# Seasonal forecasting of Bangladesh summer monsoon rainfall using simple multiple regression model

MD MIZANUR RAHMAN[1,*], M RAFIUDDIN[2] and MD MAHBUB ALAM[1]

[1]SAARC Meteorological Research Centre (SMRC), Plot No. # E-4/C, Agargaon, Dhaka-1207, Bangladesh.
[2]Department of Physics, Bangladesh University of Engineering & Technology (BUET), Dhaka-1000, Bangladesh.
*Corresponding author. e-mail: mrahman426@yahoo.com

In this paper, the development of a statistical forecasting method for summer monsoon rainfall over Bangladesh is described. Predictors for Bangladesh summer monsoon (June–September) rainfall were identified from the large scale ocean–atmospheric circulation variables (i.e., sea-surface temperature, surface air temperature and sea level pressure). The predictors exhibited a significant relationship with Bangladesh summer monsoon rainfall during the period 1961–2007. After carrying out a detailed analysis of various global climate datasets; three predictors were selected. The model performance was evaluated during the period 1977–2007. The model showed better performance in their hindcast seasonal monsoon rainfall over Bangladesh. The RMSE and Heidke skill score for 31 years was 8.13 and 0.37, respectively, and the correlation between the predicted and observed rainfall was 0.74. The BIAS of the forecasts (% of long period average, LPA) was −0.85 and Hit score was 58%. The experimental forecasts for the year 2008 summer monsoon rainfall based on the model were also found to be in good agreement with the observation.

## 1. Introduction

In Bangladesh, the agricultural economy with large growing population is closely linked with the performance of its monsoon systems, namely southwest monsoon, active during June–September. The southwest monsoon of Bangladesh, which is also known as the summer monsoon, is an important source of water for Bangladesh during June–September contributing most of the annual rainfall over Bangladesh. Nearly 65–70% of the annual rainfall is received in Bangladesh during the southwest monsoon season (Rahman 2006). Bangladesh critically depends on the southwest monsoon rainfall to grow Aman rice crop. Therefore, long-range forecasting (LRF) of summer monsoon rainfall is a high priority in Bangladesh as there is no dynamic or statistical model to give the LRF

for Bangladesh. An accurate forecast of seasonal summer monsoon rainfall over the country is an increasing demand for decision makers and planners of the country in mitigating any kind of disaster like food crisis and water scarcity.

Seasonal forecasting systems are usually categorized into two types, namely dynamical and hybrid (i.e., combination of dynamical and statistical forecasting). Dynamical techniques are based on forecasts made by combined atmosphere–ocean General Circulation Models (GCMs). Efficient dynamical long range forecasts are provided by the European Center for Medium-range Weather Forecasts (ECMWF). Statistical techniques for long range forecasting make use of the past data especially relationship between the rainfall and other weather/climate related parameters. SST is a key indication because of its relatively gradual rate of

**Keywords.** Summer monsoon rainfall; predictor; forecast; RMSE.

change and the highly effective ocean atmosphere coupling. Multiple techniques usually apply an assortment of mathematical and dynamical techniques, for example, by using mathematical techniques to forecast SSTs which are then used as reviews to a dynamical atmosphere-only GCM to generate long range rainfall forecasts.

The India Meteorological Department (IMD) has been issuing LRF of the southwest monsoon rainfall since 1886. It was, however the extensive and pioneering work of Gilbert Walker (1923, 1924), that led to the development of the first objective model based on statistical correlations between summer monsoon rainfall and antecedent global atmosphere, land and ocean parameters. There are many reviews on the LRF of Indian southwest monsoon rainfall (ISMR) (Thapliyal and Kulshreshtha 1992; Krishna Kumar *et al.* 1995; Rajeevan 2001; Gadgil *et al.* 2005). In a recent study, Gadgil *et al.* (2005) addressed the major problems of the statistical and dynamical methods for LRF of summer monsoon rainfall in view of the recent forecast failures in 2002 and 2004. Their analysis revealed that IMD's operational forecast skill based on statistical methods has not improved over seven decades despite continued changes in the operational models. For the LRF of the ISMR, three main approaches are used. The first is the statistical method, which used the historical relationship between the ISMR and global atmosphere–ocean parameters (Walker 1923; Thapliyal 1982; Singh and Pai 1996; Delsole and Shukla 2002; Rajeevan *et al.* 2005; Pai and Rajeevan 2006). The second approach is the empirical method based on a time series analysis. This method uses only the time series of past rainfall data (Kishtawal *et al.* 2003; Iyengar and Raghukanth 2004) and does not use any predictors. The third approach is based on the dynamical method, which uses general circulation models of the atmosphere and oceans to simulate the summer monsoon circulation and associated rainfall. In spite of its inherent problems, at present, statistical models perform better than the dynamical models in the seasonal forecasting of ISMR (Rajeevan *et al.* 2005). The dynamical models have not shown the required skill to accurately simulate the salient features of the mean monsoon and its interannual variability (Latif *et al.* 1994; Krishnamurti *et al.* 2000; Kang *et al.* 2002; Krishna Kumar *et al.* 2005). Linear forecast of ISMR has been studied by Delsole and Shukla (2002) and found that regression models based on two or three parameters could produce better results than models using large number of predictors. The present study is aimed to develop simple multiple regression model for southwest seasonal monsoon rainfall forecast of Bangladesh using three predictors.

## 2. Data and Methodology

The summer monsoon rainfall over the whole country has been calculated as the area weighted average of the rainfall of 28 raingauge stations during the summer monsoon season (figure 1) in Bangladesh. The long period average (LPA) (1961–2000) of the seasonal summer monsoon rainfall is 1701 mm and the coefficient of variation is about 10%. The Bangladesh summer monsoon rainfall (BSMR) has been expressed in percentage departure of long period average (LPA). The BSMR series during the period 1961–2007 has been used in this study. The monthly CPC Merged Analysis of Precipitation (CMAP) has also been used for understanding the model performance during the period (1979–2007). This data have been collected from Earth System Research Laboratory, NOAA, USA and rainfall data have been extracted point-to-point values through GrADs software.

The monthly mean datasets of various parameters such as sea surface temperature (SST), surface air temperature (SAT) and sea level pressure (SLP) have been used as predictors for summer monsoon rainfall over Bangladesh in this paper. The SST dataset have been taken from the monthly mean NOAA Extended Reconstructed Global Sea Surface Temperature version 2 (ERSST.v2) data at $2° \times 2°$ lat. × long. grid (Smith and Reynolds 2004)
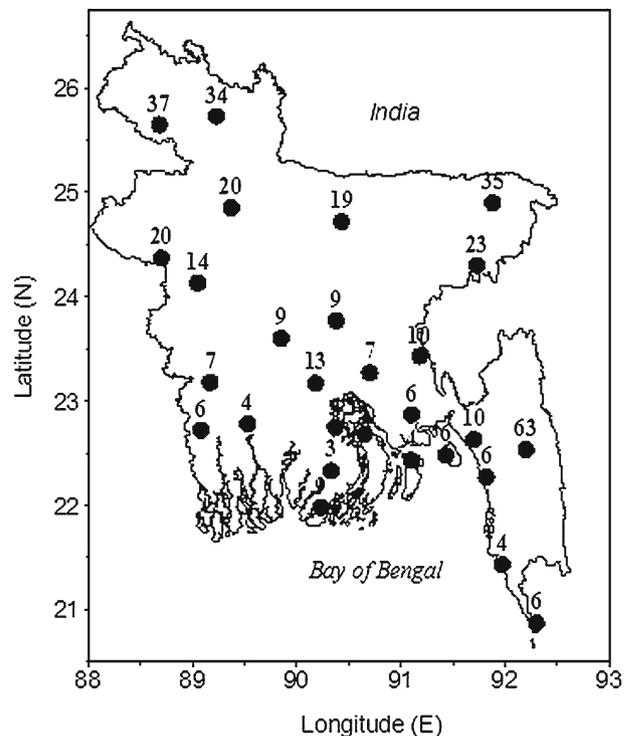


Figure 1. Solid circles represent Bangladesh Meteorological Department (BMD) observation raingauge stations and digits above the solid circles represent the elevation (in meter).

Table 1. *Details of predictors used for the Bangladesh summer monsoon rainfall forecast during the period 1961–2007.*

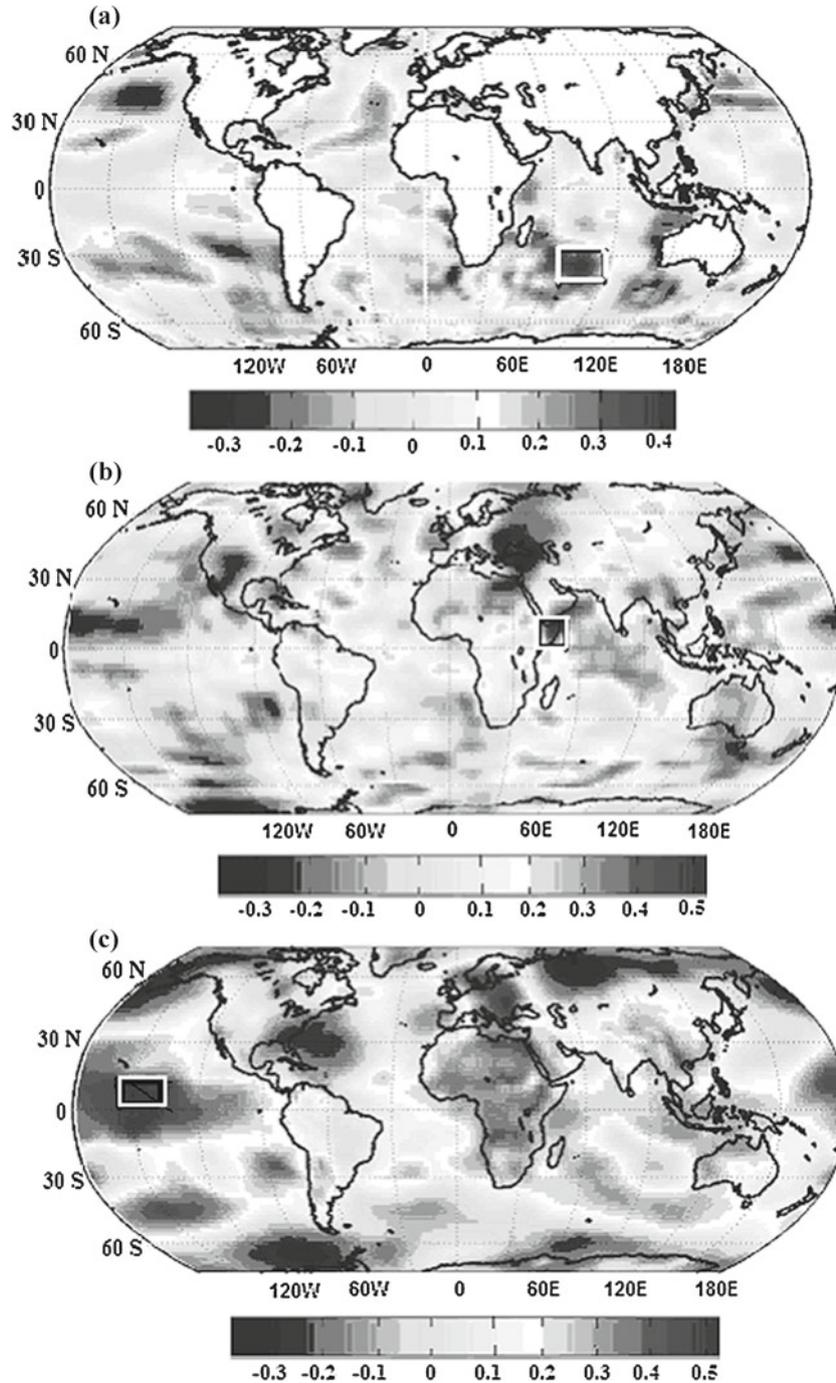| No. | Parameters | Period | Spatial domain | CC with BSMR |
|-----|------------|--------|----------------|--------------|
| P1 | Southwest Indian Ocean SST anomaly | February | 30°–36°S, 74°–78°E | 0.44 |
| P2 | Somalia surface air temperature anomaly | April | 7.5°–12.5°N, 47.5°–50°E | 0.59 |
| P3 | Central Pacific Ocean sea level pressure anomaly | April | 2.5°–7.5°N, 145°–150°W | 0.53 |

Significance at 1% level.



Figure 2. Correlation maps between Bangladesh summer monsoon rainfall and (**a**) SSTs in the month of February, (**b**) SATs in the month of April, and (**c**) SLPs in the month of April. The solid rectangular boxes indicate significant at 5% and 1% level, respectively.

([http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.ersst.html](http://www.esrl.noaa.gov/psd/data/gridded/data.noaa.ersst.html)). The monthly mean SAT and SLP data of NCEP/NCAR reanalysis (Kalnay *et al.* 1996) have been obtained from the NOAA Earth System Research Laboratory (ESRL), Boulder, CO, USA ([http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.derived.surface.html](http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.derived.surface.html)). The spatial resolutions of these datasets are $2.5° \times 2.5°$ lat. $\times$ long. All the above monthly mean datasets have been used for the period 1961–2007.

Correlation fields between BSMR and monthly mean SST, SAT and SLP at each grid point in each month have been computed by using MatLab software and areas with maximum correlation have been identified. The areas of the predictor's domain are shown in table 1.

## 3. Identification of predictors

The aim in this section is to identify predictors for the Bangladesh summer monsoon rainfall, which can then be used in statistical forecast models. The two main requirements for any useful predictors are:

1) a good relationship with the seasonal summer monsoon rainfall,
2) a reasonable lead time (i.e., months to a season).

### 3.1 *BSMR correlation with large-scale variables*

The BSMR is found to be correlated with SST, SAT and SLP during the month of February, April and April (month of pre-monsoon season), respectively, and the correlation maps are shown in figure 2(a–c). The strong shaded regions indicate correlation that is significant at 1% level. Strong positive correlations with SST's are shown in figure 2(a) over southwestern Indian Ocean regions around 35°S. The correlation map indicates persistence from the pre-monsoon season (March–May) leading up to the monsoon season, thus providing the potential for a long range lead forecast. SAT over Somalia is also selected as one of the predictors (figure 2b). This essentially captures the land–ocean gradient that gets set up by the land temperatures, especially during the pre-monsoon season before the start of monsoon. With SLPs, the correlations are strong in the central Pacific tropical region, indicating that a higher than normal tropical pressure tends to enhance the easterlies, thereby increasing the moisture transport to Bangladesh as well as to the south Asia and consequently the rainfall. The solid rectangular boxes in figure 2(a–c) show the region of high correlation from where the predictors will be developed.

## 4. Predictor selection

Based on the correlation with indices and the correlation maps with large-scale variables, predictors with high correlations with the BSMR have been identified (table 1). The selected predictors are:

- SSTs area averaged over 30.0°–36.0°S latitudes and 74.0°–78.0°E longitudes
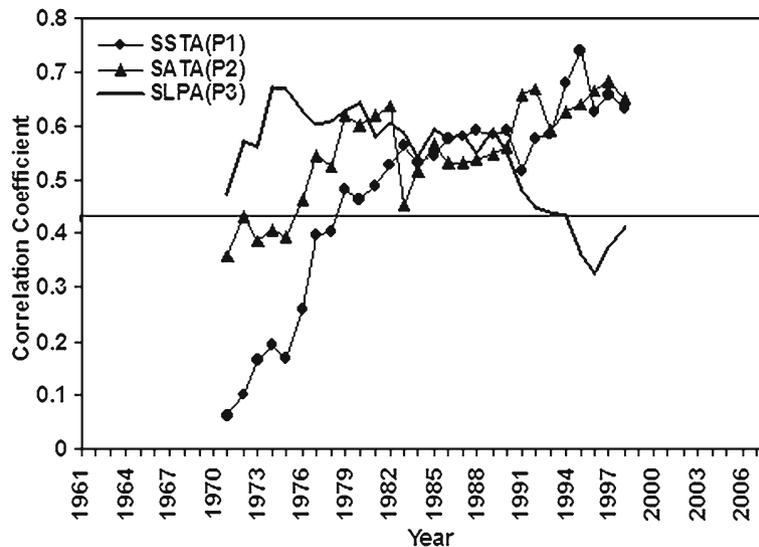- SATs area averaged over 7.5°–12.5°N latitudes and 47.5°–50.0°E longitudes



Figure 3. Sliding window (21-year) correlations between the predictors (SSTA (P1), SATA (P2) and SLPA (P3)) and BSMR during the period 1961–2007. Central year of the sliding window is shown in the figure. 5% significant level is indicated as solid line.

- SLPs area averaged over 2.5°–7.5°N latitudes and 145.0°–150.0°W longitudes

To check the temporal variability of the strength of the predictors to BSMR, moving window correlation are shown in figure 3 with Sea Surface Temperature Anomalies (SSTA), Surface Air Temperature Anomalies (SATA) and Sea Level Pressure Anomalies (SLPA), which are denoted by P1, P2 and P3, respectively. Besides, the predictors show significant correlations with the BSMR at two or three months lead time.

## 5. Forecast models

Typically, a regression (often linear) is fit between the identified predictors and a single dependent variable (i.e., BSMR). The fitted regression is then used to forecast the mean value of the variable. There is an extensive literature for fitting and testing linear regression models, and software is readily available (Helsel and Hirsch 1995). Such models have been widely used for hydro-climate forecasting in the USA (Piechota *et al.* 2001; McCabe and Dettinger 2002) and for Indian monsoon rainfall forecasting (Hastenrath 1988; Krishna Kumar *et al.* 1995). For forecasting a field of a dependent variable, such as precipitation, at several locations from fields of independent variables (e.g., tropical SST, SAT and SLP, etc.), canonical correlation analysis is typically used (Shabbar and Barnston 1996; Ntale *et al.* 2003).

## 6. Verification of the model forecasts

The model forecasts during the period 1977–2007 have been verified using a simple model statistics such as correlation coefficient between observed and predicted BSMR, bias in the model forecasts (BIAS) and RMSE. They are calculated in the following way.

$$R = \frac{\sum \left(Y - \bar{Y}'\right)\left(Y - \overline{Y}\right)}{\sqrt{\sum \left(Y - \overline{Y}'\right)^2 \sum \left(Y - \bar{Y}\right)^2}}$$

$$\text{BIAS} = \frac{\sum (Y' - Y)}{n}$$

$$\text{RMSE} = \sqrt{\frac{\sum (Y' - Y)^2}{n}}$$

where $\overline{Y}$ and $\overline{Y}'$ are the sample averages of the $Y$ and $Y'$, respectively, and $n$ is the number of training period.

The verification statistics used for the category forecasts are (a) hit score (HS), and (b) Heidke skill score (HSS). HS is the proportion of the correct forecasts and is computed as the ratio of forecasts in the correct category to the total number of forecasts. The HSS is the HS adjusted to account the proportion of forecasts that would have been correct by chance in the absence of skill and is computed as $[(3/2) \times \text{HS}] - 1/2$. For a good model, the HSS should be more than 0.33 (i.e., 1/3rd of the forecasts are categorically correct).

The significance test is carried out by using Student's $t$-distribution:

$$t_{\text{cal}} = \text{r}\sqrt{\frac{n-2}{1-r^2}}$$

where $r$ is the correlation coefficient, $n$ is the number of data and $(n-2)$ is the degree of freedom. If $t_{\text{cal}} > t_{0.05}$ or $t_{0.01}$, the correlation coefficient is significant on the basis of a one-tailed test of Student's distribution.

## 7. Results and discussion

Table 1 shows the details of the ensemble members that have been used to construct the model. The correlation field between BSMR and SSTA, SATA and SLPA in the southwest Indian Ocean, over Somalia and central Pacific Ocean, respectively is shown in figure 2(a–c). SSTA in the month of February over southwest Indian Ocean between 30°–36°S latitude and 74°–78°E longitude is positively and significantly correlated with the BSMR (CC = 0.44 and significant at 1% level) and is indicated by the rectangular box as shown in figure 2(a). The correlations are weak over rest of the Indian Ocean and Bay of Bengal. The correlations are insignificant throughout the Indian Ocean and Bay of Bengal from March to May (figure not shown). Sadhuram *et al.* (2007) reported positive and highly significant correlation (CC = 0.61) between Indian Summer Monsoon Rainfall (ISMR) and February SSTA in the northwest Australian region. Clark *et al.* (2000) also reported insignificant correlations between ISMR and SST during pre-monsoon season (March–May) for the periods 1945–1994 and 1977–1995. SATA in the month of April over Somalia is positively and highly significantly correlated with the BSMR (CC = 0.59 and significant at 1% level). The area of SATA is indicated by the rectangular box over Somalia between 7.5°–12.5°N latitude and 47.5°–50°E longitudes as shown in figure 2(b). The correlations are weak over rest of the part of the globe. Similarly, SLPA in the month of April over central Pacific Ocean is positively and highly significantly correlated with the BSMR (CC = 0.53 and significant at 1% level).

The area of surface pressure anomaly over central Pacific Ocean around equator is indicated by the rectangular box between 2.5°–7.5°N latitude 145°–150°W longitudes as shown in figure 2(c). The correlations are insignificant over rest of the part of the world during this month. The correlations are insignificant between BSMR and SLP during pre-monsoon season from March to May (figure not shown here).

Figure 3 shows the 21-year moving correlation coefficients between BSMR and the three predictors (P1, P2 and P3) selected for the model development for the period 1961–2007. Significant level at 5% (CC = 0.43) is indicated as horizontal solid

Table 2. *Forecast verification statistics of the model output computed for the period 1977–2007.*

| Sl. no. | Verification parameters | |
|---|---|---|
| 1 | CC between actual and predicted values | 0.70 |
| 2 | BIAS of the forecasts (% from LPA) | −0.85 |
| 3 | RMSE of the forecasts (% from LPA) | 8.13 |
| 4 | HS (Hit score, %) | 58 |
| 5 | HSS (Heidke skill score) | 0.37 |

Table 3. *Intercorrelation square matrix among the predictors that was used in the regression model during the period 1961–2007.*

| Parameters | SST | SAT | SLP |
|---|---|---|---|
| SST | 1.00 | 0.36 | 0.30 |
| SAT | 0.36 | 1.00 | 0.44* |
| SLP | 0.30 | 0.44* | 1.00 |

*Significance at 5% level.

line. The relationship of most of the predictors with the BSMR is stable (correlation coefficient values and significant at and above 5% level) during the entire period particularly during the recent years as shown in figure 3. The correlations between predictors and BSMR are weak during the earlier period. Clark *et al.* (2000) also found weak correlations before 1970s and strong correlations during the recent decades which were attributed to the climate change in the Indian Ocean after 1976. There are significant intercorrelations between some of the predictors (table 2). Table 3 depicts the intercorrelation square matrix of the three predictors which have been used in the regression model. The correlation between predictors is significant which indicates that the predictors are dependent on each other. Rajeevan *et al.* (2006) suggested an optimum sliding window of 23 years would minimize the RMS errors. Hence, a sliding window of 31 years has been chosen in this study as the training period to develop the linear regression model. The regression equation for the model is given below:

$$BSMR = 91.464P1 + 114.645P2 + 106.408P3 + 1740.41$$

where BSMR is expressed in percentage departure from LPA. The year-to-year performance of observed and predicted values of BSMR is shown in figure 4. In general, there is a good agreement between the observed and predicted values except in the years 1980, 1984, 1986, 1987, 1993, 1994, 1998, 2000, 2003, 2004 and 2006 out of 31 years. This is the first time that an attempt is made to develop a regression model for summer
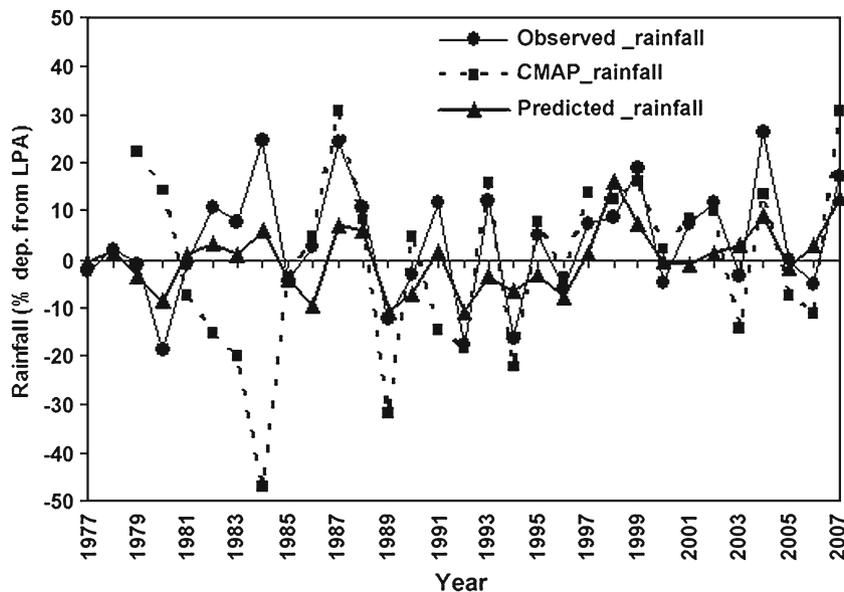


Figure 4. Predicted and observed values of BSMR (% departure from LPA) during the period 1977–2007.

monsoon rainfall in Bangladesh. The model verification statistics of the regression model given in table 2 shows that the RMSEs of the model is 8.13% of LPA. The BIAS of the model is −0.85. The Hit score (HS) and Heidke skill score (HSS) of the regression model is 58% and 0.37, respectively. The HSS is an indicator for the performance of the regression model which is within the range (0.38–0.63) of ensemble multiple linear regressions (EMR) and projection pursuit regression (PPR) models of Rajeevan *et al.* (2006). It is impressive that this simple regression model using only three parameters could capture interannual variability of BSMR satisfactorily. The CMAP data is also used for showing the model performance. During the period 1979–1984, the CMAP data shows more anomalies with opposite sign than the other years over Bangladesh because it may be due to missing data during the above period. CMAP data may be useful in verifying the broad characteristics of the rainfall patterns over the globe (Narkhedkar *et al.* 2010), but it can be seen from figure 4, that it is not at all close to the BSMR temporal analysis either qualitatively or quantitatively.

It is evident that SSTA in Indian Ocean during pre-monsoon season are playing an important role in summer monsoon rainfall over Bangladesh. Warm SST persists before a strong monsoon (Sadhuram *et al.* 2007). In order that SST directly influences the monsoon through evaporation and moisture supply, one would expect the positive correlations to persist through the pre-monsoon season (March–May) leading to the onset of rainfall, which are not shown here. Webster *et al.* (1999) and Clark *et al.* (2000) found that ISMR is strongly correlated with the Indian Ocean SST with a lead time of 3–6 months on Tropical Biennial Oscillation (TBO) time scale. Positive SST leads to the increase of surface moisture (due to enhanced evaporation) and thus a strong monsoon owing to the increased moisture in the region (Chang *et al.* 2001). Hence, warm (cold) SST in the Indian Ocean increases (decreases) the local moisture flux during pre-monsoon and fall preceding summer monsoon, and leads to strong (weak) monsoon. This may be the reason for the observed correlation between BSMR and SSTA shown in the present study.

## 8. Conclusions

From this analysis we conclude that significant correlations exist between BSMR and southwest Indian ocean SST, SLP in the central Pacific region around equator and SAT over Somalia.

- SST anomalies over southwest Indian Ocean in the preceding February are positively and strongly correlated (CC = 0.44, significant at 1% level) with the BSMR. SAT in the month of April anomalies over Somalia is also positively and highly significantly correlated (0.59) with the BSMR. Similarly, SLP anomalies over central Pacific Ocean are positively and significantly correlated (0.53) with the BSMR.
- A good agreement between the observed and predicted values except in the years 1980, 1984, 1986, 1987, 1993, 1994, 1998, 2000, 2003, 2004 and 2006 out of 31 years signifies better performance of the regression model over Bangladesh.
- The model verification statistics of the regression model: the Root Mean Square Error (RMSE) is 8.13% of LPA and the BIAS of the model is −0.85. The Hit Score (HS) and Heidke Skill Score (HSS) of the regression model are 58% and 0.37, respectively. The correlation between the predicted and observed rainfall for the 31 years during the period 1976–2007 is 0.74. Rajeevan *et al.* (2006) reported that the correlation between predicted and observed for 24 years was 0.77–0.84 for their model.
- The correlation between predictors and BSMR are weak during the earlier decades and strong during the recent decades. It is a good signal for LRF of summer monsoon rainfall over these regions.

## References

Chang C P, Harr P and Ju J 2001 Possible roles of Atlantic circulations on the weakening Indian Monsoon rainfall–ENSO relationship; *J. Climate* **14** 2376–2380.

Clark C O, Cole J E and Webster P J 2000 Indian Ocean SST and Indian summer rainfall: Predictive relationships and their decadal variability; *J. Climate* **13** 2503–2519.

Delsole T and Shukla J 2002 Linear prediction of the Indian monsoon rainfall. Centre for Ocean–Land–Atmosphere Studies (COLA); *Tech. Rep.* **46** 52.

Gadgil S, Rajeevan M and Nanjundiah R 2005 Monsoon prediction – why yet another failure?; *Curr. Sci.* **88** 1389–1400.

Hastenrath S 1988 Prediciton of Indian monsoon rainfall: Further exploration; *Bull. Am. Meteor. Soc.* **69** 819–825.

Helsel D R and Hirsch R M 1995 Statistical methods in water resources; New York: Elsevier.

Iyengar R N and Raghukanth S T G 2004 Intrinsic mode functions and a strategy for forecasting Indian monsoon rainfall; *Meteorol. Atmos. Phys.* **90** 17–36.

Kalnay E *et al.* 1996 The NCEP/NCAR 40-year reanalysis project; *Bull. Am. Meteor. Soc.* **77** 437–470.

Kang I-S, Jin K, Wang B, Lau K-M, Shukla J, Krishnamurthy V, Schubert S D, Waliser D E, Stern W F, Kitoh A, Meehl G A, Kanamitsu M, Galin V Y, Satyan V, Park C K and Liu Q 2002 Intercomparison of the climatological variations of Asian summer monsoon precipitation simulated by 10 GCMs; *Clim. Dyn.* **19** 383–395.

Kishtawal C M, Basu S, Patadia F and Thapliyal P K 2003 Forecasting summer monsoon rainfall over India using genetic algorithm; *Geophys. Res. Lett.* **30(23)** 2203, doi: 10.1029/2003GL018504.

Krishna Kumar K, Soman M K and Rupa Kumar K 1995 Seasonal forecasting of Indian summer monsoon rainfall; *Weather* **50** 449–467.

Krishna Kumar K, Hoerling M and Rajagopalan B 2005 Advancing dynamical prediction of Indian monsoon rainfall; *Geophys. Res. Lett.* **32** L08704, doi: 10.1029/2004GL021979.

Krishnamurti T N, Kishtawal C M, La Row T E, Bachiochi D R, Zhang Z, Williford C E, Gadgil S and Surendran S 2000 Multimodel ensemble forecasts for weather and seasonal climate; *J. Climate* **13** 4196–4216.

Latif M, Stockdale T, Wolff J, Burgers G, Maier-Reimer E, Junge M M, Arpe K and Bengtsson L 1994 Climatology and variability in the ECHAM coupled GCM; *Tellus* **46A** 351–366.

McCabe G and Dettinger M D 2002 Primary modes and predictability of year-to-year snowpack variations in the western United States from teleconnections with Pacific Ocean climate; *J. Hydrometeor.* **3** 13–25.

Narkhedkar S G, Sinha S K and Mukhopadhyay P 2010 Rainfall analysis using conventional and non-conventional rainfall information on monthly scale; *Atmosfera* **23(2)** 141–164.

Ntale H K, Gan T Y and Mwale D 2003 Prediction of East African seasonal rainfall using canonical correlation analysis; *J. Climate* **16(12)** 2105–2112.

Pai D S and Rajeevan M 2006 Long range prediction models for the Indian summer monsoon rainfall with different lead time periods based on the global SST anomalies; *Meteorol. Atmos. Phys.* **92** 33–43.

Piechota T C, Chiew F H S, Dracup J A and McMahon T A 2001 Development of an exceedence probability streamflow forecast; *J. Hydrol. Eng.* **6(1)** 20–28.

Rahman M M 2006 A validation of regional climate model simulation with observational data over Bangladesh: M. Phil. thesis, Bangladesh University of Engineering & Technology (BUET) Library, Dhaka, Bangladesh.

Rajeevan M 2001 Prediction of Indian summer monsoon: Status, problems and prospects; *Curr. Sci.* **11** 1451–1457.

Rajeevan M, Pai D S and Anil Kumar R 2005 New statistical models for long range forecasting of southwest monsoon rainfall over India; NCC Research Report No 1/2005, India Meteorological Department, Pune, India.

Rajeevan M, Pai D S and Anil Kumar R 2006 New statistical models for long range forecasting of southwest monsoon rainfall over India; *Clim. Dyn.*, doi: 10.1007/s00382-006-0197-6.

Sadhuram Y, Murthy T and Ramana V 2007 Simple Multiple Regression Model for long range forecasting of Indian summer monsoon rainfall; *Meteorol. Atmos. Phys.*, doi: 10.1007/s00703-007-0277.

Shabbar A and Barnston G A 1996 Skill of seasonal climate forecasts in Canada using canonical correlation analysis; *Mon. Weather Rev.* **124** 2370–2385.

Singh O P and Pai D S 1996 An oceanic model for the prediction of SW monsoon rainfall over India; *Mausam* **47** 91–98.

Smith T M and Reynolds R W 2004 Improved extended reconstruction of SST (1854–1997); *J. Climate* **17** 2466–2477.

Thapliyal V 1982 Stochastic dynamic model for long range forecasting of summer monsoon rainfall in peninsular India; *Mausam* **33** 399–404.

Thapliyal V and Kulshreshtha S 1992 Recent models for long range forecasting of southwest monsoon rainfall over India; *J. Arid Environ.* **43** 239–248.

Walker G T 1923 Correlation in seasonal variations of weather; *Meteorol. Dept.* **24** 75–131.

Walker G T 1924 Correlation in seasonal variations of weather-IV, a further study of world weather; *India Meteorol. Dept. Memoir* **24** 275–332.

Webster P J, Moore A M, Loschnigg J P and Leben R R 1999 Coupled ocean–atmosphere dynamics in the Indian Ocean during 1997–98; *Nature* **40** 356–360.