Indian Journal of Marine Sciences
Vol. 35(3), September 2006, pp. 210-220

Predictability of sea surface temperature anomalies in the Indian Ocean using artificial neural networks

K C Tripathi[1], I M L Das[*1,2] & A K Sahai[3]

[1]K.Banerjee Center of Atmospheric & Ocean Studies, University of Allahabad, Allahabad - 211002, India

[2]Department of Physics, University of Allahabad, Allahabad - 211002, India

[3]Indian Institute of Tropical Meteorology, Homi Bhabha Road, Pune - 411008, India

[*][E-mail: drimldas@yahoo.com]

Received 20 October 2005, revised 26 April 2006

Artificial Neural Networks (ANN) have been used to access the predictability of sea surface temperature (SST) anomalies for the small area of Indian Ocean Region (27° to 35° S and 96° to 104° E). Twelve networks, corresponding to each month, have been trained on the area average SST values. The performance of the networks has been evaluated and found that the models have been able to predict the anomalies with a reasonably good accuracy. The performance of ANN models has been compared with the Linear Multivariate Regression model to justify the use of a nonlinear model. It has been found that whenever the dependence of present anomalies on the past anomalies show a nonlinear relationship, the linear model such as regression models fails to make any forecast. These are the months of June, September, October and November. In such cases the nonlinear ANN models have been proved to be fairly useful and make relatively better forecasts. When the dependence is linear, the performance of the ANN models is similar to the regression models. In such cases, use of ANN models only leads to increase in complexity without significant improvement in the performance.

[**Key words**: Artificial neural networks, error back propagation, sea surface temperature, linear regression, Indian Ocean]

1. Introduction

Artificial Neural Networks (ANN) has a natural propensity for storing experiential knowledge and making it available for use[1]. The ANN is a computer model made up of simple processing units or nodes. The nodes are connected by links or weights. The node connection topologies vary from fully connected to sparsely or locally connected.

ANNs are being increasingly used for nonlinear regression and classification problems in meteorology and oceanography due to their usefulness in data analysis and prediction. Some of the recent examples include study of sea surface temperature (SST) anomalies[2], multiparameter satellite retrievals from the Special Sensor Microwave/Imager[3], an empirical atmosphere in a hybrid coupled atmosphere-ocean model of the tropical Pacific[4] and equatorial and tropical Pacific sea surface temperature prediction[5,6].

The importance of SST to accurate weather forecasting of both severe events and daily weather has been increasingly recognized over the past several years. Also, it significantly affects the monsoon locally as well as globally[7]. Recent studies indicate that the Indian Ocean impacts climate around the globe, including the widespread drought from 1998-2002 in the US, Southern Europe, and parts of Asia[8]. The importance of the Indian Ocean to remote climate variability has renewed interest in the study of SST anomalies of the Indian Ocean. A successful simulation of the SST anomaly variation of Indian Ocean can also contribute to the monsoon forecast of the Indian subcontinent. Recently, a statistical model has been developed[9] to forecast monthly sea surface temperature in Indian Ocean using a linear regression model. Further, It has been shown that Indian Ocean SST can be a good indicator for El Nino and La Nina onsets[10]. However, works on SST prediction of Indian Ocean using ANN are very scarce as compared to tropical and equatorial Pacific[5,11,12].

In the present study the multi-layer feed-forward neural network model, which performs non-linear regression and classification, has been used to predict the SST anomalies over the Indian Ocean region (27° to 35 °S and 96° to 104 °E) (Fig. 1). Among learning algorithms available, we have used the error Backpropagation[13] with delta learning rule. The ANN

[*] Corresponding author
Phone: 91-532-2460974
Fax: 91-532-2460974

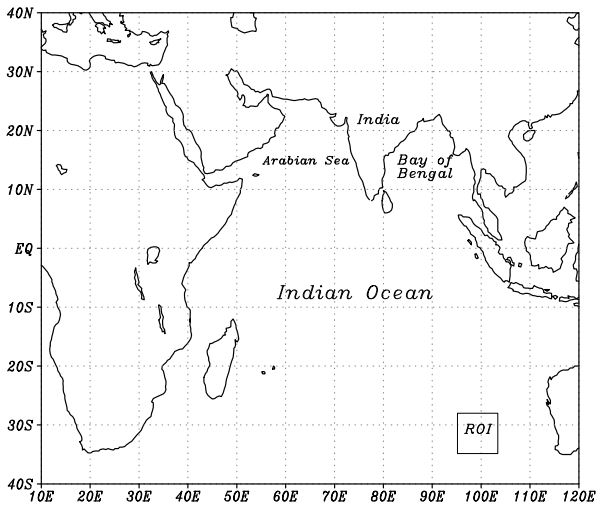

Fig. 1 — Geographical location of the region of interest depicted with rectangular box within the Indian Ocean.

architecture is 'fully connected' with each neuron in a layer connected to each neuron in the next layer although other variations are possible[14] e.g. connecting only the corresponding neurons of two layers. Also, if the data is two-dimensional, e.g. image, the corresponding neurons and the neighborhood neurons of the layers may be connected so that distant neurons do not have an impact on the input to the next layer neurons. The Reynolds' reconstructed SST[15] from 1st January 1950 to 1st December 2001 (52 years' monthly time series) has been used for the analysis. The performance of the nonlinear ANN model has been compared with a linear regression model.

2. Methodology

The method essentially involves time series prediction. The Reynolds' reconstructed SST (1950-2001), which is used in the present analysis consist of 624 (52 × 12) data points. From this series of 624 points we construct 12 series of 52 points each in such a way that each series represent the data of a particular month. Thus we get separate time series for each month, e.g. the first series consists of the SST values for January 1950, January 1951… January 2001. The anomalies of the average SST have been calculated for each month. Each time series of 52 months' anomalies has been divided into three parts namely training, validation and testing[16] (refer to section 2.3). Separate ANN architectures have been designed for the prediction of anomaly series of separate months. Making one network capable of predicting the SST anomalies of all the 12 months would have increased the number of free parameters (weights) manifold. Such a network would require a very large amount of training data. The output from each network is the value of the predicted anomaly for the corresponding month. Hence, all the ANNs have only one neuron in the output layer. The number of neurons in the input layer varies from one (for December) to four (for June) depending upon the number of predictors selected. For determining the predictors for a series of a particular month (hereby called the predictand month), the lag correlation of the series for the month was calculated with the series of all the 12 months, of the grid point belonging to the same area as well as with the grid points belonging to the neighboring area. The lag values have been taken to be 1-year and 2 years. The correlation of the series of the predictand month was also calculated with those months that precede the predictand month with no lag (0-year lag). Only those months have been taken as predictors, which gave a correlation coefficient better than 0.5 (except for the case in which none of the series gave a correlation coefficient better than 0.5) and in no case more than 4 predictors have been taken. Thus, the number of neurons in the input layer is 1 for December; 3 for April, July and August; 4 for June and 2 for the remaining months. It is a matter of fact that the best correlation was obtained with the past data of the same area and the neighboring grid points bore comparatively less correlation.

The number of hidden layers has been kept as 1 for all cases, as a single hidden layer is sufficient to identify arbitrary non-linearity in the data for continuous functions[17, 18]. The number of neurons in the hidden layer is a significant factor that contributes to the overall performance of the network and unfortunately it has always been a mystery[19]. Too few neurons in the hidden layer starve the network of the resources it needs to solve a problem and too many of them tend to learn more than they are designed for i.e. they tend to learn the noise along with the relationship. In the former case the very purpose of using the ANN, i.e. to learn the non linearity in the data is invalidated and in the latter case the performance on the test data is poor, a phenomenon called 'over training', 'over fitting' or 'memorization[19, 20]. In the present study, the number of neurons in the hidden layer has been kept at 2. The training began with two neurons in the hidden layer,

followed by evaluation of performance on the training and validation data, increasing the number of hidden neuron by one and repeating the process[20]. Two neurons gave better performance in all the cases. There was no significant improvement in the performance with 3 neurons. More than 3 neurons were not tried owing to the small data set that makes the problem solving for many parameters ill conditioned[19].

2.1 Activation functions

The logistic activation function[20]

$$f(\alpha) = 1/[1 + \exp(-\alpha)] \quad \ldots(1)$$

with bias has been used as the activation function for the neurons in the hidden layer for predicting the SST anomalies. For the neurons in the output layer the identity function[21]

$$f(\alpha) = \alpha \quad \ldots(2)$$

has been used as the activation function for all months.

2.2 Normalization

The following normalization scheme has been used:

$$X_n = [(X - X_{min})/(X_{max} - X_{min})] * 0.6 + 0.2 \quad \ldots(3)$$

where X is the SST anomaly, X_{max} is the maximum anomaly for the month, X_{min} is the minimum anomaly for the month and X_n is the normalized value of the anomaly. The factors of 0.6 and 0.2 are included so that the normalized values are not 0.0 or 1.0. As X_n approaches these extreme values, the derivative of the sigmoid function becomes 0 and no learning occurs[22].

2.3 The training data set

Since the predictors have been taken to be the lag series of some months and the maximum lag has been 2, the years 1950 and 1951 could not be used for the prediction. Thus, out of the 52 values in these series, only 50 could be used as predictands. To maintain the consistency, only 50 values have been used as the predictands for all the months. From these 50 points, last 5, i.e. years 1997, 1998, 1999, 2000 and 2001, have been used as the 'test' cases for the hindcast. Out of the remaining 45 points, 7 have been randomly selected for 'cross-validation' and the remaining points have been used as 'estimation' set data for the 'training' of the networks. Hence, the data is divided into three parts instead of the more common two parts[20].

2.4 Criteria for stopping the training

The number of cycles used for training is an important factor contributing to the performance of the ANN. If the training is stopped very early, the network will not be able to reach a minimum, i.e. the cost function will not be minimized and too long a training would lead to memorization of training data[19,20]. To avoid this threat, the 'cross-validation[16,20] of the architecture has been done during the training itself. This cross-validation is different from the 'test' phase of the network which simply runs through the data and no change in the parameters (weights) are made no matter what is the error (cost function) of the system. In the present study, the cross-validation was performed after every 10 training cycles. In short, the training algorithm proceeded as follows: the training is stopped after the weights were modified for 10 cycles, performance evaluated on the validation data and the training is resumed. When the performance on the validation data starts deteriorating, the training is stopped preventing the network from memorizing the data. To avoid monotony in the training, the patterns have been shuffled before the beginning of each cycle. The weights have been adjusted after each pattern presentation i.e. the epoch size is kept at 1.

2.5 Multivariate regression model

The performance of the ANN has been compared with the performance of the more commonly used regression model[23]. For k predictors, the value of the model output is given by the multiple regression equation

$$y' = b_0 + b_1 \times x_1 + b_2 \times x_2 + \ldots + b_k \times x_k \quad \ldots(4)$$

where k is the number of predictors, (x_1, x_2, …, x_k) is the set of predictor values, b_0, b_1, … b_k are the regression constants to be determined with the help of the available data set and y' is the calculated value of the predictand. For N given cases of the set of predictor values and the corresponding predictand values, there will be N such equations.

$$\left.\begin{aligned} y_1 &= b_0 + b_1 \times x_{11} + b_2 \times x_{12} + \ldots + b_k \times x_{1k} \\ y_2 &= b_0 + b_1 \times x_{21} + b_2 \times x_{22} + \ldots + b_k \times x_{2k} \\ &\vdots \\ y_N &= b_0 + b_1 \times x_{N1} + b_2 \times x_{N2} + \ldots + b_k \times x_{Nk} \end{aligned}\right\} \quad \ldots(5)$$

where y_i = observed output for the i^{th} case, b = regression constants and x_{ij} = value of the j^{th} predictor for the i^{th} case. It is required to estimate the k+1 parameters with the help of the above N equations. The following matrix equation is solved to obtain the regression parameters[23] for 3 predictors.

$$\begin{pmatrix} N & \sum x_1 & \sum x_2 & \sum x_3 \\ \sum x_1 & \sum x_1^2 & \sum x_1 \bullet x_2 & \sum x_1 \bullet x_3 \\ \sum x_2 & \sum x_2 \bullet x_1 & \sum x_2^2 & \sum x_2 \bullet x_3 \\ \sum x_3 & \sum x_3 \bullet x_1 & \sum x_3 \bullet x_2 & \sum x_3^2 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} \sum y \\ \sum x_1 \bullet y \\ \sum x_2 \bullet y \\ \sum x_3 \bullet y \end{pmatrix} \quad \ldots(6)$$

where $x_1 = (x_{11}, x_{21}, \ldots, x_{N1})$, $x_2 = (x_{12}, x_{22}, \ldots, x_{N2})$, $x_3 = (x_{13}, x_{23}, \ldots, x_{N3})$ and all the summations are done for N values and the operator '•' denotes dot product of two vectors. It can be easily generalized for 4 or more predictors.

For solving Eq. (6), the Guass-Jordan Elimination method[24] has been used. In the regression model, 45 data points were used to estimate the parameters and 5 were used to test the model. Thus, the value of N has been set to 45 and the value of k is varied from 1 (for December) to 4 (for June).

3. Results and Discussion

3.1 Performance of ANN

Figure 2 shows the typical plots for comparison of the ANN model output with the observed data for the months of January (showing both +ve and –ve SST anomalies with correlation coefficient (cc) = 0.71 for training and 0.97 for test data), March (showing both +ve and –ve SST anomalies but cc = 0.78 for training and 0.59 for test data) and August (showing only +ve SST anomalies with cc = 0.84 for training and 0.89 for test data). Similar graphs can be drawn for other months also.

Tables 1A and 1B show various statistics to evaluate the performance of the ANNs for the training and test cases respectively. In the training cases (Table 1A) the variance of the model output is of the same order as that of the observed data. The variance in the test cases (Table 1B) also is of the same order as that of the observed data except for the months January, May and December for which the variances are underestimated. The root mean square error (RMSE) is always less than the standard deviation (SD) of the observed data for the corresponding month both for the training and test cases. This shows the utility of the model in predicting better than the mean value. For the test cases, the RMSE is as small as 38% of the SD of observed data (for February). The worst case is for the month of October where the RMSE is 94% of the corresponding observed SD, but prediction still remains better than the mean prediction. The correlation coefficients (cc) between

Table 1 — ANN Performance measures for training and test cases

Month	Var (ob.)	Var (o/p)	SD (ob.)	SD (o/p)	cc	RMSE	Bias
			(A) Training cases				
Jan	0.19	0.10	0.44	0.32	0.71	0.31	0.01
Feb	0.29	0.12	0.54	0.35	0.66	0.40	0.00
Mar	0.23	0.14	0.48	0.37	0.78	0.30	0.00
Apr	0.16	0.12	0.40	0.34	0.85	0.21	0.00
May	0.16	0.09	0.41	0.31	0.78	0.26	0.00
Jun	0.21	0.15	0.46	0.39	0.85	0.24	0.03
Jul	0.17	0.09	0.41	0.31	0.76	0.27	0.01
Aug	0.13	0.09	0.36	0.30	0.84	0.19	0.02
Sep	0.13	0.08	0.36	0.29	0.83	0.19	0.00
Oct	0.11	0.07	0.34	0.26	0.78	0.22	0.00
Nov	0.16	0.09	0.39	0.30	0.78	0.25	0.01
Dec	0.28	0.08	0.52	0.29	0.55	0.44	0.00
			(B) Test cases				
Jan	0.26	0.04	0.51	0.20	0.97	0.32	-0.08
Feb	0.18	0.25	0.42	0.49	0.95	0.16	-0.14
Mar	0.09	0.05	0.30	0.22	0.59	0.25	0.19
Apr	0.04	0.02	0.20	0.17	0.76	0.13	0.22
May	0.13	0.01	0.36	0.11	0.90	0.26	0.18
Jun	0.06	0.04	0.26	0.22	0.92	0.10	0.33
Jul	0.05	0.01	0.24	0.08	0.77	0.18	0.09
Aug	0.02	0.03	0.14	0.17	0.89	0.08	0.02
Sep	0.03	0.05	0.18	0.21	0.87	0.10	-0.01
Oct	0.10	0.01	0.33	0.03	0.69	0.31	0.06
Nov	0.09	0.06	0.31	0.25	0.76	0.20	0.14
Dec	0.20	0.02	0.45	0.15	0.99	0.31	0.04

var (ob.) = variance in the observed data; var (o/p.) = variance in the output (predicted) data; SD(ob.) = standard deviation of the observed data; SD(o/p) = standard deviation of the output (predicted) data; cc = correlation coefficient between the observed and predicted series; RMSE = root mean square error of prediction; bias = difference in the mean of the observed and predicted series.

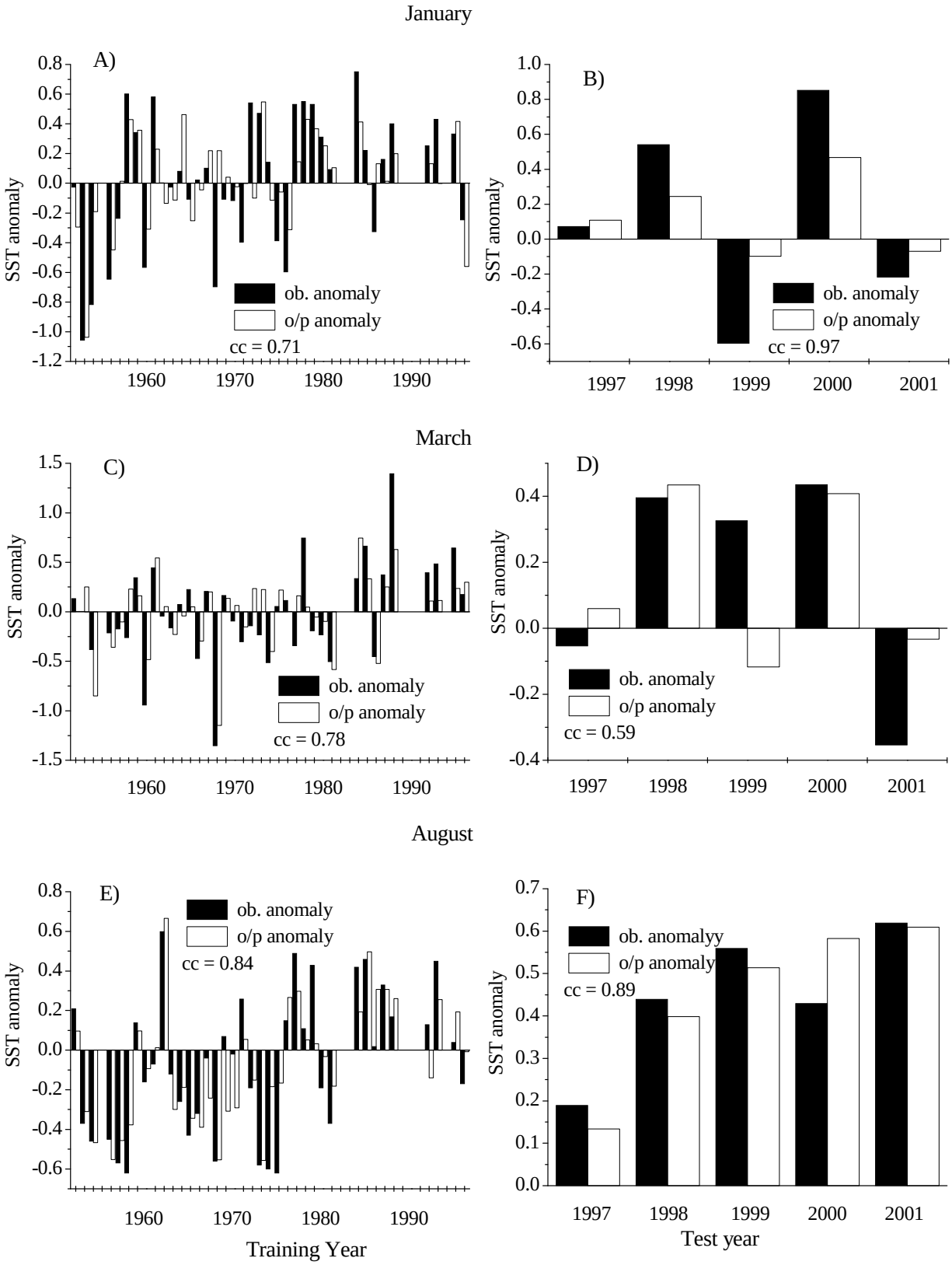

Fig. 2 — Comparison of ANN predicted anomalies with the observed anomalies for training and test cases for January (A, B), March (C, D) and August (E, F).

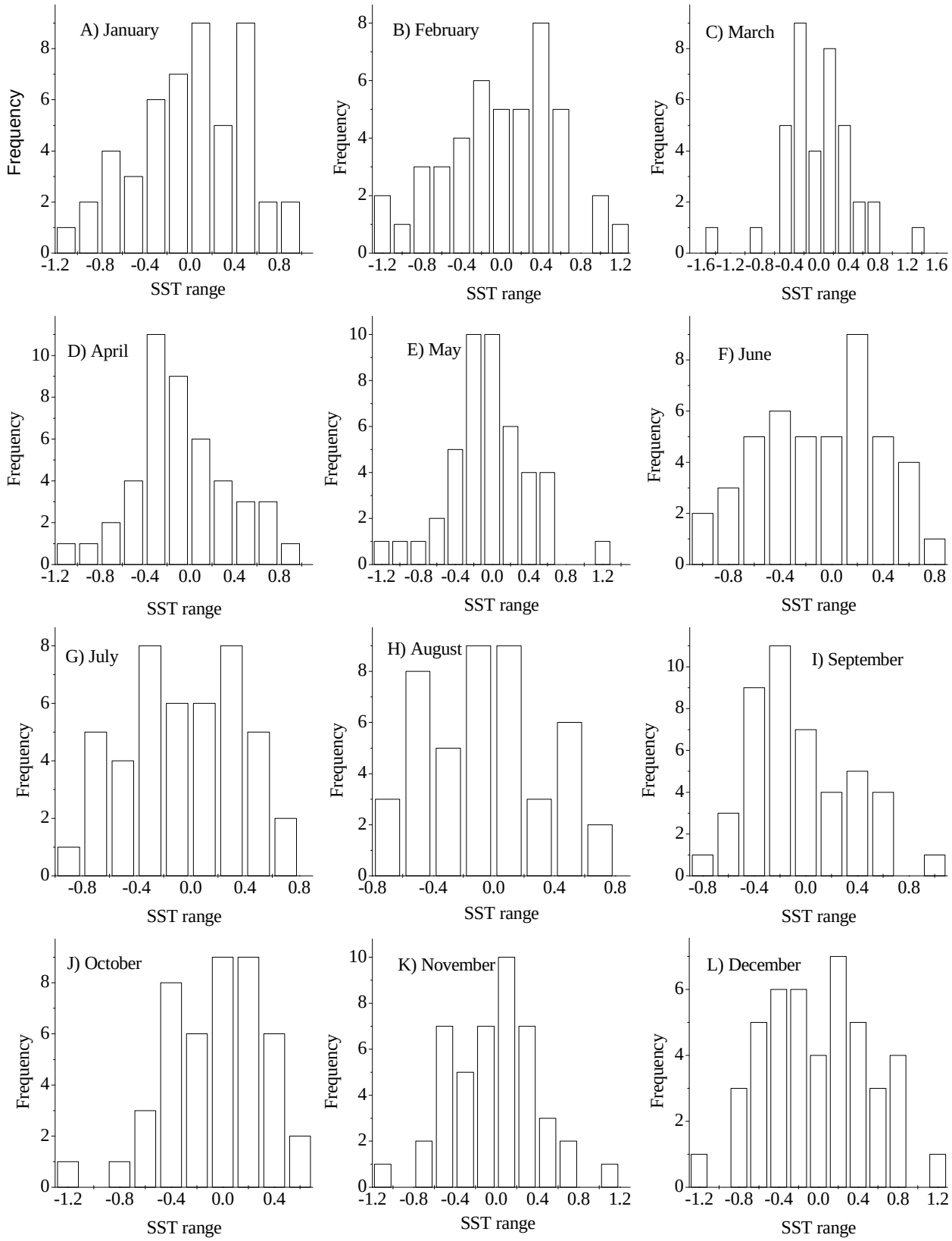

Fig. 3 — Histograms of the training data for all the months of the year.

the observed and the model output anomalies for the test cases show that the ANNs are strongly capable of predicting the anomaly trends in the SST. Five models have shown correlation coefficients > 0.9. Only for March, the correlation coefficient is relatively poor than that of the other months. This can be explained from the training case of March. Figure 3 shows the histograms of the observed data for the training cases with bin size of 0.2. The histogram of March data (Fig. 3C) has more empty bins than that of the other months. Thus, the range of training data and the test data for March is entirely different as compared to other months. The training set anomalies for March (Fig. 2C) go well beyond 1 and –1 where as the test set anomalies (Fig. 2D) are confined to the range –0.4 to 0.4. This is not so for other months. This explains the relatively poor performance of the ANN for the month of March. The ANNs have predicted the anomaly trends correctly in 53 out of 60 test cases, which is almost 90%. Seven models have predicted the trends correctly 5 out 5 times, 3 models have predicted the trends correctly 4 out of 5 times and 2 have predicted correctly 3 out of 5 times. It has been found that 3 models (for the months of March, April and September) have predicted the wrong trends for the year 1997. The reason for these fluctuations in the anomaly may lie in the presence of a dipole in the proximity of the region of interest[25]. The dipole is an anomalous event and hence the trend in the anomaly patterns changes in an unpredictable manner.

3.2 Comparison with Regression model

According to Eqs (5) and (6), 12 linear multiple regression models have been formulated corresponding to 12 months of the year. As usual, the data set for each month has been divided into two parts: training and testing. The first 45 years were used for training (estimation of parameters) and last five years were used for evaluating the performance on unseen data. The predictors used were same as those used for ANN. Figure 4 shows the comparison of the regression model output with the observed data for the months of January (with cc = 0.61 for training and 0.96 for test data), March (with cc = 0.74 for training and 0.67 for test data) and August (with cc = 0.77 for training and 0.86 for test data). Tables 2A and 2B list the performance measures of the regression models for the training and test cases respectively showing the same statistical parameters as used for ANN.

Table 2 — Regression Performance measures for training and test cases

Month	Var (ob.)	Var (o/p)	SD (ob.)	SD (o/p)	cc	RMSE	bias
(A) Training cases							
Jan	0.22	0.08	0.47	0.29	0.61	0.37	0.00
Feb	0.30	0.12	0.55	0.35	0.64	0.42	0.00
Mar	0.23	0.13	0.48	0.36	0.74	0.32	0.00
Apr	0.17	0.10	0.42	0.32	0.77	0.23	0.00
May	0.19	0.12	0.44	0.34	0.77	0.28	0.00
Jun	0.21	0.14	0.46	0.37	0.80	0.28	0.00
July	0.18	0.09	0.42	0.30	0.73	0.29	0.00
Aug	0.15	0.09	0.38	0.29	0.77	0.24	0.00
Sep	0.14	0.09	0.38	0.30	0.79	0.23	0.00
Oct	0.14	0.07	0.38	0.26	0.68	0.27	0.00
Nov	0.17	0.06	0.41	0.24	0.59	0.33	0.00
Dec	0.28	0.09	0.53	0.32	0.59	0.43	0.00
(B) Test cases							
Jan	0.26	0.05	0.51	0.22	0.96	0.31	0.01
Feb	0.18	0.14	0.42	0.38	0.93	0.16	-0.08
Mar	0.09	0.07	0.30	0.28	0.67	0.24	0.17
Apr	0.04	0.02	0.20	0.15	0.89	0.09	0.22
May	0.13	0.01	0.36	0.11	0.90	0.27	0.18
Jun	0.06	0.03	0.26	0.19	0.44	0.25	0.27
July	0.05	0.02	0.24	0.15	0.65	0.18	0.03
Aug	0.02	0.02	0.14	0.15	0.86	0.08	0.09
Sep	0.03	0.02	0.18	0.15	-0.14	0.25	-0.19
Oct	0.10	0.00	0.33	0.07	0.38	0.31	0.11
Nov	0.09	0.02	0.31	0.13	-0.29	0.37	0.04
Dec	0.20	0.06	0.45	0.24	0.98	0.22	0.02

var (ob.) = variance in the observed data; var (o/p.) = variance in the output (predicted) data;SD(ob.) = standard deviation of the observed data; SD(o/p) = standard deviation of the output (predicted) data; cc = correlation coefficient between the observed and predicted series; RMSE = root mean square error of prediction; bias = difference in the mean of the observed and predicted series.

The performance of the regression models is fairly good on the training data as can be seen from the fact that the RMSE values are always less than the standard deviation of the observed data implying that the predictions are better than the mean predictions. The correlation coefficients are also fairly good

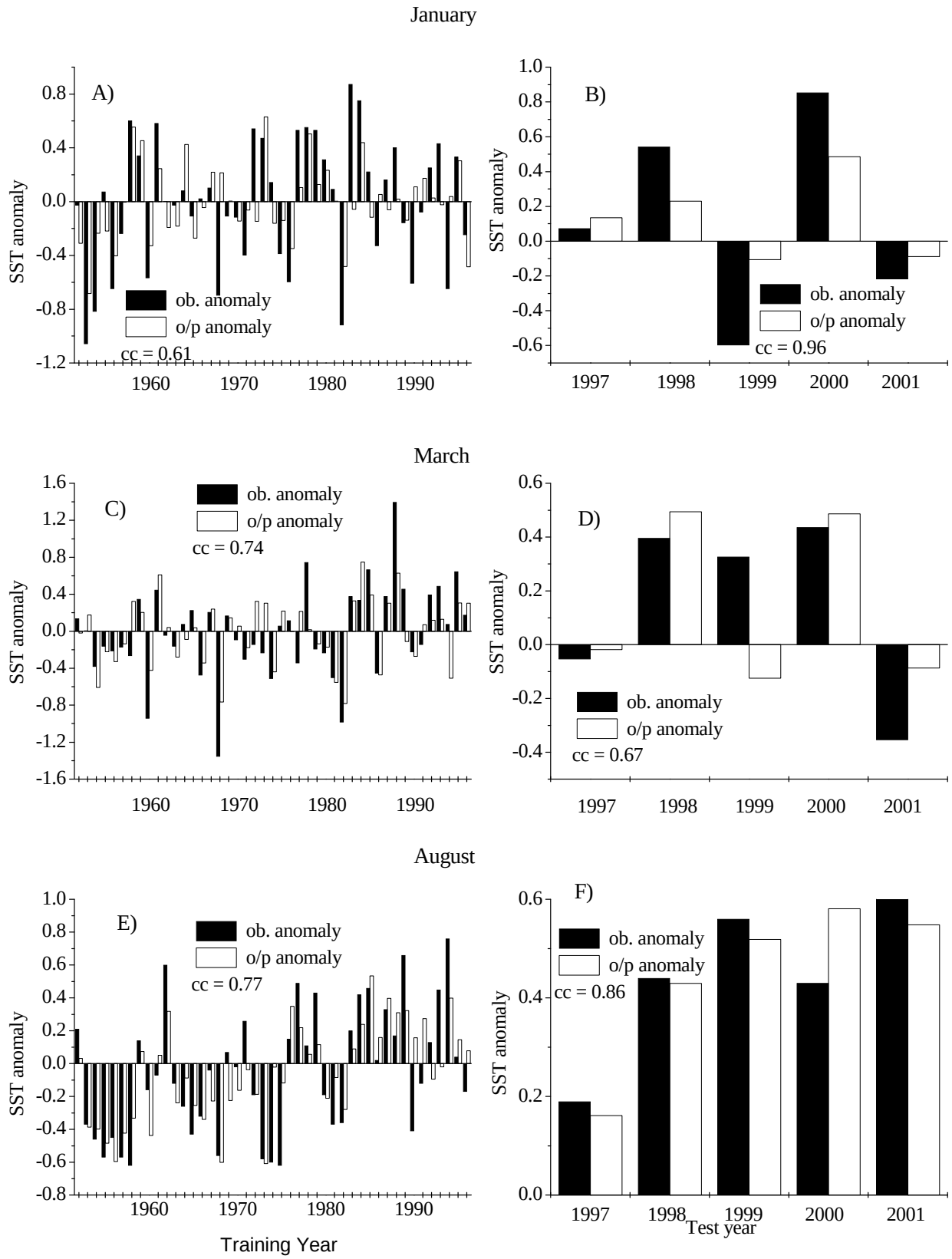

Fig. 4 — Comparison of regression predicted anomalies with the observed anomalies for training and test cases for January (A, B), March (C, D) and August (E, F).

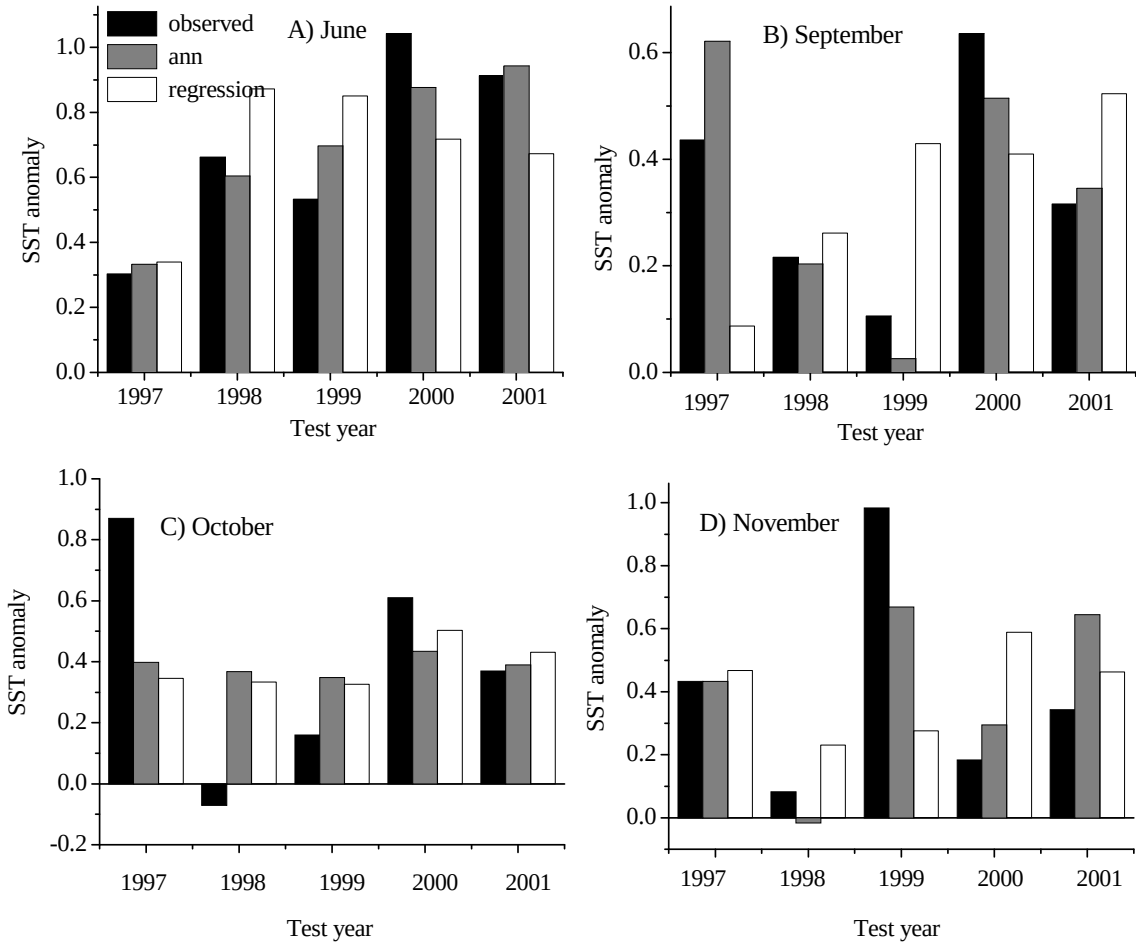

Fig. 5 — Prediction capability of the ANN model and the regression model for June (A), September (B), October (C) and November (D).

showing that the models are good enough to capture the anomaly trends for the training data. However, the performance of the regression models on the test (unseen) data is inferior to that of the corresponding ANN model (see Tables 1B and 2B for comparisons). It has been noted that the regression models fail to predict the anomalies for the months of June, September, October and November. The performance of ANN and regression models for the months of June, September, October and November is shown in Fig. 5. For June (Fig. 5A) the regression model has shown a correlation coefficient of 0.44 as against 0.92 for the ANN model. The RMSE of the regression model for June is 0.25, which is slightly better than the SD of the observed data (0.26) and that of the ANN model is 0.1. For September and November (Fig. 5B and 5D respectively) the regression model fails outright with correlation coefficients = -0.14 and –0.29 respectively. The corresponding ANN models give correlation coefficients of 0.87 and 0.76 respectively. Also, RMSE of the regression models for September (0.25) and November (0.37) exceeds the corresponding SD of observed data (0.18 and 0.31 respectively). This means that the prediction is worse than even the mean prediction. The RMSE of the ANN models for September and November are 0.1 and 0.2 respectively. These RMSE values are significantly less than the SD of the observed data for the corresponding month and the prediction can be considered to be better than the mean prediction. For October (Fig. 5C), the regression model has a correlation coefficient of only 0.38 against 0.69 for the ANN model. The RMSE for both the regression and ANN models is 0.31 for October.

The RMSE for the months of February, July, August and October are same for both the models (Tables 1B and 2B). These are the months in which the regression model and the ANN model have given

similar performances. This implies that the relationship between the past and present anomalies has a dominant linear component for these months and introducing the non-linear model does not enhance predictability. Although the RMSE of regression models is smaller than that of the ANN models for the months of January, March, April and December, the difference is significant only for December. But, even then the ANN model does not "fail" because the RMSE is still smaller than the SD of the observed data with a high correlation coefficient. The comparison of prediction skills of ANN and regression models has been made recently while predicting the equatorial Pacific[5], Indian Ocean[9] and tropical Pacific[12] sea surface temperatures. Further, recently it has been found that in general the non-linear multimodel ensemble (ANN) shows better prediction skill than other models[26]. Our objective here is not to improve the predictability to arbitrary precision. But, we are interested only in the fact whether the ANN can capture the past trend and predict the future anomalies correctly because a successful simulation of the SST anomaly variation of Indian Ocean can contribute to the monsoon forecast of the Indian subcontinent. The comparison with regression technique is made only to establish the simulation results.

4. Conclusion

ANN models have been designed for the prediction of SST anomalies over the Indian Ocean region (27° to 35° S and 96° to 104° E). Predictors are determined using the lag correlation of the series of the predictand month with the series of all the 12 months, of the grid point belonging to the same area as well as with the grid points belonging to the neighboring area. The performance of the models has been evaluated on the training and test sets. It has been observed that the ANN models have successfully simulated the observed behavior of the SST anomalies and have been able to predict the anomaly trends and give the forecast better than the mean forecast for every month. The performance of the ANN models has also been compared with the corresponding multivariate linear regression models. Except for the months of June, September, October and November, the performance of both the techniques are comparable and forecasts are fairly good. This is due to the linear relationship between the past and present anomalies. In such cases introducing the nonlinear ANN model does not significantly improve the predictability. However, it has also been found that the ANN models perform significantly better than the regression model even for the months of in which the regression model fails to give any forecast. For these months the non-linearity of the ANN models bring significant improvements in the forecast skills.

5. Acknowledgement

The authors are grateful to Dr. S.L. Srivastava and Dr A K Mittal for useful discussion and valuable suggestions. The authors (KCT and IMLD) are also grateful to the NCAOR/DOD for the financial assistance to carry out this work.

References

1 Aleksander, I. & Morton, H., *An introduction to neural computing*, (Chapman and Hall, London) 1955, pp. 1-5.

2 Tangang, F.T., Hsieh, W. W., & Tang, B., Forecasting the equatorial sea surface temperatures by neural network models, *Clim. Dynam*, 13 (1997) 135-147.

3 Gemmill, W.H. & Krasnopolsky, V.M., The use of SSM/I data in operational marine analysis, *Wea. Forecast.*, 14 (1999) 789-800.

4 Tang, Y. & Hsieh, W.W., Hybrid coupled models of the tropical Pacific-ENSO prediction, *Clim. Dynam*, 19 (2002) 343-353.

5 Tang, B.Y., Hsieh, W.W., Monahan, A.H. & Tangang, F.T., Skill comparisons between neural networks and canonical correlation analysis in predicting the equatorial Pacific sea surface temperatures, *J. Climate*, 13 (2000) 287-293.

6 Yuval, Enhancement and error estimation of neural networks prediction of Nino 3.4 SST anomalies, *J. Climate*, 14 (2001) 2150-2163.

7 Yoo, S.H., Yang, S. & Ho, C.H., Variability of the Indian Ocean sea surface temperature and its impacts on Asian-Australian monsoon climate, *J. Geophys. Res.*, 111 (2006) D03108, doi:10.1029/2005JD006001.

8 Hoerling, M. & Kumar, A., The perfect ocean for drought, *Science*, 299 (2003) 691-694.

9 Kug, J.S., Kang, I.S., Lee, J.Y. & Jhum, J.G., A statistical approach to Indian Ocean sea surface temperature prediction using a dynamical ENSO prediction, *Geophys. Res. Lett.*, 31 (2004) L09212, dio:10.1029 / 2003GL019209.

10 Kug, J.S., Soon-II An, Fie-Fie, J. & In-Sik, K., Precondition for El Nino and La Nina onsets and their relation to Indian Ocean, *Geophys. Res. Lett.*, 32 (2005) L05706, doi:10.1029./2004GL021674.

11 Chang, Y.S., Lee, D.U., Youn, Y.H. & Seo, J.W., Time series analysis of the subsurface oceanic data and prediction of the sea-surface temperature in the tropical pacific, *J.Korean Earth Sci. Soc.*, 26(2005) 706 – 713.

12 Wu, A., Hsieh, W.W. & Tang, B., Neural network forecasts of the tropical Pacific sea surface temperatures, *Neural Networks***,** 2005. (in press).

13 Bishop, C.M., *Neural networks for pattern recognition*, (Oxford University Press, New Delhi, India) 1995, pp 140-148.

14 Boskovitz, V. & Guterman, H., An adaptive neuro-fuzzy system for automatic image segmentation and edge detection, *IEEE Trans. Fuzzy Syst*, 10 (2002) 247-262.

15 Reynolds, R. W., & T. M. Smith, Improved global sea surface temperature analysis using optimum interpolation. *J. Climate*, 7 (1994) 929-948.

16 Haykin, S, *Neural networks*, (Pearson Education, Delhi, India) 2002, pp. 50-105.

17 Cybenko, G., Approximation by superposition of a sigmoidal function, *Math. Control Signal*, 2 (1989) 303-314.

18 Hornik, K., Stinchcombe, M. & White, H., Multilayer feedforward networks are universal approximators, *Neural Networks*, 2 (1989) 359-366.

19 Hsieh, W.W. & Tang, B., Applying neural network models to prediction and data analysis in meteorology and oceanography, *Bull Amer Meteorol Soc*, 79 (1998) 1855-1870.

20 Masters, T., *Practical neural network recipes in C++*, (Academic Press, California, USA) 1993, pp. 10-300.

21 Monahan, A. H., Nonlinear principal component analysis by neural networks: theory and pplication to the Lorenz system, *J.Climate*, 13 (2000) 821-835.

22 Tsintikidis, D., Haferman, J.L., Anagnostou, E.N., Krajewski, W. F. & Smith, T. F., A neural network approach to estimating rainfall from spaceborne microwave data, *IEEE Trans. Geosci Remote Sens.*, 35 (1997) 1079-1093.

23 Wilks, D.S., *Statistical methods in the atmospheric sciences*, (Academic Press, California, USA) 1995, pp. 169-369.

24 Press, W.H., Teukolsky, S.A., Vellerling, W.T. & Flannery, B.P., *Numerical recipes in C*, (Cambridge University Press, Delhi, India) 1992, pp. 36-37.

25 Saji, N.H., Goswami, B.N., Vinayachandran P.N. & Yamagata T., A dipole in the tropical Indian Ocean, *Nature*, 401 (1999) 360-363.

26 Chau, K.W., Wu, C.L. & Li, Y.S., Comparison of several flood forecasting models in Yangtze river, *J. Hydraul. Eng. - ASCE*, 10(2005) 485-491.