

An oceanic model for the prediction of southwest monsoon rainfall over India

O. P. SINGH and D. S. PAI
India Meteorological Department, Pune
 (Received 9 June 1994)

सारांश — उत्तरी हिन्द महासागर के समुद्री मौसम विज्ञानिक आंकड़ों, तथा 1961 से 1991 तक की अवधि के मानसूनी वर्षा के आंकड़ों का उपयोग करते हुए, भारत में ग्रीष्मकालीन मानसून की वर्षा के लिए दीर्घावधि पूर्वानुमान हेतु नौ नए समुद्री पूर्वसूचकों की पहचान की गई है। एक विश्वसनीय समाश्रयण निदर्श को तैयार करने के लिए मूल विविधताओं के प्रमुख घटकों का विश्लेषण (पी० सी० ए०) किया गया है। मानसूनी वर्षा के साथ सर्वाधिक घनिष्ठ सहसंबंध रखने वाले मूल घटक, जो प्रथम पांच प्राचलों को अधिकतम प्रभावित करते हैं, इस प्रकार हैं — उत्तरी हिन्द महासागर के ऊपर भूमध्यरेखा और 10° उ० के बीच वायुमंडलीय परिसीमा स्तर में पवन शक्ति, अरब सागर (0°-15° उ०) के ऊपर माध्य वाष्पीकरण, 7.5°-17.5° उ० के बीच अरब सागर के ऊपर समुद्र सतही तापमान (एस० एस० टी०) माध्य, भूमध्यरेखा तथा 10° उ० के बीच बंगाल की खाड़ी के ऊपर माध्य वाष्पीकरण तथा अरब सागर के ऊपर समुद्र सतही दाब (एस० एल० पी०) माध्य इनमें से प्रत्येक का संबंध कई माह से है। जिन प्रमुख घटकों का मानसून वर्षा से अच्छा सहसंबंध है उनका उपयोग करते हुए अखिल भारतीय दक्षिण पश्चिमी मानसून की वर्षा हेतु एक बहु समाश्रयण निदर्श विकसित किया गया। 1987 से 1991 के वर्षों के लिए निदर्श की जांच की गई तथा यह पाया गया कि 1989 को छोड़कर इन सभी वर्षों की अखिल भारतीय ग्रीष्म कालीन मानसून के पूर्वानुमानित मान वास्तविक मानों के काफी निकट हैं। जबकि 1989 की ग्रीष्मकालीन मानसून की पूर्वानुमानित और वास्तविक वर्षा के बीच पर्याप्त अंतर था।

ABSTRACT. Nine new oceanic predictors for long range forecasting of Indian summer monsoon rainfall have been identified utilising the marine meteorological data of the north Indian Ocean and the monsoon rainfall data of the period 1961-91. In order to develop a reliable regression model the principal component analysis (PCA) of original variables has been done. Five parameters having maximum influence on first principal component, which is having highest correlation with the monsoon rainfall are: wind power in the atmospheric boundary layer over the north Indian Ocean between Equator and 10°N, mean evaporation over the Arabian Sea (0°-15°N), mean sea surface temperature (SST) gradient over the Arabian Sea between 7.5°-17.5°N, mean evaporation over Bay of Bengal between Equator and 10°N and mean sea level pressure (SLP) over the Arabian Sea, each pertaining to the month of May. A multiple regression model for all India rainfall of southwest monsoon season has been developed using the principal components which have got good correlations with the monsoon rainfall. The model was tested for all the years from 1987 to 1991 and it has been found that the predicted values of all India summer monsoon rainfall of all these years except 1989 were very close to the actual values. However, there was a substantial difference between the predicted and actual rainfall of 1989 summer monsoon.

Key words — Southwest monsoon, Long range forecasting, Principal component analysis (PCA), Correlation coefficient (CC), Multiple regression.

1. Introduction

The problem of prediction of Indian summer monsoon rainfall sufficiently in advance is one of the challenging tasks faced by the Indian meteorologists. The importance and utility of an accurate forecast of seasonal rainfall over the country during the southwest monsoon is well known and requires no emphasis. Thus, it is not surprising that this prediction problem has remained at the centre stage of Indian meteorology for more than a century. Pioneering efforts in this direction were made by Blandford (1884) and Walker (1910). Some recent notable contributions in this area have been made by Bhalme and Mooley (1980), Verma (1980),

Kung and Sharif (1980, 1982), Thapliyal (1982) and others. More recently, a parametric and power regression model has been developed by Gowariker *et al.* (1989, 1991) for long range forecasting of Indian summer monsoon rainfall.

Some of the parameters used in the prediction of monsoon rainfall are inter-related. Eigen techniques like empirical orthogonal function analysis (EOF) and principal component analysis (PCA) are very useful to obtain a set of independent variables from the set of inter-related variables. Prasad and Singh (1988) have employed EOF to study large scale features of monsoon rainfall and

TABLE 1

Oceanic predictors selected on the basis of their CCs with the all India summer monsoon rainfall

S.No.	Parameters
Atmospheric pressure parameters	
1.	Mean SLP over Arabian sea
2.	Mean SLP over Bay of Bengal
Temperature parameters	
3.	Mean SST over Arabian sea between 15°-20° N
4.	Mean SST gradient over Arabian sea between 7.5°-17.5° N
5.	Mean sea minus air temperature over the equatorial north Indian ocean (0°-10° N)
6.	Mean air temperature over Bay of Bengal
Evaporation parameters	
7.	Mean evaporation over Arabian sea (0°-15° N)
8.	Mean evaporation over Bay of Bengal (0°-10° N)
Wind parameter	
9.	Wind power in the atmospheric boundary layer over the north Indian Ocean between Equator and 10° N

their association with some oceanic and atmospheric variables. Srivastava and Singh (1993, 1994) have discussed empirical orthogonal functions associated with the parameters of long range forecasting of Indian summer monsoon rainfall.

Summer monsoon is primarily a sea-dependent phenomenon which possesses large interannual variability. Similarly, the meteorological and oceanographical conditions over the north Indian Ocean before the commencement of summer monsoon season exhibit large year-to-year variability. The major part of moisture for summer monsoon rains over India comes from the north Indian Ocean. There are quantitative studies to substantiate this (Pisharoty 1965, Singh and Joshi 1993 and Singh 1994 a & b). In addition to evaporation several other oceanic parameters appear to have significant influence on summer monsoon rainfall over India. These results have been reported by Singh (1993, 1994a, 1994b).

In order to develop a reliable prediction model for the Indian summer monsoon rainfall the

predictors from the north Indian Ocean need to be considered because of their physical linkages with the monsoon rainfall. In the existing models for the prediction of Indian summer monsoon rainfall the Indian Ocean predictors are not adequately represented. There is sufficient evidence that the state of the north Indian Ocean during pre-monsoon season, particularly during the month of May, plays a key role in the performance of subsequent monsoon. Thus the Indian Ocean predictors of pre-monsoon season would be crucial in the development of a stable model for the monsoon rainfall.

The objective of the present study was two fold. Firstly, to search some new predictors for the summer monsoon rainfall having physical relationships with the rainfall in addition to usual statistical relations. Secondly, to develop a reliable model which takes care of the inter-dependence of the predictors on each other. The emphasis has been given on the Indian Ocean predictors of pre-monsoon season. It was found that the oceanic signals during March and April were not very clear. The month of May turned out to be the important month as far as oceanic predictors were concerned. A model has been developed using the principal components (PC) of nine oceanic predictors of the north Indian Ocean pertaining to the month of May. The first principal component showed remarkable relationship with the all India percentage rainfall departure of southwest monsoon season. The model was validated for five years period from 1987 to 1991 which showed promising results. An experiment was done by including two land-based parameters; namely mean sea level pressure at Jodhpur during May and 500 hPa ridge position over India during April. The model based on eleven parameters (9 ocean-based + 2 land-based) did not show any improvement upon the performance of nine parameter oceanic model. Thus, the predictors of the north Indian Ocean alone appear to be capable of providing predictive signals for succeeding monsoon.

2. Data used

About 1.25 lakh marine meteorological observations of 31 years' period, from 1961 to 1991, recorded over the north Indian Ocean during the month of May, have been processed and analysed. The entire data of sea level pressure, zonal and meridional components of the wind, SST, air and dew point temperature, etc. were arranged on a 5° grid mesh over the north Indian Ocean area bounded by 0°-25° N and 50°-100° E. All India summer

TABLE 2

Correlation matrix between the 9 predictors (in the same order as given in Table 1) and the all India area-weighted rainfall for June to September (10)

Variable No.	1	2	3	4	5	6	7	8	9	10
1	1.00	—	—	—	—	—	—	—	—	—
2	.69	1.00	—	—	—	—	—	—	—	—
3	-.31	.14	1.00	—	—	—	—	—	—	—
4	.45	.42	-.35	1.00	—	—	—	—	—	—
5	.38	.48	.10	.56	1.00	—	—	—	—	—
6	-.01	.04	.40	-.09	-.04	1.00	—	—	—	—
7	-.55	-.33	.28	-.61	-.52	.14	1.00	—	—	—
8	-.35	-.28	.38	-.58	-.31	.33	.67	1.00	—	—
9	-.61	-.63	.26	-.45	-.40	.05	.52	.68	1.00	—
10	-.66	-.53	.28	-.63	-.44	.26	.64	.47	.46	1.00

monsoon rainfall data given by Thapliyal (1990) have been utilised to compute the correlation coefficients between different oceanic parameters and the monsoon rainfall. As the emphasis of the present study was on the development of the model, the details about the computations of various derived parameters are omitted here. However, the computational details have been given by Singh and Joshi (1993).

3. The predictors

532 correlation coefficients (CC) between different probable predictors and the predictand (summer monsoon rainfall) were computed. The list of selected predictors on the basis of CCs is given in Table 1, whereas Table 2 shows the CCs between the predictors and the monsoon rainfall. The correlation matrix given in Table 2 also brings out the inter-dependence of nine predictors. It may be mentioned that all parameters given in Table 1 pertain to the month of May.

It is seen from Table 2 that mean sea level pressure (SLP) over Arabian Sea during May has highest magnitude of CC with the subsequent summer monsoon rainfall over India. Next two highly significant CCs are those of mean evaporation over Arabian Sea (0° - 15° N) and the SST gradient over Arabian Sea (7.5° - 17.5° N) with the monsoon rainfall. The parameters having lowest CCs with the monsoon rainfall are: the mean SST over Arabian Sea (15° - 20° N) and mean air temperature over Bay of Bengal. But in view of large degree of freedom

even these two parameters appear to contain sufficient predictive signals for the subsequent monsoon rainfall.

Significant inter-dependence is observed in case of SLP over the Arabian Sea, SLP over Bay of Bengal and the wind power over the equatorial north Indian Ocean. This provides an insight into the physics of Indian summer monsoon. Lower SLPs over the Arabian Sea and Bay of Bengal are definitely conducive for the stronger cross-equatorial flow which in turn enhances the rate of evaporation over the equatorial north Indian Ocean. This is reflected in the inter-correlations of mean evaporation over the Arabian Sea (0° - 15° N), mean evaporation over Bay of Bengal (0° - 10° N) and the wind power over the north Indian Ocean (0° - 10° N).

Due to their intimate physical relations with the monsoon rainfall, Indian Ocean predictors only have been identified in the present work. It is felt that the state of Indian Ocean (particularly north) and its overlying atmosphere alongwith certain meteorological conditions prevailing over the Indian sub-continent about a month before the commencement of summer monsoon season holds the answer to the prediction problem of all India rainfall during summer monsoon. There is a need to explore this aspect more instead of looking for remote parameters which appear to be purely statistical and their physical linkages with the Indian monsoon rainfall are invisible.

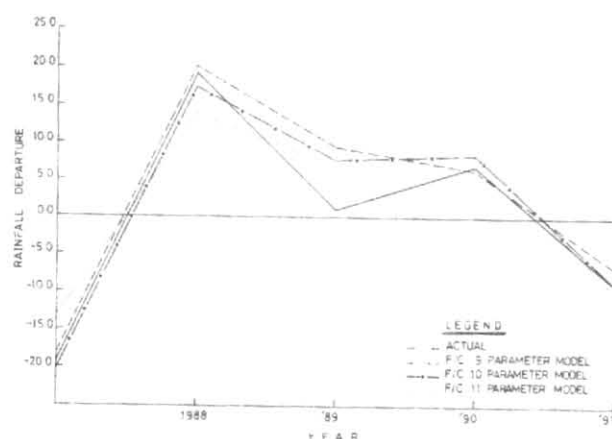


Fig. 1. Time series of first principal component score and all India area-weighted percentage rainfall departure

TABLE 3

Principal component loading matrix of nine parameter model

Variable No.	PC1	PC2	PC3	PC4	PC5
1	.76	.19	-.44	.01	.30
2	.67	.54	-.26	-.33	.06
3	-.37	.76	.31	-.25	-.27
4	.78	-.06	.28	.27	.20
5	.64	.35	.53	-.12	.22
6	-.21	.69	-.13	.67	-.13
7	-.81	.09	-.24	-.21	.05
8	-.77	.34	-.11	-.02	.47
9	-.81	-.09	.29	.07	.37
% of variance explained	46	18	10	08	07
C.C. with predictand	-.73	.03	.00	.06	-.23

4. Formulation of the model

The set of time-series of the predictors was subjected to PCA in order to generate a new set of variables which are independent of each other. An optimum number of variables from this linearly independent set have been used to develop a regression model for all India rainfall of summer monsoon.

4.1. Principal component analysis

A set of time series of m inter-correlated variables for n years can be represented by $(n \times m)$ matrix,

$$Z = [Z_{ij}; i=1, \dots, n; j=1, \dots, m] \quad (1)$$

Z is transformed into E via the matrix transformation

$$e_{ij} = Z_{ij} a_{ji} \quad (2)$$

using EOFA

In Eqn. (2) e_{ij} is the j^{th} empirical orthogonal variable for i^{th} case and a_{ji} is the i^{th} empirical orthogonal weight for j^{th} variable. In matrix form above equation can be written as:

$$E = Z A \quad (3)$$

EOFA is based on two conditions:

- Two different transformed variables are uncorrelated.
- Each transformed variable accounts for a maximum in residual total variance of the

original data set which is equivalent to maximising,

$$q E^T E = q A^T Z^T Z A = A^T R A \quad (4)$$

subject to the condition

$$A^T A = I_m \quad (5)$$

Here $q = \frac{1}{n}$, T denotes the transpose,

$I_m = m \times m$ identity matrix and R is an $(m \times m)$ correlation matrix defined as:

$$R = q Z^T Z \quad (6)$$

The solution of Eqn. (6) is unique and found from maximization of Eqn. (3) under the constraint that A has to be orthogonal using Lagrange multipliers' technique. This leads to the well known 'eigenvalue problem'

$$R A = A D \quad (7)$$

By definition, the $(m \times m)$ covariance matrix D of empirical orthogonal variables is $(m \times m)$ diagonal matrix whose elements are eigen values of R arranged in descending order of magnitude. A is

TABLE 4

Regression constants and coefficients for 1987-91

S. No.	Year	a_0	a_1	a_2	a_3	a_4	a_5
1	1987	-1.41	-8.04	-2.08	-1.52	—	—
2	1988	-2.08	-8.55	-2.21	—	-1.61	—
3	1989	-1.31	-9.46	-1.07	—	-1.67	—
4	1990	-1.24	-8.80	—	—	-1.24	-2.61
5	1991	-0.96	-8.67	—	—	0.68	-2.76

the orthogonal ($m \times m$) matrix of corresponding unit length eigen vectors.

If in addition the matrices E and A obtained from EOF analysis are rescaled according to :

$$L = AD^{1/2} \quad (8)$$

$$F = ED^{-1/2} \quad (9)$$

the complete procedure is called PCA.

Here F is a ($n \times m$) matrix known as matrix of PC scores each having zero mean and unit variance and L is ($m \times m$) matrix of PC loadings.

The results of PCA carried out on nine parameters of Table 1 have been given in Table 3.

Table 3 shows that the first principal component explains 46% of the variance and has got highest CC of -0.73 with the summer monsoon rainfall. The first five principal components together explain 89% of the variance. The CCs of PC1, PC2 and PC3 with the monsoon rainfall are very small. The fifth component has a CC of -0.23 and explains 7% of the variance. It is interesting to note that both PC1 and PC5 are negatively correlated to the monsoon rainfall.

Out of nine oceanic parameters, four; namely, mean SLPs over Arabian Sea and Bay of Bengal, mean SST gradient over Arabian Sea between 7.5°-17.5°N and mean sea minus air temperature over the equatorial north Indian Ocean (0°-10°N) had negative CCs with the monsoon rainfall, whereas remaining five were positively correlated. It is revealed by Table 3 that maximum influence on PC1 was those of wind power in the atmospheric boundary layer over the equatorial north Indian Ocean and mean evaporation over Arabian Sea

(0°-15°N). The other three parameters having significant influence on PC1 are: SST gradient over the Arabian Sea (7.5°-17.5°N) and mean evaporation over Bay of Bengal and mean SLP over Arabian Sea. Mean air temperature over Bay of Bengal and mean SST over Arabian Sea (15°-20°N) show very less influence on PC1 but have maximum influence on PC2. Evaporation over Bay of Bengal (0°-10°N), wind power in the atmospheric boundary layer over equatorial north Indian Ocean and mean SLP over Arabian Sea appear to have maximum influence on fifth principal component.

The time series of the first principal component alongwith that of all India summer monsoon rainfall departure from normal has been given in Fig. 1. There exists an inverse relation between PC1 and the monsoon rainfall which are approximately the mirror images of each other.

4.2. Multiple regression analysis

It is seen from Table 3 that PC1 has maximum CC with the monsoon rainfall followed by that of PC5. Consequently, PC1, PC5 alongwith PC4 were used to obtain the regression equation for the monsoon rainfall. General equation of the model is :

$$Y = \sum a_n P_n \quad (10)$$

where Y is all India area-weighted percentage rainfall departure from normal for June to September, a_0 is the regression constant, a_1, a_2 , etc. are the regression coefficients of corresponding principal components P_1, P_2 , etc. Table 4 gives the values of regression constants and the coefficients for the years for which model has been tested.

5. Validation of the model

As mentioned earlier the model has been validated for five years 1987-91. The results have

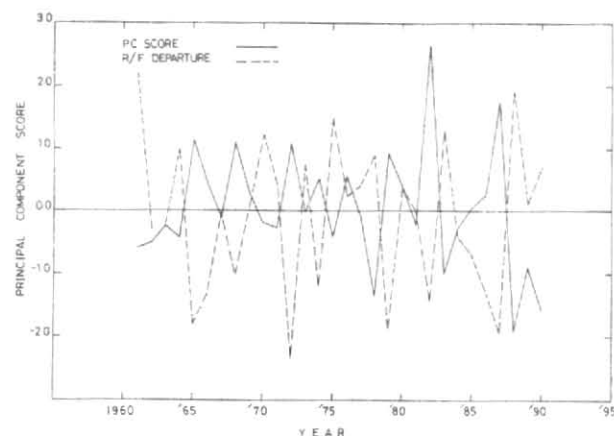


Fig. 2. Predicted and actual rainfall departure using different models

been presented in Fig. 2. The model has predicted 1987, 1988 and 1990 summer monsoon rainfall with remarkable degree of accuracy. The differences between the predicted and actual rainfall departures during these years were less or equal to 1%. The model-predicted rainfall departure of 1989 summer monsoon is 9.6%, whereas actual departure was just 1%. It may be admitted that in quantitative terms there is a considerable difference between the predicted and actual rainfall departures for summer monsoon of 1989. The most accurate prediction has been noticed for 1990 summer monsoon rainfall which is just 0.5% away from the actual rainfall departure.

The predicted and actual rainfall departures for 1991 southwest monsoon are -6.5% and -8.8% respectively. Thus the first version of the model has been able to predict four out of five monsoons very accurately. It may be pointed out that the five year (1987-91) sample contains an extremely bad summer monsoon year (1987, -19%), in extremely good monsoon year (1988, +19%), a just normal monsoon year (1989, +1%), a moderately good monsoon year (1990, +7%) and a moderately bad monsoon year (1991, -9%). Therefore, 1987-91 sample is well represented by all types of summer monsoons. The fact that the oceanic model has been able to predict all four summer monsoons of this sample having higher excess/deficient rainfall departures with very high degree of accuracy, goes to establish the flexibility of the model. It is hoped that further versions of this oceanic model will be able to produce accurate forecast on more number of occasions. The model's failure to predict 1989 summer monsoon accurately is being looked into.

An experiment was done by including two land-based parameters, namely mean SLP at Jodhpur

TABLE 5

The relative performance of models based on 9, 10 & 11 parameters

Year	Difference between predicted and actual rainfall departure (%)		
	9-parameter model	10-parameter model	11-parameter model
1987	1.0	-1.0	6.0
1988	0.9	-2.1	-5.7
1989	8.6	6.8	5.9
1990	-0.5	1.6	2.2
1991	2.3	0.1	-0.8

during May and 500 hPa ridge position over India during April. The results have been presented in Fig. 2. Table 5 shows the relative performance of the models based on 9, 10 and 11 parameters for 1987-91.

The inclusion of mean SLP at Jodhpur in the model resulted in the reduction of the difference between the predicted and actual rainfall for the year 1989 slightly. For 1991 also, difference decreased from 2.3% to 0.1%. The magnitudes of differences between predicted and actual rainfall departures for 1988 and 1990, however, increased slightly, whereas for 1987 the difference became -1% from +1%. Thus it can be said that there was no substantial difference in the model performance after the inclusion of mean SLP at Jodhpur during May. The inclusion of one more land-based parameter, namely 500 hPa ridge position over India during April rather mars the model performance instead of bringing out any improvement. The magnitudes of predicted and actual rainfall differences increased for 1987, 1988 and 1990, whereas, there was slight reduction for 1989 and 1991. Thus in general, it may be said that the inclusion of two land-based parameters in the model does not result in the improvement of the model performance.

The set of 9 oceanic parameters appears to be quite effective even if a qualitative prediction based on these parameters is considered for 1987-91. If a parameter, X_p having positive correlation with the monsoon rainfall is treated favourable (F) provided its value is more than its long-term mean and unfavourable (U) when its value is less than its mean, then the criteria for a parameter, X_n having negative correlation with the monsoon rainfall

TABLE 6

Behaviour of different oceanic parameters (same order as in Table 1) during the years 1987-91

Year	Variable									Favourable parameters %
	1	2	3	4	5	6	7	8	9	
1987	U	U	F	U	F	U	U	U	U	22
1988	F	F	F	F	F	F	F	F	F	100
1989	F	F	F	F	F	U	F	F	F	89
1990	F	F	F	F	F	U	F	F	F	89
1991	U	U	U	U	F	F	U	U	U	22

would be favourable if X_n is less than X_n and unfavourable if X_n is more than X_n . Table 6 gives the list of favourable and unfavourable parameters for each year during the period 1987-91.

It is revealed by Table 6 that during the wet monsoon year 1988 all nine parameters were favourable, whereas during the drought monsoon year 1987 only two were favourable. In 1991 also which was almost a deficient monsoon year only two parameters were favourable. During normal monsoon years 1989 and 1990 eight parameters were favourable. Thus a qualitative look at the values of different oceanic parameters can provide enough indications about the forthcoming summer monsoon.

6. Conclusion

The study has shown that the distributions of sea level pressure, sea surface temperature, evaporation rate and the wind power and instability in the atmospheric boundary layer over the north Indian Ocean during the month of May can give enough indications about the subsequent Indian summer monsoon. Due to their physical linkages with the monsoon rainfall, these parameters appear to be potential inputs to any reliable and stable model for long range prediction of summer monsoon rainfall over India. The first version of the model developed using the principal components of these oceanic parameters has shown encouraging results. The spatial and temporal distributions of north Indian Ocean parameters may have even more utility in the forecasting of monsoon rainfall on sub-divisional and monthly scales. There is a need to identify the predictors for the rainfall of each meteorological sub-division during each month of the summer monsoon season.

Acknowledgements

Authors are grateful to Dr. U. S. De, Deputy Director General of Meteorology (Weather Forecasting) and Dr. P. N. Sen, Director, Investigation and Development Section, India Meteorological Department, Pune for going through the manuscript and giving many valuable suggestions. Thanks are due to Shri A. B. Mazumdar for providing the PCA package. Shri Ajit Das is thanked for neat typing of the manuscript.

References

- Bhalme, H. N. and Mooley, D. A., 1980, *Mon. Weath. Rev.*, **108**, 1197.
- Blandford, H. F., 1884, *Proc. Roy. London*, **37**, 332.
- Gowariker, V., Thapliyal, V., Sarker, R. P., Mandal G. S. and Sikka, D. R., 1989, *Mausam*, **40**, 2, 115.
- Gowariker, V., Thapliyal, V., Kulshreshtha, S. M., Mandal, G. S., Sen Roy, N. and Sikka, D. R., 1991, *Mausam*, **42**, 125.
- Kung, E. C. and Sharif, T. A., 1980, *J. Appl. Met.*, **19**, 370.
- Kung, E. C. and Sharif, T. A., 1982, *J. Met. Soc. Japan*, Ser. II, **60**, 672.
- Pisharoty, P. R., 1965, *Proc. Sym. on Met. Results of the IIOE*, 43.
- Prasad, K. D. and Singh, S. V., 1988, *Adv. Atmos. Sci.*, **5**, 449.
- Singh, O. P., 1993, *Mahasagar*, **26**, 9.
- Singh, O. P., 1994 (a), *Proc. Sym. on Clim. Var. (TROPMET-94)*.
- Singh, O. P., 1994 (b), *Proc. CBIP Workshop (North zone)*.
- Singh, O. P. and Joshi, S. P., 1993, *Mausam*, **44**, 4, 337.

- Singh, O. P., Joshi, S. P. and Sundari, G., 1944, *Mausam*, 45, 2.
- Srivastava, H. N. and Singh, S. S., 1993, *Mausam*, 44, 1, 29.
- Srivastava, H. N. and Singh, S. S., 1994, Proc. Indian Nat. Sci. Acad., 60, 1, 181.
- Thapliyal, V., 1990, *Mausam*, 41, 339
- Verma, R. K., 1980, *Mon. Weath. Rev.*, 108, 1072.
- Walker, G. T., 1910, *Memoirs India Met. Dep.*, 21, 22.
-