Mausam, (1989), **40**, 1, 85-90

551.577.2 : 551.501.777 (547.1)

Statistical distribution of extremes of rainfall at Colaba, Bombay

NITYANAND SINGH

Indian Institute of Tropical Meteorology, Pune

(*Received 13 January 1987*)

सार — गुम्बेल और गामा बंटनों का प्रयोग करते हुए 1924 से 1984 तक के बम्बई की कोलाबा वेधशाला में रिकार्ड किये गये 1, 2, 3 6, 12, 24, 36 और 48 घंटों की अधिकतम वार्षिक वर्षा के आवृत्ति बंटन का अध्ययन किया गया है आघूर्ण और अधिकतम संभावित पर्वानुमान का प्रयोग करते हुए दोनों ही उपयुक्त पाए गये हैं। गुम्बेल बंटन आसंजित आघूर्ण विधि के साथ पूर्ण आपेक्षिक विचलन और द्विघातीय विचलन के न्यूनतम मूल्य द्वारा अत्यधिक उपयुक्त पाये गए है, किन्तु अधिकतम संभावित पूर्वानुमानों का प्रयोग करते हुए गुम्बेल बंटन के आसंजन समान रूप से सन्तोषप्रद है। बम्बई के लिये वर्षा की अवधि-गहराई-आवृत्ति सम्बन्ध बताते हुए चरम मान सम्भाव्यत: कागज पर परिणाम दिखाए गये है। चरम वर्षा राशि तथा अवधि राशि को आसानी से समझने के लिये माध्य वार्षिक अधिकतम वर्षा और अवधि के मध्य सम्बन्ध निकाला गया है। सम्बन्ध 5, 10, 15 और 30 मिनट की अत्यधिक वर्षा के लिये मान्य पाए गए हैं।

ABSTRACT. Frequency distribution of annual maximum 1, 2, 3, 6, 12, 24, 36 and 48-hr rainfall recorded at Colaba observatory, Bombay from 1924 to 1984 is studied using Gumbel and gamma distributions—both have been fitted using moments and maximum likelihood estimates separately. Gumbel distribution fitted with method of moments is found to be most suitable as indicated by minimum values of the absolute relative deviation and quadratic deviation but the fit of Gumbel distribution using maximum likelihood estimates was, by and large, equally satisfactory. The results are presented on an extreme-value probability paper giving rainfall depth-duration-frequency (DDF) relationships for Bombay. Relationship between mean annual maximum rainfall and the duration has been obtained in order to simplify the understanding of extreme rainfall amount and duration. The relationship is found valid for 5, 10, 15 and 30-min extreme rainfall also.

1. Introduction

Intense rainfall causes dislocation of normal life in the city of Bombay. The city with a population of about 9 million is situated along the west coast of Indian Peninsula (the geographical coordinates are 18° 54′N and 72° 49′E). The mean annual rainfall of Bombay is 1805 mm, 94% of which is received during southwest monsoon period from June through September. The problem of flooding in the city could be avoided by designing suitable drainage system and for this purpose knowledge of distribution of short duration extreme rainfall events is essential. Thus, the study proposes to examine the frequency distribution of 1, 2, 3, 6, 12, 24, 36 and 48-hr annual maximum rainfall.

Frequency analysis of annual maximum rainfall is done to predict the extreme rainfall amount for large return periods—larger than the period of available records. The major task of the problem is the choice of suitable probability model which could fit most closely the frequency distribution of historical rainfall data. There are numerous theoretical probability models suggested to be used for frequency analysis of extreme precipitation (Sevruk and Geiger 1981). The choice of appropriate model for a particular rainfall series depends largely on considerations of convention and experience of analysing the rainfall data. So in this process choice of more than one model is obvious. In the present study Gumbel and gamma distribution models have been utilized. The results are presented on a diagram as rainfall depth-duration-frequency (DDF) relationships for Bombay. An attempt has also been made to understand distribution of 5, 10, 15 and 30-min extreme rainfall from hourly extreme rainfall distributions.

2. Data used

Hourly (clock-hour) rainfall data for Colaba observatory, Bombay from 1924 to 1966 has been obtained from the publication entitled '*Magnetic meteorological and atmospheric electricity observations made at the government observatories, Bombay and Alibag*' and from 1967 to '84, it is collected partly from I.M.D., Poona and partly from Regional Meteorological Centre, Bombay. Thus, the study is based on 61 years (1924-84) continuous records of hourly rainfall which is longest (as it is known to author) record of hourly rainfall for any station in India. Annual maximum rainfall series for 1, 2, 3, 6, 12, 24, 36, and 48-hr durations are formed by picking highest rainfall of different durations each year and their frequency distribution have been studied. The durations are any continuous clock hours and not sticking to calendar dates. Hereafter, annual maximum rainfall series will be referred to as rainfall series.

TABLE 1

Estimates of parameters of Gumbel distribution and the *T*-yr rainfall amount (in mm) for annual maximum hourly rainfall at Colaba, Bombay. MM denotes method of moments and MML method of maximum likelihood

Duration (hour)	Method	Estimates of parameters		Return period (yr)						
		α	u	2	5	10	20	50	100	200
1	MM	16.7212	48.3466	54.5	73.4	86.0	98.0	113.6	125.3	136.9
	MML	15.5115	48.6196	53.3	71.9	83.5	94.7	109.1	120.0	130.8
2	MM	29.015	68 3813	79 0	111.9	133.7	154.6	181.6	201.8	222.0
	MML	25.2906	69.6076	78.9	107.5	126.5	144.7	168.3	185.9	203.5
3	MM	36.4407	82.6227	96.0	137.3	164.6	190.8	224.8	250.3	275.6
	MML	31.4309	84.0915	95.6	131.2	154.8	177.4	206.7	228.7	250.5
6	MM	52.7337	107.8539	127.2	186.9	226.5	264.5	313.6	350.4	387.1
	MML	42.1786	111.2883	126.7	174.5	206.2	236.6	275.9	305.3	334.7
12	MM	76.8214	138.3122	166.5	253.5	311.2	366.5	438.1	491.7	545.1
	MML	56.6774	145.7463	166.5	230.7	273.3	314.1	366.9	406 5	445.9
24	MM	97.6687	183.1853	219.0	329.7	403.0	473.3	564.3	632.5	700.4
	MML	73.2463	192.0016	218.8	301.9	356.8	409.5	477.8	528.9	579.9
36	MM	109.442	225.707	265.8	389.9	471.9	550.8	652.7	729.2	805.3
	MML	87.6752	233.525	265.6	365.0	430.8	493.9	575.6	636.8	697.8
48	MM	112.6996	257.9429	299.2	427.0	511.5	592.7	697.7	776.4	854.8
	MML	95.2424	264.1826	299.1	407.2	478.7	547.4	636.2	702.8	769.1

TABLE 2

Estimates of parameters of gamma distribution and the *T*-yr rainfall amount (in mm) for annual maximum hourly rainfall at Colaba, Bombay

Duration (hour)	Method	Estimates of parameters		Return period (yr)						
		β	υ	2	5	10	20	50	100	200
1	MM	7.9299	7.3138	55.4	74.8	86.6	97.2	110.0	119.1	127.8
	MML	6.8778	8.4326	55.7	73.8	84.6	94.3	105.9	114.2	122.1
2	MM	16.2678	5.2330	79.8	113.8	134.8	154.1	177.7	194.6	210.9
	MML	13.0827	6.5070	80.8	111.2	129.7	146.4	166.8	181.3	195.2
3	MM	21.0739	4.9187	96.7	139.6	166.2	190.5	220.5	241.9	262.7
	MML	16.9887	6.1015	98.0	136.3	159.8	180.9	206.8	225.3	243.0
6	MM	33.0782	4.1808	127.4	189.7	228.9	265.0	309.6	341.8	372.0
	MML	24.6472	5.6109	130.2	183.5	216.4	246.2	282.7	308.8	333.9
12	MM	53.1482	3.4367	165.3	256.4	314.8	368.9	436.5	485.4	533.0
	MML	35.9126	5.0816	170.8	245.0	291.1	333.0	384.5	421.5	457·2
24	MM	65.5014	3.6573	218.1	333.8	407.5	475.6	560.5	621.9	681.6
	MML	45.3847	5.2784	224.6	320.1	379.1	432.8	498.6	545.9	591.5
36	MM	68.2044	4.2355	266.5	395.7	477.0	551.6	644.0	710.6	775.0
	MML	50.7254	5.6950	272.2	382.8	450.8	512.4	587.9	641.8	694.0
48	MM	64.6862	4.9933	301.7	434.2	516.5	591.5	683.8	750.0	814.0
	MML	50.7146	6.3689	306.2	422.9	494.0	558.2	636.5	692.4	746.3

3. Frequency analysis

3.1. *Normality test*

Before applying any complicated probability model of frequency distribution, each rainfall series is examined for its normality aspect by computing Fisher's 'g' statistic (both 'g_1' and 'g_2' for skewness and kurtosis respectively) and their standard errors. If either or both of 'g_1' and 'g_2' are found more than 1. 96 times of its standard error the series is treated as significantly different from normal. The test is performed on actual as well as on square-root, cube-root and logarithmically transformed values of different rainfall series and it is found that except logarithmic transformation of 1, 2 and 3-hr rainfall series they are significantly different (at 5% level) from normal.

3.2. *Gumbel distribution*

3.2.1. *Model description* — Gumbel extreme value distribution has been widely used in hydrological frequency analysis (Sevruk and Geiger 1981). Its cumulative distribution function is given by :

$$P(x) = \exp[-\exp-(x-u)/\alpha] \quad (1)$$
$$-\infty < x < \infty \; ; \; -\infty < u < \infty \; ; \; \alpha > 0$$

where, α and u are respectively, the shape and location parameters of the distribution.

3.2.2. *Parameter estimation* — In fitting Gumbel distribution to annual maximum rainfall at Bombay, the parameters estimated by moments and maximum likelihood methods have been utilized separately.

The moment estimators of the parameters of Gumbel distribution are as follows :

$$\alpha = (\sqrt{6}/\pi)\,[\Sigma\,(x-\bar{x})^2/(N-1)]^{\frac{1}{2}} \quad (2)$$

$$\text{and} \quad u = \bar{x} - 0.577216\,\alpha \quad (3)$$

where $\bar{x}$ is the mean of the variate x, and N the sample size. This method is very popular in practical use because of computational simplicity.

Relatively complicated method of maximum likelihood for parameter estimation is used with computer aid. In this study maximum likelihood estimates of parameters of Gumbel distribution are obtained using Jenkinson method (Clarke 1973). The method involves first estimate of α and u to initiate the computation for solving the equations :

$$\bar{\alpha} = \bar{x} - [\sum_1^N x_i \exp(-x_i/\alpha)] \,/\, [\sum_1^N \exp(-x_i/\alpha)] \quad (4)$$

$$\text{and} \quad \bar{u} = -\alpha \ln [\sum_1^N \exp\{-(x_i/\alpha)\} / N) \quad (5)$$

Moments estimates of α and u are used as first estimate to initiate the computation. The estimate of $\bar{\alpha}$ and $\bar{u}$ from Eqns.(4) and (5) will give corrections to be applied to the earlier considered value of α and u. The iteration process is discontinued when $\bar{\alpha}$ and $\bar{u}$ became ≤ 0.01. The convergence is very rapid and less than 10 iterations are required to achieve the desired accuracy. The estimated parameters by moments and maximum likelihood method for different rainfall series are given in Table 1. It is interesting to note that estimate of α by method of moments is greater than that by the method of maximum likelihood, while u by maximum likelihood method is greater in all cases.

3.2.3. *Estimated T-yr rainfall* — In Eqn. (1), in terms of reduced variate y the term $(x-u)/\alpha$ can be written as follows :

$$x = u + \alpha y \quad (6)$$

$$\text{where,} \quad y = -\ln[\ln\{T/(T-1)\}] \quad (7)$$

$$\text{in which} \quad T = \frac{1}{1-P(x)} \quad (8)$$

The rainfall estimates for different durations with 2, 5, 10, 20, 50, 100 and 200-year return periods have been computed using Eqns. (6) & (7) considering estimated α and u by moments and maximum likelihood methods separately and are given in Table 1.

3.3. *Gamma distribution*

3.3.1. *Model description* — It is shown in numerous studies by Mooley and Appa Rao (1970, 1971) and Mooley [1973(a) & (b)], that gamma distribution gives good fit to rainfall of different time scales at Indian stations. This led us to investigate the suitability of gamma distribution for annual maximum hourly rainfall distribution at Colaba, Bombay. For gamma distribution the probability density function is given by :

$$P(x) = \frac{x^{\upsilon-1}\, e^{-x/\beta}}{\beta^{\upsilon}\, \Gamma(\upsilon)} \quad (9)$$

where β and υ are scale and shape parameters respectively. Gamma distribution has been fitted to different rainfall series using moments and maximum likelihood estimates of the parameters separately.

3.3.2. *Parameter estimation* — The moments estimators of the parameters of gamma distribution are as follows :

$$\bar{x} = \beta\upsilon \quad (10)$$

$$\text{and} \quad s = \sqrt{\beta^2\upsilon} \quad (11)$$

where $\bar{x}$ and s are the sample's mean and standard deviation respectively. The maximum likelihood estimates of the parameters are determined by solving the following equation numerically:

$$\phi(\upsilon) = \psi(\upsilon) - \ln(\upsilon) - \ln(\bar{x}_G/\bar{x}) = 0 \quad (12)$$

where ln is the natural logarithm, $\bar{x}_G$ and $\bar{x}$ are respectively, the geometric and arithmetic means, and $\psi(\upsilon)$ is the di-gamma function defined as :

$$\psi(\upsilon) = \frac{\partial}{\partial \upsilon} \ln \Gamma(\upsilon)$$

Eqn. (12) is solved using Newton-Raphson iteration procedure. On differentiating Eqn. (12) with respect to υ we get, :

$$\frac{d\phi}{d\upsilon} = \psi'(\upsilon) - \frac{1}{\upsilon} \quad (13)$$

where, $$\psi'(v) = \frac{\partial^2}{\partial v^2} \ln \overline{(v)}$$

is the tri-gamma function. The dia-and tri-gamma functions have been evaluated by means of the algorithm developed by Mooley (1974) which is based on truncated logarithmic form of Sterling's series.

For solving Eqn. (12) the iteration process started beginning with moments estimates as the initial estimate of v and the correction term, h, to be added to v for next iteration is evaluated from equation:

$$h = \frac{-\psi(v) + \ln(v) + \ln(x_G/\bar{x})}{\psi'(v) - (1/v)} \qquad (14)$$

The iteration process is discontinued when h becomes $\leqslant 0.0001$. Having obtained v, the scale parameter β is calculated using Eqn. (10). The estimates of β and v by moments and maximum likelihood methods are given in Table 2. It is seen from the table that for all rainfall series considered estimates of β by the method of moments is higher than that by the method of maximum likelihood whereas opposite is true in case of estimate of v.

3.3.3. *Estimated T-yr rainfall* — The estimated rainfall amount from the derived gamma distribution for different durations with 2, 5, 10, 20, 50, 100 and 200-yr return periods is obtained by evaluating the following integral :

$$P_{xt} = \int_0^{x_t} \frac{x^{v-1} e^{-x/\beta}}{\beta^v \overline{(v)}} dx \qquad (15)$$

where P_{xt} is the probability of rainfall not exceeding x_t and other terms in the equation have their usual meaning. To determine x_t for a given P_{xt} Eqn. (15) is evaluated following the procedure developed by Wilk *et al.* (1962) and used by Mooley (1973). The process involves estimating lower limit of rainfalls, x_{tL}, expected with different return periods $t = T_1, T_2, \ldots\ldots, T_n$. For this purpose Eqn. (15) has been rewritten by putting $x = x_t Z$ as:

$$P_{xt} = x_t^v \int_0^1 \frac{Z^{v-1} e^{-x_t Z/\beta}}{\beta^v \overline{(v)}} dZ$$

Since the maximum value of the exponential function is unity at $Z=0$, therefore,

$$P_{xt} \leqslant x_t^v \int_0^1 \frac{Z^{v-1}}{\beta^v \overline{(v)}} dZ$$

or

$$x_{tL} = \beta\, [\,\overline{(v+1)}\, P_{xt}]^{1/v} \qquad (16)$$

The lower limit of rainfall expected with 2, 5, 10, 20, 50, 100 and 200-yr return periods has been determined from Eqn. (16). For estimating actual rainfall amount for return period T_1 interpolation by halving is done between lower limits of rainfall for return periods T_1 and T_2 so as to satisfy Eqn. (15). Similarly for return period T_2 interpolation is done between lower rainfall limits for T_2 and T_3 return periods and so on. The rainfall amount of different durations expected with selected return periods estimated from gamma distribution fitted using moments and maximum likelihood methods, separately, are given in Table 2.

3.4. *Selection of most suitable fit*

According to χ^2-test Gumbel and gamma distribution both fitted using moments and maximum likelihood estimates gave, by and large, equally satisfactory fit to the annual maximum rainfall of different durations over Colaba. In order to select the most suitable scheme giving uniformly good fit to rainfall series of different durations absolute relative deviation and quadratic deviation have been computed which are defined as:

$$\text{Absolute relative deviation} = \sum_T \left| \frac{x_e - x_0}{x_0} \right| \qquad (17)$$

$$\text{and Quadratic deviation} = \sum_T \left(\frac{x_e - x_0}{x_0} \right)^2 \qquad (18)$$

where x_0 is the observed rainfall obtained using Tuckey's (1962) plotting position formula for selected $T=2$, 5, 10, 20 and 50-year return periods and x_e, the estimated rainfall for the given return periods.

For different rainfall series absolute relative deviation and quadratic deviation both were found minimum in case of Gumbel distribution fitted using moments estimates, implying a most suitable unified fit to annual maximum hourly rainfall at Colaba, Bombay. But results of different methods, particularly Gumbel distribution fitted using maximum likelihood estimates were quite comparable.

4. Rainfall depth-duration-frequency relationships

The T-yr rainfall amount estimated from the Gumbel distribution, fitted with method of moments (Table 1), are now plotted on an extreme value probability paper as shown in Fig. 1, which presents the depth-duration-frequency (DDF) relationships of Bombay rainfall. Eqn. (6) represents these relationships, α and u for different rainfall series are given in Table 1.

Fig. 1 can be used for interpolating the results for durations other than those considered in the analysis. However, the estimated rainfall for higher return periods (more than 200-year) may not be intuitively accepted while looking into the sample size of 61.

It may be interesting to note that for different durations the estimated rainfall for 200-yr return period is in close agreement with the highest observed rainfall. The highest observed rainfall amounts alongwith their time of occurrence and estimated return period are given in Table 3. Thus, for constructing a hydraulic structure which required information on rainfall amount for a given duration with return period upto 200-yr, the highest observed rainfall would suffice the purpose.

5. Relationship between the mean annual maximum rainfall and the duration

Increase in rainfall amount with the increase of duration is obvious. In order to simplify characteristics of extreme rainfall amount, relationship between mean annual maximum rainfall and duration is obtained. For this purpose 16 additional annual maximum rainfall series, *e.g.*, 60,72,----, 240-hr durations have been analysed by fitting Gumbel distribution using moments estima-

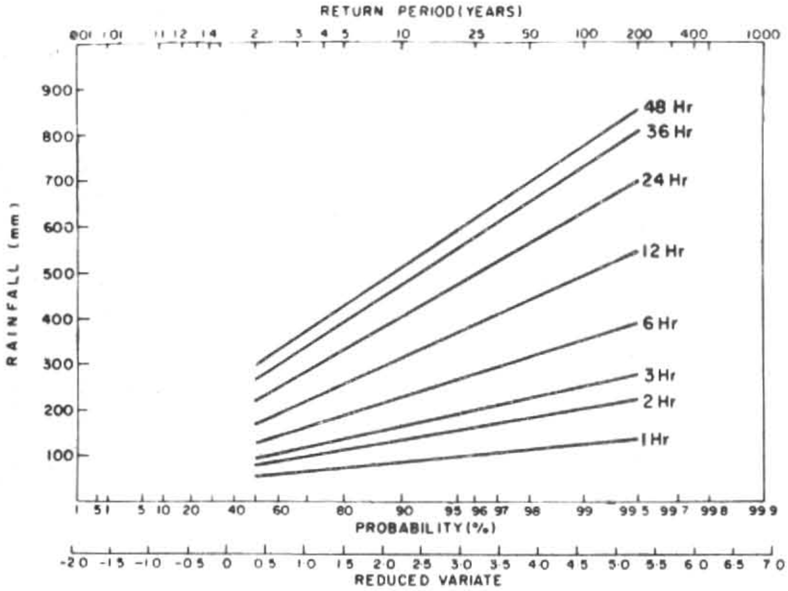

Fig. 1. Rainfall depth-duration-frequency (DDF) relationships for Colaba, Bombay

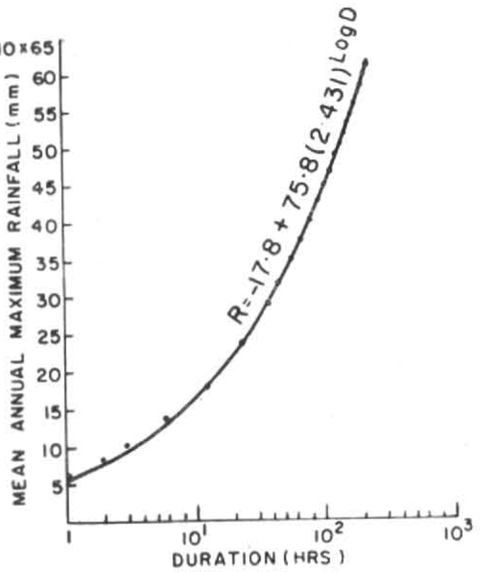

Fig. 2. Relation between the mean annual maximum rainfall of Colaba, Bombay and duration

TABLE 3

Information relating to highest observed rainfall at Colaba, Bombay for different durations

Duration (hr)	Rainfall (mm)	Starting time (hr)	Date	Estimated return period (yr)
1	128.5	1500	22-9-1949	121
2	228.6	1500	22-9-1949	250
3	269.5	1500	22-9-1949	169
6	386.1	0200	10-9-1930	196
12	571.2	0500	22-9-1949	281
24	768.3	2000	21-9-1949	400
36	798.8	1000	21-9-1949	188
48	828.3	2300	20-9-1949	158

tes as discussed earlier in this paper. The fit has been found to be good for different rainfall series. Thus, the following analysis are based on 24 selected annual maximum rainfall series.

Taking duration along abscissa on logarithmic scale and the mean rainfall on arithmatic scale along ordinate, the plot for 24 annual maximum rainfall series will follow the curve as shown in Fig. 2. The shape of the curve is of a modified exponential type (Croxton *et al.* 1975) and its equation is given by :

$$R = -17.8 + 75.8 \ (2.431)^{\log D} \qquad (19)$$

where R is the mean annual maximum rainfall in mm and D, the duration in hours.

Eqn. (19) is obtained using hourly data only to see its performance for durations less than 1-hr. The mean annual maximum rainfall for 5, 10, 15 and 30 min durations has been determined from the equation and it is compared with the mean annual maximum rainfall based on data from 1933 to '67. The results are found in very close agreement, implying the applicability of Eqn. (19) for minute rainfall also. The data for 5, 10, 15 and 30-min durations are obtained from a publication by Patel (1969), where the author has shown the annual maximum rainfall of these durations to follow the Gumbel distribution.

6. Conclusion

(*i*) Gumbel and gamma distributions both fitted with moments and maximum likelihood methods separately gave, by and large, equally satisfactory fit to annual maximum rainfall of selected durations varying from 1 to 48 hr at Colaba, Bombay. Though Gumbel distribution fitted with method of moments is found to be most suitable according to absolute relative deviation and quadratic deviation computations, the results of Gumbel distribution based on maximum likelihood estimates were, by and large, equally satisfactory.

(*ii*) There exists an exponential type relationship between mean maximum rainfall and the duration for Colaba, Bombay.

Acknowledgements

The anthor is extremely grateful to Dr. G.B. Pant for his keen interest and encouragement and to the Director, Indian Institute of Tropical Meteorology, Pune for necessary facility to pursue the study. He is also thankful to his colleagues Mr. S.S.Aralikatti and Mr. M.K. Soman for their cooperation in data aquisition and computer programming. The manuscript has been neatly typed by Miss S. S. Nandargi.

References

Clarke, R. T., 1973, '*Mathematical models in hydrology*', F. A. O., Irrigation and Drainage Paper, 19.

Croxton, F. E., Cowden, D. J. and Klein, S., 1975, '*Applied general statistics*' (3rd edition), Prentice-Hall of India Pvt. Ltd., New Delhi, p. 264.

Mooley, D. A. and Appa Rao, G., 1970, 'Statistical distribution of pentad rainfall over India during monsoon season', *Indian J. Met. Geophys.*, **21**, 2, pp. 219-230.

Mooley, D. A. and Appa Rao, G., 1971, 'Distribution function for seasonal and annual rainfall over India', *Mon. Weath. Rev.*, **99**, 10, pp. 796-799.

Mooley, D. A., 1973 (a), 'Gamma distribution probability model for Asian summer monsoon monthly rainfall', *Mon. Weath. Rev.*, **101**, 2, pp. 160-176.

Mooley, D. A., 1973 (b), 'Some aspects of Indian monsoon depressions and the associated rainfall', *Mon. Weath. Rev.*, **101**, 3, pp. 271-280.

Mooley, D.A., 1974, 'Suitable tables for application of gamma probability model to rainfall', Res. Rep. No. RR-018, Indian Institute of Tropical Meteorology, Poona.

Natural Environment Research Council, 1975, *Flood studies report—Hydrological studies*', **1**, pp. 85-87.

Patel, A. I., 1969, 'Frequency analysis of rainfall intensities for Bombay (Colaba) and Nagpur (Sonegaon),' M.Sc. (Ag.) Thesis, Maharashtra Krishi Vidyapeeth.

Sevruk, B. and Geiger, H., 1981, 'Selection of distribution types for extremes of precipitation', WMO No. 560.

Wilk, M. B., Gnanadesikan, R. and Huyett, M.J., 1962, 'Probability plots for gamma distribution', *Technometrics*, **4**, 1, pp. 1-20.