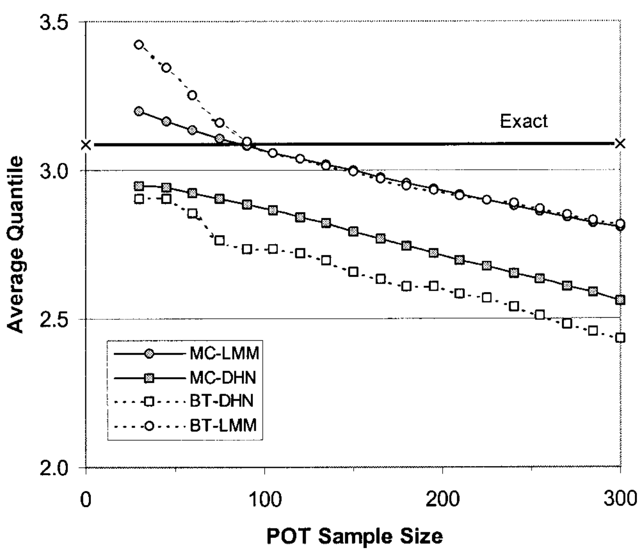

Figure 6. Comparison of Monte Carlo and bootstrap estimates of lognormal quantiles ($R = 1000$ year)

estimates is generally lower than that of DHN. For a high threshold, the SE of LMM quantiles is overestimated by the bootstrap method. For modest k values, however ($k > 100$), SEs obtained from BT and MC are in close agreement. In general, the semi-parametric bootstrap algorithm is sufficiently accurate for comparing the statistical performance of various estimation methods.

Similar simulations can be performed for parent distributions other than the Gumbel and lognormal, though the qualitative conclusions of the section are not expected to change.

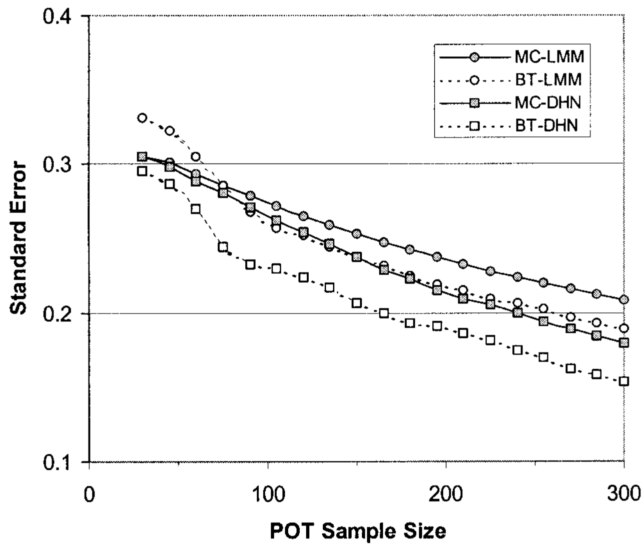

Figure 7. Comparison of standard error of lognormal quantiles ($R = 100$ year)

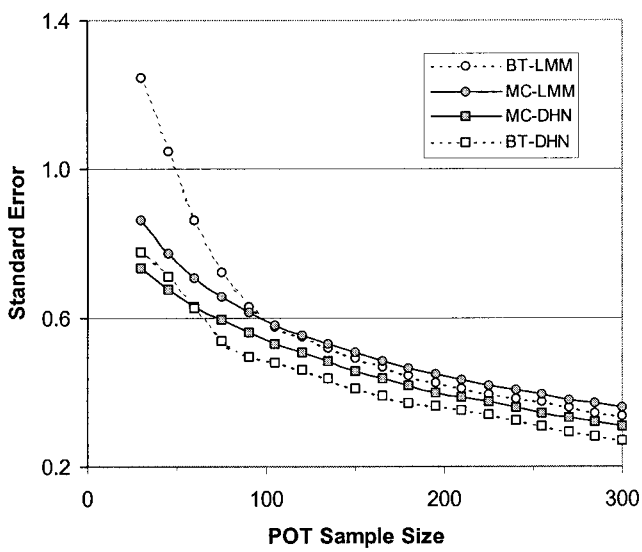

Figure 8. Comparison of standard error of lognormal quantiles ($R = 1000$ year)

5.2. Wind speed data from the U.S. stations

To compare the performance of de Haan and L-moment methods in a more realistic setting, actual wind speed data collected at various stations in the United States were analyzed. The method of L-moments was added to a computer program provided by Simiu and Heckert (1996), which generates an uncorrelated sample from daily maximum wind data and applies the de Haan method. Here, numerical results are presented for three stations, namely, Boise (ID), Moline (IL) and Toledo (OH),

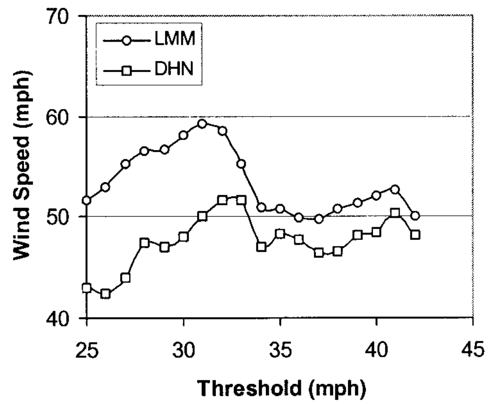

Figure 9. POT estimates of 50-year wind speed (Boise, ID)

that provide wind speed records for 23, 15 and 25 years, respectively. Bootstrap simulations consisted of 1000 iterations.

In Figure 9, POT estimates of 50 year wind speed based on Boise data are plotted for various threshold values. Note that increasing values of the threshold velocity (u) on the X-axis implies decreasing size of the POT sample. The quantile bias and standard error (SE) obtained from bootstrap simulations are shown in Figures 10 and 11, respectively. It is interesting that the bias of the L-moment estimates is fairly small, and is also insensitive to threshold value. On the other hand, quantile underestimation (i.e. negative bias) by DHN increases with lowering of the threshold (Figure 10). Based on LMM results, 50-year wind speed can be estimated as 50 mph corresponding to a minimum SE threshold of 37 mph (Figure 11). The DHN estimate based on a similar criterion is 46 mph for $u = 38$ mph. Figures 12–14 display results for 1000-year wind speed quantile with more pronounced sensitivity to threshold. LMM estimates have minimum bias and SE for $u = 35$–38 mph. It is noted that sharp peaks in the POT plot of LMM estimates (Figure 12) are also associated with large bias and SE in Figures 13 and 14, respectively. The bootstrap method, in this sense, can help to exclude such

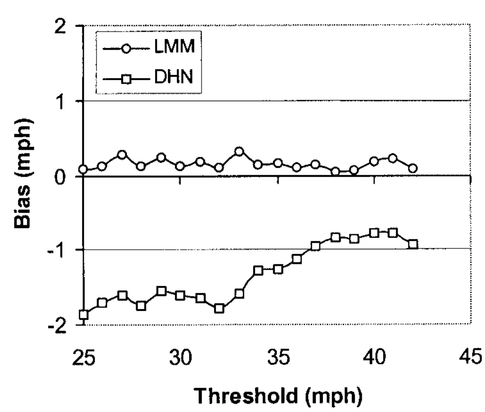

Figure 10. Bootstrap estimates of quantile bias (Boise, $R = 50$ year)

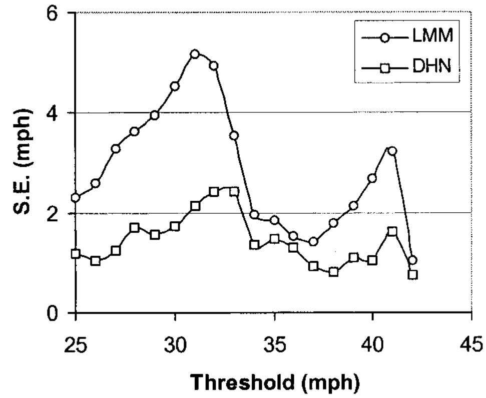

Figure 11. Bootstrap estimates of standard error (Boise, $R = 50$ year)

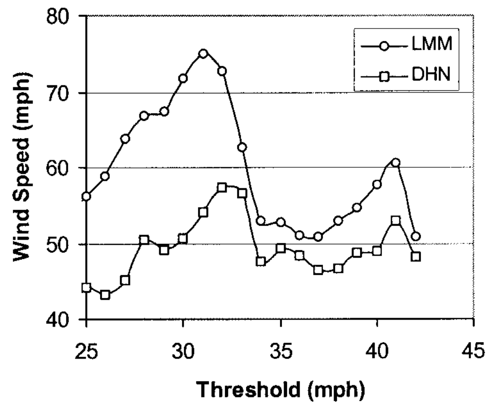

Figure 12. POT estimates of 1000-year wind speed (Boise, ID)

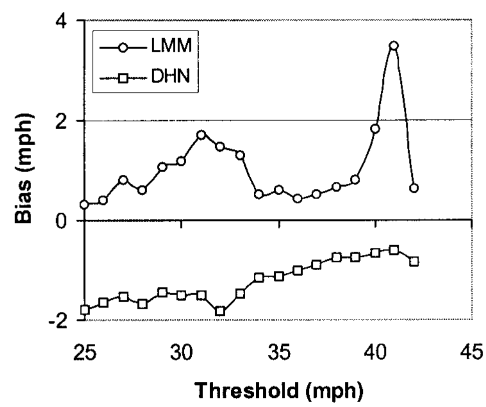

Figure 13. Bootstrap estimates of quantile bias (Boise, $R = 1000$ year)

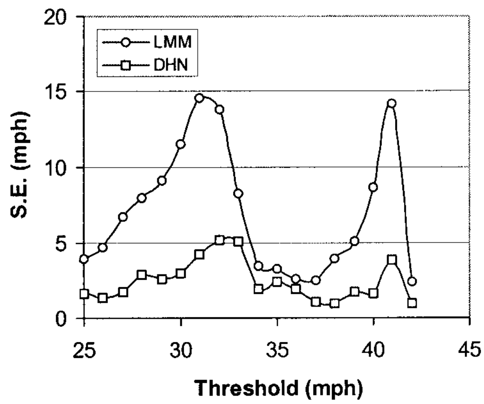

Figure 14. Bootstrap estimates of standard error (Boise, $R = 1000$ year)

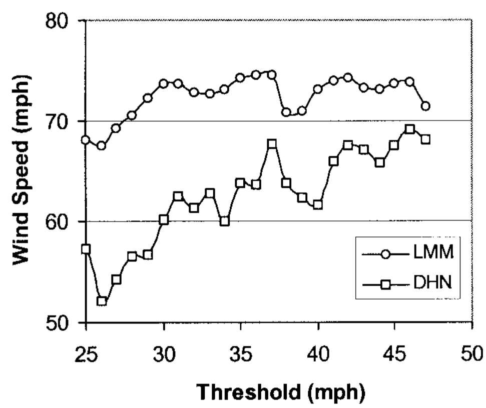

Figure 15. POT estimates of 50-year wind speed (Moline, IL)

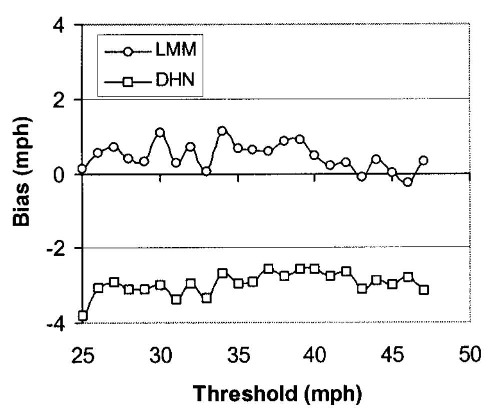

Figure 16. Bootstrap estimates of quantile biase (Moline, $R = 50$ year)

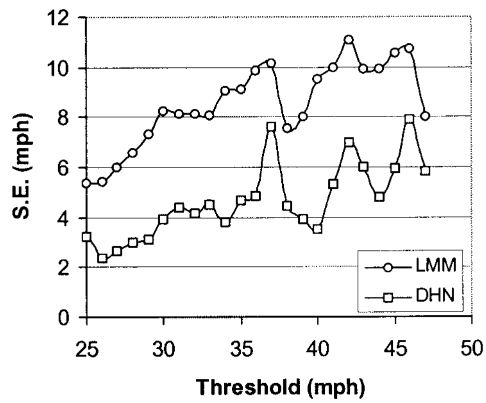

Figure 17. Bootstrap estimates of standard error (Moline, R = 50 year)

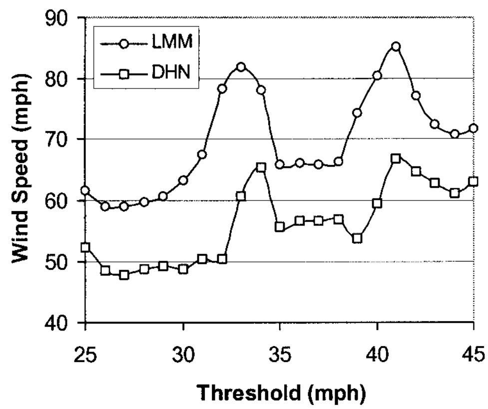

Figure 18. POT estimates of 1000-year wind speed (Toledo, OH)

suspicious points from the estimation. LMM and DHN estimates derived from the minimum SE criterion are 50 and 48 mph, respectively. Comparing Figures 9 and 12, a considerable amplification of quantile peaks is seen as the return period is increased from 50 to 1000 years.

Figure 15 shows that DHN estimates of 50-year wind speed, obtained from Moline data, are much more sensitive to threshold than LMM estimates. The bias of LMM estimates is smaller than that of DHN (Figure 16), whereas the reverse is true with reference to SE (Figure 17). Considering a reasonable threshold ($u > 30$ mph) and using the minimum SE criterion, LMM and DHN estimates are 70 ($u = 38$) and 60 mph ($u = 40$), respectively.

Estimates of 1000-year wind speed obtained using Toledo data are plotted in Figure 18. A stable variation of quantile plot is seen between $u = 35$ and 38 mph, leading to 66 mph (LMM) and 56 mph (DHN) estimates of wind speed. The bias and SE are also small in this range of thresholds for both methods (Figures 19 and 20). The quantile variation is also stable for $u < 30$ mph, which would, however, result in smaller estimates of wind speed.

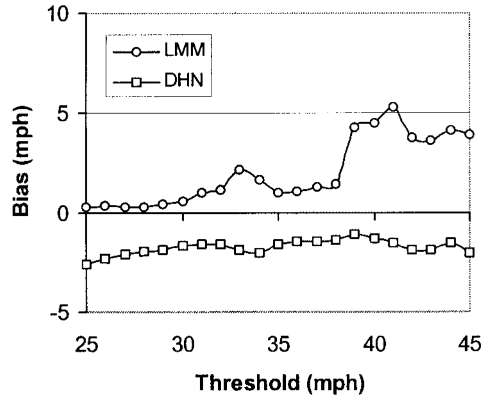

Figure 19. Bootstrap estimates of quantile bias (Toledo, $R = 1000$ year)

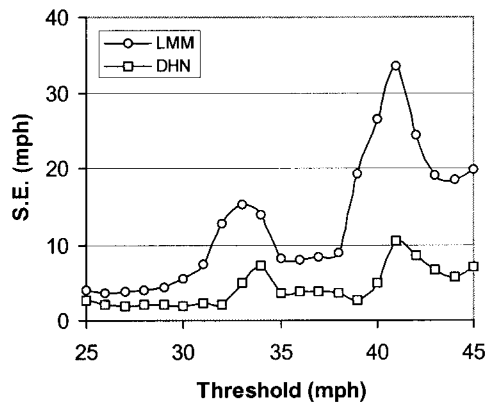

Figure 20. Bootstrap estimates of standard error (Toledo, $R = 1000$ year)

6. CONCLUSIONS

In practical applications of the POT method, the selection of a suitable threshold is critical to the estimation accuracy, since quantile estimates tend to exhibit wide variability with threshold. The article investigates the variation of quantile uncertainty (bias and variance) as a function of threshold with the purpose of identifying an optimal threshold corresponding to minimum statistical error, e.g. variance.

Since the original non-parametric version of the bootstrap method is not applicable to the extreme quantile estimation problem, the article examines the validity of a simple semi-parametric algorithm. Simulation examples considered in the article illustrate that the semi-parametric bootstrap algorithm is sufficiently accurate for comparing statistical performance of various estimation methods.

The article specifically compares the performance of the L-moment (LMM) and the de Haan (DHN) method for fitting the Pareto distribution to peak data. Simulation-based examples show that LMM can lead to unbiased estimates of 100-to 1000-year quantiles for certain threshold, whereas DHN consistently underestimates them. The LMM outperforms the DHN method in a sense that, for certain thresholds, it can lead to the least biased estimates with the same variance as that of DHN estimates.

Bootstrap uncertainty analysis of POT estimates of U.S. wind speed quantile confirms that LMM can lead to almost unbiased estimates for certain thresholds. A consistent underestimation of wind speed by DHN can be non-conservative with respect to structural safety. The standard error of DHN estimates is always less than that of LMM. However, for some thresholds the SEs of LMM and DHN are in close agreement. It is interesting to note that sharp peaks in a quantile vs. threshold plot are generally associated with large variance. This illustrates the usefulness of bootstrap method in identifying spurious quantile estimates. In general, a threshold corresponding to minimum standard error appears to provide reasonable estimates of wind speed extremes.

It is concluded that the quantification of uncertainty associated with a quantile estimate is necessary for selecting a suitable threshold and estimating a reasonable value of design wind speed. For this purpose, the semi-parametric bootstrap method has proved to be a simple, practical and effective tool.

ACKNOWLEDGEMENTS

The first author (MDP) gratefully acknowledges the financial support for this study that is provided by the Natural Science and Engineering Council of Canada.

REFERENCES

Caers J, Maes MA. 1998. Identifying tail bounds and end points of random variables. *J. Structural Safety* **20**: 1–23.

Caers J, Beirlant J, Vynckier P. 1998. Bootstrap confidence interval for tail indices. *Computational Statistics and Data Analysis* **26**: 259–277.

Davison AC, Smith RL. 1990. Models of exceedances over high thresholds. *J. Royal Statistical Society, Series B* **52**: 393–442.

de Haan L. 1994. Extreme value statistics. In *Extreme Value Theory and Applications*, Vol. 1, pp. 93–122, Galambos J, Lechner J, Simiu E (eds).

Draisma G, de Haan L, Peng L, Pereira TT. 1999. A bootstrap-based method to achieve optimality in estimating the extreme value index. *Extremes* **2**(4): 367–404.

Efron B, Tibshirani RJ. 1993. *An Introduction to the Bootstrap*. Chapman & Hall: New York, NY.

Greenwood JA, Landwehr JM, Matalas NC, Wallis JR. 1979. Probability weighted moments: definition and relation to parameters of several distributions expressable in inverse form. *Water Resources Research* **15**(5): 1049–1054.

Gross J, Heckert A, Lechner J, Simiu E. 1994. Novel extreme value estimation procedures: application to extreme wind data. *Extreme Value Theory and Applications*, Vol. 1, pp. 139–158, Galambos J, Lechner J, Simiu E (eds).

Hosking JRM. 1990. L-moments analysis and estimation of distributions using linear combination of order statistics. *J. Royal Statistical Society, Series B* **52**: 105–124.

Hosking JRM, Wallis JR. 1997. *Regional Frequency Analysis: An Approach based on L-Moments*. Cambridge University Press: Cambridge, UK.

Hutson AD. 2000. A composite quantile function estimator with applications in bootstrapping. *Journal of Applied Statistics* **27**(5): 567–577.

Kendall MG, Stuart A. 1977. *Advanced Theory of Statistics*. Griffin: London, UK.

Pandey MD, Van Gelder PHAJM, Vrijling JK. 2001a. The estimation of extreme quantiles of wind velocity using L-moments in the peaks-over-threshold approach. *J. Structural Safety* **23**: 179–192.

Pandey MD, Van Gelder PHAJM, Vrijling JK. 2001b. Assessment of an L-kurtosis-based criterion for quantile estimation. *J. Hydrologic Engineering, ASCE* **6**: 284–292.

Pickands J. 1975. Statistical inference using order statistics. *Annals of Statistics* **3**: 119–131.

Simiu E, Heckert NA. 1996. Extreme wind distribution tails: a peaks over threshold approach. *J. Structural Engineering, ASCE* **122**(5): 539–547.

Van Noortwijk JM, Van Gelder PHAJM. 1998. Bayesian estimation of quantiles for the purpose of flood prevention. *Proceedings of Coastal Engineering, ICCE* **26**: 3529–3541.

