Space-time variation and regionalization of seasonal and monthly summer monsoon rainfall of the sub-Himalayan region and Gangetic plains of India

K. K. Singh[1,*], S. V. Singh[2]

[1]Indian Institute of Tropical Meteorology, Pune-411008, India
[2]National Centre for Medium Range Weather Forecasting, Mausam Bhavan, Lodhi Road, New Delhi-110003, India

ABSTRACT: The present study analyses space-time variation of the seasonal and monthly monsoon rainfall of the sub-Himalayan region and Gangetic plains of India by using rainfall data of 90 well-distributed stations for a period of 114 yr (1871 to 1984). Principal components (PCs) and varimax-rotated principal components (RPCs) are computed from the interstation similarity matrix. The 4 leading PCs explaining 52 to 61 % of the variance respectively for seasonal and monthly rainfall are found to be statistically significant. When rotated by the varimax method these 4 PCs provide 4 distinct areas of homogeneous rainfall variability for both time scales. The most coherent region is located along the central longitude of the study area, with another coherent region over the extreme western parts. The area east of 81° E is divided into 2 coherent zones roughly along the 22° N parallel. It is noted that the regions thus identified will be useful in understanding climate variability and in weather prediction research.

KEY WORDS: Sub-Himalayan region and Gangetic plains · Principal components (PCs) · Rotated principal components (RPCs) · Spatial coherence · Monsoon rainfall regionalization

1. INTRODUCTION

The summer monsoon is of considerable scientific and economic interest for India as it contributes nearly 75 % of the annual rainfall. Rainfall is highly variable in space and time as it is the end product of complex interactions between a variety of dynamic processes with characteristic spatial and temporal scales. Different dynamic processes or different states of the same dynamic process yield different patterns of coherent precipitation. In order to understand rainfall variability over a region, it is necessary to identify these various patterns of spatio-temporal variability by some objective procedure. Principal component analysis (PCA), a multivariate statistical procedure, has commonly been used to identify the dominant patterns of spatio-tempo-ral variability in climate data. This procedure not only identifies the important patterns of spatio-temporal variability but also determines the variance explained by them.

PCA has been employed by several workers in the past to study the important modes of variability of annual (Hastenrath & Rosen 1983), seasonal (Shukla 1987, Kulkarni et al. 1992) and 5 d (Kripalani et al. 1991) rainfall over India. Shukla (1987) considered PCA as one of the methods for identifying the homogeneous regions of seasonal monsoon rainfall variability over India. For this purpose he considered the data of 31 meteorological subdivisions which covered the whole of India. Many researchers (e.g. Richman 1986, Legates 1991) have noted that the results of PCA depend on the size and shape of the study area and that the coherent regions identified by using a larger area may not contain adequate regional details. For example, Iyengar (1991) studied the rainfall over Kar-

*E-mail: frdmail@tropmet.ernet.in

nataka, a state in the southern peninsula of India, using PCA and pointed out certain regional details like the juxtaposition of the rainfall over the west coast and the dry region to the east. Such features could not be identified by previous researchers who used the rainfall over the whole country (e.g. Bedi & Bindra 1980, Shukla 1987). Further, recent research has shown that PCA by itself may not provide complete details about coherent regions of variability and that, hence, a rotation of the principal components (PCs) is desirable (e.g. Richman 1986). For the study of coherent regions of variability one may use station data or data averaged over a smaller area. Richman & Lamb (1985) have suggested that uniformly distributed station rainfall data be analysed in order to provide better details of spatial coherence. Singh (1993) recently conducted extensive eigenvectorial analysis of rainfall at 306 stations covering the whole of contiguous India and identified a few broad regions of coherent rainfall variability using rotated PCA. One of the important regions which he identified covers the sub-Himalayan region and the Gangetic plains in the north of India. This area is thickly populated and has great food production potential due to rich alluvial soils deposited by rivers such as the Ganges and its tributaries.

In the present paper our main interest is to identify the principal modes of variability of seasonal and monthly rainfall over this region and to arrive at a regionalization using rotated PCA. This will enable us to identify smaller regions of coherent variability other than those obtainable by analysing rainfall over the whole of India. Such smaller regions will be more useful for regional planning. In Section 2 below we describe the data and climatic features of the region and in Section 3 important computational aspects of the method are examined. The results are discussed in Section 4 and the important conclusions are set out in Section 5.

2. DATA AND CLIMATIC FEATURES

The study area, situated roughly between 22–29° N and 75–89° E, covers 5 states in the Gangetic plains in the north of India (Fig. 1). These states are (from west to east): Haryana, Delhi, Uttar Pradesh (excluding the hilly regions), Bihar, and West Bengal. Though the great Himalayas lie to the north of this region, the region itself is free from major topography. The mountainous northwestern region of Uttar Pradesh has been omitted from consideration. However, part of southern Bihar is covered by the Chotta Nagpur plateau, which has an average height of about 300 to 600 m. The height of a small region on this plateau containing 2 stations lies between 600 and 900 m (shaded area in

Fig. 1). The height of only 1 station, namely Darjeeling, in northern West Bengal exceeds 2000 m (shown as a solid square in Fig. 1). Monthly rainfall data for June through September collected over a period of 114 yr (1871 to 1984) at 90 stations well distributed over the study area were utilized in this study. The data are of consistently good quality as they originate mainly from the observatories manned by the India Meteorological Department. They have been processed and checked by B. Parthasarathy (see Parthasarathy et al. 1987). A few missing values were filled in by these authors using the ratio method after Rainbird (1967).

Based on this data we have computed some climatic statistics such as mean standard deviation and coefficient of variation for seasonal and monthly rainfall. The mean for the whole analysed period of monsoon season rainfall over the study area is shown in Fig. 2a. The rainfall amount over the study area varies from less than 40 cm in the western semi-arid region to over 200 cm over the humid region in the extreme east. Mean rainfall is around 40 cm over Haryana and Delhi, 70 to 100 cm over Uttar Pradesh and more than 100 cm over Bihar and West Bengal. Rainfall exceeds 100 cm over the extreme northern part of sub-Himalayan Uttar Pradesh. Over the extremely hilly northeastern part of West Bengal, in the neighbouring district of Darjeeling, Jalpaiguri and Cooch-Bihar, it is more than 200 cm. The northern part of the entire study area, i.e. the sub-Himalayan region, receives higher amounts of rainfall during weak or break monsoon conditions over central India. Though a major part of the rain over the study area is due to the monsoonal
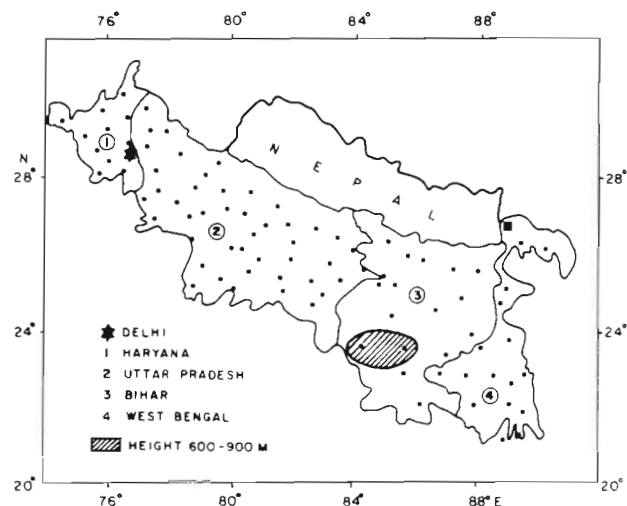

Fig. 1. Network of 90 rainfall stations in the sub-Himalayan region and on the Gangetic plains distributed over the states of Delhi, Haryana, Uttar Pradesh, Bihar and West Bengal. 114 yr (1871–1984) of rainfall data from these stations are used in this study. Shaded area is between 600 and 900 m high. (■) Darjeeling, which has a height of 2128 m

systems, the weather systems in the extra-tropical westerlies continue to influence the monsoon rainfall, albeit with reduced frequency and intensity.

The standard deviation shown in Fig. 2b is, in general, higher over the areas of high rainfall and low over the areas of low rainfall. It is higher than 50 cm over the extreme northeastern part and is lower than 10 cm over the extreme west, from where it gradually begins to increase to more than 20 cm towards the east. The coefficient of variation shown in Fig. 2c (standard deviation/mean × 100) is lowest over northern West Bengal (decreasing to the south and increasing to the west), and is 50 % over Haryana.

The spatial distribution of monthly rainfall follows the same pattern as seasonal rainfall (data not presented). The mean monthly rainfall for June, July, August and September for the whole period analysed is 60 cm over northeastern West Bengal, about 30 cm along the foothills of the Himalayas, West Bengal and most of Bihar, about 20 cm over the plains of Uttar Pradesh and decreases to less than 10 cm in western Uttar Pradesh and Haryana. The standard deviation is high (30 cm) near the foothills of the Himalayas, decreasing to less than 20 cm over most of the other parts of the study area. Some increase in standard deviation over the southern part of Uttar Pradesh is noted. The coeffient of variability is 100 % over Haryana, western Uttar Pradesh and Bihar and decreases to 50 % in the eastern and northeastern parts of West Bengal.

3. METHODS

The PCA method adopted here is too well known to require a detailed description. Therefore, we describe only those conceptual aspects of the methodology which are relevant to the objectives of the study, viz. the PC loadings, PC scores, selection of PCs and their rotation.

The PCs are the eigenvectors of the dispersion matrices (correlation or covariance matrices) which are symmetric and positive definite. These eigenvectors represent the orthogonal and the optimal modes (i.e. explaining maximum variance) of spatio-temporal variability when the elements of the matrix represent correlations/covariances between time series at 2 grid points or stations. The resultant eigenvectors or PCs can be displayed in space. This analysis procedure is called spatial (S-mode) PCA and is commonly employed in climate research. Alternatively, when the elements of the correlation matix represent correlations/covariances between 2 spatial patterns corresponding to 2 observation times, the resultant eigenvectors can be displayed in time as time series. Here,
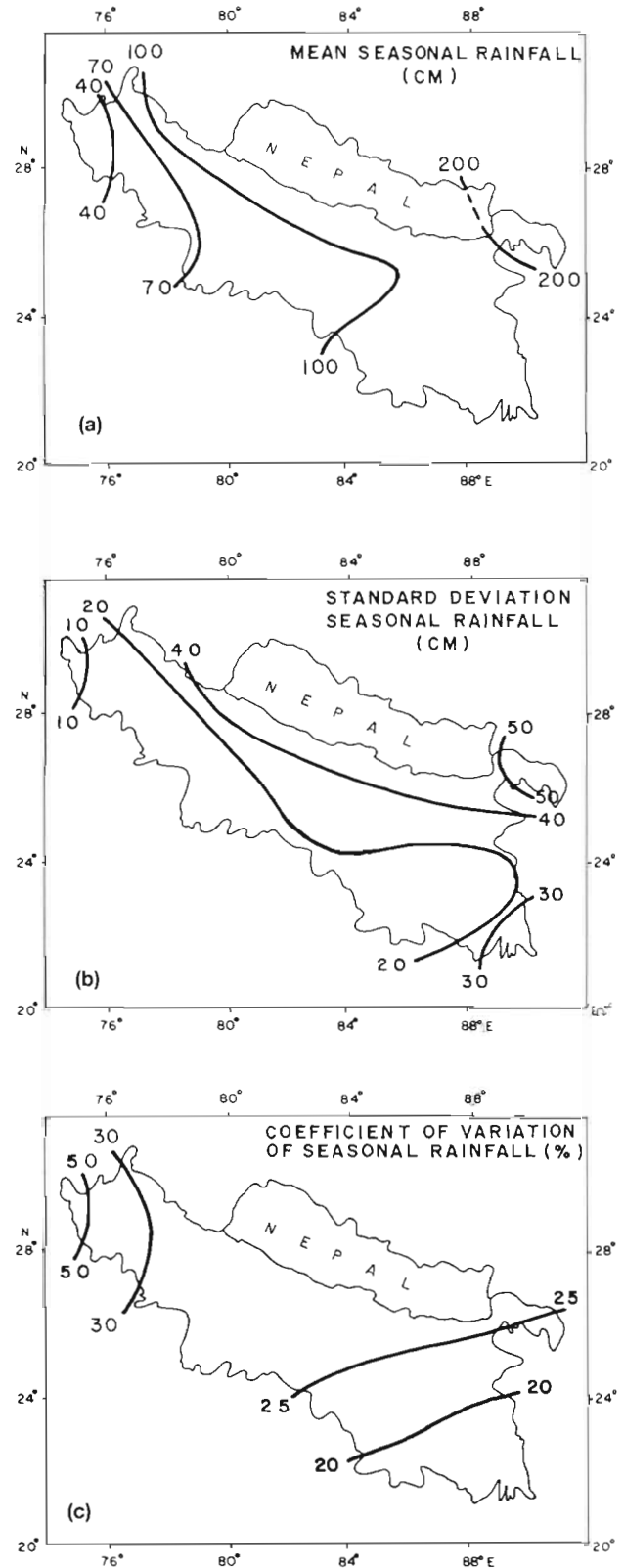

Fig. 2. (a) Mean, (b) standard deviation and (c) coefficient of variation of seasonal rainfall in the study area

we have adopted the former, namely the S-mode, type of analysis, as our objective is to identify the regions of coherent variability in space. The projection of the spatial fields of original variables at a particular time onto an eigenvector of this analysis provides a PC score for that PC for that particular time. When an element of the eigenvector is multiplied by the square root of its eigenvalue, we obtain a quantity called loading, which represents the correlation between the time series of the original variable observed at the corresponding location and the corresponding PC scores. In the present study the rainfall of 90 stations represents 90 variables and the number of observations is 114 for seasonal and 456 for monthly rainfall. The elements of the required correlation matrix represent correlations between rainfall series from the stations taken in pairs.

The PCs are ordered according to the variance explained. The first PC explains the highest variance and the subsequent PCs explain variance in decreasing order. Often the first few PCs portray the large-scale features with the same loading sign over a large contiguous area. The higher order PCs showing scattered and disorganised loadings represent mainly noise. A number of statistical significance tests to identify the important and significant PCs have been suggested and followed (Preisendorfer & Barnett 1977, North et al. 1982, Overland & Preisendorfer 1982). The test after North et al. (1982) is based on the separability of adjacent ordered eigenvalues. According to this test, a pair of PCs are unseparable when their separation lies close to the sampling error between the 2 eigenvalues; such PCs are believed to form a 'degenerate multiplet'. The PC loadings of such a multiplet should be either considered together for further analysis (e.g. rotation) or should be dropped from consideration. In the present case, the first 4 PCs were found to be separable from each other but PCs 5 and 6 were not separable and represented a degenerate multiplet. These were dropped from further analysis such as rotation described below.

Doubts have been raised about the utility of the PCs in identifying the natural modes of variability contained in the dispersion matrices (see e.g. Richman 1986). The PCs determined by the above procedure are viewed as data compression techniques that seek to maximize the variance over the study area and hence may not necessarily focus on the dominant intrinsic modes of climate variability. Under such constraints, a large number of data points may show comparable loadings on 1 PC potentially leading to improper delineation of regions of strong spatial coherence. The resultant patterns may be too smooth and too large scale. The condition of maximum variance explained for the entire area could cause distinctive

coherent rainfall areas to merge. Recent research has suggested that this difficulty can be overcome by PC rotation (e.g. Richman & Lamb 1985, 1987, Richman 1986, Pandzic 1988). Rotation tends to orient the loadings such that they are higher over smaller and more coherent regions and lower (near zero values) over the remaining parts of the area.

A good number of rotation procedures, orthogonal or oblique, have been suggested, but orthogonal varimax-rotation (Kaiser 1959) has been more commonly used and has provided satisfactory results in most climatic studies. Therefore, we have adopted this rotation procedure in the present study. It may be noted that the outcome of the rotation may depend to some extent on the number of PCs considered. We have considered the first 4 PCs for rotation which were found to be significant by the significance tests.

4. RESULTS

4.1. Principal component analysis (PCA)

PCs were determined for seasonal as well as for monthly rainfall. While analysing the monthly rainfall, anomalies were considered which remove the seasonal cycle. Three types of dispersion matrices, viz. correlation, covariance and sum of cross-products, were subjected to PCA and were found to give similar spatial patterns. Therefore, we have presented the results of correlation-based analysis only, as the loadings from this analysis are directly interpretable in terms of correlation between the original variable and the PC scores.

4.1.1. Seasonal rainfall

The variances explained by the first 4 unrotated PCs, which were found to be significant by the tests mentioned in Section 3 above, are given in Table 1. These 4 PCs, which account for a total of 52% of the spatio-temporal variance, are displayed in Fig. 3. The first PC accounts for 25.9% of the total variance of the data. This component had positive loadings throughout the study area, and loadings exceeding 0.4 cover the whole of Uttar Pradesh, most of Bihar and parts of Haryana (Fig. 3a). The area of maximum coherence (loadings exceeding 0.7) extends from the foothills of the Himalayas to the Gangetic plains of central Uttar Pradesh in a north-south direction. The Gangetic plains of Uttar Pradesh and Bihar are the regions of dominant spatial coherence on this PC.

The second PC accounts for 13% of the total variance of the data. This PC shows a dipole-type structure

Table 1. Eigenvalues and percent variance explained (VE) by 4 leading principal components for (a) seasonal rainfall and for (b) monthly rainfall anomalies (June–September). Number of observations 114 from 114 yr (1871–1984) for (a) and 456 from 114 yr (1871–1984) for (b)

| Root no.	Unrotated eigenvalue	Unrotated (VE)	Cumulative (VE)	Varimax rotated eigenvalue	Varimax rotated (VE)	Cumulative (VE)
(a) Seasonal rainfall						
1	23.29	25.87	–	16.42	18.25	–
2	11.77	13.07	38.94	12.84	14.28	32.53
3	7.49	8.32	47.26	9.52	10.58	43.11
4	4.25	4.73	51.99	8.00	8.89	52.00
(b) Monthly rainfall anomalies						
1	35.23	39.15	–	19.35	21.51	–
2	9.32	10.36	49.51	16.08	17.87	39.38
3	6.36	7.07	56.58	10.20	11.34	50.72
4	3.63	4.03	60.61	8.92	9.91	60.63
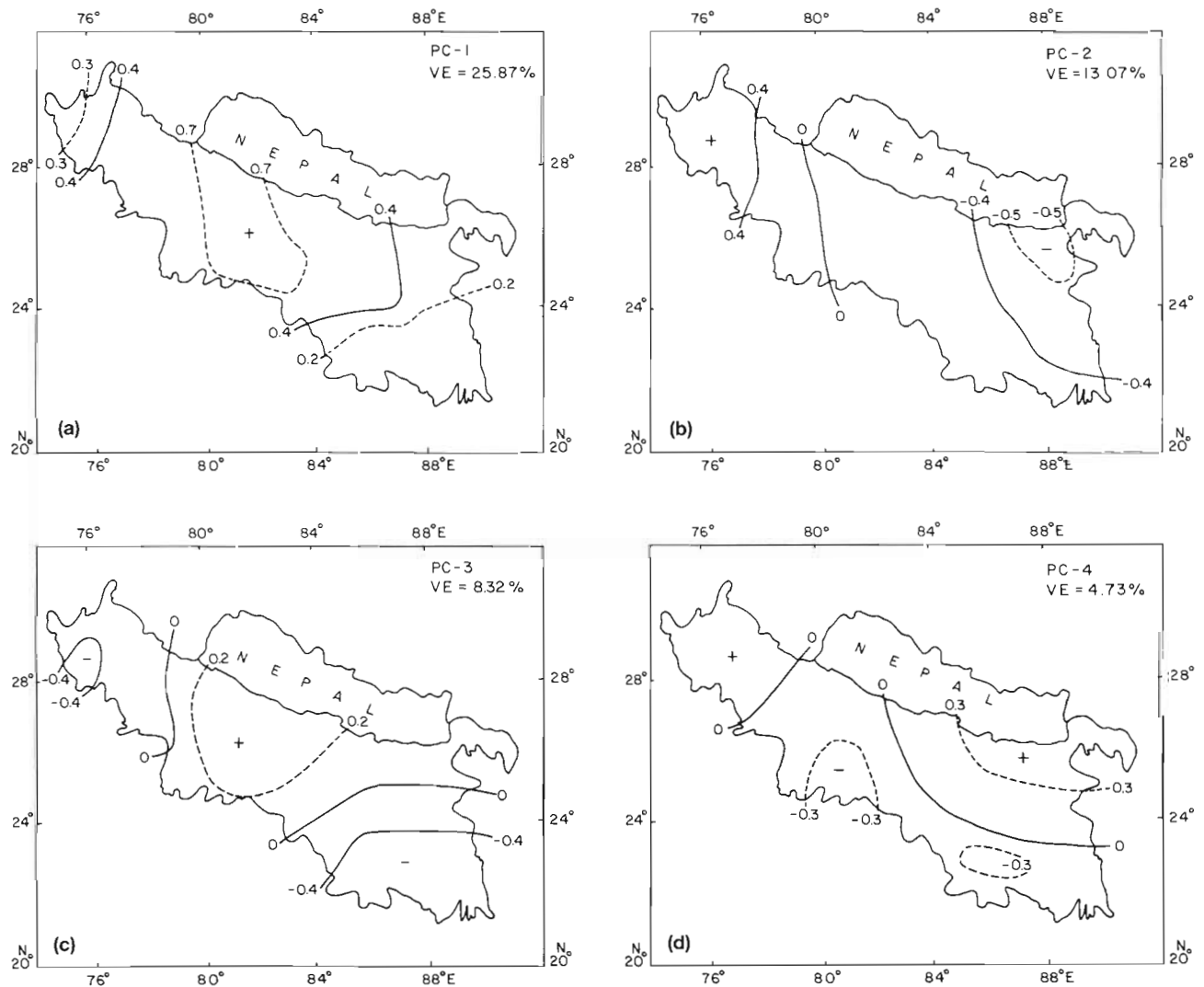
Fig. 3. Unrotated PCs 1 to 4 of the interstation correlation matrix of seasonal rainfall in the study area. VE: variance explained; +: positive loadings; –: negative loadings

with negative anomalies over the northeastern parts of Bihar and West Bengal and positive anomalies over western Uttar Pradesh and Haryana, which can be viewed as the result of seesaw rainfall activity over the 2 regions (Fig. 3b). Such an opposite relationship between the extreme western and eastern parts of the country is observed even when the data for the whole country are analysed (see for example Kulkarni et al. 1992).

The third PC, explaining 8.3% of the data variance, presents a 3-fold structure of anomalies: a negative anomaly in western Uttar Pradesh and Haryana, positive anomalies along the foothills of the Himalayas and plains of central Uttar Pradesh, and a negative anomaly over most of southern Bihar and West Bengal (Fig. 3c). Amongst these, the strongest anomalies occur over the last area, which is located close to the head of the Bay of Bengal. The general structure of this component is more complex than the first 2 PCs.

The fourth PC, which explains 4.7% of the variance, shows mainly a north-south bifurcation of loading signs over the central and eastern part of the analysis area, in contrast to the east-west partitioning of loadings in the first 3 PCs (Fig. 3d). The negative loadings here are tilted in a northwest to southeast direction akin to the surface position of the monsoon trough. The extreme western part does not show any significant loadings. Both PC 3 and PC 4 show a somewhat complex structure. The higher order components discarded by significance tests are even more complex and are therefore not considered for discussion.

The first 4 PCs, which explain a total of 52% of the data variance, were rotated by the orthogonal varimax method. The rotated patterns are shown in Fig. 4, and the variance explained by these patterns is shown in Table 1a. Note that although the total variance remains the same, there is a more uniform distribution of vari-
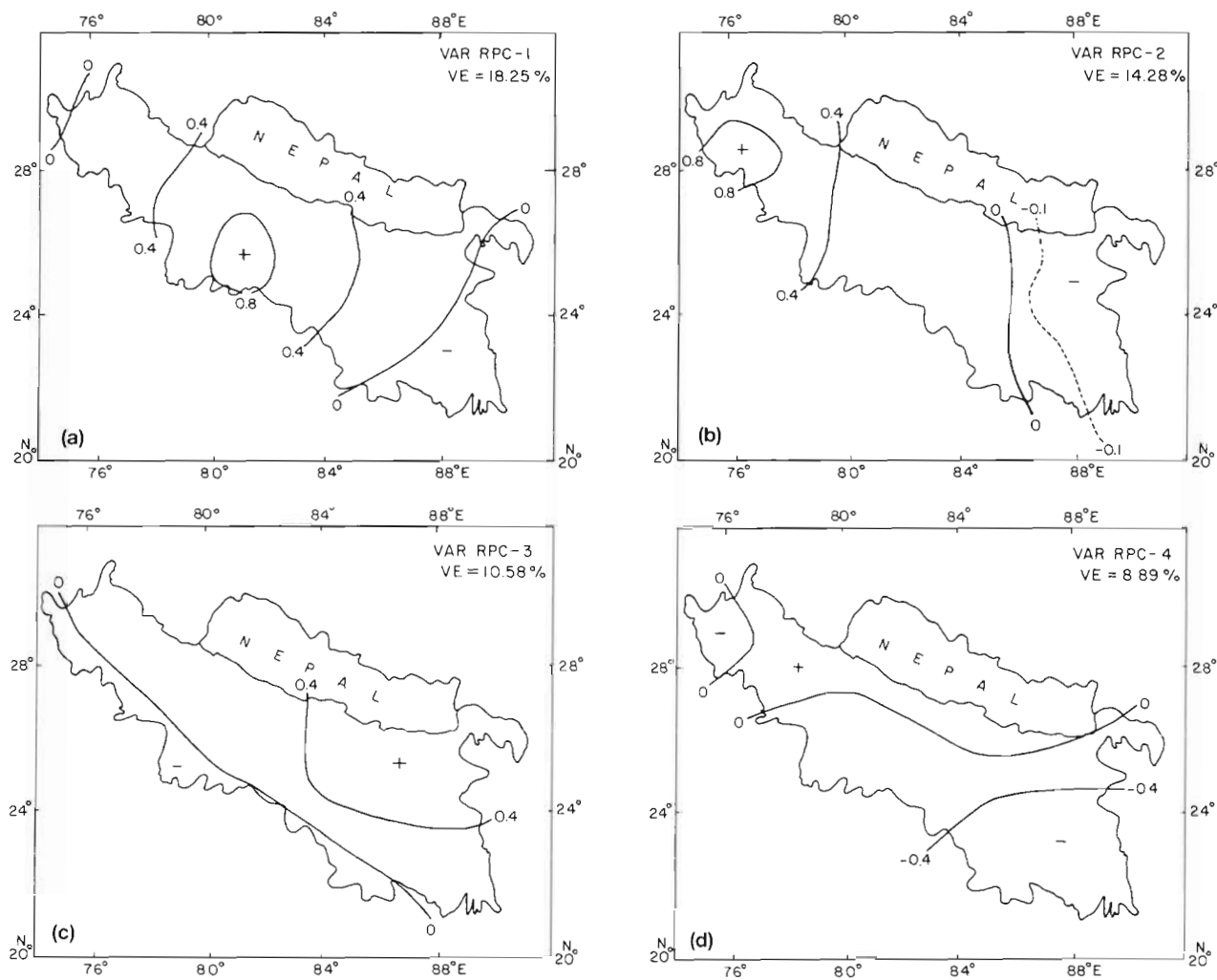

Fig. 4. Varimax-rotated PCs 1 to 4 of seasonal rainfall in the study area

ance amongst the rotated patterns. The most important pattern now explains 18.25 % of the original data variance, and the loading structure is somewhat similar to the first unrotated PC. But, after rotation, the high loadings over the central part of the study area have become more concentrated. The highest loading now exceeds 0.8. Thus, both rotated and unrotated PCs suggest that the most important and coherent mode of variability of monsoon rainfall over the Gangetic plains lies along a longitude of about 82° E, which may be weakly correlated with All India Summer Monsoon Rainfall (AISMR). Recently, B. Parthasarathy (pers. comm.) determined a correlation between AISMR and rainfall measured at different stations. He found that, whereas the seasonal rainfall over Haryana and western Uttar Pradesh shows a correlation with the AISMR ranging between 0.4 and 0.5, the eastern part of the present study area shows a correlation between 0.4 and 0.2, the correlation for Bihar and West Bengal being generally less than 0.2. This supports our belief that the coherent region found on the first PCs may not be linked with the coherent area of variability found over the extreme western part of the country identified by analysing total Indian rainfall (see e.g. Shukla 1987).

The second rotated PC, which explains 14 % of the variance, has high loadings over the western part of the area analysed, i.e. Haryana and western Uttar Pradesh (correlation with AISMR: 0.4 to 0.5), where the second unrotated PC also showed a coherent structure of strong loadings but with a coexisting opposite loading sign over the eastern part of the area. The stronger coherence over the western part of the area can be attributed to mid-latitude disturbances commonly known as western disturbances. In a season, 3 to 6 western disturbances, commonly in the form of low pressure areas migrating west to east, can influence the country and account for high precipitation in the northwestern part of the country. This region is under the continued influence of disturbances in the mid-latitude westerlies during the summer monsoon season (Rao 1976), the second important source of coherent variability. Coherence over the eastern region of the study area on the unrotated PC 2 is due to monsoonal currents and disturbances originating in the Bay of Bengal. After rotation this coherence is diluted due to redistribution of variance. The effect of western disturbances, which appeared as a dipole between the eastern and western part of the domain on the unrotated PC 2, has been clearly isolated by rotation. For the states of Haryana and western Uttar Pradesh this component provides an important source of coherent rainfall variability.

The third rotated component, which explains 11 % of the variance, also shows a simplified structure as compared to its counterpart unrotated PC 3, with high loadings oriented in an almost east-west direction and located over the eastern part of the domain. A comparison of the second and third rotated PC suggests that the western and the eastern parts of the study area witness alternatively dry or wet rainfall conditions. This east-west bifurcation of coherent precipitation was also observed on the unrotated PC 2.

The fourth rotated component, with about 9 % variance, also presents a simplified loading pattern remarkably distinct from its unrotated PC 4 counterpart, with higher loadings over the southeastern part of the study area. The high loadings over this region were noted in the third unrotated PC along with weaker positive loadings over the central and western part of the study area respectively. This PC, after rotation, orients the loadings into an organized pattern of coherence over the region influenced by monsoon disturbances (monsoon lows and depressions) originating from the Bay of Bengal and causing precipitation over the adjoining area. It is important to note here that the actual rainfall anomaly charts may not be classified into these broad categories due to the non-separability of weather patterns in a spatial domain. Still, rotated PC analysis is able to reveal the relative coherent and internal structures of rainfall variability over the various subregions of the study area.

4.1.2. Monthly rainfall

Summer monsoon rainfall (June through September) has been separately subjected to PCA in order to study the spatial patterns of variability on a monthly time scale. For computation of PCs we have considered the interstation correlations arising from rainfall anomalies for June, July, August and September pooled together for 114 yr. This enhances the observation-to-variable ratio, an important factor for the stability of the patterns. Again, the first 4 PCs were found to be significant as in the case of seasonal rainfall. The variances explained by these 4 PCs are given in Table 1b. It is interesting to note that these 4 PCs together explain about 10 % more variance than the seasonal PCs. The major part of this excess variance is contributed by the dominant first PC, which explains 39.15 % of the variance, 13 % more than the corresponding PC 1 for seasonal rainfall. This gain in variance arises obviously due to an increase in intercorrelations between stations' monthly rainfall as compared to seasonal rainfall. This is somewhat surprising, as normally one would expect higher intercorrelations for seasonal rainfall. The higher correlations for monthly rainfall obtained in the present study would suggest smoother and better organised rainfall

anomaly patterns. Although the basic structure of the
first PC is similar to that for seasonal rainfall, the load-
ings are now better organized and, in addition, load-
ings of comparable magnitude occupy a somewhat
larger area (see Fig. 5). The high loading area is now
oriented almost parallel to the surface position of the
monsoon trough in contrast to the north-south orienta-
tion in the case of seasonal rainfall. The second, third
and fourth PCs of monthly rainfall show greater east-
west orientation of loadings over the sub-Himalayan
region and Gangetic plains. Rotation has, for monthly
rainfall as well, produced a simpler structure with
high loadings concentrated over selected parts of the
study area (Fig. 6). The spatial structure of the rotated
PCs of monthly rainfall is similar to that of seasonal
rainfall, except that the third and fourth PCs are inter-
changed in terms of their contribution towards vari-
ance explained. After rotation, the excessive variance
of PC 1 is distributed over the higher order compo-
nents.

4.2. Pairwise plot of PC loadings

It has been noted above that the individual stations
have higher loadings on more than 1 unrotated PC
whereas rotated PCA produced higher loadings for a
group of stations on only 1 rotated PC. This simplifica-
tion of the structure is the primary goal of rotation. In
the first case, the simplified structure was noted by the
spatial plots of the loadings. Alternatively, the simple
structure arrived at by rotated PCA can be examined
with the help of scatter plots of loadings between pairs
of unrotated or rotated PCs. This display technique is
often used in eigenvectorial analysis (e.g. Richman &
Lamb 1985, 1987, Pandzic 1988) to assess the simple
structure obtained by the rotation of PCs. An unrotated
PC plot usually shows a broad scatter of loadings in all
4 quadrants. The data points are not aligned along a
particular axis. Such a structure indicates that the indi-
vidual stations have higher loadings on more than
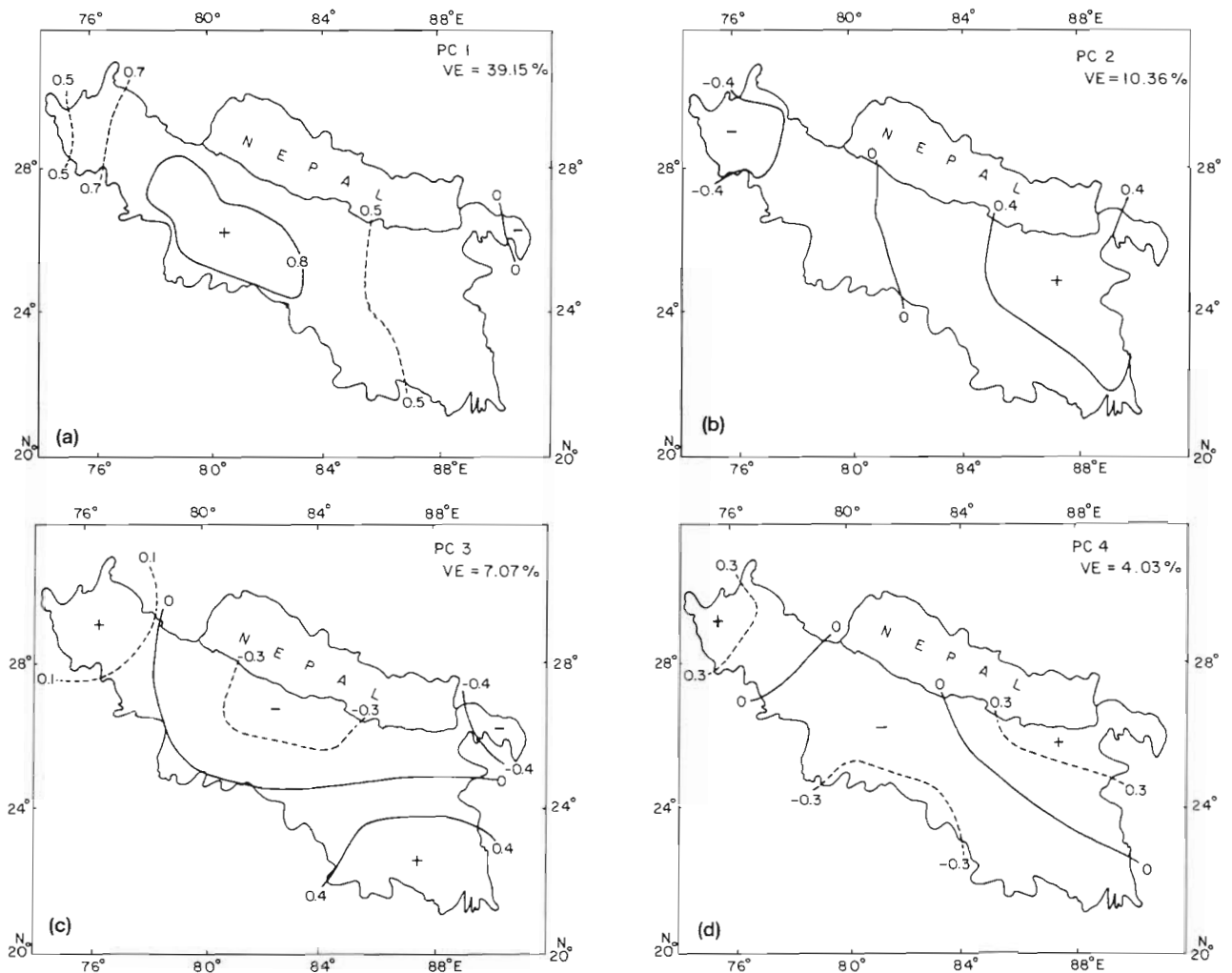1 unrotated PC (i.e. both), and such analysis is inade-

Fig. 5. Unrotated PCs 1 to 4 of the interstation correlation matrix of monthly (June–September) rainfall anomalies in the study area
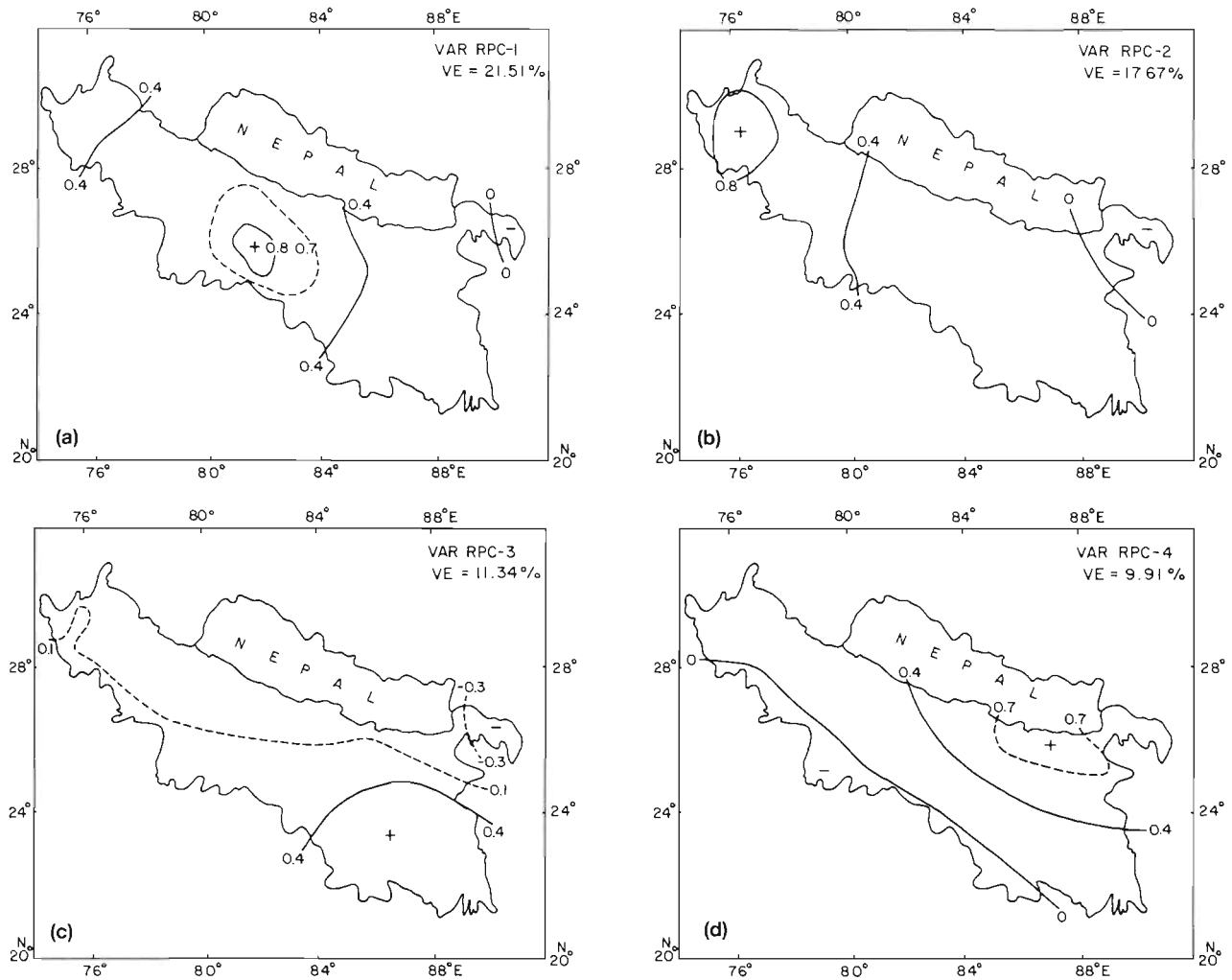
Fig. 6. Varimax-rotated PCs 1 to 4 of monthly rainfall in the study area

quate to isolate the individual areas of coherent rainfall variability.

Such pairwise plots of PC loadings are used to examine the simple structure arrived at through rotation here also. For the 4 PCs used for the study of seasonal and monthly time-scale rainfall, we have 6 pairwise plots of each unrotated and rotated PC. However, we present only 1 such case of an unrotated and rotated PC pairwise plot to illustrate the simplicity of structures arrived at through rotation using PCs 2 and 3 (Fig. 7). Fig. 7a suggests the almost random nature of scatter of unrotated PC loadings. In the scatter plot a dot represents the loadings of a single station, whereas numbers denote the number of stations having the same loading combination.

The varimax-rotated PC plot in Fig. 7b is in pronounced contrast to the plots of unrotated PCs shown in Fig. 7a. In a rotated PC pairwise plot the data points are distributed along the 2 axes, thus giving rise to a

mutually exclusive orientation of stations on the individual rotated PCs. Similar configurations are obtained for the other pair of varimax-rotated PCs for seasonal as well as for monthly rainfall (not shown). It was found that the individual stations have high loadings on only 1 rotated PC and rotated PCA has isolated the individual areas of coherent rainfall. Thus, rotated PC loading patterns are suitable for classification of stations into regions of coherent rainfall variability for the seasonal and monthly time-scale described below.

4.3. Rainfall regionalization

A cursory examination of Figs. 4 & 6 reveals that high loadings generally occur over the mutually exclusive region in different rotated PCs and that there is generally 1 center of high loadings on each rotated PC. In fact, as mentioned above, this simplification is the
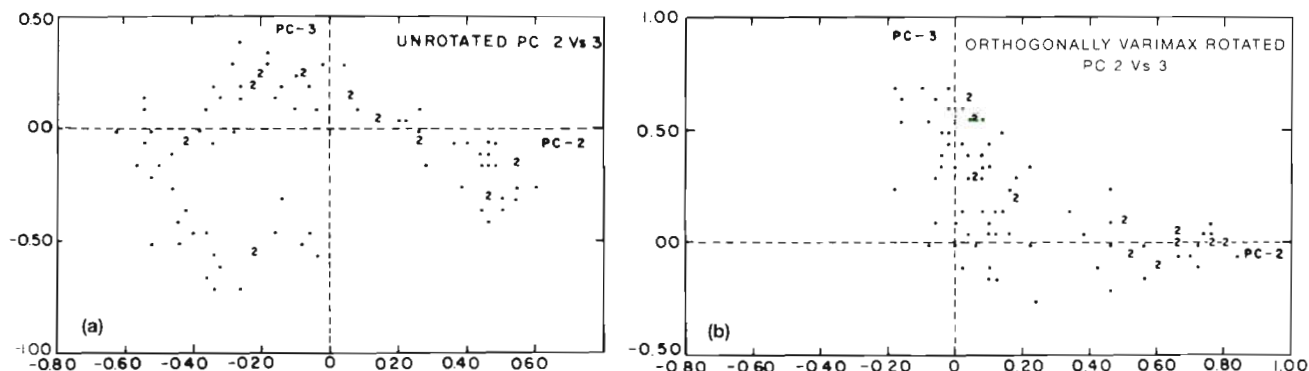
Fig. 7. Plots of station loadings for the pairs of (a) unrotated PCs and (b) orthogonally rotated PCs. PC 2 is along the x-axis and PC 3 along the y-axis. Data points represent single stations and numbers indicate multiple stations

principal objective of PC rotation. This characteristic of rotated PCs has been used extensively to delineate homogeneous regions of rainfall variation over a study area. This is usually done by selecting a loading magnitude for which a sharp gradient is observed on different rotated PC patterns. Richman & Lamb (1985, 1987) have used a loading of 0.4 for delineating homogeneous regions of precipitation over the central United States, North America and Canada. After careful examination of the spatial distribution of loadings, particularly the loading gradients, on the rotated PCs of this study, we also found loadings of 0.4 suitable for delineation of homogeneous regions. The isolines of 0.4 loadings in different rotated PCs were transferred onto a common base map for seasonal and monthly rainfall and the respective homogeneous regions are demarcated (Fig. 8). As can be seen from Fig. 8, the

0.4 loading contour is able to classify the entire study area into subregions with some overlap.

The most important coherent region lies over central and eastern Uttar Pradesh and the second coherent region covers Haryana, western and central Uttar Pradesh. The third and the fourth homogeneous regions divide Bihar, West Bengal and parts of eastern Uttar Pradesh roughly into northern and southern halves.

Between the seasonal and monthly rainfall fields the third and fourth homogeneous regions are interchanged because of the corresponding change in the rotated PCs. There is a considerable overlap between Regions I and II (overlapping regions were not shaded and wherever possible numbers were inserted), particularly for monthly rainfall. This overlap primarily arises due to the choice of loading magnitude, which is an
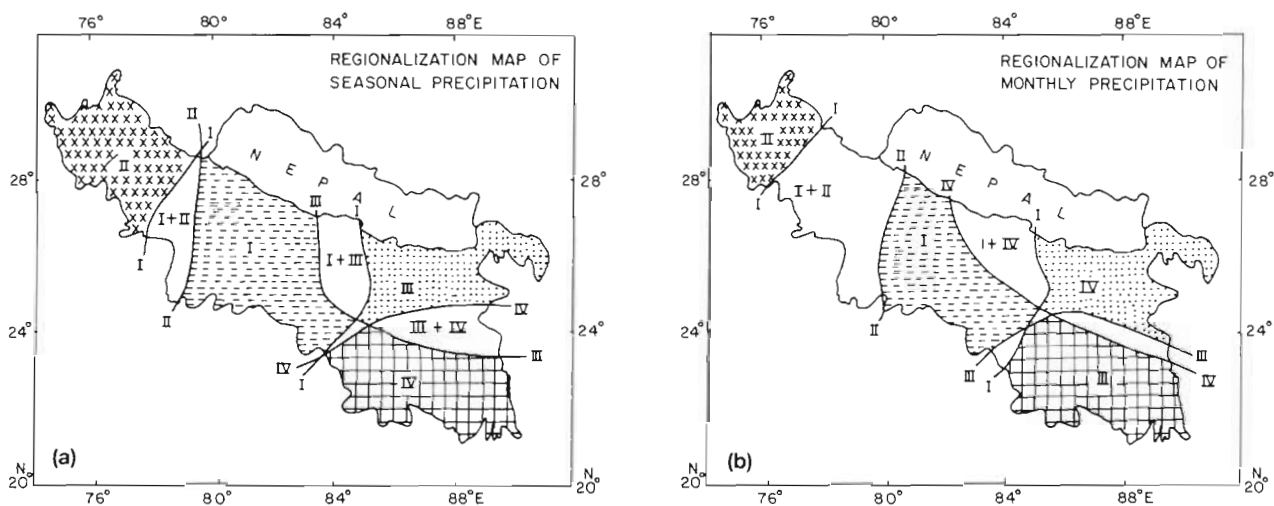

Fig. 8. Regionalization of sub-Himalayan region and Gangetic plains for (a) seasonal and (b) monthly monsoon rainfall on the basis of varimax-rotated PCs shown in Figs. 4 & 6. Regional boundaries correspond to 0.4 loading isopleths in individual rotated PCs, and the numbering of the region is consistent with Table 1. The overlapping regions were left blank and numbers inserted where space permitted, mutually exclusive regions were given separate shading patterns

objective criterion. In addition, it was not possible to draw a sharp boundary in the spatial domain between weather processes, which vary considerably in space and time from year to year. The overlap amongst regions, particularly between Regions I and II and between Regions I and III or IV (for seasonal and monthly rainfall respectively) could have been avoided by consideration of more PCs for rotation, but this could well have resulted in nonexhaustive classification and fragmented subregions of the study area.

The regionalization arrived at for the 2 time scales appears to be similar in nature except that for monthly rainfall there is a wider overlap between Regions I and II and the Regions III and IV have switched position as compared to seasonal rainfall (see Fig. 8b). In addition, the mutually exclusive regions (shaded areas) created by using loadings of rotated PCs 1 and 2 are comparable in size with the overlapped (unshaded) region for monthly rainfall. The appropriateness of the identified coherent regions needs to be further investigated by preparing composite monthly anomaly charts using a classification procedure such as the K-mean clustering method. In view of the above, the regionalization of seasonal rainfall seems to be more appropriate for use in future studies of climate variability and weather prediction research. Interpolations through the overlapping regions can be made in order to arrive at an exhaustive demarcation of the study area.

5. CONCLUSIONS

In this study we have analysed the major modes of spatial and temporal variation of seasonal and monthly monsoon rainfall over sub-Himalayan regions and the Gangetic plains of India using unrotated and rotated PCA. Four rotated PCs were used to identify 4 homogeneous subregions of coherent rainfall variability for seasonal and monthly time scales. On the basis of the results presented above the following main conclusions have been drawn.

(1) The seasonal and monthly rainfall of the sub-Himalayan region and the Gangetic plains show certain subregional coherent precipitation regimes with large-scale spatial patterns.

(2) The 4 statistically significant PCs of seasonal and monthly rainfall explain about 52 and 61 % of the data variances respectively. The corresponding first PC explained 26 and 39 % of the variance.

(3) Varimax rotation of PCs simplifies the loading structure significantly as compared to the unrotated PCs. This has enabled demarcation of the study area into homogeneous areas of coherent precipitation variability and made the relative importance of various weather processes clearer, viz. western disturbances in

the extra-tropical westerlies, cyclonic disturbances of monsoon and large-scale breaks.

(4) The most important zone of coherent rainfall in the study area lies over eastern Uttar Pradesh at a longitude of around 82° E. The second coherent region of rainfall lies over northern India, signifying the continual influence of extra-tropical systems on the mid-latitude westerlies.

(5) The area consisting of Haryana and Uttar Pradesh is divided into homogeneous regions in an east-west direction, but the region comprising Bihar and West Bengal is divided into northern and southern parts. Therefore, we infer that apparently different weather processes operate east and west of a longitude of 84° E during the summer monsoon.

The regions identified above can be of importance for studies of climate variability and weather prediction on seasonal and monthly time scales as well as for agricultural and water resource planning. The homogeneous regions identified here were stable over the study period as signified by the absence of a trend in unrotated or rotated PC scores. This implies that the internal structure (i.e. intercorrelations) of rainfall variability does not change over time. It must be added here that interannual variability, rather than multi-decadal variability or a secular trend, plays the most important role in determining the intercorrelations amongst station rainfall, a deciding factor in rainfall regionalization. The results of this study will hopefully be further validated by considering more PCs for rotation and more rainfall stations from neighboring countries like Nepal, Bhutan and Bangladesh. Similar analysis will also be conducted for other seasons as well as for shorter periods of rainfall in order to provide a better understanding of the details of climatic variability.

LITERATURE CITED

Bedi HS, Bindra MMS (1980) Principal components of monsoon rainfall. Tellus 32:296–298

Hastenrath S, Rosen A (1983) Patterns of Indian monsoon rainfall anomalies. Tellus 35 A:324–337

Iyengar RN (1991) Application of principal components analysis to understand variability of rainfall. Proc Indian Acad Sci 100(2):105–126

Kaiser HF (1959) Computer program for varimax rotation in factor analysis. Edu Psychol Meas 19:413

Kripalani RH, Singh SV, Arkin PA (1991) Large scale features of rainfall and outgoing longwave radiations over India and adjoining regions. Contr Atmos Phys 64:159–168

Kulkarni A, Kripalani RH, Singh SV (1992) Classification of summer monsoon rainfall patterns over India. Int J Climatol 12:269–280

Legates DR (1991) The effect of domain shape on principal component analyses. Int J Climatol 11:135–146

North GR, Bell TF, Calahan RF, Moeng F (1982) Sampling errors in estimation of empirical orthogonal functions. Mon Weath Rev 110:669–706

Overland JE, Preisendorfer RW (1982) A significance tests for principal components applied to cyclone climatology. Mon Weath Rev 110:1–4

Pandzic K (1988) Principal component analysis of precipitation in the Adriatic pannonian area of Yugoslavia. Int J Climatol 8:357–370

Parthasarathy B, Sontakke NA, Munot AA, Kothawale DR (1987) Droughts/floods in the summer monsoon season over different meteorological subdivisions of India for the period 1871–1984. J Climatol 7:57–70

Preisendorfer RW, Barnett TP (1977) Significance tests for empirical orthogonal functions. Preprints Fifth Conference on probability and statistics in atmospheric sciences. American Meteorological Society, Las Vegas, p 169–172

Rainbird AF (1967) Methods of estimating areal coverage precipitation. WMO/IND report no 3. WMO, Geneva

Rao YP (1976) Southwest monsoon. Meteorological Monograph No 1/76, India Met Dept, New Delhi

Richman MB (1986) Rotation of principal components. J Climatol 6:293–335

Richman MB, Lamb PJ (1985) Climate pattern analysis of three and seven day summer rainfall in the central United States, some methodological consideration and a regionalization. J Clim Appl Meteorol 24:1325–1343

Richman MB, Lamb PJ (1987) Pattern analysis of growing season precipitation in Southern Canada. Atmosphere-Ocean 25:138–156

Shukla J (1987) Interannual variability of monsoons. In: Fein JS, Stephen PL (eds) Monsoons. John Wiley & Sons, New York, p 399–463

Singh KK (1993) Principal modes of meteorological fields over monsoon region and their prediction. PhD thesis, University of Poona, India

Editor: T. Oikawa, Tsukuba, Japan