551.577:551.508.77

USE OF COEFFICIENT OF VARIATION IN DETERMINING RAINFALL PROBABILITIES IN A HUMID REGION

Coefficient of variation (C.V.), defined as ratio of standard deviation and mean of a sample, generally expressed in percentage, is an important parameter of Gaussian statistics and it is most commonly used as a criterion of temporal variability of rainfall. For small C.V. rainfall distribution can be interpreted in terms of normal probability model. But C.V. is generally large for rainfall series of shorter time-scales (monthly weekly etc.) and even for annual and seasonal rainfall at stations in arid and semi-arid regions. Hastings (1965), suggested that one need not consider C.V. as a normal statistic at all. The author found C.V. giving real degree of meaning in terms of precipitation probabilities of seasonal precipitation in arid and semi-arid regions of Baja California, U.S.A. The present study prooses to examine the utility of C. V. in determining rainfall probabilities in the humid region by taking example of Kerala over which annual rainfall is highest (around 3000 mm) among the Indian States (Ananthakrishnan *et al.* 1979).

2. The study involves 70 years (1901—70) monthly rainfall data of 80 widely spread stations (Fig. 1) in the State of Kerala. The State occupies total area of 38,864 km^2 along the west coast in the extreme south of Indian Peninsula. It is an elongated state having N-S extension of about 560 km and maximum width in the E-W of about 120 km. In the region the landscape is characterized by the highly uneven topography; it increases from the west sea coast towards inland where some of the isolated mountain peaks are over 2 km high a.s.l. The spatial distributions of rainfall are strongly influenced by these orographic features which has been discussed at length by Ananthakrishnan *et al.* (1979) using nearly the same network of stations.

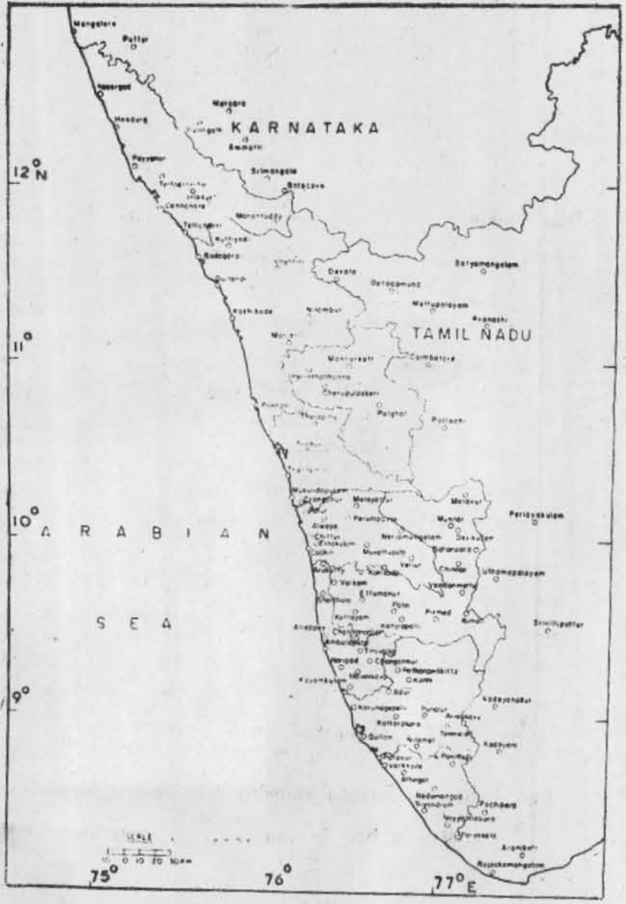

Fig. 1. Locations of rainguage stations in Kerala

Since C.V. of annual rainfall in humid region is low and in order to see the performance of higher C.V. in determining rainfall probabilities, monthly and seasonal data for different stations have also been considered independently. Data for four seasons of winter period (Jan-Feb), hot weather period (Mar-May), southwest monsoon period (June-Sept) and northeast monsoon period (Oct-Dec) form four seasonsal rainfall series for individual stations. Hence for each station there are 17 (12 months, 4 seasons and annual) rainfall series, each of size 70, and, therefore the present analysis is based on 80×17 (=1360) rainfall series for which mean rainfall varies from 4 to 4829 mm and C.V. from 15 to 457 per cent.

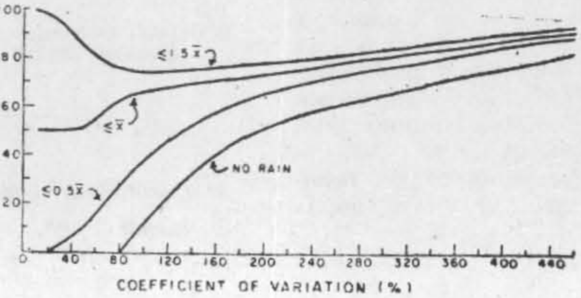

Fig. 2. Relationships between empirical probability of getting rainfall equal to or less than four stated amounts (separately) and the CV

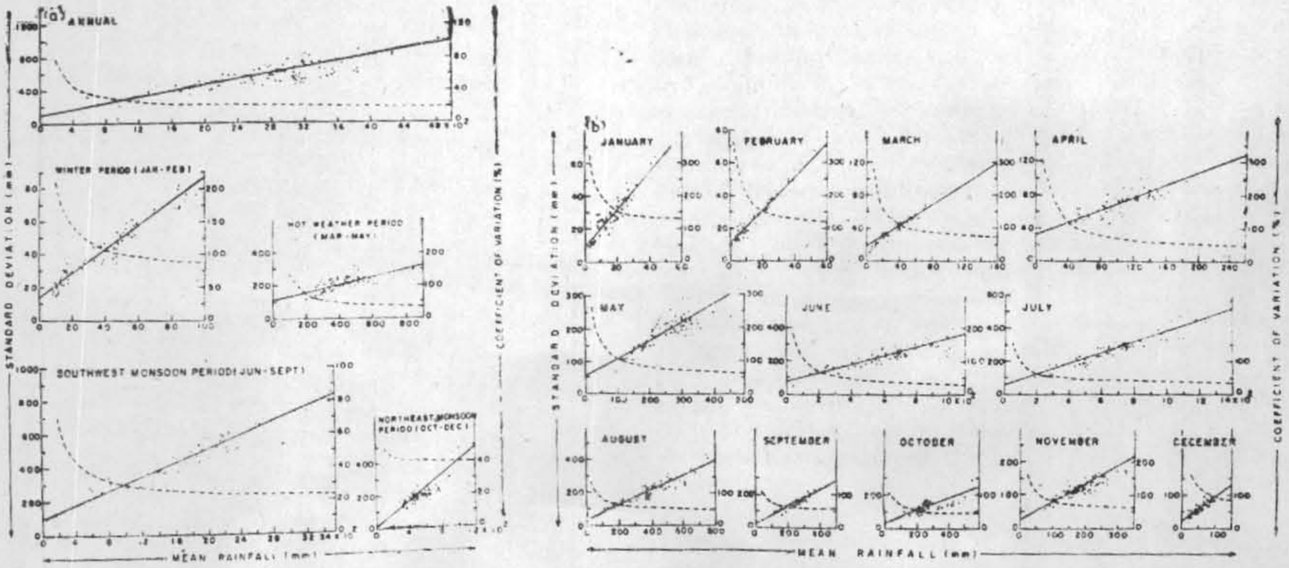

Fig. 3. Scatter diagram showing relationship between standard deviation and mean of annual, seasonal and monthly rainfall in Kerala
—line of best fit and .... variation of CV with the mean rainfall

Hershfield (1962) has shown that C.V. has no upper limit, but instead reaches a maximum value when the "series consists of N—1 items with zero value, and one other item whatsoever". In such a case C.V. $= \sqrt{(N-1)} \times 100$ and is independent of the magnitude of the finite item.

The statistical parameters like mean, standard deviation and coefficient of variation for all the 17 series of different stations have been computed. The empirical probabilities of getting no rain and less than or equal to three pre-fixed limits of rainfall, namely, (1) one-half the mean, (2) the mean and (3) one-and-a-half the

TABLE 1

Coefficients of regression (a & b) of standard deviation on mean rainfall and correlation coefficient (r) for different periods of rainfall in Kerala

Period	a	b	r
January	8.48	1.15	0.924
February	9.94	1.02	0.961
March	18.69	0.61	0.892
April	32.13	0.37	0.684
May	55.01	0.54	0.675
June	79.45	0.26	0.822
July	57.57	0.33	0.849
August	30.40	0.46	0.911
September	25.59	0.52	0.911
October	34.19	0.36	0.748
November	26.00	0.55	0.714
December	25.00	0.71	0.678
Winter	14.11	0.73	0.947
Summer	111.07	0.24	0.542
S-W monsoon	97.33	0.21	0.814
N-E monsoon	10.37	0.40	0.481
Annual	104.21	0.18	0.609

mean for individual months, seasons and annual has been calculated for all the stations. These four sets of probabilities for each rainfall series of each station have been plotted, using four different symbols, against the corresponding C.V. in the same graph. Through the scatter points eye-estimated four curves corresponding to the four specified rainfall amounts have been drawn and are shown in Fig. 2.

3. Although scattering of points between rainfall probabilities and C.V. was large, it did not show any difference among annual, seasonal and monthly rainfall series which shows dependence of rainfall probabilities upon C.V. irrespective of the time scale to which they belong. A brief description of different curves in Fig. 2 is given below :

(*i*) The temporal distribution of a series, exhibiting C.V. of less than 15%, will be between one-half and one-and-a-half of mean rainfall of the series. It may be asked pertinently when limits can be defined how normal distribution is fitted to such rainfall data.

(*ii*) For C.V. upto 80%, probability of zero rainfall in the series is slight. For higher C.V. probability of no-rain, increases non-linearly and becomes 82% at C.V. $\simeq$ 450%.

(*iii*) In case of *C.V. less than 20%* probabilities of rainfall less than one-half mean is slight, it increases non-linearly with C.V. and becomes 90% at C.V. $\simeq$ 450%.

(*iv*) The probability of occurrence of rainfall equal to or less than mean is about 50% for C.V. upto 50%. It increases suddenly from 50 to 65 per cent as C.V. increases from 50 to 95 per cent, thereafter non-linear increase is gradual and becomes 92% at C.V. $\simeq$ 450%. For higher C.V. (more than 50%) the mean starts shifting from median and consequently from the shape of normal curve.

(*v*) The probability of rainfall equal to or less than one-and-a-half times mean decreases from 100 to 74 per cent with increase in C.V. upto 100%. This probability again increase non-linearly with further increases of C.V. (above 100%) and becomes 94% at C.V. $\simeq$ 450%.

Although present results are based on annual, seasonal and monthly rainfall data, it is expected to be applicable for rainfall series of shorter time scales, *e.g.*, weekly and pentad, also. A close agreement between results of present analysis and Hastings (1965) results for semi-arid and arid regions of Baja California, U.S.A, suggests general usefulness of the technique for rainfall series of any time scale and for stations in any region.

4. Further relationship between C.V. and mean rainfall of different periods separately has been examined, since mean rainfall for different time scales for various Indian stations are published elsewhere and it would lead to enormous simplification of the task by representing rainfall probabilities in terms of mean rainfall. Therefore, again we have separated C.V. into mean rainfall and S.D. for each rainfall series. For different periods scatter diagram between mean rainfall and S.D. are given in Fig. 3. In each case straight line trend is apparent and the line of best fit has been obtained by the method of least squares. The variation in C.V. with increase in mean rainfall is also shown in Fig. 3, which is a hyperbolic curve. The coefficients of regression (S.D. $= a+b\,\bar{x}$) of S. D., on the mean rainfall for different periods are given in Table 1. Table 1 also gives simple correlation coefficient (r) between S.D. and mean rainfall which is highly significant (at 99.9% or more significance level).

The regression coefficients for different periods are different and thus for the same amount of mean rainfall S.D. consequently C.V. will be different. Hence, mean rainfall cannot be used for estimating rainfall probabilities. However, in all cases for mean rainfall more than 150 mm, C.V. is less than 100%, and the probability that the series would contain zero values is slight.

5. The important conclusions drawn from the above analysis are as follows :

(*i*) Coefficient of variation, irrespective of its magnitude, is a quite useful statistic for determining rainfall probabilities, the information which is generally required by numerous clientele for practical purposes.

(*ii*) Knowing C.V. for a rainfall series probabilities of no rain, equal to or less than one-half mean, equal to or less than mean and equal to or less than one-and-a-half mean rainfall can be obtained from Fig. 2.

A broad similarity between Hastings results for seasonal precipitation of arid and semi-arid regions of Baja California, U.S.A, and the present one for annual, seasonal and monthly rainfall for a humid region like Kerala suggests general applicability of the technique. Hence, for any rainfall series, C.V. is a useful statistic for determining rainfall probabilities and it should not be looked in the context of Gaussian statistics only.

6. The authors are grateful to the Director, Indian Institute of Tropical Meteorology, Punefor necessary facilities to pursue the study. We are also grateful to Dr. G.B. Pant, Assistant Director, Climatology and Hydrometeorology for his encouragement and keen interest in the study. The manuscript has been neatly typed by Miss S. R. Kamble.

References

Ananthakrishnan, R., Parthasarathy, B. and Pathan, J.M., 1979, 'Meteorology of Kerala', *J. Marine Sci.*, pp. 60-125.

Hastings, J.R., 1965, 'On some uses of non-normal coefficient of variation', *J. appl. Met.*, **4**, August, pp. 475-478.

Hershfield, D.M., 1962, 'A note on the variability of annual precipitation', *J. appl. Met..* **1**, pp. 575-578.

Longley, R.W., 1952, 'Measures of the variability of precipitation', *Mon. Weath. Rev.*, **80**, pp. 111-117.

NITYANAND SINGH
M. K. SOMAN
K. KRISHNA KUMAR

Indian Institute of Tropical Meteorology, Pune
31 March 1986