MAUSAM, 50, 2(April 1999), 145-152

551.509.333 : 551.553.21 : 551.577(540)

Canonical correlation analysis (CAA) model for long-range forecasts of sub-divisional monsoon rainfall over India

M. RAJEEVAN, V. THAPLIYAL, S.R. PATIL and U.S.DE

(Received 17 April 1997, Modified 16 June 1998)

सार — विहित सहसंबंध की विश्लेषण पद्धति (सी.सी.ए.) का उपयोग करते हुए भारत के 27 मौसम वैज्ञानिक उपखण्डों की मानसून (जून-सितम्बर) वर्षा के दीर्घ अवधि पूर्वानुमान के लिए पूर्वानुमान मॉडल विकसित किया गया है। भारतीय मानसून वर्षा के साथ महत्वपूर्ण सहसंबंध वाले 12 प्राचलों के समूह का उपयोग प्राग्वक्ता के रूप में किया गया है। यह मॉडल,1958-1994 की अवधि के आंकड़ों और तीन महत्वपूर्ण विहित मोडो के उपयोग से विकसित किया गया। इस मॉडल से पता चलता है कि कम त्रुटियों और उच्च कौशल वाले स्कोरों सहित भारत के मध्य भागों और उत्तर पश्चिमी भारत पर सुनिश्चित पूर्वानुमानों के मौसम वैज्ञानिक उपखण्डों के संबंध में उपयोगी प्रागुक्ति वाले कौशल थे। दक्षिण प्रायद्वीप, उड़ीसा, पश्चिमी बंगाल और बिहार में मौसम वैज्ञानिक उपखण्डों के संबंध में यह मॉडल कोई प्रागुक्त कौशल नहीं दर्शाता है। इसी प्रकार के 12 प्रागुक्तों के उपयोग से विकसित किए गए अन्य सांख्यिकीय मॉडल की तुलना में विहित सहसंबंध पद्धति मॉडल का निष्पादन बेहतर पाया गया है। विहित सहसंबंध पद्धति मॉडल से, अधिकतम वर्षा वाले वर्षों के दौरान उपखण्डों में मानसून की वर्षा के अधिक और कम वर्षा की श्रृंखलाओं के पूर्वानुमान में सामान्य कौशल का भी पता चलता है।

ABSTRACT. Using the canonical correlation analysis (CCA) approach, a forecast model for long range forecasts of monsoon (June-September) rainfall of 27 meteorological sub-divisions over India was developed. A set of 12 parameters, which have significant correlation with Indian monsoon rainfall, was used as predictors. The model was developed with the data of the period 1958-1994 and by retaining three significant canonical modes. The model showed useful predictive skill in respect of meteorological sub-divisions over central parts of India and NW India with low errors and high skill scores for categorical forecasts. The model showed no predictive skill in respect of meteorological subdivision over south peninsula, Orissa, west Bengal and Bihar. The CCA model has been also found to perform better than another statistical model developed using the 12 same predictors. The CCA model also showed moderate skill in forecasting excess and deficient rainfall categories of sub-divisional monsoon rainfall during the extreme years.

Key words -- Long-range forecast, Canonical correlation analysis (CCA), Indian Monsoon rainfall.

1. Introduction

Agriculture, power generation and industrial production substantially depend upon monsoon rainfall during June-September, which contributes about 70% of the annual rainfall over most parts of the country. In view of the critical influence of large inter-annual variability of rainfall on agricultural and industrial production, seasonal prediction of monsoon rainfall becomes very important for policy-making efforts. India was the first country to start a systematic development of long range forecast techniques for the prediction of seasonal monsoon rainfall over the country. India Meteorological Department (IMD) now operates four models namely Parametric and Power Regression models (Gowariker *et al.* 1991), Dynamic Stochastic Transfer model (Thapliyal 1982) and Multiple Regression model (Thapliyal 1986). For the review of operational models of India Meteorological Department (IMD), Thapliyal and Kulshrestha (1992) may be referred. For general review of long range forecasts for monsoon rainfall over India, Hastenrath (1995) and Krishnakumar *et al* (1995) also may be referred.

The studies on prediction of seasonal rainfall over India are mainly focused on the country as one unit. However, it is a known fact that monsoon rainfall in India exhibits large spatial and temporal variability. On temporal scale, monsoon rainfall exhibits two important quasi-periodic oscillations; 10-20 day oscillation [(Krishnamurti and Ardanuy (1980)] and 30-50 day oscillation [(Krishnamurti and Subramanyam (1982)]. Similarly, monsoon rainfall exhibits large spatial variability. Sub- divisions in NW India exhibits large inter-annual variability whereas subdivisions over NE India more or less experience normal rainfall every year. Moreover rainfall of sub-divisions in the NE parts of the country is poorly or negatively correlated with the rest of the

TABLE 1

List of predictors used in the study

S.No.	Predictor	Reference
1.	East Coast Temp. March (ECT)	Gowarikar *et al.* (1991)
2.	NW India Pressure Anomaly May (NWPA)	Krishnakumar *et al.* (1995)
3.	NW India Min. Temp. Anomaly May (NWTA)	Krishnakumar *et al.* (1995)
4.	Northern Hemispheric Temperature	Verma *et al.* (1985)
5.	Bombay MSL Pressure Tendency (MAM-DJF) (BPT)	Parthsarathy *et al.* (1991)
6.	SOI Tendency (MAM-DJF) SOI	Krishnakumar *et al.* (1995)
7.	Darwin Pressure Tendency (Apr-Jan) DPT	Shukla and Mooley (1987)
8.	Nino 3 SST Tendency (MAM-DJF) NI3T	Krishnakumar *et al.* (1995)
9.	De Bilt Mean Temp. (DBT) January	Dugam *et al.* (1993)
10.	N.H. Pressure Anomaly (NHPA)	Gowarikar *et al.* (1991)
11.	NW India and Central India (MAM)	Krishnakumar *et al.* (1995)
12.	10 hPa Zonal wind Balboa, January (ZW10)	Bhalme *et al.* (1987)

country (Parthasarathy 1984). In the best monsoon years with excess rainfall there would be always some areas with deficient rainfall. Similarly in the worst monsoon years with large deficient rainfall there would be some areas with excess rainfall. Thus it is important to have long range forecasts of monsoon rainfall on smaller spatial scales as the forecast of monsoon rainfall for India as a whole may not provide information about large rainfall variability that occurs on smaller spatial scales.

We have therefore developed a model for long range forecasts of sub-divisional monsoon rainfall over India based on Canonical Correlation Analysis (CCA) approach. This approach originally proposed by Hotelling (1936) has been used by Barnett and Preisendorfer (1987) for monthly and seasonal forecasts of US surface air temperatures, Graham *et al* (1987) and Barnston and Ropelewski (1992) for prediction of ENSO episodes. Recently, Prasad and Singh (1996) reported results on CCA model for long-range forecasts of monsoon rainfall. They have developed the model by using 8 parameters as predictors for 29 Met. sub- divisions and have obtained encouraging results. Even though we have used similar technique, we have obtained comparatively better results, probably due to the selection of better predictors and following better stringent rules in carrying out the CCA procedures. Moreover, while developing the present model, we have given more priority for the operational requirements and constraints.

2. Brief review of canonical correlation analysis (CCA)

Canonical Correlation Analysis (CCA) is at the top of the hierarchy of regression modelling approach. It is a multi-variate statistical technique that calculates linear combinations of a set of predictors that maximizes relationships in a least square error sense to similarly calculated linear combinations of set of predictands. One such combination of predictors of predictands constitute one canonical mode. CCA is a useful procedure for determining the dominant linear modes of co-variability between two data sets. This technique is more complex than multiple linear regression or discriminate analysis which treats only one predicated at a time. CCA uses eigen values and eigen vectors in a specialized way, such that the structure of the covariance between predictor and predictant variable under the constraint of maximization of cross data set correlation explained with each successive mode.

2.1. *Method of calculations*

The methods of calculations in the CCA are given below. Let $Y(P,N)$ and $Z(Q,N)$ be the predictor and predictand data sets. P and Q are the numbers of predictors and predictands respectively and P need not be equal to Q. N is the number of years of data.The predictor and predictand data sets are detrended and standardized. The predictor and predictand data sets are then separately orthogonalised and truncated using EOF analysis. This analysis provides the eigen values (k and λ) and temporal (α, β) and spatial amplitudes (e and f) for each fields respectively. The number of modes to be retained is determined by objective criterion. The resultant predictor and predicated principal component time series are normalized and then analyzed in the main portion of the CCA procedure. The details of the main CCA procedure are given below.

(a) Calculate the cross correlation matrix (C) between the predictor (α) and predictand (β) time amplitudes as follows:

$$C = \Gamma_\lambda \, \alpha' \beta \, \Gamma_k \tag{1}$$

where Γ_λ is diagonal matrix of inverse square roots of predictand EOF eigen values and Γ_k is diagonal matrix of inverse square roots of predictor EOF eigen values.

(b) Calculate $C^* = CC^t$, t means transpose of the matrix.

(c) Obtain the eigen structure of C^* by obtaining eigen values and eigen vectors R. Arrange by decreasing values of μ (the canonical correlations) and scale each R to unit length.

(d) Calculate the temporal and spatial amplitudes of the predictor side of the canonical modes (U and g) respectively.

$$U = \alpha \, \Gamma_k R \tag{2}$$

$$g = e \, \Gamma_k R$$

(e) Calculate the scaling matrix

$$S = U^t \beta \qquad (3)$$

(f) The regression models relating the predictor canonical structure to the individual predictands are given by

$$\psi = g\, S\, f^t \qquad (4)$$

For hindcast estimates of $\hat{Z}$ the following equation is used

$$\hat{Z} = U\, S\, f^t \qquad (5)$$

For more details, Barnett and Presiendorfer (1987), Graham *et al.* (1987), Barnston and Ropelewski (1992), Bretherton *et al.* (1992) and Prasad and Singh (1996) may be referred.

3. Predictor and predictand data sets

We have used a set of 12 parameters as the predictor set. The list of parameters is given in Table 1. These are regional as well as global parameters which have statistical and physical relationships with Indian monsoon rainfall. There are six temperature, five pressure and one upper air parameters as the predictors. These parameters are selected based on their stable and significant correlations with monsoon rainfall and are widely used in the long-range forecast models. These parameters are also selected on the basis of availability on real time basis by the end of May for operational purposes. One of the notable exception in the predictor data set is the 500 hPa ridge position (Banerjee *et al.* 1978) which is showing poor correlation with monsoon rainfall during the last 10 years. However, it is to be mentioned that this predictor data set need not be the best optimal predictor data set for this problem. Identification of better predictors should be an ongoing research process.

The monsoon seasonal (June to September) rainfall of 27 meteorological sub-divisions of India constitute the predictand data set. The sub-divisions considered in the model development are shown in Fig.1. We have excluded the sub-divisions in the hilly regions and the islands from the analysis. Prasad and Singh (1996) considered two more sub-divisions (North Assam and South Assam) which are more like hilly regions where the rainfall variability is very small. We have therefore excluded from our analysis.

The data for predictor data set were collected from Monthly Climatic records of the World being published by NOAA and IMD weather records. Nino-3 index data was obtained from the Climate Analysis Centre, NCEP, USA. The rainfall for the monsoon season (June-September) of 27 sub-divisions were taken from the IMD records. We have considered the period 1958-1994 for the development and testing of the model.

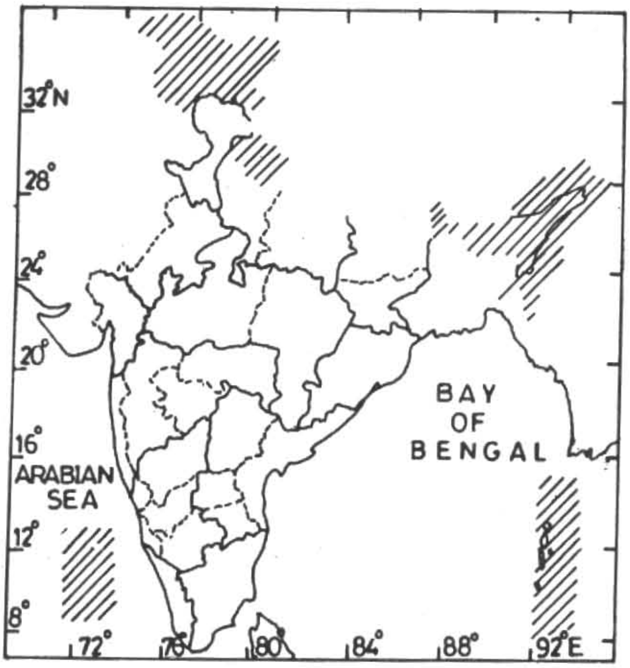

Fig.1. The 27 meteorological sub-divisions considered for the model development

4. Estimation of skill of the model

We have used the cross validation technique (Michaelsen 1987) to estimate the skill of forecasts. This technique is followed to control artificial skill inflation due to overfitting of random variability on the relatively short period of data (1958-94). The cross validation is done as follows.

Each of the $n + 1$ years is held out in turn and CCA is used to develop a prediction model from the remaining n years. For each case, a new year being held out the entire sequence of data, pre-processing and pre-orthogonalization are performed anew using the remaining years. The predictor data for the withheld year are then projected on to the predictor CCA loading patterns and predictand values are generated and verified against the observed data. We have used two skill scores to assess the performance of the model.

The first forecast skill is the root mean square of errors (RMSE) in standardized unit RMSE is calculated as

$$\text{RMSE} = \sqrt{\frac{(Z - \overline{Z})^2}{(n-1)}}$$

Here, Z and $\overline{Z}$ refer to the observed and cross-validated estimates of a predictand in standardized units respectively and the brackets refer to the expected value of the enclosed expression.

Another skill score is also calculated based on the model performance for categorical predictions. To calculate this skill score SF, each predictand is assigned values of 1, 0 or -1 according to whether the value is above normal (>19%) , normal (+ 19%) or below normal (< 19 %). The

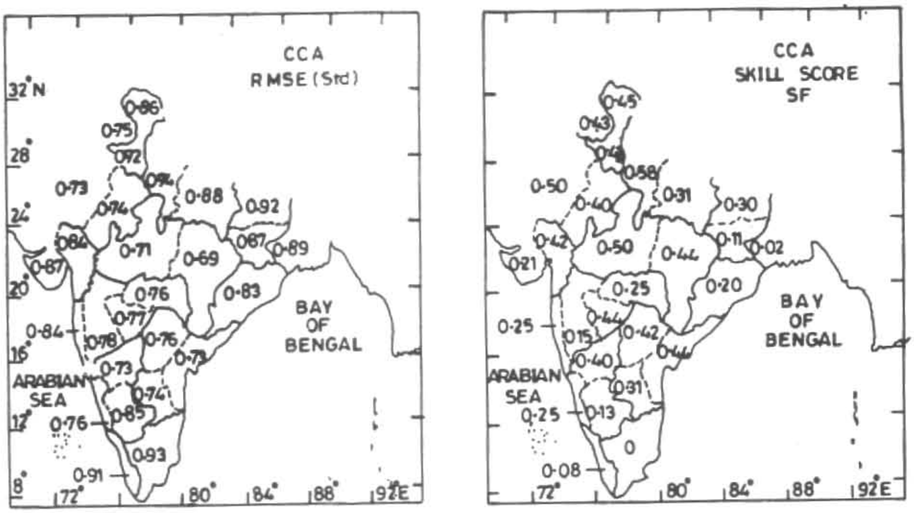

Fig.2. The RMSE (in standardized unit) and Skill score of categorical forecasts of the CCA model

TABLE 2

Details of Empirical Orthogonal Function (EOF) analysis of Predictands and Predictors

Eigen Mode	Eigen value	Cumulative Variance (%)
	(a) Predictand	
1	9.926	36.7
2	4.096	51.9
3	2.129	59.8
4	1.930	66.9
5	1.461	72.3
6	1.195	76.8
7	0.988	80.4
8	0.762	83.2
..	..	..
..	..	..
27	0.012	100.0
	(b) Predictor	
1	4.055	33.0
2.	1.665	46.6
3.	1.480	58.7
4	1.372	69.9
5	1.010	78.1
..	..	..
..	..	..
12	0.120	-100.0

cross-validated estimates of the predictand are transformed in the same way. For each year and each sub-division these classifications are compared and the number of correct forecasts are determined. The skill score SF is calculated as

$$SF = \frac{(E - C)}{(T - C)}$$

where, E is the number of correct forecasts, C is the number of correct forecasts based on 'climatology' and T is the total number of forecasts. These skill scores do not heavily penalize large errors, as do correlations and RMSE and give an idea of model performance for categorical predictions.

5. Canonical correlation analysis

Before beginning the CCA, the predictor and predictand data are detrended and then standardized so that all variables posses temporal stationarity thus ensuring that equally good predictors have equal opportunity to govern in the model regardless of their original variance. The predictor and predictand data sets are pre-orthogonalized with separate EOF analysis. The results of the EOF analysis are given in Table 2. The predictor and predictand data sets were truncated with five and six modes which together account about 78 % and 76 % variance respectively. The truncation has been done by retaining only the modes whose eigen value is more than 1.0. After the CCA, three canonical modes are retained. The first three canonical correlation values are 0.72, 0.51 and 0.46 respectively.

6. Forecast model and results

The model performance has been found to vary with the number of predictor and predictand modes retained as well as the canonical modes retained for the subsequent regression analysis. However, we have found that retaining 5 and 6 modes respectively (by the objectively criterion) for the predictor and predictand data sets gave us the best results. The results were also improved by retaining all three canonical modes in the regression equations than by including one

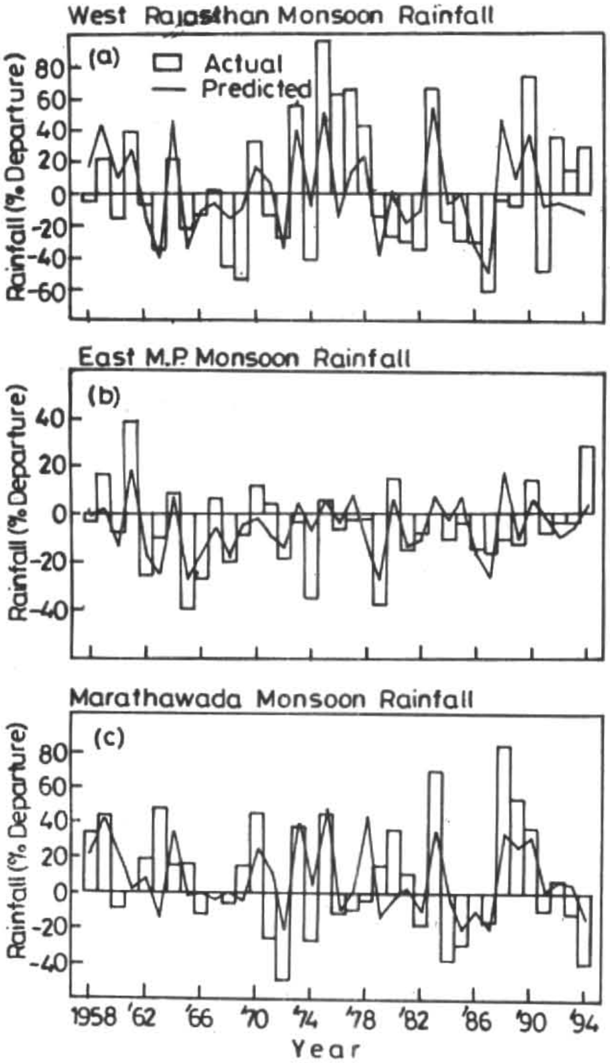

Fig.3. Actual and cross validated estimates of rainfall departures in respect of (a) West Rajasthan, (b) East Madhya Pradesh and (c) Marathawada

or two canonical modes. In this case, the third canonical modes is also found to be significantly to contributing in the predictive skill.

The RMSE for 27 meteorological sub-divisions for the period 1958-1994 are shown in Fig.2. The lowest RMSE of 0.69 is obtained for east M.P. Predictive skill in respect of south Peninsula and eastern parts of the country (Bihar, West Bengal and Orissa) is relatively low as errors are larger and close to 1.0. Root mean square errors across the central parts of the country extending up to NW India are relatively smaller. Prasad and Singh (1996) have obtained RMSE exceeding 1.0 for most of the sub-divisions during the independent period of 1969-1984.

The skill scores SF of the categorical forecasts for the period 1958-1994 using the cross validation technique for 27 meteorological sub-divisions are also shown in Fig.2. The highest skill of 0.58 is obtained for plains of west U.P. The skill scores in respect of southern parts of peninsula and eastern parts of India are generally close to zero implying very low predictive skill. Over central part of the country the skill score of Madhya Maharashtra is small (0.15). These results are comparatively better than the results obtained by Prasad and Singh (1996). While with the present model 15 out of 27 meteorological (55%) sub-divisions had skill score of >0.3 for categorical forecasts, the model developed by Prasad and Singh (1996) had 11 out 29 meteorological sub-divisions (38%) with skill score >0.3.

The actual and predicted rainfall departures for the period 1958-1994 for three sub-divisions (west Rajasthan, east M.P. and Marathwada) are shown in Fig.3. The year-to-year variation of rainfall over west Rajasthan is quite large. The model was however able to forecast the year-to-year variation quiet, reasonably with some exceptions. The two notable exceptions are for 1976 and 1988. In 1976, the model forecast was -14% but the actual was more than 60%. Similarly in 1988, the model forecast was 49% but the realized was close to zero. The model was also successful in forecasting large rainfall variations that occurred in 1970's. The forecasts for east M.P. are comparatively better with some exceptions of 1983, 1985 and 1988. The performance for Marathawada also was encouraging with some exceptions of 1990 and 1992.

The performance of the CCA model in forecasting sub-divisional rainfall is compared with yet another statistical model developed for sub-divisional monsoon rainfall. For that we have used the Principal Components Regression Analysis (Mc Cuen 1985). In this scheme, we have used the same 12 predictors which are then subjected to Principal Components Analysis (PCA). The resultant principal components which are mutually orthogonal are subjected to conventional regression analysis with sub-divisional rainfall of each sub-division separately thus developing 27 separate regression equations. This approach has the advantage in the sense that the multi collinearity of the predictor data set is avoided, because the principal components are mutually orthogonal. At the same time, the original variance explained by the predictors is also retained. This approach has been adopted by Singh and Pai (1996) for long range forecasts of monsoon rainfall using some oceanic parameters. The skill scores of principal components model for sub-divisional monsoon rainfall are shown in Fig.4. In most of the sub-divisions, the CCA model performed much better than the PCA model except over southern peninsula and eastern parts, where even otherwise the predictive skill is very low. The large difference in the predictive skill is obtained in respect of west M.P., in which the PCA model could not predict the categories accurately. Higher skill score for the PCA model was obtained for coastal Karnataka, south interior Karnataka and Tamil Nadu. However, overall the CCA model performed much better especially over NW India in both forecasting the quantum and categories of monsoon rainfall.

TABLE 3

Verification of CCA model forecasts for 1995 monsoon rainfall

S.No.	Sub division	Actual (%)	Forecast (%)	Error (%)
1	Gangetic W.B.	20	-3	23
2	Orissa	-17	3	-20
3	B. Plateau	10	-2	12
4	B. Plains	-2	-15	13
5	East U.P.	-9	1	-10
6	Plains of west U.P.	4	17	-13
7	Haryana, Chandigarh, Delhi	83	31	52
8	Punjab	44	20	24
9	Himachal Pradesh	17	-2	19
10	West Rajasthan	45	25	20
11	East Rajasthan	18	4	14
12	West M.P.	-10	7	-17
13	East M.P.	-10	0	-10
14	Gujarat Region	-21	-14	-7
15	Saurastra & Kutch	-28	-4	-24
16	Konkan and Goa	-11	12	-23
17	Madhya Maharashtra	-17	3	-20
18	Marathawada	-19	16	-35
19	Vidharbha	-13	1	-14
20	Coastal A.P.	5	10	-5
21	Telengana	-4	8	-12
22	Rayalaseema	29	10	19
23	Tamilnadu	6	12	-6
24	Coastal Karnataka	5	9	-4
25	N.I. Karnataka	-10	11	-21
26	S.I. Karnataka	-4	10	-14
27	Kerala	-6	-2	-4

TABLE 4

Average success rate (Percentage) of CCA model forecasts of sub-divisional monsoon rainfall during deficient and excess years

		Excess	Normal	Deficient	Total
			Deficient Years		
Actual	Excess	0	66	34	100
	Normal	0	89	11	100
	Deficient	0	31	69	100
			Excess Years		
Actual	Excess	84	16	0	100
	Normal	15	85	0	100
	Deficient	0	66	34	100

The performance of the model was examined by running the model for 1995 also with the data of 1995. The forecasts and realized rainfall departures are given in Table 3. In 1995, five meteorological sub-divisions (Gangetic West Bengal, Haryana and Delhi, Punjab, west Rajasthan and Rayalasema) received excess rainfall. The model could indicate the excess rainfall in respect of three meteorological sub-divisions (Haryana, Punjab and west Rajasthan). Out of two meteorological sub-divisions with deficient rainfall, the model prediction for Gujarat region, was - 14% indicating below normal. The average absolute error for 27 meteorological sub-divisions in 1995 was 17%.

Further, the skill of the CCA model in predicting the sub-divisional rainfall of extreme years also was examined. In this case, cross validation technique was applied for sub-divisional forecasts of seven deficient years (1965, 1966, 1972, 1974, 1979, 1982, 1987) and four excess years (1964, 1975, 1983 and 1988). The results are shown in Table 4. The results for two recent extreme years, 1987 (Deficient year) and 1988 (Excess year) are also shown in Fig.5.

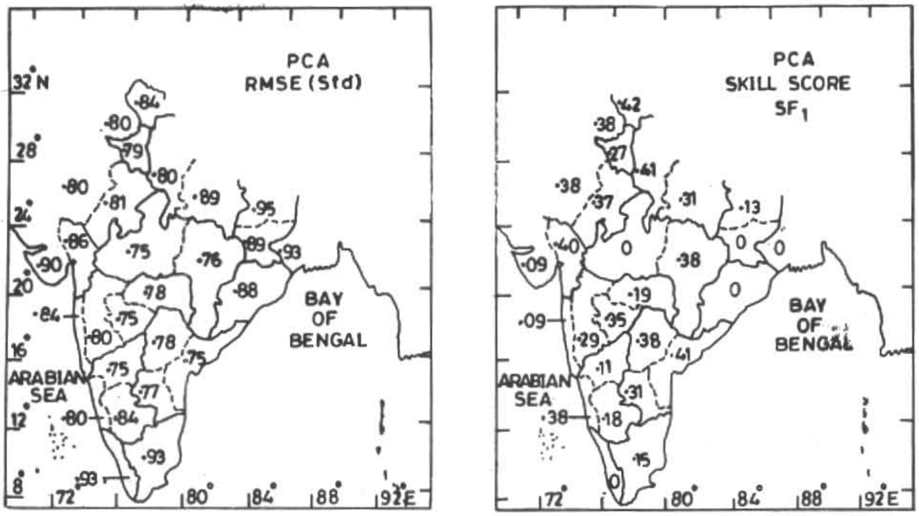

Fig. 4 . The RMSE (in standardized unit) and Skill score of categorical forecasts of the PCA model

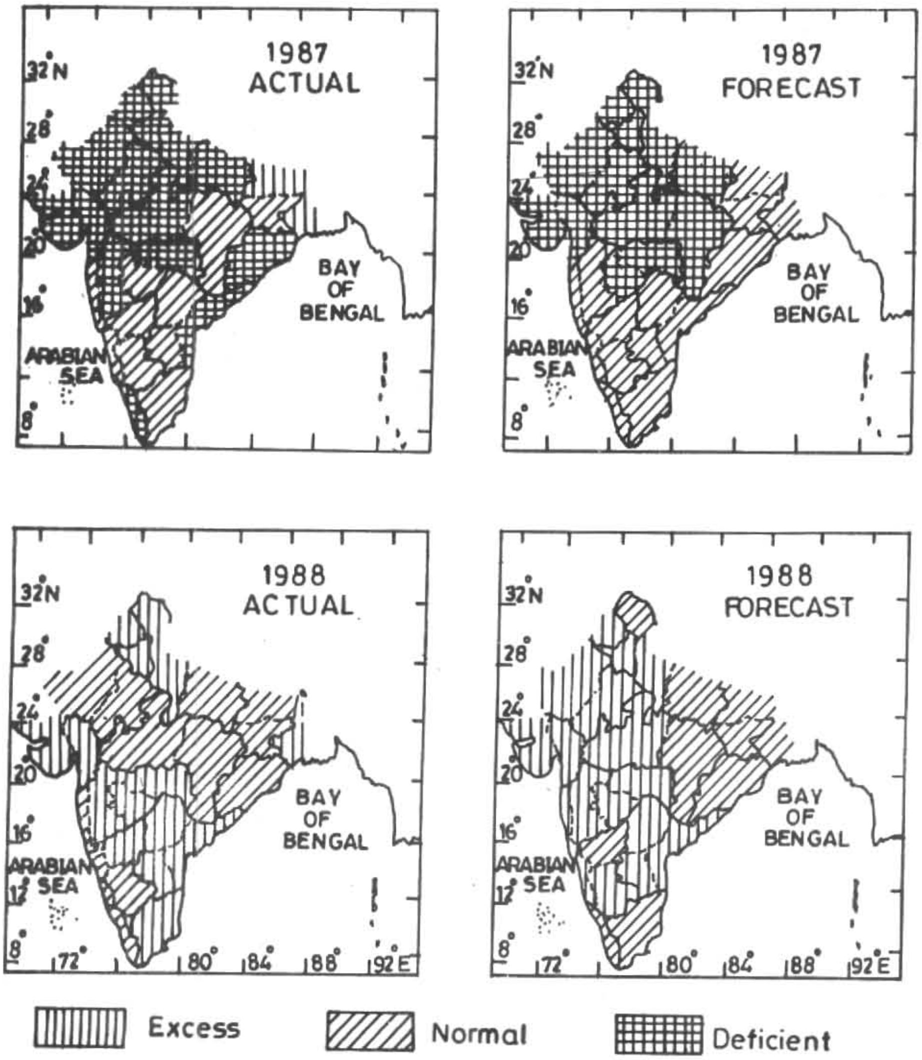

Fig.5. The actual and predicted sub-divisional rainfall departures for 1987 (deficient year) and 1988 (excess year)

In the deficient years, average success rate of predicting deficient rainfall is 69% and average success rate of predicting normal is 88%. During the excess years, average success rate of predicting excess rainfall is 84% and success rate of predicting normal rainfall is 85%. The success rate of the model in predicting deficient rainfall categories during the excess years is only 34%.

7. Conclusions

The CCA model developed thus shown useful predictive skill for the meteorological sub-divisions across the central parts of the country and NW India. The predictive skills for southern peninsula and eastern parts (Bihar, West Bengal and Orissa) are relatively very poor and not useful for operational purposes. The CCA model performed comparatively better than the PCA model probably because the CCA technique can extract the dominant modes of co-variability between the predictor and predictand data sets and use them in prediction. However, since the CCA technique considers only linear modes, constraints in better predictive skill are expected. The performance can however be improved by identifying better predictors and including in the analysis. Diagnostics of monsoon variability and identification of better predictor data set is the most important agenda in the long range forecast research.

Acknowledgements

We are thankful to Dr. N. Sen Roy, Retd DGM, and Dr. R.R. Kelkar, ADGM(S), India Meteorological Department for their keen interest in this study and encouragement and support extended to us in completion of this study. We are thankful to the anonymous referee for his constructive comments which helped to improve the quality of the paper. We are also thankful to staff members of LRF division in collection of data, computation and preparation of the diagrams and to Smt. N.S. Satbhai and Shri R.D. Rasker for typing the manuscript.

References

Banerjee, A.K., Sen, P.N. and Raman, C.R.V., 1978, "On foreshadowing southwest monsoon rainfall over India with mid-tropospheric circulation anomaly of April." *India J. Meteorol. Hydrol.*, **29**, 425-431.

Barnett, T.P. and Preisendorfer, R., 1987, "Origins and levels of monthly and seasonal forecast skill for United States surface air temperature determined by canonical correlation analysis, *Mon. Wea.Rev*, **115**, 1825-1850.

Barnston, A.G. and Rokpelewski, C.F., 1992, "Prediction of ENSO episodes using canonical correlation analysis", *J. Climate*, **5**, 1316-1345.

Bhalme, H.N., Rahalkar, S.S. and Sikdar, A.B., 1987, "Tropical quasi-biennial oscillation of 10 mb wind and Indian monsoon rainfall - Implication for forecasting, *J. Climatol*, **7**, 345-353.

Bretherton, C.S., Smith, C. and Wallace, J.M., 1992, "An inter comparison of methods for finding coupled patterns in climate data." *J. Climate*, **5**, 541-560.

Dugam, S.S., Kakde, S.B. and Verma, R.K. 1993, "A new Predictor for long-range forecasting of Indian summer monsoon," *Vatavaran*, **15**, 6-9.

Graham, N.E., Michaelsen , J. and Barnett, T.P., 1987, "An investigation of the El-Nino southern oscillation cycle with statistical models", *J Geophys. Res*, **92**, 14271-14289.

Gowarikar, V., Thapliyal, V., Kulshrestha, S.M., Mandal, G.S., Sen Roy, N., and Sikka, D.R., 1991, "A power regression model for long-range forecast of southwest monsoon rainfall over India," *Mausam*, **42**, 125-130.

Hastenrath, S., 1995, "Recent Advances in Tropical Climate Prediction," *J. Climate*, **8**, 1519-1532.

Hotelling, H., 1936, "The most predictable criterion, " *J. Ed. Physch*, **26**, 139-142.

Krishnakumar, K., Soman, M.K. and Rupakumar, K., 1995, "Seasonal forecasting of Indian summer monsoon rainfall: A review, *Weather*, **50**, 449-467.

Krishnamurti, T.N. and Ardanuy, P.A., 1980, "The 10 to 20 day westward propagating mode and "Breaks in the monsoon", *Tellus*, **32**, 15-26.

Krishnamurti, T.N. and Subramanian, D., 1982, "The 30 to 50 day mode at 850 mb during MONEX," *J. Atmos Sci*, **39**, 2088-2095.

Mc Cuen, R.H., 1985, "Statistical methods for Engineers", Prentice Hall, Inc, p. 439.

Michaelsen, 1987, "The estimation of artificial skill in forecast models", *J. Climate Appl. Meteor*, **26**, 1589-1600.

Mooley, D.A., and Paolino, D.A., 1988, "A predictive monsoon signal in the surface level thermal field over India," *Mon. Wea.Rev*, **116**, 339-352.

Parthasarathy, B., 1984, "Inter-annual and long-term variability of Indian summer monsoon rainfall," *Proc. Indian Acad. Sci, (Earth Planet Sci)*, **93**, 371-185.

Parthasarathy, B., Rupakumar, K., and Munot, A.A., 1991, "Evidence of secular variations in Indian monsoon rainfall - circulation relationships," *J. Climate*, **4**, 927-938.

Prasad, K.D. and Singh, S.V., 1996, "Forecasting the spatial variability of the Indian monsoon rainfall using canonical correlation," *Int. J. Climatol*, **16**, 1379-1390.

Shukla, J. and Mooley, D.A., 1987, "Empirical prediction of the summer monsoon rainfall over India," *Mon. Wea. Rev.*, **115**, 695- 703.

Singh, O.P. and Pai, D.S. 1996, "An oceanic model for the prediction of southwest monsoon rainfall over India," *Mausam*, **47**, 91-98.

Sundar, R.S. 1990, "Indian monsoon rainfall in relation to surface pressure anomalies." *Mausam*, **41**, 487-488.

Thapliyal, V., 1982, "Stochastic Dynamic model for long range prediction of monsoon rainfall in peninsular India," *Mausam*, **33**, 399-404.

Thapliyal, V., 1986, WMO, Long range forecasting of monsoon rainfall in India, Geneva, WMO/TD, **87**, 723-732.

Thapliyal, V. and Kulshrestha, S.M., 1992, Recent models for long range forecasting of southwest monsoon rainfall in India," *Mausam*, **43**, 239-248.

Verma, R.K. Subramanian, K. and Dugam, S.S., 1985, "Inter-annual and long term variability of the summer monsoon rainfall and its possible link with northern hemispheric surface air temperature, *Proc. Indian Acad. Sci*, **94**, 187-198.