*International Journal of Advances in Computer Science & Its Applications*
*Volume 6 : Issue 1 [ISSN 2250-3765]*
*Publication Date : 18 April, 2016*

# Rainfall Forecast of a Synoptic Station using Artificial Neural Network

Saeid Fazli, Hasan Almasi

***Abstract*— In this paper, we have utilized ANN (Artificial Neural Network) modeling for forecasting rainfall in Babolsar synoptic station in Iran. To achieve such a model, we have used daily rainfall data from 1978 to 2007 for the synoptic station. A Time Delay Series Neural Network is used in this work. Two hidden layers are considered for the neural network. Inputs are daily rainfall with variable lag to achieve optimal prediction. Using this method as a black box model, we have realized the hidden dynamics of rainfall through the past information of the system. Root Mean Square Error (RMSE) and Correlation Coefficient (r2) are evaluated for comparison purposes. Optimum delay days are calculated for rain forecasting which can be used in climatology application.**

***Keywords*—Daily Rainfall Forecasting, Back Propagation Neural Network, Time Delay Series, optimum lag.**

## I. Introduction

The rainfall as meteorological parameters is very complex nonlinear phenomena and varies along with time and place. Nevertheless, literatures show that rainfall is predictable [1], [2].

The topography in the plains and the coastal location are all Babolsar and the rippling effects not seen in topographic relief. Babolsar was almost flat with gentle slopes of the Alborz Mountains to the Caspian Sea runs. The height of seven feet above the Caspian sea and 21 meters below sea level is free. Babolsar the city level has been covered plains and coastal areas and other lands and other lands Bablrod sides of the river city with a slope of 1 to 3 percent. Daily rainfall data has collected from Babolsar's synoptic weather station. Babolsar's synoptic weather station is located in the north-east of Iran at 36o43' Northern longitude, 59o39'Eastern latitude and 21 meter elevation.

Saeid Fazli
Research Institute of
Modern Biological Techniques
University of Zanjan, Zanjan,Iran

Hasan Almasi
Electrical
Engineering Dept.
University of Zanjan, Zanjan, Iran

In Babolsar's special geographical situation, interfacing different air masses make it a region with a special continental climate. Babolsar temperate and humid climate of the rain successively and the humidity is high and the amount of rainfall during the year is between 800 to 900 mm [5]. The main objective of the research is to develop an artificial neural network for the purpose of daily rainfall forecasting. To achieve this objective, we used the daily precipitation data from 1978 to 2007 for Babolsar synoptic station. Finally, rainfall was predicted using Time Delay Series Network that is a suitable type of ANNs for meteorological predictions.

## II. Artificial Neural Networks (ANNs)

### A. *Structure for Back Propagation Neural Network*

Fig.1 depicts an example of back propagation neural network. A neural network uses a number of layers, units per layer, network inputs, and network outputs. This network has four units in the first layer (layer A) and three units in the second layer (layer B), called hidden layers. This network has one unit in the third layer (layer C) called the output layer. Finally, this network has four network inputs and one network output. Some texts consider the network inputs to be an additional layer, the input layer, but since the network inputs do not implement any of the functionality of a unit, the network inputs will not be considered a layer in this discussion.

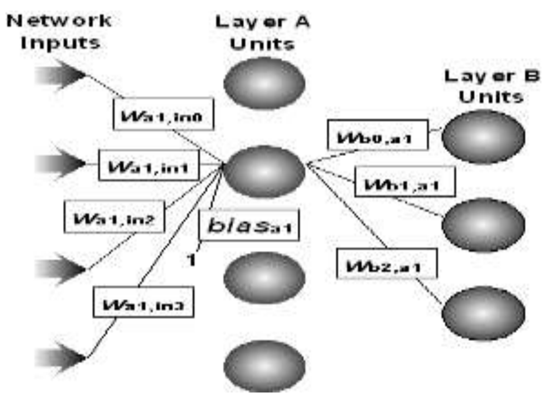



**Figure 1** A three-layer back propagation neural network.

International Journal of Advances in Computer Science & Its Applications
Volume 6 : Issue 1 [ISSN 2250-3765]
Publication Date : 18 April, 2016

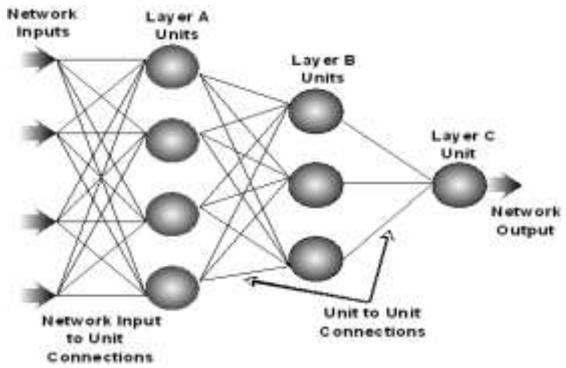



**Figure 2** Unit with its weights and bias.

If a unit is in the first layer, it has the same number of inputs as there are network inputs; if a unit is in succeeding layers, it has the same number of inputs as the number of units in the preceding layer. Each network-input-to-unit and unit-to-unit connection (the lines in Fig.1) is modified by a weight. In addition, each unit has an extra input that is assumed to have a constant value of one. The weight that modifies this extra input is called the bias. All data propagate along the connections in the direction from the network inputs to the network outputs, hence the term feed-forward. Fig.2 shows an example unit with its weights and bias and with all other network connections omitted for clarity.

This relation is explained mathematically as described in (1)

$$y = g\left[\left[\sum_{j=1}^{n} w_{k,j} f\left(\sum_{i=1}^{n}(w_{i,j}+b)\right)+b\right]\right] \qquad (1)$$

## B. *Back propagation Training*

To make meaningful forecasts, the neural network has to be trained on an appropriate data series. Examples in the form of <input, output> pairs are extracted from the data series, where input and output are vectors equal in size to the number of network inputs and outputs, respectively. Then, for every example, back propagation training consists of three steps:

1. Present an example's input vector to the network inputs and run the network: compute activation functions sequentially forward from the first hidden layer to the output layer (referencing Fig.1, from layer A to layer C).

2. Compute the difference between the desired output for that example, output, and the actual network output (output of unit(s) in the output layer). Propagate the error sequentially backward from the output layer to the first hidden layer (referencing Fig.1, from layer C to layer A) [6].

3. For every connection, change the weight modifying that connection in proportion to the error.

When these three steps have been performed for every example from the data series, one epoch has occurred. Training usually lasts thousands of epochs, possibly until a predetermined maximum number of epochs (epochs limit) is reached or the network output error (error limit) falls below an acceptable threshold. Training can be time-consuming, depending on the network size, number of examples, epochs limit, and error limit.

## C. *Time series prediction*

Time series forecasting, or time series prediction, takes an existing series of data $x_{t-n}, \ldots, x_{t-2}, x_{t-1}, x_t$ and forecasts the $x_{t+1}, x_{t+2}, \ldots$ data values. The goal is to observe or model the existing data series to enable future unknown data values to be forecasted accurately. Examples of data series include financial data series (stocks, indices, rain, rates, etc.), physically observed data series (sunspots, weather, etc.), and mathematical data series (Fib**ona**cci sequence, integrals of differential equations, etc.). The phrase "time series" generically refers to any data series, whether or not the data are dependent on a certain time increment [3].

## D. *Data Parsing*

There are two types of data on which a neural network can be trained: time series data and classification data. Examples of time series data were given in Section introduction.

To create the examples in the form of <input, output> pairs referenced in Section 1.2.2, first FORECASTER parses the time series data from the file into a one-dimensional array. Second, any preprocessing the user has specified in the Neural Network Wizard is applied. Third, as shown in Fig.3, a moving window is used to create the examples, in this case for a network with four inputs and one output (The window length is dependent on the selected delay).

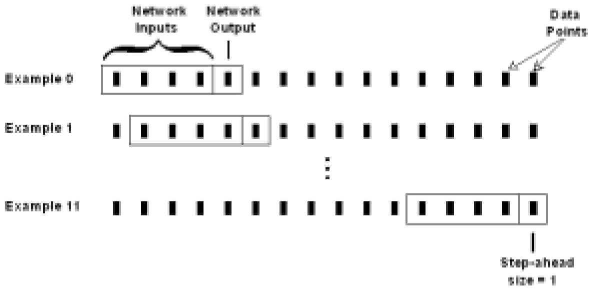
In Fig.3 (a), the network will be trained to forecast one step ahead (a step-ahead size of one). When a forecast is made, the user can specify a forecast horizon, which is the number of data points forecasted, greater than or equal to one. In this case, a forecast horizon greater than one is called iterative forecasting-the network forecasts one step ahead, and then uses that forecast to forecast another step ahead, and so on. Fig.4 shows iterative forecasting. Note that the process shown in Fig.4 can be continued indefinitely.

In Fig.3 (b), the network will be trained to forecast four steps ahead. If the network is trained to forecast n steps ahead, where n >1, then when a forecast is made, the user can only make one forecast n steps ahead. This is called direct forecasting and is shown in Fig.4 for four steps ahead. Note that the number of examples is reduced by n-1.

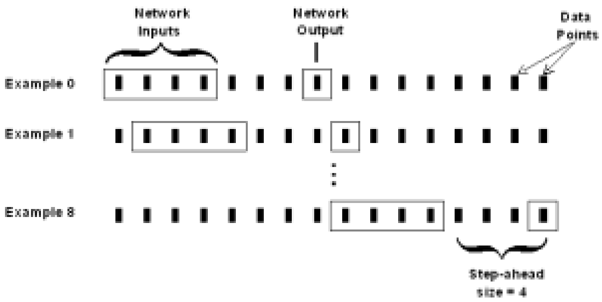
In Fig.4 and Fig.5, the last data points, saved for forecasting to seed the network inputs, are the last data points within

International Journal of Advances in Computer Science & Its Applications
Volume 6 : Issue 1 [ISSN 2250-3765]
Publication Date : 18 April, 2016

either the training set or validation set, depending on which set is "closer" to the end of the data series. By seeding the network inputs with the last data points, the network is making out-of-sample forecasts. The training set and validation set are considered in-sample, which means the network was trained on them. The training set is clearly in-sample. The validation set is in-sample because the validation error determines when training is stopped [3].

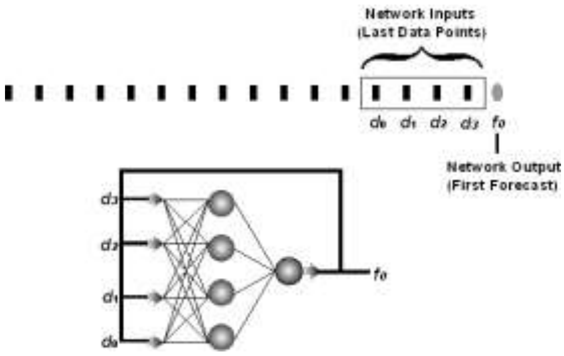



(a)

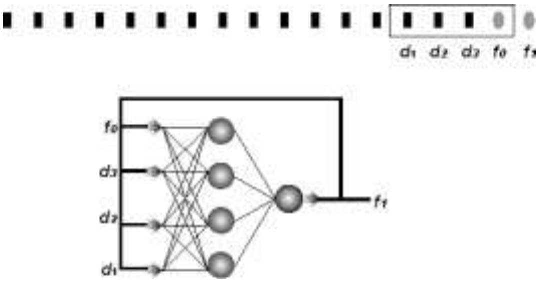



(b)

**Figure 3** Creating examples with a moving window for a network with four inputs and one output. In (a), the step-ahead size is 1; in (b), the step-ahead size is 4.




(a)



(b)

**Figure 4** Iteratively forecasting (a) the first forecast and (b) the second forecast. This figure corresponds to Fig.3 (a).

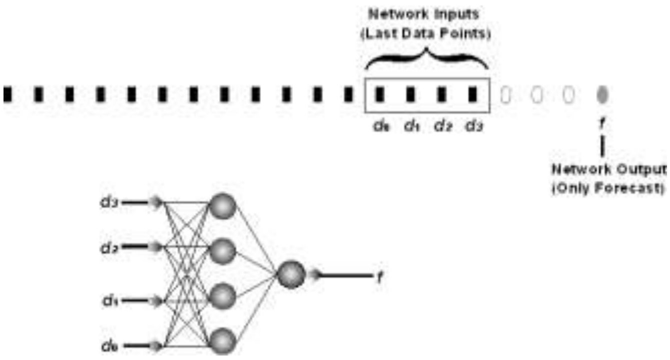



**Figure 5** Directly forecasting the only possible data point four steps ahead. This figure corresponds to Fig.3 (b).

## III. Previous work

Researchers have applied different methods for rainfall forcasting. Some of them are studied in this work. Niravesh Srikalra and Chularat Tanprasert use neural networks with online data collection. Ajith Abraham, Dan Steinberg and Ninan Sajeeth Philip work on the rainfall forecasting using soft computing models and multivariate adaptive regression splines[7].

## IV. Results and Discussions

Hurst exponent of daily data in our work was estimated 0.95. This value implies that the daily set of given time series is predictable. Then, using MATLAB software, several different structures of BP were designed with different numbers of neurons in the input and hidden layers. Each structure is introduced in the form of M (ijk), in which the indices i, j, and k stand for number of neurons in the input layer, the hidden layer, and the output layer, respectively. It should also be noted that the number of epochs was selected as 1000 and η was held at constant value of 0.1. The selected activation functions in hidden and output layer of the networks were sigmoid and linear, respectively. Among different structures, we found GS5-5-1 and GS8-5-1 and GS20-6-1 for minimum

International Journal of Advances in Computer Science & Its Applications
Volume 6 : Issue 1 [ISSN 2250-3765]
Publication Date : 18 April, 2016

RNSE and maximum r2 that renderer of optimum value lag for forecasting rain.

We used in this rain forecasting RMSE parameter and r2 parameter for compare and obtain optimum delay day for correct rain forecasting and predict for climatology application.

Equation (2) Root Mean Square Error (RMSE) [8]

$$\text{RMSE} = \sqrt[2]{\sum_{i=1}^{n}\left(\frac{\left(\hat{x}_i - x_i\right)^2}{n}\right)} \qquad (2)$$

Equation 3 Coefficient of determination comparison metric

$$r^2 = 1 - \frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \qquad (3)$$

The number of data points forecasted is n, xi is the actual value, $\hat{x}_i$ is the forecast value, and $\overline{x}$ is the mean of the actual data. The coefficient of determination can have several possible values:

$$r^2 = \begin{cases} 1 & if\ \forall i\ \hat{x}_i = x_i \\ 0 > k > 1 & if\ \hat{x}_i\ is\ a\ better\ \ forecast\ than\ \overline{x} \\ 0 & if\ generally\ \hat{x}_i = \overline{x} \\ k < 0 & if\ \hat{x}_i\ is\ a\ worse\ forecast\ than\ \overline{x} \end{cases}$$

A coefficient of determination close to one is preferable.

Final results of applying neural networks contribute to the day with more reduced RMSE, but given the time required to train the network, increases with increasing number of delays, this is not optimal. Our main objective was efficient and effective prediction of the lag.

We test several network to find the best number of the inputs and hidden layers neurons to optimize the value of lags.

| **Models** | **Specifications** | | |
|---|---|---|---|
| | **Learning epochs** | **Testin g RMS E** | **Training time (seconds)** |
| MARS | 1 | 0.0780 | 5 |
| EFuNN | 1 | 0.091 | 27 |
| ANN-SCGA | 600 | 0.0923 | 90 |
| ABFNN | 1000 | 0.093 | 75 |
| This work (optimum GS 5-5-1) | 1000 | 0.0237 | 70 |
| This work (GS 6-6-1) | 1000 | 0.0233 | 90 |
| This work (GS 8-5-1) | 1000 | 0.0229 | 205 |
| This work (GS 20-6-1) | 1000 | 0.0222 | 1250 |

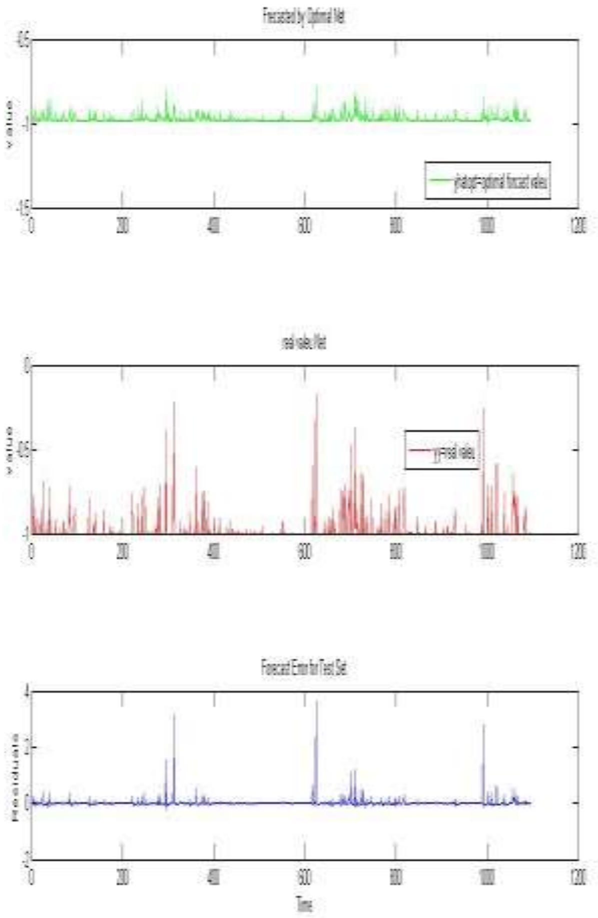


**Figure 6.**Output graphs, respectively the predicted, measured values and forecast error.

## References

[1] S. Hastenrath, and L. Greischar. (1999). Prediction of the summer rainfall over South Africa. Journal of Climate. 1999, 8: 1511-1518. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[2] K. C. Luk. An application of artificial neural networks for rainfall forecasting. Mathematical and Computer Modeling Journal. 2001, 33: 883-699.

[3] Eric A. Plummer "Time series forecasting with Feed-Forward neural network" 2000.

[4] S. P. Ninan, S. P, and J. Babo. A neural network tool for analyzing trends in rainfall. Computers & geosciences Journal. 2003, 29: 215–223.R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[5] Wikipedia, Gitashenasi Province Atlas of Iran.

[6] Qing Song; Jizhong Xiao; Yeng Chai Soh: Sch. of Electr. & Electron. Eng., Nanyang Technol. Univ "Robust back propagation training algorithm for multilayered neural tracking" 2002.

[7] Ajith Abraham, Dan Steinberg and Ninan Sajeeth Philip "Rainfall Forecasting Using Soft Computing Models and Multivariate Adaptive Regression Splines".

[8] ] Najmeh Khalili, Saeed Reza Khodashenas , Kamran Davary3 and Fatemeh Karimaldini "Daily Rainfall Forecasting for Mashhad Synoptic Station using Artificial Neural Networks" 2011.