INTERNATIONAL JOURNAL OF CLIMATOLOGY, VOL. 14, 61–70 (1994) 551.577.32:551.508.77:551.502.3(540)

OPTIMIZING A NETWORK OF RAIN-GAUGES OVER INDIA TO MONITOR SUMMER MONSOON RAINFALL VARIATIONS

NITYANAND SINGH

Indian Institute of Tropical Meteorology, Dr. Homi Bhabha Road, Pashan, Pune 411 008, India

Received 4 August 1992
Accepted 4 March 1993

ABSTRACT

A network of rain-gauges has been optimized to reconstruct the spatially representative summer monsoon rainfall series (1871–1984) of India prepared by averaging the normalized rainfall (actual divided by mean) of 306 rain-gauges. Three approaches have been attempted: (i) random selection, (ii) fitting a multiple linear regression and (iii) selecting objectively a subset of a few gauges, the mean of which shows the highest correlation coefficient (CC) with the all-India series. By averaging a randomly selected 45 rain-gauges, an average CC with the all-India rainfall series of *ca.* 0·94 was achieved, and by averaging 100 gauges the average CC was *ca.* 0·97. By fitting a multiple linear regression between the all-India series and the normalized rainfall series of 306 gauges, a maximum of 34 independent gauges (parameters) were identified and the estimated rainfall series showed a highest CC of 0·9969 with the full series. In the third approach, a subset of 34 rain-gauges was identified following an objective criterion; the mean rainfall of this subset showed a CC of 0·9931 with the all-India series. For routine updating of the all-India summer monsoon rainfall series, however, 35 rain-gauges were selected by applying the third approach to 64 randomly chosen India Meteorological Department (IMD) rain-gauges better known for their timely availability of data. The reconstructed (CC = 0·9866) all-India summer monsoon rainfall series for the period 1871–1992 is presented.

KEY WORDS India summer monsoon rainfall Optimization of rain-gauges Multiple linear regression Correlation

INTRODUCTION

The Asian south-west monsoon of the northern summer during the four months (June to September) of its duration contributes about 75–90 per cent to the annual rainfall total over major parts of India. A long homogeneous time series of the area-averaged June to September total rainfall for the entire country has been found useful for understanding and monitoring, and eventually to predict, the performance of the summer monsoon system over the Indian region. Rainfall is vital for the water-dependent agrarian economy of the country. The construction and analysis of the area-averaged rainfall series (point rainfall studies are of limited practical use) have therefore been important goals of research for the Indian meteorologists. In 1886, Blanford, for the first time, prepared the annual rainfall series of British India for a 19-year period 1867–1885 using data from 500 rain-gauges. Later Walker (1910, 1914, 1922) prepared the south-west monsoon rainfall (June to September total) series for British India for the period 1841–1908 using data from all available rain-gauges, which varied from 40 to 2000. After a gap of about 50 years, rainfall fluctuation analysis for the independent India was carried out by Parthasarathy and Dhar (1976), Parthasarathy and Mooley (1978), and Mooley and Parthasarathy (1979) considering variable (with time) networks of rain-gauges. Using data from a fixed and well-spread network of 306 rain-gauges, Mooley and Parthasarathy (1984) have prepared an all-India summer monsoon rainfall series for the period 1871–1978, which was later extended up to 1984 by Parthasarathy *et al.* (1987). Sontakke *et al.* (1992) have also constructed a summer monsoon rainfall series of India for the period 1813–1991.

For preparing representative area-averaged rainfall series, researchers have used a number of approaches. Using data of 306 rain-gauges, Singh *et al.* (1991) prepared the all-India summer monsoon rainfall series for

0899–8418/94/010061–10$10.00

(1519)

the period 1871–1984 by nine different methods and quantified their areal representation by computing an index, named the index of areal representativeness (IAR). The IAR is defined as follows:

$$\mathrm{IAR} = \frac{S_R^2}{(1/M)\sum_{i=1}^{M} s_i^2} \times 100 \tag{1}$$

or

$$\mathrm{IAR} = \frac{\text{variance of the area-averaged rainfall series}}{\text{mean variance of the rainfall series on all stations}} \times 100$$

where M indicates the number of stations averaged. The larger the value of IAR the greater the areal representation of the area-averaged rainfall series.

The IAR for the all-India summer monsoon rainfall series constructed by different methods varied between 7 per cent and 14 per cent. It was then considered whether data from all 306 rain-gauges were essential to update the all-India summer monsoon rainfall series. Using an objective approach, Sontakke *et al.* (1993) have identified 35 rain-gauges for updating the all-India summer monsoon rainfall series prepared by Mooley and Parthasarathy (1984) and 188 rain-gauges for updating the rainfall series for 29 selected meteorological subdivisions of India prepared by Parthasarathy *et al.* (1987) by averaging the district-area-weighted rainfall of 306 stations. The three main objectives of the present study are:

(i) to optimize a network of rain-gauges using different techniques for the spatially representative summer monsoon rainfall series of India prepared by averaging the normalized rainfall (actual divided by mean) of 306 rain-gauges;
(ii) to examine the validity of the optimization techniques in the light of some theoretical studies published elsewhere (Wigley *et al.*, 1984);
(iii) to select objectively a few rain-gauges for routine updating of the all-India summer monsoon rainfall series.

Through the large amounts of latent heat released, the summer monsoon influences the general circulation of the atmosphere. Climate variability is currently of great importance for atmospheric scientists. Besides being useful to Indian planners and decision makers, a quick updating of the summer monsoon rainfall series can be helpful to scientists in other parts of the globe needing near real-time data on the Indian monsoon.

DATA USED

Summer monsoon rainfall (June to September total) for the period 1871–1984 recorded at 306 rain-gauge sites (Figure 1) well spread over the contiguous and relatively flat areas of the country, excluding mountainous regions in the north and the islands in the Bay of Bengal and the Arabian Sea, were used in this study. Gaps in the data (which account for less than 2 per cent of the total data set) were filled by the ratio method, as suggested by Rainbird (1967). Details of this data set are given in Mooley and Parthasarathy (1984), and for this study the processed data set was obtained from B. Parthasarathy. For updating the all-India monsoon rainfall series, rainfall data of 35 selected stations for the period 1985–1992 were collected from the relevant records of the India Meteorological Department.

TECHNIQUES USED FOR THE OPTIMIZATION OF RAIN-GAUGES

The all-India summer monsoon rainfall series prepared by averaging non-exceedance gamma probability (Bradley *et al.*, 1987) of monsoon rainfall totals of 306 stations has previously been found to be most representative of areal rainfall over India, with an IAR of 14·3 per cent (Singh *et al.*, 1991). The second most representative (IAR = 13·2 per cent) all-India summer monsoon rainfall is that prepared by averaging a normalized (actual divided by mean) rainfall of the individual stations. The IAR for the monsoon rainfall

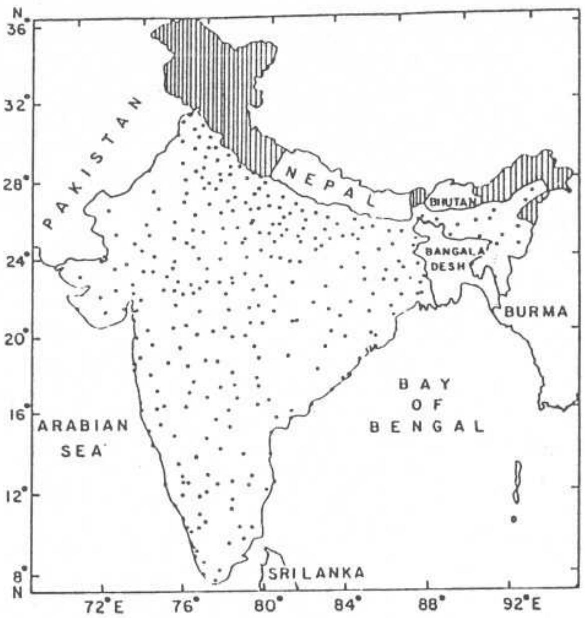

Figure 1. Location of the 306 rain-gauge stations used for preparing the representative summer monsoon rainfall series of India

series of Mooley and Parthasarathy is 9·45 per cent. For computational simplicity, and without considerable loss of information, we have optimized the network of 306 rain-gauges (Figure 1) for updating all-India summer monsoon rainfall series prepared by averaging normalized station rainfall totals. We have used three different approaches, each of which will be discussed in turn. These are:

(i) random selection;
(ii) fitting a multiple linear regression;
(iii) selecting objectively a subset of a few gauges with means showing the highest correlation coefficient with the all-India series.

Random selection

Wigley *et al.* (1984) have derived an approximate relationship for an average correlation coefficient between m stations mean series and M stations mean series ($R_{m,M}$) (m is a subset of M) as a function of m, M and mean correlation coefficient ($\bar{r}$) between all possible pairs of rainfall series in the set of M stations.

$$(\bar{R}_{m,M})^2 \approx \frac{m[1+(M-1)\bar{r}]}{M[1+(m-1)\bar{r}]} \tag{2}$$

The $\bar{r}$ for the present network of 306 (M) rain-gauges is 0·1239. Variations in $\bar{R}_{m,M}$ with the increase in m from 1 to 100 is presented in Figure 2(a). Equation (2) is derived based on the assumption of random distribution of rain-gauges. We validated equation (2) for the all-India summer monsoon rainfall series prepared using data from 306 sites. We randomly chose subsets of m stations ($m = 1, 2, 3, \ldots, 100$), calculated the m-station mean series and then calculated $R_{m,M}$. For each of the given m, 100 replications of size 114 (i.e. the entire length of the available data) are generated and $R_{m,M}$ calculated. The actual $R_{m,M}$ depends upon the particular subset of m series chosen out of the 306 stations, so besides the mean ($\bar{R}_{m,M}$) we also calculated its standard deviation as a measure of the scatter of individual realizations. The results are presented in Figure 2(b). The simulated $\bar{R}_{m,M}$ with the actual data and estimated $\bar{R}_{m,M}$ using equation (2) are, substantially, in close agreement. An inference of this result could be that the 306 rain-gauges over India are randomly distributed. For the

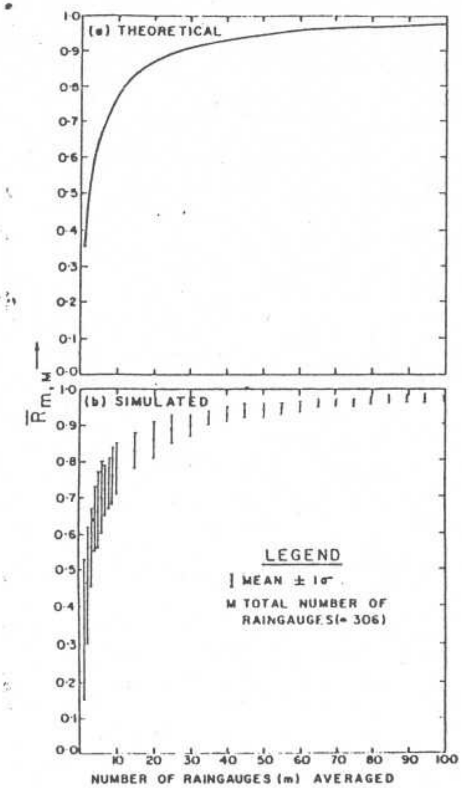

Figure 2. Increase in the correlation coefficient between mean series of m selected gauges and the all-India rainfall series ($\bar{R}_{m,M}$) with the increase in m. (a) Theoretical curve derived from equation (2), and (b) simulated curve with actual data

increase of m from 1 to 45, $\bar{R}_{m,M}$ increased sharply from 0·34 to 0·94. Thereafter, the increase in $\bar{R}_{m,M}$ with m is small. For $m=100$, $\bar{R}_{m,M}$ increased to only 0·97.

The main conclusion from this approach is that by averaging 45 randomly selected observations a series can be obtained having a correlation of about 0·94 with a representative all-India rainfall series; with 100 observations this correlation is about 0·97.

Fitting a multiple linear regression

In our second approach stations have been selected by fitting a multiple linear regression between the all-India monsoon rainfall and the normalized monsoon rainfall (actual divided by mean) of 306 stations. A forward selection scheme has been followed for the identification of independent parameters (rain-gauges) in the regression. Thirty-four gauges entered the regression by making a significant contribution to the explained variance of the all-India monsoon rainfall series (1871–1984). The F-test has been applied for statistical significance of the partial variance explained by the individual independent parameters. The test is defined as follows:

$$F=\frac{R_{p+1}^2-R_p^2}{1-R_{p+1}^2}\times(N-p-1) \tag{3}$$

where R_{p+1}^2 and R_p^2 are the coefficients of determination of the regression with $p+1$ and p parameters, respectively, and N number of observations ($=114$). The sequence in which 34 stations entered in the regression is given in Table 1, along with the multiple correlation coefficient. The location of these 34 rain-gauge stations is given in Figure 3(a). They are well spread over the region. The estimated multiple linear

Table I. The sequence of 34 rain-gauge stations selected by multiple regression and by an objective technique for the all-India summer monsoon rainfall series along with the correlation coefficient (CC) achieved at each step

Station number	Station selected by multiple linear regression (abbreviated name)		CC	Station selected by objective technique (abbreviated name)		CC
1	Ajmer	(AJM)	0·6340	Ajmer	(AJM)	0·6340
2	Bareilly	(BRL)	0·7715	Bareilly	(BRL)	0·7715
3	Asifabad	(ASF)	0·8421	Asifabad	(ASF)	0·8418
4	Hoshiarpur	(HSH)	0·8791	Dhar	(DHA)	0·8748
5	Agra	(AGR)	0·9042	Etawah	(ETW)	0·8972
6	Radhanpur	(RDP)	0·9279	Bidar	(BDR)	0·9193
7	Alibag	(ABG)	0·9378	Hoshiarpur	(HSH)	0·9315
8	Basti	(BST)	0·9497	Mandla	(MND)	0·9400
9	Dhar	(DHA)	0·9575	Kolhapur	(KLP)	0·9470
10	Hassan	(HSN)	0·9619	Dharampuri	(DRP)	0·9522
11	Cuddalore	(CDL)	0·9660	Bhopal	(BHP)	0·9589
12	Shivpuri	(SVP)	0·9701	Ahwa	(AHW)	0·9631
13	Gorakhpur	(GRK)	0·9731	Bhind	(BHD)	0·9684
14	Patiala	(PTL)	0·9762	Darjeeling	(DJG)	0·9709
15	Ongole	(ONG)	0·9786	Rajkot	(RJK)	0·9724
16	Balaghat	(BLG)	0·9814	Rai Bareilly	(RBL)	0·9769
17	Rai Bareilly	(RBL)	0·9835	Sikar	(SKR)	0·9782
18	Nagarcoil	(NRC)	0·9855	Nagpur	(NGP)	0·9803
19	Broach	(BRH)	0·9874	Udaipur	(UDP)	0·9816
20	Jodhpur	(JDP)	0·9888	Deoria	(DOR)	0·9831
21	Sibsagar	(SBG)	0·9896	Cuttack	(CTK)	0·9845
22	Akola	(AKL)	0·9908	Betul	(BTL)	0·9857
23	Gaya	(GYA)	0·9918	Fort Cochin	(CHN)	0·9868
24	Mathura	(MTR)	0·9926	Sohagpur	(SHG)	0·9875
25	Karimnagar	(KRM)	0·9931	Nalgonda	(NGD)	0·9883
26	Jatusana	(JTS)	0·9936	Gwalior	(GWL)	0·9894
27	Thana	(TNA)	0·9941	Bangalore	(BNG)	0·9899
28	Mainpuri	(MPR)	0·9945	Bhawanipatna	(BNP)	0·9906
29	Hyderabad	(HYD)	0·9949	Ranchi	(RCH)	0·9907
30	Ranchi	(RCH)	0·9952	Kakinada	(KND)	0·9911
31	Bellary	(BLY)	0·9957	Jind	(JND)	0·9920
32	Neemuch	(NMC)	0·9962	Aurangabad	(AGD)	0·9925
33	Guwahati	(GHT)	0·9965	Purulia	(PRL)	0·9927
34	Shahjahanpur	(SHJ)	0·9969	Chingelput	(CGP)	0·9931

regression is as follows:

$$\begin{aligned} Y_{AI} = {} & 0{\cdot}0968 + 0{\cdot}013X_{AJM} + 0{\cdot}0671X_{BRL} + 0{\cdot}0165X_{ASF} + 0{\cdot}0373X_{HSH} + 0{\cdot}0194X_{AGR} + 0{\cdot}0343X_{RDP} \\ & + 0{\cdot}0436X_{ABG} + 0{\cdot}0186X_{BST} + 0{\cdot}0232X_{DHA} + 0{\cdot}0355X_{HSN} + 0{\cdot}0274X_{CDL} + 0{\cdot}0202X_{SVP} \\ & + 0{\cdot}0508X_{GRK} + 0{\cdot}0279X_{PTL} + 0{\cdot}0302X_{ONG} + 0{\cdot}0507X_{BLG} + 0{\cdot}0197X_{RBL} + 0{\cdot}0136X_{NRC} \\ & + 0{\cdot}0278X_{BRH} + 0{\cdot}0216X_{JDP} + 0{\cdot}0375X_{SBG} + 0{\cdot}0354X_{AKL} + 0{\cdot}0286X_{GYA} + 0{\cdot}0151X_{MTR} \\ & + 0{\cdot}0241X_{KRM} + 0{\cdot}0094X_{JTS} + 0{\cdot}0336X_{TNA} + 0{\cdot}0228X_{MPR} + 0{\cdot}0082X_{HYD} + 0{\cdot}0495X_{RCH} \\ & + 0{\cdot}0233X_{BLY} + 0{\cdot}0219X_{NMC} + 0{\cdot}0203X_{GHT} - 0{\cdot}248X_{SHJ} \end{aligned} \qquad (4)$$

where Y_{AI} is the all-India summer monsoon rainfall and X_{stn} is the normalized rainfall of the selected stations (abbreviated name of the stations is given in Table I).

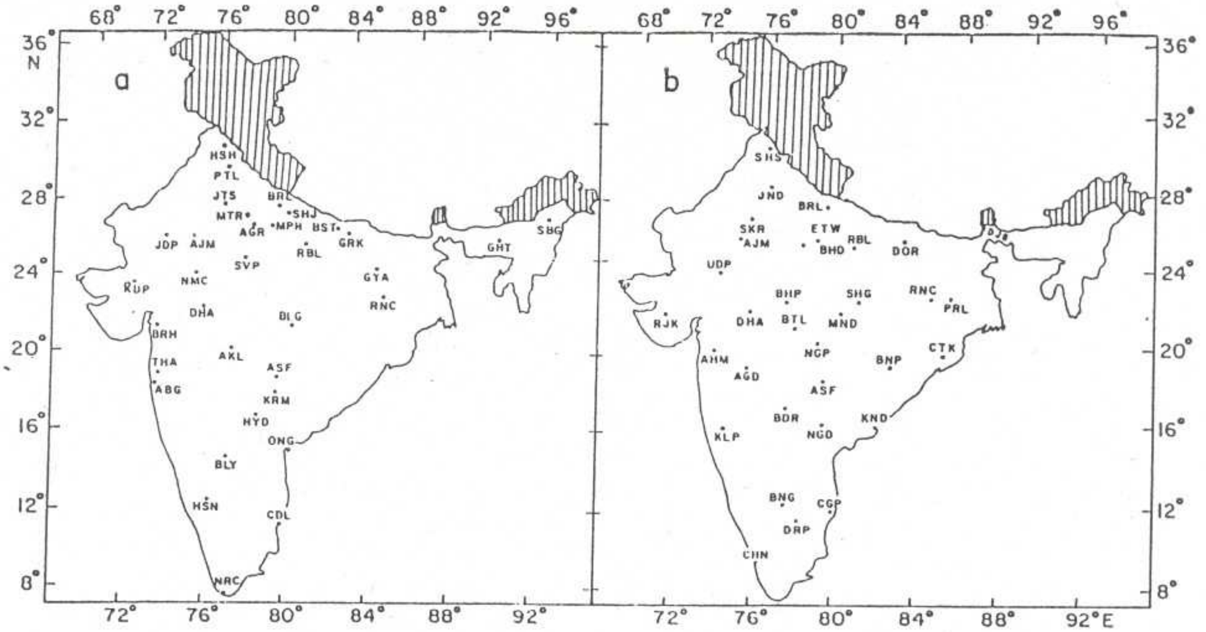

Figure 3. Location of the 34 gauges selected fitting a multiple linear regression (a), and the objectively selected 34 gauges with means showing the highest CC (b).

Objectively selecting a few gauges with means showing the highest CC

This technique has previously been used by Sontakke *et al.* (1993). The technique essentially requires the computation of a product-moment correlation coefficient between the mean series of subsets of gauges of various sizes and the all-India mean series. Gauges in the subset of a specified size with means showing the highest CC are selected one by one in a manner similar to the forward selection of independent parameters in the multiple linear regression model. The process of selection starts by correlating the all-India rainfall series with a normalized (actual divided by mean) rainfall series of the individual gauges in the network of 306 rain-gauges. The entire length of records (1871–1984) is involved in the computation of CCs. The first gauge is identified as the one which showed the highest correlation coefficient (CC) with the all-India rainfall series. The mean of the first selected gauge and one additional gauge, successively selected from the remaining 305 gauges, is then correlated with the all-India series. The pair of gauges showing the highest CC is identified. Similarly, the mean of the two selected gauges with one additional gauge, successively selected from the remaining 304 gauges, is correlated with the all-India series and a group of three gauges showing the highest CC is identified. The process continues to select a group of 4, 5, 6, ..., etc., gauges as long as CC increased due to the inclusion of an additional rain-gauge. Starting with Ajmer, with the highest CC (of 0·637), 34 gauges have been selected, giving a CC of 0·9931. These gauges are given in Table I in the sequence they were selected, and their locations are marked on Figure 3(b). In this selection, the rain-gauge stations are also found to be well spread over the region. The estimated regression is as follows:

$$Y_{AI} = 0{\cdot}0724 + 0{\cdot}9276\bar{X}_{34} \quad (5)$$

where $\bar{X}_{34}$ is the mean normalized rainfall of the 34 selected stations.

From Table I it can be seen that the multiple regression approach achieves a higher correlation compared with the objective selection technique for the same number of stations. This may be due to the multiple regression taking care of multicollinearity. It is mere coincidence that by methods (ii) and (iii) an equal number (34) of rain-gauges entered into the selection. Selection of stations following these two approaches were also carried out for 50-year and 100-year sliding periods in order to examine the stability of the selection processes. Rain-gauges entering the selection varied between 15 and 40, with the multiple regression method showing higher CCs (exceeding 0·95 by both techniques) and a lesser number of gauges compared with the mean of an objectively selected subset of gauges. On each occasion the selected rain-gauge stations were well spread over the region. It may be interesting to note that for multiple CCs of comparable magnitudes the interceptions and regression coefficients of equations (4) and (5) are comparable; that is, the interception of equation (4) is comparable to the interception of equation (5), and the regression coefficient of equation (4) is comparable the sum total of regression coefficients of equation (5). This similarity for the interceptions and regression coefficients is also seen in the selection based on 50-year and 100-year sliding periods. The use of normalized data is one of the reasons of this closeness.

SELECTING GAUGES FOR ROUTINE UPDATING OF THE ALL-INDIA SUMMER MONSOON RAINFALL SERIES

As stated previously, one of the main objectives of this study is to identify a minimum number of gauges for routine updating of the representative all-India summer monsoon rainfall series. Collection of data from 306

Table II. The geographical coordinates and the mean summer monsoon rainfall of 35 IMD rain-gauges selected for routine updating of the all-India summer monsoon rainfall series. The stations are given in the sequence they entered in the selection along with the correlation coefficient

Station number	Station (abbreviated name)		Geographical coordinates °N	°E	Correlation coefficient (CC)	Mean monsoon rainfall (mm)
1	Ajmer	(AJM)	26 27	74 37	0·6264	493·8
2	Bareilly	(BRL)	28 22	79 24	0·7672	958·1
3	Nanded	(NND)	19 05	77 20	0·8318	821·4
4	Bhopal	(BHP)	23 16	77 25	0·8682	1073·4
5	Baroda	(BRD)	22 47	73 18	0·8887	873·5
6	Nellore	(NRL)	14 27	79 59	0·9046	299·0
7	Mandla	(MND)	22 35	80 22	0·9202	1275·0
8	Patiala	(PTL)	30 20	76 28	0·9332	570·0
9	Allahabad	(ALB)	25 27	81 44	0·9408	883·3
10	Jaipur	(JPR)	26 49	75 48	0·9447	568·1
11	Bangalore	(BNG)	12 58	77 35	0·9540	493·5
12	Bhawanipatna	(BNP)	19 55	83 10	0·9596	1153·9
13	Hanamkonda	(HNK)	18 01	79 34	0·9617	729·1
14	Suri	(SRI)	23 53	87 32	0·9646	1126·7
15	Shimoga	(SMG)	13 56	75 38	0·9668	556·3
16	Mainpuri	(MPR)	27 14	79 03	0·9696	678·8
17	Udaipur	(UDP)	24 35	73 42	0·9708	593·1
18	Bombay	(BMB)	18 54	72 49	0·9720	1836·7
19	Gorakhpur	(GRK)	26 45	83 25	0·9740	1082·1
20	Akola	(AKL)	20 42	77 02	0·9752	688·3
21	Daltonganj	(DTJ)	24 03	84 04	0·9762	1008·8
22	Madras	(MDS)	13 04	80 15	0·9771	394·0
23	Hissar	(HSR)	29 09	75 44	0·9773	344·1
24	Thiruvananthapuram	(TRV)	08 29	76 57	0·9783	837·4
25	Ballia	(BLA)	25 45	84 10	0·9798	933·2
26	Mahabubnagar	(MBN)	16 44	77 59	0·9809	695·9
27	Ambikapur	(ABP)	23 07	83 12	0·9824	1370·9
28	Jalore	(JLR)	25 21	72 37	0·9832	349·6
29	Calcutta	(CAL)	22 32	88 20	0·9836	1204·9
30	Puri	(PRI)	19 48	85 49	0·9841	978·0
31	Guwahati	(GWT)	26 06	91 35	0·9843	1052·1
32	Nagpur	(NGP)	21 06	79 03	0·9848	1042·1
33	Delhi	(DLH)	28 35	77 12	0·9851	600·9
34	Bijapur	(BJP)	16 49	75 43	0·9861	383·0
35	Jabalpur	(JBP)	23 10	79 57	0·9866	1262·3

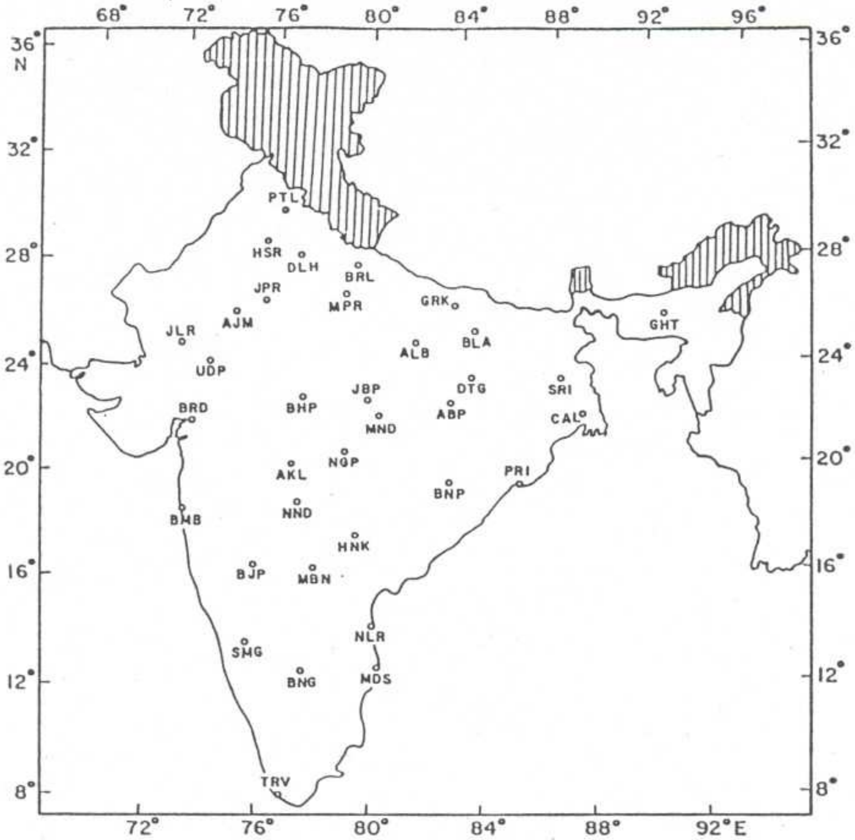

Figure 4. Location of the 35 selected IMD rain-gauges known for their timely availability of data for routine updating of the representative all-India summer monsoon rainfall series

stations is a difficult task. Obtaining data from rain-gauge stations other than those maintained by the India Meteorological Department (IMD) is particularly time-consuming. In the present network of 306 rain-gauges there are 75 non-IMD stations. We have, therefore, first chosen 64 IMD rain-gauge stations well spread over the country and known for their timely availability of data, and then selected a subset of gauges by applying method (iii). Thirty-five gauges entered the selection, with a CC of 0·9866. These gauges are given in Table II along with the increase in the CC at each step of the selection process. Their locations are marked in Figure 4. The least-squares estimated regression equation is as follows:

$$Y_{\mathrm{AI}} = 0{\cdot}0767 + 0{\cdot}9235\bar{X}_{35} \tag{6}$$

To cross-check the validity of the technique of 'selecting a group of a few gauges with their means showing the highest CC', we calculated the correlation coefficient between the mean series of m selected stations (Table II) and that of all the stations (M), ($R_{m,M}$), using the following relation given by Wigley *et al.* (1984).

$$R_{m,M} = \frac{1}{m\,s(m)} \sum_{i=1}^{m} s_i\, r_{i,M} \tag{7}$$

where m is the number of selected stations ($=35$), $s(m)$ the standard deviation of the m-station mean series ($=0{\cdot}1282$), s_i the standard deviation of the selected station i, and $r_{i,M}$ the correlation coefficient between selected station i and the all-stations mean series. Substituting m, $s(m)$, s_i and $r_{i,M}$ for stations given in Table II, $R_{m,M}$ is calculated as 0·9866 which is exactly that obtained by correlating the two series directly. Hence the simple approach of 'selecting a few gauges with their means showing the highest CC' for

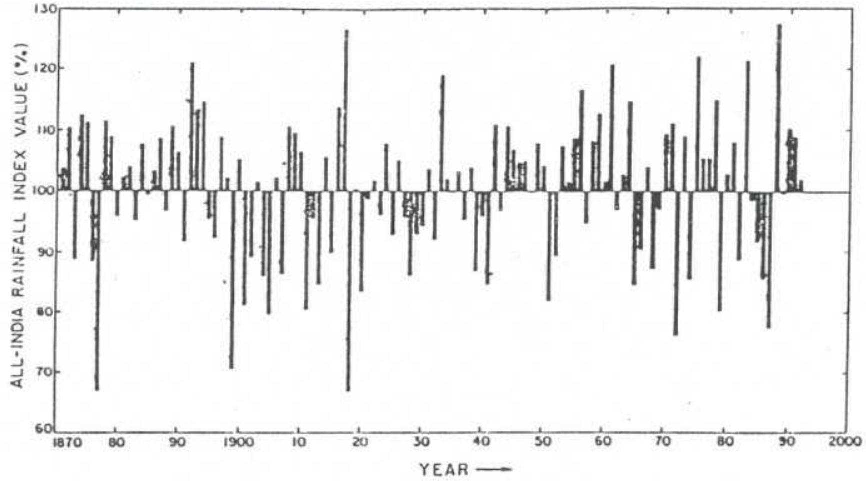

Figure 5. A reconstructed all-India summer monsoon rainfall series for the period 1871–1992

Table III. Constructed all-India summer monsoon rainfall values for the period 1871–1992 as a percentage of the 1871–1984 period mean

Year	0	1	2	3	4	5	6	7	8	9
1870		103·5	110·3	88·9	112·5	111·3	88·5	66·7	111·4	108·8
1880	95·8	102·2	103·8	94·9	107·7	99·2	103·1	108·6	96·7	110·6
1890	106·2	91·7	121·2	113·2	114·5	95·4	92·3	108·7	102·0	70·5
1900	105·1	81·0	89·0	101·3	85·8	79·5	102·1	86·3	110·6	109·6
1910	106·3	80·3	95·3	84·5	105·3	89·8	113·7	126·7	66·5	100·1
1920	83·1	98·7	101·7	96·0	107·8	92·7	105·0	95·6	85·8	92·7
1930	94·1	103·6	91·9	119·2	102·1	100·1	103·2	95·1	103·8	86·7
1940	95·8	84·3	110·9	96·7	110·9	107·0	104·7	105·0	99·9	107·9
1950	104·2	81·6	89·1	107·5	101·5	108·7	116·7	94·5	108·2	112·6
1960	101·5	120·6	96·9	102·7	114·9	84·3	90·1	103·9	87·0	96·9
1970	109·4	111·1	76·1	109·1	85·4	122·1	105·3	105·2	114·8	80·4
1980	102·7	107·9	88·6	121·4	98·4	91·6	85·5	77·4	127·5	99·6
1990	110·4	108·8	101·8							

optimizing the network of rain-gauges to reproduce a given area-averaged rainfall series is theoretically consistent.

The reconstructed all-India summer monsoon rainfall series using equation (6) for the period 1871–1992 is shown in Figure 5. The estimated values as a percentage of the 1871–1984 period mean are also given in Table III for the benefit of those interested in the Indian monsoon rainfall data.

CONCLUSIONS

The main conclusions of the present attempt to optimize the network of rain-gauges necessary to reproduce the representative all-India summer monsoon rainfall series are as follows:

(i) By averaging a randomly selected subset of 45 rain-gauges, a CC of the order of 0·94 with the all-India rainfall series can be achieved, and by averaging 100 gauges a CC of 0·97.

(ii) By fitting a multiple linear regression, a set of 34 rain-gauges have been found adequate to reproduce the all-India summer monsoon rainfall series. The estimated rainfall series from the multiple regression showed a CC of 0·9969 with the representative all-India rainfall series.

(iii) By applying a simple objective technique a maximum of 34 gauges have been identified whose mean series has shown a CC of 0·9931 with the all-India series. Investigation reveals that the result of this selection is consistent with theoretical expectations.

(iv) For routine updating of the all-India monsoon rainfall series, however, 35 rain-gauges have been selected from a group of 64 randomly chosen IMD rain-gauge stations known for their timely availability of data and whose mean series showed a CC of 0·9866 with the all-India series. Experience shows that the monsoon rainfall data for these stations are available by the middle of November of the same year and the all-India monsoon rainfall series can therefore be updated easily.

ACKNOWLEDGEMENTS

I wish to acknowledge gratefully the Director, Indian Institute of Tropical Meteorology, Pune, for making the necessary facilities available; to the Additional Director General of Meteorology (Research), India Meteorological Department, Pune, for the rainfall data; to Dr G. B. Pant, Head of the Climatology and Hydrometeorology Division for his keen interest and encouragement, and to Professor Sulochana Gadgil of the Indian Institute of Science, Bangalore, Professor C. A. S. Naidu of the National Institute of Bank Management, Pune and Dr S. V. Singh and M. K. Tandon of the Indian Institute of Tropical Meteorology, Pune, for fruitful discussion and valuable suggestions. I am indebted to one of the anonymous referees for critically reviewing and editing the manuscript.

REFERENCES

Blanford, H. F. 1886, 'Rainfall of India', *Memoirs India Meteorological Department*, **3**, 658 pp.

Bradley, R. S., Diaz, H. F., Eischeid, J. K., Jones, P. D., Kelly, P. M. and Goodess, C. M. 1987. 'Precipitation fluctuations over northern hemisphere land areas since the mid-19th century', *Science*, **237**, 171–175.

Mooley, D. A. and Parthasarathy, B. 1979. 'Poisson distribution and years of bad monsoon over India', *Arch. Meteorol. Geophys. Bioklimatol.*, **B27**, 381–388.

Mooley, D. A. and Parthasarathy, B. 1984. 'Fluctuations in all-India summer monsoon rainfall during 1871–1978', *Climatic Change*, **6**, 287–301.

Parthasarathy, B. and Dhar, O. N. 1976. 'A study of trends and periodicities in the seasonal and annual rainfall of India', *Ind. J. Meteorol. Hyd. Geophys.*, **27**, 23–28.

Parthasarathy, B. and Mooley, D. A. 1978. 'Some features of a long homogeneous series of Indian summer monsoon rainfall', *Mon. Wea. Rev.*, **106**, 771–781.

Parthasarathy, B., Sontakke, N. A., Munot, A. A. and Kothawale, D. R. 1987. 'Droughts/floods in the summer monsoon season over different meteorological subdivisions of India for the period 1871–1984', *J. Climatol.*, **7**, 57–70.

Rainbird, A. F. 1967. *Methods of Estimating Areal Average Precipitation*, WMO/IHD Report No. 3, World Meteorological Organization, Geneva, 42 pp.

Singh, N., Pant, G. B. and Mulye, S. S. 1991. 'A statistical package for constructing representative area-averaged rainfall series and its updating using selected stations data', in *Proceedings National Seminar on the Use of Computers in Hydrology and Water Resources*, 16–18 December 1991, Vol. II, pp. 122–134, Indian Institute of Technology, Delhi.

Sontakke, N. A., Pant, G. B. and Singh, N. 1992. *Construction and Analysis of all-India Summer Monsoon Rainfall Series for the Longest Instrumental Period: 1813–1991*, Reseach Report No. RR-053, Contributions from Indian Institute of Tropical Meteorology, Pune-8, 40 pp.

Sontakke, N. A., Singh, N. and Pant, G. B. 1993. 'Optimization of the raingauges for a representative all-India and subdivisional rainfall series', *Theor. Appl. Climatol.*, (in press).

Walker, G. T. 1910. 'On the meteorological evidence for supposed changes of climate in India', *Indian Meteorological Memoirs*, **21** (Part I), 1–21.

Walker, G. T. 1914, 'A further study of relationship with Indian monsoon rainfall', *Indian Meteorological Memoirs*, **21** (Part VII), 1–12.

Walker, G. T. 1922. 'Correlation in seasonal variation of weather: VII the local distribution of seasonal rainfall', *Indian Meteorological Memoirs*, **21** (Part VIII), 23–39.

Wigley, T. M. L., Briffa, K. R. and Jones, P. D. 1984, 'On the average value of correlated time series, with application in dendroclimatology and hydrometeorology', *J. Clim. Appl. Meteorol.*, **23**, 201–213.