2018 Scienceline Publication

Journal of Civil Engineering and Urbanism

Volume 8, Issue 2: 12-16; Mar 25, 2018

ISSN-2252-0430

Extreme Value Analysis of Wind Speed Data using Maximum Likelihood Method of Six Probability Distributions

Vivekanandan N

Central Water and Power Research Station, Pune, Maharashtra, India

Corresponding author's E-mail: anandaan@rediffmail.com

ORIGINAL ARTICLE
PII: S225204301800003-8
Received 20 Nov. 2017
Accepted 07 Feb. 2018

ABSTRACT: Assessment of wind speed at a region is a pre-requisite while designing tall structures viz. cooling towers, stacks, transmission line towers, etc. This can be achieved through Extreme Value Analysis (EVA) by fitting of probability distributions to the annual series of extreme wind speed (EWS) data that is derived from hourly maximum wind speed. This paper details the study on EVA of wind speed data recorded at India Meteorological Department Observatories of Delhi and Kanyakumari adopting six probability distributions such as Normal, Log Normal, Gamma, Pearson Type-3, Log Pearson Type-3 (LP3) and Extreme Value Type-1. Maximum likelihood method is applied for determination of parameters of the distributions. The adequacy of fitting of probability distributions to the series of recorded EWS data is evaluated by Goodness-of-Fit tests viz., Anderson-Darling and Kolmogorov-Smirnov and diagnostic test using D-index. Based on GoF and diagnostic tests results, the study suggests the LP3 distribution is better suited amongst six probability distributions adopted for EVA of wind speed data for Delhi ad Kanyakumari.

Keywords: Anderson-Darling test, D-index, Kolmogorov-Smirnov test, Log Pearson Type-3, Maximum likelihood method, Wind speed

INTRODUCTION

Technical and engineering appraisal of large infrastructure projects such as nuclear, hydro and thermal power plants, dams, bridges and flood control measures needs to be carried out during the planning and formulation stages of such projects. In a hydrological context, it is well recognized that whatsoever extreme the design-loading, more severe conditions are likely to be encountered in nature. Therefore, the accurate estimation of the occurrence of extreme wind speed (EWS) is an important factor in achieving the correct balance. Such estimates are commonly expressed in terms of the quantile value (x_T), i.e., the EWS which is exceeded, on average, once every T-year, the return period. For this situation, the annual series of EWS data derived from hourly maximum wind speed is generally fitted to a theoretical distribution in order to calculate the quantiles.

Probability distributions (PDs) such as Normal (NOR), 2-parameter Log Normal (LN2), Gamma (GAM), Pearson Type-3 (PR3), Log-Pearson Type-3 (LP3) and Extreme Value Type-1 (EV1) are commonly used for estimation of extreme events such as rainfall, stream flow and wind speed (Bivona et al., 2003; Della-Marta et al., 2009). Number of studies has been carried out by different researchers on adoption of different probability distributions for Extreme Value Analysis (EVA) of wind speed. Palutikof et al. (1999) expressed that the Generalized Extreme Value (GEV) distribution is better suited for EVA of wind speed for Sumburgh (Shetland). Pandey et al. (2001) applied GEV and GAM distributions for estimation of EWS for Helena, Boise and Duluth stations in United States of America. Topaloglu (2002) reported that the frequency analysis of the largest, or the smallest, of a sequence of hydrologic events has long been an essential part of the design of hydraulic structures. Guevara (2003) carried out hydrologic analysis using probabilistic approach to estimate the design parameters of storms in Venezuela.

Lee (2005) studied the rainfall distribution characteristics of Chia-Nan plain area using six PDs. Kunz et al. (2010) compared the GAM and Generalized Pareto (GP) distributions for estimation of EWS and concluded that the GP provides better estimates than

To cite this paper: Vivekanandan N. (2018). Extreme Value Analysis of Wind Speed Data using Maximum Likelihood Method of Six Probability Distributions. *J. Civil Eng. Urban.*, 8 (2): 12-16. www.ojceu.ir

GAM distribution. El-Shanshoury and Ramadan (2012) applied EV1 distribution to estimate the EWS for Dabaa area in the north-western coast of Egypt. Escalante-Sandoval (2013) applied five mixed extreme value distributions to estimate the EWS at 45 locations of the Netherlands. He also expressed that the mixed reverse Weibull and the mixture Gumbel-reverse Weibull distributions are better suited for estimation of EWS at 34 locations. Ahmed (2013) expressed that the rank regression method is the best suited amongst four methods studied for determination of parameters of Weibull distribution for estimation of EWS for Halabja region. Indhumathy et al. (2014) applied four parameter estimation methods of Weibull distribution and found that the energy pattern factor method is the best method to estimate the EWS for Kanyakumari region. Generally, when different distributional models are used for modelling EWS, a common problem that arises is how to determine which model fits best for a given set of data. This can be answered by formal statistical procedures involving Goodness-of-Fit (GoF) and diagnostic tests; and the results are quantifiable and reliable than those from the empirical procedures.

Qualitative assessment was made from the plot of the recorded and estimated EWS. For the quantitative assessment on EWS within in the recorded range, GoF tests viz., Anderson-Darling (AD) and Kolmogorov-Smirnov (KS) are applied. A diagnostic test of D-index is used for the selection of most suitable probability distribution for EVA of wind speed. In this paper, study on EVA of wind speed data adopting six PDs is presented. The applicability of GoF and diagnostic tests procedures in identifying which distribution is best for EVA of wind speed is also presented with illustrative example.

MATERIALS AND METHODS

The effort made in this study is to assess the applicability of PDs adopted in EVA of wind speed. For this, it is required to carry out various steps, which include: (i) Select six PDs such us NOR, LN2, GAM, PR3, LP3 and EV1 for EVA; (ii) select maximum likelihood method (MLM) for estimation of parameters of the distributions; (iii) select GoF and diagnostic tests and (iv) conduct EVA and analyse the results obtained thereof. Table 1 gives the quantile estimator (x_T) of six PDs that are used in EVA of wind speed.

Goodness-of-Fit tests

GoF tests viz., Anderson-Darling (AD) and Kolmogorov-Smirnov (KS) are applied for checking the adequacy of fitting of PDs to the recorded EWS data. The AD test statistic is defined by:

$$AD = (-N) - (1/N)\sum_{i=1}^{N}\left\{\begin{matrix}(2i-1)\ln(Z_i)+ \\ (2N+1-2i)\ln(1-Z_i)\end{matrix}\right\} \tag{1}$$

Here, $Z_i=F(x_i)$, for i=1,2,3,…,N with $x_1<x_2< \ldots x_N$, $F(x_i)$ is the Cumulative Distribution Function (CDF) of i^{th} sample (x_i) and N is the sample size (Zhang, 2002). The critical value (AD_C) of AD test statistic for different sample size (N) at 5% significance level is computed from:

$$AD_C = 0.757\,(1+(0.2/\sqrt{N}) \tag{2}$$

Similarly, the critical value (KS_C) of KS test statistic for different sample size (N) at 5% significance level is computed from:

$$KS_C = \underset{i=1}{\overset{N}{\mathrm{Max}}}\ (F_e(x_i) - F_D(x_i)) \tag{3}$$

Here, $F_e(x_i)$=(i-0.44)/(N+0.12) is the empirical CDF of x_i and $F_e(x_i)$ is the computed CDF of x_i.

Test criteria. If the computed value of GoF tests statistics given by the distribution is less than that of critical values at the desired significance level, then the distribution is considered to be acceptable for EVA of wind speed.

Table 1. Quantile estimator of six PDs

Distribution	PDF	X_T
NOR	$f(x;\alpha,m) = (1/\alpha\sqrt{2\pi})\, e^{-\frac{1}{2}\left(\frac{x-m}{\alpha}\right)^2}, x,\alpha>0$	$x_T = m+\alpha K_T$
LN2	$f(x;\alpha,m) = (1/\alpha x\sqrt{2\pi})\, e^{-\frac{1}{2}\left(\frac{\ln(x)-m}{\alpha}\right)^2}, x,\alpha>0$	$x_T = e^{m+\alpha K_T}$
GAM	$f(x;\alpha,\lambda) = \frac{\lvert\alpha\rvert e^{-\alpha x}(\alpha x)^{\lambda-1}}{\Gamma(\lambda)}, x,\lambda>0$	$X_T = \left(\frac{1}{\alpha}\right)(K_T\sqrt{\lambda}+\lambda)$
PR3	$f(x;\alpha,\lambda,m) = \frac{\lvert\alpha\rvert}{\Gamma(\lambda)} e^{-\alpha(x-m)}[\alpha(x-m)]^{\lambda-1}, x,\lambda>0$	$X_T = m+\left(\frac{\lambda+K_P\sqrt{\lambda}}{\alpha}\right)$
LP3	$f(x;\alpha,\lambda,m) = \frac{\lvert\alpha\rvert}{\Gamma(\lambda)}\left(\frac{e^{\alpha m}}{x^{1+\alpha}}\right)[\alpha(\ln x-m)]^{\lambda-1}, x,\lambda>0$	$X_T = e^{m+((\lambda+K_P\sqrt{\lambda})/\alpha)}$
EV1	$f(x:\alpha,m) = \frac{e^{-(x-m)/\alpha}e^{-e^{-(x-m)/\alpha}}}{\alpha}$, x, α>0	$x_T = m+\alpha Y_T$

In Table 1, α, λ and m are the scale, shape and location parameters respectively. For NOR, the values of m and α are computed from mean and standard deviation of the series of EWS. Similarly, for LN2, the values of m and α are computed from the mean and standard deviation of the log-transformed series of EWS. For EV1 distribution, the reduced variate (Y_T) corresponding to return period (T) is defined by Y_T=-ln(-ln(1-(1/T))). K_T is the frequency factor corresponding to return period and Coefficient of Skewness (CS) [CS= $2/\sqrt{\lambda}$ for GAM, CS=0.0 for NOR and LN2]. K_P is the frequency factor corresponding to CS of the original and log-transformed series of EWS for PR3 and LP3 distributions respectively (Rao and Hameed, 2000). The parameters of PDs are computed by MLM and used in estimation of wind speed. The theoretical descriptions of MLM of GAM, PR3, LP3 and EV1 are briefly described in the text book titled 'Flood Frequency Analysis' published by Rao and Hameed (2000).

Diagnostic test

The selection of most suitable probability distribution for EVA of wind speed is performed through D-index, which is defined by:

$$\mathrm{D-index} = \left(1/\bar{x}\right)\sum_{i=1}^{6}\left|x_i - x_i^*\right| \qquad (4)$$

Here, $\bar{x}$ is the mean value of the recorded EWS. Also, x_i is the i^{th} sample of the first six highest values in the series of recorded EWS and x_i^* is the corresponding estimated value by PDs. The distribution having the least D-index is considered as better suited distribution for EVA of wind speed (USWRC, 1981).

Application

In this paper, a study on EVA of wind speed adopting six probability distributions (using MLM) was carried out. HMWS data recorded at Delhi for the period 1969 to 2007 and Kanyakumari for the period 1970 to 2008 is used. The annual series of EWS is extracted from hourly wind speed data and further used for EVA. Table 2 gives the descriptive statistics of annual EWS for the regions under study.

Table 2. Descriptive statistics of annual EWS

Region	Statistical parameters			
	Mean (km/hr)	Standard deviation (km/hr)	Coefficient of Skewness	Coefficient of kurtosis
Delhi	66.1	261.1	0.047	-1.709
Kanyakumari	42.3	123.0	2.219	6.848

RESULTS AND DISCUSSIONS

By applying the procedures, as described above, computer program through R-package was developed and used for EVA of wind speed. The program computes the parameters of six PDs, GoF (AD and KS) tests statistic and D-index values for Delhi and Kanyakumari.

Estimation of EWS using six PDs

The parameters obtained from MLM were used for estimation of EWS for Delhi and Kanyakumari through quantile functions of the respective PDs and presented in Tables 3 and 4 respectively.

Table 3. Estimates of EWS given by six PDs for Delhi

Return period (year)	Estimated EWS (km/hr) using					
	NOR	LN2	GAM	PR3	LP3	EV1
2	66.1	64.2	64.7	61.2	65.0	63.3
5	79.6	79.0	79.7	80.0	80.2	78.9
10	86.6	88.1	88.4	93.1	87.6	89.2
20	92.4	96.4	96.0	105.7	93.4	99.1
50	98.9	106.7	105.0	121.8	99.6	111.9
100	103.2	114.1	111.4	133.8	103.5	121.5
200	107.2	121.4	117.4	145.6	106.9	131.1
500	112.0	130.8	124.9	161.1	110.8	143.7
1000	115.4	137.9	130.4	172.7	113.4	153.3

Table 4. Estimates of EWS given by six PDs for Kanyakumari

Return period (year)	Estimated EWS (km/hr) using					
	NOR	LN2	GAM	PR3	LP3	EV1
2	42.3	41.1	41.4	40.1	39.3	40.4
5	51.5	49.6	50.5	49.4	48.3	48.1
10	56.3	54.6	55.7	55.7	55.3	53.3
20	60.3	59.2	60.2	61.6	62.8	58.2
50	64.7	64.8	65.6	69.2	73.8	64.6
100	67.7	68.9	69.4	74.7	83.1	69.3
200	70.5	72.8	73.0	80.1	93.3	74.1
500	73.8	77.8	77.5	87.2	108.4	80.4
1000	76.1	81.5	80.7	92.4	121.3	85.1

From Table 3, it may be noted that the estimated EWS given by PR3 distribution are higher than the corresponding values of other five PDs for return period of 10-year and above for Delhi. Also, from Table 4, it may be noted that the LP3 distribution gave higher estimates for return period of 20-year and above when compared to the corresponding values of other five PDs for Kanyakumari. For qualitative assessment, the plots of recorded and estimated EWS were developed and presented in Figure 1.

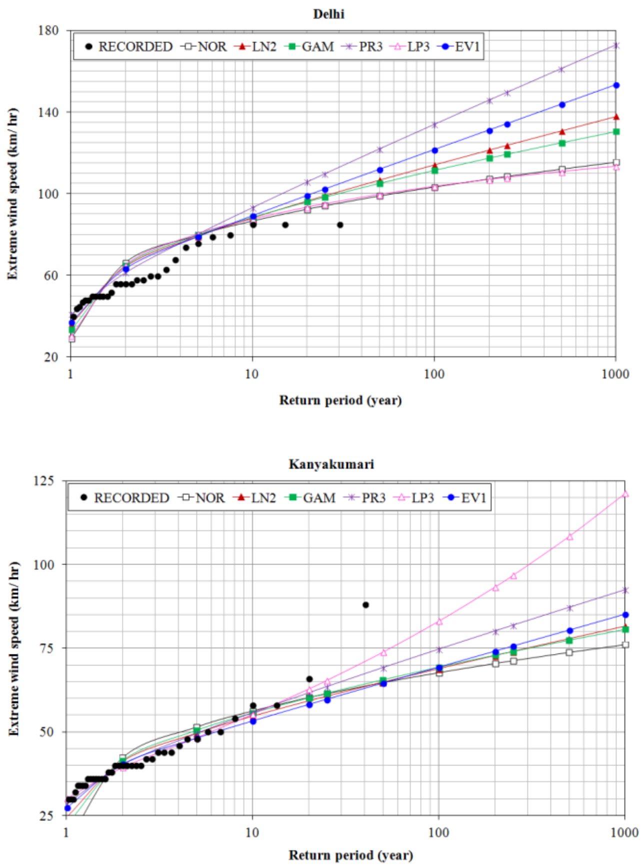

Figure 1. Plots of recorded and estimated EWS for different return periods by six PDs for Delhi and Kanyakumari

From Figure 1, it can be seen that there is no significant difference between the frequency curves of LN2 and GAM distributions for Kanyakumari. Similarly, for Delhi, it can be seen that the frequency curves of NOR and LP3 distributions are very close to each other.

Analysis based on GoF tests

By applying the procedures of GoF tests, quantitative assessment on fitting of PDs to the series of EWS was carried out; and the results are given in Table 5.

Table 5. Computed values of GoF tests statistics by six PDs

Probability distribution	Computed values of GoF tests statistics for			
	Delhi		Kanyakumari	
	AD	KS	AD	KS
NOR	2.541	0.215	1.992	0.197
LN2	2.181	0.205	0.946	0.166
GAM	2.078	0.201	1.279	0.173
PR3	1.666	0.179	0.523	0.120
LP3	2.192	0.203	0.540	0.108
EV1	2.027	0.200	0.542	0.136

From Table 5, it may be noted that the computed values of AD test statistic by six PDs are greater than the theoretical value of 0.781 at 5% significance level, and at this level, all six PDs are not acceptable for EVA of wind speed for Delhi. For Kanyakumari, it may be noted that the computed values of AD test statistic by PR3, LP3 and EV1 distributions are not greater than the theoretical value of 0.781 and therefore these three distributions are acceptable for EVA of wind speed. Also, from Table 5, it may be noted that the computed values of KS tests statistic by six PDs are not greater than the theoretical value of 0.218 at 5% significance level, and at this level, all six PDs are found to be acceptable for EVA of wind speed for Delhi and Kanyakumari.

Analysis based on diagnostic test

For the selection of most suitable PD for estimation of EWS, the D-index values of six PDs were computed and presented in Table 6.

Table 6. Indices of D-index for six PDs

Region	D-index					
	NOR	LN2	GAM	PR3	LP3	EV1
Delhi	0.428	0.646	0.622	1.203	0.471	0.805
Kanyakumari	0.857	0.918	0.826	0.707	0.600	1.014

From Table 6, it may be noted that the indices of D-index given by NOR and LP3 distributions are minimum when compared to the corresponding indices of other distributions for Delhi and Kanyakumari respectively. But, the AD test results showed that the NOR distribution is not acceptable for EVA of wind speed for Delhi. After eliminating the NOR distribution from the group of six PDs, it may be noted that the D-index value of LP3 is the second minimum next to NOR; and therefore LP3 is considered as most appropriate PD for estimation of wind speed for Delhi. On the basis of GoF and diagnostic test results, LP3 distribution is identified as better suited for estimation of EWS for Delhi and Kanyakumari.

CONCLUSIONS

The paper presented the study on EVA of wind speed adopting six PDs (using MLM). Based on the results of EVA of wind speed, GoF and diagnostic tests, the following conclusions were drawn from the study:

a) AD test results confirmed the applicability of PR3, LP3 and EV1 distributions for EVA of wind speed for Kanyakumari.

b) AD test results didn’t support the use of all six PDs for EVA of wind speed for Delhi.

c) KS test results supported the use of all six PDs for EVA of wind speed for Delhi and Kanyakumari.

d) D-index value of LP3 is found as minimum for Kanyakumari whereas the D-index value of LP3 is the second minimum for Delhi.

e) LP3 distribution is identified as better suited amongst six distributions adopted for estimation of extreme wind speed for Delhi and Kanyakumari.

The study suggested that the 1000-year return period EWS of 113.4 km/hr (for Delhi) and 121.3 km/hr (for Kanyakumari) adopting LP3 distribution could be used as the design parameters for planning and design of hydraulic structures in the regions.

DECLARATIONS

Acknowledgements

The author is grateful to Dr. (Mrs.) V.V. Bhosekar, Additional Director, Central Water and Power Research Station, Pune, for providing the research facilities to carry out the study. The author is thankful to M/s Nuclear Power Corporation of India Limited, Mumbai, for the supply of wind speed data.

Author’s contribution

Shri N. Vivekanandan, Scientist-B, Central Water and Power Research Station, Pune, carried out the data analysis and prepared the manuscript.

Competing interests

The author declares that he has no competing interests.

REFERENCES

Bivona S, Burlon R, Leone C. (2003). Hourly wind speed analysis in Sicily. Journal of Renewable Energy, 28(9): 1371-1385.

Della-Marta PM, Mathis H, Frei C, Liniger MA, Kleinn J, Appenzeller C. (2009). The return period of wind storms over Europe. Journal of Climatology, 29(3): 437-459.

Palutikof JP, Brabson BB, Lister DH, Adcock ST. (1999). A review of methods to calculate extreme wind speeds. Journal of Meteorological Applications, 6(2): 119-132.

Pandey MD, Van Gelder PHAJM, Vrijling JK. (2001). The estimation of extreme quantiles of wind velocity using L-moments in the peaks-over-threshold approach. Journal of Structural Safety, 23(2): 179-192.

Topaloglu F. (2002). Determining suitable probability distribution models for flow and precipitation series of the Seyhan River Basin. Turkish Journal of Agriculture and Forestry, 26(4): 187-194.

Guevara E. (2003). Engineering design parameters of storms in Venezuela. Hydrology Days, 80-91.

Lee C. (2005). Application of rainfall frequency analysis on studying rainfall distribution characteristics of Chia-Nan Plain Area in Southern Taiwan. Journal of Crop, Environment and Bioinformatics, 2(1): 31-38.

Kunz M, Mohr S, Rauthe M, Lux R, Kottmeier Ch. (2010). Assessment of extreme wind speeds from regional climate models – Part 1: Estimation of return values and their evaluation. Natural Hazards Earth System Sciences, 10(4): 907-922.

El-Shanshoury I, Ramadan AA. (2012). Estimation of Extreme Value Analysis of Wind Speed in the North-Western Coast of Egypt. Arab Journal of Nuclear Science and Applications, 45(4): 265-274.

Escalante-Sandoval CA. (2013). Estimation of extreme wind speeds by using mixed distributions. Journal of Engineering Investigation and Technology,14(2): 153-162.

Ahmed SA. (2013). Comparative study of four methods for estimating Weibull parameters for Halabja, Iraq. Journal of Physical Sciences. 8(5): 186-192.

Indhumathy D, Seshaiah CV, Sukkiramathi K. (2014). Estimation of Weibull parameters for wind speed calculation at Kanyakumari in India. Journal of Innovative Research in Science, Engineering and Technology, 3(1): 8340-8345.

Rao AR, Hameed KH. (2000). Flood frequency analysis. CRC Publications, Washington D.C., New York.

Zhang J. (2002). Powerful goodness-of-fit tests based on the likelihood ratio. Journal of Royal Statistical Society, Series B, 64(2): 281-294.

USWRC. (1981). Guidelines for Determining Flood Flow Frequency. United States Water Resources Council (USWRC) Bulletin No. 17B, 15-19.