Soft Comput (2005) 9: 349–354
DOI 10.1007/s00500-004-0414-3

FOCUS

Sutapa Chaudhuri · Surajit Chattopadhyay

Neuro-computing based short range prediction of some meteorological parameters during the pre-monsoon season

Published online: 29 January 2005

Abstract A Feed forward multi-layered artificial neural network model is designed in this paper to estimate the maximum surface temperature and relative humidity needed for the genesis of severe thunderstorms over Calcutta (22° 32′, 88° 20′). The performance of the model is found to be adroit. It has, thus, been discerned that the neural network technique is of great use in forecasting the occurrence of high frequency small-scale weather systems like Severe Local Storms. Filing up the missing values and extension of time series is observed to be possible with this model. Prediction error is computed and compared for single layer network and one hidden layer neural nets. Result reveals the efficiency of the one hidden layer neural net.

Keywords Single layer network · One hidden layer neural net · Severe thunderstorm · Surface temperature · Relative humidity

1 Introduction

Thunderstorm is a cumulus scale or cloud scale weather phenomenon and is generated due to instabilities created in the meso-scale. These weather events are important because they take heat and moisture near the earth's surface and transport it to the upper levels to maintain the general atmospheric circulation. But these storms occasionally lead to devastating effects when accompanied by large hail, tornado or strong horizontal wind and become the great concern for the atmospheric scientists. Such a catastrophic situation happens over northeastern part (20°N to 24°N latitude, 85°E to 93°E longitude) of India encircling Calcutta. Extensive research has already been carried out on the study of pre-monsoon severe thunderstorms since last nine decades or so. The studies on the phenomena were initially based on the analysis of the conventional surface and upper air parameters before and immediately after the occurrence of the thunderstorms [5, 12, 14]. The analysis of non-conventional parameters, like Satellite and Radar imageries were also done to mark the zone of instability [4, 9, 13,]. These studies were mainly qualitative in nature. It was established quantitatively [2] that minimization of Convective Inhibition Energy (CINE) is important rather than maximization of Convective available Potential Energy (CAPE) in the genesis of pre-monsoon severe thunderstorms.

The Soft Computation techniques are based on the information processing in biological system. The complex biological information processing system enables the human beings to survive by accomplishing tasks like recognition of surrounding, making prediction, planning and acting accordingly. Human type information processing involves both logical and intuitive information processing. Conventional computer systems are good for the former, but their capability for the later is far behind that of human beings. For a computing system to have human like information processing facility, it should be flexible enough to support three features: openness, robustness and real time processing. Openness of a system is its ability to adapt or extend itself on its own to cope with changes encountered in the real world. Robustness of a system means its stability and tolerability when confronted with distorted, incomplete or imprecise information. The real time characteristic implies the ability of the system to react within a reasonable time in response to an event. Information processing systems with all these three characteristics are known as real world computing (RWC) systems. A RWC system should, therefore, be capable of distributed representation of information, massively parallel processing, learning and

S. Chaudhuri (✉) · S. Chattopadhyay
Department of Atmospheric Sciences,
University of Calcutta,
92, APC Road, Calcutta–700 009
India
Tel: 91-033-2424 3385
Fax: 91-033-2350 9957
E-mail: sutapa_chaudhuri@hotmail.com

self-organization in order to achieve flexibility in information processing. Thus, soft computing can be viewed as the key ingredient of RWC systems. Several authors have discussed the potentials of these soft computing techniques in solving real world problems.

Artificial Neural Network (ANN), a component of Soft Computing is particularly useful in such situations, where underlying processes/relationships may display chaotic properties [6, 7, 11]. ANN's do not require any prior knowledge of the system under consideration and are well suited to model dynamic systems on a real-time basis. It is, therefore, possible to set up these systems so that they adapt to the events as they are observed and could benefit real time weather forecasting [11]. Performance of conventional statistical models very often rely on the availability of accurate real time data inputs, the quality of the technological knowledge and mathematical skills used to specify, build and operate the models, and ability of the models to respond in dynamic, unexpected and rapidly changing environment. ANNs on the other hand, offer real prospects for an effective, more flexible, less assumption dependent adaptive methodology well suited for modelling the weather system, which by nature is inherently complex because of non-linearity and chaotic effects [11]. A Multi layer Perceptron model with fuzzy logic was developed to forecast the occurrence of such storms by recognizing the pattern of the surface parameters [3].

The basic requirement for the genesis of thunderstorm is heat and moisture. The parameters, temperature and relative humidity are, thus, chosen for this study. Two different neural networks are designed for each of the parameters. One is a single layer network and the other is a three-layered non-linear Perceptron [16]. In the single layer network the transfer function is considered to be linear and for three-layer network, a sigmoidal function is assigned as the transfer function. The sigmoidal function is used to explore the non-linearity in the data set with most efficient fashion [10, 15]. The transfer function is applied between the input layer and hidden layer as well as between hidden and output layer. This three-layer network, once right network parameters are chosen, represents a non-linear model having immense potential to predict various time series. The hidden layer provides a way to improve the possibilities of finding out correct mapping between input and output. The present study considered the data of pre-monsoon thunderstorms over Calcutta between 1994 and 1998. Thus different time series corresponding to maximum surface temperature and maximum surface relative humidity are fabricated for three months of each year starting from first day of March. The entire dataset is separated into a training set and a test set. Prediction errors are computed from the difference between actual outputs and the outputs generated by employing different methods in the test cases. The results of this study may be useful to predict a day with required temperature and relative humidity, that is, with a possibility of thunderstorm occurrence. The job being basically a reconstruction of time series, some missing data can be retrieved and a series can be extended to predict a trend in the parameter values.

2 The sample data

The present experiment is executed with daily data of maximum temperature and maximum relative humidity collected over Calcutta from first day of March to last day of May during the period from 1994 to 1998. The time series, thus, consists of 460 points tested successfully as meteorologically persistent for the chosen parameters [20]. It is observed from the data set that it is possible to relate the maximum temperature and maximum relative humidity of a given day to the maximum temperature and relative humidity to the sequence of 30 previous days. This was verified after performing several tests on the data set using the back propagation method described in the next section. The situation where an input window of size less than 30 was used, prediction was found to be not good and where the size is greater than 30 we were left with too many free network parameters in order to adjust a model. No statistical way of adjusting the input window could be followed because it requires huge amount of data [1]. The training of the data is done by back propagation method [17]. 30 matrices are framed for each of the two parameters, each with 31 columns and 30 rows. Thus, from the present dataset ($30 \times 5 = 150$) matrices are possible. The 30 of the 31 columns correspond to the measures of the parameter (temperature or relative humidity) for 30 consecutive days and the last column corresponds to the observation on the next day. From each of these matrices the first row is extracted. Thus, 115 rows in total are extracted. Using these rows, (85×31) matrix is used as the training set and (30×31) matrix is used as the test set. The methods, implemented to predict maximum surface temperature and maximum relative humidity on a specific pre-monsoon day '*d*', may be described as;

$$y_d = f_d(x_1, x_2, x_3, \ldots, x_{30}) \tag{1}$$

where $x_1, x_2, x_3, \ldots, x_{30}$ represent the measures of a parameter (maximum surface temperature or maximum surface relative humidity) on 30 consecutive pre-monsoon days. It should be noted that due to change in various co-relational patterns, data outside the pre-monsoon period have not been considered. The form of the function f_d is obtained after adjusting a set of weights (or parameters) defined within the respective method using the training set of data. Within the training matrix, every row is a sample case. The quality of the performance of the procedure is tested using the test set. Percent errors of prediction (PE) are computed as;

$$PE = < |y_{tp} - y_{ta}| > / < |y_{ta}| > \tag{2}$$

where y_{tp} is the predicted value, y_{ta} is the actual value, and $<>$ means average over the 30 test cases.

3 Back propagation method

The Feed forward neural network, also known as Multilayer Perceptron, has become a popular tool for solving complex prediction as well as classification problems. The Back propagation algorithm [17], which is an adaptation of the steepest descent method, opened new avenues for the application of multilayer neural networks for many problems of practical interest.

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

$$w_{k+1} = w_k + \eta d_k \tag{3}$$

where, w_j denotes the weight vector at the j^{th} 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 sequentially performs the adaptation. Here, the positive quantity η is called the learning rate.

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 k^{th} step . Mathematically, d_k is expressed as;

$$d_k = -\nabla E(w_k) \tag{4}$$

The standard learning scheme for the back propagation algorithm in which the weights of the network are updated immediately after the presentation of each pair of input and target patterns is called on-line learning [e.g. 10]. In this learning process, the weight vector w_k contains the weights computed during the k^{th} iteration, and the output error function E is the multivariate function of the weights of the network. Mathematically, this is expressed as

$$E(w_k) = E_p(w_k) \tag{5}$$

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 pairs. Every cycle in which each one of the training patterns is presented to the neural network is called an epoch. The learning continues until E is less than a preset value at the end of an epoch.

4 The single layer network

This method [e.g. 19] assumes that the prediction obtained from a linear combination of the values of previous days is given by

$$y_d = w_{d1}x_1 + w_{d2}x_2 + w_{d3}x_2 + \cdots + w_{d30}x_{30} + w_{d0} \tag{6}$$

An interactive method is been applied according to (3) and (4), to update the weights.

5 One hidden layer neural network

In recent studies one-hidden-layer neural network has been established as a strong prediction tool [e.g., 8, 15, 18, 19]. In equation (1), if the function f_d is nonlinear, then a 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,...,n) in such a way that

$$z_{dk} = f(w_{dk1}x_1 + \cdots + w_{dk30}x_{30} + w_{dk0}) \tag{7}$$

and

$$y_d = f(v_{d1}z_1 + \cdots + v_{dn}z_{dn} + v_{d0}) \tag{8}$$

The function 'f' is defined as:

$$f(x) = (1 + exp(-x))^{-1} \tag{9}$$

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

Since the number of adjustable parameters in a one-hidden-layer feed forward neural network with n_i input units, n_0 output units, and n_h hidden units is [n_0 + n_h (n_i+ n_0 + 1)], for $n_i = 30$, $n_0 = 1$ and with 85 training cases, it is not possible to use an n_h greater than 2. Best result is found to occur for 2 hidden nodes.

6 Result and discussion

Training of the neural networks and then testing it with test data lead to the following findings;

- Prediction for maximum surface temperature and maximum relative humidity can be done by one-hidden-layer neural network more efficiently than by the single layer network (Fig. 1a, Fig. 2a). The schematic of linear relationships between the actual values and the predicted values through one-hidden-layer non-linear Perceptron model is displayed (Fig. 1b, Fig. 2b). Percent error of prediction, as described in Eq (2), by three-layer ANN is decreased by 4% for maximum temperature and by 6% for maximum relative humidity with respect to single layer network (Fig. 3, Fig. 4). A comparison between the prediction errors (Fig. 5) makes it apparent that an ANN is more efficient to predict the maximum relative humidity than maximum surface temperature.
- Forecasting yield, which is a measure of success of a model, is defined as the percentage of days in the testing set of data for which the model has performed a successful prediction. In the test set of maximum surface temperature, maximum temperature has been predicted with ± 4°C confidence range in 69% cases and with ± 6°C confidence range in 82% cases. For thunderstorm days, forecasting yield is 56% with confidence range ± 2°C. A section of actual and predicted values of maximum surface temperature are displayed in the figures. The high forecasting yield

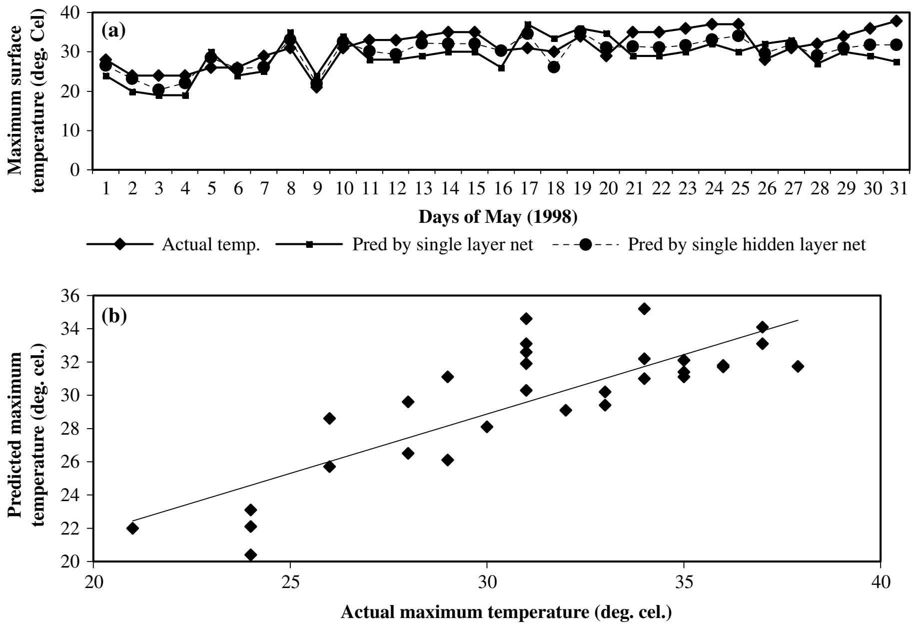

Fig. 1 a Schematic showing the comparison between outputs from the single layer and single hidden layer neural net and the actual surface temperature in the days of May 1998. **b** A scatter diagram showing the linear relationship between the actual values of maximum temperature and predicted values through one-hidden-layer ANN

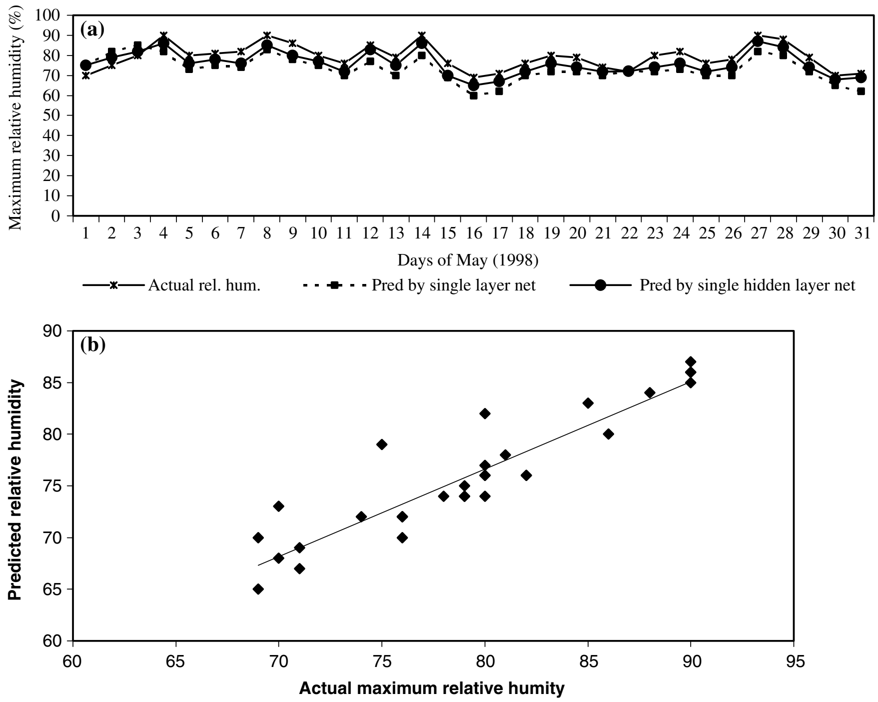

Fig. 2 a Schematic showing the comparison between the outputs from single layer and single hidden layer neural net with the actual surface relative humidity in the days of May (1998). **b** A Scatter diagram showing the linear relationship between the actual values of maximum relative humidity and that predicted by 3 layered ANN

with so small confidence range on the thunderstorm days implies the suitability of one- hidden-layer neural network as a forecasting tool.

- For maximum surface relative humidity the forecasting yield is 60% with ±8% confidence range and 80% with ± 10% confidence range. On the thunderstorm days, this forecasting yield is 64% with confidence range ± 9%. The high forecasting yield with so small confidence range on the thunderstorm days implies the suitability of one-hidden-layer neural network as a forecasting tool.
- Time series of maximum surface relative humidity and maximum surface temperature on pre-monsoon thunderstorm days can be extended using these networks and some missing data may be generated from it. Thus, a time series reconstruction is possible with this neural network model.
- The forecasting from single-layer-neural net is compared with the forecasting from the second order autoregressive process [AR(2)]. It is found that forecasting yield is better in case of single-layer-neural net than in AR(2). From the feedback of AR(2) it is

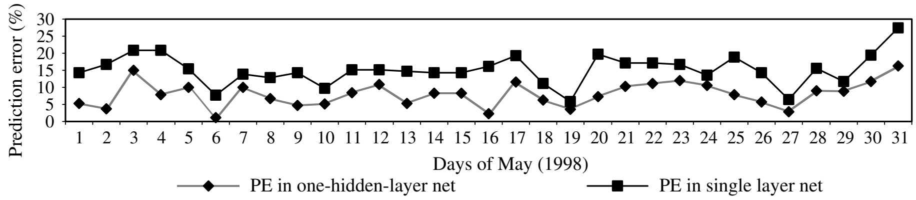

Fig. 3 Comparison between the prediction errors in case of maximum surface temperature by single layer and single hidden layer neural nets

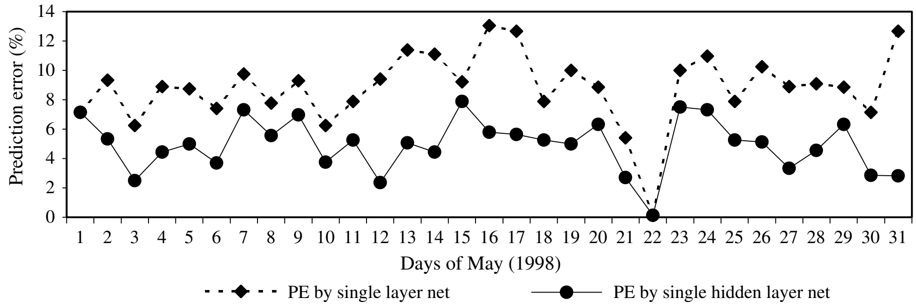

Fig. 4 Comparison between prediction errors in case of maximum relative humidity by single layer and single hidden layer neural nets

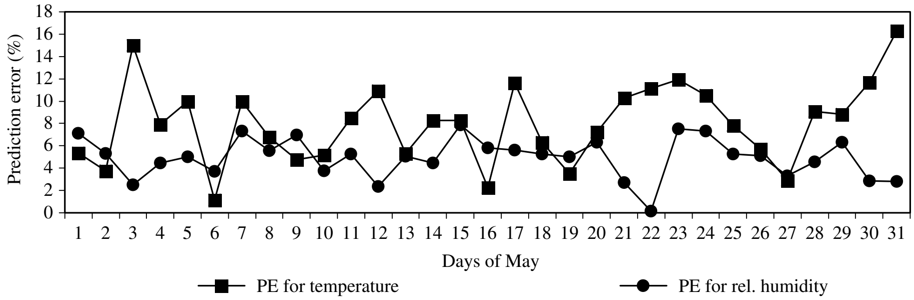

Fig. 5 Schematic showing the comparison between the prediction errors by single hidden layer network in case of relative humidity and temperature

apparent that, for maximum surface temperature, forecasting yield is 52% with confidence range ±4°C and 70% with confidence range ±6°C. For maximum relative humidity, forecasting yield is 45% with confidence range ±8% and 68% with confidence range ±10%.

7 Conclusion

From the above study it can be concluded that one-hidden-layer neural network is an efficient forecasting tool by which an estimation of maximum surface temperature and maximum relative humidity can be obtained. This estimation can help in predicting a probable thunderstorm day with one day or 24 hrs in advance. Moreover, the utility of this model lies in the fact that the time series of maximum surface temperature and relative humidity can be reconstructed which may be helpful in understanding the approximate trend of the parameters and filling up some missing points in the data archive.

References

1. Abarbanel HD, Brown I, Sidorowich RJJ, Tsimring LS (1993) The analysis of observed chaotic data in physical systems. Reviews of Modern Physics 65: 1331–1392
2. Chaudhuri S, Chattopadhyay S (2001) Measure of CINE–A relevant parameter for forecasting pre-monsoon thunderstorms over GWB. Mausam 52:679–684
3. Chaudhuri S, Chattopadhyay S (2002) Multilayer Perceptron model in pattern recognition of surface parameters during pre-monsoon thunderstorm. Mausam 53:417–424
4. De AC, Sen SN (1961) A radar study of pre-monsoon thunderstorm over GWB. Indian J. Met. Geophys. 12:57–60
5. Desai BN, RaoYP (1954) On the cold pools and their role in the development of Nor'westers over West Bengal and East Pakistan. Indian J. Met. Geophys. 5:243–248
6. Gardner MW, Dorling SR (1998) Artificial Neural Network (Multilayer Perceptron)- a review of applications in atmospheric sciences. Atmospheric Environment 32:2627–2636
7. Hsieh WW, Tang T (1998) Applying Neural Network Models to Prediction and Data Analysis in Meteorology and Oceanography. Bulletin of the American Meteorological Society 79:1855–1869
8. Hush DR, Horne BG (1993) Progress in supervised neural networks:What's new since Lippmann. IEEE Signal Processing Magazine 10:8–39
9. Kalsi SR, Bhatia RC (1992) Satellite observations of development of thunderstorm complexes in weakly forced environments. Vayu Mandal 22:65–76
10. Kamarthi SV, Pittner S (1999) Accelerating neural network training using weight extrapolations. Neural Networks12:1285–1299
11. Maqsood I, Muhammad RK, Abraham A (2002) Neurocomputing Based_Canadian Weather Analysis. Computational Intelligence and Applications. Dynamic Publishers Inc., USA :39–44
12. Mukherjee AK, Chaudhuri AK (1979) Excessive overshooting of Cb. Indian J. Met. Geophys. 30:485–492

13. Mull S, Mitra H, Kulshrestha SM (1963) Tropical thunderstorms and radar echo. Indian J. Met. Geophys. 14:23–26
14. Normand CWB (1921) Wet bulb temperature and thermodynamics of air. Indian Meteorological Memoirs 23:Part-I
15. Perez P, Reyes J (2001) Prediction of particulate air pollution using neural techniques. Neural Computing and Application 10:165–171
16. Rojas R (1996) Neural Networks:A systematic introduction. Springer-Verlag, Berlin
17. Rumelhart J, Hinton GE, Williams RJ (1986) Learning Internal Representations by Error Propagation. Parallel Distributed Processing. MIT Press, Cambridge, London:318–364
18. Weigend A, Gershenfeld N (1994) Time Series Prediction, Santa Fe Series, Addison Wesley, New York.
19. Widrow B, Winter RG, Baxter R (1987) Learning phenomena in layered neural networks. Proc. 1st IEEE Int. Conf. On Neural Networks 2:411–429
20. Wilks DS (1995) Statistical methods in Atmospheric Sciences. Academic Press.