

Comparison of Estimators of Extreme Value Distributions for Wind Data Analysis

N. Vivekanandan

Abstract--- Estimation of extreme wind speed potential at a region is of importance while designing tall structures such as cooling towers, stacks, transmission line towers, etc. Assessment of wind speed can expediently be carried out by probabilistic modelling of historic wind speed data using extreme value distribution, or by using standard procedures available under BIS code of practices for building and structures. This paper illustrates the use of extreme value distributions such as Gumbel, Frechet and Weibull for modelling wind speed data recorded at Kanyakumari. Method of Least Squares (MLS), maximum likelihood method and order statistics approach are used for determination of parameters of the distributions. Kolmogorov-Smirnov test is used for checking the adequacy of fitting the method/distribution to the recorded data. D-index is used for selection of suitable method/distribution for estimation of design wind speed. The study shows that the Gumbel distribution using MLS is better suited for estimation of design wind speed for Kanyakumari. A comparative study of wind speed estimates obtained using MLS of Gumbel and with BIS code of practices is carried out; and results presented.

Keywords--- D-index, Frechet, Gumbel, Kolmogorov-Smirnov, Least Squares, Order Statistics, Weibull

I. INTRODUCTION

WIND engineering analyses the effect of wind in the natural and built environment and the damage due to the wind. For a structural engineer, the interest is in strong winds which may cause discomfort. With the structures becoming tall and slender the effect of wind on these structures is becoming critical. The distribution of wind speed is also important in determining serviceability of buildings [1]. The basic wind speed is arrived by considering 3-seconds (sec) average wind speed at a height of 10m; and the estimated value of 1000-year (yr) return period is used to arrive at a design load, which the structure must withstand during its lifetime [2-3]. For arriving at these values, wind records over a period of time are observed and statistical analysis is carried out.

In statistical theory, generalized extreme value distribution is identified as a family of continuous probability distributions that include Gumbel, Frechet and Weibull. As defined by the extreme value theorem in statistics, the asymptotic distributions of the extremes tend to converge on certain

limiting forms for large sample; specifically to the double exponential form, or to two different single exponential forms. Since the extreme values of a random variable are invariably associated with the tails of its probability density function, the convergence of the distribution function of its extreme value to a particular limiting form will depend on the behaviour at tail end of the initial distribution in the direction of the extreme [4-6]. In the present study, 2-parameter Gumbel, Frechet and Weibull distributions are considered for estimation of design wind speed for Kanyakumari region.

Analytical procedures like Method of Least Squares (MLS), Maximum Likelihood Method (MLM) and Lieblein technique based on Order Statistics Approach (OSA) are applied for determination of parameters of the distributions. Qualitative assessment is made from the probability plot of the recorded and estimated extreme wind speed. Goodness-of-Fit (GoF) test involving Kolmogorov-Smirnov (KS) is applied for checking the adequacy of the method/distribution to the recorded wind speed data. In addition to GoF test, diagnostic test involving D-index is applied in upper tail region for selection of suitable probability distribution for estimation of design wind speed for Kanyakumari region. The 1000-yr return period value of estimated 3-sec average extreme wind speed given by Gumbel, Frechet and Weibull distributions are compared with the corresponding values of IS 875 procedure to arrive at a design parameter for Kanyakumari region. The methodology adopted in estimation of design wind speed using extreme value distributions and IS 875 procedures, GoF and diagnostic tests are briefly described in the ensuing sections.

II. METHODOLOGY

The Cumulative Distribution Function (CDF) of Gumbel and Frechet distributions are given by:

$$F(W) = e^{-e^{-\left(\frac{W_i - \alpha_G}{\beta_G}\right)}}, \quad \beta_G > 0 \text{ and } -\infty < W_i < \infty \text{ (for Gumbel)}$$

$$F(W) = e^{-\left(\frac{W_i}{\beta_F}\right)^{-\lambda_F}}, \quad \beta_F > 0 \text{ and } -\infty < W_i < \infty \text{ (for Frechet)}$$

where α_G and β_G are the location and scale parameters of Gumbel distribution. The estimates of extreme wind speed using Gumbel (W_G) for different return periods (T) are computed from $W_G = \alpha_G + Y_T \beta_G$ and $Y_T = -\ln(-\ln(1-(1/T)))$. Similarly, β_F and λ_F are the scale and shape parameters of Frechet distribution. Based on extreme value theory, Frechet distribution can be transformed to Gumbel distribution through logarithmic transformation. Under this transformation, the estimates of extreme wind speed using Frechet (W_F) for

N. Vivekanandan, Assistant Research Officer, Central Water and Power Research Station, Pune 411024. E-mail: anandaan@rediffmail.com

different return periods are computed from $W_F = \text{Exp}(W_G)$, $\beta_F = \text{Exp}(\alpha_G)$ and $\lambda_F = 1/\beta_G$.

The CDF of Weibull distribution is given by:

$$F(W) = e^{-\left(\frac{W_i}{\beta_w}\right)^{\lambda_w}}, \beta_w > 0 \text{ and } -\infty < W_i < \infty$$

where β_w and λ_w are the scale and shape parameters of Weibull distribution. The estimates of extreme wind speed (W_w) using Weibull distribution for different return periods are computed from $W_w = \beta_w [-\text{Log}(1 - (1/T))]^{1/\lambda_w}$.

A. Estimation of Distribution Parameters

OSA is applied for estimating the parameters of Gumbel and Frechet distributions of the series consisting of the largest values. In addition, MLS is applied for determining the parameters of Gumbel distribution. The procedure adopted in determining the parameters of the distributions using MLS and OSA is as follows:

Method of Least Squares: The recorded extreme wind speed data are arranged in an increasing order of magnitude by assigning rank of 1, 2, 3,, N. The non-exceedance probability for each data point W_i for rank 'i' is assigned by using the following relation:

$$P_i = \frac{i}{N+1}$$

Using MLS, the parameters of Gumbel distribution can be obtained from the following equations:

$$\beta = \frac{\left(\sum_{i=1}^N W_i\right)^2 - \left(N \sum_{i=1}^N W_i^2\right)}{\left(N \sum_{i=1}^N W_i (\ln(-\ln(P_i)))\right) - \left(\sum_{i=1}^N W_i\right) \left(\sum_{i=1}^N \ln(-\ln(P_i))\right)}$$

$$\alpha = \left(\frac{\sum_{i=1}^N W_i}{N}\right) + \frac{\left(\sum_{i=1}^N \ln(-\ln(P_i))\right) \beta}{N}$$

Here $\ln(-\ln(P_i))$ defines the cumulative probability of non-exceedance for each W_i . The Standard Error (SE) on the estimated extreme wind speed is computed by:

$$SE = \frac{\beta}{\sqrt{N}} \left(1.1589 + 0.1919 Y_T + 1.1 Y_T^2\right)^{0.5}$$

Order Statistics Approach: The approach is based on the assumption that the set of extreme values constitutes a statistically independent series of observations. The OSA parameters of Gumbel distribution are given by:

$$\alpha_G = r^* \alpha_M^* + r' \alpha_M'; \quad \beta_G = r^* \beta_M^* + r' \beta_M'$$

where r^* and r' are proportionality factors, which can be obtained from the selected values of k , n and n' using the relations $r^* = kn/N$ and $r' = n'/N$. Here, N is the sample size contains basic data that are divided into k sub groups of n elements each leaving n' remainders [7].

In OSA, α_M^* and β_M^* are the distribution parameters of the groups and α_M' and β_M' are the parameters of the remainders, if any. These can be computed from the following equations:

$$\alpha_M^* = (1/k) \sum_{i=1}^n \alpha_{ni} S_i \text{ and } \alpha_M' = \sum_{i=1}^{n'} \alpha_{ni} W_i$$

$$\beta_M^* = (1/k) \sum_{i=1}^n \beta_{ni} S_i \text{ and } \beta_M' = \sum_{i=1}^{n'} \beta_{ni} W_i$$

where $S_i = \sum_{j=1}^k W_{ij}$, $j=1,2,3,\dots,n$. The values of the weights α_{ni} and β_{ni} are given in Table 1. The standard error on the estimated wind speed is computed by:

$$SE = [\text{Var}(W_F)]^{1/2}, \quad \text{Var}(W_F) = r^* W_n + r' W_{n'}$$

where, $r^* = \frac{1}{k} \left(\frac{kn}{N}\right)^2$ and $r' = \left(\frac{n'}{N}\right)^2$. W_n and $W_{n'}$ are defined

by the general form as $W_n = (A_n Y_G^2 + B_n Y_G + C_n) \beta_G^2$. The values of A_n , B_n , and C_n are given in Table 2.

Table 1: Weights α_{ni} and β_{ni} for Computation of OSA Parameters

α_{ni} / β_{ni}	i					
	1	2	3	4	5	6
α_{2i}	0.916373	0.083627				
α_{3i}	0.656320	0.255714	0.087966			
α_{4i}	0.510998	0.263943	0.153680	0.071380		
α_{5i}	0.418934	0.246282	0.167609	0.108824	0.058350	
α_{6i}	0.355450	0.225488	0.165620	0.121054	0.083522	0.048867
β_{2i}	-0.721348	0.721348				
β_{3i}	-0.630541	0.255816	0.374725			
β_{4i}	-0.558619	0.085903	0.223919	0.248797		
β_{5i}	-0.503127	0.006534	0.130455	0.181656	0.184483	
β_{6i}	-0.459273	-0.035992	0.073199	0.126724	0.149534	0.145807

Table 2: Variance Determinators for W_n

n	A_n	B_n	C_n
2	0.71186	-0.12864	0.65955
3	0.34472	0.04954	0.40286
4	0.22528	0.06938	0.29346
5	0.16665	0.06798	0.23140
6	0.13196	0.06275	0.19117

The procedures adopted in determining the MLM parameters of Weibull distribution and standard error on the estimated wind speed for different return periods are detailed in the text book on 'Flood Frequency Analysis' by Rao and Hamed [8]. The lower and upper confidence limits (LCL and UCL) for the expected Extreme Wind Speed (EWS) estimates can be obtained from $LCL = EWS - 1.96SE$ and $UCL = EWS + 1.96SE$.

B. Goodness-of-Fit Test

The KS statistics is defined by:

$$KS = \text{Max}_{i=1}^N (F_c(W_i) - F_D(W_i)) \dots (1)$$

Here $F_c(W_i) = (i - 0.35)/N$ is the empirical CDF of W_i and $F_D(W_i)$ is the computed CDF of W_i [9]. If the computed values of KS statistics of the method/ distribution are less than that of critical value at the desired significance level then the selected method/ distribution is accepted to be adequate than any other method/ distribution for modelling wind speed data.

C. Diagnostic Test

The D-index is given by:

$$D\text{-index} = (1/\bar{W}) \sum_{i=1}^6 |W_i - W_i^*| \dots (2)$$

where W_i and W_i^* are the i^{th} highest recorded and estimated wind speed values by Gumbel (using MLS and OSA), Frechet (using OSA) and Weibull (using MLM) distributions, and \bar{W} is the average wind speed of the recorded data. The method/ distribution having the least value for D-index is considered as the best method/ distribution for estimation of design wind speed [10].

D. Concept of IS 875 Procedure

Following IS 875 procedure, the basic wind speed (W_b) for a region is obtained and subsequently modified to account for different effects and get design wind speed (W_z) at height z (in meter) for the chosen class of structure [11]. The relationship between W_b and W_z can be expressed by:

$$W_z = W_b k_1 k_2 k_3 \dots (3)$$

where k_1 is the probability-factor/ risk-coefficient, k_2 is the terrain and height factor and k_3 is the topography factor. Value of k_1 for different classes and mean probable design life of structures can be computed from Eq. (4).

$$k_1 = \frac{A - B \left[\ln \left\{ -\frac{1}{\text{NYR}} \ln(1 - P_{\text{NYR}}) \right\} \right]}{A + 4B} \dots (4)$$

Here NYR is the mean probable design life (yr) of the structure, P_{NYR} the risk level in NYR consecutive years, A and B appropriate coefficients for the basic wind speed zone.

III. RESULTS AND DISCUSSIONS

A. Estimation of Design Wind Speed using Probability Distributions

By applying the procedures, as described above, a computer program was developed and used to estimate the 3-sec average extreme wind speed for Kanyakumari region. Hourly wind speed data for the period 1970-2008 was used to estimate the 3-sec average extreme wind speeds for different return periods adopting Gumbel (using MLS and OSA), Frechet (using OSA) and Weibull (using MLM) distributions [12]. By using the monogram on normalized wind speed, the recorded 3-sec average extreme wind speed is obtained by multiplying the factor of 1.52 with hourly data. Table 3 gives the estimates of 3-sec average extreme wind speeds for different return periods given by Gumbel, Frechet and Weibull distributions for Kanyakumari region.

From Table 3, it may be noted that the 3-sec average extreme wind speed given by Gumbel distribution using MLS, is consistently higher for return periods above 2-yr when compared with the corresponding values given by Gumbel and Frechet distributions using OSA, and Weibull using MLM. Also, from Table 3, it may be noted that the standard error on the estimated extreme wind speed for different return periods by Weibull (using MLM) and Gumbel (using OSA) distributions are considerably minimum when compared with the corresponding values of Gumbel (using MLS) and Frechet (using OSA) for the region under study.

Table 3: 3-sec Average Extreme Wind Speed (EWS) Estimates with Standard Error (SE) for Different Return Periods using Gumbel, Frechet and Weibull Distributions

Return period (yr)	Estimated 3-sec average extreme wind speed (km/ hr) using							
	Gumbel		Frechet (OSA)		Weibull (MLM)			
	MLS	OSA	EWS	SE	EWS	SE	EWS	SE
2	61.6	2.9	62.2	1.8	60.8	1.7	62.8	1.8
5	79.0	4.9	73.0	2.9	71.4	3.2	74.8	2.9
10	90.6	6.5	80.1	3.8	79.5	4.7	82.1	4.0
20	101.7	8.4	87.1	4.7	88.2	6.5	88.6	4.9
50	116.0	10.6	95.9	5.9	100.8	9.4	96.4	6.1
100	126.8	12.5	102.6	7.0	111.4	12.2	101.8	6.8
200	137.4	14.1	109.3	7.9	123.1	15.4	106.9	7.8
500	151.5	16.6	118.1	9.1	140.3	20.7	112.9	8.8
1000	162.2	18.2	124.8	10.2	155.0	25.4	117.3	9.6

Figure 1 shows the plot of recorded and estimated 3-sec average extreme wind speeds for different return periods given by Gumbel, Frechet and Weibull distributions.

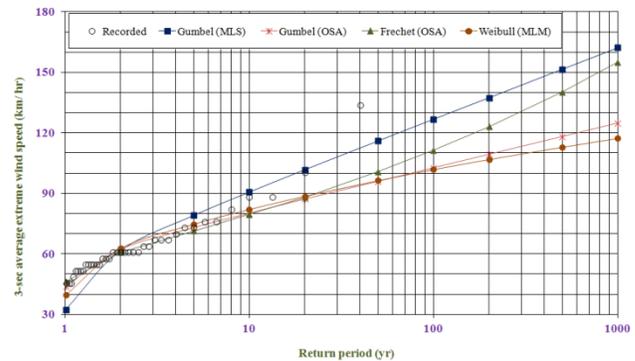


Figure 1: Plot of Recorded and Estimated 3-sec Average Extreme Wind Speeds for Different Return Periods using Gumbel, Frechet and Weibull Distributions for Kanyakumari

From Figure 1, it can be seen that the fitted curves using Gumbel and Weibull distributions show convergence while Frechet shows diverging trend on estimation of 3-sec average extreme wind speeds for the region under study.

B. Analysis Based on GoF Test

KS statistics for Gumbel (using MLS and OSA), Frechet (using OSA) and Weibull (using MLM) distributions using the series of 3-sec average extreme wind speed were computed from Eq. (1), and given in Table 4.

Table 4: Computed Values of KS Statistics Using Extreme Value Distributions

Extreme wind data series	Computed values of KS statistics using			
	Gumbel		Frechet (OSA)	Weibull (MLM)
	MLS	OSA		
3-sec average	0.176	0.150	0.109	0.179

From Table 4, it may be noted that the computed value of KS statistics given by Gumbel, Frechet and Weibull distributions are less than the critical value ($KS_{39,0.05}=0.218$) five percent level, and at this level, all three distributions are found to suitable for modelling wind speed data for Kanyakumari.

C. Analysis Based on Diagnostic Test

D-index values for Gumbel, Frechet and Weibull distributions using the series of 3-sec average extreme wind speed were computed from Eq. (2), and given in Table 5.

Table 5: Computed Values of D-index for Extreme Value Distributions

Extreme wind data series	Computed values of D-index for			
	Gumbel (MLS)	Gumbel (OSA)	Frechet (OSA)	Weibull (MLM)
3-sec average	0.697	1.098	1.063	1.005

From Table 5, it may be noted that the value of D-index given by Gumbel distribution using MLS is minimum, when compared with the corresponding value given by Gumbel and Frechet using OSA, and Weibull using MLM. Based on GoF and diagnostic tests, it is suggested that Gumbel distribution using MLS could be used for estimation of design wind speed for Kanyakumari region.

D. Estimation of Design Wind Speed using IS 875 Procedure

Based on wind speed map given in IS 875, the basic wind speed for Kanyakumari region is found to be 39 m/ s; and the coefficients of A and B corresponding to the basic wind speed is 84 and 14 respectively. Since Kanyakumari is considered to be a Terrain Category 1 i.e. exposed open terrain with few or no obstructions and in which the average height of any object surrounding the structure is less than 1.5m, the values of k_2 and k_3 are considered as 1.05 and 1.00 respectively. The values of k_1 for different return periods such as 50-yr, 100-yr, 200-yr, 500-yr and 1000-yr are computed from Eq. (4). By using the values of W_b , k_1 , k_2 and k_3 , the values of design wind speed for different return periods, at a standard height of 10m, are computed from Eq. (3). Table 6 gives a comparison of 3-sec average extreme wind speed estimates obtained from Gumbel distribution using MLS and IS 875 procedures.

Table 6: Comparison of Estimated 3-sec Average Extreme Wind Speed by Gumbel Distribution Using MLS and IS 875 Procedures

Return period (yr)	IS 875 (at standard height 10m)	Estimated 3-sec average extreme wind speed (m/ s) by Gumbel distribution using MLS	
		Wind speed	Standard Error
50	40.5	32.2	3.0
100	43.4	35.2	3.5
200	46.3	38.2	3.9
500	50.0	42.1	4.6
1000	52.9	45.1	5.0

From Table 6, it may be noted that the values of extreme wind speed for 50-yr, 100-yr, 200-yr, 500-yr and 1000-yr given by MLS of Gumbel distribution are comparatively less than the corresponding values obtained from IS 875 procedure. From the results of the data analysis, it may be noted that the percentages of variation on estimated 3-sec average extreme wind speed using MLS of Gumbel distribution, with reference to design wind speed obtained from IS 875 procedures corresponding to 50-yr, 100-yr, 200-yr, 500-yr and 1000-yr are about 21%, 19%, 18%, 16% and 15% respectively. The study recommended that the 3-sec average extreme wind speed with standard error of 50.1 m/ s (180.4 km/ hr) for 1000-yr return period may be considered for design purposes in Kanyakumari region. Figure 2 shows the plot of recorded and estimated 3-sec average extreme wind speed by Gumbel distribution using MLS together with confidence limits at 95% level for Kanyakumari region.

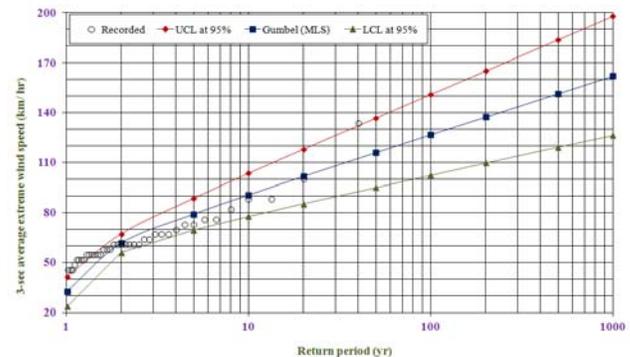


Figure 2: Plot of Recorded and Estimated 3-sec Average Extreme Wind Speed by Gumbel Using MLS with Confidence Limits at 95% Level for Kanyakumari

IV. CONCLUSIONS

The paper presented a computer aided procedure for estimation of design wind speed for Kanyakumari region adopting Gumbel (using MLS and OSA), Frechet (using OSA) and Weibull (using MLM) distributions. IS 875 procedure was also used to determine the design wind speed using basic wind speed for the region and the results were compared with the corresponding values obtained from extreme value distributions. GoF and diagnostic tests results showed that the Gumbel distribution using MLS is found to be appropriate for estimation of design wind speed for different return periods. The results also showed that the percentages of variation on

the estimated values of 3-sec average extreme wind speed with reference to design wind speed corresponding to 50-yr, 100-yr, 200-yr, 500-yr and 1000-yr are about 21%, 19%, 18%, 16% and 15% respectively. The study showed that the upper limit of 3-sec average extreme wind speed of 50.1 m/ s (180.4 km/ hr) for 1000-yr return period given by Gumbel using MLS may be adopted for design purposes for Kanyakumari region.

ACKNOWLEDGMENTS

The author is thankful to India Meteorological Department, Pune, for making available the wind speed data. The author is grateful to the Director, Central Water and Power Research Station, Pune, for providing the research facilities to carry out the study.

REFERENCES

- [1] International Atomic Energy Agency (IAEA, 2003), Meteorological events in site evaluation for Nuclear Power Plants – IAEA Safety Guide, No. NS-G-3.4, International Atomic Energy Agency, Vienna.
- [2] S. Bivona, R. Burlon and C. Leone, “Hourly wind speed analysis in Sicily, Renewable Energy”, Vol. 28, No. 9, Pp 1371-1385, 2003.
- [3] A.N. Celik, “On the distributional parameters used in assessment of the suitability of wind speed probability density functions, Energy Conversion and Management”, Vol.45, No. 11-12, Pp 1735-1747, 2004.
- [4] M. Deaves and I.G. Lines, “On the fitting of low mean wind speed data to the Weibull distribution”, Journal of Wind Engineering and Industrial Aerodynamics, Vol. 66, No. 3, Pp 169-178, 1997.
- [5] N.J. Cook, “Discussion on Modern estimation of the parameters of the Weibull wind speed distribution for wind energy analysis”, Journal of Wind Engineering and Industrial Aerodynamics, Vol. 89, No. 10, Pp 867-869, 2001.
- [6] J. Zhang and Y. Wu, “Likelihood ratio tests for normality”, Computational Statistics and Data Analysis, Vol. 49, No. 3, Pp 709-721, 2005.
- [7] Atomic Energy Regulatory Board (AERB, 2008), Extreme values of meteorological parameters, AERB Safety Guide AERB/ NF/ SG/ S-3.
- [8] R. Rao and H.K. Hamed, Flood Frequency Analysis, M/s CRC Publications, New York, London, 2000
- [9] R.B. D’Agostino and M.A. Stephens, Goodness of Fit Techniques, Marcel Dekker Inc, New York, 1986.
- [10] United States Water Resources Council (USWRC), Guidelines for Determining Flood Flow Frequency, Bulletin No. 17B, Pp 15-19, 1981.
- [11] IS: 875 (Part 3), Code of practice for design loads (other than earthquake) for building and structures, Bureau of Indian Standards, New Kanyakumari, 2003.
- [12] India Meteorological Department (IMD), Report on availability of meteorological data, 2010.



N. Vivekanandan (June 1969) post graduated in mathematics from Madurai Kamaraj University in 1991. He obtained post graduate degree in hydrology from University of Roorkee in 2000 and master of philosophy degree in mathematics from Bharathiar University in 2006. From May 1993 to March 2006, he worked as Research Assistant and analyzed studies. From April 2006 to till date, he is working Assistant Research Officer wherein carrying out hydrological studies using statistical and deterministic approaches, prediction of hydrometeorological variables using soft computing techniques and optimization of hydrometric network using spatial regression approach and information theory for various water resources projects.