Climate Dynamics (2000) 16:291–302

A. K. Sahai · M. K. Soman · V. Satyan

All India summer monsoon rainfall prediction using an artificial neural network

Received: 13 March 1999 / Accepted: 31 August 1999

Abstract The prediction of Indian summer monsoon rainfall (ISMR) on a seasonal time scales has been attempted by various research groups using different techniques including artificial neural networks. The prediction of ISMR on monthly and seasonal time scales is not only scientifically challenging but is also important for planning and devising agricultural strategies. This article describes the artificial neural network (ANN) technique with error- back-propagation algorithm to provide prediction (hindcast) of ISMR on monthly and seasonal time scales. The ANN technique is applied to the five time series of June, July, August, September monthly means and seasonal mean (June + July + August + September) rainfall from 1871 to 1994 based on Parthasarathy data set. The previous five years values from all the five time-series were used to train the ANN to predict for the next year. The details of the models used are discussed. Various statistics are calculated to examine the performance of the models and it is found that the models could be used as a forecasting tool on seasonal and monthly time scales. It is observed by various researchers that with the passage of time the relationships between various predictors and Indian monsoon are changing, leading to changes in monsoon predictability. This issue is discussed and it is found that the monsoon system inherently has a decadal scale variation in predictability.

1 Introduction

The agricultural practices and crop yields of India are heavily dependent on the summer monsoon (June to September) rainfall. Out of 142 million ha cultivated land in India, 92 million ha (i.e. about 65%) are under the influence of rainfed agriculture (Swaminathan 1998). Unlike irrigated agriculture, rainfed farming is usually diverse and risk prone. The monsoon season is the principal rain bearing season and in fact a substantial part of the annual rainfall over a large part of the country occurs in this season. Small variations in the timing and the quantity of monsoon rainfall have the potential to impact on agricultural output. Prior knowledge of monsoon behaviour will help Indian farmers and also policy makers, to take advantage of good monsoons and also to minimise crop damage and human hardship during adverse monsoons. In addition to the importance of the mean monsoon seasonal rainfall of a particular year, the intraseasonal variability is also crucial. Even if the mean annual rainfall is normal, the delay in the monsoon onset and the unusual monsoon lulls or breaks in between may cause severe disruption of agricultural activities, hydroelectric power generation and even drinking water supply, as was observed in many years in the past. What farmers need is the location and time-specific information about the monsoon behaviour. As a result, forecasting the monsoon on time scales of weeks to seasons is a major scientific issue in the field of monsoon meteorology. Our purpose is to introduce models for forecasting summer monsoon rainfall on monthly and seasonal time scales, using ANN techniques.

The forecasting of Indian summer monsoon rainfall (ISMR) over the country started more than a century ago (Blanford 1884; Walker 1908, 1910, 1933). After the MONEX program, with the increased understanding of the monsoon phenomenon, more sophisticated empirical statistical models (Thapliyal 1981, 1982; Shukla and Paolino 1983; Mooley et al. 1986; Bhalme et al. 1986; Shukla and Mooley 1987; Gowariker et al. 1989; Gowariker et al. 1991) were developed which have been used with varying degrees of success. However, successful long-range prediction on shorter than seasonal time and regional space scales still remains elusive.

A. K. Sahai (✉) · M. K. Soman · V. Satyan
Climate and Global Modelling Division,
Indian Institute of Tropical Meteorology,
Dr. Homi Bhabha Road,
Pashan, Pune-411 008, India
E-mail: sahai@tropmet.ernet.in

Apart from these statistical models the dynamical prediction models particularly general circulation models (GCMs) have also been used to predict ISMR. A comprehensive study of the performance of thirty three GCMs of the atmosphere by Sperber and Palmer (1996) has shown that the skill of the GCMs to simulate intraseasonal variations in ISMR is rather limited. A review of predictability and the prospects of prediction of monsoon by dynamical modelling has been presented by Webster et al. (1998). Current dynamical models have the advantage over the statistical models in the way that they can provide time evolution and spatial distribution of rainfall at the model's spatial and temporal resolution. Cause-and-effect relationship among various processes (like atmosphere- ocean interaction through sea surface temperature variations, atmosphere- land interactions through albedo, soil moisture and vegetation changes, etc.) represented in the model can also be analysed. The accuracy in long-range prediction of ISMR with GCMs is however not of the desired level as yet and therefore statistical models are still preferred.

The statistical models cited are based on empirical relationships of ISMR with various predictors (teleconnection parameters) detected through the analysis of linear correlations with ISMR. A large number of parameters have been identified and they have tended to fall into four general categories (Krishna Kumar et al. 1995): (1) regional conditions, (2) El Nino- Southern Oscillation (ENSO) indicators, (3) cross-equatorial flow and (4) global/hemispheric conditions. Among these predictors one or more are selected and linear regression, power regression or other statistical methods were developed for long range forecasting of ISMR.

Another statistical approach is the time-series analysis for prediction of ISMR, in which past values from ISMR seasonal mean time series are used to predict future values. Mooley and Parthasarathy (1984) have analysed periodicity in 100 years of data and found two cycles (2.8 year and 14 year). Satyan (1988) has analysed 116 years of data using the phase space approach and found that a strange attractor of dimensionality around 5.1 exists and the system has 12 relevant degrees of freedom. The phase space here is the abstract space spanned by the independent variables used to describe the dynamics of the system and the degrees of freedom is the minimum number of such variables necessary to describe the dynamics. Basu and Andharia (1992) have analysed the time series and developed a model for predicting the ISMR in which the previous seven year values of ISMR were used to predict future values. However, the performance of the model was not as good as predictions by the already available statistical models.

The conventional method of empirical modelling is to assume some form of a fitting function with unknown parameters and subsequently employ the regression or optimisation methods for their evaluation. In recent years, a new mathematical technique, known as artificial neural networks (ANNs) has been extensively used for performing non-linear function approximation and modelling of complex, non-linear, dynamical phenomena (Herz et al. 1991; Muller and Reinhardt 1991; Masters 1993). The utility of ANNs is most evident in disciplines where intrinsic non-linearities in the dynamics masks the development of exactly solvable models. In such cases, a trained neural network is synonymous with a solvable model, and even if the physical understanding is not complete, highly accurate predictions can be made if they are carefully validated. In meteorology, the criteria that the dynamics is inherently non-linear is satisfied, and predictions comprise a central goal. Elsner and Tsonis (1992, 1993) showed that multi-layer feed forward (FF) neural networks could perform better than linear statistical models when dealing with chaotic and random noise systems. So far, ANNs have been used successfully in a variety of meteorological problems ranging from cloud classification and identification (Banker 1994; Peak and Tag 1994), predictions of visibility (Peak and Tag 1989), thunderstorms (Frankel et al. 1990; McCann 1992), tornado (Marzban and Stumpf 1996), wind stress (Tang et al. 1994), SST anomaly (Tangang et al. 1997), El- Nino (Derr and Slutz 1994; Grieger and Latif 1994), ENSO events (Tangang et al. 1998) to ISMR prediction (Navone and Ceccatto 1994; Goswami and Srividya 1996; Venkatesan et al. 1997). As far as ISMR prediction is concerned Goswami and Srividya (1996) have used time series approach (in this method previous values from the time series are used to predict future values) and Venkatesan et al. (1997) have used the predictors approach (in this method various teleconnection parameters detected as predictors of ISMR are used) for ISMR prediction, while Navone and Ceccatto (1994) used both the approaches and proposed a hybrid model.

Here, we have used ANNs in a time series approach with the presumption that ISMR is not only related to previous seasonal mean ISMR values but also with the previous monthly mean (June, July, August and September) rainfall values. There are many reasons to prefer the time series approach over the predictors approach:

1. The physical processes determining the structure and variability of the monsoon are complex and the character and strengths of connection among them are not clearly known. Therefore it is not possible to predict the overall behaviour of the monsoon (evolution on monthly and regional scale) by the models developed on the basis of selecting a few parameters (even if the number be 16) because there are many more parameters (Krishna Kumar et al. 1995; Kripalani et al. 1996, 1997) which could be added to the list. However, in order to safeguard against overfitting that would jeopardise the validity of the regression relationship in the independent data-set, it has been found effective to limit the input to very few (two to three) of the most promising predictors (Hastenrath and Greischar 1993). So the prediction of ISMR by predictors depends on selection of the most promising parameters, whose physical mechanisms and cause-and-effect relationship with the ISMR are not yet clear.

2. It is known that the presence of non-linearities and instabilities make the deterministic prediction of ISMR difficult. However, the regularities of near periodic signals present in the time series can make the dynamical system predictable (Vautard et al. 1992).

3. It was found (Ramage 1983; Parthasarathy et al. 1991; Hastenrath and Greischar 1993; Annamalai 1995; Krishna Kumar et al. 1995) that the relationship between ISMR and some of the parameters (predictors) either ceased to exist or showed considerable decline with passage of time.

4. Kripalani and Kulkarni (1997) have shown that the relationship of El-Nino (a very important predictor) and ISMR is dependent on the decadal variability of the ISMR.

5. There are some studies (Normand 1953; Troup 1965) which suggest that the Indian monsoon stands out as an active, not a passive feature in the global climate system and it could be used as a predictor of various later events rather than to be predicted by earlier events.

6. Rainfall is the end product of all those various atmospheric processes which are related to different predictors. Thus if there are connections between ISMR and different predictors then, all this information should be embedded in the time series itself.

7. The lead time for predicting of ISMR using the predictors method in ANN is limited, because some of the predictors are available only by May. The lead time using time series prediction can be up to eight months. If a suitable model is developed, it may be possible to forecast two to three years in advance which is very much needed for economic planning and policies.

The motivation for communicating the present article is not only to develop models to predict ISMR on monthly and seasonal time scales but also to investigate the question that 'does the decadal variability of the monsoon lead to the decadal variation in the monsoon predictors or does the decadal variation in teleconnections with the monsoon lead to decadal variability of the monsoon?'.

2 ANN – a brief discussion

The empirical and statistical methods used in meteorology and oceanography can be broadly classified into four distinct classes: (1) linear regression (and correlation) analysis, to determine a linear relation (or correlation) between the variables x and z, (2) principal component analysis (PCA), to determine the correlation patterns within a set of variables $x_1, x_2, \ldots, x_n$, (3) canonical correlation analysis (CCA), to determine the linear relations between a set of variables $x_1, x_2, \ldots, x_n$ and another set of variables $z_1, z_2, \ldots, z_n$ and (4) artificial neural networks (ANN), to determine non linear relation between a set of variables $x_1, x_2, \ldots, x_n$ and another set of variables $z_1, z_2, \ldots, z_n$. More recently researchers have examined ANN models from a statistical perspective (Sarle 1994; Hill et al. 1994; Cheng and Titterington 1994; Connor et al. 1994) and pointed out that when ANN geometry, connectivity and parameters are changed then they are equivalent to, or are very close to existing statistical models like simple linear regression, projection pursuit regression, polynomial regression, non-parametric regression, logistic regression, linear discriminate functions, classification trees, finite mixture models, kernel regression, smoothing splines, time series models of auto regression moving average (ARMA) type, non-linear auto regression models (NAR) and non-linear auto regression moving average (NARMA) models. In addition, there are neural network approaches for computing quadratic discriminant rules, for computing PCA and for approximating Bayesian probabilities. Although ANN models are not significantly different from a number of standard statistical models, they are extremely valuable as they provide a flexible way of implementing them (Maier and Dandy 1999). Model complexity can be varied simply by altering the transfer function or network architecture. The difference in the modelling approach between ANN and traditional statistical models, coupled with a lack of strict rules governing the development of the former, are probably the major reasons for the increased popularity of neural networks. As a result, ANN can tackle more complex problems, the dimensionality of the models tends to be much higher, and methodologies are hand tailored to particular applications. As a particular reference to the time series forecasting, ANN models were found to outperform the linear Box- Jenkins models in forecasting time series with short memory and at longer lead times (Tang et al. 1991). During 1991–92, the Santa Fe Institute sponsored the Santa Fe Time series Prediction and Analysis Competition, where all prediction methods were invited to forecast several time series. For every time series in the competition, the winner turned out to be an ANN model (Hsieh and Tang 1998). Thus, since we want to forecast ISMR time series at longer lead times, the ANN model is selected.

The detailed description of feed forward neural network (FFNN) with error-back propagation (EBP) can be found in Herz et al. (1991), Muller and Reinhardt (1991), Masters (1993), Bishop (1994) and Haykin (1994). Neural networks are designed to extract existing patterns from noisy data. The procedure involves training a network with a large sample of representative data (training phase), after which the network is applied to a test data set, comprising the data not included in the training data set (validation or prediction phase) with the aim of predicting the new outputs. A FFNN has an input and an output layer with some number of input and output neurons, characterising, respectively, the independent and dependent variables of an underlying map that is to be learned by the network. There are also one or more hidden layers in between the input and the output layer with some number of neurons on each. The number of hidden layers and neurons is a quantity that must be determined empirically and for each separate situation. Typically, one experiments with a variety of architectures in order to find one that optimises the performance of the network. Due to a large number of parameters and the great flexibility of the ANN, the model output may fit the data very well during the training period yet produce poor forecasts during the test period. This results from overfitting; that is the ANN fitted the training data so well that it fitted to the noise, which of course resulted in poor forecasts over the test period. Therefore when experimenting with different architecture care should be taken that the number of patterns in the training data set should be greater than or equal to the number of connections (weights and biases) so that the network may not get overtrained.

The values β_j of the jth neuron of the first hidden layer is given by

$$\beta_j = f\left(\sum\nolimits_i w_{ij}\alpha_i + \theta_j\right) ,$$

where w_{ij} are the weights connecting the ith input neuron (whose value is α_i) to the jth hidden neuron whose activation threshold(bias) is θ_j and f is a smooth and bounded function called activation function. A similar rule applies to the neurons for the second hidden layer as well as output layer, with the values of the neurons of any layer being determined from the ones on the previous layer.

The weights are randomly initialised. The network learns the relationship between the input-output data in the training set via network training in which the weights are modified by presenting input-output patterns of the training set until a prescribed error criterion is fulfilled. Once the training is complete the weights are frozen and the data from the test set is presented as input to evaluate the performance. It is the performance of the network on this test set that is the true measure of the predictive capability of the network.

3 The network design

The network designed for the present study consists of four layers-input (25 neurons which are previous five year values from each time series of monthly mean ISMR values of June, July, August and September and the seasonal mean), output (1 neuron, the next year value from any one of the time series under consideration) and two hidden (2 and 4 neurons), and a schematic diagram is shown in Fig. 1. Thus the total number of connections 69 (62 weights and 7 biases) is less than 85 (number of training patterns) giving reduced possibility of overtraining. Networks are trained separately for seasonal, June, July, August and September mean time series. Depending on the nature of the activation function two models were developed for each time series and 10 networks were trained, two for each time series. The output for any particular year and particular time series is computed as the average of the two outputs from each models. The details of the models are given:

3.1 Input layer

There are 25 neurons (X_i , i = 1,2,..,25) in the input layer. The arrangement of the neurons in the input layer is in the following order

$$\begin{aligned} &a5_{j-4}, a5_{j-3}, a5_{j-2}, a5_{j-1}, a5_j, a4_{j-4}, a4_{j-3}, a4_{j-2},\\ &a4_{j-1}, a4_j, a3_{j-4}, a3_{j-3}, a3_{j-2}, a3_{j-1}, a3_j, a2_{j-4},\\ &a2_{j-3}, a2_{j-2}, a2_{j-1}, a2_j, a1_{j-4}, a1_{j-3}, a1_{j-2}, a1_{j-1}, a1_j \end{aligned} \tag{1}$$

Here $a5$, $a4$, $a3$, $a2$ and $a1$ denote the scaled observed values (between 0 and 1) of seasonal, September, August, July and June rainfall respectively and suffix j (j = 5,6…,124) denotes year i.e. j = 5 denotes 1875, j = 6 denotes 1876 and so on. The scaling of the seasonal series is described as

$$a5_j = a5'_j/(a5'_j + A) \ , \tag{2}$$

where $a5'_j$ is the actual value in mm for jth year and A = 852.4 mm (which is the mean of the seasonal series). A similar formula is applied for the other series with individual A, which is equal to 171.0 mm, 243.6 mm, 274.4 mm and 163.5 mm for September, August, July and June respectively.

3.2 Hidden layer 1

The first hidden layer has two neurons ($Y1_k, k = 1, 2$) and they are computed as;

$$Y1_k = f1(\Sigma_i X_i \times W1_{i,k} + B1(k)) \tag{3}$$

where $W1_{i,k}$ is the weight connecting the ith input layer neuron to the kth neuron of hidden layer 1, $B1(k)$ is the bias added to kth neuron of hidden layer 1 and the activation function $f1$ is defined as

$$f1(x) = x \ . \tag{4}$$

3.3 Hidden layer 2

There are four neurons ($Y2_m, m = 1, 2, 3, 4$) in the hidden layer 2 and they are computed with the following activation functions for the two different models.

For model I

$$Y2_m = f2(\Sigma_k Y1_k \times W2_{k,m} + B2(m)) \ , \tag{5}$$

and for model II

$$Y2_m = f3(\Sigma_k Y1k \times W2_{k,m} + B2(m)) \tag{6}$$

where $W2_{k,m}$ is the weight connecting the kth neuron of hidden layer 1 to the mth neuron of the hidden layer 2, $B2(m)$ is the bias added to the mth neuron of hidden layer 2. Activation functions $f2$ and $f3$ are defined as follows:

$$\begin{aligned} f2(x) &= x/(1+x) \quad if\ x \geq 0 \\ &= x/(1-x) \quad if\ x < 0 \end{aligned} \tag{7}$$

and

$$f3(x) = 1/(1 + e^{-x}) \tag{8}$$

3.4 Output layer

There is only one neuron (Z) in the output layer and it is computed as

$$Z = f2(\Sigma_m Y2_m \times W3_m + B3) \text{ for model II} \tag{9}$$

and

$$Z = f3(\Sigma_m Y2_m \times W3_m + B3) \text{ for model I} \tag{10}$$

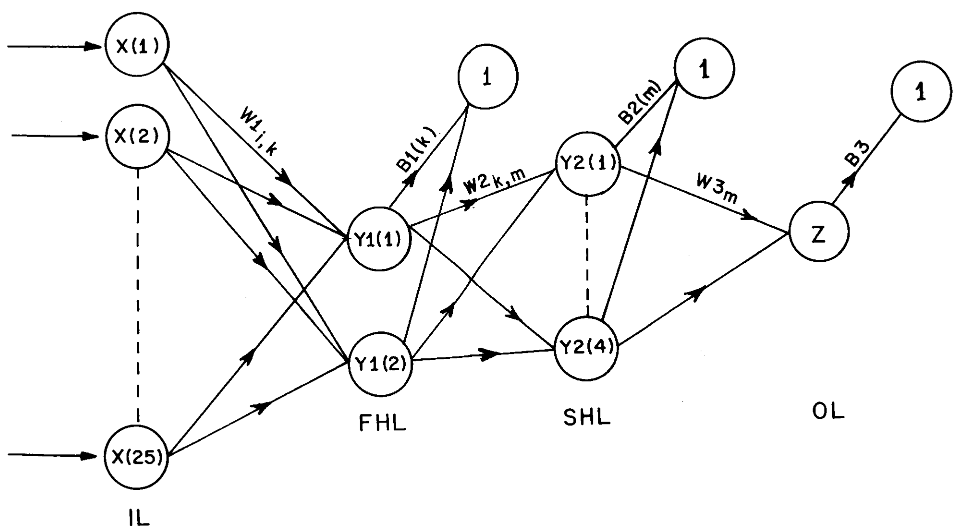

Fig. 1 Schematic diagram of ANN model. *IL*, input layer; *FHL*, first hidden layer; *SHL*, second hidden layer and *OL*, output layer

where $W3_m$ is the weight connecting the mth neuron of the hidden layer 2 to the single neuron in the output layer and $B3$ is the bias added to the output layer neuron.

In the seasonal rainfall series and the set of inputs given as in Eq. (1), the error is evaluated between the computed Z and the scaled observed value $a5_{j+1}$ for each year in the training period. The set of input neurons and their arrangement for network training is the same for each time series and both models, but the scaled observed output value for the error computation is changed, i.e., $a4_{j+1}$, $a3_{j+1}$, $a2_{j+1}$, $a1_{j+1}$ for the September, August, July and June time series respectively.

4 The data set

All India (India taken as a single unit) summer monsoon rainfall data for the period 1871–1994 for June, July, August, September and the whole season is taken from Parthasarathy et al. (1994). The data-set is prepared by area-weighting 306 well-distributed non-hilly stations and a detailed discussion for preparing these data can be obtained in Mooley and Parthasarathy (1984). Kripalani and Kulkarni (1997) have observed epochal variability in the seasonal time series and characterised the periods 1880–1895 and 1930–1963 by above normal rainfall and 1895–1930 and 1963–1990 by below normal rainfall. Therefore, we have divided the data set into training data set (1871–1960) and the test data set (1956–1994) such that data from both epochs, above and below, could be included in the training data set. There is a five year overlap in the data sets because the previous five years values are used for forecasting the next year and thus forecasts are available in the training set from 1876–1960 and also in the test set from 1961–1994.

5 Results and discussions

5.1 The training data set

As already discussed, for each model, one network was trained for each time series and the optimal weights were determined. Then one year predictions (i.e. values of rainfall of t, $t-1$, $t-2$, $t-3$ and $t-4$ years were used to predict $t+1$ year rainfall) were obtained from 1876–1960 in the training set. A graphical representation of the observed and the predicted rainfall values in the training set for seasonal, June, July, August and September is presented in Figs. 2–6. It can be seen that for most of the years, the predicted rainfall values are close to the observation on the seasonal as well as on the monthly time scales. To make the comparison more robust, various statistics, means and standard deviations (SD) of the observed and predicted values of rainfall, correlation coefficient (CC) between observed and predicted values, the root mean square error (RMSE) and the performance parameter (PP) (defined as the ratio of mean square error and the variance of observed values), were calculated. The value of PP = 1 for a network which predicts always the mean value. For a good prediction the predicted means and SDs should be close to the observed means and SDs, RMSE should be small, CC should be closer to 1 and PP should be near to zero. Not only these common statistics, but some more performance-related parameters like PC (percent correct), HSS (Heidke skill score) and biases (BIAS) are also calculated following Perrone and Miller (1985). For these, the rainfall values were categorised into five categories. The categories were defined as follows:

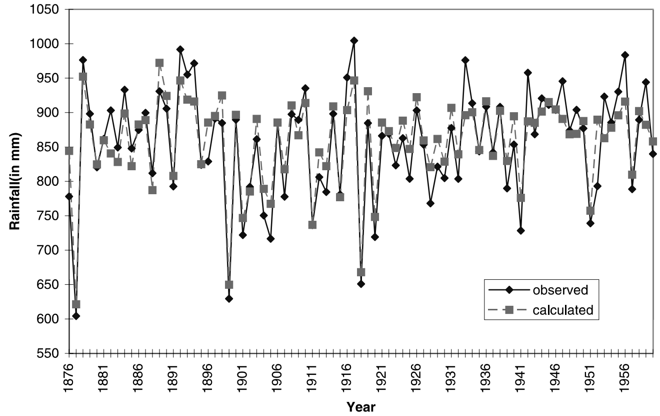

Fig. 2 Comparison of observed and predicted seasonal rainfall values for training data set

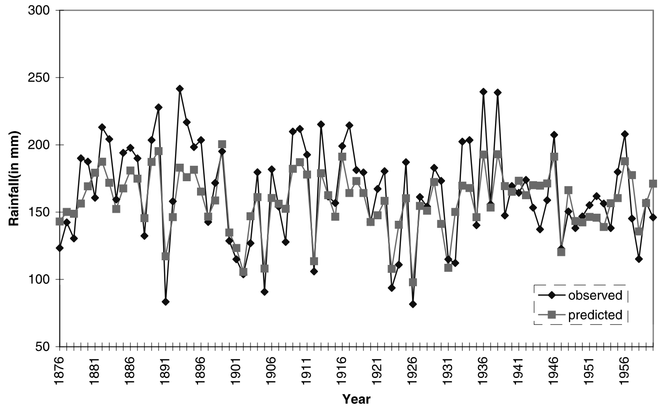

Fig. 3 Comparison of observed and predicted June rainfall values for training data set

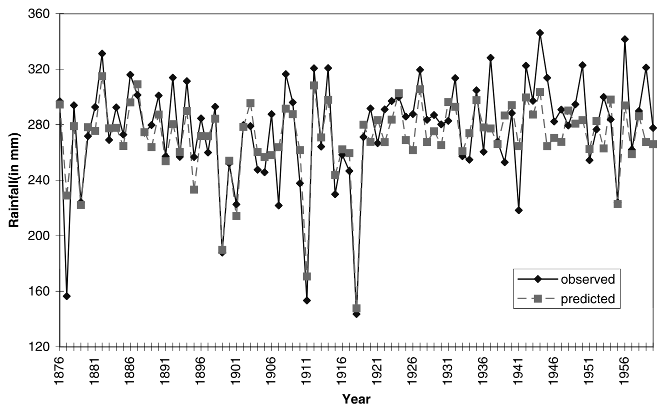

Fig. 4 Comparison of observed and predicted July rainfall values for training data set

Category 1 Anom ≤ -129 mm for season
Anom ≤ -57 mm for months

Category 2 -129 mm $<$ Anom ≤ -43 mm for season
-57 mm $<$ Anom ≤ -19 mm for months

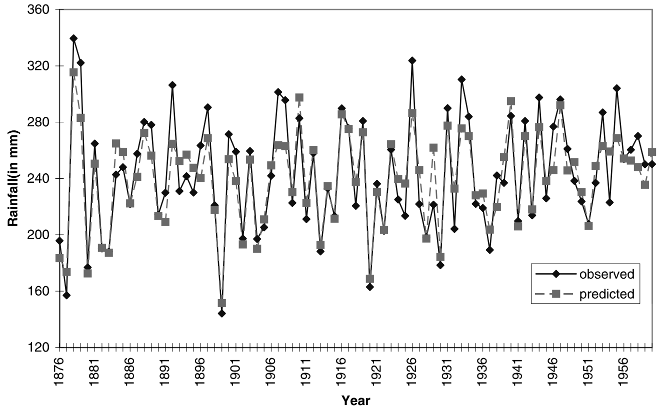

Fig. 5 Comparison of observed and predicted August rainfall values for training data set

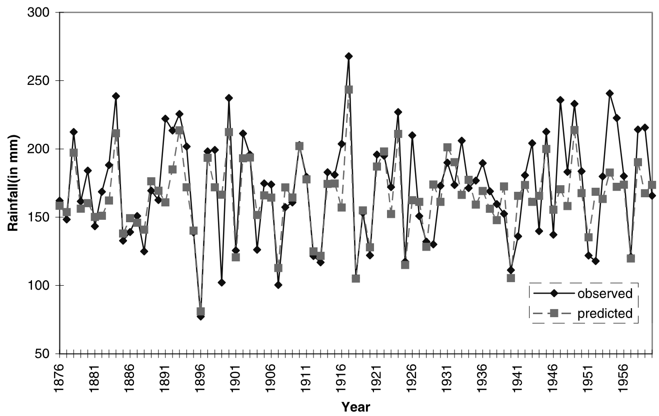

Fig. 6 Comparison of observed and predicted September rainfall values for training data set

Category 3 -43 mm $<$ Anom $\leq$ 43 mm for season
-19 mm $<$ Anom $\leq$ 19 mm for months
Category 4 43 mm $<$ Anom $\leq$ 129 mm for season
19 mm $<$ Anom $\leq$ 57 mm for months
Category 5 Anom $\geq$ 129 mm for season
Anom $\geq$ 57 mm for months

Here Anom is the rainfall anomaly from the long-term observed mean. The values 43 mm and 19 mm are approximately half of the SD of seasonal and monthly series respectively. Based on these categories, contingency tables, which are a prerequisite for calculating PC, HSS and BIAS, were prepared. The format of the contingency table is shown in Table 1 for M categories (here M = 5). The element X_{ij} in the table is the number of times the prediction is in the jth category and the observation is in ith category.

$$X_{iP} = \sum_{j=1}^{j=5} X_{ij}$$

is the total number of predictions when observation is in ith category,

$$X_{Pi} = \sum_{j=1}^{j=5} X_{ji}$$

Table 1 Format of contingency table

Observed category	Predicted category					
	1	2	3	...	M	Total
1	X_{11}	X_{12}	X_{13}	...	X_{1M}	X_{1P}
2	X_{21}	X_{22}	X_{23}	...	X_{2M}	X_{2P}
3	X_{31}	X_{32}	X_{33}	...	X_{3M}	X_{3P}
.	.	.	.	...	.	.
.	.	.	.	...	.	.
.	.	.	.	...	.	.
M	X_{M1}	X_{M2}	X_{M3}	...	X_{MM}	X_{MP}
Total	X_{P1}	X_{P2}	X_{P3}	...	X_{PM}	X_{PP}

is the total number of observations when prediction is in ith category and

$$X_{pp} = \sum_{i=1}^{i=5} X_{iP} = \sum_{i=1}^{i=5} X_{Pi}$$

is the total number of cases.

The calculation of PC, HSS and BIAS is done as follows:

Percent correct (PC): percent correct is calculated as

$$\mathrm{PC} = \left(\sum_{i=1}^{i=5} X_{ii} / X_{pp} \right) \times 100$$

PC gives the percentage of correct forecasts out of the total forecasts, regardless of category. In the present case with five categories of rainfall values, the expected percent correct with random forecasts is 20.

Heidke skill score(HSS): Heidke skill score is calculated as

$$\mathrm{HSS} = \left[\sum_{i=1}^{i=5} X_{ii} - \sum_{i=1}^{i=5} (X_{ip} \times X_{pi}) / X_{pp} \right] \Bigg/ \left[X_{pp} - \sum_{i=1}^{i=5} (X_{ip} \times X_{pi}) / X_{pp} \right]$$

This skill score measures the fraction of possible improvement over random predictions. HSS can vary from -0.33 (for all incorrect predictions) to 1.0 (for all correct predictions). The skill score is expected to be zero when the predictions are random.

Bias(BIAS): The bias for the ith category is calculated as

$$\mathrm{BIAS}_i = X_{pi} / X_{ip}$$

This gives the tendency to overforecast (BIAS > 1.0) or underforecast (BIAS < 1.0) a particular category. A BIAS of 1.0 indicates neither overforecasting nor underforecasting.

All these statistics introduced were calculated for the training data set and are shown in Table 2. The purpose of calculating this spread of verification parameters is that no single verification parameter is ideal for all purposes and we believe that the verification of results have more credibility in situations where a number of parameters point in the same direction. Various parameters shown in Table 2 indicate that the seasonal mean prediction of ISMR is good. The predictions are best for the mean category (Category-3), reasonably predicting the droughts (Catagory-1) but rather bad in predicting floods (Catagory-5). The results on monthly time scales are similar.

5.2 The test data set

The one year predicted values and the two year predicted values (i.e. using t, $t-1$, $t-2$, $t-3$ and predicted $t+1$ year rainfall values, values of $t+2$ year rainfall was predicted) were also calculated. The graphical comparison of observed, one year predicted and two year predicted rainfall values for the test set for the monsoon season means and June, July, August and September monthly means were presented in Figs. 7–11. The statistics calculated for the one year prediction in the training set were repeated for one year and two year prediction in the test set and the results were presented in Table 3. The two year predictions are only tentative because they are very sensitive to the one year prediction, but still some indication of tendencies of ISMR anomalies can be obtained two years in advance. Various parameters listed in Table 3 indicate that the prediction of the independent data set of 34 years is reasonably good on monthly and seasonal time scales.

5.3 Comparison of results for seasonal prediction with earlier works

A comparison of some of our statistics of seasonal predictions can be made with the results of NC (Navone and Ceccatto 1994), who have also reported the results of SM (Shukla and Mooley 1987) and BA (Basu and Andharia 1992). SM and BA have used regression techniques and NC neural networks. These results were reproduced in Table 4. In this table the predictions of SM and NC 2:2:1 (two neurons in input and hidden layer each and one in output layer of ANN model) were based on two predictors, BA and NC 7:4:1 (seven neurons in input layer, four in hidden layer and one in output layer of ANN model) were based on previous seven year values of ISMR itself and NC Hybrid is the combination of NC 2:2:1 and NC 7:4:1. Surprisingly, except for BA all four methods have better predictions in

Table 2 Calculation of various parameters for training set (1876–1960)

Statistics	Season	June	July	August	September
Mean observed (mm)	855.3	164.5	276.4	242.7	171.7
Mean predicted (mm)	858.8	158.3	270.9	240.1	164.7
SD observed (mm)	80.8	37.2	38.3	40.8	39.1
SD predicted (mm)	64.8	23.3	28.2	33.1	28.4
CC	0.91	0.85	0.84	0.91	0.82
RMSE (mm)	34.94	22.34	21.94	17.84	23.95
PP	0.18	0.36	0.33	0.19	0.37
PC	63.52	55.3	60.00	68.23	67.06
HSS	0.46	0.37	0.34	0.57	0.53
Bias (Category 1)	0.50	0.33	1.0	1.2	0.80
Bias (Category 2)	0.61	0.84	0.5	0.68	0.77
Bias (Category 3)	1.42	1.75	1.45	1.48	1.55
Bias (Category 4)	1.0	0.46	0.64	1.1	0.71
Bias (Category 5)	0.0	0.0	0.0	0.14	0.17

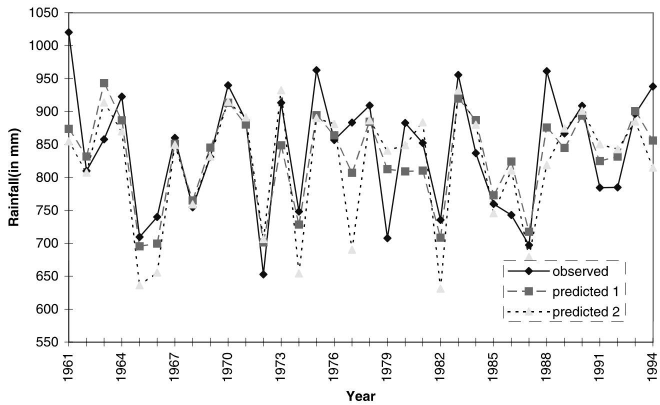

Fig. 7 Comparison of observed, 1 year predicted (predicted 1) and 2 year predicted (predicted 2) seasonal rainfall values for test data set

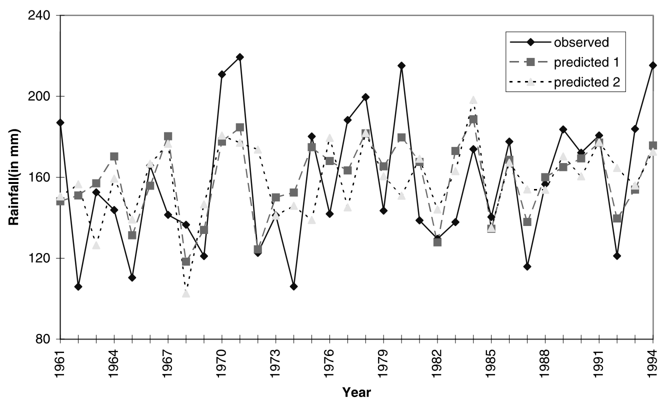

Fig. 8 Comparison of observed, 1 year predicted (predicted 1) and 2 year predicted (predicted 2) June rainfall values for test data set

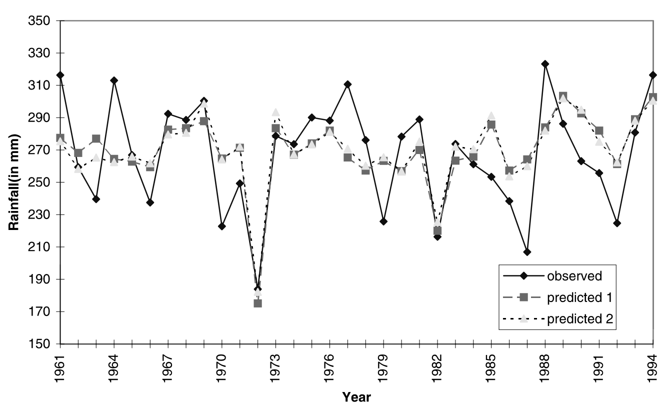

Fig. 9 Comparison of observed, 1 year predicted (predicted 1) and 2 year predicted (predicted 2) July rainfall values for test data set

the test data set than in the training data set. In principle any method should show better prediction in the training data set. Why these results are contradictory will be discussed in the next section. Although the range of training and test data sets of present study is greater than those reported in Table 4, the results are still comparable. Comparison of Figs. 2 and 7 of our study with Figs. 3 and 6 of Goswami and Srividya (1996) show similar skill. In fact, the absolute percentage error shown in Figs. 12 and 13 when compared to Figs. 5 and 8 of their study indicate even better skill in our case. This indicates that the conventional neural networks may perform equally as well as the composite neuron network (proposed by Goswami and Srividya, 1996). We cannot make a reasonable comparison with Venkatesan et al. (1997) because their test set is sparsely distributed.

However all these comparisons are only possible for the seasonal rainfall. On the monthly time scale it seems that our method is the first attempt to develop models to predict monthly rainfall values in monsoon season and the performance cannot therefore be compared with the results of previous models. However, the graphical presentation and statistics reported here show that our prediction is comparable with the observations. The predictive skill of the models is good in August, reasonably good for July and not very good for June and September.

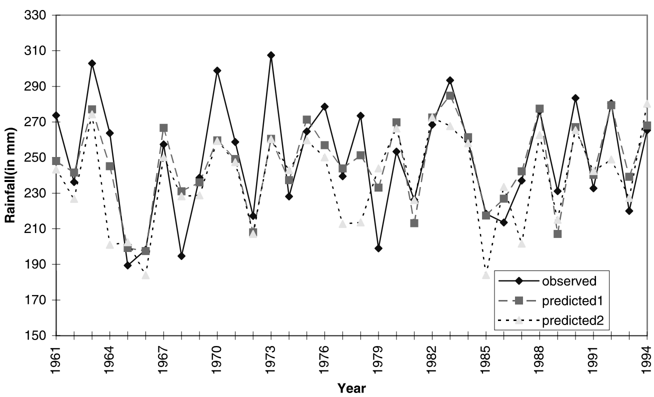

Fig. 10 Comparison of observed, 1 year predicted (predicted 1) and 2 year predicted (predicted 2) August rainfall values for test data set

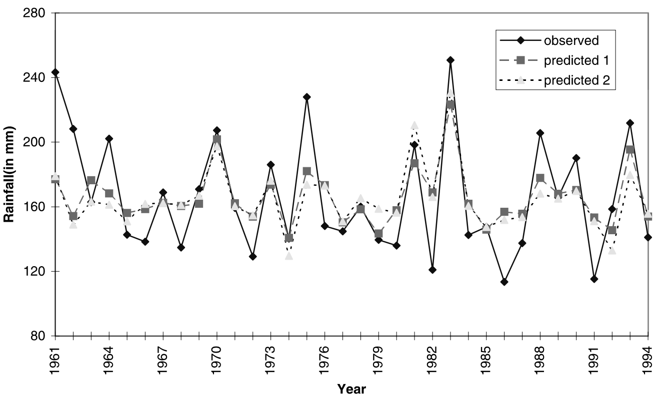

Fig. 11 Comparison of observed, 1 year predicted (predicted 1) and 2 year predicted (predicted 2) September rainfall values for test data set

6 A brief discussion on the monsoon predictability

The reliability of any empirical and statistical prediction model depends largely on the stability of the relationship between the predictant and the predictor. The relationships between predictors and ISMR vary considerably with time, in magnitude as well as in direction (Ramage 1983; Parthasarathy et al. 1991; Krishna Kumar et al.

Table 4 Various statistics reported by NC

	Method	RMSE	PP	CC
Training set 1939–68	SM	50.9	0.42	0.76
	BA	56.5	0.52	0.70
	NC 2:2:1	50.6	0.41	0.77
	NC 7:4:1	61.2	0.56	0.62
	NC hybrid	49.0	0.39	0.78
Test set 1969–84	SM	41.3	0.24	0.89
	BA	66.7	0.61	0.68
	NC 2:2:1	33.6	0.16	0.94
	NC 7:4:1	41.7	0.25	0.87
	NC hybrid	26.7	0.10	0.95

Table 3 Calculation of various parameters for test set (1961–1994)

Statistics	Season	June	July	August	September
Mean observed (mm)	840.0	157.6	267.1	249.4	166.2
Mean predicted (mm)	829.5 (819.6)	159.0 (158.4)	269.6 (270.3)	246.5 (237.9)	165.8 (164.5)
SD observed (mm)	90.4	32.9	33.6	32.0	36.0
SD predicted (mm)	69.2 (91.9)	18.8 (18.5)	22.6 (22.0)	24.0 (28.4)	17.2 (19.0)
CC	0.81 (0.69)	0.67 (0.35)	0.63 (0.65)	0.83 (0.67)	0.74 (0.66)
RMSE (mm)	54.24 (73.5)	24.76 (31.67)	26.14 (25.61)	18.33 (26.91)	26.02 (27.49)
PP	0.36 (0.66)	0.56 (0.92)	0.60 (0.58)	0.33 (0.70)	0.52 (0.58)
PC	50.00 (50.00)	47.06 (26.47)	52.94 (52.94)	55.88 (44.12)	38.24 (41.12)
HSS	0.31 (0.32)	0.25 (0.0)	0.13 (0.18)	0.34 (0.19)	0.16 (0.26)
Bias (Category 1)	1.25 (1.75)	0.0 (0.2)	0.33 (0.33)	[b] ([a])	0.0 (0.0)
Bias (Category 2)	0.63 (0.38)	0.67 (0.40)	0.13 (0.25)	0.75 (0.88)	0.33 (0.53)
Bias (Category 3)	1.82 (1.73)	2.75 (3.25)	1.76 (1.53)	1.5 (1.5)	3.25 (2.88)
Bias (Catagory 4)	0.4 (0.5)	0.22 (0.11)	0.33 (0.83)	0.83 (0.58)	0.38 (0.24)
Bias (Category 5)	0.0 (0.0)	[b] ([b])	[b] ([b])	0.0 (0.0)	0.0 (0.5)

[a] Category predicted but not observed
[b] Category neither observed nor predicted
Quantities in ( ) are for 2 year prediction

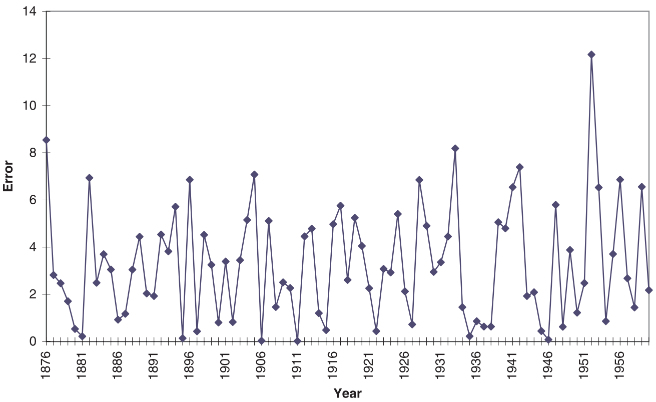

Fig. 12 Absolute % error in seasonal rainfall prediction in training data set

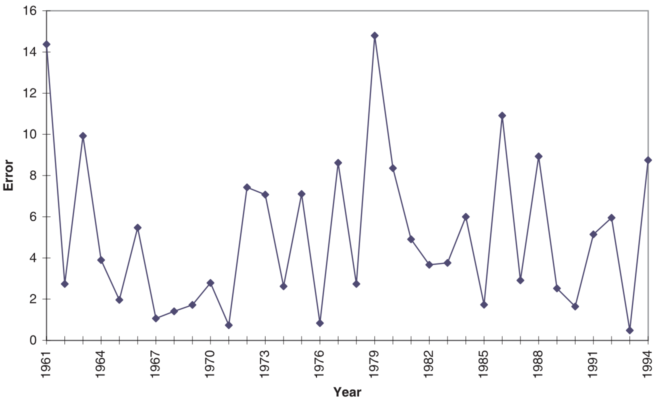

Fig. 13 Absolute % error in seasonal rainfall prediction in test data set

1995). Ramage (1983), while commenting on this issue, doubts that this may be due to absence of real physical relationships between predictors and predictant. He further pointed out that in year 1963–64 apparent reversals were observed in the statistical relationships between Canton Island sea surface temperature and Aleutian Low intensity and between the Southern Oscillation Index and Indian rainfall. In the same years deviation from normal of annual average 500–1000 mb thickness in high northern latitudes changed sign as did the Pacific surface wind anomalies. Fu and Fletcher (1988) identified large climatic variations in the Asian monsoon region by examining the zonal and meridional components of surface-winds during the monsoon months (June, July and August). They concluded that the relative dominance of the two wind components had been undergoing significant decadal- scale variability. They identified distinct climatic regimes in wind components with changes around the years 1875, 1900 and 1940. They characterised the period 1875–1900 as a "meridional monsoonal period" and the period 1900–1940 by the "zonal monsoonal period" over the Indian ocean. By examining the relationship of Bombay sea-level pressure and ISMR, Parthasarathy et al. (1991) have concluded that the Indian summer monsoon had passed through two meridional (1871–1900, 1941–90) and two zonal (1847–70, 1901–1940) circulation regimes during the past 150 years. As already mentioned, Kripalani and Kulkarni (1997) have observed epochal variability in the seasonal series and characterised the periods 1880–1895 and 1930–1963 as above normal rainfall periods and 1895–1930 and 1963–1990 as below normal rainfall periods. They concluded that the effect of El-Nino is largely dependent on below and above normal decadal epochs of ISMR.

Hastenrath and Greischar (1993) have explored the predictability of ISMR from pre- monsoon circulation indices from observations during 1939–91. They found that the relationships have varied considerably over the past half century, with strongest associations during 1950–80, and a drastic weakening of the relationships in the 1980s. They assessed the performance of various statistical models. All prediction models had 1939–68 as training period, while the data from 1969 onwards were reserved for verification on the independent data set. For the period 1969–80 of this independent data set, all the four models explained about three quarters of the observed inter-annual variance of ISMR. However, the performance of all four models deteriorated drastically after 1980, in fact so severely that the explained variance for the period 1969–1989 as a whole dropped to a quarter. Annamalai (1995) suggests that the presence of decadal-scale oscillations in the predictors themselves may possibly be responsible for the instability in the relationship between ISMR and its predictors.

It is well known that the presence of non-linearities and instabilities make deterministic prediction of ISMR difficult. On the other hand, the regularities of near periodic signals present in the time series suggests predictability in the dynamical system. Therefore, to extend the debate on predictability of ISMR we have adopted time series prediction using ANN and trained the whole data (1871–1994) in neural network models and the predicted values were obtained as described in earlier sections. The correlation coefficient (CC) over a sliding window of 11 years and 31 years between observed and predicted values were calculated and presented in Fig. 14. The 11 year and 31 year windows are selected in order to bring out features of decadal scale and climate scale variations. The odd numbers are chosen so that middle points of the window can be easily identified for plotting. A careful observation of Fig. 14 suggests that, in general, below normal rainfall epochs are associated with high CC and above normal rainfall epochs with low CC. If the CC between observed and predicted values of rainfall can be taken as a measure of predictability, we may conclude that when above normal epoch and meridional regimes are in phase (1883–1893 and 1947–1963) monsoon predictability is low and when below normal epoch and zonal regimes are in phase (1895–1924) predictability is high. From 1964–1990 when below normal epochs and zonal regimes are not in phase, the first half of this period was dominated by below normal epochs resulting in high predictability and

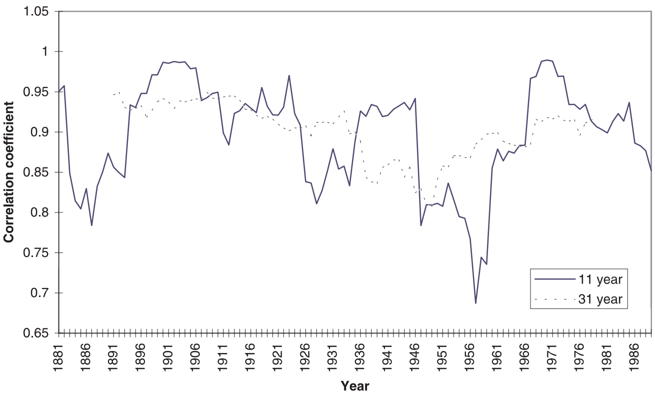

Fig. 14 Correlation coefficients between observed and predicted rainfall values over sliding windows of 11 year and 31 year

the latter half was dominated by meridional regimes of circulations resulting in low predictability. Thus the inherent lack of skill in the prediction of the ISMR during the training set of SM and NC reported in Table 4 with respect to the test set may be the reason why better predictions were not produced in the training data set.

We have not used any teleconnection relationships of monsoon and predictors to assess the predictability. This means that the decadal variation in the predictability may not be due to decadal variability in predictors as suggested by Annamalai (1995). Now let us explore the question, 'does the decadal variability of the monsoon lead to the decadal variation in the monsoon predictors or does the decadal variation in teleconnections with the monsoon lead to decadal variability of the monsoon?'. It is a well-established fact that enormous amounts of latent heat are released into the atmosphere over the Asian summer monsoon region through the processes of cloud formation and rain during a period of barely four months. The interannual variations in these amounts of latent heat are quite large. The variations in the latent heat energy released into the atmosphere may lead to the variations in the out-going longwave radiation at the surface and lower troposphere, consequently changing the circulation patterns. It is to be noticed that the variation in the characteristics of wind components and various other parameters around 1875, 1900, 1940, 1990 as identified by Fu and Fletcher (1988) and Parthasarathy et al. (1991) and variation in relationships of various circulation features around 1964 as identified by Ramage (1983) surprisingly occurred after the monsoon had entered into different epochs around 1895, 1930, 1963 and 1990 as identified by Kripalani and Kulkarni (1997). Thus, the variation in the decadal behaviour of the monsoon may lead to the variation in teleconnection relationships. This result seems to favour the conclusion of Normand (1953) that the Indian monsoon stands out as an active, not a passive feature in the world weather and it could be used as a predictor of various later events rather than to be predicted by earlier events. Also it appears that the monsoon system oscillates between two climatic states of below normal and above normal epochs and during transition from one state to another the associated circulation patterns in the tropics change. These changes feedback on the monsoon system to drive further forward into another state. This transition from one state to another some times may be rapid or slow depending upon the feedback strengths and time scales. Furthermore, it appears that the decadal variation of monsoon rainfall (above and below normal epochs) leads to the decadal variation of various teleconnections with the monsoon (not only the El-Nino and ISMR relationship as identified by Kripalani and Kulkarni 1997) and in turn the predictability of the monsoon.

Our study indicates that the monsoon system inherently has a decadal scale predictability variation, for which the cause is not clear at present. For the prediction of ISMR more emphasis should be given to predictions based on time series prediction (particularly on monthly rainfall series) so that greater lead-time, compared to the conventional methods based on teleconnections for which data of April–May months are used, could be obtained.

7 Conclusion

Artificial neural networks can be used to predict the seasonal and monthly mean rainfall over the whole of India, using only rainfall time series as inputs. Various verification statistics have shown that prediction skill is good. Since only the previous five years monthly and seasonal mean rainfall values were used to predict next year values, these predictions have much longer lead-time (8 months) compared to conventional statistical methods that use teleconnection parameters. There is also useful skill for a two-year lead-time.

The present study has, for the first time, shown that monthly rainfall during the monsoon season can be predicted with sufficient lead-time and good skill. This indicates that it may be possible to develop a suitably configured neural network for predicting monsoon rainfall on suitably defined regional scale.

Many authors have studied the epochal behaviour of the monsoon system and its relationship with the teleconnection parameters. Our study shows that the monsoon system has decadal scale variability in predictability, independent of any teleconnection as we have used only time series of monsoon rainfall in the study. Thus, the monsoon stands out as an active system whose variability influence the variability of many related features around the globe.

Acknowledgements We are thankful to Dr. G. B. Pant, Director of Indian Institute of Tropical Meteorology (IITM) for encouragement and Dr. R. H. Kripalani, IITM, for valuable suggestions and discussions. Constructive comments by three anonymous reviewers helped to improve the revised manuscript. This work was carried out under Climate Research Project funded by Department of Science and Technology, Government of India.

References

Annamalai H (1995) Intrinsic problems in seasonal prediction of the Indian summer monsoon rainfall. Meteorol Atmos Phys 55: 61–76

Banker RL (1994) Cloud classification of a AVHRR imagery in maritime regions using a probabilistic neural network. J Appl Meteorol 33: 909–918

Basu S, Andharia HI (1992) The chaotic time-series of Indian monsoon rainfall and its prediction. Proc Indian Acad Sci (Earth Planet Sci) 101: 27–34

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

Bishop CM (1994) Neural networks and their applications Rev Sci Instr 65: 1803–1832

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

Cheng JT, Titterington DM (1994) Neural networks: a review from a statistical perspective. Stat Sci 9(1): 2–54

Connor JT, Martin RD, Atlas LE (1994) Recurrent neural networks and robust time series prediction. IEEE Trans Neural Netw 5(2): 240–254

Derr VE, Slutz RJ (1994) Prediction of El- nino events in the pacific by means of neural networks AI Appl 8: 51–63

Elsner JB, Tsonis AA (1992) Non-linear prediction, chaos, and noise. Bull Am Meteorol Soc 73: 49–60

Elsner JB, Tsonis AA (1993) Corrigendum to ''Non-linear prediction, chaos, and noise''. Bull Am Meteorol Soc 74: 243

Frankel D, Schiller I, Draper JS, Barnes AA (1990) Investigation of the prediction of lightning strikes using neural networks. 16th Conf on Severe Local Storms, Kananaskis Park, AB, Canada, American Meteorological Society pp 7–11

Fu C, Fletcher J (1988) Large signals of climatic variations over the ocean in the Asian monsoon region. Adv Atmos Sci 5: 389–404

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, Sarker RP, Mandal GS, Sikka DR (1989) Parametric and power regression models – new approach to long range forecasting. Mausam 40: 115–122

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

Grieger B, Latif M (1994) Reconstruction of the El- Nino attractor with neural networks. Clim Dyn 10: 267–276

Hastenrath S, Grieschar L (1993) Changing predictability of Indian monsoon rainfall anomalies? Pro Ind Acad Sci (Earth Planet Sci) 102: 35–48

Haykin S (1994) Neural networks: a comprehensive foundation. Macmillan, New York

Herz J, Krough A, Palmer RG (1991) Introduction to the theory of neural computation. Addison-Wesley, Reading, Mass

Hill T, Marquez L, O' Connor M, Remus W (1994) Artificial neural network models for forecasting and decision making. Int J Forecast 10: 5–15

Hsieh WW, Tang B (1998) Applying neural network models to prediction and data analysis in meteorology and oceanography. Bull Am Meteorol Soc 79: 1855–1870

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

Kripalani RH, Kulkarni A, Singh SV (1997) association of the Indian summer monsoon with the Northern Hemisphere mid- latitude circulation. Int J Climatol 17: 225–240

Kripalani RH, Singh SV, Vernekar AD, Thapliyal V (1996) Empirical study on Nimbus-7 snow mass and Indian summer monsoon rainfall. Int J Climatol 16: 23–34

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

Maier HR, Dandy GC (1999) Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications, Environ Model Software (in press)

Marzban C Stumpf GJ (1996) A neural network for tornado prediction based on Doppler radar-derived attributes. J Appl Meteorol, 35: 617–626

Masters T (1993) Practical neural network recipes in C++. Academic Press, New York

McCann DW (1992) A neural network short-term forecast of significant thunderstorms. Weather Forecast 7: 525–534

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 mb level along 750 E. J Clim Appl Meteorol 25: 633–640

Muller B, Reinhardt J (1991) Neural networks: an introduction, vol 2, The physics of neural networks series. Springer, Berlin Heidelberg New York

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

Normand C (1953) Monsoon seasonal forecasting QJR Meteorol Soc 79: 463–473

Perrone TJ, Miller RG (1985) Generalised experimental markov and model output statistics: a comparative verification. Mon Weather Rev 113: 1524–1541

Pathasarathy B, Munot AA, Kothawale DR (1994) All- India monthly and seasonal rainfall series: 1871–1993. Theor Appl Climatol 49: 217–224

Parthasarathy B, Rupa Kumar K, Munot AA (1991) Evidence of secular variations in Indian summer monsoon rainfall–circulation relationships. J Clim 4: 95–126

Peak JE, Tag PM (1989) An expert system approach for prediction of maritime visibility obstruction. Mon Weather Rev 117: 2641–2653

Peak JE, Tag PM (1994) Segmentation of satellite imagery using hierarchical thresholding and neural networks. J Appl Meteorol 33: 605–616

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

Sarle WS (1994) Neural networks and statistical models. Proc 19th Ann SAS Users Group International Conference, pp 1538–1550

Satyan V (1988) Is there an attractor for Indian summer monsoon? Proc Ind Acad Sci (Earth Planet Sci) 97: 49–52

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

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

Tang B, Flato GM, Holloway G (1994) A study of Arctic sea ice and sea- level pressure using POP and neural network methods. Atmos Ocean 32: 507–529

Tang Z, Almeida C, de Fishwick PA (1991) Time series forecasting using neural networks versus Box-Jenkins methodology. Simulations 57: 303–310

Tangang FT, Hsieh WW, Tang B (1997) Forecasting the equatorial sea surface temperatures by neural network models. Clim Dyn 13: 135–147

Tangang FT, Tang B, Monahan AH, Hsieh WW (1998) Forecasting ENSO events: a neural network – extended EOF approach. J Clim 11: 29–41

Thapaliyal V (1981) ARIMA model for long-range prediction of monsoon rainfall in Peninsular India. India Meteorol Dept Monogr Climatology, 12/81
Thapaliyal V (1982) Stochastic dynamical model for long-range prediction of monsoon rainfall in peninsular India. Mausam 33: 399–404
Troup AJ (1965) The Southern Oscillation. QJR Meteorol Soc 91: 490–506
Vautard R, Yiou P, Ghill M (1992) Singular spectrum analysis: a toolkit for short, noisy and chaotic series. Physica D 58: 95–126
Venkatesan C, Raskar SD, Tambe SS, Kulkarni AD, Keshavamurty RN (1997) Prediction of all India summer monsoon rainfall using error-back propogation neural networks. Meteorol Atmos Phys 62: 225–240
Walker GT (1908) Correlation in seasonal variation of climate (Introduction). Mem Ind Meteorol Dept 20: 117–124
Walker GT (1910) On the meteorological evidence for supposed changes of climate in India. Mem Ind Meteorol Dept 21: 1–21
Walker GT (1933) Seasonal weather and its prediction. Brit Assoc Adv Sci, 103: 25–44
Webster PJ, Magana VO, Palmer TN, Shukla J, Thomas RA, Yanai M, Yasunari, T (1998) Monsoons: Processes, predictability, and prospects for prediction. J Geo Phys Res 13: 14 451–14 510