J. Ind. Geophys. Union (2003)
Vol.7, No.4, pp.173-181

Prediction of Indian Rainfall Index (IRF) using the ENSO variability and Sunspot Cycles - An Artificial Neural Network Approach

S.Sri Lakshmi, R.K.Tiwari and V.K.Somvanshi
National Geophysical Research Institute, Hyderabad – 500 007

ABSTRACT

Prediction of climate variability is one of the main concerns of geoscientists. The Indian rainfall variability is the end product of a series of complex interactions between the ocean and atmospheric processes, so any effect in this process will be reflected in the expected rainfall. Recent studies have indicated the possible role of solar and greenhouse radiative forcing in earth-ocean-atmospheric phenomenon. Hence in order to achieve the goal of forecasting of Indian rainfall variability, it is necessary to have a global data of the oceanic and atmospheric phenomena in conjunction with external forcing indices such as sunspot time series. Here in this paper, we apply the artificial neural networks (ANN) based backpropagation scheme to forecast the Indian summer monsoon rainfall (IRF) using the records of ENSO variability and sunspot cycle. Our analyses reveal a considerable degree of link among Solar/Sunspot cycle, ENSO related temperature variability (NINO3 time series) and Indian rainfall index suggesting possible role of exogenic-triggering in reorganizing the tropical ocean-atmospheric system. The analysis, may, provide useful constraints for the modeling of tele-connected tropical pacific climatic variability and Indian rainfall.

INTRODUCTION

The atmosphere, oceans and biosphere are globally coupled phenomena, which is often called the Earth system. Climatic variability is a complex environmental feature that is intimately associated with several factors such as temperature, precipitation, sunlight, and wind. The climate system varies and changes quasi-periodically because of the complexing and interactions of the atmosphere with the other Earth system components. As a result there are recurrences of short-term extreme events such as floods and droughts; and long-term episodes such as dry and wet decades, cool or warm centuries, and glacial-interglacial cycles. There are two most fundamental climate variations (i) the daily and (ii) seasonal cycles. These variations result from relationships between the sun and Earth on time scales of days and years. The third most important climate variation is El Nino-Southern Oscillation (ENSO) events, which impact the global oceanic and atmospheric circulations and can produce droughts and floods in certain regions. Climate also varies on time scales ranging from decades to millions of years. Decade-long variations result from interactions among the different components of the Earth system: atmosphere, ocean, land, biosphere, and ice. Because each of these components is characterized by different response times, their interactions produce climate variations on many time scales.

The climate system is generally considered to operate in a complex and non-linear way. Resolving the complex nature of the variability requires robust statistical investigations. Recently modern nonlinear techniques like forecasting time series analysis (based on the concept of nonlinear dynamical theory) for prediction of nonlinear geophysical time series has been exemplified by several workers (Farmer & Siderowich 1987; Sugihara & May 1990; Casdagli 1989; Tiwari et al. 2003). These methods have been used to distinguish/characterize the nature of the dynamical behavior and do not undertake part-to-part prediction. Here we attempt to predict the pattern of precipitation based on the concept of neural network. The complexity in the precipitation record seems to be primarily associated with both external forcing as well as internal atmospheric-ocean processes. It is, therefore, somewhat difficult to assign a single causative mechanism for controlling the behavior of monsoon over the Indian subcontinent. In this situation, it is therefore imperative to understand the predictive nature of this temporal variability record by modeling the data using multiple non-linear input-output relations. Neural Network Models (NNM) is an appropriate tool for investigations of the climate system. We apply the most widely used NNM architecture, the Backpropagation Network (BPN), to analyse the dynamics of the precipitation and temperature related climate system. The purpose of the present work is to: (1) gain better insights about the nature of rainfall fluctuations within the Indian sub-continent in relation to the global climate indices such as ENSO; (2) determine the signature of inter-decadal fluctuations and its relations with indices of global climate, if any; (3) develop a dynamical model using the neural networks to test the short-term climate prediction and possible linkages and coupling of ocean and atmospheric processes.

EL NINO-SOUTHERN OSCILLATION EVENT (ENSO)

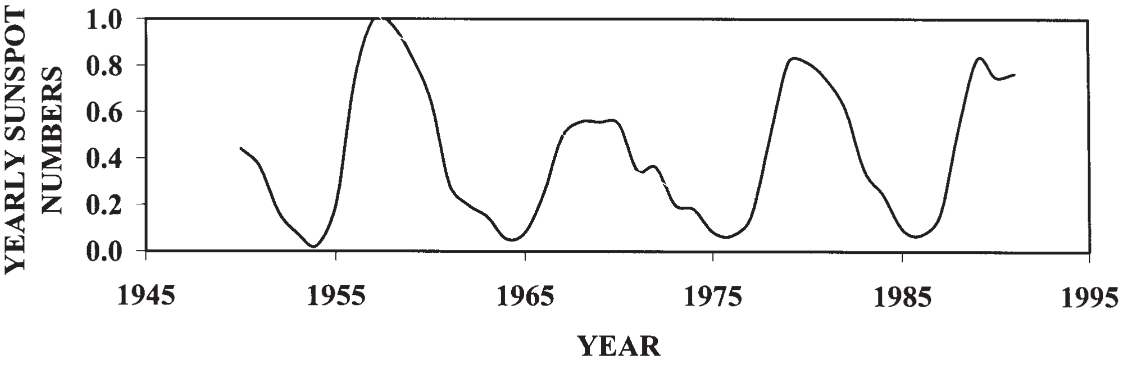
The strong coupling and interactions between the Tropical Ocean and atmosphere play a major role in the development of global climatic system. The El Nino/Southern oscillation is the outcome of such a coupling and refers to a warm inshore current annually flowing south along the coast of Ecuador but extending down the coast of Peru. The El Nino events generally recur approximately every 3-5 years with large events spaced around 3-7 years apart. The Southern Oscillation Index (SOI) is the measure of sea level atmospheric pressure difference between Darwin Australia (western Pacific) and Tahiti (eastern Pacific) (Philander 1990; Cane 1992; Bigg 1996). There is a strong coupling between the El Nino event and the Southern Oscillation Index. The El Nino-Southern Oscillation event is often referred to by the acronym ENSO. El Nino episodes (also called Pacific warm episodes or ENSO) and La Nina episodes (also called Pacific cold episodes) represent opposite extremes of the ENSO cycle. This has far reaching hazardous impact including drought, floods and intense rainfall with severe human consequences as well as distinctive "telelinked" pattern of climatic anomalies. This phenomena has also impact on the Asian monsoon that drives the surface ocean seasonality in the Indian and western Pacific oceans and Atlantic inter-tropical convergence zone (Cole, Fairbanks & Shen1993). An event of this type affects the climate of a large portion of the globe. The strongest and most reliable effects occur in the tropical Pacific Ocean where El Nino and Southern Oscillation are strongly coupled. Other parts of the world, especially in the middle Latitudes are affected through teleconnections (Philander 1990). Teleconnections are represented as statistical associations among climatic variables separated by large distances.

CLIMATIC AND SOLAR SUNSPOT CYCLES

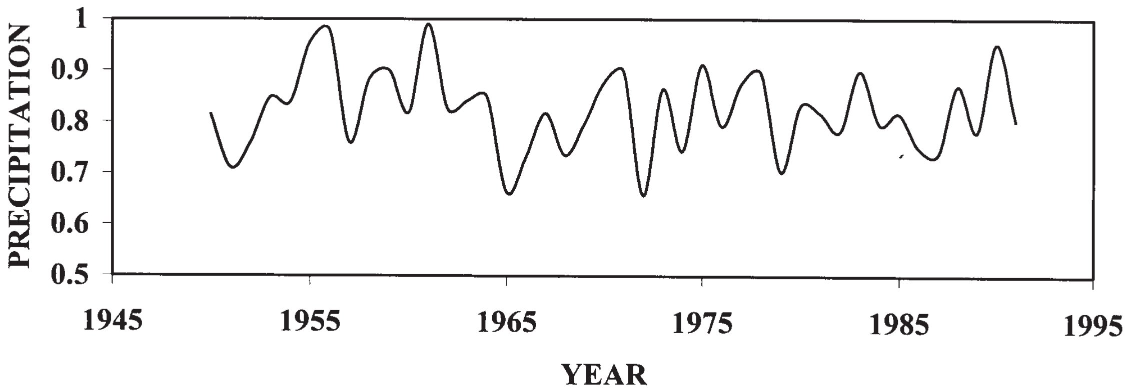
Sunspots have been observed since thousands of years. The Sunspot number is a measure for solar variability. The Wolf or Zurich Sunspot number is defined as ten times the number of sunspots plus the number of sunspots all multiplied by an observer-related constant. (Hoyt & Schatten 1997). It has been established that the sunspot activity is a cyclic phenomena with periodicities of 11, 22 and 80 years. Several recent studies of solar-climate relationship have established that the lower temperatures are associated with below average sunspot activity while the higher temperatures are associated with above average sunspot activity. This cyclic pattern also appears to be associated with the global climatic fluctuations. Recent research workers (Labitzke & Van Loon 1989, 1992, 1993) have provided intriguing evidences, which suggest that a possible link exists between solar cycles and the earth's climate. Mendoza, Perez-Enriquez & Alvarez-Madrigal (1991) reported on possible connections between solar activity and El Nino's, while Reid & Gage (1988) and Reid (1991) reported on the similarities between the 11-year running means of monthly sunspot numbers and global sea surface temperature. These findings suggest that there is possible coupling between temperature-ENSO and solar signals.

SOURCE OF DATA

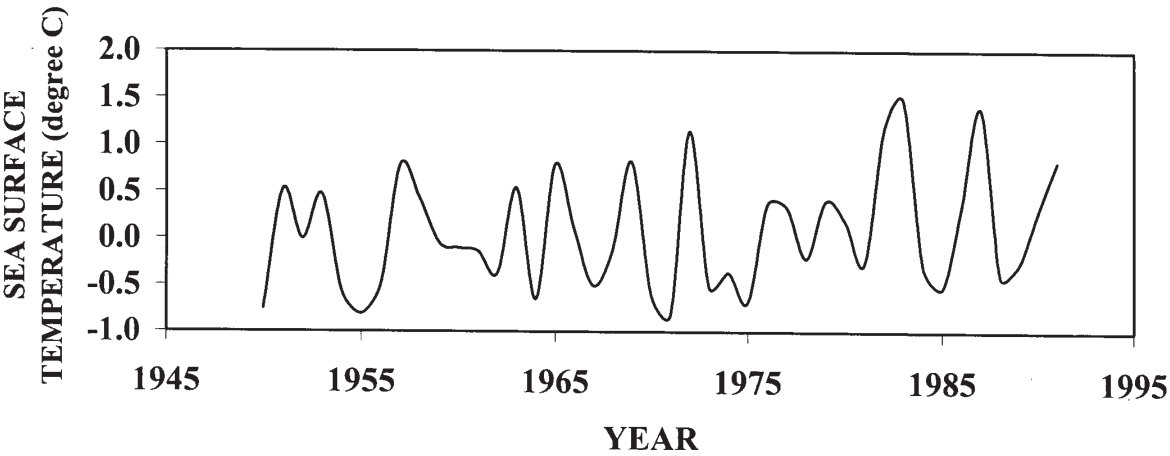
For the present analyses, we have taken here the following three sets of data for a common period of 42 years spanning over 1950 –1991: (1) Smoothed Sunspot number (2) updated Indian Rainfall (IRF) time series (in mm) of the whole country and (iii) NINO3 temperature record as proxy for the ENSO response. The smoothed yearly sunspot number is taken from the National Geophysical Data Center, Boulder, Colorado shown in Fig 1(a). The recently updated instrumental annual rainfall index of the whole country (Singh & Sontakke 1996) is shown in Fig. 1(b). The present annual IRF data is area weighted, homogeneous and cover the whole country. The third data set is NINO3 Global Sea Surface Temperature (SST) indices (in ^{o}C) which is one of the best available records of temperature variability from the Eastern equatorial Pacific (5^{o}S-5^{o}N, 150^{o}-90^{o}W) (proxy for ENSO) variability (Kaplan et al. 1998) and widely analyzed by several workers (Felis et al. 2000) (Fig. 1(c)).



(a)



(b)



(c)

Figure 1. Time series plot for the three data sets (a) sunspot numbers (b) Indian rainfall (in mm) of the whole country and (c) sea surface temperature (in °C) for the year 1950-1991.

ARTIFICIAL NEURAL NETWORKS

An artificial neural network (also referred as neural network) is a computing method works on the principle of structure of brains and nerve systems. When compared with the other analytical approaches, the neural network approach does not require human expert knowledge in terms of mathematical descriptions of the problem. A typical neural network consists of inter-connected set of processing units called neurons. Here in this present study we have used the feed-forward artificial neural network.

Backpropagation

There are number of algorithms available for the training of neural network. Among them, the Back-propagation is most commonly used [developed independently by several authors (Werbos 1974; Parker 1982; Rumelhart, Hinton & Williams 1986)] and has been applied successfully to a broad range of fields such as speech recognition, pattern recognition, and image classification. Its training procedure is intuitive because of its relatively simple concept i.e. adjust the weights to reduce the error.

Back-propagation networks topology is usually layered, with each layer fully connected to the layer before it and the one next to it. Neurons receiving input data form the input layer, while those generate output to users form the output layer. The input to the network propagates forward from the input layer, through each intermediate layer, to the output layer, resulting in the output response. When the network corrects its connecting weights, the correction process starts with the output units and propagates backward through each intermediate layer to the input layer-hence the term Back propagation. Fig.2 shows the schematic of a three-layer back-propagation neural network.

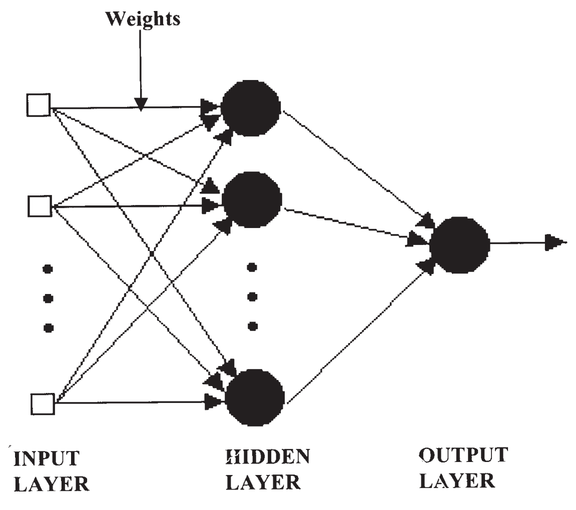


Figure 2. Architecture of three-layered feed forward back propagation neural network.

A back-propagation network may have one or more than one hidden layers.

DATA PREPARATION AND NETWORK TRAINING

This procedure is crucial to the success of applying neural network approach. The data preparation includes the selection of input variables, and normalizing between 0 and 1. The normalization is necessary for two reasons: (1) If the data used with a neural network is not scaled to an appropriate range, then the network will not converge on training, or otherwise will not produce meaningful results; (2) for the neurons' transfer functions. Since if either a sigmoid function or a hyperbolic tangent is calculated, then these can only be performed over a limited range of values. A neuron only produces output whose absolute value is less than 1, since the transfer function has asymptotes at $f(x) = 1$ and -1 (the exception to this being when a linear transformation is used).

There are two phases in its training cycle, one to propagate the input pattern and the other to adapt the output. It is the errors that are backward propagated in the network iteration to the hidden layer(s). During the neural network training each hidden and output neurons process the inputs by multiplying them with their weights. The products are thereby summed and processed using an activation function like sigmoid, tan sigmoid etc., It is the errors that are backward propagated in the network iteration to the hidden layer. (Rumelhart & McClelland 1986).

There is no universally applicable formula to be used for deciding the size of middle layers. Generally networks with too many hidden neurons tend to memorize the input patterns and with too few hidden neurons may not be able to simulate a complex system at all. A network with more hidden neurons also requires more computing power and more training time needed. It is a common practice to fix the number of hidden layers in the network and then choose the number of neurons in these layers. It has been shown that only one hidden layer is required to approximate any continuous function, given that sufficient degrees of freedom (i.e. connection weights) are provided. (Cybenko 1989). Hence one hidden layer has been used for this study. Consequently network with one hidden layer of 25 neurons converges faster (took less time) to reach the minimum error goal of 0.001.

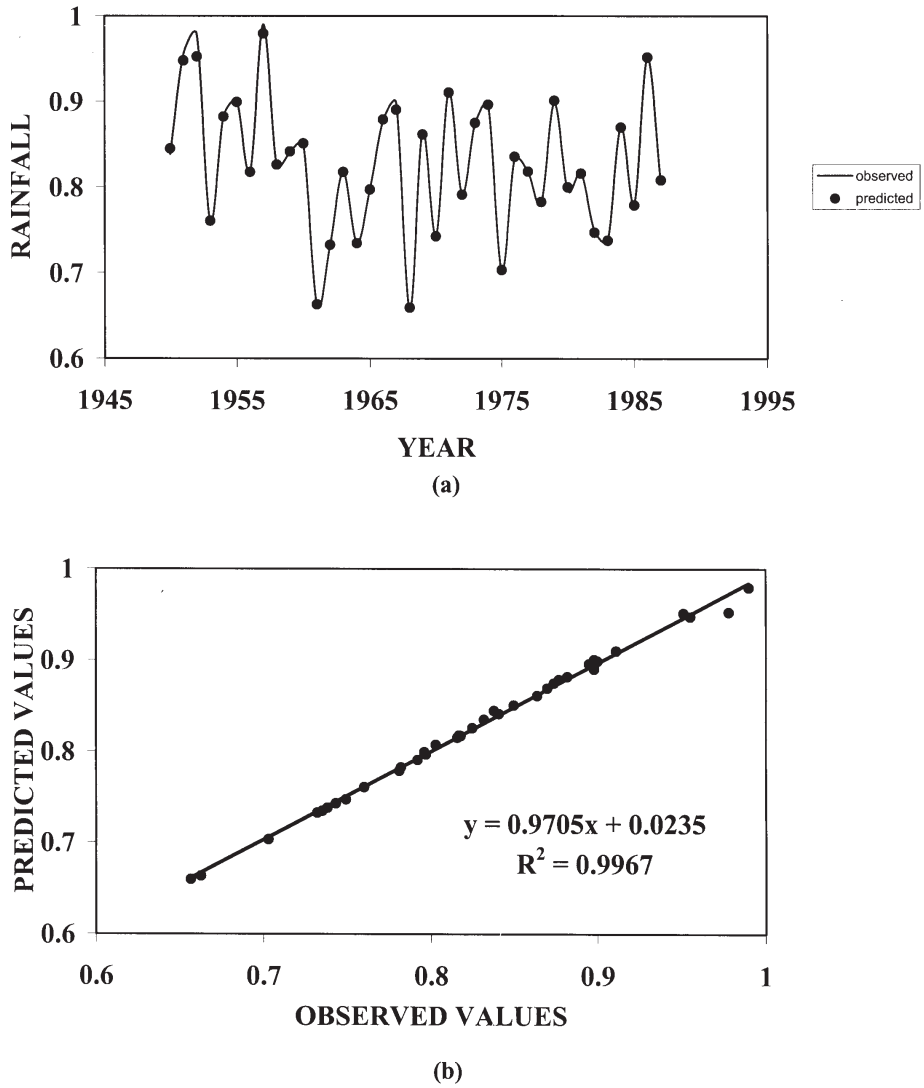


Figure 3. (a) Predicted model and (b) Scatter diagram showing the predicted Vs observed values of the IRF time series data displayed in fig. 1(b).

RESULTS

The selection of input variables is solely problem dependent. After analyze the problem, three variables were chosen for this study. They are: Southern Oscillation Index (SOI), Sunspot Number and Yearly rainfall index. The analyses includes mainly two steps: Firstly the data for the three input sets i.e. Solar/ Sunspot cycle, ENSO related temperature variability (NINO3 time series) and Indian rainfall index time series are trained individually. Three-layered feed forward neural network models were constructed to analyse the climatic variability of the Indian continent. Secondly, taking the ENSO related temperature variability (NINO3 time series) and the solar sunspot number as two input variables and the IRF time series as the output for the neural network, the network was trained and best fitted network model for the prediction of the precipitation pattern over the Indian continent has been constructed.

Fig.3 (a, b) show the fitted predicted neural network model and its corresponding scatter diagram illustrating the predicted Vs observed values respectively for the IRF time series displayed in Fig.1b. It can be seen that most of the predicted and observed values coincide along the straight line. This might suggest that the underlying IRF data have some deterministic components in it. If the data is non-deterministic, it will not fall in such an orderly pattern. However, the more perfect is the network training, more closely the predicted points organized towards the straight line. The ANN models provided

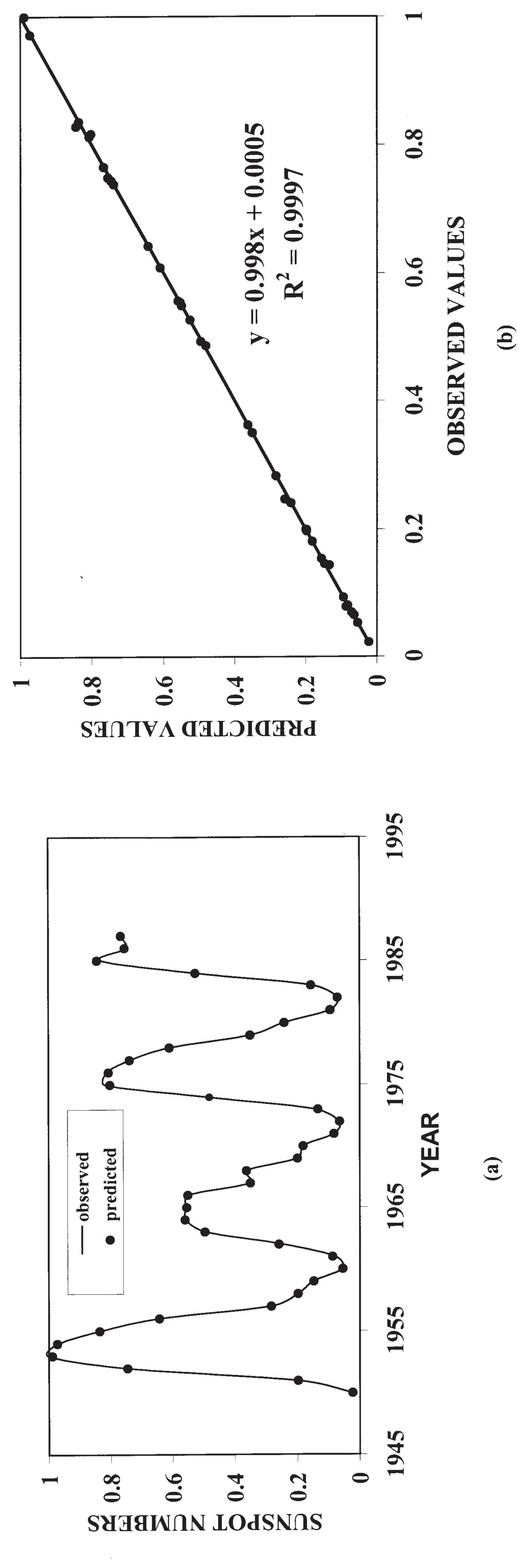


Figure 4. (a) Predicted model and (b) Scatter diagram showing the predicted Vs observed values for solar sunspot numbers.

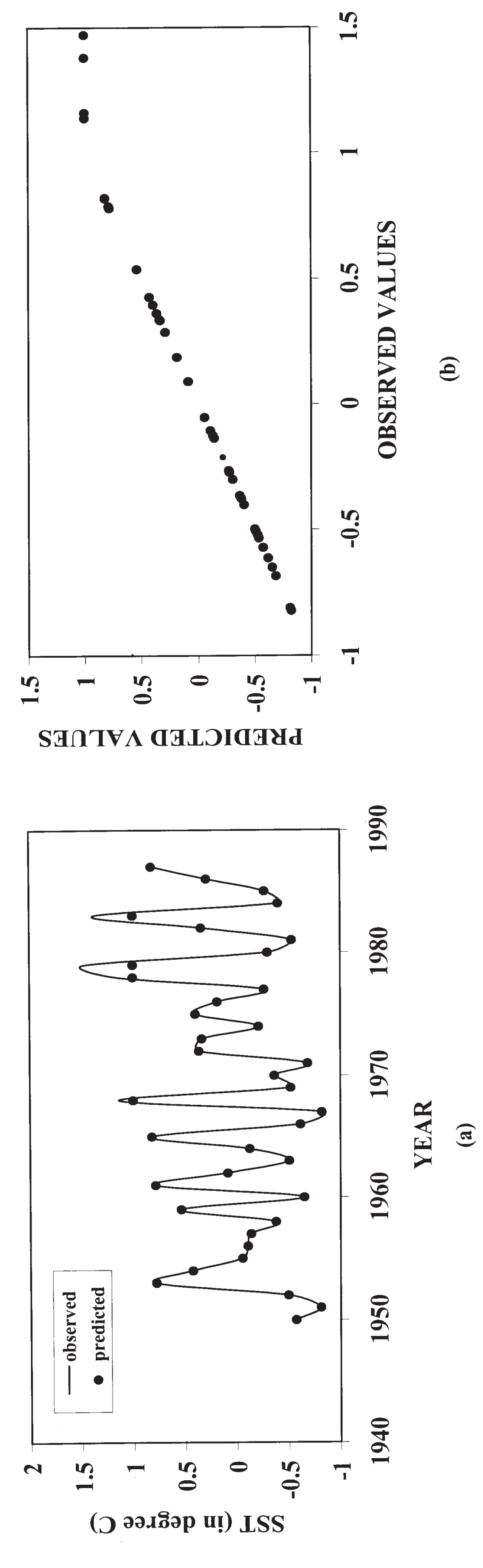


Figure 5. (a) Predicted model and (b) Scatter diagram showing the predicted Vs observed values of temperature variability (proxy for ENSO).

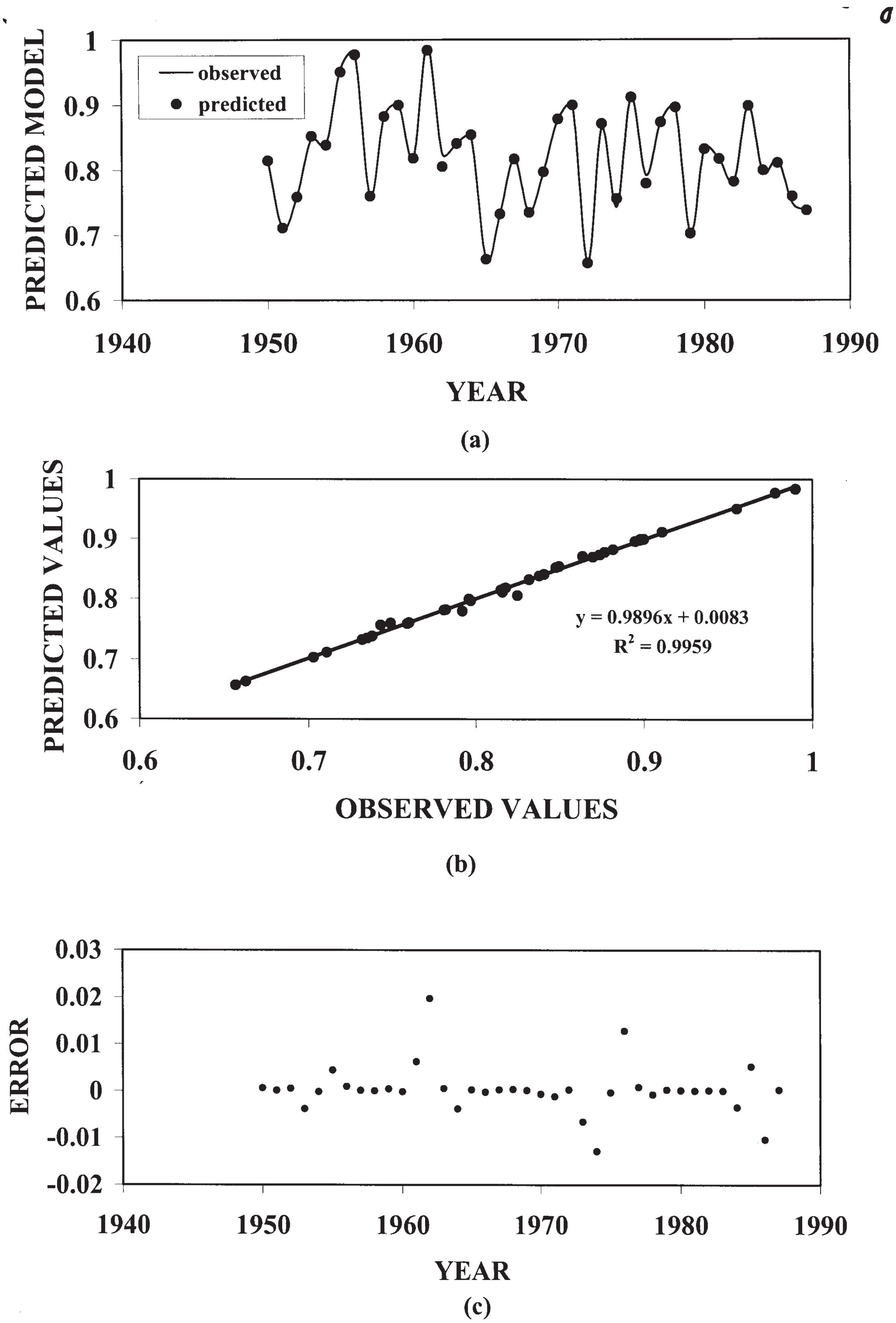


Figure 6. (a) Predicted model (b) Scatter diagram showing the predicted Vs observed values (c) Year wise error plot for the IRF time series considering the temperature variability data (NINO3 time series, proxy for ENSO) and the solar sunspot number as the two input variables for the network.

a good fit showing a high feasibility for the further prediction.

Similarly Fig.4(a) and (b) show the fitted predicted neural network model and its corresponding scatter diagram illustrating the predicted Vs observed values for the Sunspot cycles. Fig.5 (a) and (b) shows the fitted predicted neural network model and its corresponding scatter diagram illustrating the predicted Vs observed values for the ENSO variability. It is quite interesting to note that the three ANN models show good fit suggesting the applicability of ANN technique for predicting complex geophysical problems.

Keeping in view the predictive capability of the ANN technique and also considering the possible linkages among the ocean atmospheric processes as discussed in introduction, we tried here to predict one phenomenon considering another as an input parameter. Figure 6 (a) shows the predicted model for the IRF time series, considering (i) the temperature variability data (NINO3 time series, proxy for ENSO) and (ii) the solar sunspot number as the two input variables for the neural network. Figure 6 (b) shows the corresponding scatter diagram illustrating the predicted Vs observed values for the same time series obtained from the network training. Figure 6 (c) show the year wise error plot for the fitted predicted model. Here the total error obtained is around ± 0.02 for the model, which is considerably less. This shows that there is an ample depending away these indices and hence suggest a possible link among them.

The neural network modeling result shows that prediction of precipitation pattern over the Indian continent is possible within some error bounds. However, there are several other feedback factors affecting the Indian monsoon that have to be considered. A more detail analysis taking into consideration various other parameters are required for detail analysis and interpretations. This result, however, provides useful basis for developing models for the study of monsoonal behavior and prediction on rather more rational way.

CONCLUSIONS

Our results suggest that neural network technique is promising tools for modeling and prediction of the complex environmental system. Results can be summarized as follows:

(i)The ANN constructed models suggests high feasibility of the application of ANN technique for the prediction of yearly rainfall index within some reasonable error bounds.

(ii)The predictive analysis suggest a considerable degree link among Solar/Sunspot cycle, ENSO related temperature variability (NINO3 time series) and Indian rainfall index and

(iii) Finally, it provides evidence for tele-connection of tropical pacific climatic variability across time scale ranging from years to decades and also the possible role of exogenic-triggering in reorganizing the tropical ocean-atmospheric system.

ACKNOWLEDGEMENTS

One of us, Mrs. S.Sri Lakshmi is extremely thankful to CSIR for providing the senior research fellowship to carry out the research work. We are grateful to Dr. Michael Evans for suggesting the site for NINO3 data. We are also thankful to the Director, National Geophysical Research Institute, Hyderabad for his motivation and kind permission to publish this work.

REFERENCES

Bigg, G.R., 1996. The Oceans and Climate Cambridge University Press, Cambridge. pp. 1-266.

Cane, M.A., 1992. Tropical Pacific ENSO models: ENSO as a mode of the coupled system. In: Climate System Modelling, Ed: K.E. Trenberth. Cambridge University Press, Cambridge, pp. 583-614.

Casdagli, M., 1989. Non-linear prediction of chaotic time series, Physica D 35, 335-336.

Cole, J.E., Fairbanks, R.G.& Shen. G.T., 1993. Recent variability in the Southern Oscillation: Isotopic results from a Tarawa Atoll coral. Science, 260, 1790-1793.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal functions, Math. Control Signals Syst., 2, 203-314.

Farmer, J.D. & Siderowich, J.J., 1987. Predicting chaotic time series, Phys. Rev. Lett., 59, 845-848.

Felis. T., Patzold. J., Loya. Y., Fine. M., Nawar. A.H. & Wefer. G., 2000. A coral oxygen isotope record from the northern Red Sea documenting NAO, ENSO, and North Pacific teleconnections on Middle East climate variability since the year 1750. Paleoceanography, . 15 (6), 679-694.

Hoyt, V. & Schatten, K.H., 1997. The Role of the Sun in Climate change. Oxford University Press, Oxford, pp.36 (1-279).

Kaplan A., Cane M., Kushnir Y, Clement A, Blumenthal, M. & Rajgopalan, B., 1998. Analysis of global sea surface temperature. J. Geophys. Res., 103, 18567-18589.

Labitzke, K. & Van Loon, H., 1989. Association Between the 11-Year Cycle, the QBO and the Atmosphere, J..Climate, 2.

Labitzke, K. & Van Loon, H., 1992. Association between the 11-Year solar cycle and the Atmosphere. Part V: Summer, J. Climate, 5, 240-251.

Labitzke, K. & Van Loon, H., 1993. Some recent studies of probable connections between solar and atmospheric variability, Ann. Geophysicae, 11.

Mendoza, B., Perez-Enriquez, R. & Alvarez-Madrigal, M., 1991. Analysis of solar activity conditions during periods of El Nino events, Ann. Geophysicae, 9, 50-54.

Parker, D. B., 1982. Learning Logic: Stanford University, Stanford, California, Office of Technology Licensing, Invention Report S81-64, File 1, pp 119.

Philander, S.G., 1990. El Nino, La Nina and the Southern Oscillation. Academic Press, London, pp. 1-293.

Reid, G.C., 1991. Solar irradiance variations and global Ocean Temperature, J. Geomagn. Geoelectr., 43, 795-801.

Reid, G.C. &. Gage, K.S., 1988. The climatic impact of secular variations in solar irradiance, in Secular Solar and geomagnetic Variations in the Last 10000 years, Eds. F.R. and A.W. Wplfendale, NATO AS] Series, Kluwer, Dordrecht, 225-243.

Rumelhart, D. E., Hinton, G. E. & Williams, R .J., 1986. Learning internal representation by back propagating errors. Nature, 332, 533-536.

Rumelhart, D.E. & McClelland, J.L., 1986. Parallel Distributed Processing, MIT Press, Cambridge, Massachusetts, pp.318-362.

Singh, N. & Sontake, N.A., 1996. The instrumental period rainfall series of the Indian region: documentation. Contribution from Indian Institute of Tropical Meteorology, Pune, Research Report No.PR-067 ISSN 0252-1075.

Sugihara, G. & May, R.M., 1990. Non-linear forecasting as a way of distinguishing chaos from measurement error in time series, Nature, 344, 734-749.

Tiwari, R.K., Srilakshmi. S. & Rao, K.N.N., 2003. Nature of earthquake dynamics in the central Himalayan region: a nonlinear forecasting analysis. J. Geodynamics, 35, 273-287.

Werbos, P. J., 1974. Beyond regression: new tools for prediction and analysis in the behavioral sciences: unpubl. Masters thesis, Harvard Univ. Cambridge, Massachusetts, pp. 453.

(Accepted 2003 July 29. Received 2003 July 22; in original form 2002 November 28)