

Design of Neural Network Models for Daily Rainfall Prediction

N.A.Charaniya

Associate Professor, Department of Electronics and Telecommunication, B.N.College of Engg, Pusad.

S.V.Dudul, PhD.

Professor and Head Applied Electronics
Department S.G.B. Amravati University.

ABSTRACT

Rainfall is a random process and prediction of rainfall requires consistent as well as relevant information of meteorological and environmental data. In this paper, two different artificial neural networks models are proposed for consecutive daily rainfall prediction on the basis of the preceding events of rainfall data. Model designed for rainfall forecast is based on the pattern recognition methodology. In this method relevant spatial and temporal feature of rainfall series in past are extracted. These features are then utilized to predict the rainfall in future. Time lag delay neural network has capability to learn from the past event and predict the next value. Rainfall prediction is done on basis of rainfall on previous day to rainfall for the preceding six days. The proposed network is capable of forecasting daily rainfall one day in advance with accuracy of $R^2 = 0.96$ and $NMSE = 0.0005$.

General Terms

Artificial Neural network, Prediction

Keywords

Artificial neural network, Time lag neural network, Daily Rainfall, Prediction

1. INTRODUCTION

In India rainfall information is vital for crop production plan, water management and all activity plans in the nature. The incident of extended dry period or heavy rain at the critical stages of the crop growth and development may lead to note worthy reduction in the crop yield and hence affect the economy of the country. India is an agricultural country and its economy is largely based upon agricultural product. Thus rainfall forecast becomes a significant factor in agricultural countries like India [1]. A wide range of rainfall forecast methods are employed in weather forecasting at regional and national levels. Rainfall is a random process and therefore prediction of rainfall is a very difficult and cumbersome. Accurate and timely prediction of rainfall is a foremost challenge for the research community [2]. Rainfall prediction modeling involves a combination of computer models, inspection and information of trend and pattern. Using these methods, reasonably accurate forecasts can be made up. Several recent research studies have developed rainfall prediction using different weather and climate forecasting methods [3, 4, 5].

In India during monsoon season, the rainfall depends upon various parameters such as pressure level in the given region as compared to surrounding region, humidity, vapor pressure etc. If we look at the trends and the pattern of rainfall in the central part of India we find that the rainfall on one day has its effect on the rainfall for the next few days. Rainfall appears for number of days together and then disappears for number

of days together. There is a correlation between the rainfall on a given day with the rainfall on previous days.

There are two possible methods for rainfall prediction. The first one is based on the study of the rainfall processes and its dependence on other meteorological parameter such as pressure, humidity, vapor pressure, temperature etc. However, this method is complex and non- feasible because rainfall is the result of a number of complex atmospheric parameter which vary both in space and time and this parameter are limited in both the spatial and temporal dimensions [6]. Another method to forecast rainfall is based on the pattern recognition methodology. In these method relevant spatial and temporal features of rainfall series in past are extracted. These features are then utilized to predict the rainfall in future. In this technique only rainfall data is considered and therefore this approach is considered appropriate for short-duration rainfall prediction.

If rainfall pattern can be estimate then it can help in predicting the rainfall for the next day [7, 8]. Therefore in this paper an attempt has been made to estimate rainfall pattern using neural network. Neural network is the best tool for pattern recognition. The rainfall prediction model developed in this paper is based on the use of Artificial Neural Networks to understand the rainfall pattern and also to predict the future rainfall.

ANNs, which replicate the parallel distributed processing of the human nervous system, have proven to be very powerful tool in dealing with complicated problems, such as pattern recognition and function approximation. Hornik et al. [9] has shown that an ANN with adequate complexity is capable of approximating any function to a reasonable accuracy. In addition, ANNs are computationally robust, in the way that they have the ability to learn and give correct output even if the input contains noise.

There have been a number of reported studies that have used ANNs to solve problems in hydrology. For example, French et al. [10] used an ANN to forecast rainfall for a catchment with artificial rainfall inputs, while Hsu et al. [11] applied an ANN to model the rainfall-runoff process.

Rainfall process is nonlinear as well as random process and therefore most previous studies have used a Multilayer perceptron (MLP) or Generalised feedforward neural network (GFNN), without considering the other types of ANNs. Although the GFNN is a powerful nonlinear model, there are many other types of ANNs which may have been more appropriate to the problem. ANNs appropriate for rainfall forecasting include the time-delay neural networks (TDNN).

In this paper two different types of neural network models (i.e., GFNN and TDNN) are designed for forecasting rainfall one day ahead over Nagpur region in the central part of India. The ANNs is applied to forecast rainfall based on previously observed temporal rainfall patterns for a given region.

2. DATA

Data used in the present study is for Nagpur station. Data for Nagpur station is obtained from Indian meteorological department Pune. Rainfall data for 30 years (1977-2006) is used for simulation purpose. Linear transformation is used to normalize rainfall series data. The outputs of the normalization function are real numbers between 0 and 1. The equation can be described as:

$$D_{\text{norm}} = \frac{D_0 - D_{\min}}{D_{\max} - D_{\min}}$$

Where D_0 is for the observed data, D_{\max} and D_{\min} are maximum and minimum of the observed data, respectively.

3. RAINFALL FORECASTING MODEL

The scope of this paper is to develop a model to predict rainfall one day ahead. As it is known that rainfall is random process it has been modeled as a process where the probability of a k^{th} output depends upon the finite previous output in sequence. This class of random process is called as Mark off process. The rainfall value at a given location and time is a function of a finite set of previous occurrence. With this assumption, a model structure can be expressed as

$$R(t+1) = F(R(t), R(t-1), R(t-2), \dots, R(t-k+1)) \text{ ----(1)}$$

Where

$R(t+1)$ represent the next day rainfall

$R(t)$ represents rainfall today,

$R(t-1)$ represents rainfall one day before and so on

$F(\cdot)$ is a nonlinear mapping function, which will be approximated using an ANN,

k is the (unknown to be solved) number of days in past usually, k refers to the lag of the network; if $k = 1$, the rainfall at the next time-step is related only to the present rainfall, thus giving a lag-1 network.

4. Artificial Neural NETWORKS MODEL

There are a number of alternative ANNs which can be used to represent the Markovian model expressed in equation (1), where the continuous process of rainfall is being represented by a discrete process. Based on a review of these alternatives, two suitable ANN model were identified and adopted in this study for comparison. They are:

1. Generalized feedforward network (GFNN),
2. Focused time lag neural network (FTLNN).

4.1 Generalized feedforward network

In a neural network if the interconnection is restricted to feed forwarding activations (neither feedback nor self connections), the neural network is defined as feedforward. Feedforward networks are instantaneous mappers; i.e. the output is valid immediately after the presentation of an input. A special class of feedforward networks is the layered class, which is called the multilayer perceptron (MLP). This name comes from the fact that Rosenblatt's network, which was called the perceptron, consisted of a single layer of nonlinear PEs without feedback connections.

Multilayer perceptrons shown in Fig. 1.1 has an input layer of source nodes and an output layer of neurons (i.e., computation nodes); these two layers connect the network to the outside world. Layers that receive the input from the external world are called the input layers, layers in contact with the outside world are called output layers. In addition to these two layers, the multilayer perceptron usually has one or more layers of hidden neurons. The layers without direct access to the external world, i.e. not connected to the input or output, are called hidden layers. Which are so called because these neurons are not directly accessible. The hidden neurons extract important features contained in the input data. The training of an MLP is usually accomplished by using a back propagation (BP) algorithm [12, 13].

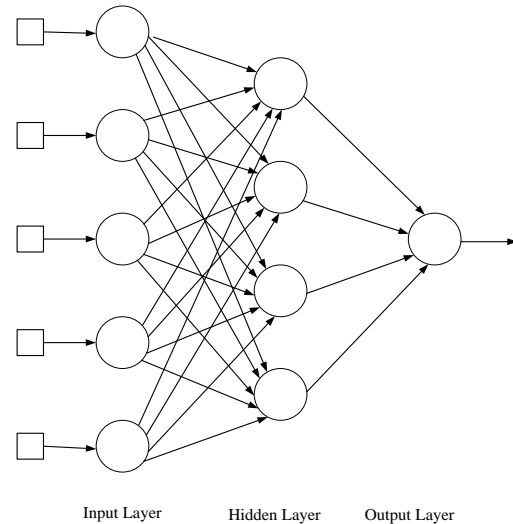


Fig 1: A simple three-layer feedforward neural network

4.2 Focused time lag neural network

To extract temporal information, time is an essential dimension of learning. First the temporal information from the rainfall series is extracted and then this information can be used to predict the rainfall one unit ahead in time. Time parameter may be incorporate into the design of a neural network implicitly or explicitly. A straightforward method of implicit representation of time is to add a short-term memory structure in the input layer of a static neural network. This memory structure is used to store the past information, which can be used to analyze the data in a more efficient way. The element in the memory structure can be varied as per requirement. The resulting network is called as focused time lagged feedforward network (TLFN).

The short-term memory structure may be implemented in different form one of Forms is Tapped-Delay-Line (TDL) Memory. This is the most commonly used form of short-term memory. It consists of p unit delays with $(p + 1)$ terminals. This may be viewed as a single input-multiple output network. Fig.3 shows a focused TLFN network using the combination of a TDL memory and multilayer perceptron. In Figs. 3 the unit-delay is denoted by z^{-1} . The memory depth of a TDL memory is fixed at p , and its memory resolution is fixed at unity, giving a depth resolution constant of p .

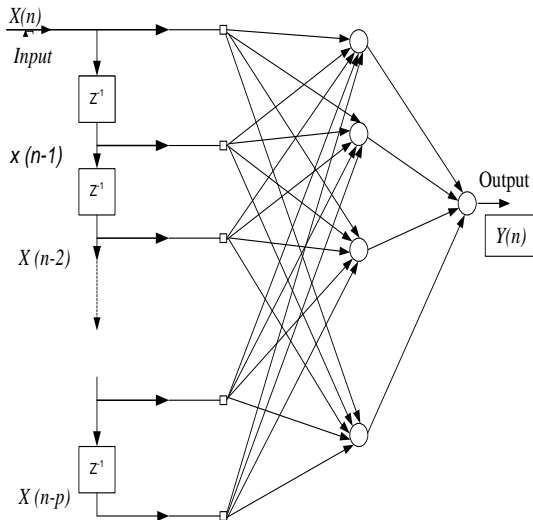


Fig 2: A focused Time Lag Feedforward neural network

5. WORK DONE

Two different models were developed for prediction of rainfall one day ahead. In the first model a generalized feedforward neural network (1-42-1) (GFNN) with one hidden layer was designed. The backpropagation (BP) algorithm used for learning was Levenberg-Marquardt (LM). The Levenberg-Marquardt (LM) algorithm is one of the most appropriate higher-order adaptive algorithms known for minimizing the mean square error of a neural network. A Tanh activation function was adopted for the hidden nodes, whereas a linear activation function was used for the output nodes. The use of a Tanh function was to enable nonlinearity of the network. The Tanh function, however, was not used for the output nodes because it would force the output to be bounded between 0.0 and 1.0

Input data applied to this network was the rainfall time series with delay. A delay means data for previous day. Number of delays for which the model was tested was taken from 1 to 7. It means data applied to the network was for last seven days to last one day. Network model was also tested for different number of neurons in the hidden layer. Optimum numbers of neurons in the hidden layer were found to be 42. At this number the mean square error was found to be minimum.

Second network model developed was a Focused time lagged feedforward neural network (1-23-1) with one hidden layer and learning algorithm LM. This model was test again for different number of lag and different number of neurons at hidden layer. Number of neurons in the hidden layer at which the model performance is optimum was found to be 23. This is less as compare to that required in GFNN network model.

6. RESULT AND DISCUSSION

Various network models have been tried to predict the rainfall one day ahead. Rainfall prediction was done based on the temporal pattern learning using neural network. Two neural network models are found to be best for rainfall prediction. Both the models have been tested with different lags and different number of neurons in the hidden layer. From the result shown in the table 1.1, it is observed that when the number of lag are three for GFFN model and two for FTLNN the normalized mean square error is low as well the

correlation between the actual rainfall and predicted rainfall series is high. The result shows that the rainfall pattern learning is better when taken for short duration. As the duration increases accuracy of prediction decreases because the rainfall on a given day is dependent on previous day or two. Also it is found that FTLNN model is better than GFFN network for learning temporal pattern.

From the result mentioned in the Table 1 another remarkable conclusion can be made that the networks with a lesser lag had a better performance than those of a higher lag. This means that the rainfall series has a short term dependent structure. Also it can be concluded that as the number of lags are increased the network become more complex and therefore learning becomes more difficult.

Another Simulation was done for prediction of rainfall number of days ahead from two days to seven days. Focused TLFN model was considered for this simulation. Input data given to the network was for last seven days and the rainfall prediction was done for two days ahead to seven days ahead. The result are as shown in table 1.2 From the result it is seen that if the prediction is done more in advance the accuracy of prediction decreases or the error of prediction increase. Therefore it is concluded that the model designed is capable of short term rainfall prediction with better accuracy.

A plot of normalized actual and predicted rainfall using both the model is shown in the fig 3. And scatter plot between actual and predicted rainfall with residual for TFLN model is shown in the fig 4.

7. CONCLUSION

Rainfall prediction using ANNs has been discussed in the paper. Two different ANNs models found suitable for this task were designed and compared.

Following points were observed.

- For each type of network, there existed an optimal solution, which was a function of the number of hidden nodes, number of neurons in the hidden node and the lag of the network.
- Networks with lower lag tends to outperform the ones with higher lag. This indicates that the rainfall time series have very short term memory characteristics.

In future ANN Models designed can be used for rainfall prediction at different location by properly training the network to estimate the temporal pattern of rainfall series at given location. Also different ANN models can be designed for different geographical area to increase the accuracy of prediction.

Table 1. Performance parameter of the network model designed

Number of lag	Normalized mean square error		Correlation coefficient	
	GFNN	FTLNN	GFFN	FTLNN
2	0.00130	0.00091	0.925	0.95
3	0.00111	0.00054	0.957	0.969
4	0.00112	0.00053	0.963	0.963

5	0.00438	0.00057	0.946	0.964
6	0.00281	0.00060	0.957	0.958
7	0.00260	0.00068	0.94	0.951

Table 2. performance parameter with N number of days ahead prediction

Number of days ahead	Normalized mean square error	Correlation coefficient
2	0.0025	0.87
3	0.0033	0.82
4	0.0046	0.78
5	0.0108	0.75
6	0.0135	0.68
7	0.0150	0.62

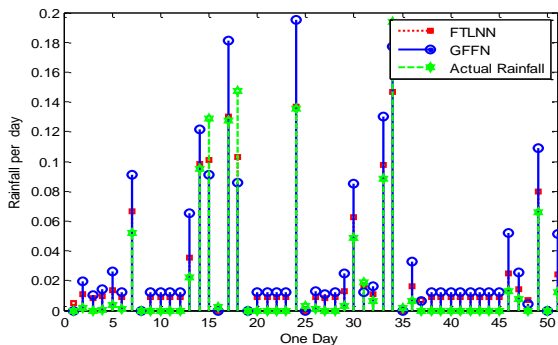


Fig 3: Actual and predicted rainfall using different neural network model .

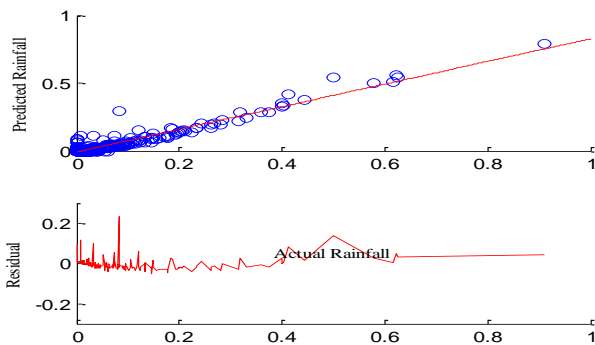


Fig 4: scatter plot between Actual and one day ahead Predicted rainfall using TLFN neural network model

8. REFERENCES

[1] Sulochana Gadgil and Siddhartha Gadgil, The Indian Monsoon, GDP and agriculture, Economic and Political Weekly, XLI, pp. 4887–4895, 2006.
[2] Rajeevan M., “Prediction of Indian summer monsoon: Status, problems and prospects”. Curr. Sci., 2001, 81, pp.1451–1457.

[3] Thapliyal, V. and Rajeevan, M., “Monsoon prediction”. Encyclopedia of Atmospheric Sciences (ed. Holton, J.), Academic Press, New York, 2003, pp. 1391–1400.
[4] Sikka, D.R., 1980, Some aspects of the large scale fluctuations of summer monsoon rainfall over India in relation to fluctuations in the planetary and regional scale circulation parameters, Proc. Ind.Acad.Sciences (Earth and Planet Sci), 89, 179-195.
[5] F. W. Zwiers, J. V. Storch, “On the role of statistics in climate research,” Int. J. Climatology, 2004 Vol.24, pp. 665-680.
[6] Kin C. Luk, J. E. Ball AND A. Sharma, 2001, An Application of Artificial Neural Networks for Rainfall Forecasting, Mathematical and Computer Modelling 33 pp. 883-699.
[7] Poggio,T., and F. Girosi, 1990, “Networks for approximation and learning,” Proc. IEEE, vol. 78, pp. 1481–1497.
[8] Sven F. Crone: A Business Forecasting Competition Approach to Modeling Artificial Neural Networks for Time Series Prediction. IC-AI 2004, pp. 207-213.
[9] K. Hornik, M. Stinchcombe and H. White, Multilayer feedforward networks are universal approximators, Neural Networks 1989 2, pp.359-366.
[10] M.French, W. Krajewski and R.R. Cuykendall, Rainfall forecasting in Space and time using a neural network, Journal of Hydrology 1992 137, 1-31.
[11] K.L. Hsu, V. Gupta and S. Soroshian, Artificial neural network modeling of the rainfall-runoff process, Water Resources Research 1995 31 (10), pp.2517-2530.
[12] Werbos, P. J., 1990, “Backpropagation through time: What it does and how to do it,” Proc. IEEE, vol. 78, pp. 1550–1560.
[13] Rumelhart, D. E., Hinton G. E., and Williams R. J., 1986, “Learning internal representations by error propagation,” in D. E. Rumelhart and J. L. McClelland, eds. (Cambridge, MA: MIT Press), vol. 1, Chapter 8.
[14] Jiansheng. Wu, Long. Jin, “Forecast Research and Applying of BP Neural Network Based on Genetic Algorithms,” Mathematics in Practice and Theory, 2005, vol. 35(1), pp.83-88.
[15] D. Silverman, J. A. Dracup, “Artificial Neural networks and long-lead precipitation prediction in California,” Journal of applied meteorology, 2000 vol.39, pp. 57-66.
[16] G. J. Bowden, G. C. Dandy and H. R. Maier. “Input determination for neural network models in water resources applications. Part 1-background and methodology,” Journal of Hydrology, 2005 vol. 301(1-4), pp. 75-92.
[17] Wan, E. A., 1994, “Time series prediction by using a connectionist network with internal delay lines,” in A. S.Weigend and N. A. Gershenfeld, eds., Time Series Prediction: Forecasting the Future and Understanding the Past (Reading, MA: Addison-Wesley), pp. 195–217.
[18] Bogardi I, Matyasovszky I, Bardossy A, Duckstein L 1993 Application of a space-time stochastic model for daily precipitation using atmospheric circulation patterns. J Geophys Res 98(D9), pp. 1653–166

- [19] Rajeevan M, Bhate J, Kale JD, Lal B 2005 Development of a high resolution daily gridded rainfall data for the Indian region: analysis of break and active monsoon spells. India Meteorological Department
- [20] R.Dean, H. Andrew and B. Fiedler, "Forecasting warm season burn-off low clouds at the San Francisco international airport using linear regression and a neural network ," *Apply Meteor*, 2002 Vol. 41(6), PP. 629-639.