Nordic Hydrology, 1983, 127-138

A New Technique for the Analysis of Extreme Rainfall with Application to Lagos Metropolis, Nigeria

E. S. Oyegoke
Hydraulics Research Unit., Univ. of Lagos
J. O. Sonuga
Enplan Group Consulting Engrs. Lagos

This paper focuses on the use of the principle of maximum entropy as an alternative technique for the parameter estimation of the Extreme Value Type – 1 (EV1) distribution or Gumbel distribution often used for the analysis and forecast of extreme events. A case study is made of storm rainfall analysis for Lagos metropolis using the available rainfall data for Ikeja, Oshodi and Lagos Mainland as obtained from Akanbi (1982).

For comparison purposes, the parameters of the EV1 distribution is also obtained using the Maximum Likelihood Method. The later being one of the most reliable techniques and perhaps the most widely used for parameter estimation of the EV1 distribution. This exercise has made it possible to demonstrate in some ways the superiority of the maximum entropy method over existing methods used for statistical simulation of extreme events.

Introduction

The extreme Value Type – 1 (EV1) distribution derived by Gumbel (1941) has found ready application in the analysis and forecast of most of the extreme occurrences found in engineering and other sciences. Up to date, several techniques are in use for estimating the parameters of the (EV1) distribution. These techniques include the method of moments, the maximum likelihood technique – NERC (1975), and others due to Kimball (1946) and Gumbel (1958).

These different methods, unfortunately, produce different forecasts with different confidence intervals. As regards the region of study, the absence of sufficiently large amount of rainfall data poses another complex problem of how best to extract optimum information from a group of inadequate data. To accomplish this objective, the maximum likelihood method which has remained the most widely recommended method for use in this instance is employed in comparison with the emerging concept of maximum entropy. Both methods are relatively complicated, but are amenable to computer applications. The Apple II microcomputer has been used to develop and run series of programmes for the studies reported in this paper.

The Extreme Value Type 1 (EV1) Distribution

The EVI distribution has the probability distribution function given as

$$p(x) = \exp[-\exp(-\frac{x-u}{\alpha})] \tag{1}$$

where $p(x)$ is the probability of an event not exceeding x, and u and α are the parameters of the distribution. Expressed in terms of a reduced variate y such that

$$x = u + \alpha y \tag{2}$$

Eq. (1) becomes

$$p(x) = \exp[-(\exp-y)] \tag{3}$$

By further defining the return period T as the reciprocal of the exceedence probability, the reduced variate y becomes

$$y = -\log_e[\log_e(\frac{T}{T-1})] \tag{4}$$

The EV1 Distribution and the Principle of Maximum Entropy

The expression for entropy was defined by Shannon (1948) as

$$H = -K\sum p(x_i)\log_e p(x_i) \qquad K > 0 \tag{5}$$

Sonuga (1972) was the first to introduce the concept of maximum entropy to hydrologic frequency anlysis. With Eq. (5) as the base, he obtained the minimally biased probability distribution of x consistent with the information that the mean, $\bar{x}$ and standard deviation s, are known as follows

$$p\left(\frac{x_i}{\bar{x}, s}\right) = \exp(-a_0 - a_1 x_i \; a_2 x_i^2 \quad \ldots .) \tag{6}$$

where a_0, a_1 and a_2 are the Lagrangian multipliers associated with the normality and main constraints respectively.

Eq. (6) can simply be written as

$$p(x_i) = \exp[-a_0 - \sum_{i=1}^{m} a_i f_i(x)] \tag{7}$$

Building on the work of Sonuga (1972), Jowitt (1979) undertook rigorous analysis that deduced the EV1 distribution from the principle of maximum entropy, when the given information relating to a random variable of unrestricted sense consists solely of the first two moments.

$$E\left(\frac{x-u}{\alpha}\right) \quad \text{and} \quad E\left[\exp\left(-\frac{x-u}{\alpha}\right)\right]$$

such that the choice of u and α assume the properties of the EV1 distribution which are mainly

$$E\left(\frac{x-u}{\alpha}\right) = 0.5772 \tag{8}$$

and

$$E\left(\exp -\frac{x-u}{\alpha}\right) = 1 \tag{9}$$

where $E[.]$ is the expectation operator.

At this stage, the procedure for obtaining the parameters u and α of the EV1 distribution by the Maximum entropy method can be outlined.

Algorithm for the Determination of u and α by the Maximum Entropy Method

STEP 1:

Obtain initial estimates α_0 and u_0 using the simpler method of moments such that

$$\alpha_0 = \frac{\sqrt{6}}{\pi}\left[\frac{\sum(x_i - \bar{x})}{(N-1)}\right]^{\frac{1}{2}} = \frac{\sqrt{6}}{\pi} s \tag{10}$$

and

$$u_0 = \bar{x} - 0.5772\alpha_0 \tag{11}$$

where mean

$$\bar{x} = \frac{1}{N}\sum_{i}^{N} x_i \tag{12}$$

and standard deviation

$$s = \frac{\sqrt{\sum (x_i - \bar{x})^2}}{N-1} \tag{13}$$

STEP 2

Transform to a new variate z such that

$$Z_0 = \frac{x_i - u}{\alpha_0} \tag{14}$$

and obtain sample moments $\bar{z}$ and $\bar{\varepsilon}_z$ where

$$\bar{Z} = \frac{1}{N}\sum_{i=1}^{N} z_i \tag{15}$$

and

$$\bar{\varepsilon}_z = \frac{1}{N}\sum_{i=1}^{N} \exp(-z_i) \tag{16}$$

STEP 3

At this stage values of u and α are required satisfying two conditions viz

$$\bar{Z} = 0.5772 \tag{17}$$

and

$$\bar{\varepsilon}_z = 1 \tag{18}$$

This can be accomplished by introducing two new variables υ and β such that

$$\bar{Z} = 0.5772\beta_0 + \nu_0 \tag{19}$$

and

$$\bar{\varepsilon}_z = -\nu_0 + 0.4228(\beta_0 - 1) \tag{20}$$

The values obtained for β_0 and υ_0 is now used to obtain new estimates of α and u such as

$$\alpha_1 = \alpha_0 \beta_0 \tag{21}$$

and

$$u_1 = u_0 + \alpha_0 \nu_0 \tag{22}$$

hence a new value of Z is obtained such that

$$Z_1 = \frac{x_i - x_1}{\alpha_1} \tag{23}$$

The process from Eqs. (14) to (22) are repeated until the values of υ and β are sufficiently close to zero and unity, respectively.

Algorithm for the Determination of u and α by the Maximum Likelihood Method

This is readily available in literature as for example Clarke (1973) and Baghirathan et al. (1978). The procedure begins also with using the simpler method of moments to provide the initial estimates of α and u as given by Eqs. (10) and (11) after which new estimates of α and u are obtained using Eqs. (24) and (25) stated as follows

$$\alpha_k = \bar{x} - \left[\frac{\sum_{i=1}^{N} x_i \exp(-x_i/\alpha_{k-1})}{\sum_{i=1}^{N} \exp-(x_i/\alpha_{k-1})} \right] \tag{24}$$

and

$$u_k = -\alpha_{k-1} \log_e \left[\sum_{i=1}^{N} \exp(-(x_i/\alpha_{k-1}))/N \right] \tag{25}$$

Updated values of u and α are obtained through an iterative procedure until the values obtained by successive iterations are practically the same.

Analysis of Individual Stations

The data available for each of the rainfall stations are suitably arranged to form a series consisting of maximum annual values. The maximum annual series for each station is then subjected to the analysis outlined above.

Tabel 1 – A comparison of the number of iterations necessary to yield accurate results by the Maximum Entropy Method (M.E.M) and the Maximum Likelihood Method (M.L.M)

Station	Duration in hrs	Sample size N	Maximum Entropy Method: Iterations I	Value of α	Value of u	Maximum Likelihood Method: Iterations I	Value of α	Value of u
Ikeja	0,40	18	5	6.6065	37.1969	11	6.9426	37.3268
	1.00	18	5	12.1275	55.8253	6	12.4706	55.9883
	3.00	25	4	18.9132	67.8115	7	18.8073	67.7668
	12.00	25	3	24.8157	79.1518	6	24.7735	79.1334
	24.00	25	4	25.9948	87.7900	7	26.7372	88.1307
Lagos Island	0.40	16	5	9.0915	36.0402	11	9.2936	36.1423
	1.00	16	2	12.1954	57.0981	11	12.6436	57.3086
	3.00	19	4	20.7219	79.5402	8	20.2109	79.2355
	12.00	19	4	35.4504	100.7437	9	33.6634	100.1019
	24.00	48	4	31.1142	111.5348	7	30.3417	110.9430
Oshodi	0.40	11	8	10.5300	33.1125	12	11.3726	33.5713
	1.00	11	4	19.4559	50.4064	6	19.3004	50.3423
	3.00	11	5	19.8630	66.4551	9	20.4369	66.7166
	12.00	11	6	24.7068	72.7752	11	25.9492	73.3912
	24.00	11	5	30.0204	77.2726	10	30.9759	77.7181

Tabel 2 – A comparison of measured and computed rainfall for Oshodi

RANK	0.4 HR. DURATION				1 HR. DURATION			
RETURN PERIOD T (YRS)	C.L.	OBS. RAIN-FALL	COMPUTED RAINFALL		C.L.	OBS. RAIN-FALL	COMPUTED RAINFALL	
			M.L.M	M.E.M			M.L.M	M.E.M
12.00	16.8	53.3	61.3	58.8	28.6	103.9	97.5	97.9
6.00	13.3	50.8	52.9	51.0	22.6	101.6	83.2	83.5
4.00	11.3	47.0	47.7	46.2	19.2	77.0	74.4	74.6
3.00	9.9	44.5	43.8	42.6	16.8	70.6	67.8	68.0
2.40	8.9	44.2	40.6	39.6	15.0	63.7	62.3	62.4
2.00	8.1	42.2	37.7	36.7	13.7	56.6	57.4	57.5
1.71	7.5	42.2	35.0	34.5	12.7	55.9	52.8	52.9
1.50	7.1	33.0	32.5	34.1	12.0	45.2	48.5	48.6
1.33	6.9	30.7	29.8	29.6	11.7	39.4	43.9	44.0
1.20	6.9	26.4	26.9	26.9	11.7	35.1	39.1	39.1
1.09	7.3	16.8	23.2	23.5	12.5	29.0	32.7	32.7

C.L = Confidence Limit = 2 × standard error thus 95 % C Limit = rainfall estimate ± C.L

M.L.M = Maximum Likelihood Method.

M.E.M = Maximum Entropy Method.

OBS. = Observed.

Regional Analysis

In view of the fact that the available station records are rather limited, a regional analysis of the data was also undertaken.

Regional analysis has the advantage of yielding floods with high return period and with a good level of confidence. This latter exercise was achieved by compounding the available data from all three stations into one series of data for each rainfall duration. The longer series of data are then analysed as earlier outlined to yield the regional forecasts.

Conclusion

The results obtained through the investigations outlined above are shown tabulated in Tables 1 to 6 and are also displayed graphically in Figs. 1 and 2. The estimates of rainfall values for various return periods provided by the two methods used indicate that both the maximum entropy approach and the maximum likeli-

Table 3 – A comparison of measured and computed rainfall for Ikeja

RANK	0.4 HR. DURATION				1 HR. DURATION			
RETURN PERIOD T (YRS)	C.L	OBS RAIN-FALL	COMPUTED RAINFALL M.L.M	COMPUTED RAINFALL M.E.M	C.L	OBS RAIN-FALL	COMPUTED RAINFALL M.L.M	COMPUTED RAINFALL M.E.M
21.00	8.9	56.9	58.3	57.2	16.0	91.7	93.7	92.5
10.50	7.3	54.1	53.3	52.4	13.1	84.7	84.7	83.7
7.00	6.4	50.8	50.3	49.6	11.5	83.3	79.3	78.5
5.25	5.7	46.2	48.1	47.4	10.3	79.0	75.3	74.6
4.20	5.2	45.7	46.4	45.8	9.4	77.5	72.2	71.6
3.50	4.8	45.2	44.9	44.4	8.6	66.5	69.6	69.0
3.00	4.5	44.5	43.6	43.1	8.0	66.0	67.2	66.7
2.63	4.2	41.9	42.5	42.1	7.5	64.4	65.2	64.8
2.33	4.0	41.7	40.8	40.5	7.1	63.8	63.2	62.9
2.10	3.8	41.1	40.4	40.1	6.7	61.5	61.5	61.2
1.91	3.6	40.6	39.4	39.2	6.4	61.0	59.7	59.5
1.75	3.4	40.1	38.5	38.3	6.2	58.9	58.1	57.9
1.62	3.3	38.1	37.6	37.5	5.9	57.7	56.5	56.3
1.50	3.2	37.9	36.7	36.6	5.8	57.4	54.9	54.7
1.40	3.1	36.8	35.7	35.7	5.6	53.3	53.1	53.0
1.31	3.1	34.6	34.8	34.8	5.6	52.6	51.4	51.3
1.24	3.1	34.3	33.9	33.9	5.6	52.1	49.8	49.8
1.17	3.2	33.8	32.7	32.8	5.7	45.5	47.8	47.8
1.11	3.3	29.2	31.5	31.7	5.9	40.6	45.5	45.6
1.05	3.5	26.7	29.6	29.9	6.3	38.9	42.2	42.4

C.L = Confidence Limit = 2 × standard error thus 95 % C Limits = rainfall estimate ± C.L
M.L.M = Maximum Likelihood Method
M.E.M = Maximum Entropy Method
OBS = Observed

hood method provide satisfactory forecasts of extreme rainfall for the region under study. Table 1 demonstrates vividly the superiority of the new maximum entropy concept when used in the simulation of extreme event. The estimates of the parameters of the EV1 produced through the principle of maximum entropy are unique and the iterative procedure for obtaining these values converges much faster than that for the commonly used technique of maximum likelihood.

It is therefore possible using the concept of maximum entropy to make considerable savings in the computation time.

Finally, the results of the local and regional rainfall depth duration frequency relationships obtained through this study should satisfy a long-felt need for a systematic forecast of extreme rainfall for a metropolis that suffers from perenial

Table 4 – A comparison of measured and computed rainfall for Lagos Island

RANK	0.4 HR. DURATION				1 HR. DURATION			
RETURN PERIOD T (YRS)	C.L	OBS RAIN-FALL	COMPUTED RAINFALL		C.L.	OBS RAIN-FALL	COMPUTED RAINFALL	
			M.L.M	M.E.M			M.L.M	M.E.M
17.00	12.6	59.4	62.2	61.5	17.2	110.0	92.8	91.3
8.50	10.2	55.6	55.5	54.9	13.9	72.4	83.6	82.4
5.67	8.8	52.3	51.4	51.0	12.0	69.3	78.0	77.1
4.25	7.9	47.8	48.4	48.0	10.7	68.6	74.0	73.2
3.40	7.1	47.0	45.9	45.6	9.7	67.6	70.6	70.0
2.83	6.5	46.5	43.9	43.6	8.9	67.3	67.8	67.2
2.43	6.0	45.0	42.0	41.8	8.2	67.3	65.3	64.8
2.13	5.7	44.5	40.4	40.2	7.7	67.7	63.1	62.7
1.89	5.3	43.4	38.8	38.6	7.2	66.5	60.9	60.6
1.70	5.1	41.7	37.3	37.1	6.9	64.8	58.8	58.6
1.55	4.9	34.8	35.8	35.7	6.6	61.0	56.9	56.7
1.42	4.7	32.5	35.0	34.3	6.4	56.5	54.8	54.7
1.31	4.6	30.5	32.8	32.7	6.3	53.3	52.6	52.6
1.21	4.7	26.9	30.9	31.0	6.4	51.6	50.2	50.3
1.13	4.8	26.5	29.0	29.0	6.5	44.7	47.6	47.7
1.06	5.2	26.2	26.3	26.5	7.0	38.1	44.0	44.2

C.L = Confidence Limit = 2 × standard error thus 95 % C Limit = rainfall estimate ± C.L.
M.L.M = Maximum Likelihood Method
M.E.M = Maximum Entropy Method
OBS = Observed

Table 5 – Regional estimates of parameters u and α of EVI distribution.

RAINFALL DURATION (HRS)	SAMPLE SIZE N	MAX. ENTROPY METHOD (M.E.M)		MAX. LIKELIHOOD METHOD (M.L.M)	
		u	α	u	α
0.40	47	35.56	8.83	35.92	9.49
1.00	47	54.48	14.57	54.72	15.06
3.00	54	71.15	20.65	71.15	20.67
12.00	54	84.44	30.13	84.26	29.69

flooding resulting from high tropical rainfall and inadequate record of rainfall data.

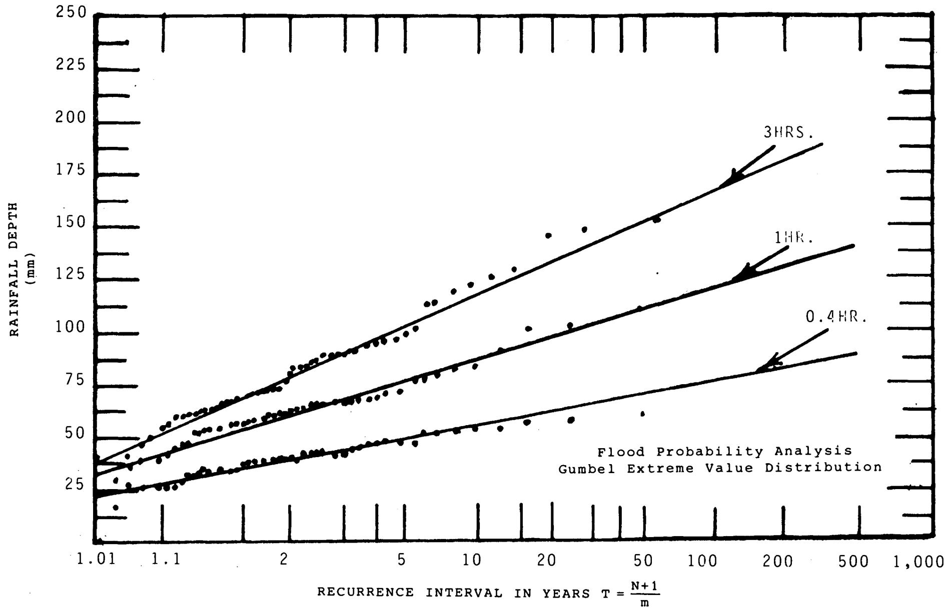

Fig. 1. Regional depth-duration-frequency relationship for Lagos Metropolis by the Maximum Entropy Method (M.E.M.).

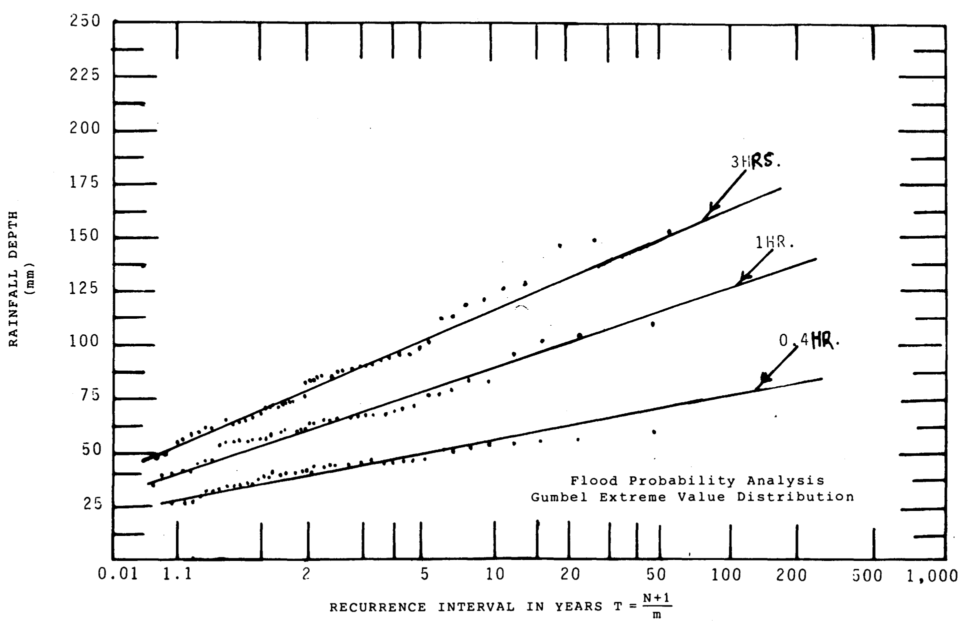

Fig. 2. Regional depth-duration-frequency relationship for Lagos Metropolis by the Maximum Likelihood Method (M.L.M.).

Table 6 – Chi-Square (X^2) test of regional estimates

RAINFALL DURATION (HRS)	SAMPLE SIZE N	X^2-VALUES BY M.E.M	X^2-VALUES BY M.L.M	X^2 0.01	X^2 0.05
0.40	47	9.3045	10.4712	26.7	31.48
1.00	47	6.3594	6.0725	26.7	31.48
3.00	54	7.2508	7.2098	32.04	37.32
12.00	54	12.6231	14.3393	32.04	37.32

Acknowledgement

The authors acknowledge the noble assistance of Messrs. C. B. Nkemdrim and Tunde Olasubulumi for the successful completion of this project. The assistance of the later in running the computer programmes developed for this project is well appreciated.

The painstaking efforts of Mrs. Toyin Oyegoke in typing this paper is also appreciated.

References

Akanbi, A. A. (1982) The Empirical, Probability and Regional Analysis of some Water Resources Parameters for Nigeria. A University of Lagos M. Sc. (Engineering) Dissertation.

Baghirathan, V. R., and Shaw, E. M. (1978) Rainfall depth-duration-frequency studies for Sri Lanka, *J. Hydrol., Vol. 37,* 223-239.

Bell, F. C. (1969) Generalized rainfall-duration-frequency relationship, *J. Hydraulic Div. ASCE 95 (HY1)*, 311-327.

Clarke, R. T. (1973) Mathematical models in hydrology. F.A.O., Irrig. Drain. Pap., 19.

Gumbel, E. J. (1941) The return period of flood flows. *Ann. Math. Statistics, Vol. 12 (2),* 163-190.

Gumbel, E. J. (1958) *Statistics of Extremes.* Columbia University Press, New York, N.Y., 375 pp.

Jowitt, P. W. (1979) The extreme-value type – 1 distribution and the principle of maximum entropy, *J. Hydrol., Vol. 42,* 23-38.

Shannon, C. E. (1948) A mathematical theory of communication. Bell System Techn. J., July and October.

Sonuga, J. O. (1972) Principle of Maximum Entropy in Hydrologic Frequency Analysis, *J. Hydrol., Vol. 17,* 177-191.

Received: 2 January, 1983

Address:
E. S. Oyegoke,
Hydraulics Research Unit,
University of Lagos,
Faculty of Engineering,
Akoka, Yaba,
Lagos, Nigeria.

J. O. Sonuga,
Enplan Group Consulting Engineers,
Lagos, Nigeria.