Mausam, (1990), **41**, 4, 523-530

551.501.777:551.577

Space correlation structure of rainfall over India

D.S. UPADHYAY, SURINDER KAUR, M.S. MISRA and M.K. GUPTA

Meteorological Office, New Delhi

(*Received* 18 *August* 1989)

सार — इस शोधपत्र में, 21 मौसम वैज्ञानिक एक समान क्षेत्रों जिनमें लगभग सम्पूर्ण देश आ जाता है, के लिए वर्षा क्षेत्र की सहसम्बंध संरचना का पता लगाया गया है। इस उद्देश्य के लिए लगभग 2000 स्टेशनों के 70 वर्षों (सन् 1901–1970) के वर्षा आंकड़ों का उपयोग किया गया है।

ABSTRACT. In this paper, the correlation structure of rainfall field has been worked out for 21 meteorological homogeneous regions which cover almost entire country. For this purpose 70 years (1901-1970) rainfall data for about 2000 stations have been utilised.

1. Introduction

The rainfall varies in space (x, y, z) and in time (t). It is possible that a significant correlation exists between two sets of rainfall observations. Such correlation may be of two types :

(*i*) $r(t)$ — correlation between two sets of data of one station recorded in different period of time.

(*ii*) $r(s)$ — correlation between two data sets recorded at different stations during same time span.

$r(t)$ refers to the persistence effect. To eliminate the affect of persistence on estimation of areal rainfall its standard error is reduced by a factor $[1-r(t)]^{1/2}$. However the analysis of this type of correlation is not under the scope of present work. In succeeding paras the spatial structure of correlation field has been analysed for various meteorologically homogeneous regions. The entire India has been divided into 16 such regions and the spatial correlation maps have been produced for these.

Some of the important hydrologists who applied the concept of correlation between gauges for solving operational problems in hydrology are :

(1) Hershfield (1965) — to design proper spacing between gauges,

(2) Kagan (1966, 1972) — for computing errors in interpolation,

(3) Rodriguez-Iturbe and Mejia (1974) – for separation of spatial and temporal factors of the error in estimating areal rainfall from point observations,

(4) Ramanathan *et al.* (1981) — applied work for determining optimum network density using the correlation structure of rainfall.

The information on correlation between rainfall series located at two stations and their variation with the distance between the stations has large practical utility in the following areas :

(*i*) Estimation of point to areal rainfall relationships,

(*ii*) Network design,

(*iii*) Transfer of rainfall information from one station to the other including interpolation of the missing data in spatial series, and

(*iv*) Regional rainfall analysis for specific purpose like drainage and designing small hydraulic structures.

In many statistical studies relating rainfall variable x_i recorded at various points (i) assuming that the series x_i are statistically independent. The significant correlation existing between x_i's is bound to revise the results of such studies. In this context also, the importance of present work is apparent.

2. Method of computation of space correlation

Suppose 'A' is a meteorologically homogeneous area. Let $x_1, x_2 \ldots\ldots x_T$ and $y_1, y_2 \ldots\ldots y_T$ be the annual precipitation series at the two stations in this area having correlation coefficient r.

If there are n existing stations in this area, there will be ${}^{n}c_2$ correlation coefficients (r) and the same number of distances (s) between the stations. It is possible to establish a functional relationship $r=f(s)$ between the two variables. This relationship describes the correlation structure of rainfall over an area. It is generally expected that the correlation decreases as the distance increases. Rodriguez – Iturbe and Mejia (1974) have shown that the functional relation $f(s)$ can be any one of the following :

$$\left.\begin{array}{lll} (i) & r = a\,e^{-bs} & \text{(exponential)} \\ (ii) & r = a - bs & \text{(linear)} \\ (iii) & r = bs\,k_1\,(bs) & \text{(modified Bessel)} \end{array}\right\} \quad (1)$$

where a and b are constants.

2.1. *The probability density function of 's'*

The distance 's' may be regarded as a random variable as it may assume different values between O and D (maximum possible distance between the two points in the area of interest) with frequency function $\phi(s)$. Form of $\phi(s)$ depends on the shape of the area/catchment. Erikson (1972) suggested the following function $\phi(s)$ for a circular area of radius R as :

$$\phi(s) = 2s/R^2,\ 0 \leqslant s \leqslant R \quad (2)$$

2.2. *Mean spatial correlation*

The mean correlation over a given area A is :

$$E\,(r/A) = \bar{r} = \int r\,(s).\ \phi(s)ds \quad (3)$$

If $r(s)$ follows an exponential law given by Eqn. (1), then

$$\bar{r} = \int_0^R a\,e^{-bs}\,\frac{2s}{R^2}\,ds$$

$$= \frac{2a}{R^2\,b^2}\left[1 - (1 + Rb)e^{-bR}\right] \quad (4)$$

If $r(s)$ follows a linear law of variation, then :

$$\bar{r} = \int_o^R (a - b\,s)\,.\,\frac{2s}{R^2}\,ds = a - \frac{2}{3}\,Rb \quad (5)$$

It may be noted that Eqns. (4) & (5) are valid only for circular regions. For other shapes, p.d.f.$\phi(s)$ will acquire different functional form. Eagleson (1967) has derived these forms for some other geometrical shapes like squares and rectangles.

$\bar{r}$ depends on (*i*) parameters a and b which describe the characteristic of the correlation structure in that area and (*ii*) the longest distance possible in that area (radius, in case of circular area) or area of the region A.

It is also apparent from the relations (*i*) to (*iii*) of Eqn. (1) that theoretically a tends to 1 and b has a dimension of L^{-1}. Hence Ab^2 is a dimensionless quantity expressing combined effect of area and correlation characteristic. It could be interesting to express $\bar{r}$ as a function of Ab^2.

3. Data used

Long term (1901-1970) rainfall data for about 2000 stations distributed over the plain of India have been used in this study. For computing correlations between the pairs of rainfall series, effort has been made to omit minimum quantity of data due to missing observations. The distances between two stations have also been regarded as random variable which are calculated from the coordinates (Lat., Long.) of stations.

TABLE 1

r-s frequency table

s (km)	r: 0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1.0	$\bar{r}_i$
Circle 1						
0-40	0	0	1	5	0	0.67
40-80	0	1	8	7	1	0.59
80-120	0	1	7	3	0	0.54
120-160	0	1	1	0	0	0.40
Circle 2						
0-40	0	0	3	10	2	0.60
40-80	0	3	16	25	1	0.61
80-120	0	4	30	12	1	0.54
120-160	0	5	9	6	0	0.51
160-200	0	2	6	1	0	0.48
Circle 3						
0-40	0	0	5	11	2	0.67
40-80	0	3	31	29	2	0.59
80-120	0	9	51	22	1	0.54
120-160	0	12	39	12	0	0.50
160-200	0	14	24	7	0	0.47
200-240	0	9	8	0	0	0.39
240-280	0	4	1	0	0	0.34
280-320	0	1	2	0	0	0.43
320-360	0	1	0	0	0	0.30
Circle 4						
0-40	1	0	8	18	4	0.65
40-80	3	11	47	46	2	0.55
80-120	4	22	104	36	1	0.50
120-160	6	40	97	19	0	0.46
160-200	3	41	83	11	0	0.43
200-240	5	47	45	0	0	0.35
240-280	3	31	22	0	0	0.36
280-320	5	7	8	0	0	0.30
320-360	0	5	0	0	0	0.30

TABLE 2

Parameters of correlation structure

Circle No.	a	b (km^{-1})	Radius R (km)	Ab^2	$\bar{r}$
1	0.80	0.00313	69	0.15	0.69
2	0.76	0.00263	98	0.21	0.64
3	0.76	0.00222	130	0.26	0.64
4	0.72	0.00227	162	0.43	0.57

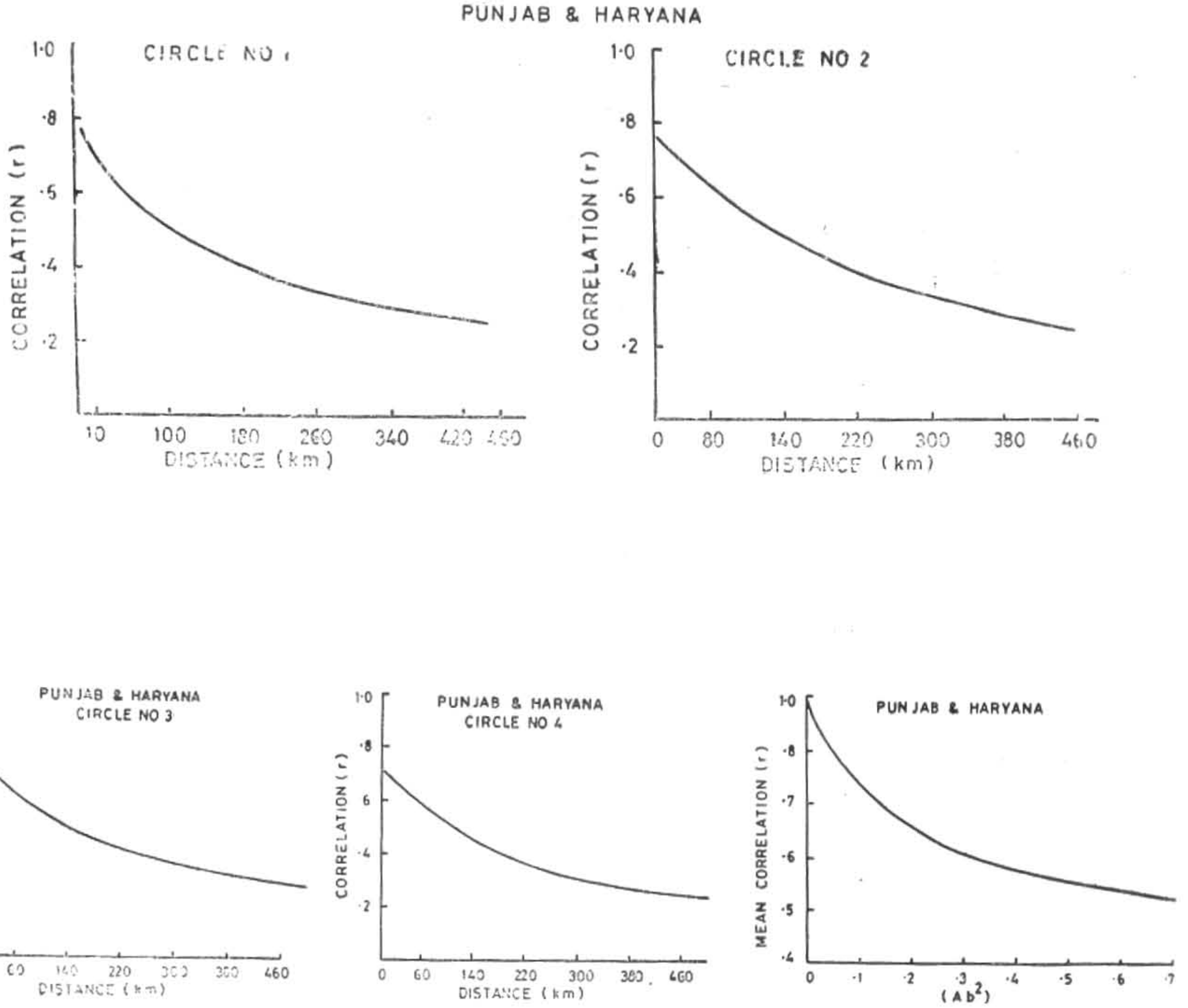

Fig. 1. Space correlation structure (Punjab & Haryana)

4. A practical illustration for the computation of correlation structure

The study outlined in preceding paras have been carried out for different homogeneous regions covering almost the entire country. The regions have been demarcated by circular lines so that the density function given by Eqn. (2) is applicable.

Under each circle there will be different number of stations which will provide varying pattern of r–s relationship. In other words the mean correlation ($\bar{r}$) of the rainfall field can be treated as a function of area (A) of the circle and the parameter (b) of the functional relationship between r and s.

In the succeeding paras the procedures are illustrated in respect of one homogeneous region, namely, Punjab and Haryana. The major parts of Punjab and Haryana have been classified by 4 circles covering the areas about 15,000, 30,000, 53,000 and 82,000 km^2. The number of stations falling under these areas are 9, 17, 25 and 41 respectively.

For circle 1, $^9C_2$ (=36) pairs of stations provide 36 values of correlations (r) and distances (s) as given in the bivariate frequency Table 1.

It is clearly indicated that (r) decreases with distance (s) and follows an exponential pattern [Eqn. (1)] with parameters a and b given in the Table 2.

A graph showing the variation of $\bar{r}$ with a dimensionless quantity (Ab^2), A being the area of the circle is given at Fig. 1.

It determines mean correlation structure for a given area if the parameter b is known.

5. Regional correlation structure of rainfall over India

The aforesaid study has been conducted for the entire country dividing it in 21 climatically homogeneous regions. The important results and inferences are summarised in the Table 3.

The graphs showing $\bar{r} - Ab^2$ relationship for different regions are given at Figs. 2-21.

TABLE 3

Parameters of correlation structure

Circle No.	No. of stations (n)	R (km)	a	b(10^{-4} km^{-1})	$\bar{r}$	Ab^2
(1) North Andhra Pradesh (139800 km²)						
1	33	110	0.59	47	0.46	0.48
2	54	143	0.63	45	0.42	1.33
3	68	171	0.68	47	0.40	2.08
4	99	193	0.66	43	0.38	2.21
5	127	211	0.65	43	0.35	7.64
(2) South Andhra Pradesh (89700 km²)						
1	11	44	0.77	53	0.67	0.17
2	27	86	0.70	83	0.44	1.60
3	50	120	0.68	83	0.35	3.11
4	71	145	0.73	71	0.37	3.33
5	86	169	0.70	67	0.34	4.02
(3) Bihar (205800 km²)						
1	8	51	0.80	156	0.48	1.99
2	29	90	0.75	100	0.42	2.54
3	64	134	0.75	81	0.38	3.70
4	94	182	0.70	68	0.32	4.75
5	128	213	0.75	74	0.28	7.80
6	159	238	0.82	30	0.51	1.63
7	176	256	0.82	29	0.52	1.78
(4) Eastern India (23250 km²)						
1	5	36	0.80	178	0.53	1.29
2	23	52	0.80	182	0.44	2.79
3	37	60	0.75	147	0.39	3.23
4	60	86	0.78	128	0.38	3.82
(5) Gujarat (157700 km²)						
1	8	64	0.90	48	0.74	0.29
2	[illegible]	111	0.92	38	0.69	0.57
3	44	151	0.93	28	0.71	0.56
4	63	189	0.88	26	0.64	0.78
5	76	224	0.88	25	0.61	0.98
(6) Himachal Pradesh (80400 km²)						
1	6	39	0.75	238	0.42	2.46
2	12	70	0.72	160	0.35	3.94
3	23	100	0.70	56	0.49	0.97
4	36	130	0.73	114	0.29	6.85
5	43	160	0.68	100	0.25	8.04
(7) Jammu & Kashmir (37300 km²)						
1	5	44	0.60	10	0.43	0.742
2	13	77	0.58	66	0.42	0.820
3	19	92	0.60	76	0.38	1.520
4	29	109	0.60	80	0.34	2.380
(8) North Karnataka (85550 km²)						
1	14	66	0.72	63	0.55	0.53
2	31	107	0.64	67	0.40	1.60
3	56	139	0.64	59	0.38	2.10
4	69	165	0.62	55	0.34	2.64
(9) South Karnataka (101750 km²)						
1	20	48	0.59	100	0.43	0.72
2	51	85	0.60	70	0.39	1.11
3	56	122	0.57	60	0.37	1.68
4	114	147	0.60	60	0.35	2.44
5	140	180	0.56	59	0.28	3.52
(10) Kerala* (38700 km²)						
1	17	53	0.75	37	0.66	0.12
2	30	71	0.70	47	0.56	0.36
3	49	89	0.70	31	0.58	0.25
4	73	111	0.70	26	0.58	0.27
(11) West Madhya Pradesh (159100 km²)						
1	27	115	0.90	77	0.51	2.46
2	34	150	0.85	62	0.47	2.72
3	52	188	0.80	53	0.42	3.12
4	67	225	0.75	50	0.37	3.98
(12) East Madhya Pradesh (225750 km²)						
1	15	110	0.75	36	0.58	0.48
2	26	137	0.75	30	0.57	0.54
3	40	170	0.72	34	0.50	1.01
4	53	198	0.72	33	0.48	1.06
5	70	234	0.74	28	0.47	1.33
6	93	268	0.70	25	0.45	1.41
(13) Maharashtra (174900 km²)						
1	14	65	0.63	63	0.48	0.53
2	26	101	0.70	76	0.43	1.84
3	46	132	0.59	29	0.46	0.45
4	66	157	0.60	27	0.44	0.60
5	80	177	0.65	25	0.48	0.61
6	101	194	0.64	29	0.45	1.02
7	120	212	0.63	28	0.43	1.09
8	142	226	0.64	26	0.43	1.11
9	153	236	0.64	18	0.42	0.56
(14) Orissa (88650 km²)						
1	13	68	0.92	53	0.73	0.40
2	30	101	0.92	45	0.68	0.64
3	46	130	0.90	41	0.63	0.91
4	59	150	0.90	36	0.63	0.90
5	69	168	0.90	34	0.62	1.02
(15) Rajasthan (311850 km²)						
1	19	128	0.90	23	0.59	0.265
2	33	176	0.90	20	0.55	0.390
3	46	224	0.83	16	0.53	0.410
4	61	271	0.82	16	0.52	0.560
5	92	315	0.78	21	0.46	0.720
(16) Tamil Nadu (91850 km²)						
1	28	66	0.80	100	0.52	1.37
2	74	107	0.81	83	0.46	2.50
3	133	142	0.83	77	0.41	3.19
4	171	171	0.83	91	0.31	7.61
(17) East Uttar Pradesh (138475 km²)						
1	13	72	0.85	63	0.63	0.64
2	31	105	0.80	45	0.59	0.71
3	47	140	0.80	36	0.58	0.78
4	62	172	0.80	33	0.55	1.03
5	76	210	0.75	31	0.49	1.35
(18) Plains of west Uttar Pradesh (68800 km²)						
1	7	51	0.95	53	0.79	0.23
2	19	86	0.90	59	0.65	0.80
3	45	116	0.90	50	0.62	1.06
4	71	148	0.86	67	0.46	3.05
(19) Hills of west Uttar Pradesh (43000 km²)						
1	24	70	0.69	100	0.44	1.54
2	33	90	0.71	102	0.39	2.65
3	43	117	0.68	81	0.37	2.79
(20) West Bengal (66925 km²)						
1	21	61	0.72	131	0.43	2.00
2	39	94	0.65	116	0.32	3.73
3	65	120	0.66	100	0.30	4.52
4	80	146	0.70	11	0.26	7.50

*Due to elongated structure of the region, circles were not appropriate for the division of area. So region divided into 4 zones.

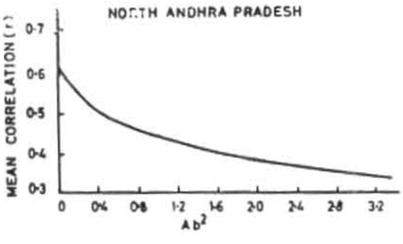

Fig. 2

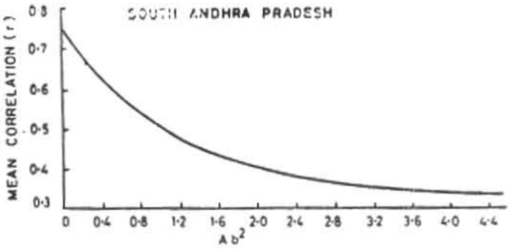

Fig. 3

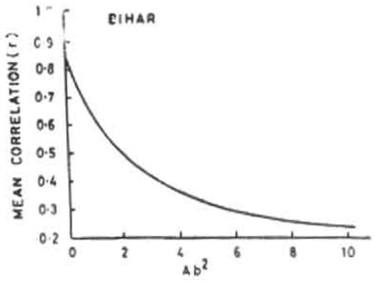

Fig. 4

Figs. (2-4). Space correlation structure

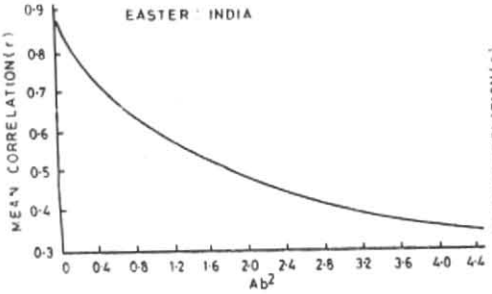

Fig. 5

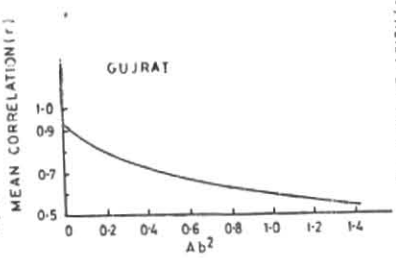

Fig. 6

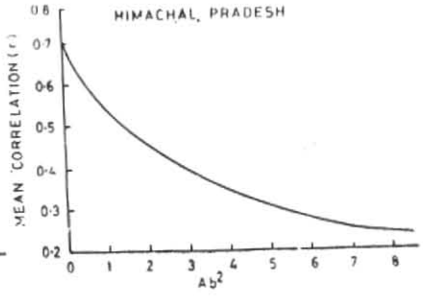

Fig. 7

Figs. (5-7). Space corelation structure

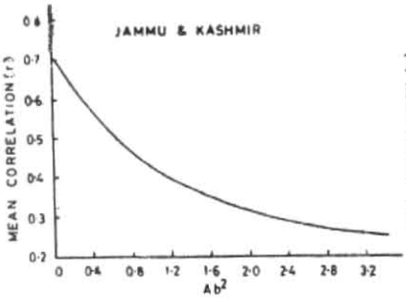

Fig. 8

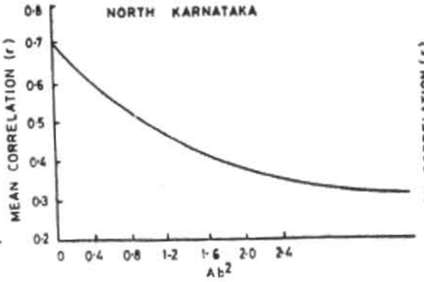

Fig. 9

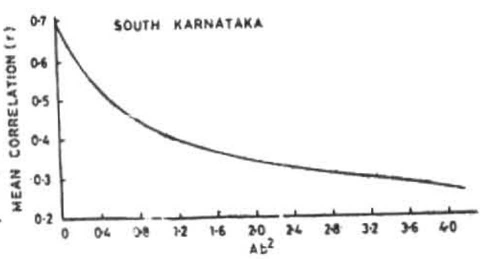

Fig. 10

Figs. (8-10). Space corelation structuture

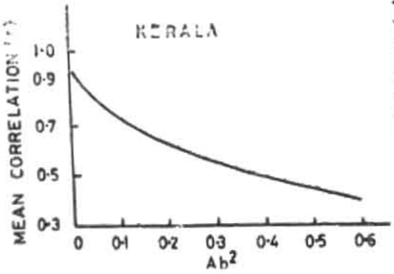

Fig. 11

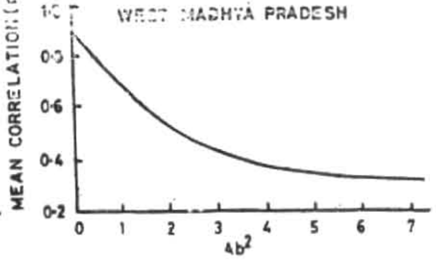

Fig. 12

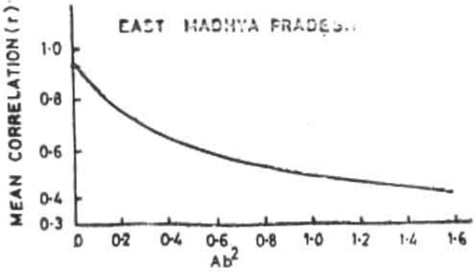

Fig. 13

Figs. (11-13). Space correlation structure

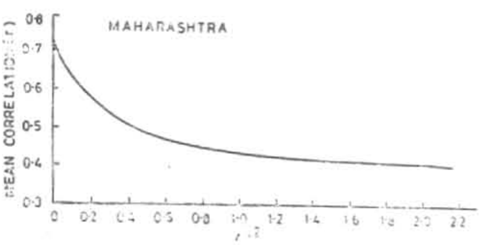

Fig. 14

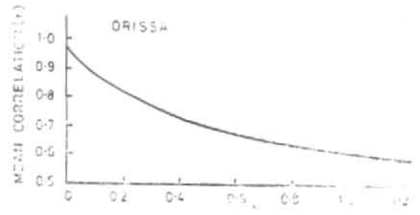

Fig. 15

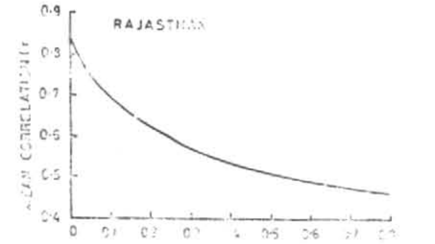

Fig. 16

Figs. (14-16). Space correlation structure

Fig. 17

Fig. 18

Fig. 19

Figs. (17-19). Space correlation structure

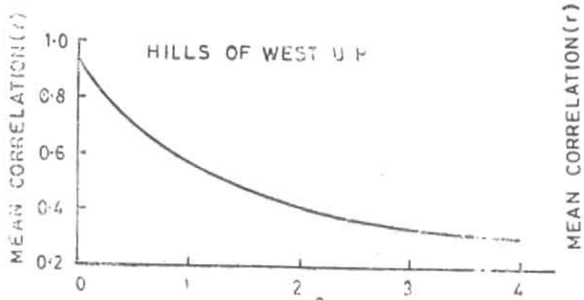

Fig. 20

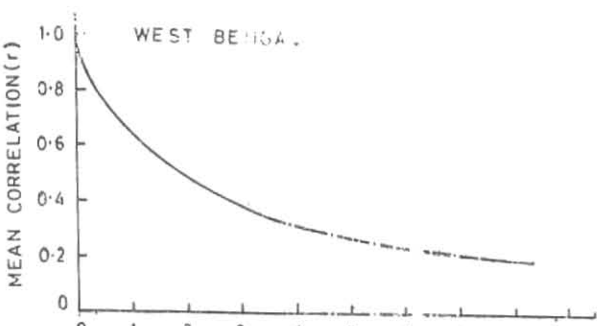

Fig. 21

Figs. (20-21). Space correlation structure

6. Some areas of application

6.1. *Network design*

Many authors namely, Rodriguez-Iturbe and Mejia (1974), Kaur and Upadhyay (1987), etc have indicated that the variance of estimate from point to areal rainfall is given by :

$$V(P_A) = \sigma_P^2 \ f_1(T) \cdot f_2(n) \qquad (6)$$

where, $f_1(T)$ is the temporal reduction factor depending on the length of observation (T) and $f_2(n)$ the spatial reduction factor depending on the number of raingauges n.

If we assume that rainfall between two stations is not independent then :

$$f_2(n) = [1 + (n-1)\ \bar{r}\ ]/n$$

where $\bar{r}$ is the mean correlation of rainfall series. As $n \to \infty, f_2(n) \to \bar{r}$ which means that the error of estimate can not be reduced significantly by simply increasing the number of raingauges. The concept given in Eqn. (6) is more appropriate for evaluating network density of an area for a desirable accuracy in estimation of areal mean.

6.2. *Point to areal rainfall*

The areal reduction factor P_A/P_o may be expressed as a function of $\bar{r}$. This function could be a power function, *i.e.*,

$$P_A/P_o = k\,(\bar{r})^{1/m}$$

As an illustration (Upadhyay *et al.* 1984), let $P_A/P_o = (\bar{r})^{1/2}$

Area A (km²)	$\bar{r}$	P_A/P_o
250	0.64	0.80
500	0.62	0.79
1000	0.61	0.78
1500	0.60	0.77
2000	0.59	0.77
3000	0.58	0.76
4000	0.57	0.75
5000	0.55	0.74

6.3. *Regional transfer of information*

Since the stations are not statistically independent in rainfall observations, a part of information of a station is contained in the observations taken over other stations. Thus, the equivalent number of independent stations (n_e) having same information content will be less than n. If μ_i and σ_i^2 are the mean and variance of ith station, then the regional mean μ and variance σ_μ^2 is given by :

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\mu_i$$

$$\sigma_\mu^2 = \frac{1}{n^2}\sum_{i=1}^{n}\sigma_i^2 + \frac{2}{n^2}\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} r_{ij}\,\sigma_i\,\sigma_j$$

$$= \frac{\sigma^2}{n}\left[1 + (n-1)\,\bar{r}\right]$$

Then the information content regarding μ is given by :

$$I_\mu = \frac{\sigma^2/n}{(\sigma^2/n)\,[1+(n-1)\,\bar{r}\,]} = [1+(n-1)\,\bar{r}\;]^{-1}$$

$$n_e = n/\,[1 + (n-1)\;\bar{r}\,)]$$

6.4. *Interpolation of missing data*

The concept of interpolation is based directly on the relationship of one observation with the other as spatial or temporal. Kagan (1972) have derived the errors in interpolation for a triangular grid network as a function of the rainfall variability, network density, the spacing between two stations and the correlation structure, it is given by :

$$Z_{\text{int}} = Cv\;[\{(1-a)/3\} + 0.52\,ba\,(A/n)^{\frac{1}{2}}]^{\frac{1}{2}}$$

where a and b are the parameters of the correlation structure; Cv is the coefficient of rainfall variation and A is the area of the catchment.

7. Conclusion

The statistical survey based on country *vide* voluminous data set presents important features of space correlation between rainfall series recorded at various points. Some of the features are :

(*i*) The rainfall correlation $r(s)$ between two stations decreases with distance (s) between them and follows generally an exponential pattern.

(*ii*) For a homogeneous area, concept of mean correlation ($\bar{r}$) is applicable which can be dervied by :

$$\bar{r} = \int r\,(s)\;\phi\,(s)\;ds$$

where, $\phi\,(s)$ is a density function for the random variable s. The $\phi\,(s)$ assumes simplest form for a circular area given by :

$$\phi\,(s) = 2s/R^2$$

R being radius of the circle.

(*iii*) $\bar{r}$ depends on area (A) and the parameter b of the correlation pattern. The combined effect of these two factors is indicated by a dimensionless quantity Ab^2. It is found for all homogeneous regions that $\bar{r}$ decreases with Ab^2 and generally follows an exponential pattern.

(*iv*) The concept of $r(s)$ and $\bar{r}$ have vide range of application in the areas of network design, interpolations, point to areal rainfall relationships, information transfer, regression analysis etc.

Acknowledgements

The authors are grateful to Dr. S.M. Kulshrestha, Director General of Meteorology and Dr. N. Sen Roy, Additional Director General of Meteorology for their keen interest and encouragement. They also express their gratitude to Shri S.R. Puri, Director for useful discussions and suggestions. Thanks are due to S/Shri Mukhtiar Singh, Vinod Kumar and Ms. Asha for technical and secretarial assistance.

References

Eagleson, P.S., 1967, "Optimum density of rainfall networks", *Water Resour. Res.*, **3,** 4, pp. 1021-1023.

Erikson, E., 1972, "The mean error of areal precipitation estimates in the valon representative basin", Report AC-7, Int. Met. Inst. in Stockholm.

Hershfield, D.M., 1965, "On the spacing of Raingauges", Symp. on Design of Hydrol. Networks, *Int. Assoc. Sci. Hyd. Publ.*, **67,** pp. 72-75.

Kagan, R.L., 1966, "An evaluation of the representativeness of the precipitation", *Data works of the Geophys. Obs.*, USSR, **191.**

Kagan, R.L., 1972, "Planning the spatial distribution of Hydro-meteorological stations to meet an error criterion", WMO case book of Hydrological Network Design Practice, III, 1.2.

Kaur, S. and Upadhyay, D.S., 1987, "Precipitation network design for mountainous catchment", *Mausam*, **38,** 1, pp. 73-78.

Ramanathan, A.S. *et al.*, 1981, "Precipitation network design for hydrological purposes", Part-I, I.M.D., Pre-publ. Sci. Rep.

Rodriguez-Iturbe, I. and Mejia, J.M., 1974 (a), "The design of rainfall networks in time and space", *Water Resour. Res.*, **10,** pp. 713-728.

Rodriguez-Iturbe, I. and Mejia, J.M., 1974 (b), "On the transformation of point rainfall to areal rainfall', *Water Resour. Res.*, **10,** pp. 729-735.

Upadhyay, D.S. *et al.*, 1984, "The depth area relationship— a theoretical approach", Proc. of "Seminar on Hydrology", Pune.