Long-range forecast of monthly rainfall over India during summer monsoon season using SST in the north Indian Ocean

Most of the Indian subcontinent receives 70–90% of its annual rainfall during summer monsoon season (June–September), which influences the lives of more than one billion people. Since India's major income comes from agriculture, inter-annual variability of Indian summer monsoon rainfall (ISMR) plays a vital role not only in the economy of the country but also in the socio-economic conditions of the people. Droughts and floods greatly affect India's economy. Because of the drought in the year 2002, the khariff production was 30 million tons lower than the previous year[1]. Drought in the year 2002 could not be foreseen by any of the existing statistical or dynamical models[2]. Long-range forecasting of summer monsoon rainfall was reported through linear models by Delsole and Shukla[3]. They showed that minimum number of predictors are sufficient for accurate forecasts. Recent studies[4,5] reported long-range prediction of summer monsoon rainfall using Artificial Neural Network (ANN) approach. IMD developed power regression models[1] using 16, 8 and 10 parameters and issued the forecast for the 2003 summer monsoon season, which was accurate, but the forecast went wrong in 2004. This study used the rainfall data of 36 sub-divisions (including islands and hilly regions), while all the above studies used rainfall data (weighted average of 29 sub-divisions) of Parthasarathy *et al.*[6]. Goswami and Xavier[7] reported the prediction of rainfall under active and break cycles of the monsoon. Rajeevan *et al.*[1] mentioned that they could successfully develop a model to forecast the July rainfall and this was not possible for other months (June, August and September). Since the predictors need data up to June, it will be possible to give a forecast by the end of June. In the present study, simple regression equations are proposed for long-range forecast of ISMR (June–September) and monthly (June, July, August and September) rainfall over India using SST anomalies (SSTA) in the Arabian Sea (AS) and Central Equatorial Indian Ocean (CEIO). Possibility of predicting monthly rainfall and ISMR, about 3 months before the onset of summer monsoon has been shown here using only SST.

Utilizing the rainfall data of Parthasarathy *et al.*[6] and SSTA data (5 × 5 degree grids) of Kaplan *et al.*[8], it is found that the SSTA over the AS during winter (DJF –1/0 year) in the region 15–20°N; 60–70°E (AS), and 0–5°N; 80–85°E (CEIO) during fall (September–November, –1 year) are useful for long-range forecast of ISMR[9]. Strong and positive correlations are observed between the above SSTA and ISMR. Based on the above study and keeping in view the climate change over the Indian Ocean after 1976, correlations between the rainfall (monthly and ISMR) and SSTA in the above regions starting from September (–1 year) to May (0 year) preceding the summer monsoon season have been computed for the periods I (1978–93; $N = 15$) and II (1978–98; $N = 20$).

Figure 1 shows the monthly variations of correlations between SSTA (AS, top and CEIO, bottom) and rainfall (monthly

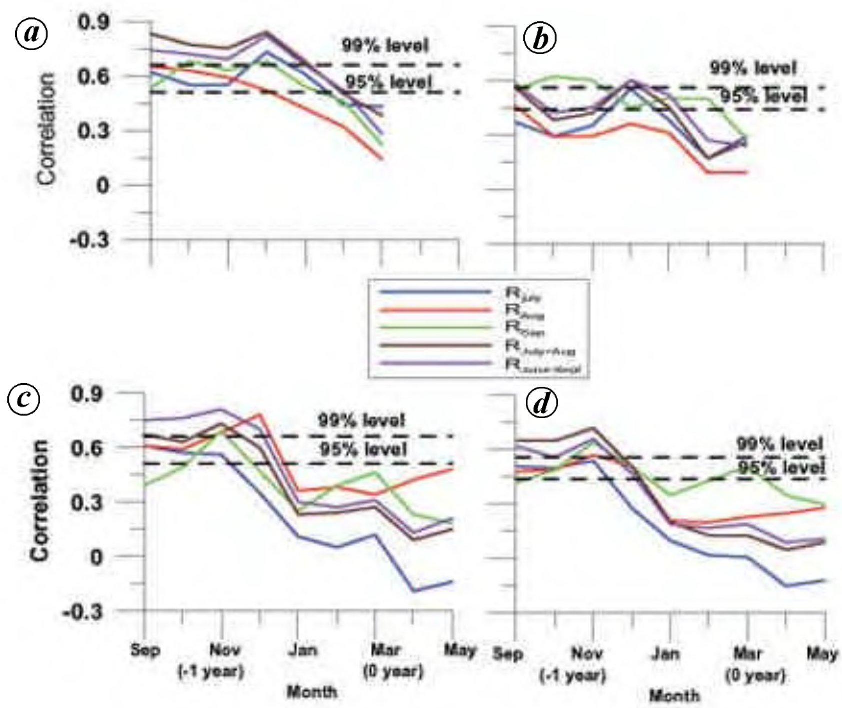

Figure 1. Monthly variation of correlations between SST indices (AS – top; CEO1 – bottom) and rainfall for the periods 1978–93 (***a***, ***c***) and 1978–98 (***b***, ***d***).

Table 1. Correlation between SST indices and rainfall for the periods 1978–93 (I) and 1978–98 (II)

SST index/rainfall		R_{July}	R_{Aug}	R_{Sep}	$R_{July+Aug}$	ISMR
AS	I	0.65	0.48	0.59	0.75	0.73
	II	0.51	0.43	0.57	0.47	0.54
CEO1	I	0.64	0.65	0.56	0.84	0.73
	II	0.57	0.56	0.56	0.75	0.68
AS + CEO1	I	0.79	0.60	0.65	0.90	0.85
	II	0.54	0.48	0.66	0.67	0.65
CEO2	I	0.69	0.64	0.58	0.87	0.78
	II	0.61	0.59	0.57	0.78	0.73
AS + CEO2	I	0.75	0.58	0.65	0.85	0.89
	II	0.54	0.44	0.64	0.70	0.65

Table 2. Regression coefficients (*A* and *B*) and SST index for periods I and II, to predict ISMR and monthly rainfall

		A	*B*	SST index
R_{July}	I	5.36	23.50	AS + CEO1
	II	8.96	22.61	CEO2
R_{Aug}	I	–	–	–
	II	7.66	21.68	CEO2
R_{Sep}	I	5.62	13.72	AS + CEO1
	II	5.41	12.70	AS + CEO1
$R_{Jul+Aug}$	I	9.026	46.52	AS + CEO1
	II	16.63	44.29	CEO2
ISMR	I	15.92	75.54	AS + CEO1
	II	28.37	72.10	CEO2

Table 3. Comparison between actual and forecast values of rainfall (cm) during June–September, July, July + August and September for the training and testing (bold) periods

	June–September ($R_{June-Sept}$)			July (R_{July})			July–August ($R_{July+Aug}$)			September (R_{Sept})		
		Forecast			Forecast			Forecast			Forecast	
Year	Actual	I	II	Actual	I	II	Actual	I	II	Actual	I	II
1979	70.8	73.9	69.7	22.6	22.9	21.8	42.5	45.6	42.9	13.9	13.2	13.0
1980	88.3	85.0	82.6	27.8	26.7	25.9	53.2	51.9	50.5	13.6	17.1	15.4
1981	85.2	86.1	85.1	28.9	27.1	26.7	51.5	52.5	51.9	19.8	17.5	15.7
1982	73.5	78.3	77.7	21.6	24.4	24.4	48.5	48.1	47.5	12.1	14.7	14.2
1983	95.6	88.3	90.8	27.4	27.7	28.5	56.7	53.7	55.2	25.1	18.2	17.0
1984	83.7	78.5	87.8	26.1	24.5	27.6	52.0	48.2	53.5	14.2	14.8	15.0
1985	76.0	75.0	77.8	25.3	23.3	24.4	47.2	46.2	47.6	14.7	13.6	13.3
1986	74.3	70.9	74.7	23.8	21.9	23.4	45.2	43.9	45.8	11.3	12.1	13.1
1987	69.7	73.4	78.8	20.7	22.8	24.7	44.4	45.3	48.2	13.7	13.0	14.0
1988	96.1	99.0	95.3	32.3	31.4	29.9	59.9	59.8	57.8	20.6	22.0	19.6
1989	86.7	87.1	89.4	28.6	27.4	28.1	51.7	53.1	54.4	16.6	17.8	15.8
1990	90.9	88.3	83.5	26.3	27.8	26.2	54.6	53.8	51.0	19.0	18.3	16.2
1991	78.5	85.3	86.3	25.6	26.8	27.1	48.5	52.1	52.6	12.1	17.2	16.1
1992	78.5	84.6	88.0	22.5	26.5	27.6	50.5	51.6	53.6	15.7	17.0	16.0
1993	86.6	83.7	83.2	28.1	26.3	26.1	50.7	51.2	50.8	20.4	16.7	15.7
1994	95.2	**86.4**	84.2	31.7	**27.1**	26.4	58.1	**52.7**	51.3	14.4	**17.6**	15.7
1995	79.0	**82.5**	81.1	27.7	**25.9**	25.4	50.5	**50.5**	49.5	16.5	**16.3**	15.2
1996	85.2	**83.3**	81.2	24.4	**26.1**	25.5	51.8	**50.9**	49.6	15.6	**16.5**	15.3
1997	87.1	**78.0**	83.5	26.7	**24.3**	26.2	52.4	**47.9**	51.0	16.6	**14.6**	13.8
1998	85.1	**94.3**	87.2	25.4	**29.8**	27.4	49.6	**57.2**	53.1	20.9	**20.4**	19.0
1999	82.1	**89.5**	**86.3**	25.3	**28.2**	**27.1**	46.2	**54.4**	**52.6**	19.0	**18.7**	**15.6**
2000	77.3	**83.0**	**82.8**	24.8	**26.9**	**26.0**	45.7	**52.2**	**50.6**	12.9	**17.3**	**15.4**
2001	80.6	**85.5**	**87.2**	26.4	**27.8**	**27.4**	45.4	**53.7**	**53.1**	11.5	**18.3**	**16.4**
2002	66.2	**83.7**	**91.2**	11.8	**26.3**	**28.6**	35.7	**51.2**	**55.5**	13.3	**16.6**	**14.4**
2003	85.6	**87.6**	**89.7**	28.6	**27.6**	**28.2**	50.8	**53.4**	**54.6**	17.8	**19.8**	**16.7**
2004	74.7	**87.1**	**88.2**	22.0	**27.4**	**27.7**	45.3	**53.1**	**53.7**	12.9	**17.8**	**16.6**

and ISMR) for periods I (*a* and *c*) and II (*b* and *d*). (Correlations of 0.51 and 0.66 are significant at 95 and 99% levels respectively, for period I; correlations of 0.44 and 0.56 are significant at 95 and 99% levels respectively, for period II). AS SSTA during fall (September–November) and winter (December–February) is strongly and positively correlated with rainfall for both the periods (Figure 1 *a* and *b*). The correlations are almost zero during April and May, which is not shown. SSTA during fall in the CEIO is strongly and positively correlated with the seasonal and monthly rainfall. It is weak during other months (Figure 1 *c* and *d*) under periods I and II. The correlations with rainfall during June are insignificant and hence are not shown. Based on the above, the correlations have been computed between winter SSTA over the AS (15–20°N; 60–70°E – AS) with rainfall. Correlations between the SSTA during fall in two regions in the CEIO (0–5°N; 80–85°E- CEO1; 5°S- 5°N; 80–90°E – CEO2) and rainfall have been computed. The results are presented in Table 1. SST indices AS and CEO1 are significantly correlated with ISMR and monthly rainfall for periods I and II, except with R_{Aug}. It could be seen that the combined index (AS + CEO1) is strongly and significantly (>99% level) correlated with the monthly rainfall and ISMR for periods I and II. Very high and significant correlations ($r = 0.85$ and 0.90; significant at >99% level) are observed between AS + CEO1 with the ISMR and $R_{July + Aug}$. Rajeevan *et al.*[10] reported positive and significant correlations ($r = 0.44$; significant at 99% level) between the ISMR and SSTA during January + February in the region 15–25°N; 50–70°E, for the period 1958–99. Rainfall during July (R_{July}) and September (R_{Sept}) is correlated at 0.79 and 0.66 respectively, for period I (1978–93). It is interesting to see that CEO2 and AS + CEO2 are also strongly and significantly correlated with rainfall for period II. CEO2 is a new predictor which has not been reported earlier (Table 1).

Based on these correlations, the regression equations have been proposed using the following relationship:

$$Y = AX + B, \quad (1)$$

where Y is the rainfall (in cm) and X is the SST index. A and B are the regression coefficients shown in Table 2.

Since correlations with rainfall in June and August are weak, they are estimated using the following equations:

Table 4. Comparison between actual and forecast rainfall (cm) during June and August for training and testing (bold) periods

	June (R_{June})			August (R_{Aug})		
		Forecast			Forecast	
Year	Actual	I	II	Actual	I	II
1979	14.4	15.1	13.8	19.9	22.7	21.0
1980	21.5	16.0	16.8	25.3	25.2	24.5
1981	13.9	16.1	17.5	22.6	25.5	25.2
1982	13.0	15.5	15.9	26.9	23.6	23.2
1983	13.8	16.3	18.5	29.3	25.9	26.7
1984	17.4	15.5	19.3	25.9	23.7	25.9
1985	14.0	15.2	16.9	21.9	22.9	23.2
1986	17.8	14.9	15.8	21.4	21.9	22.4
1987	11.6	15.1	16.6	23.7	22.5	23.5
1988	15.7	17.1	18.1	27.6	28.4	27.9
1989	18.4	16.2	19.2	23.1	25.7	26.3
1990	17.2	16.3	16.3	28.3	26.0	24.8
1991	18.0	16.0	17.6	22.9	25.3	25.5
1992	12.3	16.0	18.3	27.0	25.1	26.0
1993	18.2	15.9	16.7	19.7	24.9	24.7
1994	21.5	**16.1**	17.1	27.0	**25.5**	24.9
1995	12.9	**15.8**	16.3	23.4	**24.6**	24.1
1996	19.1	**15.9**	16.3	27.5	**24.8**	24.1
1997	17.8	**15.5**	18.7	26.3	**23.6**	24.7
1998	14.7	**16.8**	15.0	24.0	**27.3**	25.7
1999	17.5	**16.4**	**18.1**	21.3	**26.2**	**25.5**
2000	18.6	**13.5**	**16.9**	20.9	**25.3**	**24.6**
2001	23.1	**13.6**	**17.7**	19.0	**25.9**	**25.8**
2002	17.2	**16.0**	**21.3**	23.9	**24.9**	**26.8**
2003	16.4	**14.5**	**18.4**	22.2	**25.8**	**26.4**
2004	16.4	**16.2**	**17.9**	23.3	**25.7**	**26.0**

Table 5. RMS errors (RMSE) for methods I (1978–93) and II (1978–98). Values of mean and standard deviation (SD) for the training, testing and whole periods are also presented

		Training period		Testing period		Whole period	
		I	II	I	II	I	II
R_{June}	Mean	15.8	17.0	15.5	17.6	15.7	17.2
	SD	0.6	1.4	1.1	1.0	0.9	1.3
	RMSE	2.6	3.0	3.5	3.0	4.6	3.2
R_{July}	Mean	25.8	26.2	27.2	27.2	26.4	26.4
	SD	2.5	1.0	1.4	0.9	2.2	1.7
	RMSE	1.9	2.4	2.6	3.6	2.2	2.6
R_{Aug}	Mean	24.6	24.7	25.5	25.6	25.0	24.9
	SD	1.7	1.6	1.0	0.8	1.5	1.5
	RMSE	2.5	2.3	3.0	4.4	3.9	3.0
R_{Sep}	Mean	16.2	15.4	17.8	16.4	16.9	16.7
	SD	2.6	1.7	1.6	0.6	2.3	1.6
	RMSE	3.0	2.6	3.4	3.9	3.0	3.2
$R_{Jul+Aug}$	Mean	50.5	50.9	52.7	52.7	51.4	51.9
	SD	4.2	3.4	2.4	1.7	3.7	3.2
	RMSE	1.9	2.7	5.9	5.5	3.9	3.5
ISMR	Mean	82.5	83.4	85.9	86.5	83.9	84.1
	SD	7.4	7.0	4.3	2.3	6.4	6.4
	RMSE	4.4	5.3	6.9	6.0	5.2	5.5

Table 6. Observed and predicted rainfall (cm) during 2005 summer monsoon season

		Predicted	
	Observed	I	II
June	13.8	16.3	13.8
July	33.5	27.7	28.6
August	19.9	25.9	24.7
September	20.4	18.2	16.7
June + July	47.3	44.0	42.4
August + September	40.3	44.1	41.4
June–September	87.7	88.1	83.4

$$R_{\text{June}} = \text{ISMR} - [R_{\text{July+Aug}} + R_{\text{Sept}}], \quad (2)$$

$$R_{\text{Aug}} = R_{\text{July+Aug}} - R_{\text{July}}. \quad (3)$$

Equation (2) is used for both the periods for forecasting R_{June} and eq. (3) is used to compute rainfall during August (R_{Aug}) for period I.

Comparison between the observed and forecast values of rainfall (cm) during the training (1978–93, I; 1978–98, II) and testing periods (1994–2004, I; 1999–2004, II; indicated in bold), is shown in Table 3. Rainfall during July in 2002 was 51% below normal, which is three times the standard deviation regarded as an extreme climatic event and beyond the scope of a statistical model[1]. Though the data have been presented for the year 2002, it was excluded while computing root mean square errors (RMSE). In general, there is good agreement between the actual and forecast values, except in 1998 and 2004 in case of ISMR for period I, while the errors are high during 1992, 1994 and 2004 for period II. Maximum error is observed in 2002 which is an extreme year. In case of R_{July}, the forecast values are in good agreement with actual values during the whole period for both I and II, except 2002. Similarly, in case of $R_{\text{July + Aug}}$, there is good agreement between the actual and forecast values, except in only two years (1998 and 2001) and (2001 and 2004) for I and II respectively. The errors are high during 1983 and 2001 in case of R_{Sept} for both I and II.

The estimated rainfall in June (R_{June}) and August (R_{Aug}) is shown in Table 4. There is good agreement between the actual and forecast values, except in the year 2001 in case of R_{June} and R_{Aug}. It is interesting to see that the rainfall during June, August and September during 2002 could be predicted accurately. The RMS errors for the testing, training and the whole periods (excluding 2002) for I and II are presented in Table 5. Mean and standard deviations are also shown. The errors are ±2 to ±4 cm and ±3 to ±7 cm for the training and testing periods respectively, for period I. RMSE for the whole period (excluding 2002) were of the order of ±2 to ±5 cm for period I. For period II, the RMSE are ±2 to ±5 cm and ±3 to ±6 cm for the training period and testing periods respectively. RMSE for the whole period (26 years; excluding 2002) were of the order of ±3 to ±5 cm for II. There is not much variation in RMSE between I and II.

Comparison between the observed and predicted rainfall (cm) for 2005 summer monsoon season using data provided by Rajeevan[11] is shown in Table 6. There is good agreement between the observed (87.7 cm) and predicted (88.1 cm) seasonal rainfall from the equation for period I. The proposed simple equations for both the periods (I and II) could reasonably predict the monthly rainfall also, even though there were some extreme rainfall events (94.4 cm rainfall in one day over Mumbai and 49 cm in Khammam district, etc.). Rainfall during the first half (June + July) and second half (August + September) of the season could be predicted accurately.

Using the present method, it will be possible to predict monthly rainfall over India well in advance. The July rainfall can be predicted about 120 days in advance, which is crucial for irrigation. The SST indices (AS, CEO1 and CEO2) appear to have great potential for long-range forecasting of monthly rainfall and ISMR. However, further studies are desirable to confirm these results.

1. Rajeevan, M., Pai, D. S., Dikshit, S. K. and Kelkar, R. R., *Curr. Sci.*, 2004, **86**, 422–431.
2. Gadgil, S., Srinivasan., J., Nanjundiah. R. S., Krishna Kumar, K., Munot, A. A. and Rupa Kumar, K., *Curr. Sci.*, 2002, **83**, 394–403.
3. Delsole, T. and Shukla, J., *J. Climate*, 2002, **15**, 3645–3658.
4. Sahai, A. K., Pattanaik, D. R., Satyan, V. and Grimm, A. M., *Meteorol. Atmos. Phys.*, 2003, **84**, 217–227.
5. Iyengar, R. N. and Raghu Kanth, S. T. G., *Meteorol. Atmos. Phys.*, 2004.
6. Parthasarathy, B., Munot, A. A. and Kotawale, D. R., Research Report No. 65, IITM, Pune, 1995.
7. Goswami, B. N. and Xavier, P. K., *Geophys. Res. Lett.*, 2003, **30**.
8. Kaplan, A., Cane, M. A., Kushnir, Y., Clement, A. C., Blumenthal, M. B. and Rajagopalan, B., *J. Geophys. Res.*, 1998, **103**, 18,567–18,589.
9. Sadhuram, Y., Invited talk in the National Seminar on 'Advances in the frontier areas of meteorology at the turn of the millennium' Andhra University, Visakhapatnam, 29–31 October 2002.
10. Rajeevan, M., Pai, D. S. and Thapliyal, V., *Mausam*, 2002, **53**, 337–348.
11. Rajeevan, M., Personal communication, 2005.

ACKNOWLEDGEMENTS. I thank IITM, Pune, for making the rainfall dataset available on their website (www.tropmet.res.in). Thanks are also due to Dr Kaplan and his team for making SST anomalies data available. I thank Dr S. R. Shetye, Director, NIO, Goa and Dr K. S. R. Murthy, NIO, Regional Centre, Visakhapatnam, for encouragement. Thanks are also due to the reviewer for suggestions which helped improve the original version of the manuscript. This is NIO contribution number 4133.

Received 6 October 2005; revised accepted 2 May 2006

Y. SADHURAM

National Institute of Oceanography,
Regional Centre,
176, Lawson's Bay Colony,
Visakhapatnam 530 017, India
e-mail: sadhuram@.nio.org