1134

Climate Dynamics (2003) 20: 855–863
DOI 10.1007/s00382-003-0306-8

A. K. Sahai · A. M. Grimm · V. Satyan · G. B. Pant

Long-lead prediction of Indian summer monsoon rainfall from global SST evolution

Received: 31 May 2002 / Accepted: 3 December 2002 / Published online: 13 March 2003
© Springer-Verlag 2003

Abstract A methodology is presented for making optimum use of the global sea surface temperature (SST) field for long lead prediction of the Indian summer monsoon rainfall (ISMR). To avoid the node phase of the biennial oscillation of El Niño Southern Oscillation (ENSO)-monsoon system and also to include the multiyear ENSO variability, the ISMR-SST relationship was examined from three seasons lag prior to the start of the monsoon season up to four years lag. First, a correlation analysis is used to identify the regions and seasonal lags for which SST is highly correlated with ISMR. The correlation patterns show a slow and consistent temporal evolution, suggesting the existence of SST oscillations that produce significant correlation even when no direct physical relationship between SST and ISMR, at such long lags, is plausible. As a second step, a strategy for selecting the best 14 predictors (hot spots of the global oceans) is investigated. An experimental prediction of ISMR is made using the SST anomalies in these 14 hot spots. The prediction is consistent for the 105 years (1875 to 1979) of the model development and the 22 years (1980 to 2001) of the model verification period (these last years were not included in the correlation analysis or in the computation of the regression model). The predicted values explain about 80% of the observed variance of ISMR in the model verification period. It is shown that, to a large extent, the behavior of the ensuing ISMR can be determined nine months in advance using SST only. The consistent and skillful prediction for more than a century is not a product of chance. Thus the anomalously strong and weak monsoon seasons are parts of longer period and broader scale circulation patterns, which result from interactions in the ocean–atmosphere coupled system over many seasons in the past. It is argued that despite the weakening of the ENSO-ISMR relationship in recent years, most of the variability of ISMR can still be attributed to SST in a global perspective.

A. K. Sahai (✉) · A. M. Grimm
Dept. de Fisica, Universidade Federal do Parana,
Centro Politecnico, Jardim das Americas,
Caixa Postal 19044, CEP 81531-990,
Curitiba-PR, Brazil
E-mail: sahai@tropmet.res.in

V. Satyan · G. B. Pant
Indian Institute of Tropical Meteorology,
Pune, India

1 Introduction

The Indian summer monsoon is the major component of the Asian summer monsoon. India receives about 80% of its total annual rainfall during the summer monsoon season, from June to September. Indian agriculture is largely controlled by rainfall in this season. Small variations in the monsoon onset, in the spatio-temporal variability during the season and in the seasonal mean rainfall have a potential for significant economic and social impacts. Therefore, accurate forecasting of all India summer monsoon rainfall (ISMR) is beneficial to more than a billion people and has profound influence on agricultural planning and economic strategies of the country (Swaminathan 1998).

Forecasting of any climatic event requires the knowledge of its spatial and temporal variability. The yearly phenomenon of the monsoon occurs as a spectacular change in convective activity, especially between India and Australia. The Indian Ocean monsoon winds blow from the southwest during summer (wet) and from the northeast in winter (dry). This annual cycle of monsoon exhibits variability on time scales ranging from intraseasonal to decadal (Webster et al. 1998), as happens in most monsoon regions of the globe. It is well known that, even during a particular monsoon season, large-scale spatial and intraseasonal variability of the monsoon rainfall over India is evident (Krishnamurthy

and Shukla 2002). On interannual time scales, the ISMR exhibits a fairly distinct biennial cycle (Mooley and Parthasarathy 1984) and the multiyear ENSO frequency (Walker 1923). Occasionally long-term (3–4 decade) periods of persistent weak or strong monsoons occur over India (Kripalani and Kulkarni 1997). A large number of studies have analyzed the influence of intraseasonal, biennial and decadal variations on the interannual variability of ISMR (Webster et al. 1998). How the longer-term climate fluctuations modulate the interannual variability of ISMR are still not clear and need to be examined as part of the broader effort to develop useful long-lead prediction.

The long range forecasting of ISMR was started more than a century ago (Blanford 1884). Since then, many statistical (Thapliyal 1981; Shukla and Paolino 1983; Mooley et al. 1986; Bhalme et al. 1986; Shukla and Mooley 1987; Gowariker et al. 1991; Navone and Ceccatto 1994; Goswami and Srividya 1996; Sahai et al. 2000) and also dynamical (Manabe et al. 1974; Hahn and Manabe 1975; Palmer et al. 1992; Chen and Yen 1994; Ju and Slingo 1995; Sperber and Palmer 1996; Soman and Slingo 1997; Harrison et al. 1997) forecasting models have been developed and used. The principal scientific basis of these models or any seasonal climate forecasting model is the premise that the lower-boundary forcings (SST, sea-ice cover, land-surface temperature and albedo, vegetation cover and type, soil moisture and snow cover, etc.), which evolve on a slower time scale than that of the weather systems themselves, can give rise to significant predictability of the statistical characteristics of large-scale atmospheric events (Charney and Shukla 1981), especially in the tropics. Empirical forecasting of ISMR has been performed using combinations of climatic parameters including atmospheric pressure, wind, and snow cover, SST, and the phase of ENSO. Though the performance of a few statistical models is found to be better than that of dynamical ones, successful long-range prediction still remains elusive (Webster et al. 1998; Krishna Kumar et al. 1995). The performance of even the most successful acclaimed model of 16 parameters (Gowariker et al. 1991) has not been better than the climatological forecast in recent years (Kulkarni 2000; DelSole and Shukla 2002). The secular variation in parameters used as predictors in statistical models is identified as the main problem (Ramage 1983; Krishna Kumar et al. 1995; Sahai et al. 2000).

The classical relation between ENSO and ISMR, which has been observed, is that in the majority of years during the ENSO warm (cold) events the ISMR was below (above) normal. Almost all the statistical seasonal prediction schemes of ISMR rely heavily on the change in magnitude of various ENSO indices from winter (December to February) to spring (March to May) prior to the start of monsoon season. The above average ISMR in 1997, during the greatest El Niño event of the last century, and excess during the moderate El Niño year of 1994 have prompted many studies to reexamine the ENSO-ISMR relationship. It has been shown that this relationship is weakening (Kripalani and Kulkarni 1997) and it was proposed that it might be due to global warming (Krishna Kumar et al. 1999). It was also observed that this weakening of the relationship is not peculiar to the Niño-3 index only, but appears to be with any ENSO index. This weakening has been observed not only in recent times, but also in earlier periods, e.g. 1920–1960 (Webster et al. 1998). The ISMR time-series is remarkably stable during the last 130 years and the behavior of the decadal variations and the interannual variability in recent years is similar to that in the past (Sahai et al. 2002). Thus the weakening of the relationship cannot be explained by the global warming alone.

Therefore, a question arises about origin of the interannual variability of ISMR in recent years as well as in periods in the past when the ENSO-ISMR relationship was also weak. Actually, it is also associated with other oceanic regions in addition to the eastern tropical Pacific, such as the warm pool of the west Pacific Ocean, the northwest Pacific Ocean (Ju and Slingo 1995; Soman and Slingo 1997) and the Indian Ocean (Saji et al. 1999). Several studies have documented empirical links between Indian Ocean SST anomalies and monsoon variability. Negative correlations exist between the ISMR and 16 months earlier SST near Indonesia (Nicholls 1983). Furthermore, the links between the ISMR, the Indian Ocean and the tropical eastern Pacific have shown a biennial variability (Meehl 1987, 1997). Therefore, instead of calculating various indices from the tropical Pacific Ocean and using only a six month lag (December to May) for prediction of the ISMR, it is logical to include all oceans in different seasons with longer lags. An attempt in this regard has been made by Clark et al. (2000), who combined various indices from three regions in the Indian Ocean in different seasons, to develop a combined index with long lead time, which shows stable relationship throughout the period 1945–1997. Here we extend this attempt by considering global oceans with sufficient lag.

Section 2 deals with the data used in this study and describes the methodology. In Sect. 3, results are presented and discussed. The last section summarizes the major results and discusses the future impact of this study.

2 Data and method

2.1 Data

The ISMR data, from June to September, for the period 1871–2001 are obtained from the Indian Institute of Tropical Meteorology (IITM) data set (Parthasarathy et al. 1994). The monthly SST data used in this study are from the GISST 2.3b (Rayner et al. 1996) for the period January 1871 to December 2001. The SST data were originally on 1° × 1° resolution. They were averaged over boxes of 10° latitude × 20° longitude whose centers are 5° latitude × 10° longitude apart. Thus, there are overlapping regions between two neighboring boxes. The purpose of this is to achieve a good spatial

resolution (5° × 10°) while working with regions of larger extent (10° × 20°).

2.2 Correlation analysis between ISMR and SST

The monsoon exhibits variability in three major frequency bands longer than the annual cycle: the biennial oscillation, the multiyear ENSO-related variability (3–7 years) and the interdecadal variability. Webster et al. (1998) have noted that the biennial oscillation in the ENSO-monsoon system is an oscillation, which tends to have a strong seasonality, with the maximum in boreal winter and the node in boreal spring to early summer. By node we mean the transition from positive to negative values or vice-versa. During this period there is a change in the sign of correlation between various indices of ocean–atmosphere system (Yasunari 1990; Webster and Yang 1992). Webster et al. (1998) have further noted that the implication of the changes in correlation is that an anomalous state of the ocean–atmosphere system in the equatorial Pacific Ocean basin tends to decay in the boreal spring and another state with opposite sign tends to develop at the time of the next summer Asian monsoon onset. This is referred to as the predictability barrier of the climate system in the tropics. Thus, to avoid the node phase of the biennial oscillation in the ENSO-monsoon system on the one hand and to introduce the multiyear ENSO variability on the other hand, the ISMR-SST relationship was examined from three seasons prior to the start of monsoon season up to four years lag. For instance, to predict the 1990 ISMR, SST data used were from March 1986 to November 1989. The correlation coefficients (CC) between the ISMR and the seasonal SST anomalies and also with the tendency of seasonal SST anomalies (change in seasonal SST anomalies from the previous season) were calculated for various ocean basins in the

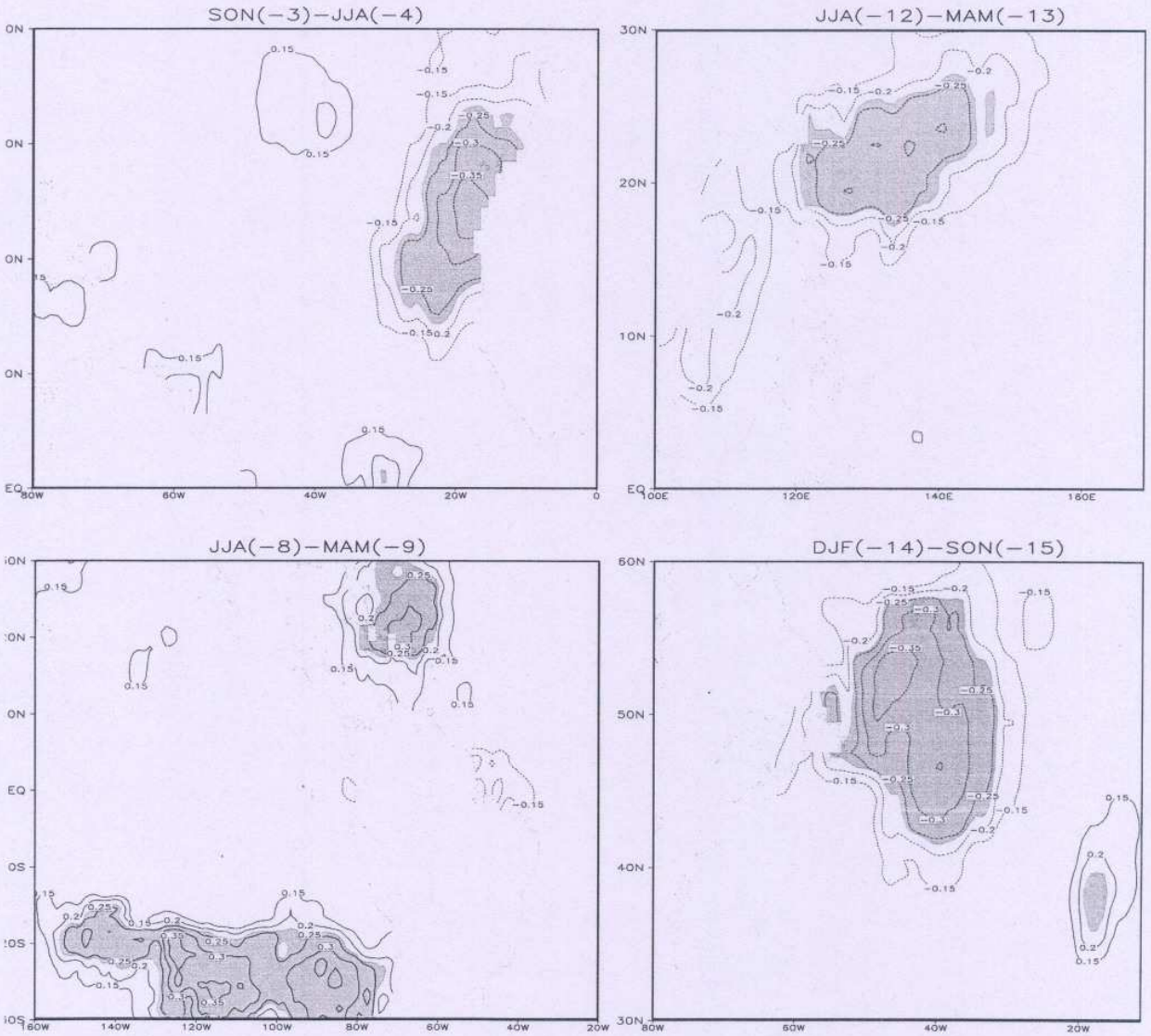

Fig. 1 Correlation coefficients between ISMR and tendency of seasonal SST anomalies for the period 1875–1979 are shown in some regions. The *labels* on the *top* of *each panel* show the seasonal lags for SST tendency. The *number* in the *braces* are the season lags from the start of ISMR. Contour intervals are 0.05 and the contours of –0.1, –0.05, 0.0, 0.05, 0.1 are dropped. *Dotted* (*continuous*) line represents negative (positive) values. The regions where CC is greater than 1% significance level are *shaded*

tropics and the northern extra-tropics (25°S to 55°N). The lag correlation is very significant in some regions even 4 years prior to the monsoon season (CC for some lags is shown in Fig. 1).

The correlation patterns show a slow and consistent temporal evolution, suggesting the existence of SST oscillations that produce significant correlations even when no direct physical relationship between SST and ISMR, at such a long lags, is plausible. This indicates the possibility of longer lead-time prediction. In many regions, the tendency of seasonal SST anomalies was found to be more correlated with ISMR than the seasonal SST anomalies themselves. The tendency in the seasonal SST anomalies represent more appropriately the anomalous response of the ocean–atmosphere coupled system to the seasonal march of the solar radiation than the SST anomalies. This may be the reason that ISMR is more related to the SST evolution than the SST itself.

Since we are more concerned with the prediction in recent years, the correlation analysis and also the cross-validated multiple regression analysis (discussed in the next section) were performed for the 105 years of data from 1875 to 1979 (model development set), while the recent 22 years of data from 1980 to 2001 (verification set) are kept to verify the performance of the proposed method.

2.3 Methods of selecting the best predictor set

When there are many predictors and their physical relationships with the predictant are not well defined, then a few best among them have to be selected based on robust statistical methods. We have adopted the most commonly used screening procedure known as stepwise regression (Wilks 1995). In this procedure one begins computing simple linear regressions between each of the available L predictors and the predictant. The predictor whose linear relationship is the best among all candidate predictors is chosen as first predictor. Different ways are proposed to detect the best relationship. One way is to select the predictor with minimum root mean square error (RMSE) in the model development set. DelSole and Shukla (2002) have used a cross validation scheme (Stone 1974; Michaelsen 1987) for doing this. In this scheme, the model development set (N data) is successively divided into pairs of mutually exclusive sets, the independent and the dependent. A regression model is developed with each dependent set and then used to predict the corresponding independent set. Repeating this procedure for all pairs, N predicted values are obtained with different regression models for each predictor variable. Values of RMSE are calculated for each predictor, by comparing the N predicted and observed data of the development set. Then the best predictor is that one, for which the RMSE is the minimum. DelSole and Shukla (2002) have applied the cross-validation technique for prediction of ISMR and shown that there is no qualitative difference between the results of using cross-validation with one year or five years in the independent data set. Therefore, the data of the model development period (from 1875 to 1979) were divided into N (=105) mutually exclusive dependent and independent sets in which each independent set consists of one year and the remaining N–1 years are in the corresponding dependent set. This procedure is called 'leave-one-out' cross-validation and is often confused with the 'jackknife' method. Both involve omitting each time one case in the model development period and obtaining the prediction model on the remaining subset. From a practical point of view, the major difference lies in their application. Cross validation is used for model selection and assessment whereas jackknife provides bias and variance estimates.

After selecting the first predictor, trial multiple regression equations are constructed using the first selected predictor in combination with each of the remaining L–1 predictors, and using the above criteria the second predictor is selected. Subsequent steps follow this pattern. The selection procedure is terminated when there is no further significant reduction in the RMSE for the independent development set. However, the principle of parsimony (Box et al. 1994) requires that an empirical model should employ the number of predictors as small as possible. The most often used statistics, to select the smallest and the best fit number of predictors, is Mallows' C_p:

$$C_p = \left(\frac{e_A^2}{\frac{e_B^2}{N-M-P-1}}\right) - (N - 2(P+1)) \quad (1)$$

where A is the reduced model with P predictors, B is the full model with $P+M$ predictors and $e_A{}^2$ and $e_B{}^2$ represent the mean squared error of models A and B, respectively. The correct model size can be determined by plotting C_p against $P+1$ (DelSole and Shukla 2002). A sensible strategy is to look for a model with a low C_p value, which is below but close to the 45° line. If there are several such points, the model with the smallest value of C_p is chosen. This holds for any P and M values, but the particular case of $M=1$ is used in this study.

2.4 Measures of predictive skills

The performance of the model is assessed through several indicators, like correlation coefficient (CC) between predicted and observed values, root mean square error (RMSE) and the absolute error (AE) defined as the absolute value of the difference between predicted and observed values for each pair of data. Moreover, another parameter, called the performance parameter (PP) was calculated:

$$\mathrm{PP} = 1 - (\mathrm{RMSE}/\mathrm{SD})^2 \quad (2)$$

where SD is the standard deviation of the ISMR. PP is the skill score defined in Wilks (1995) when the reference forecast is climatology. When PP > 0 (PP < 0), the forecast will be better (worse) than the climatological forecast (always mean value). The closer PP is to 1, the better the forecast.

3 Results and discussion

3.1 Assessing SST-ISMR relationship and identifying predictors

To begin with, 107 regions of 10° latitude × 20° longitude were identified as having their SST very significantly and consistently correlated with ISMR in different season lags. The seasonal SST anomalies or tendency of seasonal SST anomalies in these regions are correlated with ISMR at better than 1% significance level. To ensure the consistency, the 105 years of model development period was divided into two parts, the first 53 years and the last 52 years. The correlations were verified to have 10% significance level in these two parts in addition to have 1% significance level for the whole development period. Furthermore, all the correlations have not changed their sign between the two parts of the development period. This set of 107 regions can be seen as a set of potential predictors for the ISMR. However, the number of independent variables may be much less. Therefore, it is necessary to select the best independent predictor variables from these by using stepwise regression as described earlier. The values of RMSE and CC, for the predicted data of the independent set of the model development period, obtained from successively increasing the number of predictors in the cross-validation procedure, are shown in Fig. 2. It is difficult to judge from this figure when to stop the selection pro-

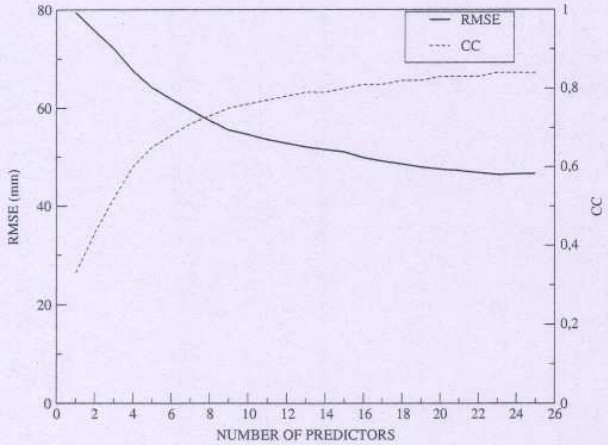

Fig. 2 The RMSE and CC versus the number of predictors for the independent set of the model development period, obtained by using the regression equations developed with dependent sets within the development period

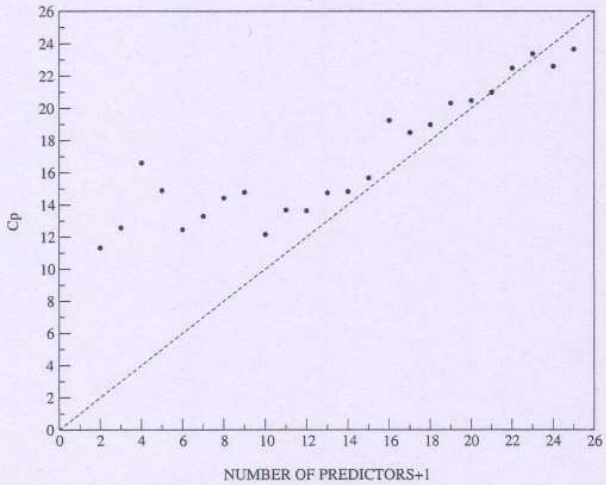

Fig. 3 Mallows' C_p statistic for the selected parameters when screening was made using the stepwise regression (Fig. 2). Also shown is the 45° line (*dash*)

cedure, because at each step there is some improvement. At this stage the Mallows' C_p statistic was plotted against $P+1$ in Fig. 3. Following the criterion described in the previous section, the number of predictors plus one is 21, and therefore the model with 20 predictors has the most favorable C_p statistic. This number is still very high for a prediction model. It could over-fit the model development data but may perform poorly on the verification data. The careful observation of Fig. 3 for the next closest point to the 45° line with fewer predictors gives $P+1$ as 15 and thus the model with 14 predictors is retained for further calculation. The 14 selected regions (1 from Indian Ocean, 8 from Pacific and 5 from Atlantic) are shown in Fig. 4 and the details are given in Table 1. It is interesting to note that this selection scheme has not selected any predictor from the east

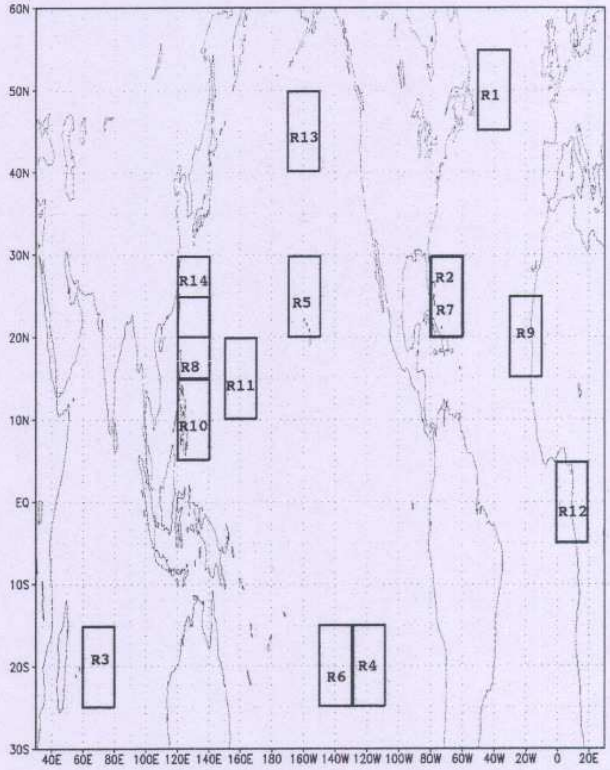

Fig. 4 Location of the 14 selected regions

Pacific and the north Indian Ocean, though there are many from these regions in the set of 107 predictors. However, there are several regions where SST anomalies are associated with ENSO events that appear among the selected regions in Fig. 4, such as the west Pacific and the subtropics of the north and south. Therefore, it can be seen that although the east Pacific is not among the most important regions, there are other regions where the SST is affected by ENSO events that are selected. The higher correlation between the ISMR and the west Pacific SST if compared with the east Pacific SST is coherent with the findings of Yasunari (1990) (Fig. 19b of Webster et al. 1998). There is no region among the selected 14 ones, whose SST anomaly is associated with the ISMR at less than four seasons lag (the nearest lag is JJA of the previous year). Thus it is possible to forecast ISMR using SST up to previous year JJA season, i.e., nine months in advance. This lead-time is even longer than the time series forecast (Goswami and Srividya 1996; Sahai et al. 2000), because the value of ISMR can only be obtained in the first week of October while the JJA SST is available by the first week of September.

3.2 Empirical prediction of ISMR

The seasonal SST anomalies and tendency in seasonal SST anomalies have been calculated for the 14 selected

Table 1 Table showing details of 14 regions

Region	Latitude		Longitude		Season lag for SST/ SST tendency	Value of CC
	Starting	Ending	Starting	Ending		
R1	45	55	310	330	DJF(–14)–SON(–15)	–0.37
R2	20	30	280	300	JJA(–8)–MAM(–9)	+0.30
R3	–25	–15	60	80	MAM(–17)	–0.30
R4	–25	–15	230	250	JJA(–8)–MAM(–9)	+0.31
R5	20	30	190	210	MAM(–5)–DJF(–6)	–0.23
R6	–25	–15	210	230	JJA(–4)–MAM(–5)	–0.24
R7	20	30	280	300	MAM(–9)–DJF(–10)	–0.27
R8	15	25	120	140	JJA(–12)–MAM(–13)	–0.31
R9	15	25	330	350	JJA(–4)	+0.30
R10	5	15	120	140	SON(–7)–JJA(–8)	–0.25
R11	10	20	150	170	JJA(–16)–MAM(–17)	+0.24
R12	–5	5	0	20	JJA(–8)	–0.23
R13	40	50	190	210	SON(–15)–JJA(–16)	–0.24
R14	20	30	120	140	JJA(–12)–MAM(–13)	–0.24

Note: (1) Values in braces ( ) under season lag represent the number of seasons prior to ISMR
(2) When there are two seasons in the season lag, the predictor is the SST tendency and when there is only one season the predictor is SST in the indicated region

regions in the given lag and then each series is standardized. The standardization is achieved by subtracting the mean of each series from each value of the respective series, and then dividing the resulting value by the standard deviation of that series. The following multiple regression equation for prediction of the ISMR anomalies in mm was derived using the model development period (1875–1979):

$$\begin{aligned}\mathrm{ISMR} = &-23.432 \times \mathrm{R1^T} + 15.189 \times \mathrm{R2^T} - 24.296 \\ &\times \mathrm{R3^A} + 25.161 \times \mathrm{R4^T} - 22.229 \times \mathrm{R5^T} \\ &- 15.047 \times \mathrm{R6^T} - 14.092 \times \mathrm{R7^T} - 27.294 \\ &\times \mathrm{R8^T} + 17.993 \times \mathrm{R9^A} - 10.988 \times \mathrm{R10^T} \\ &+ 13.334 \times \mathrm{R11^T} - 11.056 \times \mathrm{R12^A} \\ &- 11.899 \times \mathrm{R13^T} + 18.753 \times \mathrm{R14^T}\ . \qquad (3)\end{aligned}$$

Rj^T- are the standardized values of SST anomaly tendency and Rj^A is the SST anomaly in the regions Rj for the given season as per Table 1. The predicted anomalies were calculated using this equation. The predicted anomalies, the observed ones, and the absolute error expressed as the percentage of the long term mean (the long term mean is 853.3 mm and the SD is 84.4 mm), are shown in Fig. 5 for the model development period, and in Fig. 6 for the model verification period. The predicted values are very close to the observed ones. There are only eight occasions in the entire period when absolute error is greater than 10% of the ISMR long term mean (Table 2). For three of these years the predicted and observed anomalies show the same sign, and in four years the predicted and observed anomalies are both normal (within 10% of the long term mean). Thus in only one year (1975) has the model performed very poorly. For the 105 years of the model development period, CC = 0.85, RMSE = 44.3 mm and PP = 0.72 and for the 22 years of model verification period, CC = 0.89, RMSE = 38.8 mm and PP = 0.71. The 21-year sliding CC and PP for the entire period are shown in Fig. 7. The CC between the observed and predicted ISMR is almost always greater than the 0.01% significance level and the

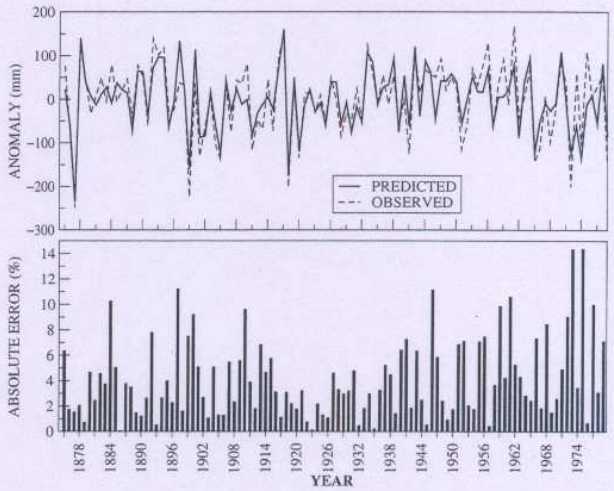

Fig. 5 Predicted and observed ISMR anomalies (*upper panel*) and % absolute error (*lower panel*) for model development period

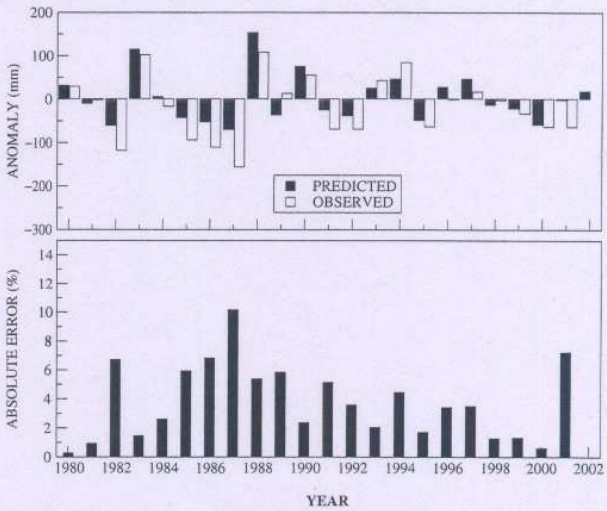

Fig. 6 Predicted and observed ISMR anomalies (*upper panel*) and % absolute error (*lower panel*) for model verification period

Table 2 Table showing the predicted and observed ISMR values for the years when absolute error (AE) is greater than 10% of long term mean ISMR

Year	Predicted anomaly (mm)	Observed anomaly (mm)	AE (%)
1884	−8	80	10.27
1897	133	38	11.21
1946	−44	51	11.14
1961	77	167	10.59
1973	−62	60	14.36
1975	−13	109	14.38
1977	−55	30	10.01
1987	−69	−1555	10.16

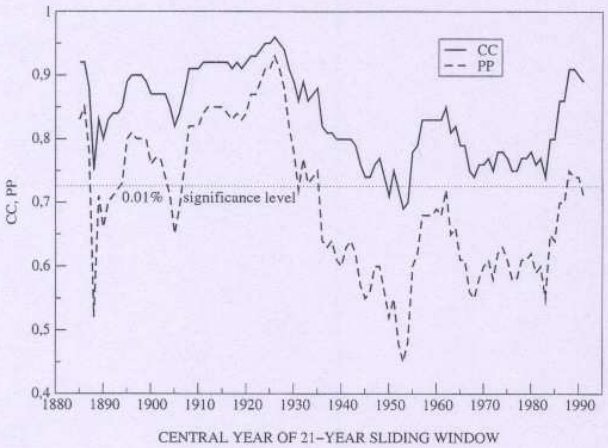

Fig. 7 The 21-year sliding CC and PP are shown for the entire data period. In this case prediction is done using Eq. (3). The *line* for 0.01% significance level of CC is also *drawn*

PP is almost always greater than 0.5. A careful examination reveals an interdecadal variability in the predictability. Similar interdecadal variability was noted by Sahai et al. (2000) when no predictor was used for ISMR prediction but the ISMR time series only. They concluded that the monsoon system inherently has interdecadal predictability variation and its variability influences the variability of many related features around the globe. The very low predictability seen in Fig. 7 from mid 1930s to mid 1950s coincides with a period of very low variability of ISMR (Fig. 1 of Sahai et al. 2002). The period of highest predictability (Fig. 7) and low absolute errors (Fig. 5), from the early 1910s to the late 1920s, is also a period of high variability. When there is a transition from high to low variability, the predictability decreases, as happened from the mid 1920s to the early 1930s. When the transition is from low to high variability, the predictability increases, as from mid 1950s to mid 1960s. The changes in variability of ISMR explain the changes in correlation between the predicted and observed values. The period of lower variability of ISMR is also the period of lower variability of SST in many regions related with the ISMR (Webster et al. 1998). This means that when variability of monsoon changes, SST variability has also changed, which confirms the reliability of the proposed method of prediction.

The physical mechanism that can explain why these 14 predictors capture the ISMR oscillations of different periods is not clear. We can see that in the South Pacific, the positive correlation at R4 (two years prior to the monsoon season) becomes negative after one year and slightly displaced to R6. Similarly, in the northwestern Pacific the positive correlation at R11 (four years prior to the monsoon season) becomes negative and slightly displaced to R8 after one year and one year later it is displaced slightly southward to R10 with same sign. In the North Atlantic, the negatively correlated region at R1, about three and half years prior to the start of monsoon season, changes its sign of CC and is displaced to R2 in one and half years. Thus, the SST anomalies (or their tendency) in these 14 regions, if taken with sufficient time lag, capture the spatial and temporal variation of the SST-ISMR relationship. It seems that these 14 predictors are able to capture biennial oscillation modulated by longer period oscillations. We are not claiming that SST anomalies or tendencies in the 14 selected regions can only influence the ISMR. A selection scheme, different from that presented here, may select a different set of predictors with equal performance. Thus, there may be regions, which are highly influencing ISMR, but if there are, either they are highly correlated with some of the selected regions or their influence is short lived. We can say that, as far as longer term variability and predictability is concerned, these 14 regions contain most of the information which can be obtained from the SST field nine months prior to the start of the monsoon season.

4 Summary and conclusions

The relationship between the SST and the ISMR for the prediction of the later has been examined in a new perspective. This study focuses on two issues: (1) the changing SST-ISMR relationship and (2) the prediction of the ISMR using SST only with long-lead time. It is shown that despite the weakening of the relationship of ENSO-ISMR in recent years, the relationship between SST in some regions of the global oceans and the ISMR is consistent for more than a century and any small variation in this relationship is part of natural oscillations. For the first time the important role of the South Pacific and the North Atlantic Ocean in the ISMR variability has been shown. It is shown that while the relationship between the ISMR and SST in some oceans is decreasing, it is strengthening in other ocean basins. It is shown that, to a large extent, the behavior of the ensuing summer monsoon rainfall in India as a whole can be determined nine months in advance using SST only. The secular variation of the predictor–predictant rela-

tionship and the inter-dependency of predictors do not hinder the performance of the statistical model presented in here. The consistent and skillful prediction for more than a century in the model development period and in the 22 years of the model verification period (these later years were not included in the correlation analysis for identifying significant regions or for computing the regression Eq. 3) cannot have happened just by chance. Thus, the anomalously strong and weak monsoon seasons are not stochastic summer patterns, but are parts of longer period and broader scale circulation patterns which result from the interactions in the ocean–atmosphere coupled system in many seasons in the past. In different decades, about 55–85% of the variance associated with the ISMR is explained by SST alone (Fig. 5) and so only 15–45% is explained by other boundary forcings and internal dynamics.

The 14 predictors used for prediction of the ISMR involve all the ocean basins with sufficient time lag to encompass the evolution of important heat sources and sinks of the coupled ocean–atmosphere system from tropics to extra-tropics. Thus it can be concluded that they capture: (1) the spatial and temporal heterogeneity of the interaction between the global oceans and the monsoon system, and (2) the biennial oscillation of the coupled system modulated by longer-term oscillations. That is why they are able to represent so closely the ISMR. Our study enhances hope in dynamical models like coupled general circulation models, which are designed to capture the evolution of the ocean–atmosphere system for climate prediction. It may also provide a new perspective for the discussion on global warming and the ENSO-monsoon relationships and the consequent changes in precipitation patterns. It is not possible to establish cause and effect relationship from an empirical analysis like this, but this indicates that the coupling and uncoupling of various ocean basins with the monsoon system are parts of the natural variability. Therefore, the observed changes in ENSO-monsoon relationships in recent years can be attributed to the decadal and longer term natural climate variability rather than longer term trends related to anthropogenically induced global warming climate changes.

The central point of the present work is that a long lead and skillful linear statistical prediction of ISMR can be made nine months prior to the start of the monsoon season using SSTs only. This forecast may be further improved by using non-linear statistical methods. Inclusion of other slowly varying boundary forcings, pre-monsoon upper air circulation features and other atmospheric parameters may also improve the forecast. We plan to use artificial neural networks as non-linear statistical methods and pre-monsoon geopotential height in this regard in future.

Acknowledgements We gratefully acknowledge the help of Prof. J. Shukla (COLA, USA) for valuable suggestions and constructive comments on the primary draft of this paper. We had fruitful discussions with R. H. Kripalani, R. Krishnan, K. Rupakumar, K. Krishna Kumar and M. Mujumdar, all from IITM, India. Thanks are also due to two anonymous reviewers whose constructive comments have greatly improved the revised manuscript. Financial support from CNPq, Brazil to one of the authors (A. K. S.) with a Visiting Researcher Fellowship is acknowledged. A.K. Sahai is currently on leave from the Indian Institute of Tropical Meteorology, Pune, India.

References

Blanford HF (1884) On the connection of the Himalayan snowfall with dry winds and seasons of draughts in India. Proc R Soc London 37: 3–22

Bhalme HN, Jadhav SK, Mooley DA, Ramana Murty Bh.V (1986) Forecasting of monsoon performance over India. J Climatol 6: 347–354

Box GEP, Jenkins GM, Reinsel GC (1994) Time series analysis, 3rd edn Prentice Hall, Eaglewood cliffs, N.J., USA, pp 598

Charney JG, Shukla J (1981) Predictability of monsoons. In: Lighthill J, Pearce RP (eds) Monsoon dynamics. (University Press), pp 99–109

Chen T-C, Yen M-C (1994) Interannual variation of the Indian monsoon simulated by the NCAR Community Climate Model: effect of the tropical Pacific SST. J Clim 7: 1403–1415

Clark CO, Cole JE, Webster PJ (2000) Indian Ocean SST and Indian summer rainfall: predictive relationships and their decadal variability. J Clim 13: 2503–2519

DelSole T, Shukla J (2002) Linear prediction of Indian monsoon rainfall. J Clim 15: 3645–3658

Goswami P, Srividya (1996) A novel neural network design for long-range prediction of rainfall pattern. Curr Sci 70: 447–457

Gowariker V, Thapliyal V, Kulshrestha SM, Mandal GS, Sen Roy N, Sikka DR (1991) A power regression model for long-range forecast of southwest monsoon rainfall over India. Mausam 42: 125–130

Hahn DG, Manabe S (1975) The role of mountains in the south Asian monsoon circulation. J Atmos Sci 32: 1515–1541

Harrison M, Soman MK, Davy M, Evans T, Roberson K, Ineson S (1997) Dynamical seasonal prediction of the Indian summer monsoon. Experimental Long-lead Bull 6: 29–32

Ju J, Slingo JM (1995) The Asian summer monsoon and ENSO. Q J R Meteorol Soc 122: 1133–1168

Kripalani RH, Kulkarni A (1997) Climate impact of El-Niñno/ La-Nina on the Indian monsoon: a new perspective. Weather 52: 39–46

Krishna Kumar K, Rajagopalan B, Cane MA (1999) On the weakening relationship between the Indian monsoon and ENSO. Science 284: 2156–2159

Krishna Kumar K, Soman MK, Rupakumar K (1995) Seasonal forecasting of Indian summer monsoon rainfall: a review. Weather 50: 449–467

Krishnamurthy V, Shukla J (2000) Intraseasonal and Interannual variability of Rainfall over India. J Clim 13: 4366–4377

Kulkarni A (2000) A note on the performance of IMD prediction model for ISMR. Paper presented at Annual Monsoon Workshop of Indian Meteorol Soc, Pune Chapter, 20 December, 2000

Manabe S, Hahn DG, Hollaway J (1974) The seasonal variation of tropical circulation as simulated by a global model of atmosphere. J Atmos Sci 32: 43–83

Meehl GA (1987) The annual cycle and interannual variability in tropical Pacific and Indian Ocean region. Mon Weather Rev 115: 27–50

Meehl GA (1997) The south Asian monsoon and the tropospheric biennial oscillation. J Clim 10: 1921–1943

Michaelsen J (1987) Cross-validation in statistical climate forecast models. J Clim Appl Meteorol 26: 1589–1600

Mooley DA, Parthasarathy B (1984) Fluctuations in All-India summer monsoon rainfall during 1871–1978. Clim Change 6: 287–301

Mooley DA, Parthasarathy B, Pant GB (1986) Relationship between all-India summer monsoon rainfall and location of ridge at 500-hPa level along 750 E. J Clim Appl Meteorol 25: 633–640

Navone HD, Ceccatto HA (1994) Predicting Indian monsoon rainfall: a neural network approach. Clim Dyn 10: 305–312

Nicholls N (1983) Predicting Indian monsoon rainfall from sea-surface temperature in the Indonesia-north Australia area. Nature 306: 576–577

Palmer TN, Brankowic C, Viterbo P, Miller MJ (1992) Modeling interannual variations of summer monsoons. J Clim 5: 399–417

Parthasarathy B, Munot AA, Kothawale DR (1994) All India monthly and summer rainfall indices. Theor Appl Climatol 49: 219–224. This data set is available at website http://www.tropmet.res.in/data and referred to as IITM All India Rainfall Series

Ramage CS (1983) The teleconnections and the seige of time. J Climatol 3: 223–231

Rayner NA, Horton EB, Parker DE, Folland CK, Hackett RB (1996) Version 2.2 of the global sea-ice and sea surface temperature data set, 1903–1994. Climate research technical note 4 available from Hadley centre for climate and prediction research, Meteorological Office, London Rd, Bracknell, Berkshire, RG12 2SY, England

Sahai AK, Soman MK, Satyan V (2000) All India summer monsoon rainfall prediction using an artificial neural network. Clim Dyn 16: 291–302

Sahai AK, Grimm AM, Satyan V, Pant GB (2002) Prospects of prediction of Indian summer monsoon rainfall using global SST anomalies. IITM Res Rep, India, RR-093: 1–44

Saji NH, Goswami BN, Vinayachandran PN, Yamagata T (1999) A dipole mode in the tropical Indian Ocean. Nature 401: 360–363

Shukla J, Mooley DA (1987) Empirical prediction of the summer monsoon rainfall over India. Mon Weather Rev 115: 695–703

Shukla J, Paolino DA (1983) The Southern Oscillation and long-range forecasting of the summer monsoon rainfall over India. Mon Weather Rev 111: 1830–1837

Sperber KR, Palmer TN (1996) Interannual tropical rainfall variability in general circulation model simulations associated with the Atmospheric Model Intercomparison Project. J Clim 9: 2727–2750

Soman MK, Slingo J (1997) Sensitivity of Asian Summer monsoon to aspects of sea surface temperature anomalies in the tropical pacific ocean. Q J R Meteorol Soc 123: 309–336

Stone M (1974) Cross-validatory choice and assessment of statistical predictions (with discussion). J R Statis Soc Series B 36: 111–147

Swaminathan MS (1998) Padma Bhusan Prof. P. Koteswaram First Memorial Lecture-23rd March 1998: Climate and sustainable food security. Vayu Mandal 28: 3–10

Thapaliyal V (1981) ARIMA model for long-range prediction of monsoon rainfall in Peninsular India. India Met Dep Monograph Climatology 12/81

Walker GT (1923) Correlation in seasonal variations of weather, VIII, A preliminary study of world weather. Mem Indian Meteorol Dept 24: 75–131

Webster PJ, Yang S (1992) Monsoon and ENSO: selectively interactive systems. Q J R Meteorol Soc 118: 877–926

Webster PJ, Magana VO, Palmer TN, Shukla J, Thomas RA, Yanai M, Yasunari T (1998) Monsoons: processes, predictability, and prospects for prediction. J Geophys Res 13: 14,451–14,510

Wilks DS (1995) Statistical methods in atmospheric sciences. Academic Press, New York, pp 467

Yasunari T (1990) Impact of Indian monsoon on the coupled atmosphere/ocean system in the tropical Pacific. Meteorol Atmos Phys 44: 29–41