Winter shower in India - A Neurocomputing based predictive model

***S. S. De*[1], *Goutami Chattopadhyay*[1] *and Surajit Chattopadhyay*[2]**

[1]Centre of Advance Study in Radio Physics and Electronics, University of Calcutta
1, Girish Vidyaratna Lane, Kolkata- 700 009, India
[1]e-mail: de_syam_sundar@yahoo.co.in

[2]Department of Information Technology
Pailan College of Management and Technology, India

Abstract

The development of a neurocomputing technique to forecast the average winter shower in India has been modeled from 48 years of records (1950 to 1998). The complexities in the rainfall-sea surface temperature relationships have been statistically analyzed where linear as well as polynomial trend equation are obtained. The coefficient of determination for linear trend is very low and it remains low, where polynomial of degree six is used. For this reason, Artificial Neural Net Model as predictive tool for the said meteorological event in the form of Multiple Layer Perceptron has been generated with sea surface temperature anomaly and monthly average winter shower data over India during the above period. After proper training and testing, a Neural Net model with small prediction error is developed and supremacy of Artificial Neural Net over conventional statistical predictive procedure has been established statistically.

1. Introduction

Accurate rainfall predictions are necessary for planning day-to-day activities [1]. Sea Surface Temperature (SST) anomalies influence the atmosphere by altering the flux of latent heat and sensible heat from the ocean [2]. In tropics, positive SST anomalies are associated with enhanced convection and the resulting heating is balanced by adiabatic cooling [3]. SST anomalies also play an important role in producing rainfall [4]. El-Niño-Southern Oscillation (ENSO) is a coupled Ocean-atmosphere phenomenon that has worldwide impact on climate in general and Indian monsoons in particular [5]. The oscillations in wind stress owing to the Southern Oscillation are associated with changes in the circulation of the ocean and the SST anomally that are referred to as El Nino. Warm ENSO episodes are characterized by increased number and intensity of tropical storms over the Bay of Bengal and hence enhanced winter monsoon rainfall [5].

Present paper endeavors to develop an Artificial Neural Network (ANN) model to forecast average winter shower in India. The model is particularly useful when the underlying physical processes are not fully understood or display chaotic properties [6]. The outcome of this model work has been compared with statistical model Multiple Linear Regression (MLR). ANN model shows lower prediction error than MLR forecast.

2. Data Analysis

The monthly average winter shower data from 1950 to 1998 have been collected from the IITM rainfall data series available at www.tropmet.res.in. The SST anomaly data pertaining to the same years are collected from http://jisao.washington.edu/datasets/global sstanomts/

The autocorrelation function of the predictand has been calculated from the following equation [7]

$$C(T) = \sum_{n=0}^{N} s(n)s(n+T) \qquad (1)$$

Where, *s(n)* is the average winter shower at time *n*, *N*=48 and T=0,1,2,……..,*n*

In Figure 1, it is apparent that the time series of average winter shower has no deterministic pattern. Because of highly chaotic nature of the data series, six predictors paired with the predictand have been generated giving six trend equation and six coefficient of determination (CD). The CD's for linear trend are too low. When polynomials of degree six are used in trend equation, the CD's still remained low. This proves very high degree of non-linearity between the predictor and predictand although they are physically related. Therefore, the necessity of neural net became inevitable.

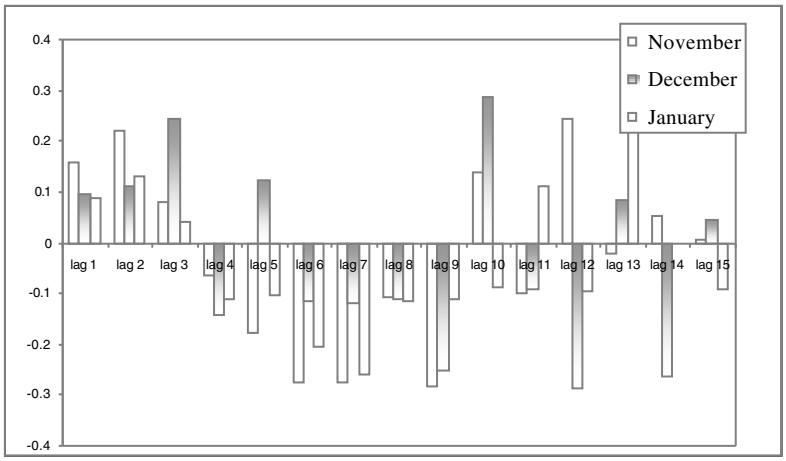

Figure 1. Schematic showing the autocorrelation function corresponding to the rainfall amounts in the winter months.

All data are scaled to provide values between 0.1 and 0.9 as follows:

$$z_i = 0.1 + 0.8 \times \left(\frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \right) \quad (2)$$

Where, z_i denotes the transformed appearance of the raw data x_i. After the modeling is completed, the scaled data are reverse scaled according to

$$P_i = x_{\min} + \left(\frac{1}{0.8} \right) \times [(y_i - 0.1) \times (x_{\max} - x_{\min})] \quad (3)$$

Where, P_i denotes the prediction in original scale. The training of the data is done by Backpropagation method. The method, implemented to predict average winter rainfall on a specific winter day 'd,' may be described as:

$$y_d = f_d(x_1, x_2, \ldots\ldots\ldots\ldots\ldots\ldots x_{24}) \quad (4)$$

Where, $x_1, x_2, \ldots\ldots\ldots\ldots\ldots\ldots x_{24}$ represent the measures of a parameter on 24 consecutive winter shower days. The form of the function f_d is obtained after adjusting a set of weights by using the training set of data. Within the training matrix, every row is a sample case. Percent error of the prediction (PE) has been computed [8] as:

$$PE = \frac{\langle | y_{predicted} - y_{actual} | \rangle}{\langle y_{actual} \rangle} \quad (5)$$

Where, $\langle \; \rangle$ implies the average over the whole test set.

3. Methodology

In this algorithm, an initial weight vector w_0 of a feed forward neural network is iteratively adopted according to the recursion,

$$w_{k+1} = w_k + \eta d_k \quad (6)$$

Where, w_j denotes the weight vector at the jth step. This recursion relation is used to find an optimal weight vector. Presenting a set of pairs of input and target vectors to the network, the adaptation is performed sequentially. The quantity η is called the learning rate [1]. The direction vector d_k is the negative of the gradient of the output error function E, which is the mean squared error at the kth step. Mathematically d_k is expressed as,

$$d_k = -\nabla E(w_k) \quad (7)$$

In the learning scheme for the back propagation [9] algorithm, the weight vector w_k contains the weights computed during the kth iteration and the output error function E is the multivariate function of the weights of the network. Mathematically this is expressed as

$$E(w_k) = E_p(w_k) \quad (8)$$

Where $E_p(w_k)$ represents the half-sum-of-squares error function of the network outputs for a certain input pattern p. The objective of this supervised learning is to select the set of weights that minimizes E, which is the deviation between the network output and the target pattern over the complete set of training patterns. This is presented to the neural network, called an epoch [10]. The learning continues until E is less than a present value at the end of an epoch.

In equation (1) if the function f_d is nonlinear, then non-linear perceptron is achieved. Additional room for a good fitting of data is obtained by introducing a set of hidden nodes $z_{dk}(k = 1,2,\ldots\ldots\ldots\ldots\ldots, n)$ in such a way that,

$$z_{dk} = f(w_{dk_1} x_1 + \dots\dots\dots\dots + w_{dk_{24}} x_{24} + w_{dk_0}) \quad (9)$$

And

$$y_d = f\left(v_1 z_1 + \dots\dots\dots\dots + v_{d_n} z_d + v_{d0}\right) \quad (10)$$

The function f is defined as,

$$f(z) = \frac{1}{1+\exp(-z)} \quad (11)$$

In order to find w's and v's, the back propagation method is to be used.

4. Results and comparison with conventional model

Equations (6) to (11) are implemented to train the ANN constructed with the sets of predictor and predictands. After on-line learning with 500 epochs, the ANN model is tested for the validation or test set. The predictions are shown in Figure 2. It is found that in many cases, there are close associations between actual and predicted winter shower over India. The prediction made by the Neural Net is then compared with the Multiple Linear Regression (MLR) forecast. Overall prediction error (PE) from ANN model comes out to be 0.17.

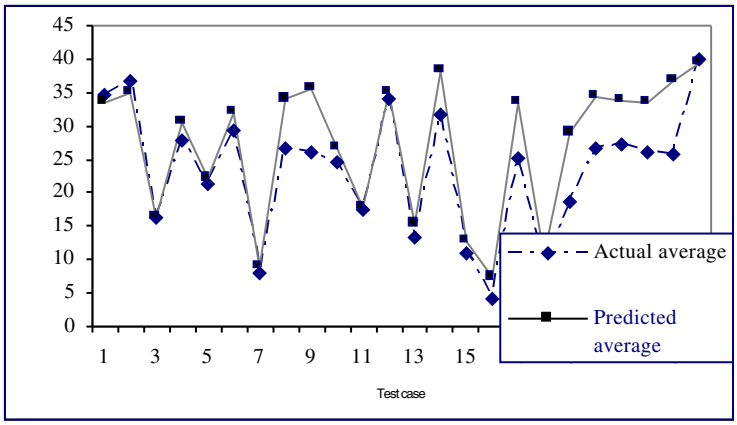

Figure 2. Actual versus predicted average winter shower in the test cases. The prediction is made by Artificial Neural Network (ANN).

A MLR equation is now fit to the set of six predictors and one predictand. Just like Neural Net model, first 50% data are used for training and last 50% data are used for testing [7, 11]. The predictive MLR equation is then fitted using the method of least squares and the equation comes out to be,

$$\hat{y} = 2.17E-14x_1 + 0.334x_2 + 0.334x_3 + 0.334x_4 + 1.497E-16x_5 + \\ +1.95E-16x_6 + 4.14E-17 \quad (12)$$

Where the first three variables in the right hand side imply the rainfall amounts in the months of November, December, January of the year n, the last three variables imply SST anomalies in the months of November, December, January of the year n and the left hand side implies the average winter rainfall predicted for the year (n+1). The PE produced by the MLR model for the test years is found to be approximately equal to 2.21, which is very much higher than the other one. All the PE's are presented in the Figure 3.

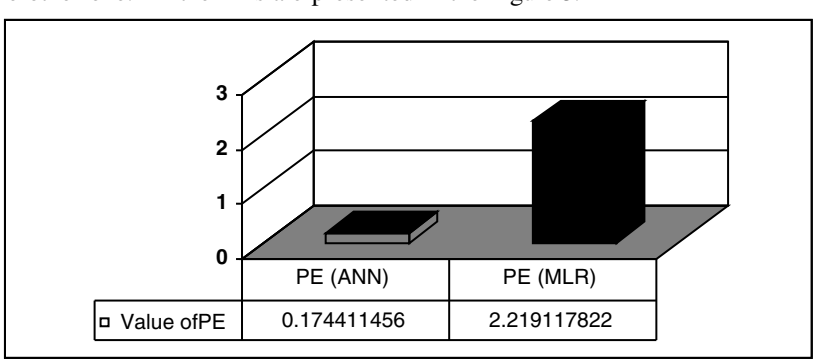

Figure 3. This figure compares the prediction errors (PE) produced by two competitive models in predicting average winter shower over India. The computation is made over the test cases.

5. Conclusion

ANN model has been implemented to predict average winter rainfall over India. Six predictors are generated in the input matrix for the Neural Net. After 500 epochs, the ANN has been found to produce a forecast with small prediction error. Finally the model has been compared with

statistical model MLR. It has been found that ANN has produced lower prediction error than MLR forecast.

6. References

1. Mart´ýn del Br´ýo B, Serrano Cinca C., "Self-organizing neural networks for analysis and representations of data: Some financial cases," *Neural Computing and Applications,* **1,** 1993, pp.193–206

2. Holton J. R., "An Introduction to Dynamic Meteorology," *Academic Press, San Diego*, 1972.

3. Reddy P. R. C. and Salvekar P.S., "Equatorial East Indian Ocean sea surface temperature: A new predictor for seasonal and annual rainfall," *Current Science,* **85,** 2003, pp.1600-1604.

4. Clark O. C., Cole J. E. and Webster P. J., "Indian Ocean SST and Indian summer monsoon rainfall: predictive relationships and their decadal variability,"*Journal of Climate,* **14,** 2000, pp.2503-2519.

5. Kumar B., "Impact of ENSO on Winter Monsoon Rainfall over South India," *Geophysical Research Abstract*, **7**, 2005.

6. Sivakumar B., "Chaos theory in hydrology: important issues and interpretations," *Journal of Hydrology,* **227,** 2000, pp.1–20.

7. Wilks D. S., "Statistical Methods in Atmospheric Sciences," *Academic Press, USA,* 1995.

8. Perez P., Reyes J., "Prediction of particulate air pollution using neural techniques," *Neural Computing and Application,* **10,** 2001, pp.165-171.

9. Widrow B., Lehr M. A., "30 years of Adaptive Neural Networks: Perceptron, Madaline and Backpropagation," *Proceedings of IEEE,* **78,** 1990, pp.1415-1442.

10. Gardner M. W., Dorling S. R., "Artificial Neural Network (Multilayer Perceptron) - a review of applications in atmospheric sciences," *Atmospheric Environment,* **32,** 1998, pp. 2627-2636.

11. Wilks D. S., "Representing serial correlation of meteorological events and forecasts in dynamic decision – analytic models," *Monthly Weather Review,* **119,** 1991, pp.1640-1662.