Arch. Met. Geoph. Biokl., Ser. A, 30, 261–277 (1981)

ARCHIVES
FOR METEOROLOGY,
GEOPHYSICS, AND
BIOCLIMATOLOGY

551.577.22:551.509.314

Indian Institute of Tropical Meteorology, Pune, India

Persistence in Daily and 5-Day Summer Monsoon Rainfall over India

S. V. Singh, R. H. Kripalani, Priya Shaha, P. M. M. Ismail, and **S. D. Dahale**

With 2 Figures

Received December 2, 1980

Summary

Persistence in daily rainfall occurrence during the peak summer monsoon months of July–August has been studied for 22 stations, widely spread over the country. The logarithmic, Markov chain order 1 and 2 and modified logarithmic models are fitted to the runs of wet and dry days to examine the nature of persistence and the fitness tested by χ^2 tests. Markov chain order 2 (modified logarithmic model) fits to the runs of dry spells (wet spells) better than the other models. The logarithmic model is found to be a poor fit for both types of spells. Persistence in daily and 5-day precipitation amounts and in daily expected areal coverage of precipitation is examined by computing auto-correlation coefficients for various lags. Persistence in 5-day precipitation amount and in 5-day expected areal coverage of precipitation is also examined by constructing 3×3 contingency tables. All these results show that persistence is more prominently exhibited in certain parts of central and western India.

Zusammenfassung

Andauer täglicher und 5-tägiger Sommermonsunregen über Indien

Die Andauer des Vorkommens täglicher Regenfälle in den Hauptsommermonsunmonaten Juli und August wurde für 22 über das Land verstreute Stationen untersucht. Markovsche Ketten erster und zweiter Ordnung und modifizierte logarithmische Modelle wurden dem Verlauf von feuchten und trockenen Tagen zur Prüfung der Art der Andauer angepaßt und ihre Brauchbarkeit wurde mit χ^2-Tests untersucht. Die Markovsche Kette zweiter Ordnung (modifiziertes logarithmisches Modell) paßt sich dem Verlauf von Trockenperioden (Feuchtperioden) besser an als andere Modelle. Es wurde gefunden, daß das logarithmische Modell eine schlechte Anpassung für beide Typen von Perioden gibt. Die Andauer von täglichen und fünftägigen Niederschlagsbeträgen und von fünftägig erwarteter räumlicher Verbreitung des Niederschlags wird auch durch Konstruktion von 3 × 3 Kontingenztabellen untersucht. Alle diese Ergebnisse zeigen, daß die Andauer sich besonders hervorragend in bestimmten Teilen von Zentral- und West-Indien erweist.

0066-6416/81/0030/0261/$ 03.40

1. Introduction

In the present study we shall be concerned with examination of persistence in daily and 5-day precipitation in summer monsoon season. Since a study of persistence of daily precipitation is more attractive because of readily available long and reliable records, better understanding of behaviour of daily atmospheric circulation systems and direct applicability of such results to the major operational forecasting programs, the major part of the present

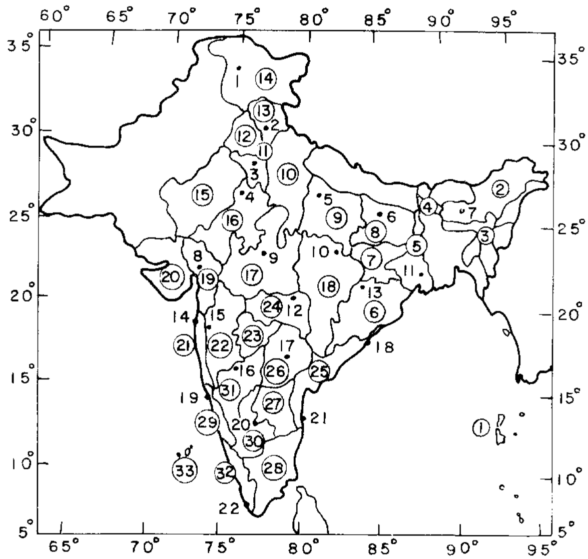

Fig. 1. Locater map of the 22 stations and 33 meteorological subdivisions (SD's). The circled numbers are the serial numbers of the SD's. The dots with uncircled numbers represent the 22 rainfall stations

study will be devoted to the persistence of daily rainfall occurrence. We shall investigate in some detail the extent and nature of persistence in daily precipitation occurrence during the peak summer monsoon season (July–August) utilising data of 19 years (1955–73). As the results of such previous studies [16, 19] show smooth large scale distribution of persistence parameters, we shall limit these investigations to 22 stations (Fig. 1) only, which are more than adequate to represent various climatic regimes. We shall present in the following section conditional probabilities (CP's) for two threshold criteria (2.5 mm and 10 mm, the 2.5 mm criterion represents official definition of rainy day as defined by the India Meteorological Department, while 10 mm represents approximately the average daily rain intensity over major parts of the country during the height of summer monsoon season) for various lags and the results of t-tests for testing the significance of difference from respective climatic probabilities. In the same section we shall present average length of wet or dry spells obtained easily from CP's. In section 3 we shall fit various models to the frequency distribution of wet and dry spells and test their fitness. In section 4 we shall present some results of our investigations on persistence in daily rainfall amount and daily expected areal coverage (EAC) of precipitation (PPN) for meteorological subdivisions and the main monsoon trough zone (lying between 15°N and 30°N and West of 87°E). In section 5 we shall examine persistence in 5-day rainfall amounts and occurrences.

These results will obviously be of some use in characterising the climate, particularly the frequency distribution of lengths of dry or wet spells, for agricultural or water resources planning as argued by most of the previous workers who have worked on these lines (e. g. [5, 22]). In our opinion these results will be equally useful in actual weather forecasting and forecast evaluation. Although some of the results (viz. auto-correlation coefficients) presented below can be applicable in determining number of independent observations utilised in making the time averaged climatic means, a subject which has drawn considerable attention in recent years. We shall, however, focus our attention mainly on the forecast applicability of the results.

2. Persistence in Rainfall Occurrence

The four CP's representing transition rates into two rainfall states (D or W) have been computed for 1 to 10 days lag between conditional and verifying periods for both the threshold criteria. The persistence, which is a measure of probability of occurrence of an event (W or D) above climatological occurrence [12] is defined as follows:

Wet Persistence $PrW = PW/W - PW$
Dry Persistence $PrD = PD/D - PD$

where the symbol / is read as given that, PW (PD) is climatological probability of PPN (no PPN) and PW/W and PD/D are respectively CP's of occurrence of PPN or no PPN in the verifying period given that PPN or no PPN occurred in conditional periods. The values of PW, PD, PW/W, PD/D, PrW and PrD are presented in Table 1 along with other parameters to be discussed later. These measures of persistence viz. PrW and PrD are also presented in the map form in Fig. 2 for 2.5 mm threshold only. While drawing the isolines in the figure we received considerable guidance from the previous results obtained for 221 stations [16]. The general conclusion which can be easily drawn from these diagrams is that the dry regions of

Table 1. *Statistical Parameters.* Climatic (*PW*, *PD*) and conditional probabilities (*PW/W*, *PD/D*) of dry and wet days for threshold 2.5 mm and 10.0 mm. The *t*-values are obtained from testing significance of *PW/W*–*PW* and *PD/D*–*PD* by student's *t*-test. *ALS* is the average length of spell. *X* is the parameter of the LOG model and *Q* and *A* are the parameters of the MODLOG model (see text)

a) Threshold = 2.5 mms

S. No.	Station	Wet spells							Dry spells						
		PW	*PW/W*	*t*	*ALS*	*X*	*Q*	*A*	*PD*	*PD/D*	*t*	*ALS*	*X*	*Q*	*A*
1.	Srinagar	.15	.32	5.5	1.47	.47	.39	4.90	.85	.88	1.9	7.11	.92	.96	4.10
2.	Ambala	.33	.50	6.3	1.99	.71	.61	1.50	.67	.77	4.2	3.96	.87	.81	5.50
3.	Delhi	.34	.52	6.4	2.08	.70	.61	3.00	.66	.76	4.2	3.75	.88	.88	0.80
4.	Jaipur	.36	.57	7.7	2.33	.73	.74	0.90	.64	.77	5.7	3.93	.87	.91	0.70
5.	Lucknow	.40	.54	5.7	2.20	.72	.70	0.90	.60	.70	4.3	3.22	.85	.77	3.00
6.	Patna	.39	.53	5.2	2.11	.70	.59	4.60	.61	.71	4.1	3.23	.82	.82	1.70
7.	Gauhati	.46	.56	4.0	2.26	.73	.72	0.80	.54	.64	3.7	2.61	.77	.68	5.00
8.	Ahmedabad	.36	.58	8.2	2.41	.74	.76	0.70	.64	.77	5.9	4.06	.89	.74	4.30
9.	Bhopal	.47	.62	6.2	2.65	.76	.75	2.50	.53	.68	6.0	2.97	.83	.78	1.20
10.	Ambikapur	.59	.66	3.3	2.92	.80	.78	2.10	.41	.53	4.1	2.05	.72	.65	0.90
11.	Calcutta	.53	.59	2.5	2.43	.73	.69	3.70	.47	.55	2.7	2.11	.70	.59	5.00
12.	Nagpur	.49	.58	4.0	2.38	.75	.70	1.90	.51	.61	3.8	2.50	.77	.78	0.50
13.	Jharsaguda	.56	.64	3.6	2.76	.77	.71	7.00	.44	.56	4.2	2.17	.70	.59	7.00
14.	Bombay	.69	.80	5.6	4.88	.90	.95	0.40	.31	.58	8.9	2.24	.77	.77	0.00
15.	Poona	.37	.61	8.9	2.55	.77	.71	2.60	.63	.78	6.8	4.29	.88	.93	0.20
16.	Bijapur	.20	.41	6.9	1.69	.60	.45	7.00	.80	.85	3.2	6.12	.93	.88	2.00
17.	Hyderabad	.33	.48	5.4	1.92	.67	.72	0.00	.67	.75	3.6	3.73	.89	.80	2.10
18.	Visakhapatnum	.30	.40	4.0	1.68	.57	.46	6.60	.70	.76	2.5	3.87	.85	.97	0.10
19.	Honavar	.86	.90	3.4	9.67	.95	.95	1.50	.14	.47	9.5	1.75	.65	.66	0.00
20.	Bangalore	.28	.38	3.6	1.61	.56	.43	5.00	.72	.76	1.9	3.93	.87	.96	0.00
21.	Madras	.28	.36	2.8	1.55	.55	.48	1.10	.72	.75	1.5	3.78	.86	.84	2.30
22.	Trivandrum	.38	.58	7.5	2.37	.77	.72	1.00	.62	.76	6.0	3.86	.86	.85	1.70

Table 1 (continued)

b) Threshold = 10.0 mms

S. No.	Station	Wet spells							Dry spells						
		PW	*PW/W*	*t*	*ALS*	*X*	*Q*	*A*	*PD*	*PD/D*	*t*	*ALS*	*X*	*Q*	*A*
1.	Srinagar	.06	.19	4.4	1.24	.36	.37	0.00	.94	.95	0.9	14.75	.96	1.00	5.00
2.	Ambala	.23	.43	6.8	1.75	.63	.60	0.40	.77	.84	3.5	5.54	.91	0.89	7.00
3.	Delhi	.20	.35	4.9	1.53	.52	.39	6.00	.80	.84	2.1	5.43	.91	0.89	4.90
4.	Jaipur	.20	.39	6.3	1.64	.55	.44	7.00	.80	.85	3.0	5.86	.91	0.91	4.00
5.	Lucknow	.24	.40	5.3	1.66	.60	.49	2.20	.76	.81	2.7	4.82	.90	0.86	3.20
6.	Patna	.23	.37	5.1	1.60	.58	.46	2.40	.77	.82	2.6	5.11	.90	0.86	5.00
7.	Gauhati	.26	.38	4.1	1.60	.52	.43	6.00	.74	.79	2.1	4.28	.87	0.84	5.00
8.	Ahmedabad	.18	.38	6.8	1.62	.54	.44	5.00	.82	.87	2.8	6.77	.93	0.80	7.00
9.	Bhopal	.29	.44	5.6	1.79	.59	.50	7.00	.71	.78	3.4	4.31	.87	0.84	5.00
10.	Ambikapur	.36	.50	5.3	2.01	.67	.55	7.00	.64	.72	3.6	3.38	.84	0.76	6.30
11.	Calcutta	.29	.37	2.8	1.58	.55	.56	0.20	.71	.74	1.6	3.67	.83	0.80	7.00
12.	Nagpur	.26	.34	3.1	1.52	.47	.40	7.00	.74	.78	1.7	4.20	.88	0.79	6.00
13.	Jharsaguda	.34	.42	3.1	1.73	.57	.48	6.00	.66	.71	2.1	3.26	.80	0.78	6.00
14.	Bombay	.41	.62	8.2	2.66	.79	.71	3.10	.59	.76	7.0	3.79	.87	0.92	0.10
15.	Poona	.14	.38	7.8	1.61	.56	.59	0.10	.86	.90	3.0	8.97	.95	0.96	0.70
16.	Bijapur	.08	.30	7.5	1.44	.43	.35	7.00	.92	.94	1.9	14.52	.96	0.99	6.00
17.	Hyderabad	.18	.36	6.2	1.56	.57	.57	0.00	.82	.87	2.5	6.70	.94	0.99	0.40
18.	Visakhapatnum	.13	.21	3.0	1.27	.34	.25	6.90	.87	.89	1.0	7.85	.93	1.00	2.60
19.	Honavar	.65	.75	5.1	4.03	.88	.82	3.00	.35	.59	7.6	2.22	.72	0.75	0.30
20.	Bangalore	.11	.16	2.1	1.20	.30	.36	0.00	.89	.90	0.5	8.75	.95	1.00	0.30
21.	Madras	.14	.23	3.3	1.31	.39	.33	1.40	.86	.88	1.1	7.23	.93	1.00	0.30
22.	Trivandrum	.19	.40	7.3	1.68	.59	.57	0.60	.81	.87	3.4	6.86	.93	0.91	2.10

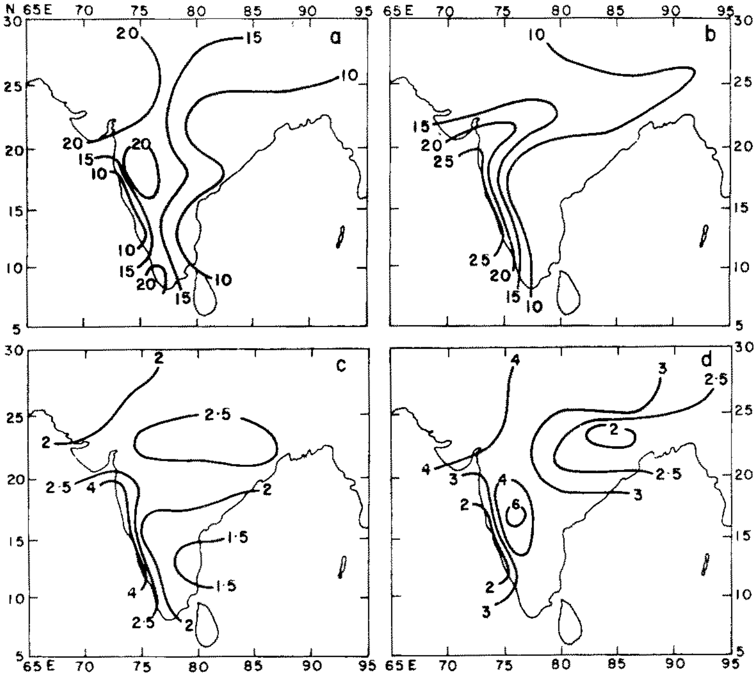

Fig. 2. a) Wet persistence $PrW = PW/W-PW$, b) dry persistence $PrD = PD/D-PD$. Isolines labelled in terms of percentages are drawn at intervals of 5. c) Average length of wet spells in days. d) Average length of dry spells in days

northwest India and midwest peninsula show more wet persistence while the wet regions of west coast, upper eastcoast, central parts of the country and northeast India show less wet persistence. Likewise wet regions of west-coast, central parts of the country and northeast India show more dry persistence than the dry regions of northwest India or southeast peninsula. The persistence measure PrW thus suggests entirely different picture of wet persistence distribution over the country than that shown by Srinivasan [19] with the help of X-parameter of a fitted logarithmic (LOG) model. The distribution of X-parameter in his presentation (and also by analysis of our results of fitting LOG distribution, see Table 1), was similar to that of PW or PW/W. The results of the present study may appear contradictory or even incorrect at first sight but it is not difficult to appreciate them if we understand that the difference between PW/W (PD/D) and PW (PD) will be small where PW (PD) itself is high. Such a geographical distribution of persistence measures reflect on the nature of circulation systems. In the dry regions PPN is generally caused by synoptic systems with influence period exceeding one day, likewise in wet regions PPN interrupting synoptic situations continue to last for more than a day.

The t-values obtained from testing the difference between CP's and the climatic probabilities (PW or PD) for 1 day lag (t-test is also performed for other lags) are significant for all stations for wet persistence and for all but 3 stations for dry persistence (Table 1). The distribution of these values is similar to the respective distribution of PrW or PrD. Besson's coefficient of persistence, persistence ratio and their confidence interval [3] and χ^2-values from 2×2 contingency tables prepared for four possible transitions between conditional and verifying periods lead to similar conclusions. These results are however not presented.

Length of dry and wet spells: It is trivial to show that average length of wet and dry spell is given by $1/(1-PW/W)$ and $1/(1-PD/D)$ respectively. The average length of wet and dry spells are given in Table 1 and Fig. 2 (only for 2.5 mm threshold) for better geographical exposition. The average length of wet spell varies from 2 to 3 days over major parts of the country except over west coast and southeast peninsula where it is more and less than this range respectively. The average length of dry spell is higher than average length of wet spell and varies from 2 to 4 days over most of the country excluding west coast. This leads us to conclude that the average length of the weather cycle varies from 4 to 6 days over major parts of the country. Power spectrum analysis of daily rainfall amounts for each of the 19 years and for all the 33 meteorological subdivisions (SD's) carried out by Singh and Prasad [17] has shown significant peaks in this range of the period.

3. Fitting Models to Examine Nature and Extent of Persistence

To examine over all behaviour of runs of dry or wet weather various types of mathematical models are fitted. Although all the desired characteristics of the rainfall occurrence can be obtained empirically by gathering large sets of data, fitting models is advantageous as it helps in understanding the nature of persistence and gives relatively smooth and stable estimates of various parameters (e. g. [6, 7]) from smaller length of data than what can be obtained empirically. A survey by the authors of more than 100 papers published in meteorological journals during the last two decades revealed that the Markov chain, first used by Gabriel and Neumann [8] in meteorology and LOG model first proposed by Williams [24] are the most commonly used models, to investigate the nature of persistence. In LOG model the persistence factor increases with the length of the spell of particular type of weather occurred and in Markov chain the probability of occurrence of an event is assumed to be constant after a particular length of the spell of similar event has occurred, depending on the order of the chain desired to be fitted. Green [10] proposed a modified LOG (MODLOG) model which has intermediate characteristics between these two models.

Considerable but sporadic work has been done in this line on Indian rainfall [1, 5, 19, 20, 22]. These workers have not compared performance of various models. Any one model may not fit to data of all stations or it may fit to only one kind of spell wet or dry. Further the fitness of a model may change if the threshold criterion of test procedure is changed [2]. Moreover results obtained by various workers for different stations cannot be compared as they are based on different length and period of data and some on different threshold. In this section we shall compare the fitness of various popular models to the runs of dry or wet spells of 22 stations for 2.5 mm and 10 mm threshold criteria.

The following models are fitted to the frequency distribution of dry and wet spells:

Model	Frequency/probability of spells of length r	Parameters of model to be determined	Normalizing factor
1. LOG model [24]	$\alpha X^r/r$	α, X	$\sum_{r=1}^{62} \alpha X^r/r$ = total no of spells of $\leqslant 62$ days
2. Markov chain [23]	For dry days $PW/D\,(PD/D)^{r-1}$	$PW/D, PD/D$	$\sum_{r=1}^{62} PW/D\,(PD/D)^{r-1} = 1$
	For wet days $PD/W\,(PW/W)^{r-1}$	$PD/W, PW/W$	$\sum_{r=1}^{62} PD/W\,(PW/W)^{r-1} = 1$
3. Markov chain order 2 [15]	For dry days $f_2\,(1-PD/DD)^{r-2}$	$f_2, PD/DD$	$\sum_{r=3}^{62} f_2\,(1-PD/DD)^{r-2}$
	For wet days $f_2\,(1-PW/WW)^{r-2}$	$f_2, PW/WW$	$\sum_{r=3}^{62} f_2\,(1-PW/WW)^{r-2}$ = Total spells of 3 or more days
4. Green's MODLOG [9]	$\frac{Q^r}{r+A}$	Q, A	$\sum_{r=1}^{15} \frac{Q^r}{r+A}$ = Total spells of $\leqslant 15$ days

For further details of other notations and fitting procedures the readers may refer to [19] for LOG model, [23] for Markov chain order 1, [15] for Markov chain order 2 (however see below), and [9] for MODLOG model, which have been followed in this study. The frequencies of spells of various lengths are normalized such that the sum of frequencies of all computed spells under consideration is same as the sum of frequencies of all observed spells [2]. This is essential but seems to have been ignored by Sakamoto [15], Lawrence [13] and Caskey [4] who build up frequencies of the higher length spells from the frequency of some lower length of spell without normalising the final distribution. To test the fitness of the models, χ^2 values are computed between observed and model expected frequencies of wet and dry spells of various lengths. While computing the χ^2 values the frequency distribution is truncated upto the spell length beyond which the model computed frequency falls below five and all the higher spell frequencies are pooled in this class for both observed as well as model computed frequency distributions.

Table 2. *Frequencies of Different Ranges of Probability of Exceeding or Equalling Computed χ^2 by Chance, or Symbolically P ($\chi^2 \geqslant \chi_c^2$)*

	Threshold = 2.5 mms								Threshold = 10.0 mms								
Ranges	0.05	0.05 0.10	.1 .2	.2 .3	.3 .5	.5 .7	.7 .8	.8 .9	0.05	0.05 0.10	.1 .2	.2 .3	.3 .5	.5 .7	.7 .8	.8 .9	Total
Models	Dry spells																
LOG	16	1	3	1	1	0	0	0	20	2	0	0	0	0	0	0	22
Markov-1	10	2	4	1	3	1	1	0	11	0	0	3	7	1	0	0	22
Markov-2	8	0	2	3	5	3	0	1	5	0	5	4	3	3	2	0	22
MODLOG	14	1	3	1	1	2	0	0	19	2	0	0	0	1	0	0	22
	Wet spells																
LOG	16	2	3	0	1	0	0	0	15	1	2	1	2	1	0	0	22
Markov-1	3	2	5	5	5	2	0	0	6	2	4	3	4	2	1	0	22
Markov-2	9	2	3	1	5	0	1	1	19	0	2	1	0	0	0	0	22
MODLOG	4	0	7	6	2	3	0	0	2	1	7	4	3	3	1	1	22

The χ^2 results of testing fitness of various models are given in Table 2 in a summary form. The figures shown in this Table are the number of stations (out of total 22) for which computed χ^2 value will be exceeded by chance with the probability range shown above the column. For better fitness more stations should lie in the columns to the right, as for these columns the computed χ^2 is smaller and probability of its exceeding by chance is larger. The table reveals that Markov chain order 2 is better fit for dry spells for both types of threshold while MODLOG model is better fit for wet spells. Markov chain order 1 is a close competitor for both types of spells. LOG model is generally a poor fit and as such its parameter X may not display the right geographical distribution of persistence. It should be remembered that MODLOG model is fitted through an iterative process involving search of parameters Q and A which give minimum χ^2 and as such does not have unique persistence specification as in the case of Markov chains or LOG model.

On examination of computed frequencies of spells of various lengths from different models, in comparison with observed frequencies (Table 3 for 5 stations for 2.5 mm threshold), it was found that frequencies of 1 day spell computed from LOG model are more than observed in almost all cases and for 2 day spell are equal or less than observed. Frequencies of 1 day (2 day) spell computed from Markov chain order one are equal or less (generally equal) as compared to observed ones while the frequencies obtained from MODLOG model are generally equal for both 1 or 2 day spell. Markov chain order 2 and MODLOG model are equally efficient in reproducing frequencies of 3 days length of wet or dry spells. Beyond 3 days the frequencies of spells fall considerably, not allowing to draw any firm conclusion. This suggests that the frequency distribution of 1, 2 or 3 day spell can be generated by a model having characteristics between LOG and Markov chain. While judging this equality in frequency distribution a subjective tolerance of ± 5 spells was allowed which is quite reasonable for the spells of 1, 2 or 3 days length.

4. Auto-Correlation in Rainfall Amount and EAC

Persistence in rainfall amount is examined by computing auto-correlation coefficients in daily rainfall values for 1 to 10 day lags. Since point rainfall amounts are highly variable, relatively smooth space averaged rainfall values are considered for this purpose, as our main interest is to investigate the general synoptic scale behaviour of monsoon rainfall. Further, very high values of space averaged rainfall (values more than mean plus three times standard deviation of the series concerned) if obtained are equalled to the sum of the mean of the series plus 3 times standard deviation. The auto-correlation coefficients are computed for each monsoon month and for each of the 19 years (1955–73) separately for 33 subdivisions covering India. Number of occasions (years) out of total 19, when the autocorrelation coefficients for lag 1, 2 and 3 days exceed or equal 0.35, the significant value at 5% level, are shown in Table 4. It should be remembered that a correlation value of 0.27 is significant at 5% level if one tailed t-test is applied. Thus, the criteria set for significance is quite stringent. For lag of 1 day all the SD's show significant persistence at 5% level although they are not significant for each year. A noteworthy feature is that for lag upto 3-days some subdivisions in western and central India show significant persistence in all the summer monsoon months showing that persistence can be a very useful parameter in predicting or evaluating the forecast of not only rainfall

Table 3. *Frequencies of Observed and Computed Spells of Dry and Wet Days for Various Lengths for 2.5 mms Threshold.* Spell lengths of more than or equal to 7 days are pooled

Wet spells								Dry spells						
Spell length	1	2	3	4	5	6	≥ 7	1	2	3	4	5	5	≥ 7
New Delhi:														
Observed	99	39	31	9	7	4	4	66	47	23	11	16	10	34
LOG	111.9	39.3	18.4	9.7	5.4	3.2	5.1	86.8	38.0	22.2	14.6	10.2	7.5	27.5
Markov-1	92.6	48.2	25.1	13.0	6.8	3.5	3.7	49.5	37.6	28.6	21.7	16.5	12.5	37.5
Markov-2	99.0	39.0	24.7	13.6	7.5	4.1	5.3	66.0	47.0	21.5	16.6	12.8	9.9	33.2
MODLOG	96.2	42.1	22.0	12.5	7.4	4.6	7.2	64.4	38.4	25.5	18.1	13.3	10.1	33.3
Nagpur:														
Observed	106	62	24	20	13	8	8	108	61	26	10	10	8	19
LOG	130.1	48.9	24.5	13.8	8.3	5.2	10.2	126.4	48.8	25.1	14.5	9.0	5.8	12.5
Markov-1	101.2	58.7	34.1	19.7	11.4	6.6	9.2	94.4	57.6	35.1	21.4	13.1	8.0	12.6
Markov-2	106.0	62.0	28.5	17.4	10.6	6.5	9.9	108.0	61.0	26.1	16.8	10.8	6.9	12.4
MODLOG	105.3	53.5	30.3	18.3	11.5	7.4	14.6	108.6	49.7	28.1	17.4	11.5	7.8	18.9
Madras:														
Observed	142	49	11	9	2	0	2	67	44	28	22	15	11	36
LOG	147.9	40.8	15.0	6.2	2.7	1.3	1.1	97.8	42.0	24.1	15.5	10.6	7.6	25.3
Markov-1	137.6	49.5	17.8	6.4	2.3	0.8	0.4	55.7	41.8	31.3	23.5	17.6	13.2	38.9
Markov-2	142.0	49.0	14.6	5.7	2.3	0.9	0.5	67.0	44.0	26.5	20.2	15.4	11.8	36.1
MODLOG	141.7	43.3	16.7	7.1	3.2	1.5	1.5	67.4	41.9	28.3	20.2	14.9	11.2	36.2
Bombay:														
Observed	48	30	15	14	8	9	43	93	28	14	7	3	6	11
LOG	64.3	29.1	17.5	11.9	8.6	6.5	29.0	85.1	32.7	16.7	9.7	5.9	3.8	8.0
Markov-1	33.4	26.7	21.3	17.1	13.7	10.9	42.5	68.1	39.5	22.9	13.3	7.7	4.5	6.1
Markov-2	48.0	30.0	15.3	12.7	10.5	8.7	41.6	93.0	28.0	13.2	9.0	6.1	4.1	8.9
MODLOG	47.4	25.3	17.1	12.7	10.1	8.3	40.1	80.9	32.4	17.3	10.3	6.6	4.4	10.1
Calcutta:														
Observed	115	40	51	19	14	10	10	123	66	40	13	6	5	7
LOG	144.5	52.7	25.6	14.0	8.2	5.0	9.0	150.8	52.9	24.8	13.0	7.3	4.3	7.0
Markov-1	106.2	62.7	37.0	21.8	12.9	7.6	10.8	117.0	64.3	35.4	19.5	10.7	5.9	7.1
Markov-2	115.0	40.0	39.8	24.6	15.2	9.4	15.0	123.0	66.0	33.0	17.7	9.5	5.1	5.8
MODLOG	104.9	61.5	36.5	21.9	13.3	8.1	12.8	124.0	61.4	32.4	17.8	10.1	5.8	8.2

Table 4. *The Number of Years 1955–1973 in Which the Autocorrelation Coefficients in Daily PPN Amounts for Lags 1, 2 and 3 Days Exceed or Equal 0.35* (i. e. $r \geqslant 0.35$)

Month	June			July			August			September		
Subdivision	Lag 1	Lag 2	Lag 3	Lag 1	Lag 2	Lag 3	Lag 1	Lag 2	Lag 3	Lag 1	Lag 2	Lag 3
1. Bay Islands	11	4	4	11	2	–	13	2	2	11	4	–
2. North Assam	13	5	1	9	1	1	7	2	1	8	2	–
3. South Assam	12	6	1	8	1	–	5	1	–	8	2	–
4. Sub-Himalayan W. Bengal	8	2	1	9	2	–	5	–	–	8	–	–
5. Ganetic W. Bengal	10	2	1	7	1	–	9	–	–	15	3	1
6. Orissa	16	6	3	14	1	–	9	–	–	14	3	–
7. Bihar Plateau	10	6	2	10	1	–	8	2	1	13	–	–
8. Bihar Plains	12	2	1	11	2	–	10	–	–	13	3	1
9. U.P. East	13	3	1	11	3	2	11	3	–	14	5	3
10. U.P. West	12	4	3	11	2	2	9	1	–	15	4	4
11. Haryana	8	2	–	10	4	–	10	1	–	12	4	–
12. Punjab	5	4	–	8	–	–	2	–	–	13	4	1
13. Himachal Pradesh	9	2	1	6	2	1	6	1	1	12	4	1
14. Jammu & Kashmir	7	3	–	4	–	–	6	1	–	5	1	1
15. Rajasthan W.	6	3	1	11	2	1	10	1	2	10	5	2
16. Rajasthan E.	11	7	3	13	4	1	16	7	4	17	6	3
17. M. P. West	15	10	7	15	6	2	14	5	3	18	5	2
18. M. P. East	15	10	7	14	3	2	14	2	–	18	6	2
19. Gujarat	12	7	3	15	5	2	14	4	3	17	6	1
20. Saurashtra Kutch	14	8	6	16	10	3	16	3	2	15	4	1
21. Konkan	18	11	8	16	7	5	17	9	5	18	13	5
22. Madhya Maharashtra	12	3	1	12	5	4	12	6	2	16	9	6
23. Maratha wada	5	2	–	9	2	1	6	1	–	11	6	4
24. Vidarbha	16	6	5	10	4	3	10	4	1	15	3	1
25. Coastal A. P.	7	1	1	10	2	1	4	1	–	5	1	–
26. Telengana	4	2	1	11	3	3	9	1	1	10	3	2
27. Rayal Seema	4	1	–	8	2	–	10	4	–	12	6	2
28. Tamilnadu	2	1	–	5	–	–	7	1	–	8	3	–
29. Coastal Mysore	16	7	5	15	4	3	18	9	4	15	9	5
30. Int. Mysore N.	5	2	–	12	4	2	9	3	2	9	6	3
31. Int. Mysore S.	3	2	1	12	1	–	8	5	2	11	5	4
32. Kerala	13	4	1	15	3	3	13	7	3	12	5	4
33. Lakshwadeep	8	1	–	16	4	1	14	4	2	7	1	–

occurrence but also of rainfall amounts in these regions. We must however remember that autocorrelation coefficients do not consider the wet or dry periods of a series separately and thus continuance of dry periods will enhance the correlation value. This factor is more important for some subdivisions in June and September where monsoon influence does not last for

Table 5. *Yearwise Autocorrelation upto Lag 3 for Area Weighted EAC for 20 SD's Lying over the Monsoon Trough Zone.* The 20 SD's used are from serial Nos. 4 to 11 and 15 to 26 (for SD's name refer Table 4)

Lag Year	1 July	2 July	3 July	1 Aug.	2 Aug.	3 Aug.	Year	1 July	2 July	3 July	1 Aug.	2 Aug.	3 Aug.
1958	.65	.26	.01	.39	−.20	−.19	1966	.90	.83	.74	.87	.78	.73
1959	.70	.29	.09	.42	−.10	−.04	1967	.80	.55	.39	.37	.00	−.14
1960	.82	.62	.45	.67	.31	.17	1968	.80	.44	.05	.84	.73	.69
1961	.54	.31	.48	.43	.14	.01	1969	.54	.23	.30	.77	.62	.52
1962	.74	.37	.23	.67	.30	.21	1970	.77	.57	.29	.23	.32	−.11
1963	.76	.56	.52	.55	.07	−.03	1971	.61	.23	.04	.78	.54	.33
1964	.52	−.05	−.21	.59	.10	−.04	1972	.92	.82	.75	.81	.50	.14
1965	.77	.41	.22	.88	.70	.57	1973	.55	.36	.37	.33	−.18	−.19

the entire month. Similar autocorrelation coefficients for daily EAC of all subdivisions for all months and years are also computed but not presented as they yield similar results. It may be of some interest to know about the fluctuations in autocorrelation coefficients from year to year. Yearwise autocorrelation-coefficients upto lag 3 for area weighted EAC for 20 SD's lying in the main monsoon trough zone are shown in Table 5. All autocorrelation coefficients for lag 1, barring August 1970, are significant at 5% level. Most of the correlations particularly in July are highly significant, demonstrating high persistence in large scale behaviour of monsoon. In fact correlation as high as 0.92 is obtained for July 1972. The persistence influence lasts upto several days in some years (particularly during dry years) although generally vanishes after 3 days in most of the years.

5. Persistence in 5-Day Rainfall

Mooley and Apparao [14] from a study of 5-day (pentad) rainfall of 11 stations concluded that monsoon rainfall over time scales of 5 and 10-day is not pairwise independent for consequetive periods. Serial correlation coefficients (SCC) between pentad rainfall, for 1 to 5 pentad lags have been computed for each monsoon month (June through September) by pooling 19 years of data. Very high pentad rainfall values representing noise have been truncated as done in the case of daily rainfall amounts. The value of SCC for lag of 1 pentad only have been shown in Table 6, as for other lags the SCC's are not significantly different from what could be obtained by chance. It may be seen that out of 33 SD's 23, 22, 27 and 19 demonstrate significant SCC for June, July, August and September respectively. Most of the signifcant SCC's pertain to contiguous SD's lying under the influence

Table 6. *Values of Serial Correlation Coefficients in Pentad Rainfall for Lag 1* (for SD's name refer Table 4)

S.D. No.	1	2	3	4	5	6	7	8	9	10	11
June	.25	.03	.03	.14	.29	.25	.32	.25	.22	.40	.20
July	.13	−.02	.08	.05	.02	.10	.02	.05	.21	.18	.24
Aug.	.27	.25	.16	.35	−.01	.13	.42	.09	.25	.18	.07
Sept.	.19	.13	.15	.17	.04	.10	.03	.06	.18	.11	.11
S.D. No.	12	13	14	15	16	17	18	19	20	21	22
June	.17	.38	.08	.19	.25	.51	.40	.29	.22	.40	.25
July	.08	.21	.02	.33	.23	.27	.33	.14	.27	.20	.36
Aug.	.10	.19	.16	.29	.17	.16	.19	.33	.39	.57	.31
Sept.	.25	.19	.25	.32	.25	.35	.04	.23	.32	.25	.23
S.D. No.	23	24	25	26	27	28	29	30	31	32	33
June	.27	.39	.27	.28	.08	.14	.26	.11	.05	.03	.09
July	.25	.36	.18	.29	.31	.22	.17	.41	.21	.37	.31
Aug.	.45	.18	.33	.41	.52	.34	.28	.31	.41	.16	.14
Sept.	.11	.10	.16	.12	.09	.09	.39	.26	.41	.41	.26

of main monsoonal trough/synoptic systems outside north east hilly regions and northern and southern most parts of the country. By squaring the SCC's we find that for most cases as shown in Table 6 only 4 to 16% of variance of pentad rainfall is explained by previous pentad rainfall and thus there is not strong reason to support that reasonable predictions can be made by considering pentad rainfall as autoregressive process. However, the variances explained appear to be reasonable for including the preceding rainfall as a predictor in statistical methods concerned with prediction of 5-day rainfall employing classical predictor-predictand relationships.

For most of the applications in reality, it may not be necessary to know the exact amount of precipitation. It may rather be sufficient to know the general character of rainfall during a 5-day period. Singh et al. [18] studied the persistence in 5-day subdivisional rainfall by preparing 3 × 3 contingency tables between 3 equally probable classes (terciles) of rainfall. They presented for each of the summer monsoon month (averaged over all 33 SD's) the percentage of occasions of a particular character (tercile) of pentad rainfall being followed by similar character of pentad rainfall. These aggregate figures do not show significant skill above chance where prediction can be treated as a binomial process with 1/3 probability of success and 2/3 probability of failure in each trial. In view of the geographically unequal distribution of persistence over the country, it may be worthwhile to examine the persistance in character of pentad rainfall for individual SD's. For this purpose 19 years (1955–73) of rainfall data is used to find terciles of pentad rainfall distribution in each month. Contingency table (3 × 3) between successive pentad rainfall characters have been prepared and Heidke-

Table 7. *Heidke-Skill Scores for Regions Showing High Persistence in 5-Day PPN Amount*
Significant and highly significant values are marked by one and two asterisks respectively (for SD's name refer Table 4)

S.D. No.	15	16	17	18	19	20	21	22	23
June	$.17^{+}$	$.20^{++}$	$.21^{++}$	$.23^{++}$	.09	$.23^{++}$	.09	$.21^{++}$	$.20^{++}$
July	$.23^{++}$	$.19^{++}$	.07	.11	.10	.09	.08	$.16^{+}$	$.18^{+}$
Aug.	–.01	.10	–.04	$.21^{++}$	$.19^{++}$	.11	$.34^{++}$	$.16^{+}$	$.26^{++}$
Sept.	$.31^{++}$	$.18^{+}$	$.14^{+}$	$.15^{+}$	.13	$.23^{++}$	.07	$.18^{+}$	$.21^{++}$
S.D. No.	24	25	26	27	28	29	30	31	32
June	$.21^{++}$	.07	$.21^{++}$	.10	–.03	.07	.00	–.12	–.05
July	$.26^{++}$	.08	.12	.12	.10	–.06	$.22^{++}$	.13	$.18^{++}$
Aug.	–.01	.23	.18	.31	.13	$.18^{+}$	$.23^{++}$	$.23^{++}$	.09
Sept.	.05	.13	$.16^{+}$	.05	.13	$.16^{+}$	.08	$.19^{++}$	$.18^{+}$

skill scores (SS) have been evaluated. The SS's for the geographical regions which generally show high persistence have been presented in Table 7. Significant SS are also obtained for SD's 6, 7, 9 and 10 in June and SD's 11, 12, 13 and 14 in September but they appear to arise due to late arrival of monsoon in June and early withdrawal in September for these SD's. Most of the SD's in Table 7 show positive SS and many of them exceed 0.2, suggesting use of persistence as unskilled standard in evaluating skill of forecast techniques, 5-day EAC show similar results.

6. Discussion and Conclusions

Several characteristics of rainfall can be computed mathematically from the parameters obtained by fitting Markov chain models (e. g. [6, 7, 20]). Some of the reasons why this has not been attempted will be described below. Fitting models on data for entire season ignores the vital synoptic information. Diametrically opposite synoptic processes operate during the heart of the monsoon season causing extreme dry or wet weather over the country during their life cycles. Persistence may change its character from one synoptic process to another synoptic process. Hopkins and Robillard [11] point out the difficulty of treating the PPN arising due to different synoptic processes together and recommend the use of negative binomial compound poisson distribution as a remedy in place of Markov chain. From above it is clear that as far as forecasting is concerned optimum skill from fitting the models cannot be obtained unless at least the conditional periods are stratified according to certain large scale synoptic processes. Such stratification however demands larger data. Todorovic and Woolhiser [21] also raise doubt about physical explanation of various features of rainfall without considering climatological factors. Clearly much is left to be done if these and other new ideas presented here are to be proved useful in precipitation forecasting.

Acknowledgements

The authors are grateful to the Director, Indian Institute of Tropical Meteorology, Pune for providing facilities to pursue the work and to Shri D. R. Sikka for constant encouragement. Thanks are due to Mrs. Mane for typing the manuscript.

References

1. Basu, A. N.: Fitting of a Markov Chain Model for Daily Rainfall. Indian J. Met. Geoph. *22*, 67–74 (1971).
2. Blairfish, J. A.: An Investigation into Spells of Wet and Dry Days by Region and Season for Great Britain. Met. Mag. *104*, 360–375 (1975).
3. Brooks, C. E. P., Carruthers, N.: Handbook of Statistical Methods in Meteorology, 413 pp. London: Her Majesty's Stationery Office 1953.
4. Caskey, J. E.: Markov Chain Model of Cold Spells at London. Met. Mag. *93*, 136–138 (1964).
5. Chaudhary, A., Gokhale, S. S., Rentale, G. S.: Spells of Dry Days Related to Agricultural Drought in India. Mausam *30*, 501–510 (1979).
6. Eriksson, B.: A Climatological Study of Persistency and Probability of Precipitation in Sweden. Tellus *17*, 484–497 (1965).
7. Gabriel, K. R., Neumann, J.: A Markov Chain Model for Daily Rainfall Occurrence at Tel Aviv. Quart. J. R. Met. Soc. *88*, 90–95 (1962).
8. Gabriel, K. R., Neumann, J.: On the Distribution of Weather Cycles by Length. Quart. J. R. Met. Soc. *83*, 375–379 (1957).
9. Green, J. R.: A Generalised Probability Model for Sequences of Wet and Dry Days. Mon. Weath. Rev. *98*, 238–241 (1970).
10. Green, J. R.: A Model for Rainfall Occurrence. J. R. Statist. Soc. B *26*, 345–353 (1964).
11. Hopkins, J. W., Robillard, P.: Some Statistics of Daily Rainfall Occurrence for the Canadian Prairie Provinces. J. Appl. Met. *3*, 600–602 (1964).

12. Jorgensen, D. L., Klein, W. H.: Persistence of Precipitation at 108 Cities in the Conterminous United States. ESSA Tech. Memo. WBTM TDL 31, 64 pp. (1970).
13. Lawrence, E. N.: Application of Mathematical Series to the Frequency of Weather Spells. Met. Mag. *83*, 195–200 (1954).
14. Mooley, D. A., Appa Rao G.: Independence of Rainfall over India over Time Scales of Less than a Month during Monsoon Season. IITM Sc. Rep. 129, 11 pp. (1970).
15. Sakamoto, C. M.: Markov Chain Models for Probabilities of Hot and Cool Day Sequences and Hot Spells in Nevada. ESSA Tech. Report EDS 9, 26 pp. (1970).
16. Singh, S. V., Kamte, P. P., Dahale, S. D., Upasani, P. U.: Rainfall Persistency over India during Summer Monsoon. Proc. Symp. on Tropical Monsoons, Indian Inst. of Trop. Met., 418–425 (1976).
17. Singh, S. V., Prasad, K. D.: Periodicities in Daily Summer Monsoon Rainfall of India. Proc. Symp. on Tropical Monsoons, Indian Inst. of Trop. Met., 427–435 (1976).
18. Singh, S. V., Prasad, K. D., Mooley, D. A.: Climatology of Five-Day Upper Level Summer Monsoon Flow over India and Its Relationship with Rainfall. Rivista Ital. Geofis. E Science Affini *IV*, 233–240 (1977).
19. Srinivasan, T. R.: Rainfall Persistence in India during May–October. Indian J. Met. Geofis. e Scienze Affini *IV*, 233–240 (1977).
20. Sundarraj, N., Ramachandra, S.: Markov Dependent Geometric Models for Weather Spells and Weather Cycles. Indian J. Met. Geoph. *26*, 221–226 (1975).
21. Todorovic, P., Woolhiser, D. A.: A Stochastic Model for *n*-Day Precipitation. J. Appl. Met. *14*, 17–24 (1975).
22. Victor, U. S., Sastry, P. S. N.: Dry Spell Probability of Markov Chain Model and Its Application to Crop Development Stages. Mausam *30*, 479–484 (1979).
23. Weiss, L. L.: Sequences of Wet or Dry Days as Described by a Markov Chain Probability Model. Mon. Weath. Rev. *92*, 169–176 (1964).
24. Williams, C. B.: The Log Series and Its Application to Biological Problems. J. Ecology *34*, 253–272 (1947).

Authors' address: Dr. S. V. Singh, R. H. Kripalani, Mrs. Priya Shaha, P. M. M. Ismail, and S. D. Dahale, Indian Institute of Tropical Meteorology, Ramdurg House, University Road, Pune-411 005, India.