Pakistan Journal of Applied Sciences 2(7): 774-776, 2002

Rain Fall Intensity Estimates by Generalized Pareto Distribution

[1]M. Arshad, [2]M. T. Rasool and [3]M. I. Ahmad
[1]University College of Administrative Sciences, Kotli, Azad Kashmir
[2]Government College, Samundri, District Faisalabad, Pakistan
[3]Department of Mathematics and Statistics, Sultan Qaboos University, Musqat, Oman

Abstract: The Generalized Pareto Distribution is considered as a model for a maximum rainfall data in Pakistan. The rain fall data available for ten stations in Pakistan is utilized. The test statistics of Kolmogrov Smirnov, Cramer Von Mises, Anderson Darling and Modified Anderson Darling are obtained. The rainfall intensity estimates are derived with different return periods.

Key Words: Intensity Estimates, Pareto Distribution, statistics

Introduction

Classical analysis of statistical data in most fields including meteorology and hydrology has assumed that the data being analysed may be reasonably modeled by distribution which are some what "light tailed" where the tail of the density function approaches zero like some kind of exponential function. The frequency analysis of the largest, or the smallest, of a sequence of meteorological and hydraulic events has long been an essential part in the design of the hydraulic structures. For instance, sizing the capacity of a diversion dam is based on the peak discharge corresponding to a given return period or exceedance probability. Estimates of flood discharge with various risks of exceedance are needed for a wide range of engineering problems. Examples are culvert and bridge design and construction in major projects. Decision making values usually lie at the tail end of a distribution and a three parameter distribution may give a better fit in the tail region as compared to a two parameter distribution. So the Generalized Pareto (GP) distribution is considered as a model for maximum rainfall data in Pakistan. The application of GP distribution include use in the analysis of extreme events (Hosking and Wallis, 1987) e.g. for the analysis of the precipitation data, in the flood frequency analysis, in the analysis of data of greatest wave heights or sea levels, maximum winds loads on building, in the maximum rain fall analysis, in the analysis of greatest values of yearly floods, breaking strengths of materials air craft loads etc.

The main objective of maximum rainfall analysis is to estimate the future probabilities of occurrence of rain fall events based on past records. These estimates are made by using annual maximum data of 24 hours duration from 10 stations across the country.

Materials and Methods

The Generalized Pareto Distribution: As defined by Van Montfort and Witter (1986), a random variable X is said to be distributed as Generalized Pareto (GP) distribution:

$$X(F) = b + a/c\,[1 - (1-F)^c] \qquad c \neq 0$$
$$\quad\;\; = b - a \log(1-F) \qquad C = 0$$

where a = scale parameter, b = location parameter, c= shape paramter

Reduced Variate: Let the distribution function of the Generalized Pareto distribution is

$$F = 1 - e^{-y}$$

where

$$y = -c^{-1} \log\{1 - c/a\,(x-b)\} \qquad c \neq 0$$
$$\text{and} \qquad = (x-b)/a \qquad c = 0$$
$$\Rightarrow \qquad y = -\ln(1-F)$$

where y is called reduced variate

The quantile function can be written as a function of reduced variate:

$$x(F) = b + 1/c\,(1-e^{-cy})$$

Return period: The return period T is defined as the reciprocal of the probability of exceedance i.e.

$$T = (1 - F)^{-1}$$

The quantile function of the GP distribution can be expressed in terms of return period as

$$x(T) = b + 1/c\,(1-T^{-c})$$

This expression is used for the rainfall intensity estimates.

Probability Weighted Moments (PWM): Let X be a real valued random variable with distribution function F, Greenwood *et al.* (1979) defined the probability weighted moments of X to be the quantities.

$$M_{p,r,s} = E\,[X^p\{F(x)\}^r\,\{1-F(x)\}^s]$$
$$= \int X^p\{F(x)\}^r\,\{1-F(x)\}^s\,dF(x)$$
$$= \int_0^1 \{x(F)\}^p F^r\,(1-F)^s\,dF$$

where p, r and s are real numbers and X(F) is the quantile function of X.

Estimation of Rainfall Intensities: In this study, the data available for the 10 stations is utilized. These rainfall data were extracted from the records of autographic rain guages maintained by the Regional Meterological Officer, Lahore. A computer programme was developed in Minitab package for the analysis of annual maximal data which performed the following function:

- Fitting a GP distribution to the annual maximum series by using PWM estimation method.
- Performing the Empirical Distribution Function (EDF) goodness of fit test.
- Computing rainfall intensity estimates with 2, 5, 10, 25 and 50 years return periods by using the quantile function of the GP distribution in terms of return period T.

Results and Discussion

The Generalized Pareto distribution is selected for the study because it is flexible enough to represent adequately the data arising from extreme natural events, since it is three parameters distribution. As compared to other methods of estimation, the method of PWM is regarded as better method, because this method gives more accurate estimates of the parameters for small samples and are less subject to bias. This method has greater recognition in maximum rainfall analysis and in flood frequency analysis than any other method. So PWM is used to estimate the parameters of the GP distribution.

To find the rainfall intensity estimates by GP distribution, the 24 hour annual maximum data for Faisalabad, Mianwali, Khanpur, Khushab, Muree, Bahawalpur, Jhelum, Islamabad, Bahawalnagar and Sialkot is utilized. The basic statistics of these maximum rainfall series are given in Table 1. This Table shows that all the rainfall series are skewed. The skewness ranges from 0.62 to 2.66. The test statistics of Kolmogrov Simirnov (D), Crammer Von Misos (W), Anderson Darling (A^2_n) and Modified Anderson Darling (AU^2_n) are determined. The parameter estimates and goodness of fit statistics are given in Table 2. By testing the goodness of fit, it is found that in over 80% of these cases, the GP distribution have been found to fit at a 0.05 level of significance. Especially the Anderson Darling, Modified Anderson Darling and Cramer Von Mises tests are attractive and powerful means of assessing goodness of fit of GP distribution. The rainfall intensity estimates with 2, 5, 10, 25 and 50 years return period for each city considered are computed. These estimates are presented in Table 3. Such estimates are required for several engineering purposes, including agriculture, drainage, estimation of erosin and reservoir sedimentation and designing the dimensions of storm sewers, flood channels, dams, spillwags, bridges, culverts and flood control works. This satisfactory goodness of fit results obtained justify the use of GP distribution. Finally the given estimates are recommended for design purposes.

Table 1: Basic Statistics of the Maximum Rainfall Data for Different Stations in Pakistan

Station	Count	Mean	Median	SD	Skewness
Faisalabad	48	64.67	61.00	27.59	1.42
Mianwali	28	64.74	51.50	45.78	2.66
Khanpur	21	32.22	22.90	26.94	1.12
Khoushab	32	63.25	53.95	34.03	1.16
Muree	40	80.18	75.45	19.67	0.09
Bahawalpur	41	38.36	33.00	20.85	0.62
Jhelum	20	76.47	69.50	32.92	1.45
Islamabad	20	98.54	95.00	38.23	0.29
Bahawalnagar	12	31.08	30.50	13.61	0.34
Sialkot	21	90.10	71.10	56.40	1.22

Table 2: Parameter Estiamtes and EDF Statistics for the GP Distribution for Individual Stations

	Parameter estimates			EDF statistics			
Station	b^	a^	c^	D	W	A^2_n	AU^2_n
Faisalabad	39.54	36.30	0.209	0.066	0.035	0.292	0.162
Mianwali	18.30	56.38	0.214	0.157	0.115	0.694	0.524
Khanpur	3.77	28.01	-0.015	0.174	0.014	0.119	0.156
Khoushab	18.26	66.25	0.473	0.166	0.118	0.888	0.529
Muree	62.93	30.03	0.335	0.103	0.054	0.375	0.181
Bahawalpur	27.34	17.24	-0.083	0.112	0.059	0.360	0.201
Jhelum	35.42	54.39	0.325	0.113	0.042	0.382	0.239
Islamabad	49.23	79.37	0.500	0.114	0.033	0.220	0.105
Bahawalnagar	27.48	4.40	-0.521	0.092	0.079	0.489	0.267
Sialkot	41.03	42.93	-0.187	0.101	0.035	0.215	0.100

Table 3: Rainfall Intensity Estimates (In Millimeters) for the Individual Statios for Selected Return Periods

	Return periods (years)				
Station	2	5	10	25	50
Faisalabad	62.96	89.16	105.90	124.62	136.58
Mianwali	54.62	95.07	120.80	149.45	160.69
Khanpur	23.29	49.42	69.43	85.22	106.74
Khoushab	57.42	92.92	111.22	127.82	136.37
Muree	81.89	101.18	109.68	116.31	119.24
Bahawalpur	37.99	59.11	69.26	77.86	81.99
Jhelum	69.18	103.58	123.59	143.98	165.84
Islamabad	95.72	136.95	160.73	186.15	205.44
Bahawalnagar	72.81	121.67	164.62	210.66	258.70
Sialkot	35.16	38.57	47.07	64.22	83.86

Conclusion

The GP distribution is selected for the study because it has been quite popular not only for flood frequency analysis but for fitting the distribution of extreme natural events in general. The parameters of the GP distribution are estimated by PWM method because this method is regarded as the best method and has a greater recognition in maximum rainfall analysis and in flood frequency analysis than any other method.

The major use of this study was to find the maximum rainfall estimates at certain probability levels. To illustrate the use of GP distribution, annual maximum rainfall data for different cities in Pakistan is utilized. The parameters of GP distribution are estimated by PWM test statistics of EDF tests are obtained. By testing the goodness of fit, it is found that in over 80% of the cases, the GP distribution have been found to fit at 0.05 level of significance, implying a satisfactory fit. The GP distribution is used to derive the rainfall intensity estimates for different return periods, for each city considered. The satisfactory goodness of fit results obtained justify the use of GP distribution. Finally, the given estimates are recommended for design purposes. It is hoped that the results presented in this study will help to increase the understanding of the statistical structure of extreme rainfalls.

References

Ahmad, M.I., C.D. Sinclair and B.D. Spurr, 1988. Assessment of Flood Frequency Models Using Empirical Distribution Function Statistics. Water Resources Res. 24: 1323-1328.

Arshad, M., M.T.Rasool and M.I. Ahmed, 1996. Cramer Von Mises Test for Generalized Pareto Distribution. Pak. J. Sci. 48: 76-79.

Greenwood, J.A., J.M. Landwehr and W.C. Matalas, 1979. Probability Weighted Moments. Definition And Relation to Parameters of Distributions Expressible in Inverse Form. Water Resour. Res., 15: 1049-1054.

Hosking, J.R.M. and J.R. Willis, 1987. Parameters and Quantile Estimators for Generalized Pareto Distribution. Technometrics, 29: 339-349.

Moharram, S.H., A.K. Gosian and P.N. Kapoor, 1993. A Comparative Study for the Estimators of The Generalized Pareto Distribution. J. Hydrology, 150: 169-185.

Van Montfor, M.A.J. and J.V. Witter, 1986. The Generalized Pareto Distribution Applied to Rainfall Depths. Hydrol. Sci. J. 31: 151-162.