ISSN 0252-1075
Research Report No. RR-088

Contribution from
Indian Institute of Tropical Meteorology

REDUCTION OF AGCM SYSTEMATIC ERROR BY ARTIFICIAL NEURAL NETWORK : A NEW APPROACH FOR DYNAMICAL SEASONAL PREDICTION OF INDIAN SUMMER MONSOON RAINFALL

by

A.K. SAHAI
and
V. SATYAN

PUNE – 411 008
INDIA

DECEMBER 2000

CONTENTS

Reduction of AGCM Systematic Error by Artificial Neural Network: A New Approach for Dynamical Seasonal Prediction of Indian Summer Monsoon Rainfall

A. K. Sahai and V. Satyan

ABSTRACT

The ability of state-of-the-art Atmospheric General Circulation Models (AGCMs) to extract all possible information from the data available (i.e., initial conditions and boundary conditions etc.) and produce seasonal forecasts is limited by several sources of uncertainties and errors. One way to tackle this problem is to use ensemble forecasts and multi-model multi-ensemble. The other approach is to combine statistical and dynamical models. Some studies using this approach are based on empirical orthogonal function analysis, canonical correlation analysis and singular value decomposition. These methods are designed to identify the linear combinations of variables in one field that are most strongly correlated with linear combinations of variables in another field. If the structure of the data is inherently linear then these methods are best in feature extraction. However, if the data contain nonlinear structure, it will not be detected. Recently Artificial Neural Network (ANN) has been used for nonlinear principal component analysis to extract non-linear features. The objective of this paper is to combine the nonlinear feature extracting capability of ANNs for reduction of systematic errors in simulation of AGCMs.

The most challenging, and of course, a litmus test for any GCM is to simulate the variability of the Indian summer monsoon rainfall (ISMR). It has been shown that models exhibit greater fidelity in capturing the large-scale dynamical fluctuations than the regional scale rainfall variation. Therefore AGCM produced parameters, which include dynamical parameters along with rainfall, are selected and ANN models have been developed. The region selected is from 5°N to 35°N and from 65°E to 95°E. Various skill-evaluating parameters were calculated do demonstrate the skill of the method. It is shown that the corrected AGCM forecasts are better, beyond reasonable doubt, than that of random and climatological forecasts. The improvement is also in accordance with El Nino Southern- Oscillation (ENSO) and ISMR relationship. Some improvements are also reported on real time forecasts.

1. Introduction

Asian monsoon is a major component of the general circulation, which influences over one third of the tropical region for several months during northern summer. Most parts of south and East Asia receive bulk of their annual rainfall during the summer monsoon season. Therefore forecasting space-time distribution of this rainfall is intimately linked to the economy, hydropower generation, drinking water availability etc. of more than a billion people.

Indian summer monsoon, which is a part of Asian monsoon, is distinct in many aspects and particularly due to orographic features in the north by the Himalayas and along West Coast by Western Ghats. The agricultural practices and crop yields of India are heavily dependent on the rainfall in this season. About 65% of the cultivated land in India is under the influence of rainfed agriculture (Swaminathan, 1998). Unlike irrigated agriculture, rainfed agriculture is usually diverse and risk prone. Even the small variations in the timing and the quantity of monsoon rainfall have the potential to impact on agricultural output. Every year we hear of terrible human suffering as long periods of drought or floods devastate whole country and it is very difficult to deal with such problems. Prior knowledge of monsoon behavior will help Indian farmers and policy makers, to take advantage of good monsoons and also to minimize crop damage and human hardship during adverse monsoons. These are the issues which has motivated many studies on forecasting of Indian Summer Monsoon Rainfall (ISMR) since more than a century ago (Blanford, 1884, Walker 1908) and has been continued in recent years (Thapaliyal, 1981; Shukla and Paolino, 1983; Mooley et. al., 1986; Bhalme et. al. 1986; Shukla and Mooley, 1987; Gowariker et. al., 1991; Navino and Ceccatto, 1994; Goswami and Srividya, 1996; Venkatesan et. al., 1997; and Sahai et. al.; 2000). The techniques used in these studies are empirical statistical and Artificial Neural Networks (ANNs). However the successful and reliable forecasting of ISMR on intraseasonal to seasonal time scale and regional to country as a whole on spatial scale is still a great challenge for meteorologists.

Apart from the statistical and ANN models for prediction of ISMR the dynamical models particularly Atmospheric General Circulation Models (AGCMs) have also been used to simulate various aspects of the monsoon system in order to predict ISMR. Since the pioneering investigations by Manabe et. al. (1974) and Hahn and Manabe (1975) there have been several studies of Indian monsoon with AGCM (Palmer et. al., 1992; Chen and Yen, 1994; Ju and Slingo, 1995; Sperber and Palmer, 1996, Soman and Slingo; 1997). An understanding of how well the models can simulate the monsoon variations, when forced by observed SST, is important for assessing the potential for generating predictions of the Indian monsoon on seasonal to interannual scale. The simulation of ISMR over Indian region by various AGCMs differ not only considerably from one model to another but also deviates considerably from observations in great details such as the location and intensity of the major rainfall belts, the total rainfall over Indian land mass and the interannual variability. Sperber and Palmer (1996) evaluated the performance of variability of rainfall using 32 models for the period 1979- 1988 as a part of the Atmospheric Model Inter-comparison Project (AMIP). Their results showed that the precipitation variation over India is less well simulated. But by evaluating the simulation of interannual variability of a wind shear index over the summer monsoon region they pointed out that the models exhibit greater fidelity in capturing the large-scale dynamical fluctuations than the regional scale rainfall variation. In another study while comparing the performance of 30 models from AMIP runs Gadgil and Sajni (1998) have pointed out that the rainfall pattern over the Indian longitudes is extremely complex and not surprisingly, the simulations of the mean ISMR pattern has proved to be a difficult task.

Modeling and observational evidences suggest that the slowly varying boundary conditions of Sea Surface Temperature (SST), sea ice, soil moisture and snow at the earth's surface can influence the interannual variability of the atmospheric circulation (Charney and Shukla, 1981, Yang and Lau, 1998). So if the other boundary forcings are kept fixed and the AGCM is forced only by observed SST, then it is believed that the AGCM should nonlinearly transform the SST information from around the globe and produce a set of solutions over a particular region. The AGCM used in the present study is UKMO model. This model has shown systematic errors in producing spatial anomaly patterns and geographically shifting rainfall anomaly dipoles and their strength relative to observations in the region considered. Such errors degrade the skill of the model when grid-point comparisons between simulations and observations are made as is very common. The skills of more complicated CPU time consuming GCMS are still not better than comparatively negligible CPU time consuming empirical methods when compared for seasonal prediction of ISMR. But the current dynamical models have the advantage over the empirical models in the way that they can provide time evolution and spatial distribution of rainfall at desired spatial and temporal resolutions. Also the cause-and-effect relationship among various processes (like atmosphere-ocean interaction through SST variations, atmosphere-land interaction through albedo, soil moisture and vegetation changes etc.) represented in the model can also be analyzed.

The ability of state-of-the-art models to extract all possible information from the data available and produce seasonal forecasts is limited by several sources of uncertainties and errors. One way to tackle this problem is to use ensemble forecasts and multi-model multi-ensemble forecasts (Stern and Miyakoda, 1995; Kumar et. al., 1996; Brankovic and Palmer, 1997; Krishnamurti et. al., 1999). The other approach is to combine statistical and dynamical models (Sarda et. al., 1996 and Feddersen et. al., 1999). This approach is based on empirical orthogonal function analysis, canonical correlation analysis and singular value decomposition. These methods are designed to identify the linear combinations of variables in one field that are most strongly correlated with linear combinations of variables in another field. If the structure of the data is inherently linear then these methods are best in feature extraction. However, if the data contain nonlinear structure, it will not be detected. Recently ANNs have been used (Monahan, 2000) for nonlinear principal component analysis to extract non- linear features. The objective of this paper is to combine the nonlinear feature extracting capability of ANNs for reduction of systematic errors in simulation of AGCMs in order to achieve realistic simulation of rainfall over Indian region. For this purpose we have selected six parameters produced by model (rainfall, vorticity at 700 and 850 mb, relative humidity at 700 mb, vertical velocity and surface pressure) and have forced them to get observed rainfall. The method in detail is discussed in section 3. The reason for taking dynamical parameters lies in the conclusion by Sperber and Palmer (1996) that models exhibit greater fidelity in capturing the large-scale dynamical fluctuations than the regional scale rainfall variation. The region selected is from 5°N to 35°N and from 65°E to 95°E. Though the region selected seems to be very small, the skillful prediction of summer monsoon rainfall in this region is vital for the socio-economic activities of nearly 30% of the global population.

When an AGCM is run with the observed SST, it is pre- requisite that it should well simulate the events associated with major variation of SST field, such as ENSO. It is well known that there is a correspondence between deficit (excess) in the ISMR and the occurrence of El- nino (La- nina) events associated with warm (cold) events in the pacific. Recent observational studies of Kripalani and Kulkarni (1997) and Krishna Kumar et. al. (1999) have shown the weakening of ENSO- ISMR relationship. How AGCM behaves in this context and how far ANN corrected simulations of AGCM captured ENSO- ISMR relationship is also one concern of this communication.

One of the main objectives of the present communication is to examine the success of the ANN correction on real time forecast of ISMR. If ANN correction is able to improve the skill of AGCM at least comparable with statistical models, then it will be an achievement. For this, experimental prediction with five different initial conditions and May SST anomaly persisted for the whole monsoon season has been done from the period 1998 to 2000. Though the results cannot be compared with the performance of ANN corrections when model is run by observed SST, still the results were found encouraging.

2. Model Description and Experimental and observational Data

The model used in this study is UKMO Unified Model (HadAM2b). It is a global grid point AGCM (Cullen, 1993) with conservation split-explicit integration scheme (Cullen and Davies, 1991), which may be configured for numerical weather prediction or climate modeling. It has horizontal resolution of 3.75°long×2.5°lat when used for climate modeling. The model uses hybrid vertical coordinate system with 19 levels in the vertical and sophisticated parameterizations of radiation, boundary layer, large-scale cloud and precipitation, gravity- wave drag and convection. Version HadAM2b (Stratton, 1999) is an improved version of the atmospheric component of the coupled ocean-atmosphere climate model HadCM2, described by Jones et. al. (1997).

The model is integrated with observed SST for a period of 17 years from 1979-95 as a part of the CLIVAR Intercomparison Project of Asian-Australian Monsoon Climatology and Variability by the use of Atmospheric GCM. The monthly mean global SST and sea- ice as prescribed by Atmospheric Model Intercomparision Project II (AMIP II) were used. Other boundary conditions such as soil moisture, albedo, snow cover etc. are prescribed from climatology at the initial time and in the course of integration the model updates these parameters. Carbondioxide mixing ratio and monthly vertical distribution of ozone are also prescribed. The daily output fields of precipitation, surface pressure, vertical velocity at 850mb, relative humidity at 700mb and components of horizontal velocity at 850 and 700mb were retained from 1st June to 30th September for each of the 17 years and for whole globe. The vorticity at 850 and 700mb was calculated from the components of horizontal velocity. Data for six fields (precipitation, surface pressure, vertical velocity at 850mb, relative humidity at 700mb and vorticity at 850 and 700mb) were interpolated from 3.75°long × 2.5°lat model grid to the centre of 2.5°long x 2.5°lat grid. From this data set the data for the region from 5°N to 35°N and from 65°E to 95°E were extracted. And finally monthly means for June, July, August and September were prepared.

The observational data set of the global precipitation monthly mean fields are derived from the Climate Prediction Center Merged Analysis Precipitation (CMAP) data (Xie and Arkin, 1997) available from 1979 to 1999 on 2.5°long x 2.5°lat grid. From this data set monthly means for June, July, August and September were extracted for the region under consideration.

The error correction technique proposed in the present communication was tested for real time forecast. Five-member ensemble integration was carried out for which initial conditions were taken from the model dumps corresponding to 1st April of the last five years of 17-year integration of the same model with climatological SST over global oceans. Boundary conditions of SST used are created by persisting May SST anomalies on the climatology of the remaining months of the season. The model was integrated from April to September for the year 1998, 1999 and 2000.

3. Methodology

3.1 ANN- a brief review

Artificial Neural Networks are a class of computational models and were conceived from the logical processing of biological nervous system. They are composed of a series of parallel layers, which are interconnected computational nodes or neurons and can perform a variety of statistical modelling tasks. The traditional statistical methods invariably have inherent assumptions regarding the distribution of data and the functional relationships and their violation may severely affect the performance of the model. On the other hand, ANNs are extremely robust with regard to a priori distribution of data. Also in ANN analysis no assumptions are made regarding the underlying functions and in fact by varying the number of nodes in hidden layers one may effectively parameterize the space of all functions. ANNs are the systems that have the property of adaptation or learning. These properties make ANNs an ideal choice in many cases and they have been successfully used in nonlinear regression, classification, clustering, time-series prediction and computing non-linear PCA (Sarle, 1994; Hill et. al., 1994, Cheng and Titterington, 1994; Connor et. al., 1994, Sahai et. al., 2000 and Monahan, 2000). 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).

Feed forward ANN model has been most widely used to perform generalized nonlinear regression. This type of model is composed of a series of parallel layers (the first one is input layer, last one is output layer and there may be one or more intermediate layers called hidden layers), each of which contains a number of processing elements, or neurons, such that the output of the *i*th layer is used as input to the (i+1)th. In this study we have used a model with one hidden layer. The *j*th neuron in this hidden layer is assigned a value y_j, given in terms of the input values x_i by

$$y_j = \tanh\left(\sum_i w_{ij} x_i + b_j\right), \tag{1}$$

and the *k*th neuron in the output layer (z_k) is calculated in terms of y_j by

$$z_k = \tanh\left(\sum_j \widetilde{w}_{jk} y_j + \widetilde{b}_k\right), \tag{2}$$

Where w's and b's are the weight and bias parameters, respectively, and the hyperbolic function is used as the inner- layer transfer function. Now the problem is to determine weights and biases. This is done through network training. For this the data is divided into two mutually exclusive parts - training and test sets. The weights are randomly initialized. To construct an ANN model for forecasting, the predictor variables are used as the input, and the predictands as the output. With o_k denoting the observed data, the ANN is trained by finding the optimal values of the weights and bias parameters, which will minimize the cost function:

$$J = \sum (z_k - o_k)^2, \tag{3}$$

where the rhs of the equation is simply the total squared error of the output. The optimal parameters can be found by back- propagation algorithm. A detailed discussion on feed-forward ANN model with error- back propagation can be found in Herz et. al. (1991), Muller and Reinhardt (1991), Masters (1993), Bishop(1994) and Haykin (1994). Once the training is complete the weights are frozen and the data from test set is presented as input to evaluate the performance. It is the performance of the network on this test data set that is the true measure of the predictive capability of the network. Due to large number of parameters and the great flexibility of the ANN, the model output may fit the data very well on training data set yet producing poor forecasts on the test data set.

3.2 The network design

The network consists of one hidden layer besides input and output layers. The input layer has 24 neurons (x_i , i=1,2,…,24), which are the model-produced values of six parameters (rainfall, vorticity at 700 and 850 mb, relative humidity at 700 mb, vertical velocity and surface pressure) kept in a sequence and for the four months June to September. The output layer consists of four neurons, which are observed rainfall values for the same months. Since the inner- layer transfer function is hyperbolic tangent, the input and output values were normalized such that they lie between +1 and -1. There are 50 neurons in the hidden layer. Thus the total numbers of parameters (weights and biases) which are to be determined through training are 1454 (24x50+50+50x4+4). Since there is one pattern for each of the 17 seasons and for each of the 144 grids thus the total number of patterns is 2448. These patterns are divided into two parts- training and test. The training set consists of 2160 patterns (from the year 1979 t0 1993) and the test set of 288 (year 1994 and 1995). Since the number of parameters (1454) in the ANN model is much less than the number of training patterns (2160), there is less possibility of over fitting. However, to avoid the over fitting if any we have trained 11 networks with different initializations and out of this, 6, which were performing best on the test set, have been selected. The final output is taken as the average of outputs from these 6 networks. This is called the ensemble method (Perrone and Cooper, 1993).

3.3 The evaluation performance statistics

Let $F_{i,j}$ (i=1,2,,...,M; j=1,2,...,N; where M (=144) is the number of grid points and N (=17) is the number of years) be the forecasted precipitation field (either raw model output or ANN corrected model output) and $O_{i,j}$ be the observed precipitation field. The anomalies in these fields are defined as

$$A'_{i,j} = A_{i,j} - \overline{A}_i \qquad i = 1,2,.., M \tag{4}$$

where $\overline{A}_i$ is the time mean at each grid i and is given as

$$\overline{A}_i = \frac{1}{N}\sum_{j=1}^{N} A_{i,j} \qquad i = 1,2,.., M \tag{5}$$

Here A may be F or O. Only 15-year values (1979-1993) were used when calculating means and standard deviations. The following performance evaluation were calculated at each grid point:

For anomaly forecast

(i) Correlation coefficient (r_i): One possible choice of verification measure is the correlation coefficient (r_i) defined as follows:

$$r_i = \frac{\frac{1}{N}\sum_{j=1}^{N} F'_{i,j} \times O'_{i,j}}{(\sigma_i)_F \times (\sigma_i)_O} \qquad i = 1,2,.., M \tag{6}$$

where $$(\sigma_i)_A = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(A'_{i,j}\right)^2} \tag{7}$$

is the standard deviation of A.

This describes the strength of the linear relationships between forecasts and corresponding observations. The range of r_i is $-1 \le r_i \le 1$. For good forecasts r_i should be significantly greater than 0, for perfect forecast $r_i=1$, if there is no predictability $r_i=0$ and for inverse forecasts $r_i<0$. But correlation coefficients are not true measure of variability, that is, the correlation coefficients will not tell whether, for example, a prediction is near climatology in absolute value. Therefore another parameter, skill score, is calculated.

(ii) Skill Score (SS_i): The skill score is defined as

$$SS_i = 1 - \left(\frac{RMSE_i}{(\sigma_i)_O} \right)^2 \qquad i = 1,2..,M \tag{8}$$

The *RMSE* is traditionally calculated from the difference of the forecasted and observed fields, but since the difference of model simulated mean values and observed mean values (fig. 1 and fig. 2) are of the order of the magnitude of observed values on many grid points, therefore it will be better to calculate *RMSE* in the following way for reasonable comparison:

$$RMSE_i = \sqrt{\frac{1}{N} \sum_{j=1}^{N} \left(F'_{i,j} - O'_{i,j} \right)^2} \qquad i = 1,2,..,M \tag{9}$$

This parameter is a good measure of forecast. This measures the accuracy of the forecasts of interest relative to the climatological forecasts. For a perfect forecast $SS_i = 1$ and for always climatological forecasts it is equal to 0. If it is less than zero it implies that the forecast is even worse than the always-climatological forecasts.

For categorical Forecast

Apart from these, two more parameters were also calculated based on the categorical forecast. It is customary to issue precipitation forecast in categorical form like above normal, normal and below normal. Therefore to verify the results a 3x3 contingency table is prepared for each grid point. The categories are defined as follows

$$\left. \begin{array}{ll} \text{Category 1} & A'_{i,j} > (\sigma_i)_A \\ \text{Category 2} & (\sigma_i)_A \ge A'_{i,j} \ge -(\sigma_i)_A \\ \text{Category 3} & A'_{i,j} < -(\sigma_i)_A \end{array} \right\} \tag{10}$$

Where A takes the values of F forecast category and the values of O for observed category. Then Heidke Skill Score (HSS) and Percent Correct (PC) values were calculated for each grid points (the details of making contingency table and calculating HSS and PC can be found in Sahai et. al., 2000, Barnston, 1992 and Perrone and Miller, 1985). The value of HSS is 1 for all correct forecasts and 0 for random forecasts. Values greater than 0 indicate the improvement of forecast over random forecast. Values less than 0 indicate forecasts are worse than random forecast. The values of PC are 100 for all correct forecasts and with 3 categories its value for random forecast is 33.33.

All these above performance evaluation parameters have been calculated for each grid points. Now some parameters must be calculated for assessing performance year by year. For this following two parameters were calculated:

(i) Spatial anomaly correlation (SAC):

$$SAC_j = \frac{\frac{1}{M}\sum_{i=1}^{M} F''_{i,j} \times O''_{i,j}}{(\sigma_F)_j \times (\sigma_O)_j} \qquad j = 1,2,\ldots,N \qquad (11)$$

$$\left.\begin{aligned} \text{where} \quad & A''_{i,j} = A'_{i,j} - \frac{1}{M}\sum_{i=1}^{M} A'_{i,j} \\ \text{and} \quad & (\sigma_A)_j = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(A''_{i,j}\right)^2} \end{aligned}\right\} \quad j = 1,2,\ldots,N \qquad (12)$$

are spatial anomaly mean and standard deviation for particular year j.

(ii) Spatial skill score (SSS)

$$SSS_j = 1 - \frac{\frac{1}{M}\sum_{i=1}^{M}\left(F_{i,j} - O_{i,j}\right)^2}{(\sigma_O)_j^{\,2}} \qquad j = 1,2,\ldots,N \qquad (13)$$

It has been shown that if $SAC_j > 0.6$ and also $SSS_j > 0$ then the forecast field are useful. These parameters were calculated for the whole region under consideration 5°N to 35°N and from 65°E to 95°E and also for the Indian landmass as shown in figure 8.

4. Improvement due to ANN correction

4.1 Discrepancies between model and observed climatology

The 15-year (1979 to 1993) observed (CMAP) mean patterns of seasonal rainfall for monsoon season are shown in Fig. 1(b). This figure shows two rainbelts one along the eastern part of foothills of Himalayas (rainbelt I) and other along the West Coast of peninsula (rainbelt II). These rainbelts are associated with the occurrence of heavy rainfall due to orography of the Himalayas and the Western Ghats. There is also a third rainbelt associated with the continental ITCZ (commonly known as monsoon trough in Indian region) and extends from the Bay of Bengal to the west- northwest across the Indo- Gangetic plain (rainbelt III). During the summer season this rainbelt fluctuates with monsoon trough and gives rise to the variation in monsoon rainfall on intra- seasonal time scale and in turn on interannual time scale. Particularly, the variation of the western part of this rainbelt, i.e., west of 80°E is the prime contributor to the interannual variation.

Apart from these three rainbelt regions, there is a rain shadow region along the East Coast of the peninsula. The variation of rainfall in this region also contributes significantly to the interannual variation.

Fig. 1(a) shows the model-simulated climatology for 15 years (from 1979 to 1993). When compared with the observations the model simulation shows that rainbelt I is weaker, rainbelt II is more intense and its maxima shifted northward, while rainbelt III is also more intense at the head Bay of Bengal and shifted southward at both the ends. The rain shadow region is not simulated on the land but it is eastward and on the sea. Thus the model is successful in capturing almost all the important features of the observation but with changed geographical location and changed intensity.

Fig.1 (c) shows the ANN corrected model climatology. It is difficult to find any difference between observed climatology and corrected climatology. Therefore difference figure was plotted. Fig. 1(d) is the difference between model and observed climatology (model - observed) and fig. 1(e) is the difference between ANN corrected model climatology and observed climatology (corrected - observed). We find that model shows excess rainfall over most of the peninsula and along West Coast the difference is well above 10 mm/day. The model also shows less rainfall over head Bay and northern and eastern parts of the country. The ANN corrected rainfall has shown maximum difference along Gujrat coast, but its magnitude is just above 1 mm/day, which is negligible when compared with deficiency in models. In other parts of the region the difference is below 1 mm/day. On the monthly scale also the ANN corrected climatology is very close to observed climatology (fig. 2). These improvements seem to be dramatic but it is not really so. Because if we use the following formula at each grid point:

$$\text{Corrected model value} = (\text{standardized model value}) \times (\text{Standard deviation of observed values}) + \text{Mean of observed values} \qquad (14)$$

then the climatology and variance of corrected model output will be exactly matching with observation. This formula has been used in some form while normalizing (putting values between -1 and 1) the values for ANN correction and then denormalizing (getting back the values from -1 to 1 to the actual values). Thus the actual comparison will be on interannual and intra-seasonal time scales where the formula (14) has nothing to do with sign of anomalies.

4.2 Interaannual variability

Grid point Comparison: The correlation coefficients (r) between model and observed values for 17-year period (1979-1995) is shown in Fig. 3(a). Except for some regions over the Arabian Sea and Bihar Plateau, where model shows some significant correlation, there is practically no correlation and even some regions over Bay of Bengal are showing significant negative correlation. While r between ANN corrected and observation (Fig. 3(c) shows that over almost all places correlation has improved linearly, i. e., where model was showing significant positive correlation it has improved to greater than 0.8 and where model was showing significant negative correlation it is between 0.1 and 0.4.

Nowhere correlation is negative. For categorical forecast Fig. 3(b) shows the Heidke Skill Score (HSS) for model and Fig. 3(d) for ANN corrected. It can be observed that over about half of the region model forecast is slightly better than random forecast and over rest of the region it is even worse. It has slightly better skill in some parts of the Arabian Sea, while ANN corrected is everywhere better than the random forecast except for some parts of coastal Andhra and Orissa. Fig. 3(e) and (g) show the values of percentage of correct categorical forecast (PC) for model and ANN corrected respectively. Almost everywhere ANN corrected forecast is corrected more than 50% and many places over 70 to 80 % and even more, while the model forecast is either less than 50% correct or slightly above. Fig. 3(f) and (h) show the values of Skill Score (SS) for model and corrected values respectively. Except along some part of eastern region everywhere the values of SS for GCM simulations are negative and even less than -10 along peninsula and Bay of Bengal, while the values of SS are slightly lower than 0 in some places for ANN corrected forecast. Thus from Fig. 3 we can easily see that on seasonal scale there is a significant improvement in model forecast by ANN correction on the interannual time scale.

On monthly scale, the above-discussed statistics are shown in Fig.4 to 7 and we can find similar improvement on intra-seasonal time scale also. Over all we can conclude that northern Gujrat, Rajasthan, western Madhya Pradesh, some parts of eastern states, middle peninsula, and most parts of Arabian Sea are more predictable regions when model prediction is corrected with ANN method. While tip of peninsula, southern Gujrat, east Maharashtra, regions over Orissa and Andhra coast and most parts of Bay of Bengal are comparatively less predictable.

Time series comparison: As already mentioned, the time series for whole region (WR) and Indian land (IL) only were made for anomaly, skill score and spatial correlation. The anomaly is shown in fig. 9 and 10. It is clear that on seasonal scale though there is an improvement (correlation coefficient increased from -0.093 to 0.233 for IL and from 0.084 to 0.312 for WR) but still the correlation is not significant. On monthly scale except for July there is a significant improvement with ANN correction when compared with model. Fig. 11 shows the skill score and spatial anomaly correlation. It can be seen in Fig. 11(a) and (b) that for ANN corrected values skill score is always positive and having high values (close to 1) while in case of model it is either negative or close to 0. Spatial anomaly correlation fig. 11(c) and (d) shows that except for the year 1990 not a single year the model forecast is useful while ANN corrected forecast is useful in most of the years.

Thus there is a significant improvement when model forecast is corrected by ANN method.

4.3 Comparison for some individual Years (ENSO/ISMR relationship)

For evaluating the improvement of the ANN correction in GCM performance on seasonal rainfall in individual years we have selected 4 years from the training - contrasting years of 1982/1983 and 1987/1988. These years are associated with ENSO. There are only 2 years in test set, 1994 and 1995. These years are also considered.

Fig. 12 shows the model simulated rainfall, corrected rainfall and observed rainfall for the year 1982. From the fig. 12f (observed rainfall anomaly), it is seen that except some parts of Tamil Nadu, Karnataka and north India the rainfall anomaly is negative. The model-simulated rainfall (fig. 12a), however, is not satisfactory. Like mean rainfall simulation (fig. 1), the West Coast maxima is located to the north of its observed position. The corresponding anomaly from the GCM output (fig. 12d) however shows negative anomalies over most part of the country except some parts of western and central India. The corrected model output (fig. 12b) has improved rainfall distribution over the country, when compared with actual rainfall distribution (fig. 12c). West Coast maxima is very close to the observed position. In the corrected anomaly (fig. 12e), the positive anomaly over north India is in accordance with the observation. However the positive anomaly area is extended to larger region. In uncorrected case the positive anomaly over north India was not seen instead there was negative anomaly. The magnitude of negative anomaly over north- east India in ANN corrected has increased and is closer to observed anomaly. Although the seasonal rainfall is deficient during 1982, this feature is not distributed uniformly over the entire country. If we see the rainfall distribution in much smaller spatial scale, the rainfall is above normal over some part of north India (fig. 12i). Below normal rainfall is observed over an east- west oriented rainbelt from Maharashtra coast to Orissa coast. In rest of the country it is normal. Here the above normal, normal and below normal rainfall are defined as in eqn. (10). There is significant improvement in the ANN corrected output (fig 12h) compared to the raw model output (fig. 12g), in the sense that above normal rainfall over north India is well produced, in ANN corrected case where as it was not there in model case. A small pocket of below normal rainfall over East Coast is also reported in corrected case. Again like the anomaly (fig. 12e), here also the above normal rainfall has larger spatial extension.

Unlike 1982, in 1983 the observed rainfall (fig. 13c) and its anomaly (fig. 13f) show well distributed rainfall over the country. Except the region of NE India and a small pocket of southern India observed rainfall anomaly is positive. The worst performance of GCM is seen for the year 1983. The rainfall anomaly simulated by the GCM for 1983 (fig. 13d) shows negative anomaly over entire India. In most of the GCM simulations studied by Gadgil and Sajani (1998) it is seen that models have not captured the extensive rainfall during 1983. This GCM also has not captured the extensive rainfall during 1983 and has simulated very little rainfall. However, the corrected anomaly simulated by the GCM has shown tremendous improvement. It has captured almost satisfactorily the negative anomaly over NE India, a small pocket of negative anomaly over southern India and positive anomaly over the rest part of the country. Figure 13g, h and i show the area of above normal, normal and below normal rainfall over different parts of the country as simulated by model, ANN corrected and that from the observation. Although the spatial distributions of above normal, normal and below normal rainfall are not matching in the model output (fig. 13g) with observation (fig. 13i), a very significant improvement can be seen in ANN corrected rainfall categories (fig. 13h).

Fig. 14 depicts the spatial distribution of actual, anomaly and different categories for GCM, ANN corrected and observed rainfall of 1987. This year was the one of the severe drought years in recent time. From fig. 14f it can be seen that the observed rainfall shows negative anomaly over entire country except NE India. From fig. 14d we can see that model performance is good in the sense that most of the GCMs have simulated the variability during 1987 and 1988. But the model has simulated positive anomaly over most part of the peninsula, which is not the case for observation. This negative anomaly is almost disappeared after ANN correction. The distribution of positive and negative anomaly in ANN corrected (fig.14e) is very much identical to that of observed (fig. 14f). While comparing the categorical forecast (fig. 14g, h and i) similar improvement can be seen. The normal rainfall over peninsula in model simulation (fig.14g) became below normal after correction (fig. 14h) and it is in accordance with the observation (fig. 14i).

Year 1988 was also a good monsoon rainfall like 1983. However the difference is that in 1988 many GCMs well captured the rainfall distribution where in 1983 it was not so. If we compare the model anomaly (fig. 15d) with observed anomaly (fig. 15f) there is a large discrepancy in rainfall distribution over West Coast and southern India with a large negative anomaly in model (fig. 15d) whereas observation shows positive anomaly (fig. 15f). This feature of observed anomaly has clearly brought out in the ANN corrected case (fig. 15e). Similarly the negative anomaly over some parts of north India as seen in the observation, which was not simulated by GCM is also well produced in the ANN corrected output. Similar improvements can also be seen for categorical forecast (fig. 15g, h and i).

So far we have discussed the performance on some years from training set. Now the improvement of performance by ANN correction in the year 1994 and 1995 of the test set will be the real validation of the correction technique. The most critical year is 1994 when the ENSO was very much similar to 1991 but 1994 was a good monsoon year where as 1991 was a bad one. Most of the statistical models and GCMs have failed in the prediction of rainfall anomaly for 1994. During this year except East Coast of India, some parts of central India and NE India (fig. 16f), the rainfall anomaly is positive in the rest of the country. This pattern of rainfall distribution is also seen in fig. 16c. Although the rainfall is well simulated by the GCM (fig. 16a), scanty rainfall over Tamil Nadu and Andhra Pradesh is not demonstrated. However, this feature has been brought out in corrected output (fig. 16b). Of course this improvement is not reflected well in the anomaly pattern (fig. 16d and e). But when we compare fig. 16d and 16e with fig. 16f we clearly find that negative anomaly over northern India produced by the GCM is replaced with positive anomaly after correction and positive anomaly over NE India is replaced with negative one. These improvements are in accordance with the observation. Also the categorical distribution has improved when we compare the uncorrected (fig. 16g) and corrected (fig. 16h) output with observed (fig. 16i). Above normal rainfall over NW and north India, seen in the observation (fig. 16i) is clearly brought out in ANN corrected (fig. 16h) where it was mostly below normal in GCM simulation (fig. 16g). The below normal rainfall over some parts of east coast and adjoining Bay of Bengal as shown in fig. 16i is also improved in corrected output, whereas this was not seen in GCM simulation.

During 1995, the model rainfall shows a large excess rainfall over West Coast (fig. 17a). The West Coast maxima is very much extended. The same feature is also reflected in model anomaly (fig. 17d). The observed anomaly during 1995 shows positive and negative anomalies distributed uniformly over the country. In the corrected model anomaly (fig. 17e), the very extensive rainfall of West Coast has improved with reduction in magnitude and well compared with observed anomaly (fig. 17f). In the categorical distribution of rainfall (fig. 17 g, h and i) of course there is not much improvement in the ANN corrected output (fig. 17h). Only the normal rainfall over central India starting from Maharashtra, Gujrat Coast to Orissa Coast is reflected in fig. 17h. However the east- west oriented above normal belt is not desirable in corrected output.

All the discussions above are mainly concentrated on Indian landmass. Similar improvements are also noticed on oceanic regions with greater improvement in Arabian Sea than in the Bay of Bengal. Thus we see that there is a significant improvement in the model simulations when they are corrected with ANN method as far as spatial distribution and intensity is concerned. The proposed corrections have very well captured the teleconnection link between ENSO and ISMR in simulating the rainfall anomaly distribution for the years 1982,1983,1987,1988. The recent observational studies of Kripalani and Kulkarni (1997) and Krishna Kumar et. al. (1999) have indicated that the ENSO- ISMR relationship may be weakening. But the performance of the ANN corrected model output is not affected during 1990s and therefore we may conclude that though there may be weakening in ENSO/ISMR relationship, the GCMs can still produce good simulations with observed SST as boundary conditions. This may also imply that the dominant SST forcing may not be from the central Pacific alone.

The improvements may be summarized in the following table:

Year	Spatial Correlation with CMAP			Spatial Anomaly correlation with CMAP anomaly	
	Model	ANN corrected	CMAP Climatology	Model	ANN corrected
1982	0.72	0.95	0.95	-0.26	0.49
1983	0.71	0.97	0.93	-0.07	0.74
1987	0.66	0.97	0.91	0.21	0.87
1988	0.69	0.96	0.96	0.09	0.55
1994	0.75	0.94	0.94	-0.26	0.24
1995	0.55	0.96	0.92	-0.03	0.70

These values are shown for the whole region. We find that in terms of spatial distribution of rainfall for every year discussed above the climatological forecast is better than model forecast but ANN corrected model forecast is always either better or at par. When we compare the spatial anomaly forecast we find that there is a very significant improvement in simulating the anomalies after ANN correction and this is not so only in training set years but also in test set years. But the improvement depends on the performance of GCM itself.

4.4 Improvement in real time forecast:

The improvement in GCM forecast with observed SST when corrected with ANN will be validated for the case for May anomaly persistence forecast for years 1998 to 2000. For this instead of taking observed SST as boundary conditions the anomaly in SST in the month of May was persisted for the four monsoon months. We have used ensemble method. The model was integrated with 5 different initial conditions. These initial conditions were taken from the model dumps corresponding to last five years of 17-year integration of the same model with climatological SST. The model was integrated from April to September. Here we have to keep in mind that neither the SST is observed nor the initial conditions are real.

The model simulated rainfall, ANN corrected rainfall and CMAP rainfall for 1998 seasonal mean is shown in fig. 18a, b and c respectively. There is a good agreement in spatial distribution between ANN corrected and observation. When comparing the anomalies in fig. 18d, e and f we find that model simulated negative anomalies throughout the Indian land and Bay of Bengal except some positive anomalies were present in some parts of Arabian Sea and northern and eastern states. However in observations (fig. 18f) there were negative anomalies only in central India and Orissa and head of the Bay of Bengal. The ANN corrected rainfall (fig. 18e) shows some improvement in replacing the negative anomaly over southern peninsula, Bay of Bengal and also intensifying the magnitude of anomalies along eastern states produced by the model in association with observation. There is a significant improvement in categorical forecast also (fig. 18g, h and i).

For the year 1999 we again find the distribution of rainfall in ANN corrected (fig. 19b) is very much close to observation (fig.19c) than the GCM simulation (fig. 19a). The model is showing positive anomalies almost everywhere (fig. 19d) except some regions in NE India and a small pocket along the tip of peninsula. The observed anomalies (fig. 19f) are, however, negative over the entire region except some regions over Uttar Pradesh, Bihar and NE India. The ANN correction (fig. 19e) has shown some improvements in the sense that the magnitude of positive anomalies were reduced along West Coast, positive anomalies were replaced with negative ones along East Coast, Bay of Bengal and western parts of the country and also the negative anomaly over the head Bay and NE India is replaced by positive anomaly in accordance with observation. But the correction has spoiled the simulation along the tip of peninsula where it has replaced negative anomaly produced by model in a positive one. Though these improvements can be seen in categorical distribution (fig. 19g, h and i) when we compare the performance between model and corrected one, but still the corrected one is not very much encouraging if we compare with observation.

Fig. 20a and b shows the model and corrected forecast for year 2000, the CMAP data for this year is not yet available. But we can compare figures 20c to 20f with India Meteorology Department observation figure available on the Internet at the site http://www.tropmet.ernet.in/~kolli/MOL/Monsoon/frameindex.html. The model has shown negative anomalies throughout the peninsula and positive anomalies over rest of the India (fig.

20c). The observations show positive anomaly along the peninsula, some parts of Uttar Pradesh, Bihar and NE India while negative anomaly over rest of the India with strong negative anomaly along western and central part. The ANN corrections have replaced negative anomaly over peninsula as positive one and also strong positive anomalies along western part of the country were replaced with small negative ones (fig. 20e). These improvements were also reflected in categorical distribution (fig. 20f).

Thus we can easily obtain good improvement in the spatial distribution of anomalies and categories when the model forecast is corrected with ANN technique. One important improvement can be noticed in the categorical distribution that the difference between observed and corrected category is not of 2 categories, i. e., never the excess was forecasted as below neither below was forecasted as above, as it is observed in many places for GCM forecast.

5. Discussions and Conclusion

The principal scientific basis of seasonal forecasting is founded on the premise that lower-boundary forcing, which evolves on a slower time-scale than that of the weather systems themselves, can give rise to significant predictability of atmospheric events (Palmer and Anderson, 1994). These boundary conditions include sea surface temperature (SST) apart from others like sea-ice cover and, land-surface temperature and albedo, soil moisture and snow cover etc. The strongest evidence for long-term predictability comes from the influence of persistent SST anomalies on the atmospheric circulation, which in turn, induces seasonal climate anomalies. Based on these premises there are two approaches for seasonal prediction- statistical and dynamical. Using statistical methods, the relationship between regional atmospheric anomalies and pre-season SST anomaly patterns are studied and where significant links are found, current SST patterns are used to predict future atmospheric anomalies. The other approach is the dynamical one using AGCMs or coupled ocean-atmosphere GCMs. The results obtained from statistical methods are very encouraging and remain a target to be achieved by the dynamical models. It is likely that the best forecast will be produced by an objective combination of both methods (Carson, 1998).

The most challenging and, of course, a litmus test for any GCM is to simulate the interannual variability of the summer monsoon (JJAS) rainfall over Indian region. While comparing the performance of 32 GCMs for the period 1979-1988, Sperber and Palmer (1996) showed that except for showing the enhanced precipitation in the year 1988 in comparison to 1987, little or no consensus among GCM simulations exists with regard to Indian monsoon rainfall. They further concluded that GCMs exhibit greater fidelity in capturing the large-scale dynamical fluctuations than the regional-scale rainfall variations. This paper deals with the most challenging issue of prediction of Indian summer monsoon rainfall by combining the ANN technique with GCM. The novel idea of this paper is inspired by the non-linear feature extraction capability of ANN (Monahan, 2000) and the conclusion of Sperber and Palmer (1996) and therefore dynamical parameters produced by the GCM were used to obtain the correction using ANN technique. Our aim is not to demonstrate that the GCM used is the best one nor the model produced dynamical parameters selected here for ANN model development are optimal, but to demonstrate the potential of the proposed method. We have clearly demonstrated the skill of the method and showed that the corrected GCM forecasts are better, beyond reasonable doubt, than that of random forecasts and climatological forecasts. The correction by this method is most effective in some parts of Arabian Sea where the GCM's performance was also better. This implies that if the performance of the GCM will be good then the proposed correction method will be best. The present paper is focused not only on correcting the geographical errors in simulating the mean pattern of Indian summer monsoon rainfall but also on the variability on intra-seasonal (monthly) to interannual time scales.

When forced by observed SST as boundary conditions we find that there is a very significant improvement in simulating the interannual variability of anomaly distribution. Since the ANN network was obtained using monthly data sets, similar improvements are also found on monthly scales. Thus we see that there is a significant improvement in the model simulations when they are corrected with ANN method as far as spatial distribution is concerned. The proposed corrections have very well captured the teleconnection link between ENSO and ISMR in simulating the rainfall anomaly distribution for the years 1982,1983,1987,1988. The recent observational studies of Kripalani and Kulkarni (1997) and Krishna Kumar et. al. (1999) have shown the weakening of ENSO- ISMR relationship. But it seems that the performance of ANN corrected model output is not affected during 1990s. Therefore we may conclude that although there may be weakening in ENSO/ISMR relationship, the GCMs when combined objectively with some statistical model can still produce good simulations with SST as boundary conditions. All these improvements are achieved for a single run of GCM, which can be even more improved with multi- runs. For real time forecast with persistence of May SST anomaly over the entire season, we also find some improvement. What we understand that if the GCM can be run with a few initial conditions with May SST anomaly persistence as boundary conditions for many summer seasons and if the ANN will be trained on these simulations then we may get forecasts better than the purely statistical model forecasts. Also the linear increase in correlation coefficient, as seen in fig. 3 to 7, indicates that different ANN models can be developed for different regions based on positive and negative correlation between model and observation.

Acknowledgements: We are thankful to Dr. G. B. Pant, Director of Indian Institute of Tropical Meteorology for encouragement. Thanks are also due to U. K. Meteorological Office, U. K. for providing HadAM2b GCM used in this study and to Stuttgart University for using its neural network simulator (SNNSv4.1) for ANN model development. We are deeply indebted to Late Dr. M. K. Soman of our Division who had basically set up all the GCM experiments and with whom we had many valuable discussions.

REFERENCES

Barnston, A. G., 1992: Correspondence among the correlation, RMSE, and Heidke forecast verification measures; Refinement of the Heidke score, *Weat. Forcast.*, 7, 699-709.

Barnston, A. G., 1994: Linear statistical short-term climate predictive skill in the northern hemisphere, *J. Clim.*, **7**, 1513-1564.

Bhalme, H. N., Jadhav, S. K., Mooley, D. A. and Ramana Murty, Bh. V., 1986: Forecasting of monsoon performance over India, *J. Climatol.*, **6**, 347-354.

Bishop, C. M., 1994: Neural Network and their applications, *Rev. Sci. Instr.*, **65**, 1803-1832.

Blanford, H. F., 1884: On the connection of the Himalayan snowfall with dry winds and seasons of draughts in India, *Proc. Roy. Soc. London*, **37**, 3-22.

Brankovic, C. and Palmer, T., 1997: Atmospheric seasonal Predictability and estimates of ensemble size. *Mon. Wea. Rev.*, **121**, 1786- 1802.

Carson, D. J., 1998: Seasonal forecasting, *Q. J. R. Meteorol. Soc.*, **124**, 1-26.

Charney, J. G. and Shukla, J., 1981: Predictability of monsoons, *Monsoon Dynamics*, J. Lighthill and R. P. Pearce, Eds., University Press, 99-109.

Chen, T. -C. and Yen, M. -C., 1994: Interannual variation of the Indian monsoon simulated by the NCAR Community Climate Model: Effect of the tropical Pacific SST, *J. Climate*, **7**, 1403-1415.

Cheng, J. T. and Titterington, D. M., 1994: Neural networks: A review from a statistical perspective, *Statistical Science*, **9(1)**, 2-54.

Connor, J. T., Martin, R. D. and Atlas ,L. E., 1994: Recurrent neural networks and robust time series prediction, *IEEE Trans. Neural Networks*, **5(2)**, 240-254.

Cullen, M. J. P., 1993: The unified forecast/climate model, *Meteorol. Mag.*, **122**, 81- 94.

Cullen, M. J. P., and Davies, T., 1991: A conservative split- explicit integration scheme with fourth- order horizontal advection, *Q. J. R. Meteor. Soc.*, **117**, 993-1002.

Feddersen, H., A. Navarra and M. N. Ward, 1999: Reduction of model systematic error by statistical correction for dynamical seasonal prediction. *J. Climate*, **12**, 1974-1989.

Gadgil, S. and Sajani, S., 1998: Monsoon precipitation in AMIP runs, *Clim. Dyn.*, **14**, 659-689.

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

Gowariker, V., Thapliyal, V., Kulshrestha, S. M., Mandal, G. S., Sen Roy, N. and Sikka, D. R., 1991: A power regression model for long-range forecast of southwest monsoon rainfall over India, *Mausam*, **42**, 125-130.

Hahn, D. G. and Manabe, S., 1975: The role of mountains in the south Asian monsoon circulation, *J. Atmos. Sci.*, **32**, 1515-1541.

Haykin, S., 1994: *Neural Networks: A comprehensive Foundation*, New york, Macmillan.

Herz, J., Krough, A. and Palmer, R. G., 1991: *Introduction to the theory of Neural Computation,* Addison- Wesley.

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

Jones, T. C., Carnell, R. E., Crossley, J. F., Gregory, J. M., Mitchell, J. F. B., Senior, C. A., Tett,S. F. B. and Wood, R. A., 1997: The second Hadley Centre coupled ocean-atmosphere GCM: Model description, spinup and validation, *Clim. Dyn.*, **13**, 103-134.

Ju, J. and Slingo, J., 1995: The Asian summer monsoon and ENSO, *Q. J. R. Meteor. Soc.*, **121**, 1133- 1168.

Kripalani, R. H. and Kulkarni, A., 1997: Climate impact of El- Nino/ La- Nina on the Indian monsoon: A new perspective, *Weather*, **52**, 39- 46.

KrishnaMurti et. al., 1999: Improved weather and seasonal climate forecasts from multimodel superensemble, *Science*, **285**, 1548-1550.

Kumar, A., M. Hoerling, M. Ji, A. Leetma and P. Sardeshmukh, 1996: Assessing a GCM's suitability for making seasonal predictions. *J. Climate*, **9**, 115- 129.

Krishna Kumar, K., B. Rajagopalan and M. A. Cane, 1999: On the weakening Relationship Between the Indian Monsoon and ENSO, *Science*, **284**, 2156-2159.

Maier, H. R. and Dandy, G. C., 2000: Neural networks for the prediction and forecasting of water resources variables: A review of modelling issues and applications, *Environmental Modelling and Software*, **15**, 101-124.

Manabe, S., Hahn, D. G., Hollaway, J., 1974: The seasonal variation of tropical circulation as simulated by a global model of atmosphere, *J. Atmos. Sci.*, **32**, 43-83.

Masters, T., 1993: Practical Neural Network Recipes in C++, Academic Press.

Monahan, A. H., 2000: Nonlinear principal component analysis by neural networks: Theory and application to the Lorenz system, *J. Clim.*, **13**, 821-835.

Mooley, D. A., Parthasarathy, B. and Pant, G. B., 1986: Relationship between all- India summer monsoon rainfall and location of ridge at 500 mb level along 750 E, *J. Climate Appl. Meteor.*, **25**, 633- 640.

Muller, B. and Reinhardt, J.,1991: *Neural Networks: An Introduction*, Vol.2, The physics of Neural Networks Series, Spriger-Verlag.

Navone, H. D. and Ceccatto, H. A., 1994: Predicting Indian monsoon rainfall: A neural network approach, *Clim. Dyn.*, **10**, 305- 312.

Palmer, T. N., Brankowic, C., Viterbo, P., Miller, M. J., 1992: Modeling interannual variations of summer monsoons, *J. Clim*, **5**, 399-417.

Palmer, T. N. and Anderson, D. L. T., 1994: The prospects for seasonal forecasting- A review Paper, *Q. J. R. Meteorol. Soc.*, **111**, 947-975.

Perrone, T. J. and Miller, R. G., 1985: Generalised experimental markov and model output statistics: A comparative verification, *Mon. Wea. Rev.*, **113**, 1524- 1541.

Perrone, M. P. and Cooper, L. N., 1993: When networks disagree: Ensemble methods for hybrid neural networks, in J. Mammone (ed.), *Artificial Neural Networks for Speech and Vision*, Chapman and Hall, London, 126-142.

Rowell, D. P., Folland, C. K., Maskell, K. and Ward, M. N., 1995: Variability of summer rainfall over tropical north Africa (1906-92): Observations and modelling, *Q. J. R. Meteorol. Soc.*, **121**, 669-704.

Sahai, A. K., M. K. Soman and V. Satyan, 2000: All India summer monsoon rainfall prediction using an artificial neural network, *Clim. Dyn.*, **16**, 291- 302.

Sarda, J., G. Plaut, C. Pires and R. Vautard, 1996: Statistical and dynamical long-range atmospheric forecasts: Experimental comparison and hybridization. *Tellus*, **48A**, 518-537.

Sarle, w. S., 1994: Neural networks and statistical models, *Proceedings of the Nineteenth Annual SAS Users Group International Conference* , 1538- 1550.

Shukla, J. and Mooley, D. A., 1987: Empirical prediction of the summer monsoon rainfall over India, *Mon. Wea. Rev.*, **115**, 695- 703.

Shukla, J. and Paolino, D. A., 1983: The Southern Oscillation and long-range forecasting of the summer monsoon rainfall over India, *Mon. Wea. Rev.*, **111**, 1830-1837.

Soman, M. K. and Slingo, J., 1997: Sensitivity of Asian Summer monsoon to aspects of sea surface temperature anomalies in the tropical pacific ocean, *Q. J. R. Meter. Soc*, **123**, 309-336.

Sperber, K. R., Palmer, T. N., 1996: Interannual tropical rainfall variability in general circulation model simulations associated with the Atmospheric Model Intercomparison Project, *J. Clim.*, **9**, 2727- 2750.

Stern, W. and K. Miyakoda, 1995: Feasibility of seasonal forecasts inferred from multiple GCM simulations, *J. Climate*, **8**, 1071- 1085.

Stratton, R. A., 1999: A high resolution AMIP integration using Hadley Centre model HadAM2b, *Clim. Dyn.*, **15**, 9-28.

Swaminathan, M. S., 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*, No. 12/81.

Venkatesan, C., Raskar, S. D., Tambe, S. S. Kulkarni, A. D. and Keshavamurty, R. N., 1997: Prediction of All India Summer Monsoon Rainfall Using Error- Back Propogation Neural Networks, *Meteorol. Atmos. Phys.*, **62**, 225- 240.

Walker, G. T.,1908: Correlation in seasonal variation of climate(Introduction), *Memoirs of Ind. Met. Dep.*, **20**, 117-124.

Ward, M.N. and Folland, C. K., 1991: Prediction of seasonal rainfall in the north Nordeste of Brazil using eignvectors of sea-surface temperature, *Int. J. Climatol.*, **11**, 711-743.

Webster, P. J., Magana, V. O., Palmer, T. N., Shukla, J., Thomas, R. A., Yanai, M. and Yasunari, T., 1998: Monsoons: Processes, predictability, and prospects for prediction, *J. Geo. Res.*, **13**, 14,451-14,510.

Xie, P. and Arkin, P. A., 1997: Global precipitation: A 17-year analysis based on gauge observations, sattelite estimates, and numerical model outputs, *Bull. Am. Met. Soc.*, **78**, 2539-2558.

Yang, S. and Lau, M. K., 1998: Influences of SST and ground wetness on Asian summer monsoon, *J. Clim.*, **11**, 3220-3246.

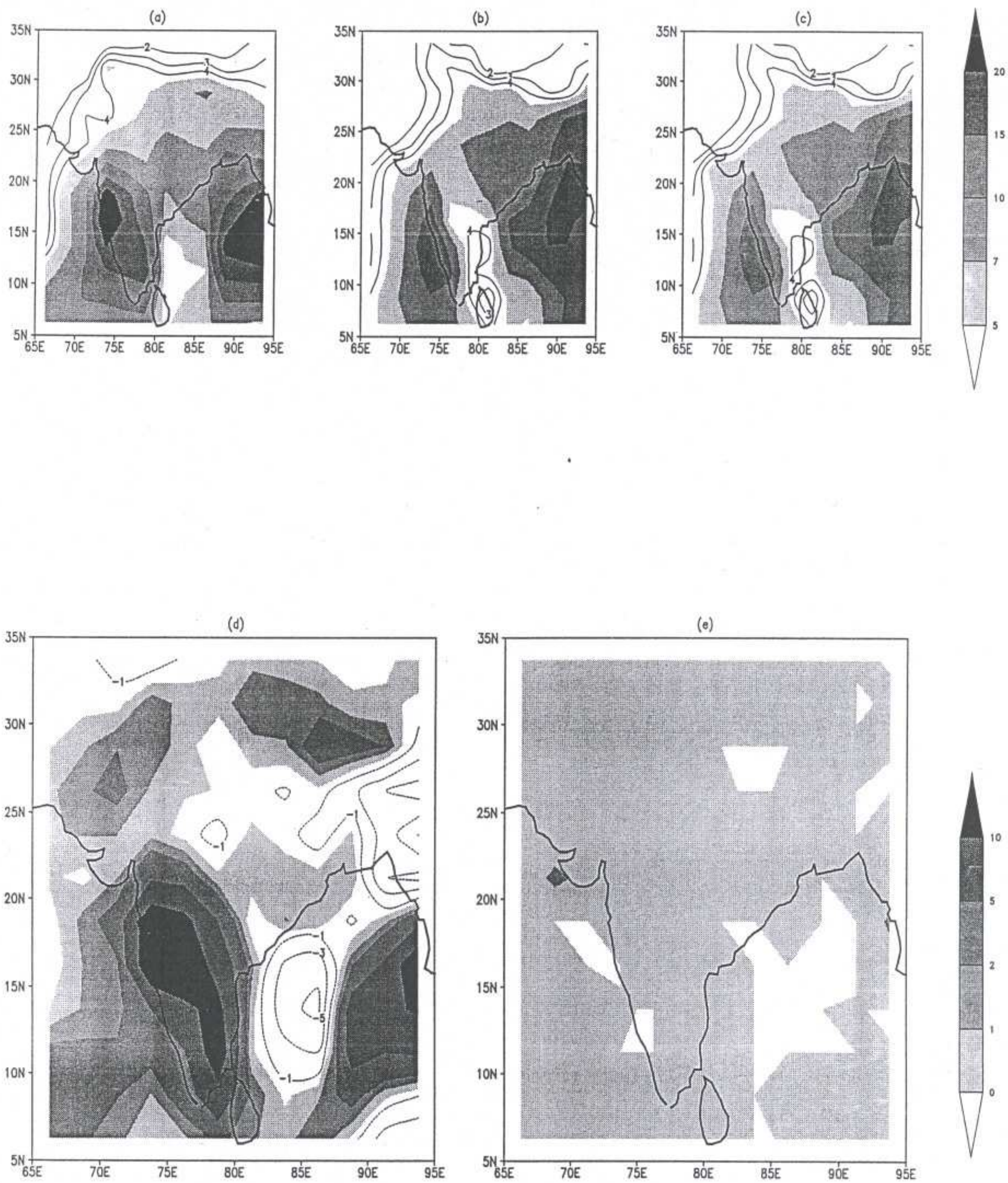


Figure 1. 15 year (from 1979 to 1993) climatology of summer monsoon season (June to September) mean rainfall (a) as simulated by model (b) for CMAP (observation) data and (c) for ANN corrected model output. Difference in climatology (d) (model - CMAP) (e) (corrected - CMAP). All these values are in mm day^{-1}.

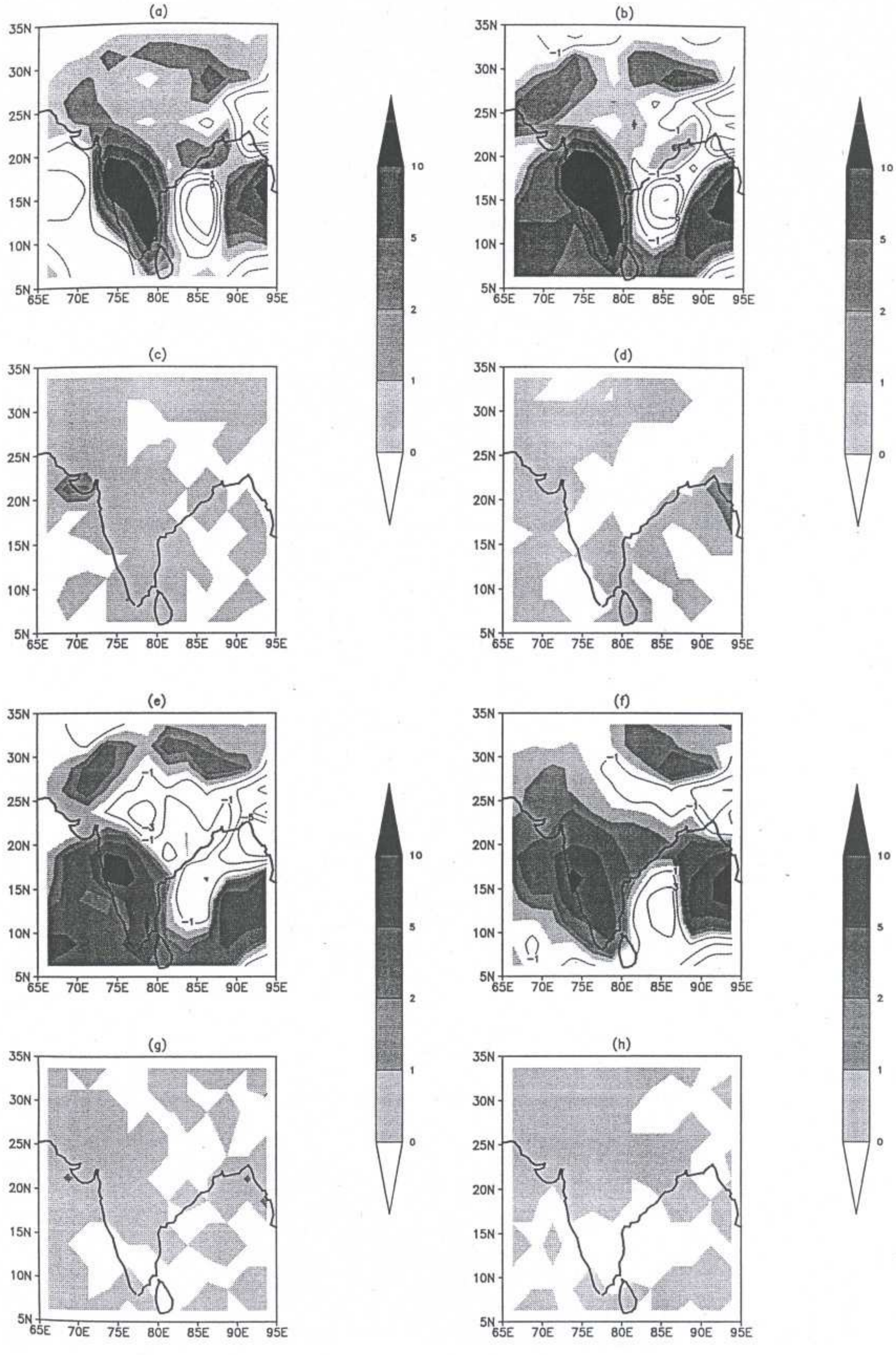


Figure 2. Difference in climatology (a) (model - CMAP) for June (b) (model -CMAP) for July (c) (corrected - CMAP) for June (d) (corrected - CMAP) for July (e) (model - CMAP) for August (f) (model - CMAP) for September (g) (corrected - CMAP) for August (h) (corrected - CMAP) for September. All these values are in mm day^{-1}.

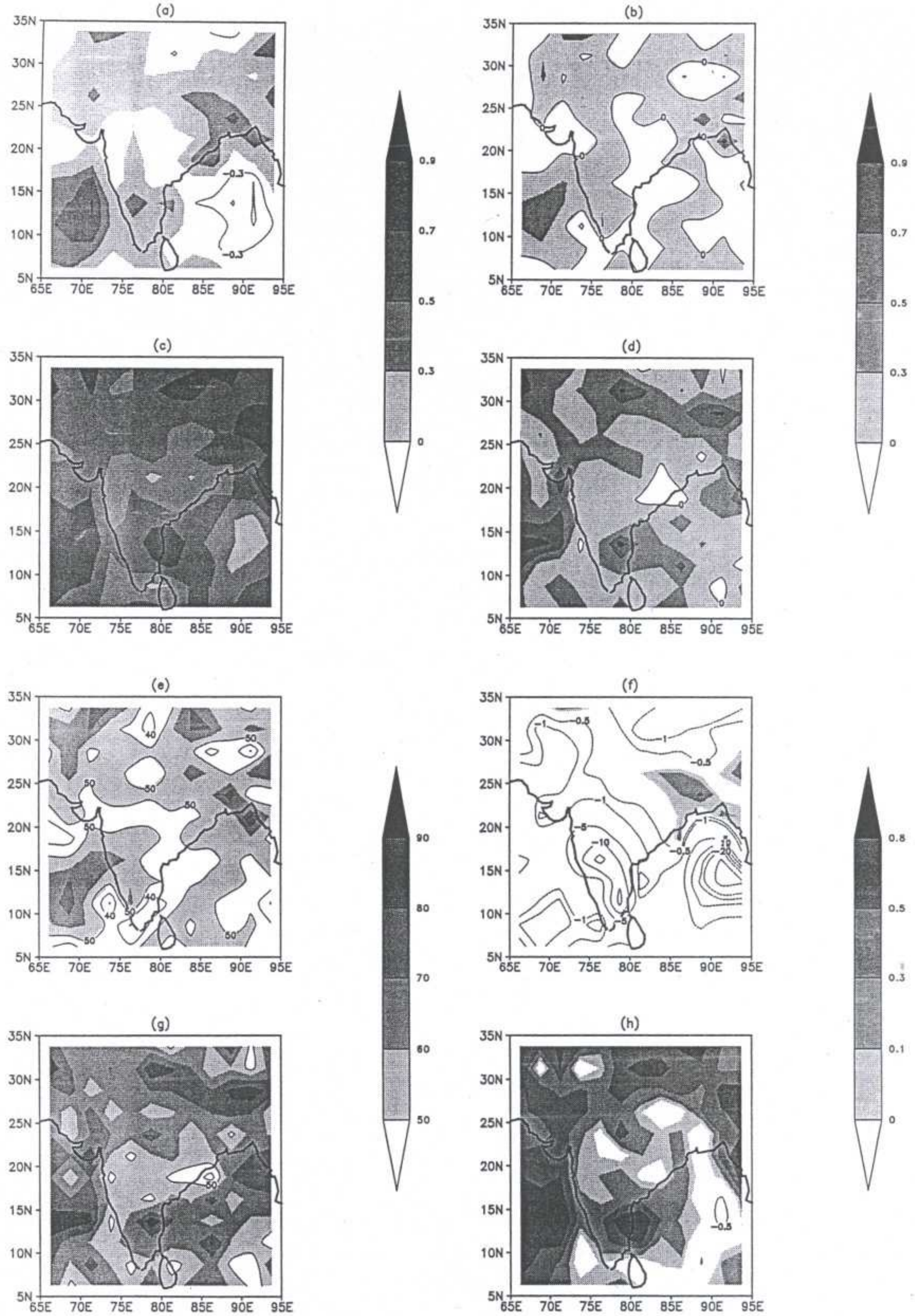


Figure 3. (a) Correlation coefficient (r) between model and CMAP (b) HSS for model (c) r between corrected and CMAP (d) HSS for corrected (e) PC for model (f) SS for model (g) PC for corrected and (h) SS for corrected. All these values are for seasonal mean rainfall and for 1979 to 1995 period.

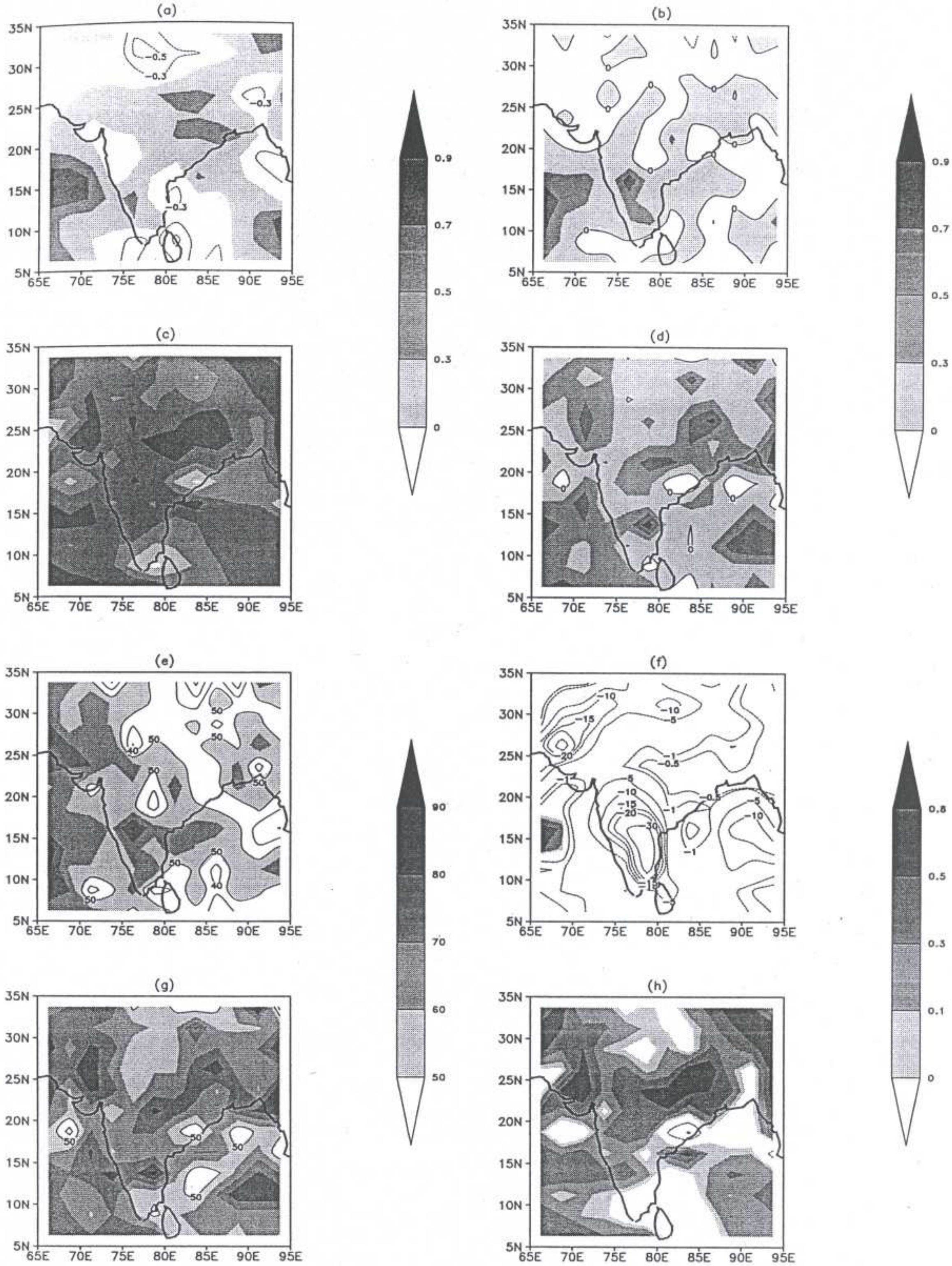


Figure 4. Same as Fig. 3 but for June.

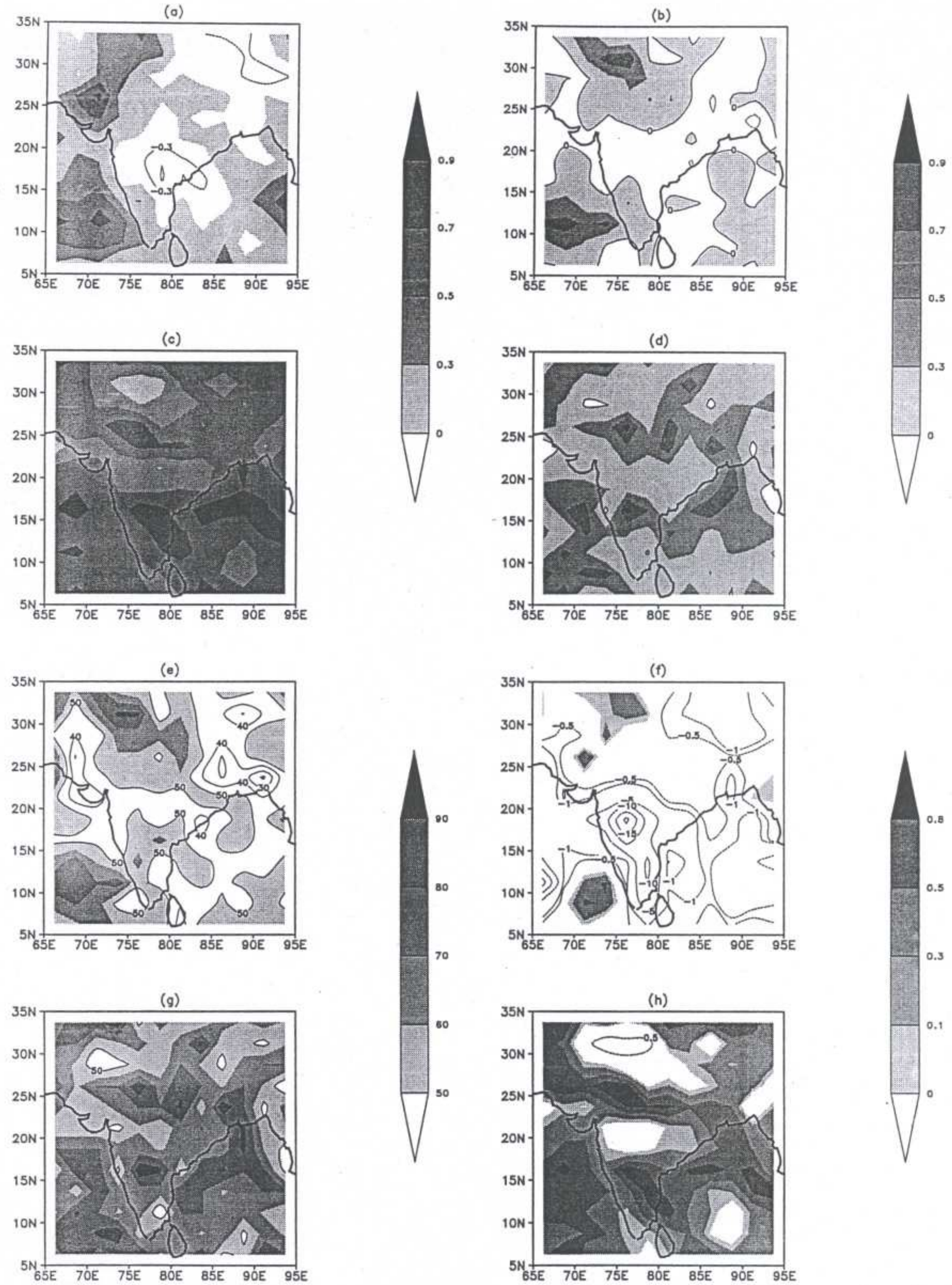


Figure 5. Same as Fig. 3 but for July.

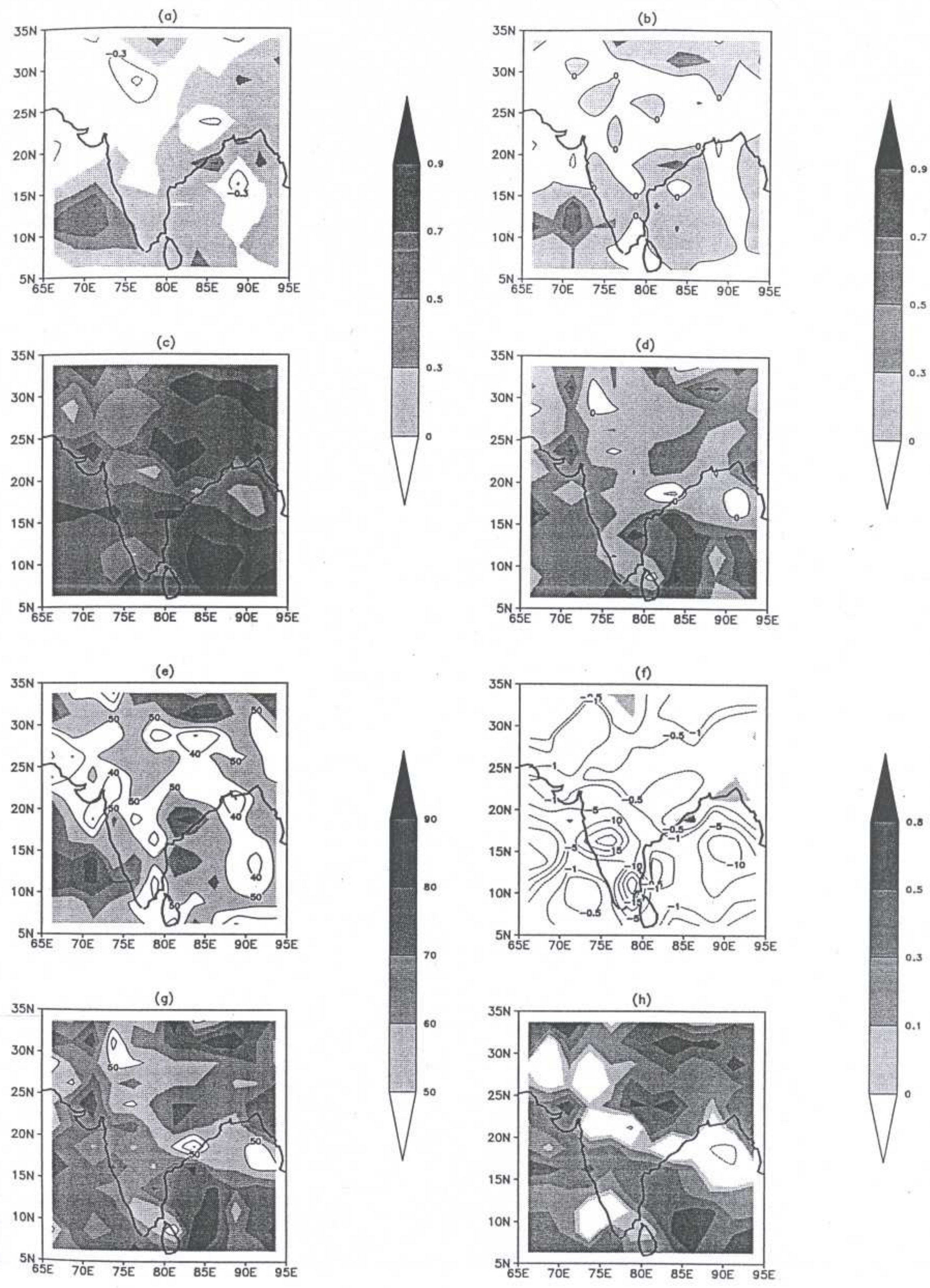


Figure 6. Same as Fig. 3 but for August.

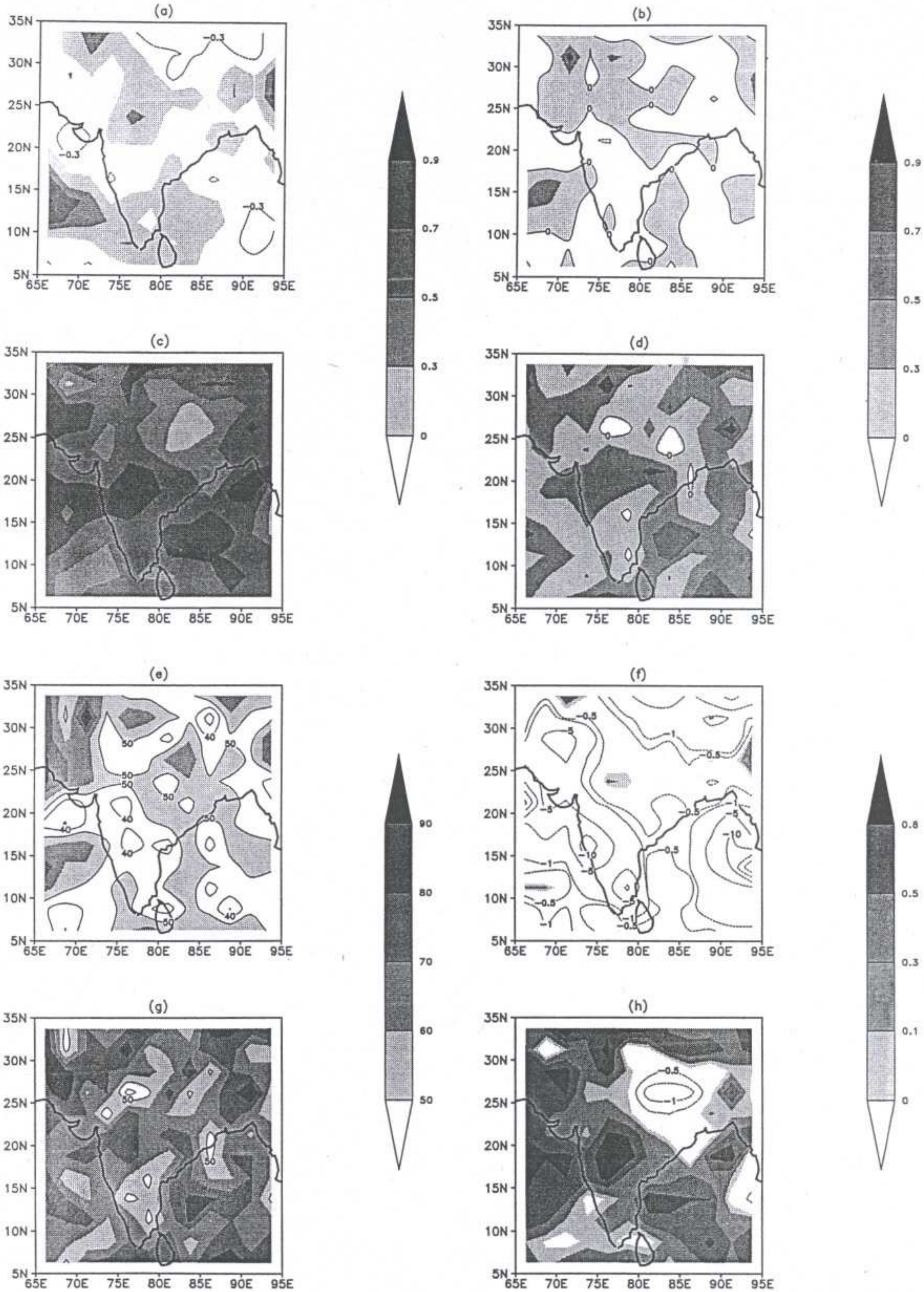


Figure 7. Same as Fig. 3 but for September.

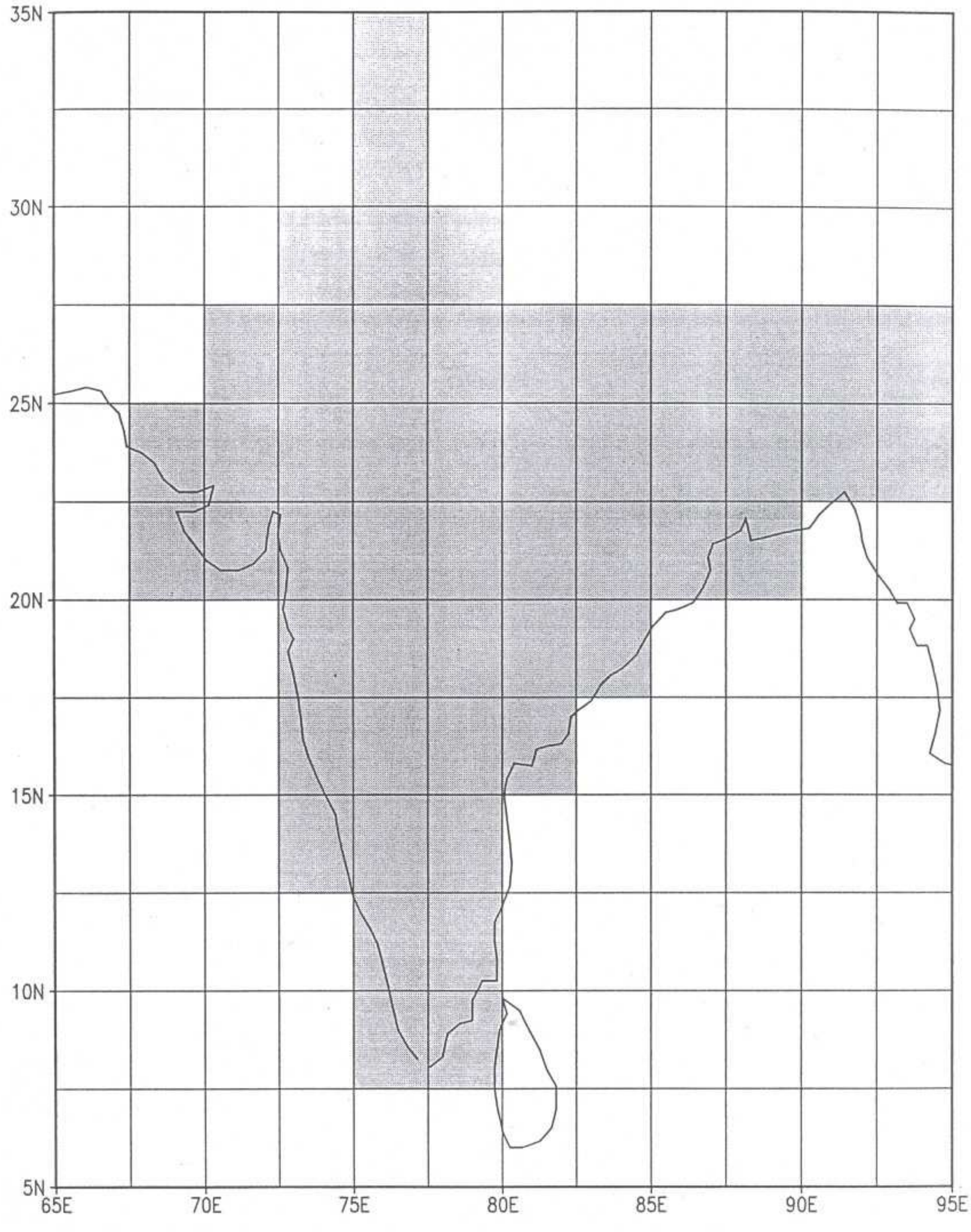


Figure 8. Shaded grids are showing the Indian land mass.

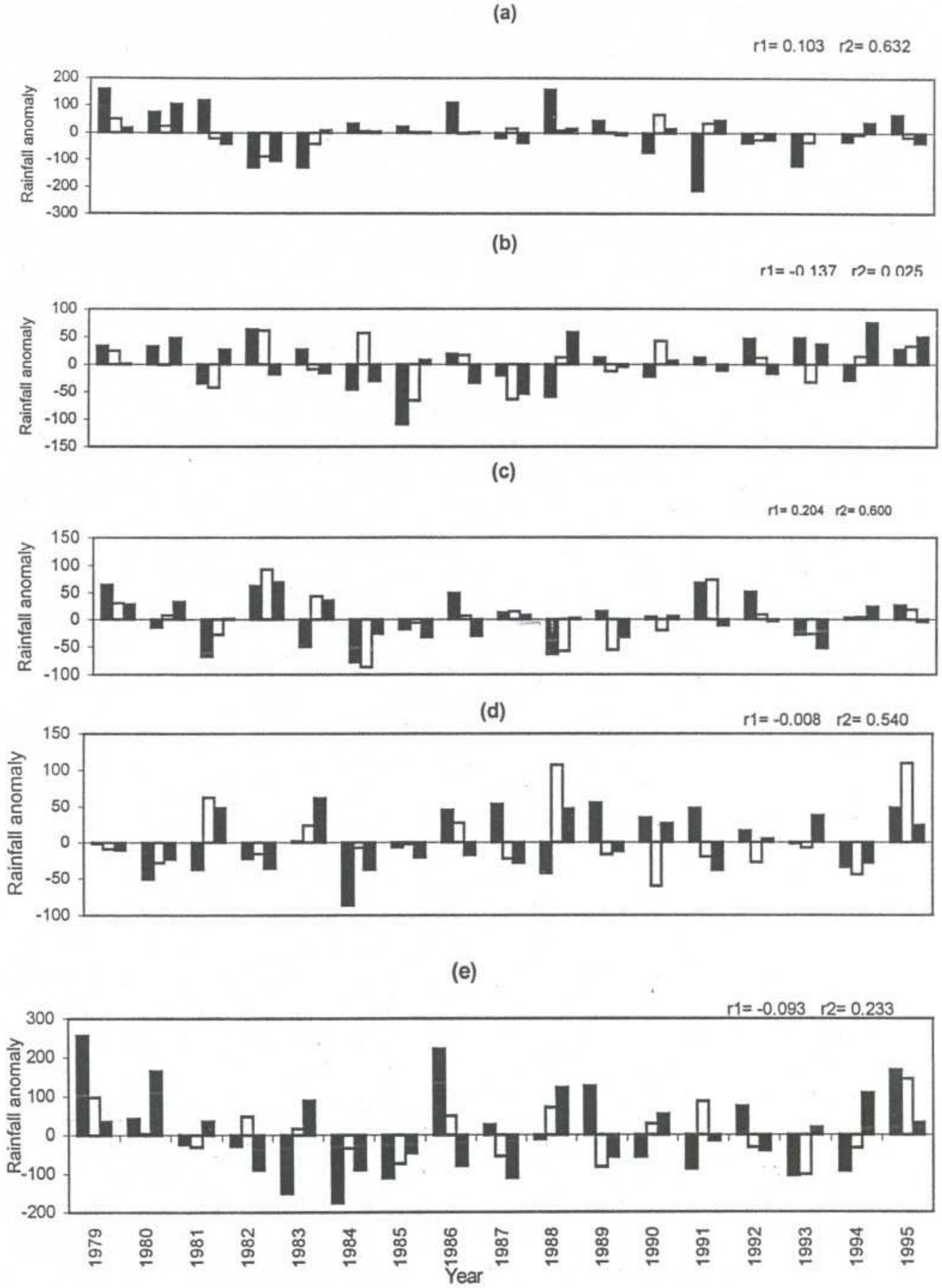


Figure 9. Time series showing departure of rainfall from normal on Indian land grids. (a) June (b) July (c) August (d) September and (e) season. Values are in mm. Partly filled bars are for model output, unfilled are ANN corrected model and fully filled are CMAP values. The values r1 and r2 showed in each figure are correlation coefficient between model and CMAP and ANN corrected model and CMAP respectively.

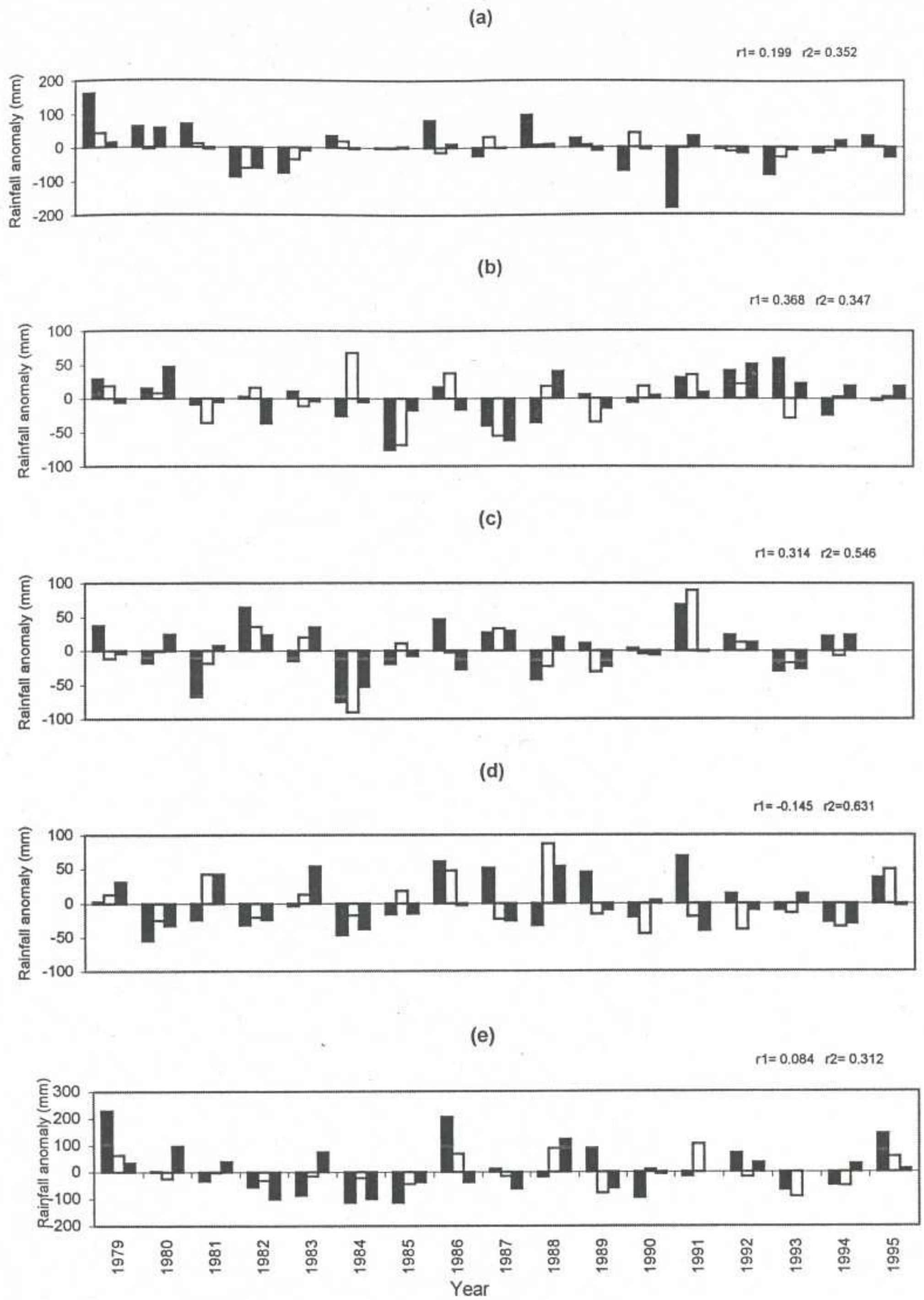


Figure 10. Same as Fig. 9 but for whole region.

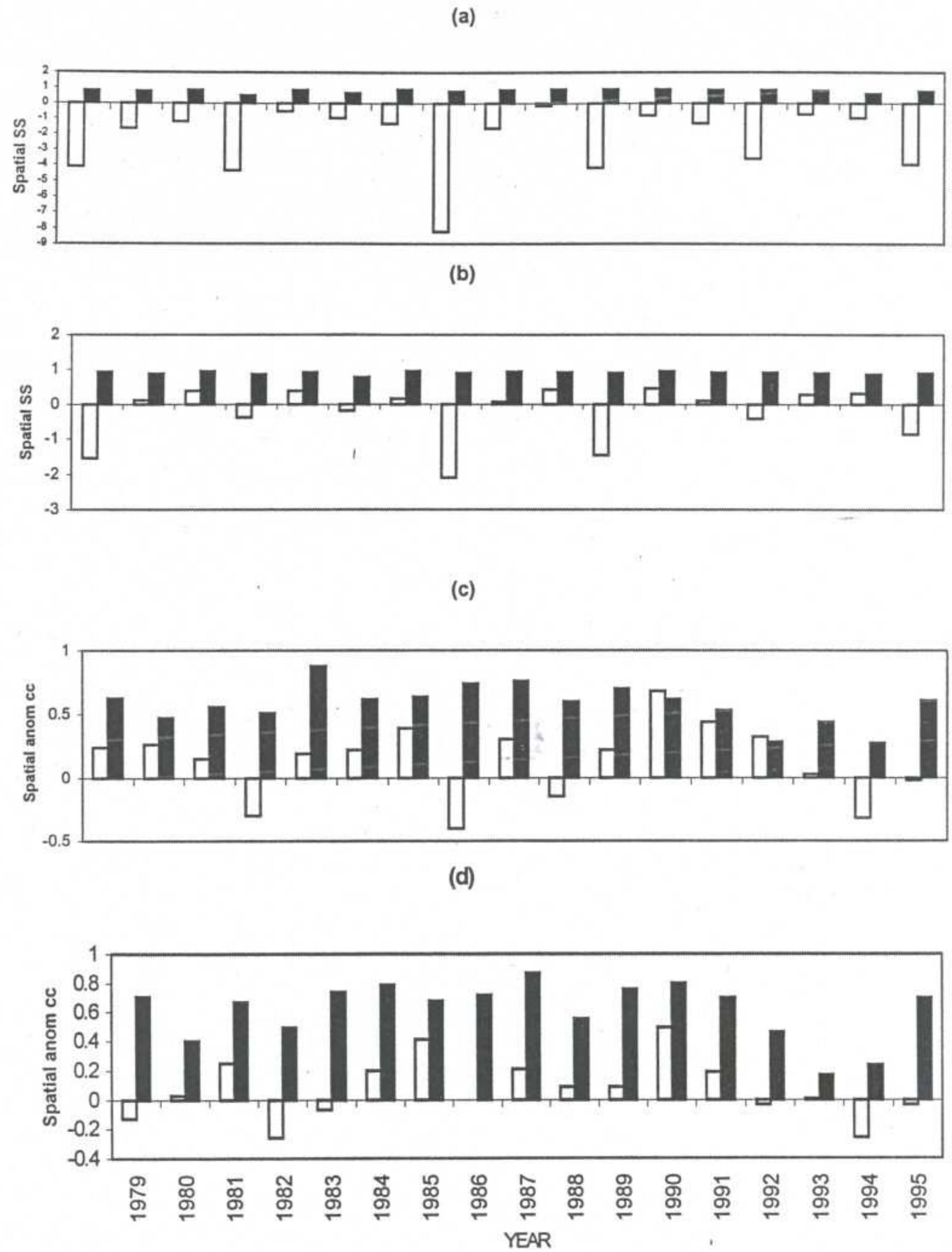


Figure 11. Time series of the values of spatial skill score (a) Indian land grids (b) whole region and spatial anomaly correlation (c) Indian land grids (d) whole region. Filled bars are for ANN corrected model and unfilled are for model.

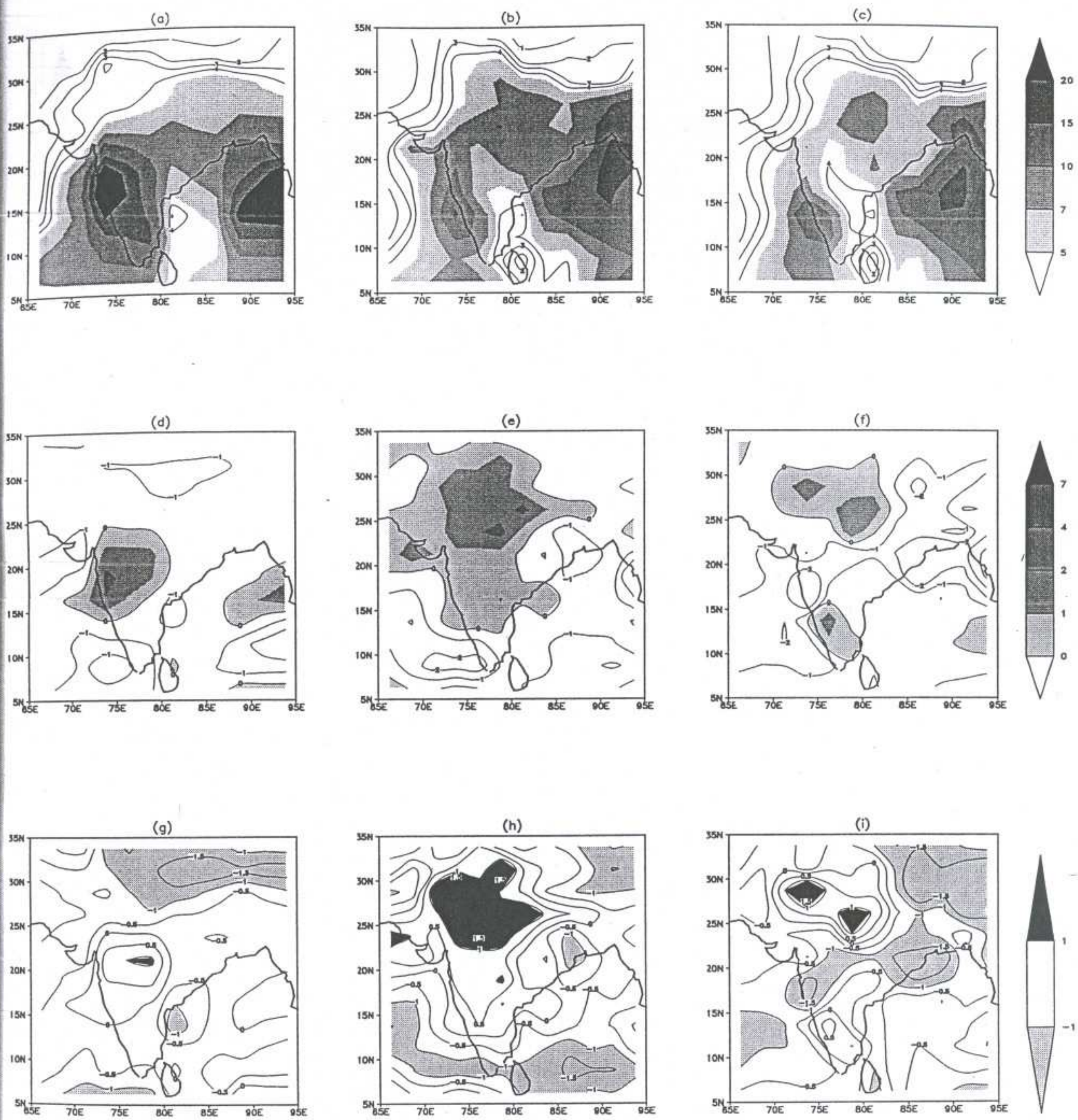


Figure 12. Spatial distribution of (a) actual model output (b) corrected model output (c) actual CMAP (d) model anomaly (e) corrected anomaly (f) observed anomaly. Values are in mm day^{-1}. Categorical spatial distribution. Dark shaded regions are above normal, light shaded below normal and not shaded are normal. (f) for model (g) for corrected and (h) for CMAP. All these values are for seasonal mean rainfall for year 1982.

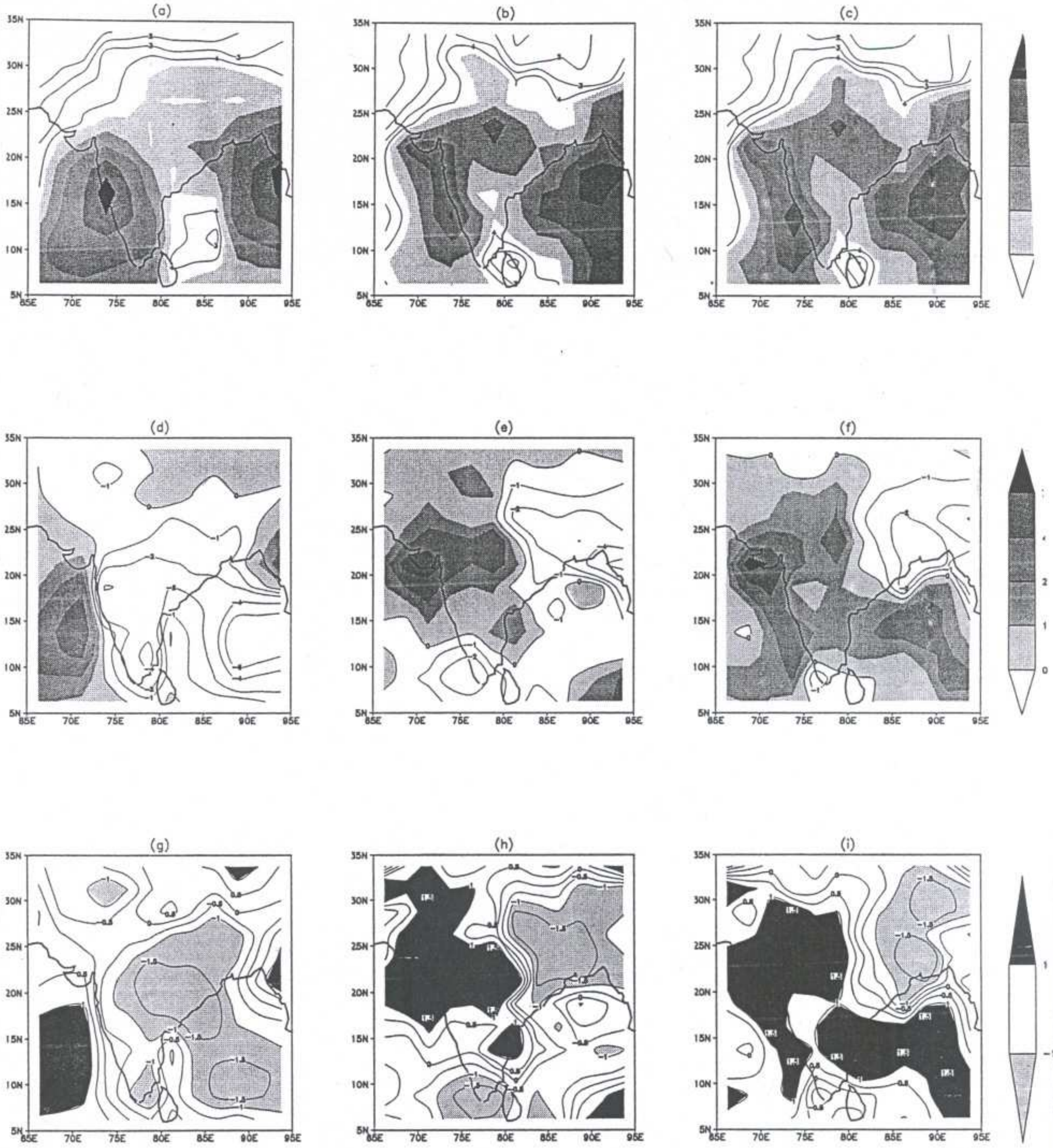


Figure 13. Same as Fig. 12 for the year 1983.

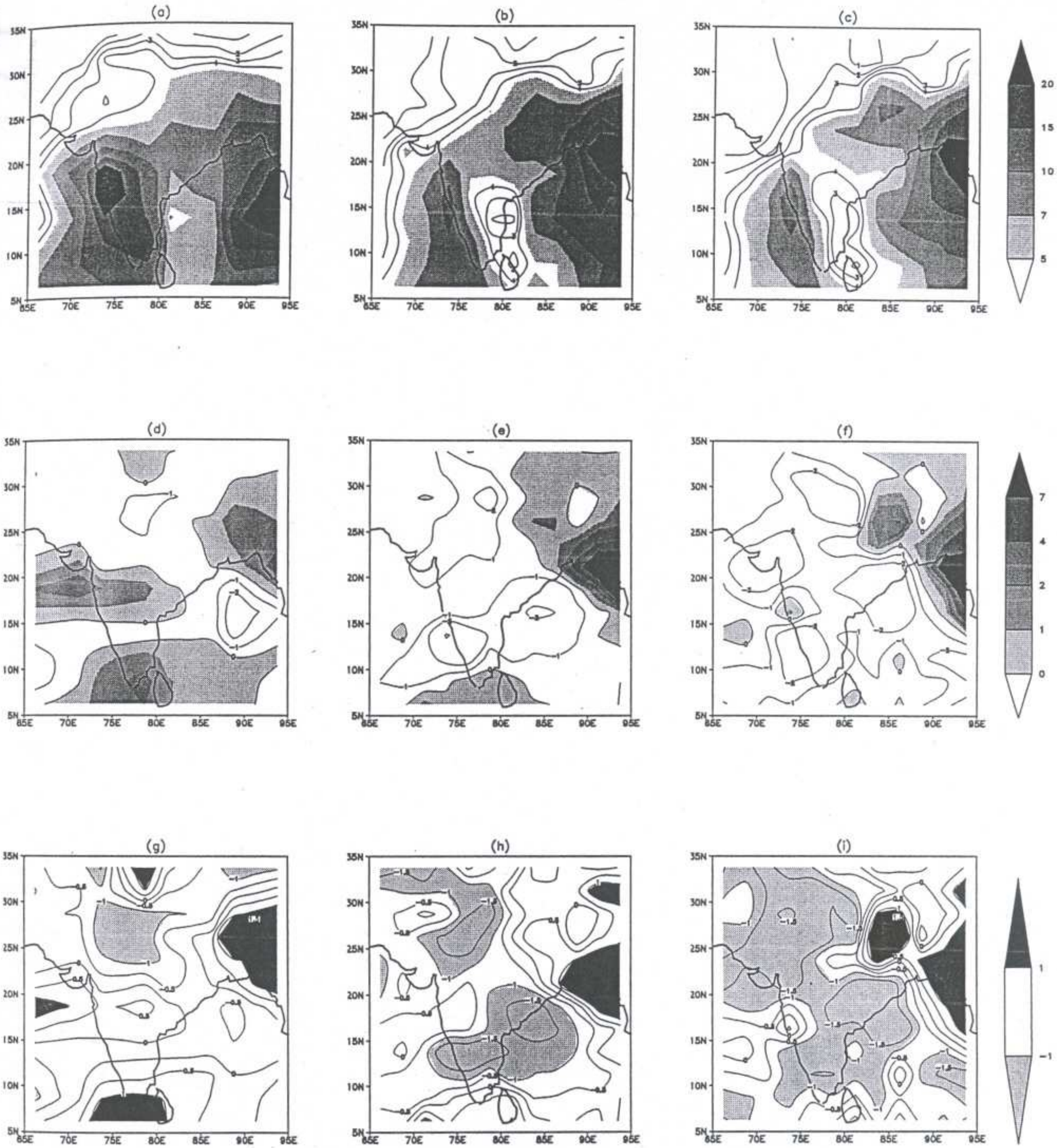


Figure 14. Same as Fig. 12 for the year 1987.

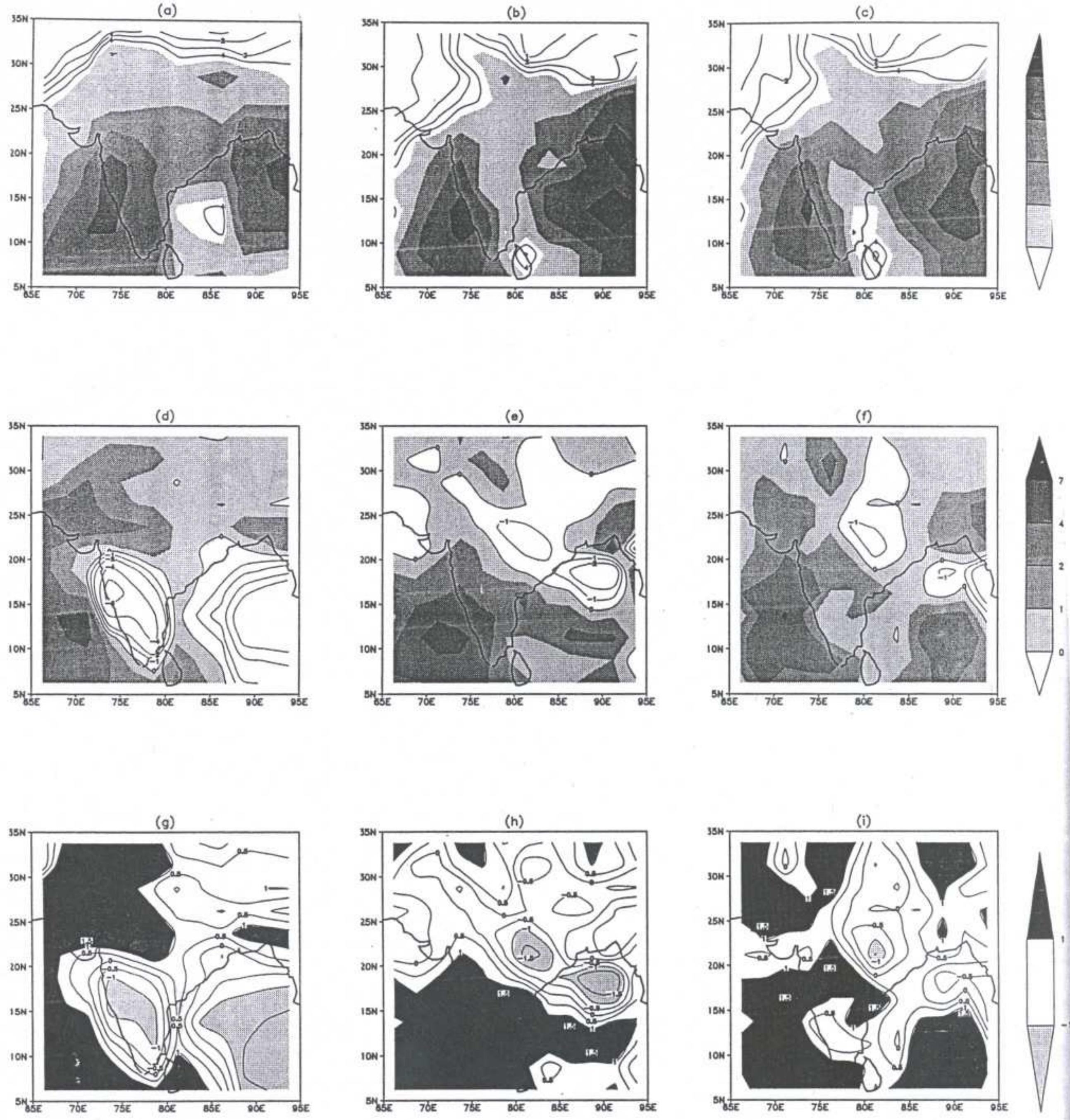


Figure 15. Same as Fig. 12 for the year 1988.

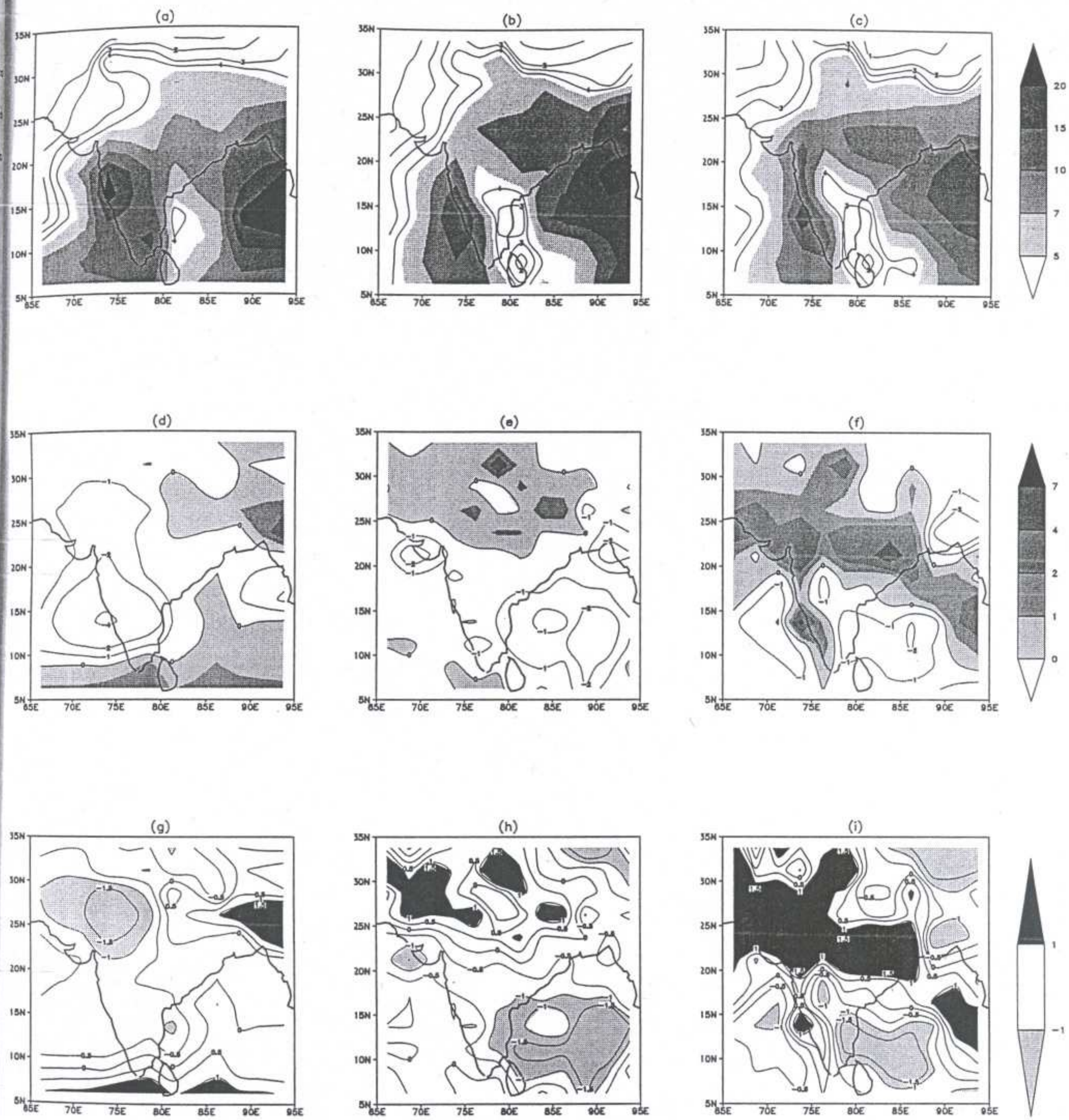


Figure 16. Same as Fig. 12 for the year 1994.

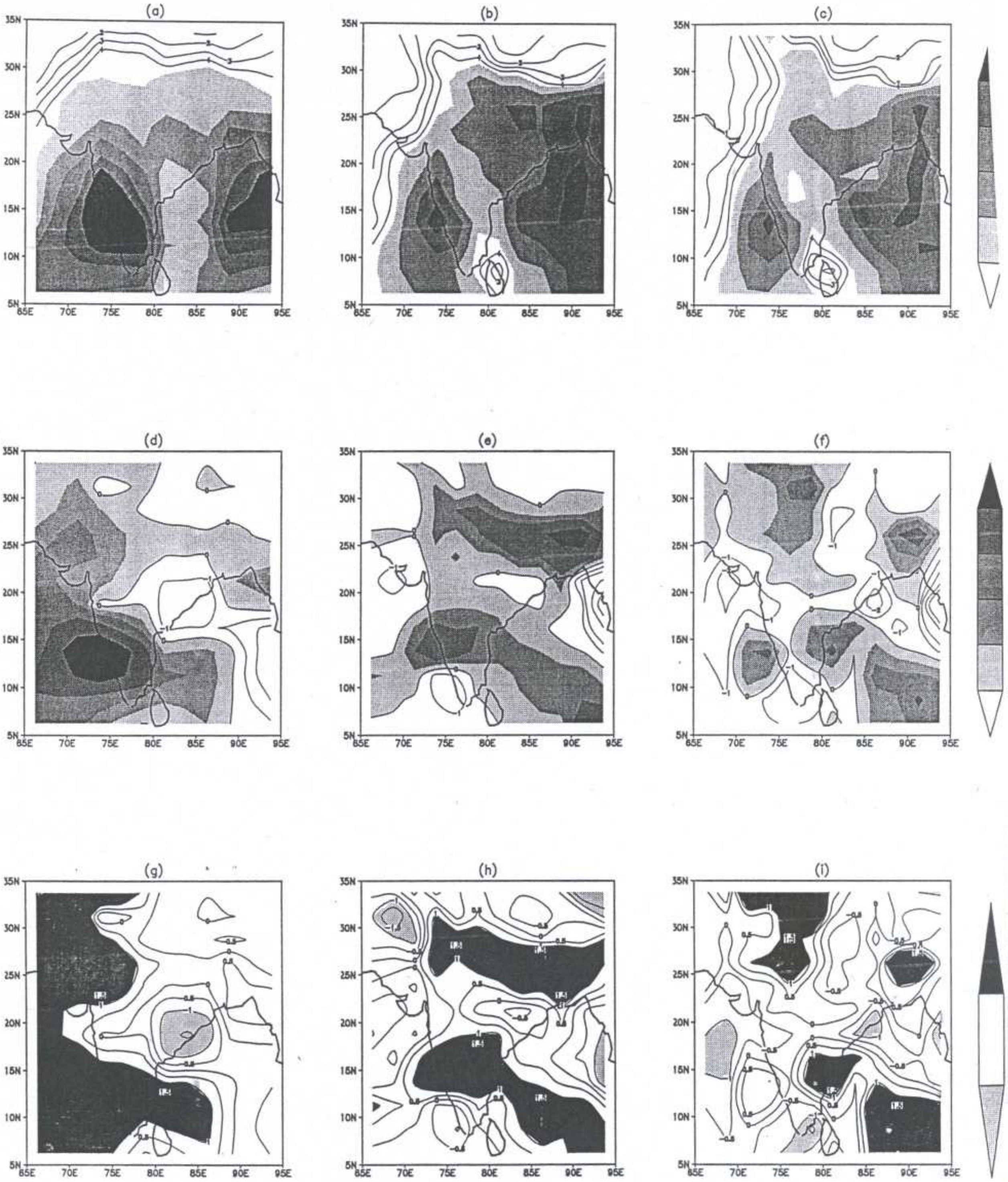
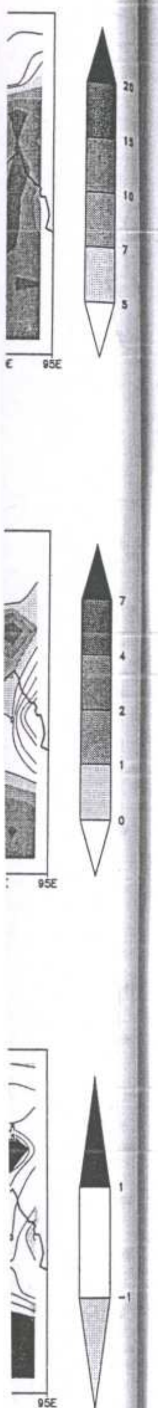


Figure 17. Same as Fig. 12 for the year 1995.



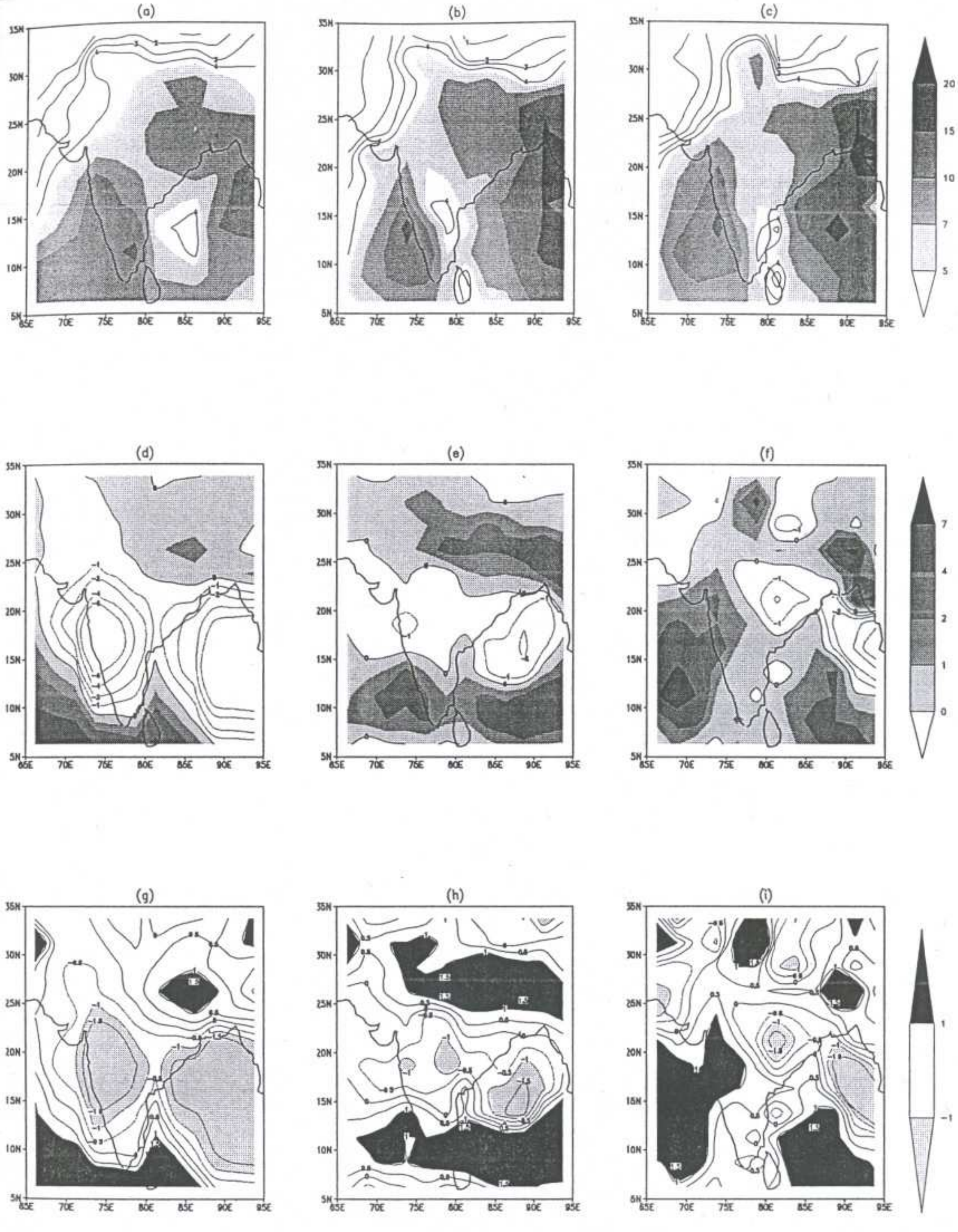


Fig. 18 Spatial distribution of (a) actual model ensemble mean (b) corrected model ensemble mean (c) actual CMAP. (d) model ensemble mean anomaly (e) corrected ensemble mean anomaly (f) observed anomaly. Values are in mm day^{-1}. (f) categorical spatial distribution of model ensemble mean. Dark shaded regions are above normal, light shaded below normal and not shaded are normal. (g) same as (f) for corrected ensemble mean and (h) same as (f) for CMAP. All these values are for seasonal mean rainfall for year 1998.

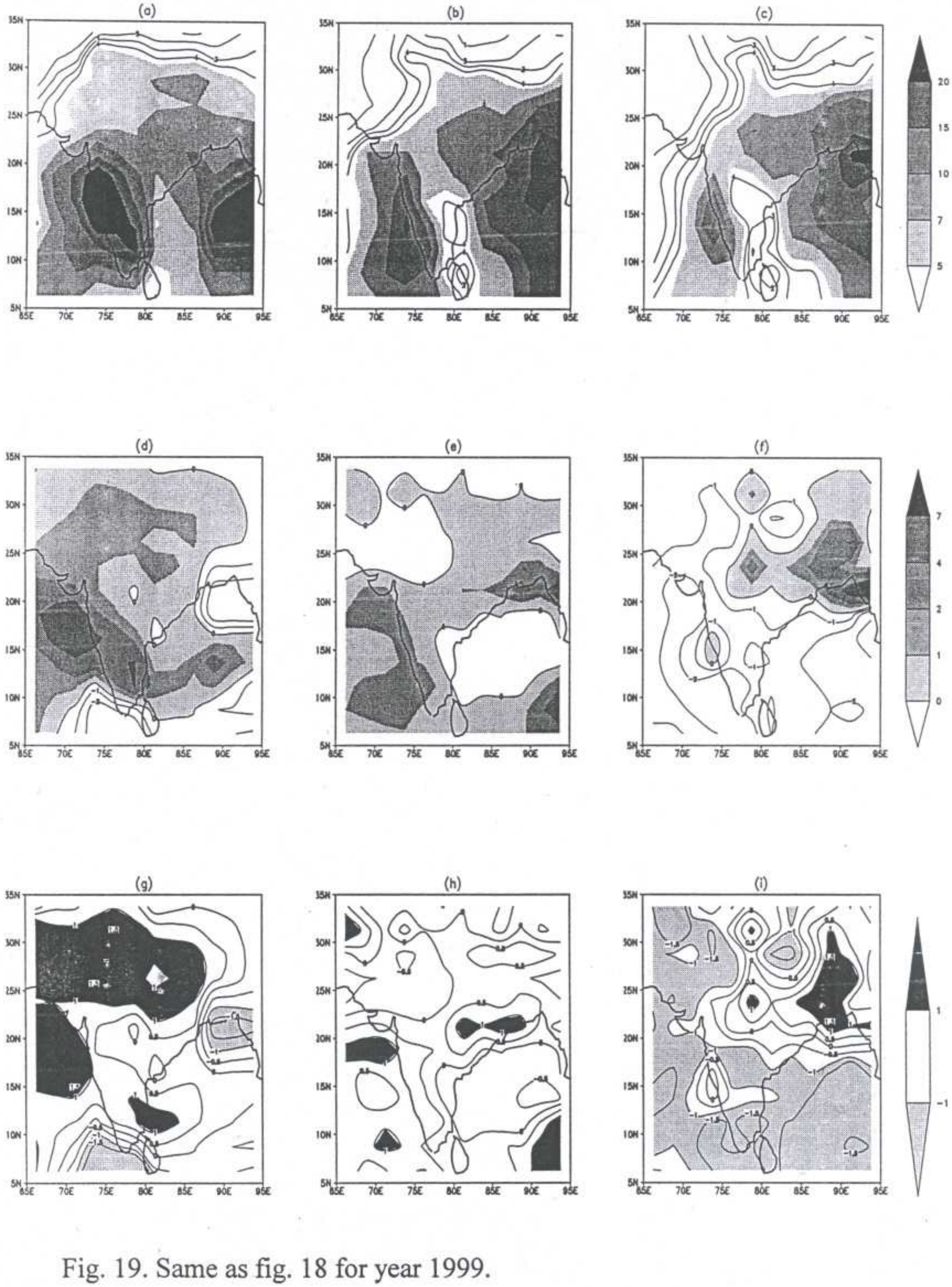


Fig. 19. Same as fig. 18 for year 1999.

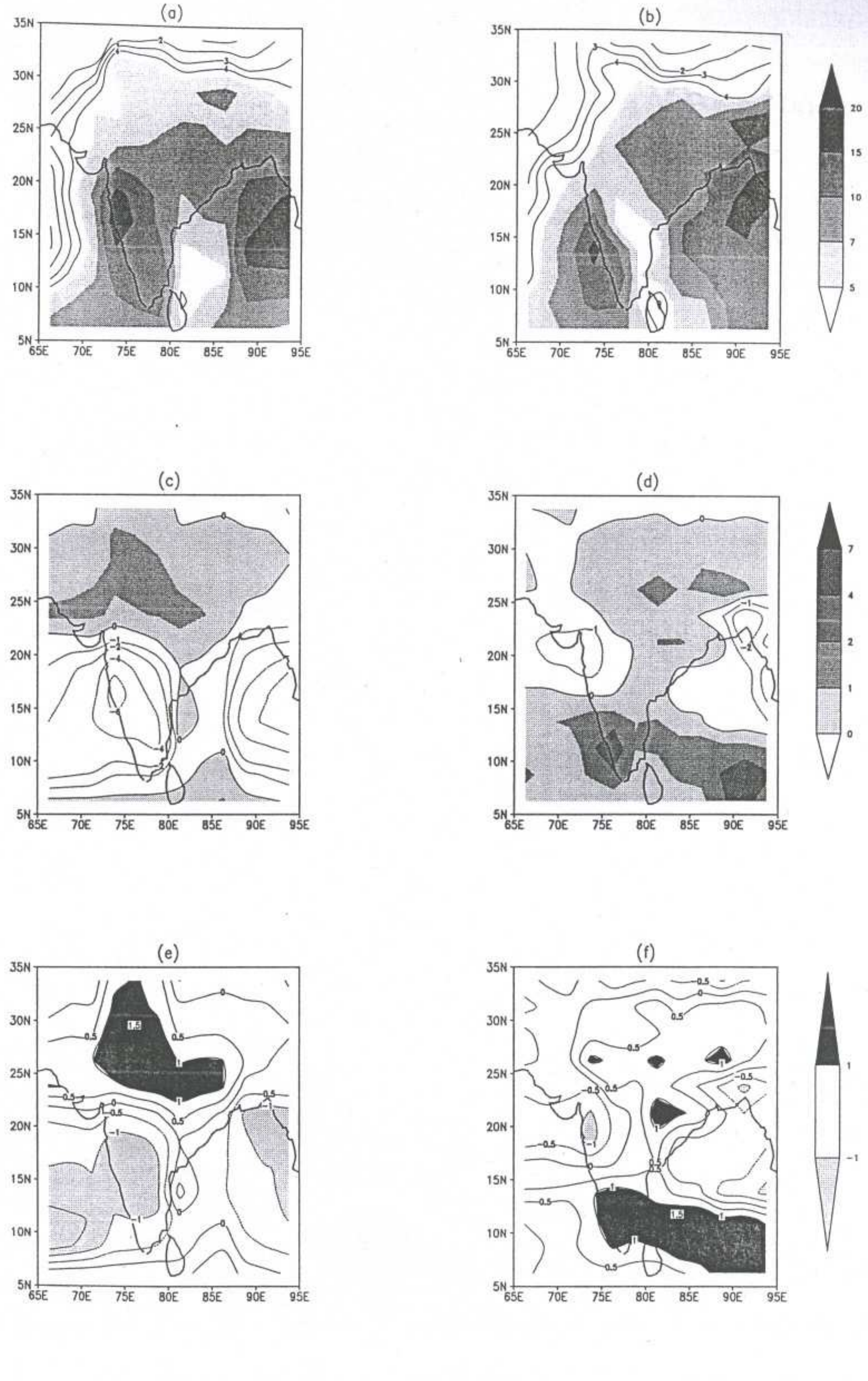


Fig. 20. Spatial distribution of (a) actual model ensemble mean (b) corrected model ensemble mean. (c) model ensemble mean anomaly (d) corrected ensemble mean anomaly. Values are in mm day^{-1}. (e) categorical spatial distribution of model ensemble mean. Dark shaded regions are above normal, light shaded below normal and not shaded are normal. (f) same as (e) for corrected ensemble mean. All these values are for seasonal mean rainfall for year 2000.

I. I. T. M. RESEARCH REPORTS

1. Energetic consistency of truncated models, *Asnani G.C.*, August 1971, RR-001.

2. Note on the turbulent fluxes of heat and moisture in the boundary layer over the Arabian Sea, *Sinha S.*, August 1971, RR-002.

3. Simulation of the spectral characteristics of the lower atmosphere by a simple electrical model and using it for prediction, *Sinha S.*, September 1971, RR-003.

4. Study of potential evapo- transpiration over Andhra Pradesh, *Rakhecha P.R.*, September 1971, RR-004.

5. Climatic cycles in India -1: Rainfall, *Jagannathan P. and Parthasarathy B.*, November 1971, RR-005.

6. Tibetan anticyclone and tropical easterly jet, *Raghavan K.*, September 1972, RR-006.

7. Theoretical study of mountain waves in Assam, *De U.S.*, February 1973, RR-007.

8. Local fallout of radioactive debris from nuclear explosion in a monsoon atmosphere, *Saha K.R. and Sinha S.*, December 1972, RR-008.

9. Mechanism for growth of tropical disturbances, *Asnani G.C. and Keshavamurty R.N.*, April 1973, RR-009.

10. Note on "Applicability of quasi-geostrophic barotropic model in the tropics", *Asnani G.C.*, February 1973, RR-010.

11. On the behaviour of the 24-hour pressure tendency oscillations on the surface of the earth, Part-I: Frequency analysis, Part-II: Spectrum analysis for tropical stations, *Misra B.M.*, December 1973, RR-011.

12. On the behaviour of the 24 hour pressure tendency oscillations on the surface of the earth, Part-III: Spectrum analysis for the extra-tropical stations, *Misra B.M.*, July 1976, RR-011A.

13. Dynamical parameters derived from analytical functions representing Indian monsoon flow, *Awade S.T. and Asnani G.C.*, November 1973, RR-012.

14. Meridional circulation in summer monsoon of Southeast Asia, *Asnani G.C.*, November 1973, RR-014.

15. Energy conversions during weak monsoon, *Keshavamurty R.N. a* *Awade S.T.*, August 1974, RR-015.

16. Vertical motion in the Indian summer monsoon, *Awade S.T. ar* *Keshavamurty R.N.*, August 1974, RR-016.

17. Semi-annual pressure oscillation from sea level to 100mb in the northe hemisphere, *Asnani G.C. and Verma R..K.*, August 1974, RR-017.

18. Suitable tables for application of gamma probability model to rainfa *Mooley D.A.;* November 1974, RR-018.

19. Annual and semi-annual thickness oscillation in thenorthern hemisphere *Asnani G.C. and Verma R.K.*, January 1975, RR-020.

20. Spherical harmonic analysis of the normal constant pressure charts in the northern hemisphere, *Awade S.T., Asnani G.C. and Keshavamurty R.N.*, May 1978, RR-021.

21. Dynamical parameters derived from analytical function representing normal July zonal flow along 87.5 °E, *Awade S.T. and Asnani G.C.*, May 1978, RR-022.

22. Study of trends and periodicities in the seasonal and annual rainfall of India, *Parthasarathy B. and Dhar O.N.*, July 1975, RR-023.

23. Southern hemisphere influence on Indian rainfall, *Raghavan K.., Paul D.K.. and Upasani P.U.*, February 1976, RR-024.

24. Climatic fluctuations over Indian region - Rainfall : A review, *Parthasarathy B. and Dhar O.N.*, May 1978, RR-025.

25. Annual variation of meridional flux of sensible heat, *Verma R..K. and Asnani G.C.*, December 1978, RR-026.

26. Poisson distribution and years of bad monsoon over India, *Mooley D.A and Parthasarathy B.*, April 1980, RR-027.

27. On accelerating the FFT of Cooley and Tukey, *Mishra S.K..*, February 1981, RR-028.

28. Wind tunnel for simulation studies of the atmospheric boundary layer, *Sivaramakrishnan S.*, February 1981, RR-029.

29. Hundred years of Karnataka rainfall, *Parthasarathy B. and Mooley D.A.*, March 1981, RR-030.

30. Study of the anomalous thermal and wind patterns during early summer season of 1979 over the Afro-Asian region in relation to the large-scale performance of the monsoon over India, *Verma R.K. and Sikka D.R.*, March 1981, RR-031.

31. Some aspects of oceanic ITCZ and its disturbances during the onset and established phase of summer monsoon studied with Monex-79 data, *Sikka D.R., Paul D.K. and Singh S.V.*, March 1981, RR- 032.

32. Modification of Palmer drought index, *Bhalme H.N. and Mooley D.A.*, March 1981, RR-033.

33. Meteorological rocket payload for Menaka-II/Rohini 200 and its developmental details, *Vernekar K.G. and Brij Mohan*, April 1981, RR-034.

34. Harmonic analysis of normal pentad rainfall of Indian stations, *Anathakrishnan R. and Pathan J.M.*, October 1981, RR-035.

35. Pentad rainfall charts and space-time variations of rainfall over India and the adjoining areas, *Anathakrishnan R. and Pathan J.M.*, November 1981, RR-036.

36. Dynamic effects of orography on the large scale motion of the atmosphere Part I : Zonal flow and elliptic barrier with maximum height of one km., *Bavadekar S.N. and Khaladkar R.M.*, January 1983, RR-037.

37. Limited area five level primitive equation model, *Singh S.S.*, February 1983, RR-38.

38. Developmental details of vortex and other aircraft thermometers, *Vernekar K.G., Brij Mohan and Srivastava S.*, November 1983, RR-039.

39. Note on the preliminary results of integration of a five level P.E. model with westerly wind and low orography, *Bavadekar S.N., Khaladkar R.M., Bandyopadhyay A. and Seetaramayya P.*, November 1983, RR-040.

40. Long-term variability of summer monsoon and climatic change, *Verma R.K., Subramaniam K. and Dugam S.S.*, December 1984, RR-041.

41. Project report on multidimensional initialization for NWP models, *Sinha S.*, February 1989, RR-042.

42. Numerical experiments with inclusion of orography in five level P.E. Model in pressure-coordinates for interhemispheric region, *Bavadekar S.N. and Khaladkar R.M.*, March 1989, RR-043.

43. Application of a quasi-lagrangian regional model for monsoon prediction, *Singh S.S. and Bandyopadhyay A.*, July 1990, RR-044.

44. High resolution UV-visible spectrometer for atmospheric studies, *Bose S., Trimbake H.N., Londhe A.L. and Jadhav D.B.*, January 1991, RR-045.

45. Fortran-77algorithm for cubic spline interpolation for regular and irregular grids, *Tandon M.K.*, November 1991, RR-046.

46. Fortran algorithm for 2-dimensional harmonic analysis, *Tandon M.K.*, November 1991, RR- 047.

47. 500 hPa ridge and Indian summer monsoon rainfall : A detailed diagnostic study, *Krishna Kumar K., Rupa Kumar K. and Pant G.B.*, November 1991, RR-048.

48. Documentation of the regional six level primitive equation model, *Singh S.S. and Vaidya S.S.*, February 1992, RR-049.

49. Utilisation of magnetic tapes on ND-560 computer system, *Kripalani R.H. and Athale S.U.*, July 1992, RR-050.

50. Spatial patterns of Indian summer monsoon rainfall for the period 1871-1990, *Kripalani R.H., Kulkarni A.A., Panchawagh N.V. and Singh S.V.*, August 1992, RR-051.

51. FORTRAN algorithm for divergent and rotational wind fields, *Tandon M.K.*, November 1992, RR-052.

52. Construction and analysis of all-India summer monsoon rainfall series for the longest instrumental period: 1813-1991., *Sontakke N.A., Pant G.B. and Singh N.*, October 1992, RR-053.

53. Some aspects of solar radiation, *Tandon M.K.*, February 1993, RR-054.

54. Design of a stepper motor driver circuit for use in the moving platform, *Dharmaraj T. and Vernekar K.G.*, July 1993, RR-055.

55. Experimental set-up to estimate the heat budget near the land surface interface, *Vernekar K.G., Saxena S., Pillai J.S., Murthy B.S., Dharmaraj T. and Brij Mohan*, July 1993, RR-056.

56. Identification of self-organized criticality in atmospheric total ozone variability, *Selvam A.M. and Radhamani M.*, July 1993, RR-057.

57. Deterministic chaos and numerical weather prediction, *Selvam A.M.*, February 1994, RR-058.

58. Evaluation of a limited area model forecasts, *Singh S.S., Vaidya S.S, Bandyopadhyay A., Kulkarni A.A, Bawiskar S.M.,Sanjay J., Trivedi D.K. and Iyer U.*, October 1994, RR-059.

59. Signatures of a universal spectrum for atmospheric interannual variability in COADS temperature time series, *Selvam A.M., Joshi R.R. and Vijayakumar R.*, October 1994, RR-060.

60. Identification of self-organized criticality in the interannual veriability of global surface temperature, *Selvam A.M. and Radhamani M.*, October 1994, RR-061.

61. Identification of a universal spectrum for nonlinear veriability of solar-geophysical parameters, *Selvam A.M., Kulkarni M.K., Pethkar J.S. and Vijayakumar R.*, October 1994, RR-062.

62. Universal spectrum for fluxes of energetic charged particles from the earth's magnetosphere, *Selvam A.M. and Radhamani M.*, June 1995, RR-063.

63. Estimation of nonlinear kinetic energy exchanges into individual triad interactions in the frequency domain by use of the cross-spectral technique, *Chakraborty D.R.*, August 1995, RR-064.

64. Monthly and seasonal rainfall series for all-India homogeneous regions and meteorological subdivisions: 1871-1994, *Parthasarathy B., Munot A.A. and Kothawale D.R.*, August 1995, RR-065.

65. Thermodynamics of the mixing processes in the atmospheric boundary layer over Pune during summer monsoon season, *Morwal S.B. and Parasnis S.S.*, March 1996, RR-066.

66. Instrumental period rainfall series of the Indian region: A documentation, *Singh N. and Sontakke N.A.*, March 1996, RR-067.

67. Some numerical experiments on roundoff-error growth in finite precision numerical computation, *Fadnavis S.*, May 1996, RR-068.

68. Fractal nature of MONTBLEX time series data, *Selvam A.M. and Sapre V.V.*, May 1996, RR-069.

69. Homogeneous regional summer monsoon rainfall over India: Interannual variability and teleconnections, *Parthasarathy B., Rupa Kumar K. and Munot A.A.*, May 1996, RR-070.

70. Universal spectrum for sunspot number variability, *Selvam A.M. and Radhamani M.*, November 1996, RR-071.

71. Development of simple reduced gravity ocean model for the study of upper north Indian ocean, *Behera S.K. and Salvekar P.S.*, November 1996, RR-072.

72. Study of circadian rhythm and meteorological factors influencing acute myocardial infraction, *Selvam A.M., Sen D. and Mody S.M.S.*, April 1997, RR-073.

73. Signatures of universal spectrum for atmospheric gravity waves in southern oscillation index time series, *Selvam A.M., Kulkarni M.K., Pethkar J.S. and Vijayakumar R.*, December 1997, RR-074.

74. Some example of X-Y plots on Silicon Graphics, *Selvam A.M., Fadnavis S. and Gharge S.P.*, May 1998, RR-075.

75. Simulation of monsoon transient disturbances in a GCM, *Ashok K., Soman M.K. and Satyan V.*,August 1998, RR-076.

76. Universal spectrum for intraseasonal variability in TOGA temperature time series, *Selvam A.M., Radhamani M., Fadnavis S. and Tinmaker M.I.R.*, August 1998, RR-077.

77. One dimensional model of atmospheric boundary layer, *Parasnis S.S., Kulkarni M.K., Arulraj S. and Vernekar K.G.*, February 1999, RR-078.

78. Diagnostic model of the surface boundary layer - A new approach, *Sinha S.*, February 1999, RR-079.

79. Computation of thermal properties of surface soil from energy balance equation using force - restore method, *Sinha S.*, February 1999, RR-080.

80. Fractal nature of TOGA temperature time series, *Selvam A.M. and Sapre V.V.*, February 1999, RR-081.

81. Evolution of convective boundary layer over the Deccan Plateau during summer monsoon, *Parasnis S.S.*, February 1999, RR-082.

82. Self - organized criticality in daily incidence of acute myocardial infarction, *Selvam A.M., Sen D., and Mody S.M.S.*, Feburary 1999, RR-083.

83. Monsoon simulation of 1991 and 1994 by GCM : Sensitivity to SST distribution, *Ashrit R.G., Mandke S.K. and Soman M.K.*, March 1999, RR-084.

84. Numerical investigation on wind induced interannual variability of the north Indian Ocean SST, *Behera S.K., Salvekar P.S. and Ganer D.W.*, April 1999, RR-085.

85. On step mountain eta model, *Mukhopadhyay P., Vaidya S.S., Sanjay J. and Singh S.S.*, October 1999, RR-086.

86. Land surface processes experiment in the Sabarmati river basin: an overview and early results, *Vernekar K.G., Sinha S., Sadani L.K., Sivaramakrishnan S., Parasnis S.S., Brij Mohan, Saxena S., Dharamraj T., Pillai, J.S., Murthy B.S., Debaje, S.B., Patil, M.N. and Singh A.B.*, November 1999, RR-087.