Modelling of Annual Extreme Rainfall, Temperature and Wind Speed Using OSA of EV1 and EV2 Distributions

N. Vivekanandan

***Abstract*— Prediction of rainfall, temperature and wind speed is of utmost importance for planning, design and management of the civil structures at the project site. This can be carried out by fitting of probability distributions to the observed data. This paper illustrates the adoption of Gumbel (EV1) and Frechet (EV2) distributions for modelling of annual extreme rainfall, temperature and wind speed for Kanyakumari. Order Statistics Approach is used for determination of parameters of EV1 and EV2 distributions. Goodness-of-Fit (GoF) tests viz., Anderson-Darling and Kolmogorov-Smirnov are used for checking the adequacy of fitting of the distributions. Model Performance Indicators (MPIs) such as root mean square error and correlation coefficient are used for evaluating the performance of the probability distributions adopted in modelling of rainfall, temperature and wind speed. Based on GoF tests results and MPIs, the study recommends the EV1 distribution is better suited distribution for modelling of annual extreme rainfall and temperature whereas EV2 for annual extreme wind speed for Kanyakumari.**

***Key words*— Anderson-Darling, Correlation, Frechet, Gumbel, Kolmogorov-Smirnov, Mean square error**

I. INTRODUCTION

Technical and engineering appraisal of large infrastructure projects such as nuclear, hydro and thermal power plants, dams, bridges and airports needs to be carried out during the planning and formulation stages of such projects [1]. For the reason, modelling of rainfall, temperature, wind speed, etc., relating to the geographical region where the project is located as a basic requirement for assessing such phenomena, and arriving at structural and other related design parameters for the project [2]. This can be carried out by fitting of probability distributions to the observed data.

Out of number of probability distributions, Gumbel (EV1), Frechet (EV2) and Weibull are generally adopted for modelling of meteorological variables [3-4]. Esteves [5] applied EV1 distribution to estimate the extreme precipitation depths for different rain-gauge stations in southeast United Kingdom. Olumide et al [6] applied normal and EV1 distributions for prediction of rainfall and runoff at Tagwai dam site in Minna, Nigeria. They have also expressed that the normal distribution is better suited for rainfall prediction while Log-Gumbel for runoff. Lee et al. [7] applied EV1 and Weibull probability distributions for estimation of EWSs using the Korea wind map. They have observed that the EV1 distribution gives better results than Weibull. Daneshfaraz et al. [8] carried out the wind speed frequency analysis adopting 2-parameter log-normal, truncated extreme value, truncated logistic and Weibull probability distributions and found that the truncated extreme value is the most appropriate distribution for Urmia synoptic station in Iran.

Manuscript received 27 July 2015
N. Vivekanandan, Assistant Research Officer, Hydrometeorology Division, Central Water and Power Research Station, Pune 411024, Maharashtra; Phone: (020) 24103367; E-mail: anandaan@ rediffmail.com

Atomic Energy Regulatory Board (AERB) guidelines described that the Order Statistics Approach (OSA) is considered to be more appropriate for determination of parameters of EV1 and EV2 distributions [9]. AERB guidelines also described that the OSA estimates are popular owing to less bias and minimum variance though number of methods are available for parameter estimation. In this paper, Weibull distribution is not considered for modelling of meteorological variables due to non-existence of OSA for determination of parameters of Weibull distribution. Goodness-of-Fit (GoF) tests viz., Anderson-Darling (A^2) and Kolmogorov- Smirnov (KS) statistic is applied for checking the adequacy of fitting of the distributions to the observed data. Model Performance Indicators (MPIs) such as Root Mean Square Error (RMSE) and Correlation Coefficient (CC) are used for evaluating the performance of the EV1 and EV2 distributions used in modelling of meteorological variables. This paper details the procedures adopted in modelling of annual extreme rainfall, temperature and wind speed for Kanyakumari using OSA of EV1 and EV2 distributions through GoF tests and MPIs.

II. METHODOLOGY

A) Probability Distributions

The Cumulative Distribution Functions (CDFs) of EV1 and EV2 distributions for the series containing largest and smallest values are given in Table 1.

Table 1: CDFs of EV1 and EV2distributions

Distribution	CDFs for the series containing	
	Largest values	Smallest values
EV1	$F(X)=e^{-e^{-\left(\frac{X_G-\alpha_G}{\beta_G}\right)}}$	$F(X)=1-e^{-e^{-\left(\frac{X_G-\alpha_G}{\beta_G}\right)}}$
EV2	$F(X)=e^{-\left(\frac{X_F}{\beta_F}\right)^{(-\lambda_F)}}$	$F(X)=1-e^{-\left(\frac{X_F}{\beta_F}\right)^{(-\lambda_F)}}$

Here α_G and β_G are the location and scale parameters of EV1 distribution. The value of X_G for a given return period (T) adopting EV1 distribution is computed from:

$$X_G = \alpha_G + Y_T \beta_G \quad \ldots (1)$$

where $Y_T = -\ln(-\ln(1-(1/T)))$. Similarly, β_F and λ_F are the scale and shape parameters of EV2 distribution. Based on extreme value theory, EV2 distribution can be transformed to EV1 distribution through logarithmic transformation. Under this transformation, the value of X_F for a given return period (T) adopting EV2 distribution is computed from $X_F = \mathrm{Exp}(X_G)$, $\beta_F = \mathrm{Exp}(\alpha_G)$ and $\lambda_F = 1/\beta_G$ [10].

A1) Order Statistics Approach

The parameters of EV1 distribution using OSA are as follows:

$$\alpha_G = r^*\alpha^*_M + r'\alpha'_M \text{ and } \beta_G = r^*\beta^*_M + r'\beta'_M \quad \ldots (2)$$

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 and N can be written as N=kn+n′. 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 Eqs. (3) and (4), and are:

$$\alpha^*_M = (1/k)\sum_{i=1}^{n}\alpha_{ni}S_i \text{ and } \alpha'_M = \sum_{i=1}^{n'}\alpha_{n'i}X_i \quad \ldots (3)$$

$$\beta^*_M = (1/k)\sum_{i=1}^{n}\beta_{ni}S_i \text{ and } \beta'_M = \sum_{i=1}^{n'}\beta_{n'i}X_i \quad \ldots (4)$$

where $S_i = \sum_{i=1}^{k} X_{ij,}$ j=1,2,3,..,n. Here, X_i is the i^{th} observation in the remainder group having n′ elements, X_{ij} is the i^{th} observation in the j^{th} group having n elements. Table 2 gives the weights of α_{ni} and β_{ni} used in determination of parameters using OSA.

B) Goodness-of-Fit Tests

For the purpose of quantitative assessment, A^2 and KS tests are applied for checking the adequacy of fitting of EV1 and EV2 distributions to the observed data.

The A^2 statistic is defined by:

$$A^2 = (-N) - (1/N)\sum_{i=1}^{N}\left\{\begin{matrix}(2i-1)\,\mathrm{Ln}(Z_i) + \\ (2N+1-2i)\,\mathrm{Ln}(1-Z_i)\end{matrix}\right\} \quad \ldots (5)$$

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 CDF of i^{th} sample (X_i) and N is the sample size. The theoretical value (A^2_C) of A^2 statistic for different sample size (N) at 5% percent significance level is computed from $A^2_C = 0.757\,(1+(0.2/\sqrt{N}))$.

The KS statistic is defined by:

$$KS = \underset{i=1}{\overset{N}{\mathrm{Max}}}\ (F_e(X_i) - F_D(X_i)) \quad \ldots (6)$$

Here, $F_e(X_i)$ is the empirical CDF of X_i and $F_D(X_i)$ is the computed CDF of X_i by probability distributions. The theoretical value (KS_C) of KS statistic for different sample size (N) at 5% percent significance level is computed from $KS_C = 1.36/\sqrt{N}$ [11]. For empirical CDF, Weibull plotting position formula is generally used for presentation of results in the form of probability plots.

Test criteria: If the computed values of GoF tests statistic given by the distribution is less than that of theoretical values at the desired significance level, say 5% then the distribution is considered to be acceptable for modelling of meteorological variables at that level.

Table 2: Weights of α_{ni} and β_{ni} for determination of parameters using OSA

α_{ni} (or) β_{ni}	i					
	1	2	3	4	5	6
α_{2i}	0.9164	0.0836				
α_{3i}	0.6563	0.2557	0.0880			
α_{4i}	0.5110	0.2639	0.1537	0.0714		
α_{5i}	0.4189	0.2463	0.1676	0.1088	0.0584	
α_{6i}	0.3555	0.2255	0.1656	0.1211	0.0835	0.0489
β_{2i}	-0.7214	0.7214				
β_{3i}	-0.6305	0.2558	0.3747			
β_{4i}	-0.5586	0.0859	0.2239	0.2488		
β_{5i}	-0.5031	0.0065	0.1305	0.1817	0.1845	
β_{6i}	-0.4593	-0.0360	0.0732	0.1267	0.1495	0.1458

C) Model Performance Analysis

Theoretical descriptions of CC and RMSE are given by:

$$CC = \frac{\sum_{i=1}^{N}(X_i - \overline{X})(X^*_i - \overline{X^*})}{\sqrt{\left(\sum_{i=1}^{N}(X_i - \overline{X})^2\right)\left(\sum_{i=1}^{N}(X^*_i - \overline{X^*})^2\right)}} \quad \ldots (7)$$

$$RMSE = \left(\frac{1}{N}\sum_{i=1}^{N}(X_i - X^*_i)^2\right)^{0.5} \quad \ldots (8)$$

where X_i is the observed value of i^{th} observation, X^*_i is the predicted value of i^{th} observation by probability distribution, $\overline{X}$ is the average of X_i and $\overline{X^*}$ is the average of X^*_i. The distribution provides minimum RMSE and better CC is considered as the best suited distribution for modelling of meteorological variables [12].

III. APPLICATION

In this paper, the study on modelling of annual extreme rainfall, temperature and wind speed was carried out adopting EV1 and EV2 distributions (using OSA) for Kanyakumari. The series of Annual 1-day Maximum Rainfall (AMR) and Annual Maximum and Minimum Temperature (AT_{max} and AT_{min}) was extracted from the daily data. Similarly, the annual series of Annual Maximum wind speed (AMWS) was extracted from the hourly wind speed. The extracted series of AMR, AT_{max}, AT_{min} and AMWS was further used for modelling adopting EV1 and EV2 distributions. Daily rainfall, temperature and wind speed data observed at Kanyakumari for the period 1970 to 2008 was used. Table 3 gives the descriptive statistics of the observed meteorological data of Kanyakumari.

Table 3: Descriptive statistics of observed meteorological data of Kanyakumari

Descriptive statistics	AMR	AT_{max}	AT_{min}	AMWS
Average	92.2 mm	35.7 °C	20.7 °C	42.3 km/hr
SD	47.0 mm	0.9 °C	0.8 °C	11.1 km/hr
Skewness	1.317	1.578	-0.598	2.223
Kurtosis	1.389	6.066	0.593	6.891
SD: Standard Deviation				

IV. Results and Discussions

A) *Modelling of Extreme Values using EV1 and EV2*

Based on the parameter estimation procedures of OSA for EV1 and EV2 distributions, as given in the AERB guidelines, computer codes were developed in FORTRAN language and used for modelling of meteorological variables. These programs compute the distribution parameters, predicted values for different return periods, GoF tests statistic and MPIs. The plots of predicted values of AMR, AT_{max}, AT_{min} and AMWS using OSA of EV1 and EV2 distributions together with the plots of observed values are presented in Figures 1 to 4.

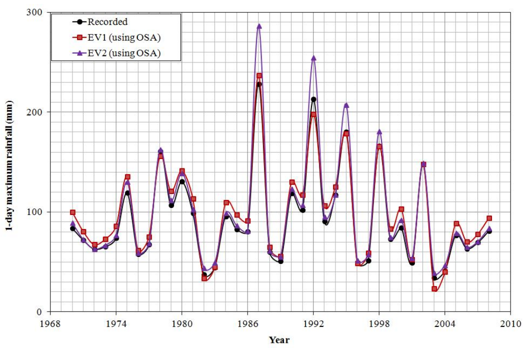

Figure 1: Plots of observed and predicted AMR

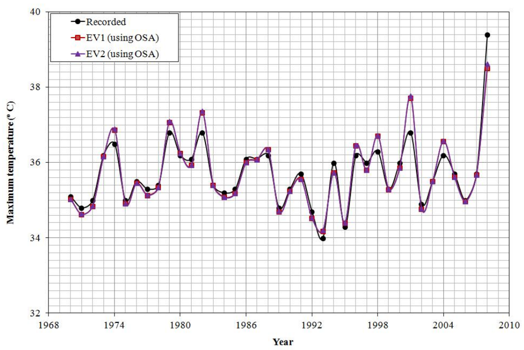

Figure 2: Plots of observed and predicted AT_{max}

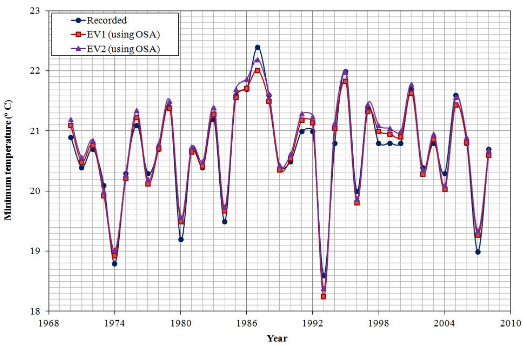

Figure 3: Plots of observed and predicted AT_{min}

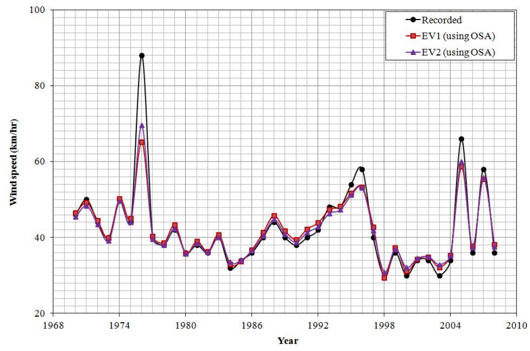

Figure 4: Plots of observed and predicted AMWS

For Figures 1, 3 and 4, it can be seen that the predicted values of AMR, AT_{min} and AMWS using EV2 distribution is consistently higher than the corresponding values of EV1. Also, from Figure 2, it can be seen that there is no appreciable between the predicted AT_{max} values using EV1 and EV2 distributions.

B) *Analysis based on GoF tests*

The GoF tests such as A^2 and KS were applied for checking the adequacy of fitting of EV1 and EV2 distributions (using OSA) to the data series of observed meteorological variables; and the results are presented in Table 4. After summarizing the A^2 and KS tests results, the following observations were made from the study:

i) A^2 test supported the use of EV1 and EV2 distributions for modelling the series of AMR and AT_{max}.
ii) KS test supported the use of EV1 and EV2 distributions for modelling the series of AMR, AT_{max}, AT_{min} and AMWS.
iii) A^2 test confirmed the EV1 and EV2 distributions were not found to be acceptable for modelling the series of AT_{min}.
iv) A^2 test didn't suggest the use of EV1 distribution for modelling the series of AMWS.

C) *Analysis Based on MPIs*

MPIs using CC and RMSE were used for the selection of a suitable probability distribution for modelling of AMR, AT_{max}, AT_{min} and AMWS. The CC and RMSE values computed for EV1 and EV2 distributions (using OSA) are presented in Table 5.

From Table 5, it may be noted that the RMSE values obtained from EV1 distribution are minimum when compared with the corresponding values of EV2 for the data series of AMR, AT_{max} and AT_{min}. For the data series of AMWS, the RMSE value obtained of EV2 distribution was found to be minimum when compared with the corresponding values of EV1. Also, from Table 5, it may be noted that there is a good correction between the observed and predicted values adopting EV1 and EV2 distributions for rainfall and temperature. For wind speed, it may be noted that the CC value of EV2 distribution is higher than EV1. Based on the findings obtained through GoF tests results and MPIs, the study suggested that the EV1 (using OSA) is better suited probability distribution for modelling the series of rainfall and temperature whereas EV2 distribution for modelling the series of wind speed for Kanyakumari.

Table 4: Computed and theoretical values of GoF tests statistic using EV1 and EV2 distributions

GoF tests	Computed values of A^2 and KS								Theoretical value at 5 % level
	EV1				EV2				
	AMR	AT_{max}	AT_{min}	AMWS	AMR	AT_{max}	AT_{min}	AMWS	
A^2	0.533	0.690	1.223	1.031	0.483	0.713	1.014	0.771	0.781
KS	0.104	0.100	0.164	0.150	0.067	0.102	0.175	0.109	0.218

Table 5: Computed values of MPIs using EV1 and EV2 distributions

MPIs	Computed values of MPIs							
	EV1				EV2			
	AMR	AT_{max}	AT_{min}	AMWS	AMR	AT_{max}	AT_{min}	AMWS
CC	0.986	0.955	0.981	0.960	0.953	0.957	0.981	0.985
RMSE	10.285	0.275	0.162	4.114	13.015	0.278	0.185	3.401

V. CONCLUSIONS

The paper describes briefly the study carried out for modelling of annual extreme rainfall, temperature and wind speed adopting EV1 and EV2 probability distributions (using OSA) for Kanyakumari. The following conclusions are drawn from the study:

i) For AMR, AT_{min} and AMWS, it was found that the predicted values by EV2 distribution are higher than the corresponding values of EV1 distribution.
ii) The A^2 test results showed that EV1 and EV2 distributions are acceptable for modelling the series of AMR and AT_{max}.
iii) The A^2 test results showed that the EV1 and EV2 distributions are not acceptable for modelling the series of $AT_{min.}$
iv) From KS test results, it was observed that the EV1 and EV2 distributions for modelling the series of AMR, AT_{max}, AT_{min} and AMWS.
v) For the series of AMR, AT_{max} and AT_{min}, it was found that the RMSE values obtained of EV1 distribution are minimum than EV2. For the series of AMWS, the RMSE value obtained of EV2 distribution is minimum than EV1.
vi) The CC values showed that there is a good correction between the observed and predicted values for the series of meteorological variables.
vii) The study showed that the EV1 (using OSA) is better suited probability distribution for modelling the series of AMR and temperature (AT_{max} and AT_{min}) whereas EV2 (using OSA) for modelling the series of AMWS for Kanyakumari.
viii) The results presented in the paper would be helpful to the decision makers for planning, design and management of civil structures near Kanyakumari.

ACKNOWLEDGMENTS

The author is grateful to the 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 and India Meteorological Department, Pune, for supply of meteorological data used in the study.

REFERNCES

[1] International Atomic Energy Agency (IAEA), Meteorological events in site evaluation for Nuclear Power Plants – IAEA Safety Guide, No. Ns-G-3.4, Vienna, 2003.
[2] A.H. Ang and W.H. Tang, Probability concepts in engineering planning and design, Vol. 2, John Wiley & Sons, 1984.
[3] M.C. Casas, R. Rodriguez, M. Prohom, A. Gazquez, and A. Redano, "Estimation of the probable maximum precipitation in Barcelona (Spain)", Journal of Climatology, 2011, Vol. 31, No. 9, pp. 1322–1327.
[4] H. Hasan, N. Salam and M.B. Adam, "Modelling extreme temperature in Malaysia using generalized extreme value distribution, International Journal of Mathematical, Computational, Natural and Physical Engineering, 2013, Vol. 7, No. 6, pp. 618-624 .
[5] L. S. Esteves, "Consequences to flood management of using different probability distributions to estimate extreme rainfall", Journal of Environmental Management, 2013, Vol. 115, No. 1, pp. 98-105.
[6] B. A. Olumide, M. Saidu and A. Oluwasesan, "Evaluation of best fit probability distribution models for the prediction of rainfall and runoff volume (Case Study Tagwai Dam, Minna-Nigeria)", Journal of Engineering and Technology, 2013, Vo1. 3, No. 2, pp. 94-98.
[7] B.H. Lee, D.J. Ahn, H.G. Kim and Y.C. Ha, "An estimation of the extreme wind speed using the Korea wind map", Renewable Energy, 2012, Vol. 42, No. 1, pp. 4–10.
[8] R. Daneshfaraz, S. Nemati, H. Asadi and M. Menazadeh, "Comparison of four distributions for frequency analysis of wind speed: A case study", Journal of Civil Engineering and Urbanism, 2013, Vol. 3, No. 1, pp. 6-11.
[9] Atomic Energy Regulatory Board (AERB), "Extreme values of meteorological parameters", AERB/ NF/ SG/ S-3, 2008.
[10] N. Cook, "Comments on plotting positions in extreme value analysis", Journal of Applied Meteorology and Climatology, 2011, Vol. 50, No.1, pp. 255-266.
[11] P.E. Charles Annis, "Goodness-of-Fit tests for statistical distributions", [http://www.statistical engineering.com/goodness.html], 2009.
[12] J. Chen and B. J. Adams, "Integration of artificial neural networks with conceptual models in rainfall-runoff modelling", Journal of Hydrology, 2006, Vol. 318, No, 1-4, pp. 232-249.