551·577·2/·36

The Cumulative Frequency Distribution of Rainfall of different intensities

N. S. KULKARNI and M. B. PANT

Institute of Tropical Meteorology, Poona

(*Received 4 November 1966*)

ABSTRACT. The cumulative frequency distribution of rainfall of different intensities during southwest monsoon is studied for twenty stations in India. The distribution is found to be exponential in all the cases.

The equation $Y = ae^{-bX}$ (where Y is the cumulative frequency of daily rainfall exceeding an amount X and a, b are constants) is very satisfactorily found to fit the data. The curve has been fitted to the observed data by the method of least squares. It has also been shown that the constants a and b can be found from the number of rainy days and the total seasonal rainfall.

The variations of the parameters a and b with rainfall and number of rainy days are discussed.

1. Introduction

An idea of how rainfall at any place is distributed according to intensity is important and some investigations have been made on this line. To mention a few, Ashmore (1944) studied the contribution to the total rainfall by the wettest n days (7, 12, 15, 20, 30, 40, 50, 60) at a station in Wrexham District, UK and suggested the relation $P = \sqrt{90n - 300}$ where P is the percentage of annual rainfall and n the number of wettest days. Olascoaga (1950) studied on similar lines the rainfall distribution in a number of regions in Argentina and suggested the equation $Y = ae^{-bX}$ where Y is the cumulative percentage rainfall and X is the cumulative percentage rainy days and a, b are contants.

In India a few studies have been made in this direction. Doraiswamy Iyer and Zafar (1937) have prepared the maps of frequency of occurrence of falls of 3″ and 5″ and above. Rao (1959) presented similar maps of falls of 10″ and above. Jagannathan and Raghavendra (1966) studied the frequency distribution of rainfall of different intensities during the southwest monsoon in Rajasthan by the frequency function $f_r = \alpha \dfrac{x^r}{r}$ where α and x are parameters representing the intrinsic character of the frequency distribution and r is the rainfall interval. Another important paper in this respect is by Subbaramayya and Rao (1964), in which a theoretical approach has been made to find the frequency distribution of rainfall in different parts of India. They found that the distribution is hyperbolic in all the cases and the equation $Y = \dfrac{a}{x+c} + b$ where x is the rainfall per day, ydx is the frequency of rainfall in the interval x to $x+dx$ and a, b and c are constants, is very satisfactorily found to fit into the data. The constants a, b and c have been found theoretically from the rainfall amount and the number of rainy days. Also the equation has been applied to daily rainfall data of individual stations and the fit is found to be correct by more than 90 per cent accuracy.

We present here a similar study on the cumulative frequency distribution of rainfall of different intensities for some stations in India. The distribution is found to be exponential and the curve $Y = ae^{-bX}$ is found to fit into the data satisfactorily where Y is the cumulative frequency of rainfall exceeding an amount X. The variation of the parameters a and b with rainfall, number of rainy days have also been discussed.

2. Method of analysis

The country receives most of its annual rainfall during the summer monsoon (southwest monsoon) season. As such the daily rainfall in this season (June to September) only is examined in the present investigation.

We have restricted our study to a few stations. Their positions are shown in Fig. 1. We have stations on windward side as well as on lee side. We have taken coastal stations and also inland stations.

The data for the 50-year period 1901–50 is studied. The daily (24-hr) intensity of rainfall in cents is divided into class intervals namely, ⩾ 10 cents but less than 40 cents; ⩾40 cents but less than 80 cents; ⩾80 cents but less than 120 cents and so on. The frequencies for each class are found

TABLE 1

Observed (O) and computed (C) cumulative frequency distribution

Rainfall amounts ⩾ (cents)	Trivandrum		Cochin		Bhagamandala		Karwar		Vengurla		Kodaikanal		Mercara		Poona		Buldhana		Hoshangabad	
	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)
10	53·0	43·6	83·0	71·7	105·0	93·8	89·0	73·6	86·0	73·0	43·0	35·8	96·9	80·8	35·0	29·2	42·0	33·9	49·0	41·0
20	39·5	37·6	71·0	64·7	98·0	89·3	77·0	68·4	73·0	67·6	31·0	29·2	85·0	74·3	23·0	24·7	31·0	29·5	41·0	37·1
30	31·0	32·4	62·0	58·3	92·0	84·9	70·0	63·6	64·0	62·5	23·5	23·8	76·0	68·2	17·0	21·0	24·5	25·6	31·5	33·6
40	25·6	27·9	54·4	52·5	87·1	80·7	62·4	59·1	56·0	57·9	18·5	19·4	68·6	62·7	13·9	17·9	20·5	22·2	29·3	30·4
80	13·2	15·1	33·1	34·7	68·3	66·3	42·1	44·2	36·5	42·5	8·1	8·6	45·5	44·6	6·8	9·3	10·7	12·6	18·3	20·4
120	7·2	8·5	21·8	22·9	53·5	54·3	30·1	33·0	26·0	31·3	3·5	3·8	31·0	31·7	3·7	4·8	6·0	7·2	12·0	13·6
160	4·1	4·7	14·1	15·1	42·7	44·5	22·9	24·6	20·0	23·0	1·4	1·7	20·2	22·6	2·2	2·5	3·5	4·1	8·8	9·1
200	2·3	2·6	9·5	10·0	35·7	36·5	17·1	18·4	15·0	16·9	0·7	0·7	14·6	16·1	1·2	1·3	2·0	2·3	6·5	6·1
240	1·5	1·4	6·1	6·6	29·5	29·9	12·9	13·7	11·4	12·4	0·4	0·3	10·4	11·4	0·6	0·7	1·7	1·3	4·5	4·1
280	1·0	0·8	4·1	4·3	24·3	24·6	9·7	10·3	8·7	9·1			7·7	8·1	0·5	0·3	1·0	0·7	3·8	2·7
320			2·9	2·9	19·3	20·1	7·4	7·7	6·5	6·7			5·3	5·8	0·2	0·2			2·4	1·8
360			2·0	1·9	15·1	16·5	5·9	5·7	4·9	4·9			3·9	4·1					1·9	1·2
400			1·3	1·3	12·5	13·5	4·2	4·3	3·4	3·6			2·7	2·9					1·3	0·9
440					10·5	11·1	3·2	3·2	2·5	2·7			2·2	2·1						
480					8·7	9·1	2·4	2·4	1·8	1·9			1·6	1·5						
520					7·3	7·5	1·9	1·8	1·5	1·4			1·2	1·1						
560					6·3	6·1	1·3	1·3	1·0	1·1										
600					5·3	5·0	1·1	1·0												
640					4·3	4·1														
680					3·4	3·4														
720					2·9	2·8														
760					2·4	2·3														

Computed from the equation $Y = ae^{-bX}$

out from the daily data for each month June to September and then totalled. The frequencies in each class interval are found for each year from 1901 to 1950. From the actual frequencies were found for each year the cumulative frequencies exceeding amount 40 cents, 80 cents etc. The mean of the cumulative frequencies for the years 1901–50 which form the basic data of our analysis are given in Table 1 for each station.

2.1. *Cumulative frequency distribution* — A typical graph of cumulative frequency distribution against intensity for Bhagamandala is shown in Fig. 2. It is clear from the graph that the cumulative frequency distribution is exponential in nature. It has been ascertained that the distribution is similar for all the stations considered. Accordingly we have tried to fit an exponential curve of the form —

$$Y = ae^{-bX} \qquad (1)$$

where Y is the frequency of rainfall exceeding an amount X and a and b are constants in the range $0 < X < \infty$.

We have found the constants a and b by fitting to the observed Y and X values the straight line —

$$\log_{10} Y = \log_{10} a - Xb \log_{10} e \qquad (2)$$

by the method of least squares. The expected distribution for Bhagamandala is shown by the continuous curve in Fig. 2. Also the expected values for the ten stations for which we have fitted the equation (1) are shown side by side the observed values in Table 1.

It would be seen from Table 1 that the curve does not fit well for rainfall below 20 cents for stations having less rainfall and for rainfall below 80 cents for stations having more rainfall. In general, the curve does not fit well below 40 cents of rainfall. The computed cumulative frequency is an underestimate of the actual cumulative frequency below this range. Apparently, the cumulative frequency distribution below about 40 cent is not strictly governed by the exponential curve we have presented. This requires further examination.

The actual observed frequencies in the different ranges of rainfall and the calculated frequencies are given in Table 2. The correlation coefficients between the actual and calculated frequencies and the percentages of variations accounted for by the curve are also shown in this table. The correlation coefficients are in general highly significant. The percentages of variations for 6 stations vary from 80 to 96 per cent. For the remaining 4 stations, namely, Trivandrum, Poona, Buldhana and Hoshangabad the percentages of

Fig. 1. Observatory stations under study

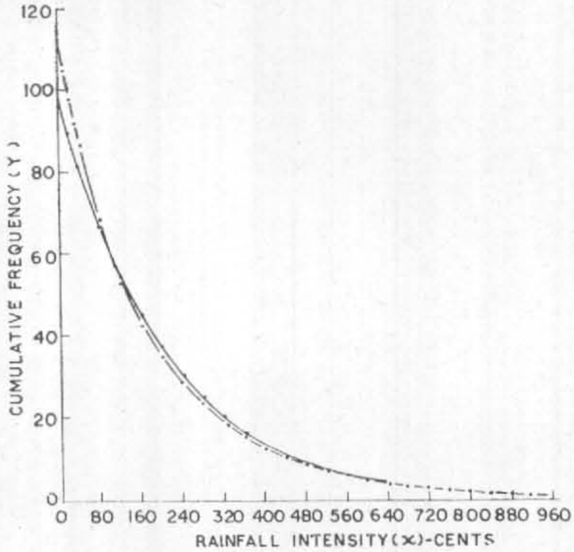

Fig. 2. Rainfall intensity *vs* cumulative frequency—Bhagamandala

The dotted curve is an actual plot of the observed cumulative frequencies. The continuous curve represents the curve $Y = a e^{-bX}$ with the constants computed from the dotted curve

variations accounted for by the curve are respectively 69, 53, 63 and 65. However if we omit the frequency in the range upto 20 cents for which fit is not good, the percentages of variations for these stations become 95, 97, 91 and 98 respectively. One is thus reasonably justified to infer that the exponential curve (1) is a good fit for cumulative frequency for rainfall exceeding 20 cents or so.

Also in order to see how good the fit is when applied to independent data, we have examined the actual frequencies in the different ranges during the period 1951-60 for the two stations, *viz.*, Karwar and Hoshangabad. It is seen that the values of the parameters a and b computed from the mean frequencies for the period 1951-60 for both the stations compare fairly well with those obtained for the period 1901-50. For example, values of a and b for Karwar for the period 1951-60 are 84·45 and 0·0073 respectively compared to the values 79·16 and 0.0073 for the period 1901-50. For Hoshangabad these values for 1951-60 are 41·54 and 0·0100 compared to the values 45·36 and 0·0100 for the period 1901-50. Alternatively, comparison of the actual mean frequencies during the period 1951-60 with the computed frequencies in the different ranges for the period 1951-60 by the curve (1) shows that there is good agreement. The correlation coefficients of the actual frequencies during 1951-60 with the frequencies observed from the curve (1) are 0·9550 for Karwar and 0·9522 for Hoshangabad, which are highly significant. It thus appears that the curve (1) fits quite well when applied to independent data also.

2.2. *Relation between observed rainy day and computed rainy day*—We have seen before that cumulative frequency curve (1) is an underestimate of the actual cumulative frequency for rainfall below 40 cents or so. In the India Meteorological Department a rainy day is defined as a day with rainfall $\geqslant$ 10 cents. According to the curve (1) the number of rainy days is given by $N_c = ae^{-10b}$. We compare this computed number of rainy days with observed rainy days N. These values along with the ratio N/N_c are given in Table 3 for the ten stations for which we have fitted the curve (1).

It would be clear from the table that the ratio N/N_c can be approximately taken to be constant and equal to 1.2. The exponential curve (1) thus underestimates the number of observed rainy days by about 16 per cent.

2.3. *Determination of a and b from seasonal rainfall R and the number of rainy days N*—We now utilise the above relation to find a and b from seasonal rainfall R and rainy days N in the following way. On the cumulative frequency curve let (X, Y) and $(X+\delta x, Y+\delta y)$ be two points. Then the actual frequency of rainfall between X and $X+\delta x$ is $-\delta y$ and the amount of rainfall for intensities between X and $X+\delta x$ is $-X\,\delta y$. Thus

TABLE 2

Observed (O) and computed (C) actual frequency distribution

Rainfall (cents)	Trivandrum		Cochin		Bhagamandala		Karwar		Vengurla		Kodaikanal		Mercara		Poona		Buldhana		Hoshangabad	
	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)	(O)	(C)
10—20	13·5	6·0	12·0	7·0	7·0	4·5	12·0	5·2	13·0	5·4	12·0	6·6	11·9	6·5	12·0	4·5	11·0	4·4	8·0	3·9
20—30	8·5	5·2	9·0	6·4	6·0	4·4	7·0	4·8	9·0	5·1	7·5	5·4	9·0	6·1	6·0	3·7	6·5	3·9	9·5	3·5
30—40	5·2	4·5	7·6	5·8	4·9	4·2	7·6	4·5	8·0	4·6	5·0	4·4	7·4	5·5	3·1	3·1	4·0	3·4	2·2	3·2
40—80	12·6	12·8	21·3	17·8	18·8	14·4	20·3	14·9	19·5	15·4	10·4	10·8	23·1	18·1	7·1	8·6	9·8	9·6	11·0	11·0
80—120	6·0	6·6	11·3	11·8	14·8	12·0	12·0	11·2	10·5	11·2	4·6	4·8	14·5	12·9	3·1	4·5	4·7	5·4	6·3	6·8
120—160	3·1	3·8	7·7	7·8	10·8	9·8	7·2	8·4	6·0	8·3	2·1	2·1	10·8	9·1	1·5	2·3	2·5	3·1	3·2	4·5
160—200	1·8	2·1	4·6	5·1	7·0	8·0	5·8	6·2	5·0	6·1	0·7	1·0	5·6	6·5	1·0	1·2	1·5	1·8	2·3	3·0
200—240	0·8	1·2	3·4	3·4	6·2	6·6	4·4	4·7	3·6	4·5	0·3	0·4	4·2	4·7	0·6	0·3	0·3	1·0	2·0	2·0
240—280	0·5	0·6	2·0	2·3	5·2	5·3	3·2	3·4	2·7	3·3			2·7	3·3	0·1	0·3	0·7	0·6	0·7	1·4
280—320			1·2	1·4	5·0	4·5	3·2	2·6	2·2	2·4			2·4	2·3	0·3	0·2			1·4	0·9
320—360			0·9	1·0	4·2	3·6	1·5	2·0	1·6	1·8			1·4	1·7					0·5	0·6
360—400			0·7	0·6	2·6	3·0	1·7	1·4	0·5	1·3			1·2	1·2					0·6	0·3
400—440					2·0	2·4	1·0	1·1	0·9	0·9			0·5	0·8						
440—480					1·8	2·0	0·8	0·8	0·7	0·8			0·6	0·6						
480—520					1·4	1·6	0·5	0·6	0·3	0·5			0·4	0·4						
520—560					1·0	1·4	0·6	0·5	0·5	0·3										
560—600					1·0	1·1	0·2	0·3												
600—640					1·0	0·9														
640—680					0·9	0·7														
680—720					0·5	0·6														
720—760					0·5	0·5														
760—800																				
C. C. between actual & computed frequency	**0·83		**0·96		**0·98		**0·93		**0·89		**0·90		**0·97		*0·72		*0·79		**0·81	
Percentage of variation of actual frequency accounted for by the exponential curve (1)	69·29		92·85		96·20		86·99		78·85		80·44		94·81		52·61		62·87		65·51	

** Significant at 1 per cent level * Significant at 5 per cent level C.C. = Correlation Coeff.

the total precipitation with intensities between X_1 and X_2 is —

$$-\int X\,\delta y = -\int_{X_1}^{X_2} X \frac{dy}{dx}\,dx$$

and the total seasonal rainfall

$$R = -\int_0^{X_0} X \frac{dy}{dx}\,dx \qquad (3)$$

where X_0 is a sufficiently large intensity exceeding which the cumulative frequency distribution curve will have negligible frequency. Thus if we define $Z = -x\dfrac{dy}{dx}$, then the curve Z *versus* X will be the rainfall distribution curve. The area bounded by the curve and the X-axis is the total rainfall and the area between the curve, X-axis and any two ordinates, X_1 and X_2 is the precipitation contributed by the falls of intensities between X_1 and X_2. A typical computed rainfall distribution curve is shown in Fig. 3 for Bhagamandala. It is seen that the rainfall distribution curve invariably shows maximum at a point where $X = R/a$. Imposing this condition on the rainfall distribution curve, *viz.*, $dZ/dx = 0$ when $x = R/a$,

$$\text{we get, } b = a/R \qquad (4)$$

Thus utilising the relations $N/N_c = 1.2$ and $b = a/R$,

$$\text{we get, } N = 1{\cdot}2\ ae^{-10a/R} \qquad (5)$$

Knowing N and R, a can be found from Eq. (5) by graphical method. Then b can be found from Eq. (4), *viz.*, $b = a/R$.

The values of a and b so computed for the ten stations are given side by side with the values computed by the statistical method in Table 4. It would be seen that the values from the two methods compare very well. The correlation coefficient between the two sets of values of a is 0·9965 and that for b is 0·9825.

For our subsequent study with a and b, we have computed a and b for ten more stations from rainfall and number of rainy days utilising the above method.

3. Behaviour of the frequency curves with a and b

In order to see how the cumulative frequency changes when a and b vary, we have examined a set of hypothetical curves for b varying when a is held constant at 80 and another set of curves for a varying when b is held constant at 0·01 (figures not presented).

It is seen that as b increases the cumulative frequency for any intensity decreases. Also the limit of maximum intensity decreases as b increases. For example for $b = 0·030$ the maximum intensity can just exceed 160 cents whereas for $b = 0·005$, the intensity can exceed 880 cents.

We also see that as a increases the cumulative frequency of rainfall exceeding lower intensity rainfall increases in a greater proportion than the frequencies of higher intensity rainfalls.

4. Relations of a and b with N and R

We have seen before that the parameters a and b can be computed from the number of rainy days N and the seasonal rainfall R. We examine below how a and b vary with N and R separately.

(*i*) *Variation of a with N* — Fig. 4 represents the variation of a with N for windward and leeward stations combined. It is seen that a increases with N. The straight line between a and N has the equation —

$$a = 4·6282 + 0·8606\ N$$

with correlation coefficient 0·9965.

It is also noted that for stations having rainfall less than 50″, a is very nearly equal to N and for more rainfall, a is slightly less than N.

(*ii*) *Variation of a with R*—An examination of the relation of a with R (figure not illustrated) shows that a increases with R. The relation is, however, more prominent for coastal and windward stations than that for leeward and inland stations. The straight line and parabolic fittings for the two cases are as follows —

Windward and coastal	$a = 51·6186 + 0·2503R$
	$a = 37·3719 + 0·5367R - 0·001211R^2$
Leeward and inland	$a = 26·4698 + 0·4333R$
	$a = 13·1566 + 1·4713R - 0·0173R^2$

The multiple correlation coefficients of straight line and parabolic curve fitting are 0·91 and 0·96

TABLE 3

Observed and computed number of rainydays

	N	N_c	N/N_c
Trivandrum	53·0	43·60	1·21
Cochin	83·0	71·71	1·16
Bhagamandala	105·0	93·80	1·12
Karwar	89·0	73·58	1·21
Vengurla	86·0	72·97	1·18
Kodaikanal	43·0	35·84	1·20
Mercara	96·9	80·83	1·20
Poona	35·0	29·16	1·20
Buldhana	42·0	33·95	1·24
Hoshangabad	49·0	41·04	1·19

(Rainy days is one with rainfall ⩾ 10 cents)

TABLE 4

	Computation from statistical method		Computation from rainfall R and rainy days N	
	a (days)	b (cent^{-1})	a (days)	b (cent^{-1})
Trivandrum	50·59	0·0137	51·50	0·0152
Cochin	79·56	0·0104	76·50	0·0099
Bhagamandala	98·58	0·0050	91·50	0·0045
Karwar	79·16	0·0073	79·70	0·0073
Vengurla	78·79	0·0077	77·70	0·0081
Kodaikanal	43·97	0·0205	43·40	0·0190
Mercara	88·00	0·0085	87·50	0·0081
Poona	34·35	0·0164	34·80	0·0177
Buldhana	39·10	0·0141	40·50	0·0141
Hoshangabad	45·36	0·0100	45·00	0·0097

C.C. between values at Col. 2 and Col. 4 is 0·9965 and that between Col. 3 and Col. 5 is 0·9825

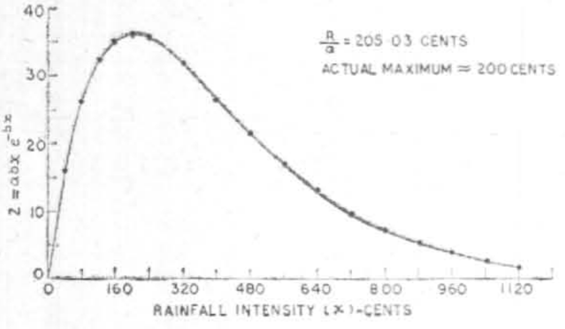

Fig. 3. Rainfall distribution curve — Bhagamandala

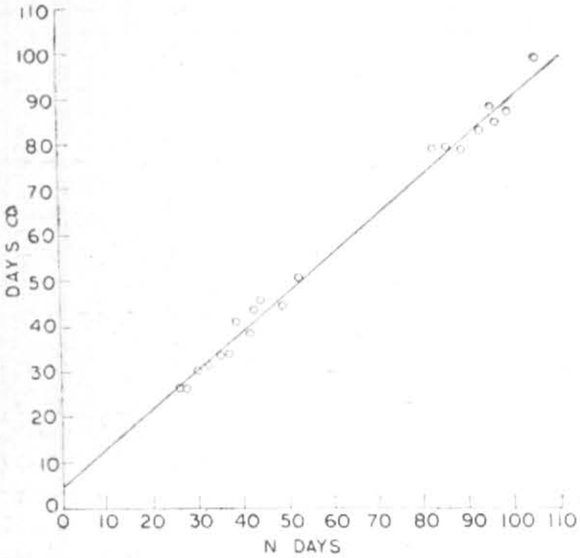

Fig. 4. Variation of a with N

respectively for windward stations. The corresponding values for leeward stations are 0·55 and 0·60.

(*iii*) *Variation of b with N*—It is seen that for coastal and windward stations b decreases as N increases (figure not presented) and the straight-line fit between b and N is —

$$b = 0{\cdot}0231 - 0{\cdot}000172\ N$$

with correlation coefficient 0·97. As we know that $b = a/R$ which can be approximately taken to be N/R, this result appears to indicate that the average rate of daily rainfall increases with the number of rainy days. No such regular pattern appears to exist for leeward and inland stations.

(*iv*) *Variation of b with R*—It is observed that for windward stations b decreases as R increases. Following the argument in the previous paragraph it appears that average daily rainfall increases with total seasonal rainfall. The leeward stations also show in general a similar pattern. The linear relations between b and R for the two cases are as follows —

Windward : $b = 0{\cdot}013444 - 0{\cdot}000048\ R$

Leeward : $b = 0{\cdot}024985 - 0{\cdot}000340\ R$

with the correlation coefficients 0·85 and 0·71 respectively.

5. Conclusions

We can draw the following conclusions from the present study —

(*i*) The cumulative frequency distribution during southwest monsoon is exponential in nature. It is seen that there curve $Y = ae^{-bX}$ applies equally well for coastal as well as inland stations. The constants a and b can be found from the seasonal rainfall R and the number of rainy days N. The exponential curve, however, does not fit well for rainfall below 40 cents or so.

(*ii*) The parameter a increases with number of rainy days and the amount of seasonal rainfall for both coastal and inland stations.

(*iii*) The parameter b decreases both with number of rainy days and amount of seasonal rainfall for coastal stations. From this one can draw an indirect conclusion that average daily increases with number of rainy days as well as with seasonal rainfall. The variation of b for inland stations is irregular.

6. Acknowledgement

We are thankful to Shri P. Jagannathan, Assistant Director and Shri R. P. Sarker, Senior Scientific Officer for encouragement and guidance in this investigation and for help in the preparation of the manuscript.

REFERENCES

Ashmore, S. E.	1944	*Quart. J.R. met. Soc.*, **70**, p. 241.
Doraiswamy Iyer, V. and Zafar Mohamad	1937	*Sci. Notes*, India met. Dep., **7**, 77, pp. 109-118.
Jagannathan, P. and Raghavendra V. K.	1966	*Indian J. Met. Geophys.*, **17**, Spl. No., pp. 197-206.
Olascoaga, M. J.	1950	*Tellus*, **2**. 4, p. 312.
Rao, K. N.	1959	Proc. Symp. on Meteorological and Hydrological Aspects of Floods and Droughts in India, India met. Dep., pp. 11-14.
Subbaramayya, I. and Jaganmohan Rao, N.	1964	*J. met. Soc., Japan*, Ser. II, **42**, 5.