Time Series based Temperature Prediction using Back Propagation with Genetic Algorithm Technique

Shaminder Singh[1], Pankaj Bhambri[2] and Jasmeen Gill[3]

[1] **CSE, Punjab Technical University, BCET**
Ludhiana, Punjab 141206, India

[2] **CSE, Punjab Technical University, GNE**
Ludhiana, Punjab 141002, India

[3] **CSE, Punjab Technical University, RIMT-IET**
Mandi Gobindgarh, Punjab 147301, India

Abstract

Temperature prediction is a temporal and time series based process. Accurate forecasting is important in today's world as agricultural and industrial sectors are largely dependent on the temperature. Due to non-linearity in climatic physics, neural networks are suitable to predict these meteorological processes. Back propagation integrated with genetic algorithm is the most important algorithm to train neural networks. In this paper, in order to show the dependence of temperature on a particular data series, a time series based temperature prediction model using integrated back propagation with genetic algorithm technique is proposed. In the proposed technique, the effect of under training and over training the system is also shown. The test results of the technique are enlisted along with.

Keywords: *Artificial Neural Networks, Back Propagation Algorithm, Genetic Algorithms, Time Series Prediction.*

1. Introduction

Forecasting is a phenomenon of knowing what may happen to a system in the next coming time periods [4]. Temporal 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 data values $x_{t+1}, x_{t+2}, \ldots, x_{t+m}$. The goal is to observe or model the existing data series to enable future unknown data values to be forecasted accurately [11]. As weather is a continuous, data-intensive and dynamic process, the parameters required to predict temperature are enormously complex such that there is uncertainty in prediction even for a short period [8]. These properties make temperature forecasting a formidable challenge. The property of artificial neural networks that they not only analyze the time series data but also learn from it for future predictions makes them suitable for time series based temperature forecasting. Neural networks provide a methodology for solving many types of non-linear, time based problems that are difficult to be solved through traditional techniques [11]. Hence these characteristics of neural networks guided this research work to use them for the prediction of the meteorological processes.

Inspired by the brain, neural networks are an interconnected network of processing elements called neurons. Neural networks learn by example i.e. they can be trained with known examples. One of the most popular training algorithms in the domain of neural networks used so far, for temperature forecasting is the back propagation algorithm (BPN). As the algorithm suffers from many problems, attempts have been made by various researchers to solve these problems using genetic algorithms [2], [5]. Due to the temporal nature of weather processes, time series prediction has been introduced, but, no advancement beyond feed-forward neural networks trained with integrated back propagation and genetic algorithm has revolutionized the field. Therefore, much work still waits.

So, the motivation of this work is firstly, to develop time series based temperature forecasting model using hybrid neural network approach i.e. integrated BPN with GA and secondly, to estimate the effects of under-training and over-training the model.

The remainder of the article is organized as follows. Section 2 introduces the artificial neural networks. Next, a brief description of the time series predictor, neural network predictor and data series partitioning is given in Section 3. The details of the time series based temperature forecasting based on integrated BP/GA technique are shown in Section 4, followed by results in Section 5. Finally, conclusions are summarized in Section 6.

2. Artificial Neural Networks

Artificial Neural Network can be defined as a pool of simple processing units (neurons) which communicate among themselves by means of sending analog signals. These signals travel through weighted connections between neurons. Each of these neurons accumulates the inputs it receives, producing an output according to an internal activation function. This output can serve as an input for other neurons, or can be a part of the network output [3].
Back propagation is a systematic method of training multilayer artificial neural networks. It is built on sound mathematical base. The back propagation is a gradient descent method in which gradient of the error is calculated with respect to the weights for a given input by propagating the error backwards. The combination of weights which minimizes the error function is considered to be a solution of the problem [1].
Although Back propagation algorithm is an efficient technique applied to classification problems, system modeling, adaptive robotics control, but it suffers from local minima problem, scaling problem, long training time etc [1].
Genetic Algorithms developed in 1970 by John Holland, are computerized search and optimization algorithm that mimic the principle of natural genetics and natural selection. Genetic Algorithms perform directed random searches through a given set of alternatives to find the best alternative with respect to given criteria of fitness [6].
To eliminate the problems of back propagation algorithm, integrated BP/GA technique was developed. This technique is an efficient approach if only the requirement of a global searching is considered. It is good at global search (not in one direction) and it works with a population of points instead of a single point. Also it blends the merits of both deterministic gradient based algorithm BP and stochastic optimizing algorithm GA [1].
Neural Networks have been widely used as time series forecasters: most often these are feed-forward networks which employ a sliding window over the input sequence. Typical examples of this approach are market predictions, meteorological and network traffic forecasting [10]. Two important issues must be addressed in such systems: the frequency with which data should be sampled, and the number of data points which should be used in the input representation.

3. Time Series Prediction

A time series is a sequence of vectors, $\boldsymbol{x}(t)$, t = 0,1,… , where t represents elapsed time. For simplicity we will consider here only sequences of scalars, although the techniques considered generalize readily to vector series. Theoretically, x may be a value which varies continuously with t, such as a temperature. In practice, for any given physical system, x will be sampled to give a series of discrete data points, equally spaced in time.
Work in neural networks has concentrated on forecasting future developments of the time series from values of x up to the current time. Formally this can be stated as: find a function f: $R^N \rightarrow R$, such as to obtain an estimate of x at time t + d, from the N time steps back from time t, so that:

$$x(t+d)=f(x(t), x(t-1), \ldots, x(t-N+1)) \quad (1)$$

and

$$x(t+d)=f(\mathbf{y}(t)) \quad (2)$$

where $\mathbf{y}(t)$ is the N - ary vector of lagged x values
Normally d will be one, so that f will be forecasting the next value of x [10].
Time series forecasting has several important applications. One application is preventing undesirable events by forecasting the event, identifying the circumstances preceding the event, and taking corrective action so the event can be avoided. Another application is forecasting undesirable, yet unavoidable, events to preemptively lessen their impact. At this time, the sun's cycle of storms, called solar maximum, is of concern because the storms cause technological disruptions on Earth and other meteorological processes like interplanetary shocks, volcano eruption, cyclones, earth quakes, tsunamis etc. can also be predicted with the help of time series based forecasting. Also the daily weather forecasting of any place useful for agricultural purposes can be predicted through it. Finally, many people, primarily in the financial markets, would like to profit from time series forecasting. Whether this is viable is most likely a never-to-be-resolved question. Nevertheless many products are available for financial forecasting [11].
The standard neural network method of performing time series prediction is to induce the function f using any feed-forward function approximating neural network architecture using a set of N-tuples as inputs and a single output as the target value of the network. This method is often called the *sliding window technique* as the N-tuple input slides over the full training set. Figure 1 gives the basic architecture [10].
One typical method for training a network is to first partition the data series into three disjoint sets.

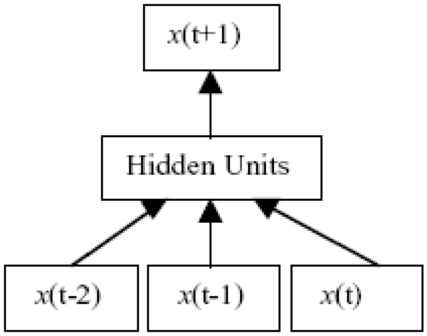

Fig. 1 The standard method of performing time series prediction using a sliding window of, in this case, three time steps.

These sets are: the training set, the validation set, and the test set. The network is trained (e.g., with back propagation) directly on the training set, its generalization ability is monitored on the validation set, and its ability to forecast is measured on the test set. A network's generalization ability indirectly measures how well the network can deal with unforeseen inputs, in other words, inputs on which it was not trained [11]. A network that produces high forecasting error on unforeseen inputs, but low error on training inputs, is said to have over-fit the training data. Over-fitting occurs when the network is blindly trained to a minimum in the total squared error based on the training set. A network that has over-fit the training data is said to have poor generalization ability.
To control over-fitting, the following procedure to train the network is often used:

1. After every 'n' epoch, sum the total error for all examples from the training set.
2. Also, sum the total squared error for all examples from the validation set. This error is the validation error.
3. Stop training when the trend in the error from step 1 is downward and the trend in the validation error from step 2 is upward.

When consistently the error in step 2 increases, while the error in step 1 decreases, this indicates the network has over-learned or over-fitted the data and training should stop. When using real-world data that is observed (instead of artificially generated), it may be difficult or impossible to partition the data series into three disjoint sets. The reason for this is the data series may be limited in length and/or may exhibit non-stationary behavior for data too far in the past (i.e., data previous to a certain point are not representative of the current trend) [9], [11].

4. Time Series based Temperature Prediction

In this section, the features of temporal data to the integrated BP/GA technique are introduced. The data used in this research are the daily weather data for the Ludhiana city of Punjab (India). The data in the un-normalized form have been collected from the "Meteorological Department of Punjab Agriculture University, Ludhiana (Punjab)" of the year 2009. The first thirty days data (month of January, 2009) have been used in this research. For training, the first twenty five days data have been used and next five days data have been used for testing purposes [2].
Before feeding the data to the network, it is to be converted to normalized form so as to provide improved performance or otherwise the use of original data to network may cause convergence problem. All the weather data sets were, therefore, transformed into values between 0 and 1 through dividing the difference of actual and minimum values by the difference of maximum and minimum values [7].
The neural network architecture along with the inputs required to feed to the network to perform time-series based temperature prediction and the outputs are shown below in fig. 2.

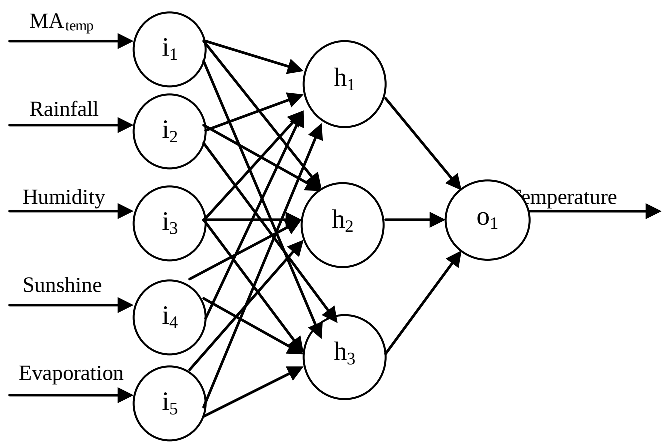

Fig. 2 Neural Network Architecture for the Proposed Model.

In this research, 5-3-1 neural network architecture has been used for BP/GA technique. The number of input neurons is 5 representing the moving average of mean air temperature of the previous temperature data, daily rainfall, relative humidity, sunshine and evaporation for the day for which the mean air temperature is to be predicted, the number of hidden neurons is 3 for processing and the number of outputs is 1 representing the weather variable i.e. mean air temperature to be forecasted.
The proposed time series based model starts with the collection of weather related data, selecting the weather parameters to be forecasted, extracting the relation

between the different weather parameters, formation of training data set (containing inputs and outputs) and test data set (containing inputs).

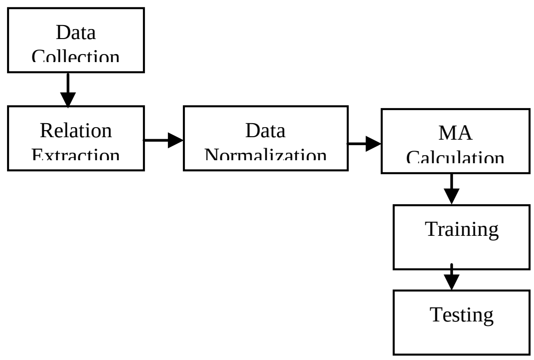

Fig. 3 Methodology for the Proposed Model.

For performing the time series prediction, a sliding window of size 5 has been moved over the full data set to obtain the moving average. This along with the dependent parameters act as an input to the system and have been used to train the network.

For training the network, an initial population of chromosomes is randomly generated. Then the weights are extracted from each chromosome depending upon the number of genes a chromosome is having. The cumulative error and the fitness are calculated over the inputs obtained from forecasting data. Then the crossover operator is applied for preparing the new population. This process is repeated till the stopping condition has been reached [1], [2].

5. Results and Discussions

The temporal weather forecasting technique based on BP/GA technique has been implemented by taking different population sizes.

Table 1: Selection of appropriate NN architecture

Population Size	***Hidden Neurons***	***Iterations***	***MAPE***
30	1	110	1.10
60	2	140	0.86
90	3	220	0.42
120	4	282	1.47
150	5	425	1.85

For each value of population, the program has been executed and the error has been calculated. Table 1 shows the variations in population size, number of neurons in hidden layer and the corresponding mean absolute percentage error values.

From the table 1, it is clear that the MAPE value is the lowest corresponding to population size 90 and number of hidden neurons as 3. So the present setup will use this population size for further research. The error vs. iteration graph corresponding to population size 90 is shown in fig. 4.

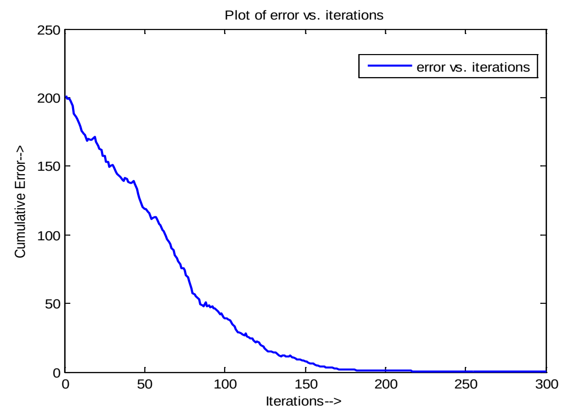

Fig. 4 The cumulative error values corresponding to iterations for population size 90.

The error values corresponding to mean air temperature are shown in fig. 5 along with the Series 1, Series 2 and Series 3 of the inputs are shown. Error values are shown after 200 epochs, 400 epochs and 600 epochs. Clearly Series 3 shows the minimum error values in all the cases and it shows the lowest value after 400 epochs.

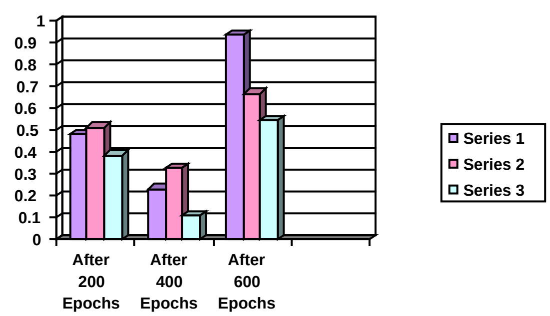

Fig. 5 Temporal prediction for mean air temperature.

Below is shown the actual prediction of mean air temperature using the proposed method along with the desired output as recorded with the help of instruments in fig. 6.

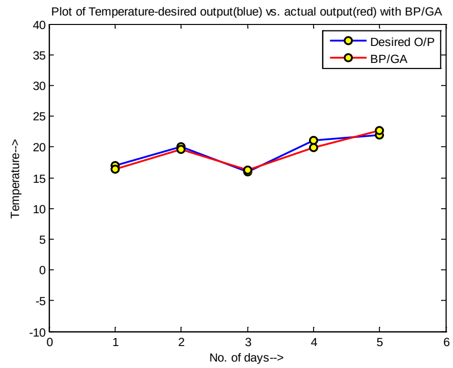

Fig. 6 The five-day mean air temperature prediction using the proposed model.

It is clearly shown in the fig. 6 that the time series based temperature prediction model using integrated BP/GA technique is suitable to predict the temperature. Secondly, dependence of weather parameters on the time series data is clearly shown in fig. 5. For the same weather parameter, the error values come out to be different when the network is trained with different data series. Thirdly the effect of under training the network through 200 epochs and over training the network through 600 epochs is easily visible as the error values are lowest after 400 epochs.

6. Conclusions

From the analysis above, it is easy to observe the compensability between time series based BP/GA technique and the back propagation alone. The proposed technique can learn efficiently by combining the strengths of GA with BP. It is good at time series data, global search (not in one direction) and it works with a population of points instead of a single point. Also it blends the merits of both deterministic gradient based algorithm BP and stochastic optimizing algorithm GA. Hence the use of the time series based temperature prediction model using integrated BP/GA technique is proposed.

References

[1] S. Rajasekaran and P. Vijayalakshmi, Neural networks, Fuzzy Logic and Genetic Algorithms, New Delhi: Prentice Hall of India, 2004.

[2] J. Gill, B. Singh and S. Singh, "Training Back Propagation Neural Networks with Genetic Algorithm for Weather Forecasting", 8th IEEE International Symposium on Intelligent Systems Subotica Serbia, 2010, vol. 8, pp. 465-469.

[3] A. Azadeh and et al., "Integration of ANN and GA to Predict Electrical Energy consumption", 32nd IEEE Conference on Industrial Electron., IECON, Paris, France, 2006.

[4] P. Sarangi and et al., "Short Term Load Forecasting using Artificial Neural Network: A Comparison with Genetic Algorithm Implementation", J. of ARPN Engineering and Applied Sciences, vol. 9, 2009.

[5] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Machine Learning, Addison - Wesley. Reading, MA, 1989.

[6] H.S. Rao, V.G. Ghorpade and A. Mukherjee, "A genetic algorithm based back propagation network for simulation of stress-strain response of ceramic-matrix-composites", J. of Computers and Structures, vol. 84, no. 5-6, 2006.

[7] I. Maqsood, M.R. Khan and A. Abraham, "Weather Forecasting Models Using Ensembles of Neural Networks", 3rd International Conference on Intelligent Systems, Design and Applications Germany, 2003, pp. 33-42.

[8] A. Abraham et. al., "Soft Computing Models for Weather Forecasting" J. of Applied Sciences and Computations, USA, vol. 11, no. 3, pp. 106-117, 2004.

[9] M. Peter and Roth, "Temporal Pattern Recognition with Neural Networks" Pattern Recognition & Machine Learning, 1998.

[10] R. J. Frank, N. Davey and S. P. Hunt, "Time Series Prediction and Neural Networks", J. of Intelligent and Robotic Systems, vol. 31, no. 1-3, 2001.

[11] E. A. Plummer, "Time Series Forecasting with Feed-forward Neural Networks: Guidelines and Limitations", M.S. thesis, Department of Computer Science, The Graduate School of The University of Wyoming, Laramie, USA, 2000.