INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **19**: 1561–1578 (1999)

FORECASTING ALL-INDIA SUMMER MONSOON RAINFALL USING REGIONAL CIRCULATION PRINCIPAL COMPONENTS: A COMPARISON BETWEEN NEURAL NETWORK AND MULTIPLE REGRESSION MODELS

ALEX J. CANNON* and IAN G. MCKENDRY
Department of Geography, University of British Columbia, Vancouver, BC, Canada, V6T 1Z2

Received 12 June 1998
Revised 26 February 1999
Accepted 1 March 1999

ABSTRACT

Pre-monsoon principal components (PCs) of circulation fields covering the South Asian subcontinent were used as predictors for all-India summer monsoon rainfall (AISMR) over the period 1958–1993. Predictive skill of non-linear neural network models and linear multiple regression models was compared using a bootstrap-based resampling procedure. Monsoon precursor signals represented by PCs were investigated and comparisons made with a recent observational and general circulation modelling study.

Pre-monsoon PCs of the 200 hPa geopotential height field in May formed a compact, interpretable, and significant set of predictors for AISMR. Predictive skill was comparable to or better than that reported in prior modelling studies, each of which used optimized sets of regional and global predictors. No improvement was noted when using data from multiple atmospheric levels, and skill at lead times more than 1 month prior to monsoon onset in June was poor. For May predictors there were only small differences in skill between the neural network and multiple regression models, although the neural network results at longer lead times tended to be better than those shown by multiple regression. Interestingly, the 850 hPa PCs in January showed a maximum in predictive skill that was only evident in the neural network model results. The strength of this relationship suggests that further investigation into the use of January 850 hPa predictors for the long-range forecasting of AISMR is warranted. Copyright © 1999 Royal Meteorological Society.

KEY WORDS: all-India summer monsoon rainfall; long-range forecasting; neural network model; principal component analysis

1. INTRODUCTION

The Indian monsoon is characterized by winter–summer reversals in various aspects of the atmospheric circulation over the South Asian subcontinent (Pant and Kumar, 1997). Associated with these seasonal changes in circulation patterns is the abrupt onset of heavy precipitation over India in June; this heavy rainfall tends to persist through the month of September and accounts for the majority of the country's total annual rainfall. Since the late 1800s several studies have attempted the long-range prediction of all-India summer monsoon rainfall (AISMR) (Kumar *et al*., 1995). The primary method of forecasting during this period has been through the use of empirical and statistical models. Recent attempts at dynamical simulations of the monsoon circulation with general circulation models (GCMs) have met with some success (Lau and Yang, 1996; Yang *et al*., 1996). However, these models are currently not able to give accurate long-range forecasts of AISMR (Pant and Kumar, 1997).

Most statistical models for long-range forecasting of monsoon rainfall are based on the linear regression of one or more predictors with AISMR. Recently, two studies have successfully applied non-linear methods to this problem. Gowarikar *et al*. (1991) used a parametric, multiple power regression model with

* Correspondence to: Department of Geography, University of British Columbia, 217-1984 West Mall, Vancouver, BC, Canada, V6T 1Z2. Tel.: +1 604 8222663; fax: +1 604 8226150; e-mail: acannon@geog.ubc.ca

CCC 0899–8418/99/141561–18$17.50

16 input variables to predict AISMR. In this model non-linear interactions are considered through the iterative determination of the exponent for each predictor in the regression equation. As Kumar *et al.* (1995) pointed out, this model suffers from the presence of significant intercorrelations between the predictors, possible overfitting, and the use of predictors originally developed through the analysis of linear correlations with AISMR. Using a different approach, Navone and Ceccatto (1994) applied feed-forward neural network techniques to the prediction of monsoon rainfall. The first of their neural networks uses two atmospheric predictors—500 hPa ridge position in April and sea-level pressure (SLP) tendency from January to April at Darwin. The second uses summer rainfall totals from the 7 previous years as predictors. Limited verification of the results of these models has been completed, and there is some doubt as to the ability of the network to make accurate predictions based on previous summer rainfall totals alone. Kumar *et al.* (1995) called for further work to verify that this method does in fact outperform linear methods.

Irrespective of the type of forecast model used, identification of robust and statistically significant pre-monsoon predictors for AISMR continues to be of great importance. Historically, predictors have tended to fall into four general categories (Kumar *et al.*, 1995): (1) regional conditions; (2) El Niño–Southern Oscillation (ESNO) indicators; (3) cross-equatorial flow; and (4) global/hemispheric conditions. Of these four, recent studies show regional circulation conditions to be of greatest importance in forecasting AISMR (Parthasarathy *et al.*, 1988; Kumar *et al.*, 1997). Recognizing that significant intercorrelations exist between circulation predictors, methods for reducing redundant information in the set of predictor variables have been sought. Principal component analysis (PCA) provides a promising means by which this may be achieved. For example, Kumar *et al.* (1995) used PCA on 19 AISMR predictors and showed that a single PC may account for 46% of the predictor set variance as well as being strongly and significantly correlated with AISMR. In a different context, Tangang *et al.* (1998) have successfully predicted ENSO conditions using PCs of SLP anomaly fields.

Due to the promise shown by the combination of PCA-reduced predictors and non-linear neural network models in the forecasting of ENSO-related sea-surface temperature anomalies (Tangang *et al.*, 1998), this study investigates the applicability of such methods in the context of AISMR. Specific goals are:

1. determination of robust PCA-reduced circulation predictors for AISMR;
2. comparison of standard multiple regression and feed-forward neural network models as forecast tools for AISMR;
3. examination of circulation patterns associated with AISMR anomalies.

Two aspects of the approach used are novel in the context of AISMR prediction. First, regional predictors based on gridded circulation fields from several atmospheric levels are used to investigate the extent to which a three-dimensional representation of the atmospheric circulation improves forecasts (Kidson, 1997). Second, predictive skill of the neural network and linear regression models is evaluated using a bootstrap-based resampling technique (LeBaron and Weigend, 1998).

2. DATA AND PRELIMINARY ANALYSES

2.1. Sea-level pressure and geopotential height fields

Several studies identify significant correlations between surface and upper level circulation parameters in the Indian region and summer monsoon rainfall. Parthasarathy *et al.* (1988) found that SLP tendency (for months March, April, May–December, January, Febraury, MAM–DJF) at surface stations in India is negatively correlated with AISMR. Bhalme *et al.* (1986) identified negative correlations between both May SLP and 850 hPa heights and AIMSR. Several studies show the importance of the circulation at other pressure levels. For example, Mooley *et al.* (1986) found that the pre-monsoon position of the 500 hPa ridge over India shows a positive correlation with AISMR. More recently, Parthasarathy *et al.*

(1991a) investigated the relationship between 200 hPa meridional wind indices over India and summer monsoon rainfall. In these studies the state of the circulation at each level of the atmosphere was represented by a single index value. While these indices have been reasonably effective in the forecasting of AISMR, they offer only approximate representations of the actual atmospheric circulation over the region.

In this study, gridded data from four atmospheric levels over the entire Indian region were considered. Daily averaged SLP, 850 hPa, 500 hPa, and 200 hPa geopotential height data (2.5° × 2.5° resolution) were obtained from NCEP/NCAR reanalysis output (Kalnay *et al.*, 1996) provided by the Climate Diagnostics Center (NOAA). A subset covering the region from 62.5°E to 95°E and 7.5°N to 35°N was extracted and monthly averages computed for the 36-year period from 1958 to the end of 1993. The data were then spatially averaged to a resolution of 5° × 5°. SLP and geopotential height anomalies were computed by subtracting the climatological monthly mean (1958–1993) from each of the monthly average values. Pre-monsoon months (DJFMAM) were extracted from the set, leaving a 216-month series of anomaly fields (42 grid points) for each atmospheric level.

From the anomaly fields, 15 sets of data were formed covering all possible combinations of the four included atmospheric levels (Table I). An S-mode PCA was then performed on the correlation matrix $\boldsymbol{R}$ of each set of pre-monsoon circulation anomalies:

$$\boldsymbol{R} = \frac{1}{m-1}\boldsymbol{X}_s'\boldsymbol{X}_s \tag{1}$$

where m is the number of months and $\boldsymbol{X}_s$ is the matrix of standardized circulation anomalies under consideration. The correlation matrix was used to ensure that the different levels included were weighted equally in the PCA.

For each set of circulation data, the correlation matrix was decomposed into a matrix of eigenvectors $\boldsymbol{U}$ and a diagonal matrix of eigenvalues $\boldsymbol{D}$. PC loadings $\boldsymbol{F}$ and standardized PC scores $\boldsymbol{Z}_s$ were then computed:

$$\boldsymbol{F} = \boldsymbol{U}\boldsymbol{D}^{0.5}. \tag{2}$$

$$\boldsymbol{Z}_s = \boldsymbol{X}_s\boldsymbol{U}\boldsymbol{D}^{-0.5}. \tag{3}$$

The number of PCs to retain as predictors was determined using the rule-N test outlined in Overland and Preisendorfer (1982). In this procedure, a large sample of independent variables is generated from a Gaussian distribution of random numbers. These variables are formed into matrices of the same

Table I. Results of the PCA of SLP, 850 hPa, 500 hPa, and 200 hPa standardized circulation anomaly fields

Included level(s)	Retained PCs	Cumulative variance (%)
SLP	3	87.9
SLP, 850 hPa	3	87.4
SLP, 500 hPa	6	91.5
SLP, 200 hPa	6	91.6
SLP, 850, 500 hPa	6	91.6
SLP, 850, 200 hPa	6	90.9
SLP, 500, 200 hPa	6	90.8
SLP, 850, 500, 200 hPa	6	90.7
850 hPa	2	88.6
850, 500 hPa	5	92.5
850, 200 hPa	5	92.0
850, 500, 200 hPa	6	93.1
500 hPa	4	95.7
500, 200 hPa	5	94.3
200 hPa	4	96.2

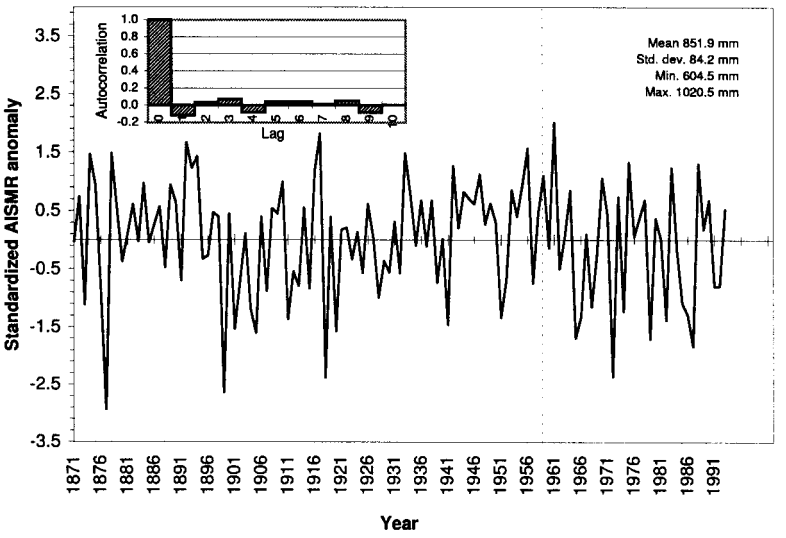


Figure 1. Standardized all-India summer monsoon rainfall anomalies (1871–1993) with autocorrelations to lag 10. Years to the right of the dotted line (1958–1993) were used in the current study

dimension as the data being analysed and the correlation matrix for each set is computed. Eigenvalues are then calculated for each correlation matrix, and the percent variance accounted for by each component is determined. PCs of the actual data are retained if the percent variance accounted for by a given component exceeds the percent variance accounted for by the 95th percentile of the corresponding resampled component distribution.

The number of significant PCs found for each set of circulation anomaly fields analysed is given in Table I, as is the cumulative variance accounted for by the retained components. In all cases the variance accounted for by the retained PCs exceeds 87%.

Preliminary inspection of the retained component loadings showed patterns characteristic of those found by Buell (1975). Given that the primary purpose of the PCA was data reduction, no further attempt at interpretation of the components was attempted in the initial data processing step. Interpretation of selected PC solutions is offered in the 'Pre-monsoon circulation relationships' section.

Each of the 15 sets of PC scores was then separated by month, forming predictors at lead times from 0 (using May PC scores) to 5 months (using December PC scores).

2.2. AISMR

The AISMR data set consists of area-weighted averages of summer (JJAS) rainfall totals from 306 district rain gauge stations. Approximately 90% of the country is accounted for, with a number of hilly regions removed from the analysis. Originally developed and described by Parthasarathy *et al.* (1987) for the period 1871–1984, the data set was recently extended in Parthasarathy *et al.* (1994) to include years through 1993. Inspection of the time series over this period and of the autocorrelations of data (Figure 1) reveals the significant interannual variability and lack of persistence of AISMR. This lack of persistence has hindered attempts at pure time series forecasting of the phenomenon.

AISMR totals for years from 1958 to 1993 were extracted from the entire time series. The mean value over this series was removed from each summer's reported total. These anomalies were then standardized using the sample S.D. over the subset data period. For use in the 'Pre-monsoon circulation relationships' section, years with AISMR values greater than or less than 1 S.D. from the subset mean are defined as having had excess or deficient rainfall totals, respectively.

3. NEURAL NETWORK MODELS

3.1. Feed-forward neural network design

Neural networks have been successfully used in non-linear regression, classification, and clustering applications (Sarle, 1994). In this study, a single hidden layer feed-forward backpropagation neural network model was used to perform non-linear regression. Details on this form of model can be found in Bishop (1995).

A hyperbolic tangent activation function was used in the hidden layer allowing the neural network to represent arbitrary non-linear functions. For regression purposes the activation function for the output layer is linear. Training of the network involves minimizing the sum of squared error E_{SSE} over the set of training cases,

$$E_{SSE} = \frac{1}{2} \sum (z - z_{obs})^2 \tag{4}$$

where z is the output of the neural network and z_{obs} is the observed output value for a given input pattern. The minimization process, which involves the iterative adjustment of the network weights and biases, can be performed by a variety of numerical optimization algorithms (see Gibb, 1996, for a review).

For this study, the resilient backpropagation (Rprop) algorithm, developed by Riedmiller and Braun (1993), was used in training. Rprop, a local optimization technique with a sign-dependent and weight-specific update rule, performs well on a variety of problems and is relatively insensitive to user-adjustable parameters (Riedmiller, 1994). The default value for the initial update parameter Δ_0 (as described in Riedmiller, 1994) was used in this study. The $\Delta_{\max}$ parameter, the maximum step size for weight and bias adjustment, was set to 0.001, lower than the suggested default value. Limiting the maximum step size tended to slow model training but allowed the minimum validation set error (described in the 'Avoid overfitting' section) to be estimated with greater accuracy.

In each of the neural network models considered, five neurons were used in the hidden layer. The initial weight and bias values for each network layer were chosen randomly from a uniform distribution ranging from -0.5 to 0.5.

3.2. Ensemble methods

Unlike standard multiple regression models, neural network outputs tend to be relatively unstable. The degree of instability depends on the input variables, the choice of training set data, the number of hidden neurons, the training method, and the initial weight and bias values. Recognition of the inherent instability of this class of models has led to the development of ensemble neural network methods (Perrone and Cooper, 1993). In ensemble methods outputs from a number of neural networks, each trained from different initial conditions, are combined to help improve model stability and accuracy.

In this study, the final output of each trained ensemble was taken to be the average of the outputs of the individual members. A resampling method known as bootstrap aggregation (bagging) was used (Breiman, 1996) to generate random splits of the data into training, validation, and test sets. In bagging, ensemble training sets are obtained by sampling with replacement from all available input/output pairs. Resampling is stopped when the number of cases in the training set equals the number of available cases. In using this method approximately 37% of available cases are excluded from a given training set. These excluded or out-of-bag cases are then randomly partitioned into validation and test data sets which can be used to aid in training or to estimate generalization error.

3.3. Avoiding overfitting

Given a sufficient number of hidden neurons, a feed-forward backpropagation neural network with a single hidden layer can represent any continuous function with arbitrary accuracy (Hornik *et al.*, 1989). In regression problems, where the transfer function is developed on a training set of known inputs and

outputs, care must be taken to ensure that the optimized network produces reasonable outputs for previously unseen input patterns.

Given the limited amount of circulation data available in this study, strict measures were taken to avoid overfitting. The one-step weight decay method outlined in Finnoff *et al.* (1993) was used to this end.

In this method, the error function minimized during training is modified from Equation (4) to include two penalty terms (Bishop, 1995):

$$E_c = E_{SSE} + \frac{\alpha_1}{2}\sum ({}^1w)^2 + \frac{\alpha_2}{2}\sum ({}^2w)^2 \tag{5}$$

where ${}^1w$ and ${}^2w$ are the hidden and output layer weights and α_1 and α_2 are weight decay constants that control the degree to which the penalty terms influence the network solution. These extra terms penalize large weights in the hidden and output layers and lead to final weight values that are small in magnitude. This has been shown to improve the generalization properties of the trained network transfer function (Hinton, 1987).

Initially, training of the network is performed with the weight decay coefficients α_1 and α_2 set to zero. During training, E_{SSE} is monitored on both the training data set and an independent data set used to estimate the generalization performance of the model. The network is trained to convergence on the training set while the point at which the minimum error on the independent validation set occurs is noted. The network is then reinitialized with the weights and biases corresponding to this minimum and training is restarted with α_1 and α_2 set to non-zero values. Training is then continued until E_c converges.

Prior to the training of the neural network ensembles, optimum values for α_1 and α_2 were determined. For each of the 15 predictor sets, five neural networks were initialized using the procedure outlined in the 'Ensemble methods' section. Values of the weight decay coefficients were then chosen such that the average validation set error over these members was minimized.

4. RESULTS

4.1. Linear relationships between regional circulation PCs and AISMR

Values of the multiple correlation coefficient R were calculated for each of the 15 circulation PC predictor sets at each lead time (Table II) in order to assess the strength of linear relationships between the circulation PCs and AISMR over the period 1958–1993. Given the unequal number of components retained in each set, care should be taken in the interpretation of these values.

Considering all possible combinations of atmospheric levels and lead times, only the circulation PCs in May showed statistically significant correlations with AISMR. In each case, large reductions in R were noted as the lead time was increased. Of the single level predictors, 200 hPa PCs in May showed the highest correlation with AISMR. The combined level predictor sets in May exhibited similar multiple correlation coefficient values. In no case, however, did the addition of data from multiple levels significantly increase R above the value shown by May 200 hPa PCs alone. While not significantly correlated with AISMR, values of R in December and January tended to be higher than at the other lead times greater than 1 month. This pre-monsoon maximum in the strength of the circulation–AISMR relationship occurred at the surface in December and in January for the upper atmospheric fields.

To assess the stability of May circulation–rainfall relationships, values of R for the predictor PC sets and AISMR were calculated using 15- and 21-year sliding windows—the same as in Parthasarathy *et al.* (1991a). The results for single level PC predictors are shown in Figure 2. As noted in other studies (for example, Parthasarathy *et al.*, 1991b; Hastenrath and Greischar, 1993) secular variations in the strength of the circulation AISMR relationship appear to be present on decadal time scales. In recent years, declines in correlation between various circulation parameters and AISMR have been noted (Parthasarathy *et al.*, 1991a; Singh *et al.*, 1995). Based on the 21-year window width, however, the relationship between May circulation PCs and AISMR appears to have remained stable since 1958. Only

the 500 hPa PC set showed marked variability on this time scale. Decreases in R for the 200 hPa and 500 hPa PCs were noted for the 15-year sliding windows centred on the early to mid-1980s. In contrast, SLP and 850 hPa PCs have remained relatively stable predictors of AISMR over the latter part of the data record.

4.2. Ensemble neural network and multiple regression predictions of AISMR

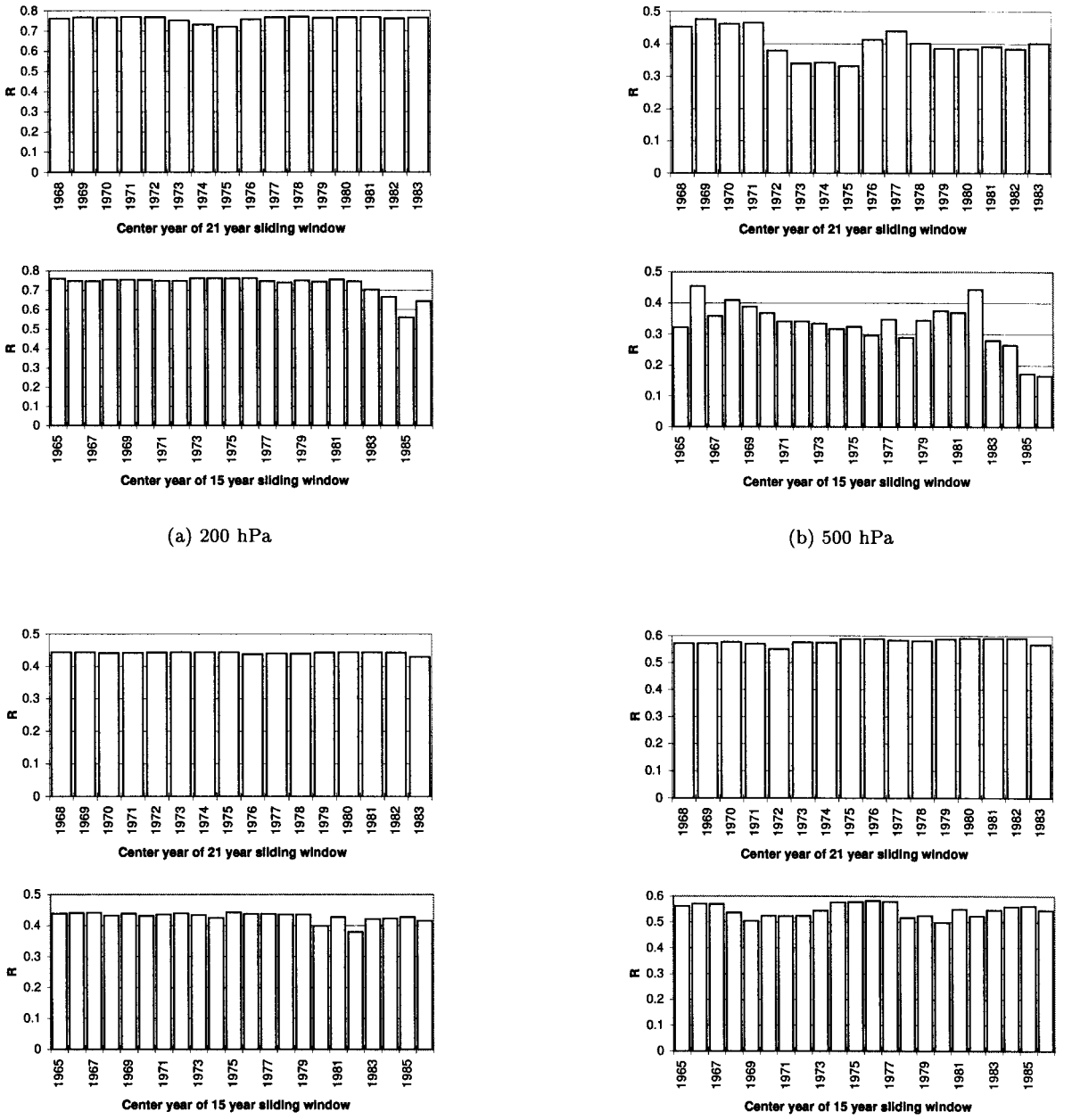
Results of the previous section show that significant linear correlations between May circulation PCs and AISMR were present over the period 1958–1993. No attempt was made to determine whether the circulation–AISMR relationships at any of the lead times were non-linear. In addition, rigorous assessments of the predictive usefulness of the PC predictor sets were not made. While high correlations between various predictors and AISMR have been noted on the data used in the development of forecast models for AISMR, similar results on independent test data have often been lacking (Parthasarathy *et al.*, 1988). Accurate assessments of predictive skill require that predictions be made on data not used in model development. In this study, the bagging procedure outlined in the 'Ensemble methods' section was used to assess the predictive skills of both the linear multiple regression and non-linear neural network models. Unlike previous studies (for example, Parthasarathy *et al.*, 1988; Hastenrath and Greischar, 1993; Navone and Ceccatto, 1994; Singh *et al.*, 1995), estimates of predictive skill were made for a large number of randomly selected independent test data sets.

For each of the 15 circulation PC predictor sets at each lead time, ensembles consisting of 100 members were generated using bagging. Neural network ensemble members were initialized with randomly assigned weight and bias values. For each ensemble member, half of the out-of-bag cases were used to form an independent validation set for the one-step weight decay procedure; the other half were used to test model performance. It should be noted that, as in Tangang *et al.* (1998), the PC analyses were completed using the entire data set. As such, the test data were not, in a strict sense, independent of the data used to estimate the model parameters. To determine the effects of this lack of independence, predictions were made on five test sets consisting of single years taken from 1980 to 1984. PCs based only on data prior to each test year and PCs based on the entire data set were both used to predict AISMR for the given year. In each case network training was started using the same initial weights and biases and was performed using PC scores preceding the test year. Using PCs based on the entire data set did not increase predictive skill.

Table II. Multiple correlation coefficient R between circulation PC scores and AISMR (1958–1993)

Included level(s)	PCs	December	January	February	March	April	May
SLP	3	0.40 (1.91)	0.29 (0.96)	0.19 (0.40)	0.16 (0.28)	0.33 (1.32)	0.59 (5.77)*
SLP, 850 hPa	3	0.34 (1.33)	0.29 (1.01)	0.16 (0.27)	0.12 (0.16)	0.36 (1.61)	0.55 (4.54)**
SLP, 500 hPa	6	0.49 (1.47)	0.43 (1.11)	0.42 (1.01)	0.28 (0.42)	0.42 (1.04)	0.75 (6.34)*
SLP, 200 hPa	6	0.53 (1.78)	0.49 (1.48)	0.41 (0.96)	0.40 (0.92)	0.46 (1.31)	0.79 (7.76)*
SLP, 850, 500 hPa	6	0.45 (1.20)	0.43 (1.08)	0.43 (1.11)	0.28 (0.41)	0.49 (1.49)	0.75 (6.21)*
SLP, 850, 200 hPa	6	0.50 (1.51)	0.48 (1.41)	0.39 (0.87)	0.40 (0.90)	0.52 (1.82)	0.79 (8.05)*
SLP, 500, 200 hPa	6	0.48 (1.36)	0.46 (1.29)	0.41 (0.97)	0.39 (0.86)	0.45 (1.24)	0.78 (7.54)*
SLP, 850, 500, 200 hPa	6	0.46 (1.27)	0.46 (1.28)	0.40 (0.93)	0.39 (0.87)	0.49 (1.50)	0.78 (7.50)*
850 hPa	2	0.28 (1.38)	0.33 (1.98)	0.23 (0.95)	0.11 (0.19)	0.29 (1.50)	0.44 (4.05)***
850, 500 hPa	5	0.35 (0.83)	0.41 (1.19)	0.42 (1.26)	0.28 (0.52)	0.46 (1.60)	0.74 (7.05)*
850, 200 hPa	5	0.30 (0.56)	0.45 (1.51)	0.38 (1.00)	0.40 (1.11)	0.47 (1.74)	0.79 (9.72)*
850, 500, 200 hPa	6	0.41 (0.92)	0.45 (1.20)	0.38 (0.83)	0.44 (1.18)	0.48 (1.45)	0.78 (7.51)*
500 hPa	4	0.27 (0.59)	0.37 (1.21)	0.40 (1.44)	0.24 (0.47)	0.24 (0.47)	0.48 (2.31)
500, 200 hPa	5	0.27 (0.47)	0.42 (1.26)	0.38 (1.02)	0.43 (1.34)	0.43 (1.38)	0.77 (8.88)*
200 hPa	4	0.31 (0.82)	0.39 (1.40)	0.37 (1.19)	0.36 (1.17)	0.34 (1.00)	0.77 (11.27)*

* Significant at 0.5% level.
** Significant at 1% level.
*** Significant at 5% level.
Values in parentheses are the F-test statistics for the significance of R^2.



(a) 200 hPa (b) 500 hPa

(c) 850 hPa (d) SLP

Figure 2. Multiple correlation coefficients between May PC scores (200 hPa, 500 hPa, 850 hPa, and SLP) and AISMR based on 15- and 21-year sliding windows (1958–1993)

To compare performance of the non-linear neural networks with linear methods, a standard linear multiple regression was run on each ensemble member at each lead time. As suggested by LeBaron and Weigend (1998), the linear model parameters were estimated using the union of the training and validation sets. For each trained ensemble member, the multiple correlation coefficient R, the index of agreement d (Willmott *et al.*, 1985), and the root mean squared error ($RMSE$) were computed between the test set estimates of AISMR and the observed values. For comparison, values of d and $RMSE$ were

also computed using the climatological mean AISMR value over the validation and test sets as a predictor.

As noted in the previous section, May 200 hPa PCs showed the strongest relationship with AISMR. Results of neural network, multiple regression, and climatological test set predictions at this lead time are given in Table III. No improvements over the 200 hPa results were noted with the addition of data from more than one atmospheric level. To help compare results of the different model types, histograms of model performance statistics for May 200 hPa PCs are shown in Figure 3. At this lead time, both neural network and multiple regression models performed better than climatology. Little difference between the neural network and multiple regression models is evident. However, the dependence of results on the particular cases included in the training, validation, and test sets is apparent in the wide spread of histograms for each statistic.

As a portion of the data available for model development was used to stop training of the neural networks and was not actively used in the adjustment of weights and biases, test results for the individual ensemble members might have favoured the multiple regression model. This may account for the differences in the performance statistics shown in Table III. For comparison, plots of ensemble-averaged results for the neural network and multiple regression models are given in Figure 4. As the test set used for each ensemble member was drawn randomly from all available AISMR values, average results for each year were derived from models using different training data sets. In terms of yearly predictions at this lead time, little difference between the models was evident. Values of R, d, and $RMSE$ were almost identical for the neural network (0.70, 0.80, and 63.0 mm) and linear multiple regression models (0.71, 0.83, and 62.8 mm). Plots of best fit lines using least-squares regression and locally-weighted regression (Figure 5) indicate that relationships between each of the PCs and AISMR are generally linear at this lead time.

Values of R and $RMSE$ (approximately 0.7 and 63 mm, respectively) for May 200 hPa ensemble results on the independent test data (1958–1993) equal or exceed reported results by Parthasarathy *et al.* (1988), Hastenrath and Greischar (1993) and Singh *et al.* (1995), each of whom used optimized sets of regional and global predictors on test sets drawn from approximately the same time period used in this study. Given the specific differences in training and test sets, as well as in methodology (for example, the use of randomly resampled test sets to estimate predictive skill in the current study), comparisons between the various studies are not strictly valid. Hastenrath (1987) compared cross-validated estimates of the predictive skill of AISMR forecast models with estimates based on fixed data periods and showed that use of training data from after the test year does not significantly inflate AISMR forecast skill. One would expect the same to apply to the resampling procedure used in this study. Consequently, it is likely that May 200 hPa PCs form an important set of predictors for AISMR.

A general decline in model skill was evident at all lead times prior to May. Figure 6 shows the climatological forecast $RMSE$ from 1958 to 1993 and the minimum $RMSE$ values from ensemble forecasts at lead times from December to May. Excluding the May predictor sets, only 850 hPa PCs in January showed any improvement over climatology. Interestingly, this maximum in predictive skill was only captured by the neural network model. Inspection of histograms of differences in $RMSE$ between neural network, multiple regression, and climatological predictions (Figure 7) indicates that the neural network model performed better than either of the other models for this predictor set at this lead time, although the improvement over climatology was small. Values of R, d, and $RMSE$ for the ensemble averaged results (0.46, 0.57, and 78.4 mm, respectively) are such that further investigation into the use of this predictor set, especially considering the early lead time, is recommended.

5. PRE-MONSOON CIRCULATION RELATIONSHIPS

The 'Results' section shows that 200 hPa geopotential height anomalies in May were important in forecasting AISMR over the period 1958–1993. In this section the nature of the circulation–AISMR relationships at this level and lead time are examined.

Table III. Model performance statistics for the individual neural network, multiple regression, and climatological forecasts based on the May PC sets (1958–1993)

Included level(s)	Model type	*R*	*d*	*RMSE* (mm)
SLP	Neural network	0.38 (−0.38, 0.9)	0.53 (0.19, 0.81)	82.9 (47.7, 123.1)
	Multiple regression	0.4 (−0.29, 0.88)	0.59 (0.22, 0.85)	79.1 (42.9, 115.0)
	Climatology	n/a	0.29 (0.05, 0.47)	87.3 (55.1, 117.9)
SLP, 850 hPa	Neural network	0.37 (−0.42, 0.9)	0.52 (0.09, 0.82)	81.5 (45.0, 117.4)
	Multiple regression	0.36 (−0.43, 0.87)	0.55 (0.12, 0.83)	82.1 (40.4, 125.5)
	Climatology	n/a	0.31 (0.04, 0.45)	87.7 (58.1, 118.1)
SLP, 500 hPa	Neural network	0.53 (−0.29, 0.91)	0.66 (0.26, 0.91)	75.0 (35.6, 107.9)
	Multiple regression	0.6 (−0.12, 0.96)	0.72 (0.39, 0.93)	72.0 (40.6, 107.8)
	Climatology	n/a	0.3 (0.06, 0.46)	90.9 (66.5, 115.5)
SLP, 200 hPa	Neural network	0.64 (0.05, 0.94)	0.71 (0.42, 0.92)	69.9 (36.9, 104.6)
	Multiple regression	0.66 (0.09, 0.94)	0.75 (0.43, 0.94)	68.9 (32.6, 96.5)
	Climatology	n/a	0.3 (0.07, 0.47)	88.5 (55.6, 118.1)
SLP, 850, 500 hPa	Neural network	0.55 (−0.16, 0.95)	0.67 (0.26, 0.91)	72.5 (35.0, 102.5)
	Multiple regression	0.54 (−0.08, 0.93)	0.69 (0.31, 0.92)	73.8 (36.9, 108.2)
	Climatology	n/a	0.3 (0.03, 0.46)	87.9 (54.7, 123.5)
SLP, 850, 200 hPa	Neural network	0.59 (−0.09, 0.9)	0.69 (0.32, 0.92)	71.1 (39.2, 99.0)
	Multiple regression	0.65 (0.08, 0.92)	0.75 (0.47, 0.93)	65.7 (33.2, 96.7)
	Climatology	n/a	0.33 (0.1, 0.46)	90.8 (56.5, 117.0)
SLP, 500, 200 hPa	Neural network	0.57 (−0.07, 0.92)	0.68 (0.38, 0.93)	73.4 (36.5, 111.4)
	Multiple regression	0.6 (0.05, 0.91)	0.73 (0.42, 0.93)	69.0 (40.4, 101.9)
	Climatology	n/a	0.3 (0.01, 0.47)	88.9 (57.6, 115.6)
SLP, 850, 500, 200 hPa	Neural network	0.62 (−0.04, 0.91)	0.7 (0.37, 0.93)	71.7 (37.8, 102.2)
	Multiple regression	0.64 (−0.07, 0.94)	0.74 (0.34, 0.94)	68.0 (35.8, 100.2)
	Climatology	n/a	0.31 (0.06, 0.46)	90.6 (60.4, 124.1)
850 hPa	Neural network	0.26 (−0.5, 0.77)	0.41 (0.08, 0.62)	89.3 (57.1, 120.2)
	Multiple regression	0.29 (−0.48, 0.86)	0.48 (0.15, 0.72)	88.5 (57.0, 117.8)
	Climatology	n/a	0.31 (0.03, 0.46)	89.7 (61.4, 114.9)
850, 500 hPa	Neural network	0.63 (0.16, 0.9)	0.7 (0.44, 0.91)	72.3 (45.4, 100.7)
	Multiple regression	0.66 (0.1, 0.92)	0.74 (0.4, 0.92)	70.5 (42.1, 107.9)
	Climatology	n/a	0.29 (0.05, 0.47)	89.9 (58.9, 117.1)
850, 200 hPa	Neural network	0.59 (−0.15, 0.94)	0.69 (0.34, 0.92)	72.3 (41.2, 98.9)
	Multiple regression	0.64 (−0.1, 0.93)	0.74 (0.29, 0.93)	67.7 (37.8, 96.5)
	Climatology	n/a	0.31 (0.03, 0.46)	91.9 (58.6, 120.3)
850, 500, 200 hPa	Neural network	0.61 (0.13, 0.93)	0.69 (0.35, 0.93)	73.7 (39.1, 109.4)
	Multiple regression	0.63 (−0.12, 0.95)	0.74 (0.36, 0.95)	71.4 (36.0, 104.0)
	Climatology	n/a	0.32 (0.04, 0.45)	92.5 (59.3, 120.8)
500 hPa	Neural network	0.21 (−0.57, 0.74)	0.44 (0.14, 0.67)	89.7 (55.3, 117.1)
	Multiple regression	0.27 (−0.55, 0.81)	0.49 (0.12, 0.76)	89.4 (52.6, 124.2)
	Climatology	n/a	0.3 (0.06, 0.45)	87.7 (59.2, 116.7)
500, 200 hPa	Neural network	0.63 (0.08, 0.92)	0.71 (0.39, 0.92)	69.9 (40.8, 100.2)
	Multiple regression	0.67 (0.17, 0.93)	0.76 (0.48, 0.93)	66.6 (35.7, 95.9)
	Climatology	n/a	0.28 (0.03, 0.45)	89.9 (59.3, 116.7)
200 hPa	Neural network	0.67 (0.19, 0.94)	0.73 (0.43, 0.91)	65.1 (38.9, 96.5)
	Multiple regression	0.7 (0.26, 0.94)	0.78 (0.49, 0.94)	61.8 (33.5, 87.7)
	Climatology	n/a	0.3 (0.05, 0.46)	89.5 (59.7, 119.0)

The climatological mean value used as the predictor for each ensemble member was taken over the training and validation set data. Each statistic was calculated using only the independent test data for each ensemble member. The mean value over the ensemble members is given for each statistic; numbers in parentheses correspond to the values of the lower and upper 5% of each ensemble.

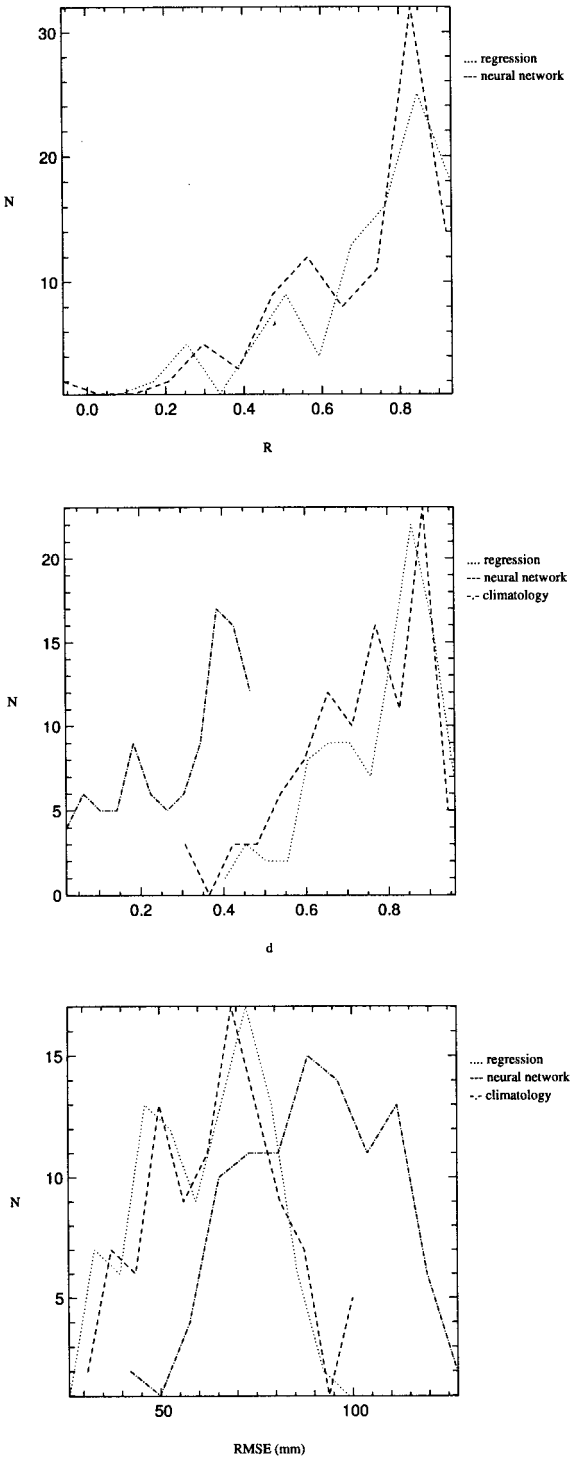


Figure 3. Empirical distributions of the model performance statistics for the May 200 hPa ensemble (100 members)

Given that the neural network and linear multiple regression models performed almost identically and that the predictor–predictand relationships do not appear non-linear, an analysis of only the linear relationships between AISMR and the May 200 hPa PCs was considered appropriate. Table IV shows the linear correlations between AISMR and each of the retained 200 hPa PC scores in May. It should be noted that, as PCs were extracted from all pre-monsoon (DJFMAM) months, small correlations were present between each of the May PC scores. Of the four retained PCs, scores of PC1, PC3, and PC4 in May were highly correlated ($|r| > 0.4$) with summer monsoon rainfall. The correlation between PC2 and AISMR was negligible. Each PC loading is graphically depicted in Figure 8. Contour values show the magnitude of linear correlations between PC scores and standardized anomaly series at each grid point. As the PC solutions were not subjected to rotation, interpretation of loading patterns as dominant modes of regional variability was not attempted (Richman, 1986); even so, interpretation of the unrotated solutions in terms of the pre-monsoon circulation features captured by the statistical models was still possible.

Inspection of the PC loadings shown in Figure 8 and the correlations in Table IV show that the two most important pre-monsoon circulation characteristics are (1) the magnitude and sign of the 200 hPa geopotential height anomaly field over the entire Indian region, and (2) contrasts between 200 hPa height anomalies in the eastern and western portions of the domain. Specifically, May months with positive

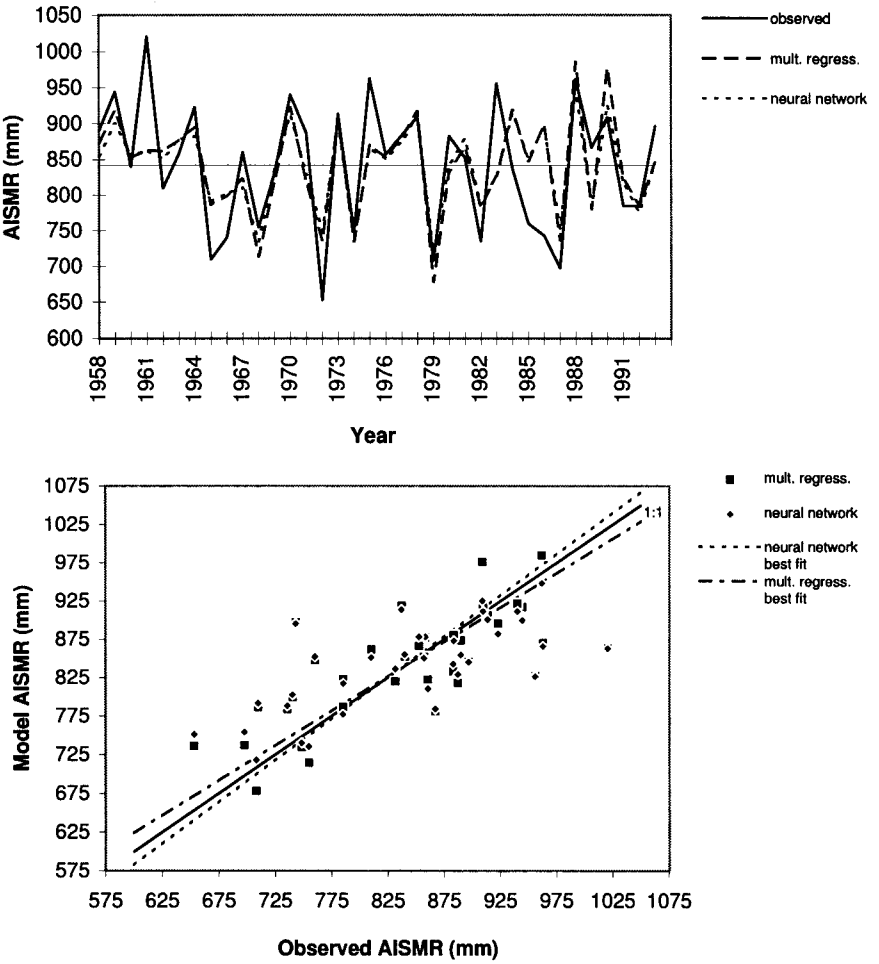


Figure 4. Ensemble averaged neural network and linear multiple regression predictions (from the May 200 hPa PC test set data) and observed AISMR values (1958–1993)

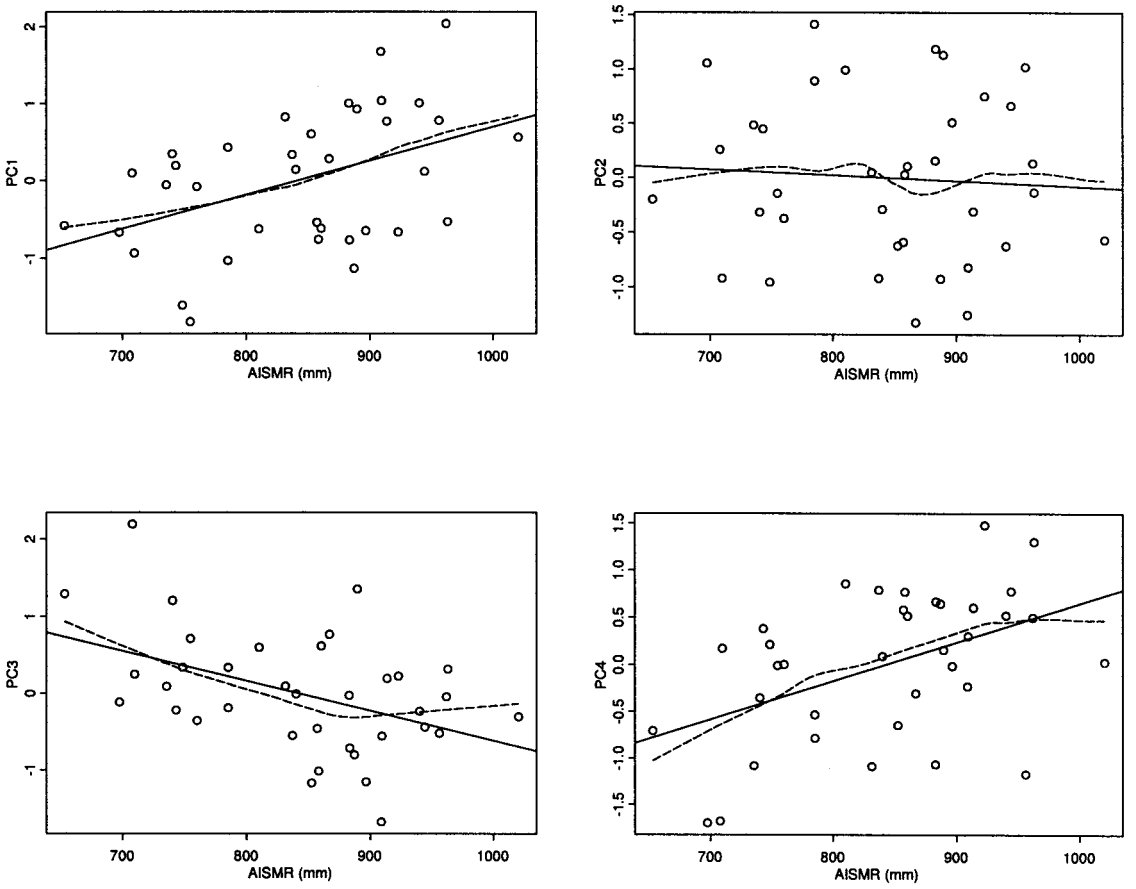


Figure 5. Scatterplots between May 200 hPa PC scores and AISMR with best-fit least-squares (solid) and locally-weighted regression (dashed) lines

(negative) height anomalies over the entire domain were associated with positive (negative) AISMR anomalies (PC1). After removing the first two dominant sources of variability (PC1 and PC2), May months with positive (negative) height anomalies centred over northeastern India tended to be associated with negative (positive) AISMR anomalies (PC3). Finally, after removing the first three dominant sources of variation (PC1, PC2, and PC3), May months with positive (negative) height anomalies centred over northwestern India were generally associated with positive (negative) AISMR anomalies (PC4).

Composite maps of 200 hPa geopotential height anomalies in May for excess and deficient AISMR years, as defined in the 'AISMR' section, are given in Figure 9. Observations during these extreme monsoon years confirm the two relationships outlined above. Excess AISMR years tended to have positive height anomalies over the entire region, with a maximum centred over the northwestern corner of the domain. Conversely, deficient years were associated with negative height anomalies over much of the region, with positive values over the northeastern corner of the domain. The negative centre was positioned in the northwest, shifted slightly east of the position of the positive anomaly centre present during the excess years.

The two monsoon precursors identified here appear to be related to (1) nature of the winter–summer transition in upper level circulation regimes, from westerly flow dominated by the subtropical westerly jet stream to easterly flow dominated by the tropical easterly jet stream; and (2) changes in strength and position of the upper tropospheric Tibetan anticyclone. Both of these features have been identified as important controls on the summer monsoon (Pant and Kumar, 1997). Early onset or increased strength of the tropical easterly jet stream could provide dynamic forcing capable of enhancing summer rainfall

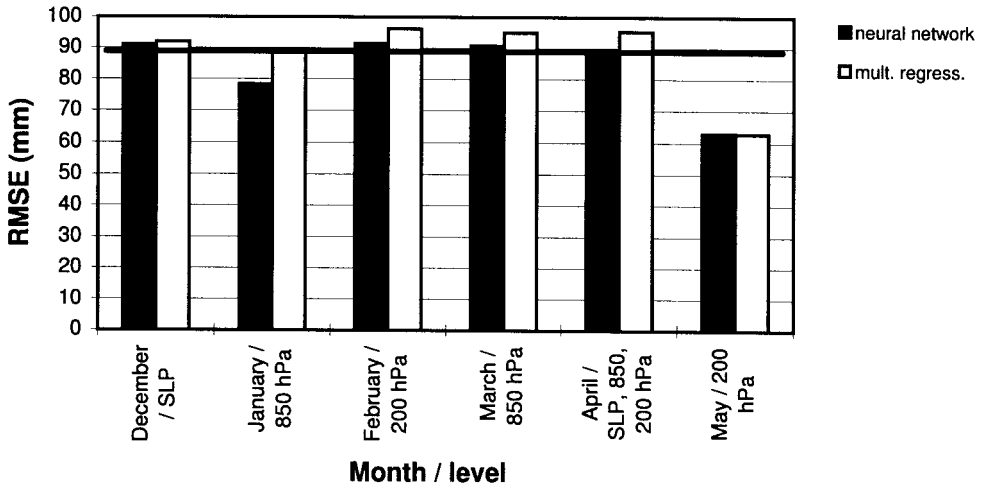


Figure 6. Minimum *RMSE* values from the neural network and multiple regression ensemble predictions at each lead time. The thick horizontal line shows the *RMSE* obtained by the climatological forecast over the study period (1958–1993)

over the Indian region (Webster, 1987). Easterly shifts in position of the Tibetan anticyclone have been linked observationally to summer periods of weakened monsoon conditions (Pant and Kumar, 1997).

In a recent observational and GCM study, Yang *et al.* (1996) considered the response of the Indian summer monsoon to spring geopotential height, wind, and temperature anomalies at the 850 hPa and 200 hPa levels. They found evidence of precursor relationships similar to those shown here and consistent with the upper level circulation features identified above. Strong monsoon years were associated with positive 200 hPa height anomalies over the entire Indian region, while weak monsoon years were associated with negative anomalies. Differences in positioning of the anomaly centres between strong (northwest) and weak (northeast) years were also noted. Easterly wind anomalies in spring were associated with strong monsoon years, while enhanced upper level westerlies were associated with weak monsoon years. Prior to strong monsoon years an anomalous anticyclonic circulation was present to the north of India, while an anomlous cyclonic circulation existed before weak monsoon years.

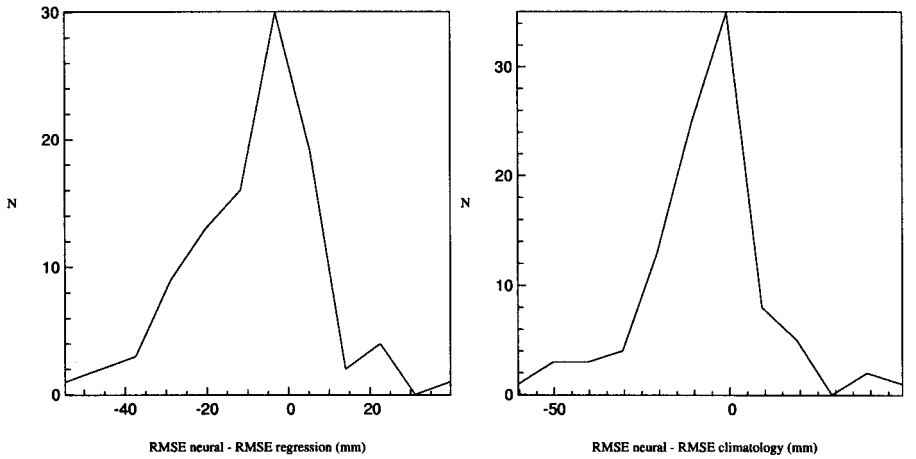


Figure 7. Empirical distributions of the difference in *RMSE* between the three model types (neural network, linear multiple regression, and climatology) for the January 850 hPa ensemble (100 members)

Table IV. Correlation matrix of the 200 hPa PC scores in May and AISMR

	AISMR	PC1	PC2	PC3	PC4
AISMR	+1.00				
PC1	+0.45	+1.00			
PC2	−0.06	−0.12	+1.00		
PC3	−0.44	−0.13	+0.16	+1.00	
PC4	+0.47	−0.18	−0.17	−0.22	+1.00
% var.		62.2	18.5	8.4	7.0

6. DISCUSSION AND CONCLUSION

Predictive skills of two statistical models for AISMR were evaluated using data from 1958 to 1993. A bootstrap-based resampling procedure was used to compare AISMR predictions made by neural networks, linear multiple regression, and climatological models using pre-monsoon (DJFMAM) circulation PCs as predictors. The 200 hPa geopotential height PC set in May was found to yield forecasts that were better than climatology, and that were comparable to or better than prior forecasts made using optimized sets of regional and global predictors (Parthasarathy *et al.*, 1988; Hastenrath and Greischar, 1993; Singh *et al.*, 1995).

Coherent and interpretable precursor signals related to the magnitude and sign of May 200 hPa geopotential height anomalies over the entire region, and of contrasts between 200 hPa height anomalies in the eastern and western portions of the domain, were found to be important in forecasting AISMR. May months with positive 200 hPa geopotential height anomalies over the Indian region were associated with excess AISMR years, while negative 200 hPa geopotential height anomalies were associated with deficient AISMR years. After removing this source of variability, positive anomalies in the northwestern portion of the domain and negative anomalies in the northeastern portion of the domain were associated with excess AISMR years. In deficient AISMR years the opposite circulation situation tended to be present. These circulation precursor signals are similar to those identified in Yang *et al.* (1996) and appear to be related to the winter–summer transition between upper level westerly and easterly flow regimes over the Indian region and the strength and position of the Tibetan anticyclone.

Predictors based on PCs from other atmospheric levels during May were also found to be significant, but yielded lower skill scores than the 200 hPa PC set. The use of data from multiple atmospheric levels at this lead time did not increase forecast skill. A possible explanation for the lack of improvement in predictive skill observed when using data from multiple atmospheric levels was put forward by Yang *et al.* (1996). In their study, structure of the precursor signals to both strong and weak summer monsoon circulations was found to be essentially barotropic. In addition, monsoon precursor signals were found to be strongest in the upper troposphere and weaker in the lower troposphere. This corresponds well with the general decline in predictive skill noted in this study for PC predictors at lower levels of the troposphere.

Comparisons between neural network and multiple regression ensemble runs showed that significant differences between the two classes of model were not apparent when using May predictors. For the May 200 hPa PCs, relationships with AISMR appear to be linear. For months prior to May, model performance exceeded climatology only when using 850 hPa PCs in January as predictors. While weak, this relationship was only identified with the neural network model. The general lack of predictive skill at lead times longer than 1 month found in this study is somewhat surprising given recent work by Webster and Yang (1992) and Yang *et al.* (1996). Coherent precursor signals were found to be present more than a season in advance of monsoon onset. However, both studies considered the strength of the monsoon circulation in slightly different ways. In Yang *et al.* (1996) the monsoon index was based on strength of summer 200 hPa zonal winds, while Webster and Yang (1992) used a measure of zonal wind shear between 850 hPa and 200 hPa. The use of AISMR as the monsoon index here prevented closer comparisons between the three studies.

Further investigation into long-range forecasting of AISMR using May 200 hPa PCs and January 850 hPa PCs combined with predictors from outside the Indian region is required. Given that the neural network ensemble model always performed at least as well as the linear multiple regression model, and that the relation between the January 850 hPa PCs and AISMR was only captured by the neural network, future use of this class of model is recommended.

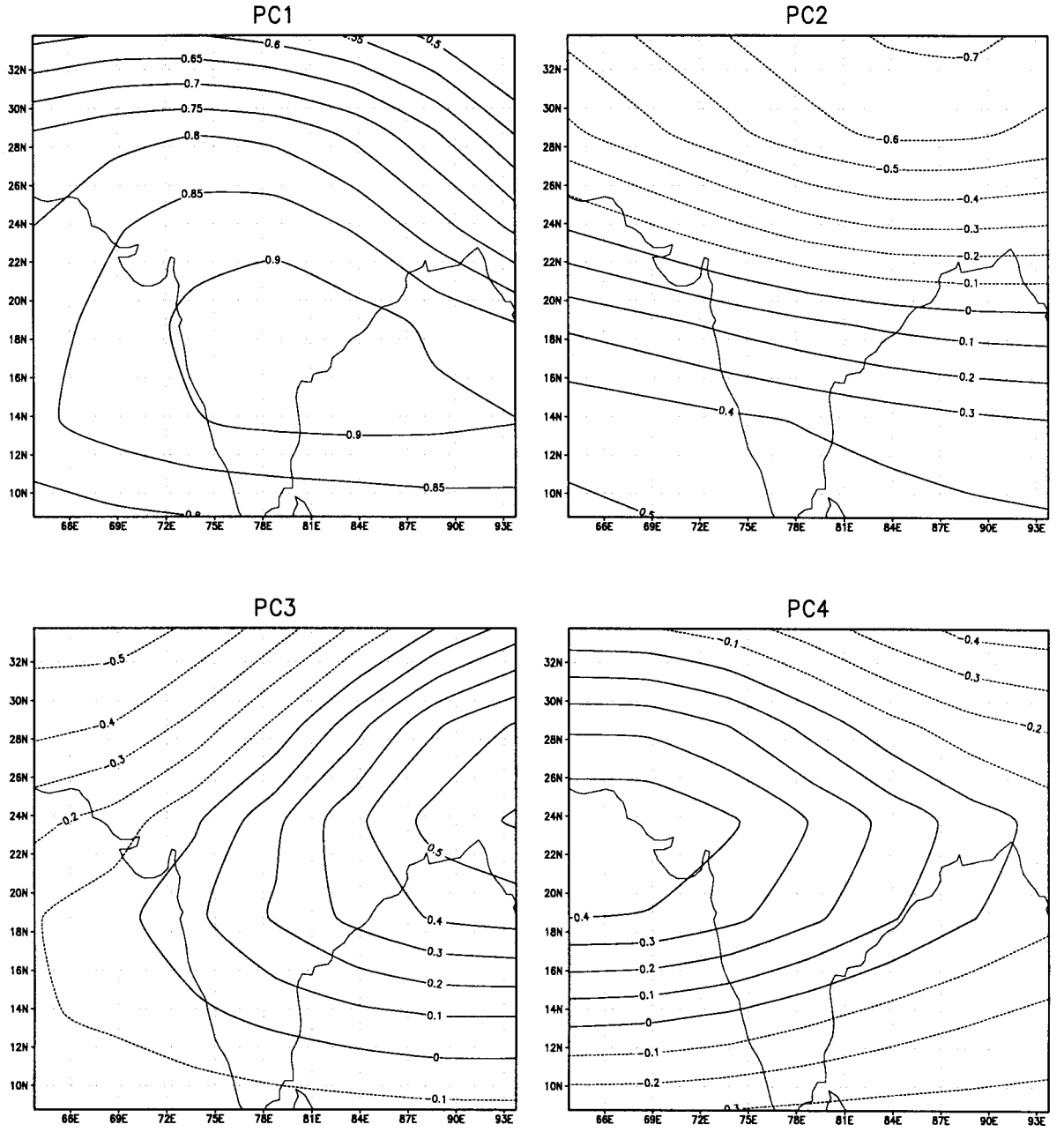


Figure 8. Pre-monsoon (DJFMAM) PC loadings of the 200 hPa geopotential height anomaly field

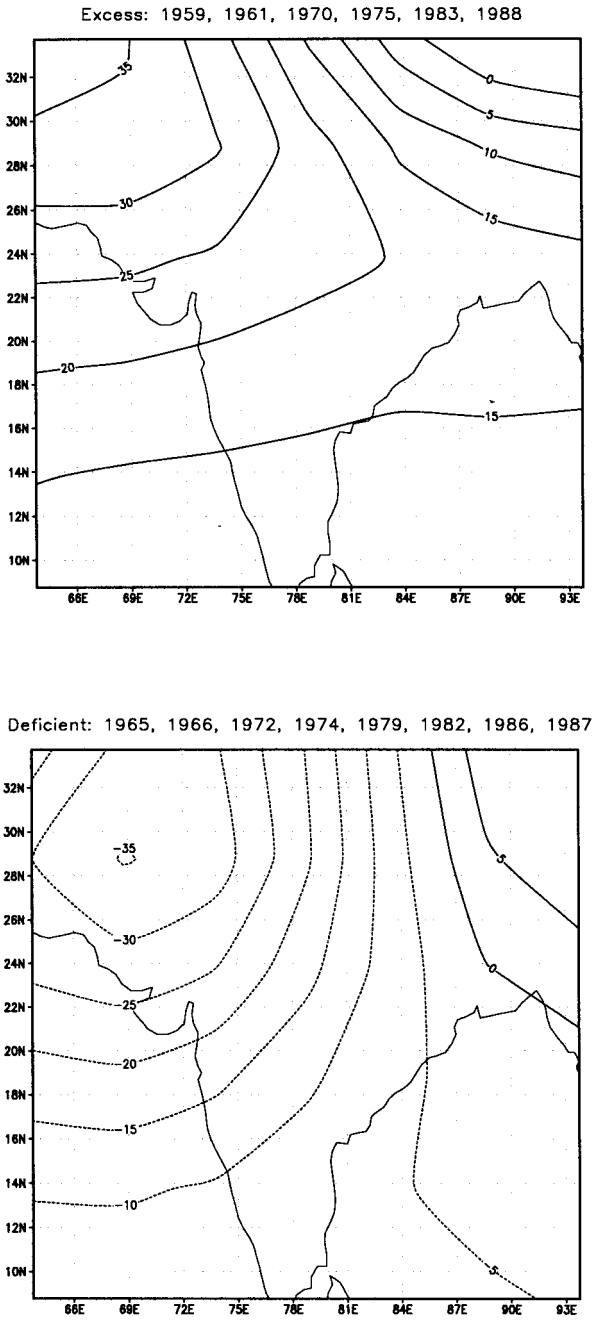


Figure 9. Composite maps of 200 hPa geopotential height anomalies in May for excess and deficient AISMR years

ACKNOWLEDGEMENTS

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Bhalme, H.N., Jadhav, S.K., Mooley, D.A. and Murty, B.V.R. 1986. 'Forecasting of monsoon performance over India', *J. Climatol.*, **6**, 347–354.

Bishop, C.M. 1995. *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, p. 482.

Breiman, L. 1996. 'Bagging predictors', *Mach. Learn.*, **24**, 123–140.

Buell, C.E. 1975. 'The topography of the empirical orthogonal functions', in *Fourth Conference on Probability and Statistics in Atmospheric Sciences*, American Meteorological Society, Boston, pp. 188–193.

Finnoff, W., Hergert, F. and Zimmermann, H.G. 1993. 'Improving model selection by nonconvergent methods', *Neural Netw.*, **6**, 771–783.

Gibb, J. 1996. *Back Propagation Family Album*, Technical Report C/TR96-05, Macquarie University. p. 65.

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

Hastenrath, S. 1987. 'On the prediction of India monsoon rainfall anomalies', *J. Clim. Appl. Meteorol.*, **26**, 847–857.

Hastenrath, S. and Greischar, L. 1993. 'Changing predictability of Indian monsoon rainfall anomalies?', *Proc. Indian Acad. Sci. (Earth Planet. Sci.)*, **102**(1), 35–47.

Hinton, G.E. 1987. 'Learning translation invariant recognition in massively parallel networks', in de Bakker, J.W., Nijman, A.J. and Iteleaven, P.C. (eds.), *Proceedings PARLE Conference on Parallel Architectures and Languages Europe*, Springer-Verlag, Berlin, pp. 1–13.

Hornik, K., Stinchcomb, M. and White, H. 1989. 'Multilayer feedforward neural networks are universal approximators', *Neural Netw.*, **2**, 359–366.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R. and Joseph, D. 1996. 'The NCEP/NCAR 40-year reanalysis project', *Bull. Am. Meteorol. Soc.*, **77**(3), 437–471.

Kidson, J. 1997. 'The utility of surface and upper air data in synoptic climatological specification of surface climatic variables', *Int. J. Climatol.*, **17**, 399–413.

Kumar, K.K., Kumar, K.R. and Pant, G.B. 1997. 'Pre-monsoon maximum and minimum temperatures over India in relation to the summer monsoon rainfall', *Int. J. Climatol.*, **17**, 1115–1127.

Kumar, K.K., Soman, M.K. and Kumar, K.R. 1995. 'Seasonal forecasting of Indian summer monsoon rainfall: a review', *Weather*, **50**(12), 449–467.

Lau, K.M. and Yang, S. 1996. 'Seasonal variation, abrupt transition, and intraseasonal variability associated with the Asian summer monsoon in the GLA GCM', *J. Clim.*, **9**, 965–985.

LeBaron, B. and Weigend, A.S. 1998. 'A bootstrap evaluation of the effect of data splitting on financial time series', *IEEE Trans. Neural Netw.*, **9**, 213–220.

Mooley, D.A., Parthasarathy, B. and Pant, G.B. 1986. 'Relationship between Indian summer monsoon rainfall and location of the ridge at the 500-mb level along 75 degrees E', *J. Clim. Appl. Meteorol.*, **25**, 633–640.

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

Overland, J.E. and Preisendorfer, R.W. 1982. 'A significance test for principal components applied to a cyclone climatology', *Mon. Wea. Rev.*, **110**(1), 1–4.

Pant, G.B. and Kumar, K.R. 1997. *Climates of South Asia*, Wiley, West Sussex, p. 344.

Parthasarathy, B., Diaz, H.F. and Eischeid, J.K. 1988. 'Prediction of all-India summer monsoon rainfall with regional and large-scale parameters', *J. Geophys. Res.*, **93**(D5), 5341–5350.

Parthasarathy, B., Kumar, K.R. and Deshpande, V.R. 1991a. 'Indian summer monsoon rainfall and 200-mbar meridional wind index: application for long-range prediction', *Int. J. Climatol.*, **11**, 165–176.

Parthasarathy, B., Kumar, K.R. and Munot, A.A. 1991b. 'Evidence of secular variations in Indian monsoon rainfall–circulation relationships', *J. Clim.*, **4**, 927–938.

Parthasarathy, B., Munot, A.A. and Kothawale, D.R. 1994. 'All-India monthly and seasonal rainfall series: 1871–1993', *Theor. Appl. Climatol.*, **49**, 217–224.

Parthasarathy, B., Sontakke, N.A., Munot, A.A. and Kothawale, D.R. 1987. 'Droughts/floods in the summer monsoon season over different meteorological subdivisions of India for the period 1871–1984', *J. Climatol.*, **7**, 57–70.

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

Richman, M.B. 1986. 'Rotation of principal components', *J. Climatol.*, **6**, 293–335.

Riedmiller, M. 1994. 'Advanced supervised learning in multi-layer perceptrons—from backpropagation to adaptive learning algorithms', *Comput. Stand. Interfaces*, **16**(3), 265.

Riedmiller, M. and Braun, H. 1993. 'A direct adaptive method for faster backpropagation learning: the Rprop algorithm', in Ruspini, H. (ed.), *Proceedings of the IEEE International Conference on Neural Networks*, San Francisco, pp. 586–591.

Sarle, W.S. 1994. 'Neural networks and statistical models', in Proceedings of the Nineteenth Annual SAS Users Group International Conference, SAS Institute Inc., Cary, pp. 1538–1550.

Singh, D., Bhadram, C.V.V. and Mandal, G.S. 1995. 'New regression model for Indian summer monsoon rainfall', *Meteorol. Atmos. Phys.*, **55**, 77–86.

Tangang, F.T., Tang, B., Monahan, A.H. and Hsieh, W. 1998. 'Forecasting ENSO events: a neural network-extended EOF approach', *J. Clim.*, **11**, 29–41.

Webster, P.J. 1987. 'The variable and interactive monsoon', in Fein, J.S. and Stephens, P.L. (eds.), *Monsoons*, Wiley, New York, pp. 269–330.

Webster, P.J. and Yang, S. 1992. 'Monsoon and ENSO: selectively interactive systems', *Q. J. R. Meteorol. Soc.*, **118**, 877–926.

Willmott, C.J., Ackleson, S.G., Davies, R.E., Feddema, J.J., Klink, K.M., Legates, D.R., O'Donnell, J. and Rowe, C.M. 1985. 'Statistics for the evaluation and comparison of models', *J. Geophys. Res.*, **90**, 8995–9005.

Yang, S., Lau, K.M. and Sankar-Rao, M. 1996. 'Precursory signals associated with the interannual variability of the Asian summer monsoon', *J. Clim.*, **9**, 949–964.

