Abhishek Agrawal, Vikas Kumar, Ashish Pandey, Imran Khan / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue 2,Mar-Apr 2012, pp.974-980

AN APPLICATION OF TIME SERIES ANALYSIS FOR WEATHER FORECASTING

Abhishek Agrawal*, Vikas Kumar[1],Ashish Pandey**[2],Imran Khan**[3]**
*(M. Tech Scholar, Department of Computer Science, Bhagwant University, Ajmer, India)
** (Asst. Professor, Department of Computer Science, Bhagwant University, Ajmer, India)

ABSTRACT
Weather forecasting has been one of the most challenging problems around the world for more than half a century. Not only because of its practical values in meteorology, but it is also a typical unbiased time series forecasting problem in scientific research. This paper utilizes artificial neural network (ANN) simulated in MATLAB to predict two important weather parameters i.e. maximum and minimum temperature. The model has been trained using past 60 years of data (1901-1960) and tested over 40 years to forecast maximum and minimum temperature. The results based on mean square error function (MSE) confirm, this model which is based on multilayer perceptron has the potential to successful application to weather forecasting.

Keywords - **Artificial neural network, Multilayer perceptron, Time series analysis, Mean Square error function, MATLAB.**

I. INTRODUCTION
WEATHER forecasting for the future is one of the most important attributes to forecast because agriculture sectors as well as many industries are largely dependent on the weather conditions [1]. The present technology uses numerical model based weather forecasting which involves a large number of calculations made very rapidly using supercomputers. Mohsen Hayati and Zahra Mohebi worked on one day ahead prediction of weather parameters [14].Suarjit Chattopadhyay, Department of Mathematics, Techno Model School proposed an Artificial Neural Network based predictive model for predicting average summer-monsoon rainfall [12]. Chaw Thet Zan and Thinn Thu Naing of University of Computer Studies, Yangon presented Hidden Markov Models (HMMs) approach for forecasting rainfall time series data for Myanmar rainy season [13]. Paras, Sanjay Mathur, Avinash Kumar, and Mahesh Chandra found that the feature based forecasting model can make predictions with high degree of accuracy [1].

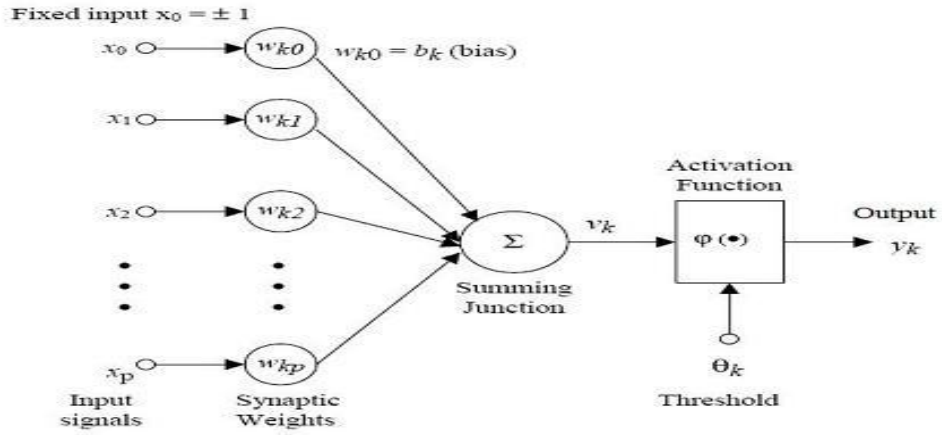
There is a scope of improvement in the accuracy of forecasts made by these models when considering the benefits that it has. The idea of incorporating human intelligence together with above discussed model can result in enormous improvement in accuracy of weather forecasting [2]. The objective of this study is to develop a model which comprises this intelligent behavior. The variables defining weather conditions like temperature (maximum or minimum), rainfall etc., vary continuously with time, forming time series of each parameter which can be used to develop a forecasting model that uses this time series data.

This paper is organized as follows: a brief overview of Artificial Neural Network. Multilayer perceptron, Backpropagation, Time Series Analysis and MATLAB are discussed in next section, the data groups and data variables are described in the later section, Experimental results obtained using the proposed network and generalisation capacity of model are listed in second last section, and finally we conclude this paper in last section.

II. NEURAL NETWORK MODEL
The development of Artificial Neural Network started 50 years ago to mimic the human ability to adapt the changing circumstances and the current environment. It is composed of a large number of highly interconnected processing elements (neurons) working in union to solve specific problems, as shown in Fig1. An artificial neuron is characterized

Fig1: Artificial Neural Network.

by, architecture (connection between neurons, training or learning determining weights on the connections, activation function [3]. A neural network resembles the human brain in two respects:



1. Knowledge is acquired by the network through a learning process.
2. Inter-neuron connection strengths known as synaptic weights are used to store the knowledge.

The most interesting fact about neural networks is the possibility of learning. A neural network is an adaptive system that changes its structure based on external or internal information that flows through the network during the learning phase [4]. It is used to learn patterns and relationships in data. The learning rules enables a network to 'gain knowledge' from available data and apply that knowledge to train the network [3].

There are three major learning paradigms, each corresponding to a particular abstract learning task. These are supervised learning, unsupervised learning and reinforcement learning. In Supervised Learning, a commonly used function is the mean-squared error, which tries to minimize the average squared error between the network's output, f(x), and the target value y over all the pairs. In order to minimize this function using gradient descent for the class of neural networks called multilayer perceptrons, we obtain the common and well-known backpropagation algorithm for training neural networks [4].

III. MULTI LAYER PERCEPTRON NETWORK

A multilayer perceptron (MLP) is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate output. An MLP consists of multiple layers of nodes in a directed graph, with each layer fully connected to the next one, as shown in Fig 2. It utilizes a supervised learning technique called backpropagation for training the network [5][6]. It is a modification of the standard linear perceptron, which can distinguish data that is not linearly separable [7].

The two main activation functions used in current applications are both sigmoid, and are described by

$$\phi(y_i) = \tanh(v_i) \quad \text{and} \quad \phi(y_i) = (1 + e^{-v_i})^{-1}$$

in which the former function is a hyperbolic tangent which ranges from -1 to 1, and the latter is equivalent in shape but ranges from 0 to 1. y_i is the output of the ith node (neuron) and v_i is the weighted sum of the input synapses[8].

It can implement arbitrary complex input/output mappings or decision surfaces separating different patterns. Multiple-layer networks are quite powerful. For instance, a network of two layers, where the first layer is sigmoid and the second layer is linear, can be trained to approximate any function (with a finite number of discontinuities) arbitrarily well. This kind of two-layer network is used extensively in Backpropagation [3].

Back propagation is a systematic method for training multi layer artificial neural networks. It has a mathematical foundation that is strong if not highly practical. It is a multi-layer forward network using extend gradient descent based delta learning rule,

commonly known as Backpropagation rule. Being a gradient descent method it minimises the total squared error of the output computed by the network. The network is trained by supervised learning method. The aim of this network is to train the net to achieve a balance between the ability to respond correctly to the input patterns that are used for training and the ability to provide good responses to the input that are similar. This algorithm covers a wide area of applications like, Optical character Recognition, Image Compression, Data Compression, Face Recognition etc [3].

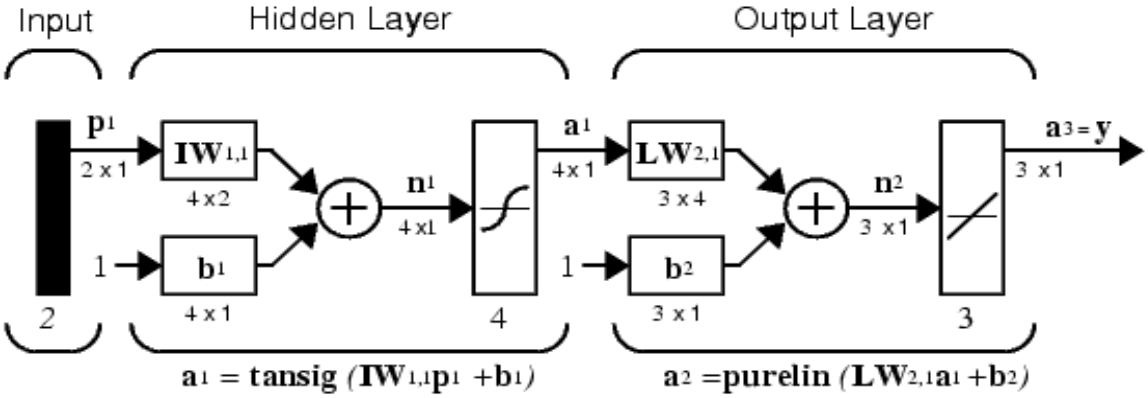


Fig 2: Multilayer Perceptron Network

IV. Time Series Analysis

A Time series (TS) is a sequence of observations ordered in time. Mostly these observations are collected at equally spaced discrete time intervals. A basic assumption in any time series analysis modeling is that some aspects of the past pattern will continue to remain in the future. Also under this setup often the time series process is assumed to be based on past values of the main variable but not on explanatory variables which may affect the system. So the system act as a black box and we may only be able to know 'what' will happen rather than 'why' it happens. Here it is tactically assumed that information about the past is available in the form of numerical data. Ideally at least 50 observations are necessary for performing time series analysis/modeling as propounded by Box and Jenkins who were pioneers in TS modeling [9].

Time series analysis can be used more easily for forecasting purposes because historical sequences of observations upon study variables are readily available from published sources. These successive observations are statistically dependent and Time series modeling is concerned with techniques for the analysis of such dependencies [9].

V. Matrix Laboratory(MATLAB)

The name MATLAB stands for matrix laboratory. MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. MATLAB is an interactive system whose basic data element is an array that does not require dimensioning. This allows you to solve many technical computing problems, especially those with matrix and vector formulations, in a fraction of the time [3].

VI. Implementation Of Proposed Work

Weather parameters like maximum temperature, minimum temperature, extent of rainfall, cloud conditions, wind streams and their directions, are used to predict weather conditions, but temperature affects all other parameters. Temperature affects the ability of air to absorb water vapor (warmer air can contain more vapor) and ultimately it is the latent heat in water vapor which is the energy source for thunderstorms, hurricanes etc., which affects extent of rainfall, cloud conditions, wind streams and their directions. The rate at which temperature decreases with height affects the stability of the atmosphere. Increasing temperature is likely to lead to increasing precipitation.

The change in the weather takes place because of unequal heating of the atmosphere by the sun. As the sun rises and the sun's rays warm the earth's surface, it heats the air close to it. This makes the air temperature rise and the day becomes increasingly hot. So the amount of sunlight the surface receives

determines the hotness or coldness of the air, which in turn decides the state of weather of any place.

The proposed system takes minimum temperature and maximum temperature as input parameters and predicts the same. Weather data of hundred years for the meteorological variables as shown in Table 1 were collected from the Indian Meteorological Department (IMD). The ellicitated data refers to two sets i.e. average maximum temperature and average minimum temperature for 12 months of each year. The general Structure of inputs/outputs for the network is shown in Fig3. The global data is divided into two groups, the training group corresponding to 60% of the data and the test group, corresponding to 40% of the data; so that the generalization capacity of the network could be verified after the training phase.
Mean Squared Error (MSE) is used as a measure of error made by the proposed network.

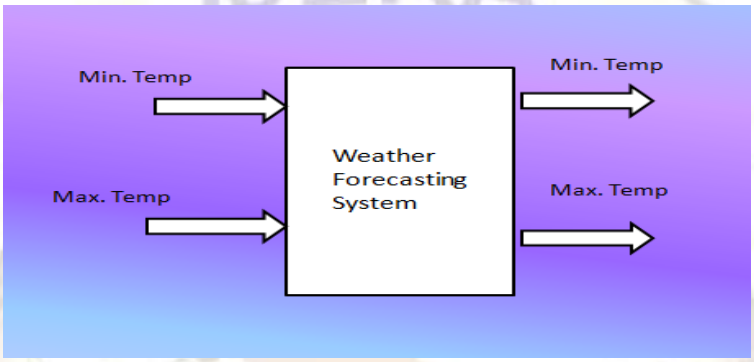


Fig 3: General Structure of inputs/outputs for the network.

Table 1: Meteorological Variables.

S.NO	VARIABLE	UNIT
1.	Minimum temperature	Deg.C
2.	Maximum Temperature	Deg.C

VII. RESULTS

The obtained optimal network structure is shown in Table 2. The network is trained with 60 years of data for both maximum and minimum temperature. After training phase the network is simulated by using 40% of data so that the predicted result can be verified. The graphs depicting the stabilization of different weight and bias metrics are shown in Fig 4, 5 and 6.

Table 2: The optimal Network Structure.

Network Structure	
Number of Hidden neurons	**3**
Number of epochs	**100**
Activation function used in hidden layer	**tan-sig**
Activation function used in output layer	**Pure linear**

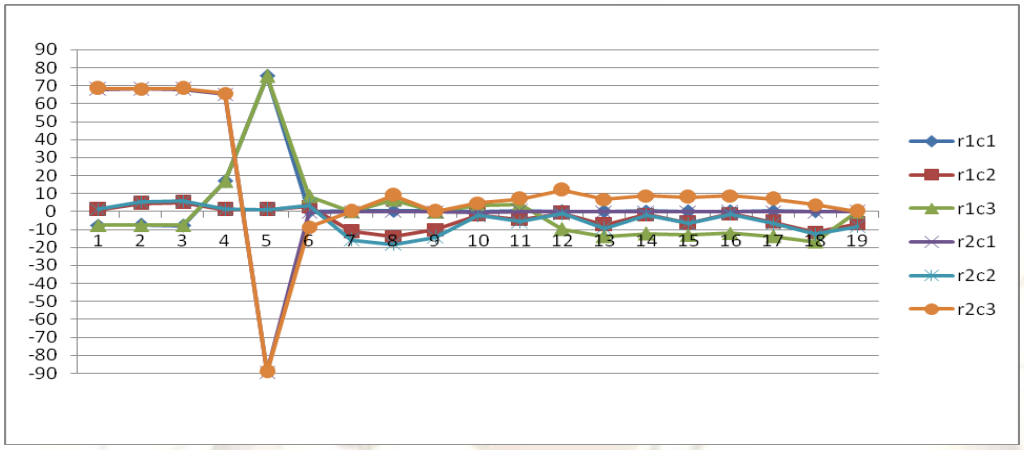


Fig 4: Graph showing stabilisation of Weight metric for layer 2.

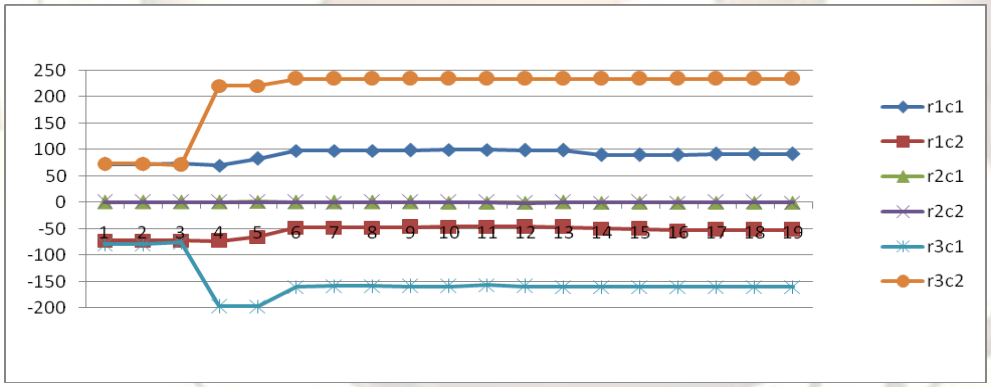


Fig 5: Graph showing stabilisation of Weight metric for layer 1.

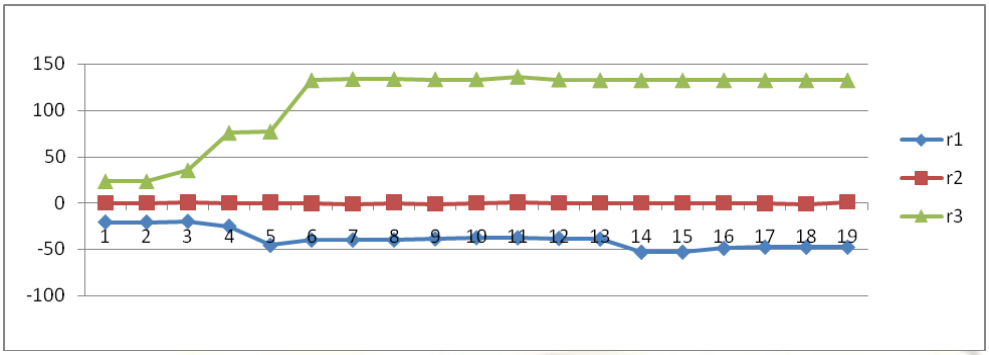


Fig 6: Graph showing stabilisation of bias metric

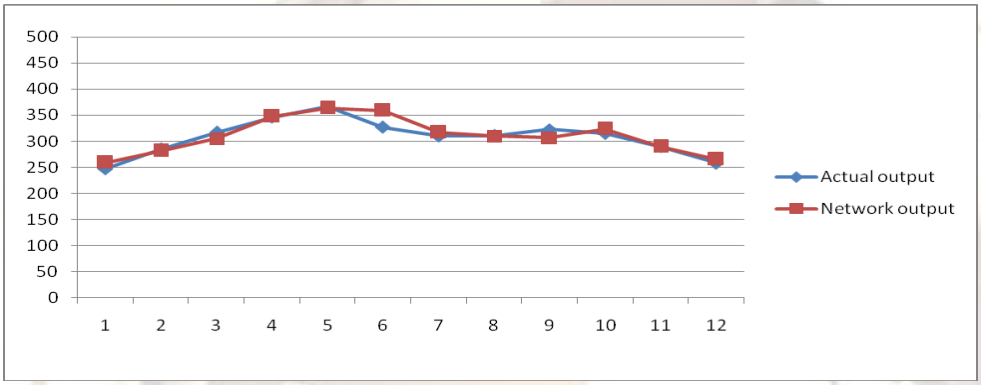
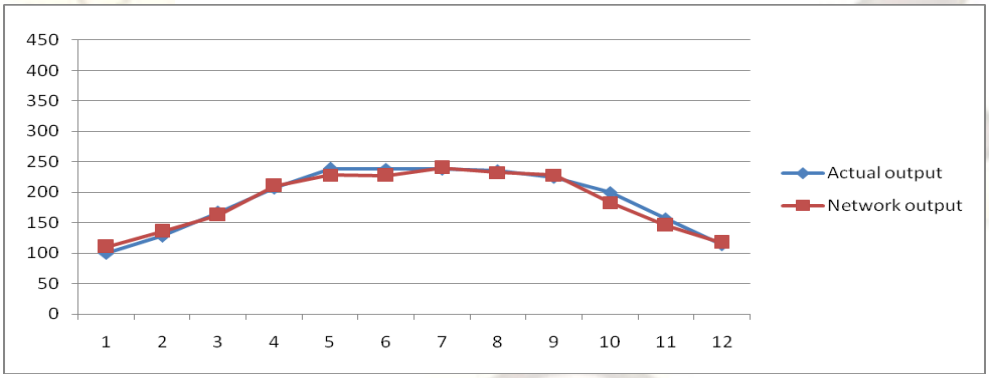


Fig 8: Comparison between actual and predicted maximum temperature values for year 2001



CONCLUSION

The results conclude that Multilayered Neural Network can be an effective tool in weather prediction. As we can see in Fig-7 & Fig-8, this type of Network can correctly map input to the output using historical data. Once the number of layers and number of units in each layer , has been selected , the network's weights and thresholds must be set so as to minimise the prediction error made by the network. The forecasting reliability was evaluated by comparing the actual and predicted temperature values. The results show that the network can be an important tool for temperature forecasting.

REFERENCES

[1] Paras, S.Mathur, A.Kumar and M.Chandra (2007), "A Feature Based Neural Network Model for Weather Forecasting",World Academy of Science, Engineering and Technology.
[2] Mr.R.C vashishtha, director IMD.
[3] S N Sivanandam, S Sumathi, S N Deepa, "Introduction to neural Networks using MATLAB".
[4] http://en.wikipedia.org/wiki/Neural_network.
[5] Rosenblatt, Frank. x. (1961),"Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms", Spartan Books, Washington DC.
[6] Rumelhart, David E., Geoffrey E. Hinton, and R. J. Williams (1986),"Learning Internal Representations by Error Propagation". David E. Rumelhart, James L. McClelland, and the PDP research group. (editors), Parallel distributed processing: Explorations in the microstructure of cognition, Volume 1: Foundations. MIT Press.
[7] Cybenko, G. (1989), "*Approximation by superpositions of a sigmoidal function Mathematics of Control, Signals, and Systems (MCSS)*", 2(4), 303–314.
[8] http://en.wikipedia.org/wiki/Multilayer_perceptron.
[9] Ramasubramanian V."*Time Series Analysis*", I.A.S.R.I,Library Avenue,New Delhi.
[10] http://www.tropmet.res.in
[11] http://www.imd.gov.in.
[12] Chattopadhyay S. (2006), "*Multilayered feed forward Artificial Neural Network model to predict the average summer-monsoon rainfall in India*",IEEE ,Vol.:11, pp.:125-130
[13] Zan C. (2009) "*Myanmar Rainfall Forecasting Using Hidden Markov Model*",IEEE International Advance Computing Conference.
[14] Hayati M. and Mohebi Z.(2007) , "*Temperature Forecasting Based on neural Network Approach*", World Applied Sciences Journal 2 (6):613-620.ISSN 1818-4952.