Training Back Propagation Neural Networks with Genetic Algorithm for Weather Forecasting

Er. Jasmeen Gill*, Er. Baljeet Singh ** and Er. Shaminder Singh***
*Assistant Professor, RIMT-IET, Mandi Gobindgarh
er.jasmeengill@gmail.com
** Assistant Professor, BBSBEC, Fatehgarh Sahib
baljitkhehra@rediffmail.com
*** Assistant Professor, BCET, Ludhiana
er_shaminder@yahoo.co.in

Abstract- **Accurate weather forecasting is important in today's world as agricultural and industrial sectors are largely dependent on the weather conditions. Secondly, it is used to warn about natural disasters. Due to non-linearity in climatic physics, neural networks are suitable to predict these meteorological processes. Back Propagation algorithm using gradient descent method is the most important algorithm to train a neural network for weather forecasting. Back propagation algorithm suffers from several problems. In this paper, in order to overcome some of these problems, an integrated back propagation based genetic algorithm technique to train artificial neural networks is proposed. In the proposed technique, back propagation is combined with genetic algorithm in such a way that the pitfalls of the algorithm get converted to benefits. The results and comparison of the technique with the previous one are enlisted along with.**

I. INTRODUCTION

Weather is a continuous, data-intensive, dynamic and chaotic process. The parameters required to predict weather are enormously complex such that there is uncertainty in prediction even for a short period [10]. These properties make weather forecasting a formidable challenge. The property of artificial neural networks that they not only analyze the data but also learn from it for future predictions makes them suitable for weather forecasting. Neural networks provide a methodology for solving many types of non-linear problems that are difficult to be solved through traditional techniques. Furthermore neural networks are capable of extracting the relationship between inputs and outputs of a process without the physics being explicitly provided [8]. Hence these characteristics of neural networks guided us to use them for the prediction of the weather 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 [7]. One of the most popular training algorithms in the domain of neural networks used so far, for weather forecasting is the back propagation algorithm. It is a gradient descent method. The algorithm suffers from many problems. Several attempts have been made by various researchers to solve these problems using genetic algorithms, the computerized search and optimization algorithms that mimic the principle of natural genetics and natural selection [5], [6]. But, in the field of weather forecasting, such efforts are still to be put up.

So, the motivation of this work is firstly, to integrate BPN with GA in such a way that the disadvantages of back propagation algorithm get converted to benefits, used further for an accurate weather forecasting model and secondly, to perform comparison between the two main techniques- back propagation network based on gradient descent technique and back propagation network based on genetic algorithm to train the neural network.

The remainder of the article is organized as follows. Section 2 and Section 3 introduce the back propagation algorithm and genetic algorithms respectively. The details of the integrated BP/GA technique for weather forecasting model are shown in Section 4, followed by results in Section 5. Finally, conclusions are summarized in Section 6.

II. BACK PROPAGATION ALGORITHM

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 from output layer to hidden layer and further to input layer. This method adjusts the weights according to the error function. So, the combination of weights which minimizes the error function is considered to be a solution of the problem [4].

The merits of back propagation algorithm are that the adjustment of weights is always toward the descending direction of the error function and that the adjustment only needs some local information. Secondly, the mathematical formula present here, can be applied to neural network with any architecture and does not require any special mention of the function to be learnt. Also the computing time is reduced if the weights chosen are small at the beginning [1].

Although back propagation algorithm is an efficient technique applied to classification problems, system modeling, adaptive robotics control, but it does have some pitfalls. For one, back propagation algorithm suffers from the scaling problem. It works well on simple training problems. However, as the problem complexity increases, the

978-1-4244-7395-3/10/$26.00 ©2010 IEEE

performance of back propagation falls off rapidly because gradient search techniques tend to get trapped at local minima. When the nearly global minima are well hidden among the local minima, back propagation can end up bouncing between local minima [2]. A second shortcoming is that the convergence of the algorithm is very sensitive to the initial value. So, it often converges to an inferior solution and gets trapped in a long training time. Furthermore, the required precision is so high that it is difficult to realize the weight storage [11].

III. GENETIC ALGORITHM

Genetic algorithm developed by John Holland in 1970, is computerized search and optimization algorithm that mimics the principle of natural genetics and natural selection. Genetic algorithm performs directed random searches through a given set of alternatives to find the best alternative with respect to given criteria of fitness. Fitness is defined as a figure of merit which is to be either maximized or minimized. An initial population of chromosomes is taken to generate offspring that competes for survival to make up the next generation of population [6]. Three main inheritance operators used in genetic algorithm are reproduction, crossover and mutation. Successive generation of chromosome improve in quality provided that the criteria used for survival is appropriate [5]. This process is referred to as Darwinian natural selection or survival of the fittest [1].

As compared with back propagation algorithm, genetic algorithm is a parallel stochastic optimizing algorithm and is good at global searching (not in one direction). Also it works with a population of points instead of a single point. It is a population based search algorithm and multiple optimal solutions can be captured thereby reducing the effect to use the algorithm many times [12]. Secondly, genetic algorithm works with a string coding of variables instead of the variables. The advantage of working with a coding of variables is that coding makes the search space discrete even though the function may be continuous. Thus, a discrete function can be handled with no extra cost [1], [9].

The price one pays for genetic algorithm is its slowness. The slowness is mainly due to the slow but crucial exploration mechanisms employed i.e. reproduction, crossover and mutation. In addition, genetic algorithm starts searching from random genes, which will cost a lot of time [11], [12].

IV. INTEGRATED BACK PROPAGATION BASED GENETIC ALGORITHM (BP/GA TECHNIQUE)

The proposed weather forecasting model based on BP/GA technique starts with the collection of weather related data, selecting the weather parameters to be forecasted, formation of training data set and testing data set. Finally the methodology and its simulation are provided.

A. Weather Parameters

The daily weather parameters collected from PAU Ludhiana are shown in Table I. along with their units of measurement. The parameters chosen for prediction in this setup are mean air temperature (ºC), relative humidity (%) and daily rainfall (mm). There is no particular reason behind this choice of weather parameters. The choice is made just to predict three main weather variables.

TABLE I.

Sr. No.	Meteorological Variables	Unit
1.	Air Temperature	ºC
2.	Soil Temperature	ºC
3.	Relative Humidity	%
4.	Vapor Pressure	mm
5.	Wind Speed	Km/h
6.	Wind Direction	---
7.	Sunshine	hrs
8.	Rainfall	mm
9.	Evaporation	mm

B. Research Data

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. Thirty days data (month of January, 2009) have been used in this research. First twenty five days data have been used for training and next five days data have been used for testing purposes.

C. Normalization of Data

After the collection of data and selection of the weather parameters, next issue is normalization of data. Neural networks generally provide improved performance with normalized data. The use of original data to train the neural 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 [8].

D. Methodology

Chromosomes form the initial population which is randomly generated and consist of some number of genes. Every gene is encoded of a randomly chosen string length. A particular number of weights are extracted from each chromosome depending upon the number of genes a chromosome have [1]. This calculation is done as follows:

Let the network configuration of the network is l-m-n. Therefore, the numbers of weights (genes), W that are to be determined are:

$$W = (l + n) * m \qquad (1)$$

With each gene being a real number, and taking the gene length as d, the string representing the chromosomes of weights will have a length of:

$$C = W * d \qquad (2)$$

It represents the weight matrices of the input-hidden-output layers. With this weight set, the network is trained for the first cycle and the cumulative error is calculated over the inputs obtained from weather forecasting data.

The search for selecting an individual is guided by the fitness of each individual i.e. evaluating the quality of each chromosome [9]. So, the fitness function is evaluated by reciprocating the cumulative error value as follows.

$$F = 1/MSE \qquad (3)$$

More is the fitness value; more are the chances of a chromosome to be selected for reproducing an offspring. With this criterion in mind, mating pool is prepared by replacing the individual with minimum fitness value by individual having maximum fitness value. For cross over, a two-point cross over method will be used to prepare the new population and the network is considered to be trained when 95% of the individuals have same fitness value [1], [13]. The various steps of weather forecasting model based on BP/GA algorithm are explained below and are shown in fig.1:

1. Start: Generate random population of ‘p’ chromosomes (suitable solutions for the problem).
2. Extraction: Extract weights for input-hidden-output (l-m-n) layers from each chromosome x.
3. Fitness: Evaluate the fitness f(x) of each chromosome x in the population by reciprocating the cumulative error values obtained for each input set (weather forecasting data).
4. New population: Create a new population by repeating following steps until the new population is complete
   - Selection: Select two parent chromosomes from a population according to their fitness (the better fitness, the bigger chance to be selected)
   - Crossover: Cross over the parents to form new offspring (children). If no crossover was performed, offspring is the exact copy of parents.
   - Mutation: With a mutation probability mutate new offspring at each position in chromosome.
   - Acceptance: Place the new offspring in the new population.
5. Repeat: Repeat steps 3 to 5 until stopping condition is met.
6. Test: Return the best solution in current population using the test set inputs (weather forecasting data) and the weights (obtained in the above five steps).

The block diagram for the algorithm is given in fig. 1.

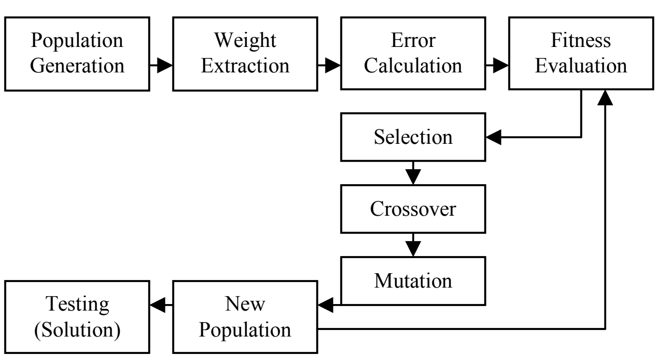

Figure 1. Methodology for integrated BP/GA technique

E. Simulation

In this research, 3-2-3 neural network architecture has been used. The number of input neurons is 3 representing the date fields, the number of hidden neurons is 2 for processing and the number of outputs is 3 representing the weather variables to be forecasted. A real coding system has been adopted for coding the chromosomes. As the network configuration is 3-2-3, therefore, the numbers of weights (genes) to be determined are 12, as in (1). Taking the gene length as 5, the string representing the chromosomes of weights will have a length of 60, as in (2). This is the weight matrix of the input-hidden-output layers. For cross over, we have used a two-point cross over selected at random and the selection is made on the basis of fitness value, as in (3). The stopping criterion for the training is when fitness converges.

V. RESULTS AND DISCUSSION

The BP/GA technique has been implemented by taking different population sizes. For each value of population, the program has been executed and the error has been calculated. Table II shows the variations in population size, number of neurons in hidden layer and the corresponding mean absolute percentage error values for the integrated BP/GA technique.

TABLE II

Population Size	Hidden Neurons	Iterations	MAPE
30	1	50	0.1043
60	2	120	0.0430
90	3	200	0.2018

Table II indicates that the MAPE value comes out to be the lowest corresponding to population size 60 and number of hidden neurons as 2. So the present setup will use this population size for further research. The error vs. iteration graph corresponding to population size 60 is shown in fig. 2.

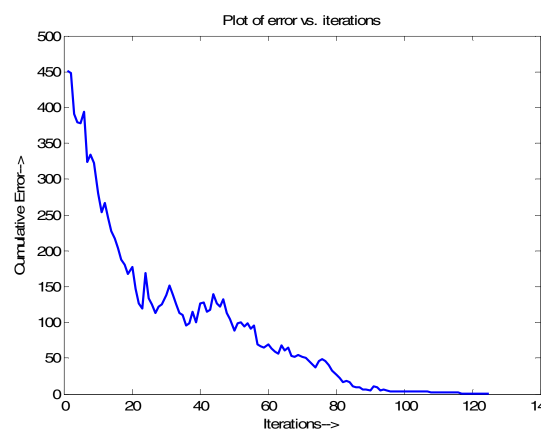

Figure 2. The cumulative error values corresponding to iterations for population size 60.

The error values corresponding to mean air temperature are shown in Table III for the last five days of January 2009.

TABLE III
TEMPERATURE FOR BP/GA

Day	Desired Output	Forecasted Output	Error Value
1	15.0	16.8	-1.8
2	16.0	17.2	-1.2
3	17.0	17.7	-0.7
4	18.0	18.7	-0.7
5	19.0	19.0	0.0

These error values are calculated as desired output – actual output and are represented graphically in fig. 3.

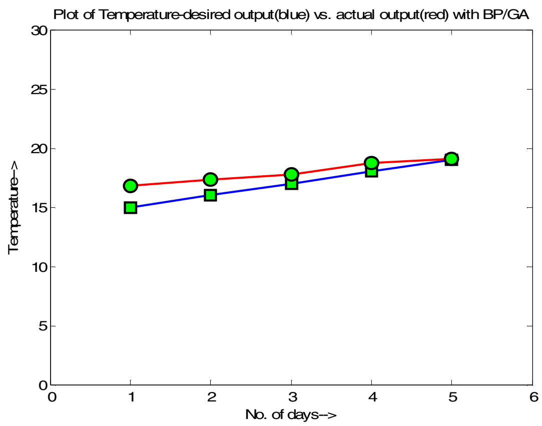

Figure 3. The 5-day-mean air temperature prediction for the BP/GA(red) technique and the desired output(blue).

Table IV shows the prediction of relative humidity along with the error values.

TABLE IV
RELATIVE HUMIDITY FOR BP/GA

Day	Desired Output	Forecasted Output	Error Value
1	97	98.9	-1.9
2	98	98.8	-0.8
3	98	98.8	-0.8
4	99	99.0	0.0
5	100	99.0	1.0

Graphical representation of Table IV is given below in fig. 4.

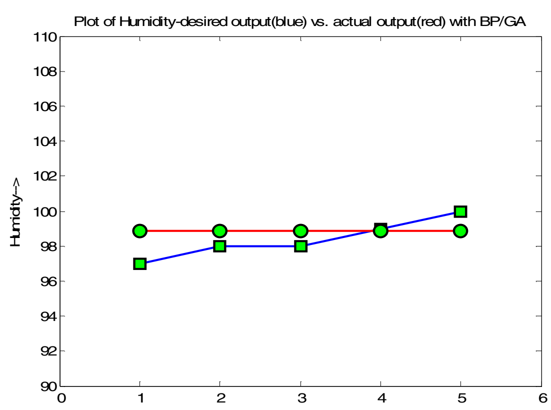

Figure 4. The 5-day-relative humidity prediction for the BP/GA(red) technique and the desired output(blue).

The error values corresponding to daily rainfall parameter is shown in table V along with the desired output and the forecasted output for the BP/GA technique.

TABLE V
RAINFALL FOR BP/GA

Day	Desired Output	Forecasted Output	Error Value
1	1.9	2.0	-0.1
2	2.1	2.3	-0.2
3	2.3	2.3	0.0
4	2.5	2.3	0.2
5	3.0	2.4	0.6

The above values are represented graphically in fig. 5. It shows the desired and forecasted outputs of the mean air temperature corresponding to five days for the month of January 2009.

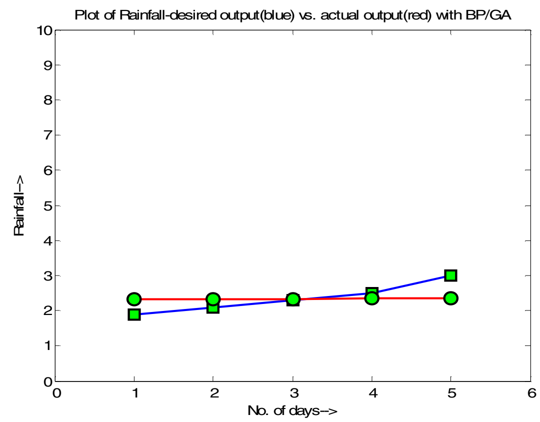

Figure 5. The 5-day-daily rainfall prediction for the BP/GA(red) technique and desired output(blue).

The comparison of proposed integrated BP/GA technique with the traditional gradient descent based back propagation algorithm is done for the same five days data for the month of January 2009. The results of the comparison for all the three parameters- mean air temperature, relative humidity and daily rainfall are illustrated in fig. 6, 7and 8 respectively. Both the techniques are compared with the desired outputs.

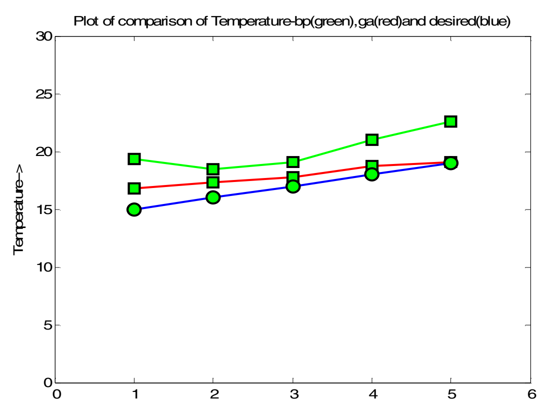

Figure 6. The comparison of the BP/GA technique (red) and gradient descent based BP (green) with the desired output(blue) for temperature.

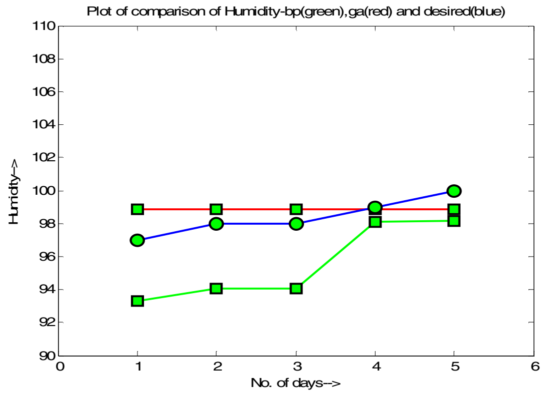

Figure 7. The comparison of the BP/GA technique (red) and gradient descent based BP (green) with the desired output(blue) for humidity.

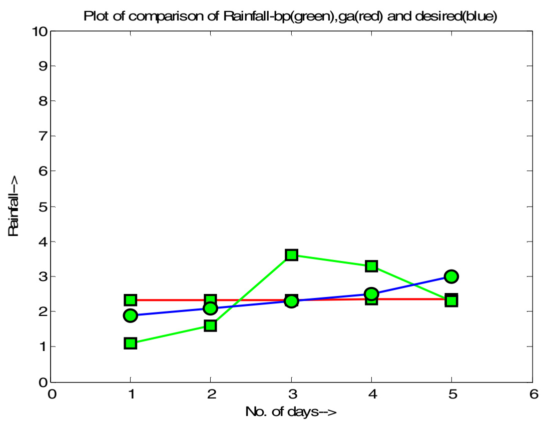

Figure 8. The comparison of the BP/GA technique (red) and gradient descent based BP (green) with the desired output (blue) for rainfall.

The above comparison shows clearly that the integrated BP/GA technique is more suitable to predict weather than the traditional gradient based back propagation algorithm because in all cases- temperature, humidity and rainfall; proposed BP/GA technique is more close to the desired output than the back propagation algorithm.

VI. CONCLUSION

From the analysis above, it is easy to observe the compensability between back propagation algorithm and genetic algorithm. The proposed BP/GA technique can learn efficiently by combining the strengths of genetic algorithm with back propagation. The proposed approach is more qualified for neural networks while training them for weather forecasting data if only the requirement of a global searching is considered.

It works with a population of points instead of a single point. Also it blends the merits of both deterministic gradient based algorithm and stochastic optimizing algorithm. By using local gradient information advantageously, the BP/GA is more speed efficient than genetic algorithm alone. Hence the use of the integrated BP/GA technique in weather forecasting system is suggested.

REFERENCES

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

[2] David J. Montana and Lawrence Davis, "Training Feedforward Neural Networks Using Genetic Algorithms", BBN Syst. and Tech. Corp. 10 Mouiton St. Cambridge, MA 02138

[3] Azadeh A., et al., "Integration of ANN and GA to Predict Electrical Energy consumption", Dept. of Industrial Eng. and Research Inst. of Energy Manag. and Planning, Univ. Tehran Iran Islamic Azad Univ., IEEE 2552, 2006.

[4] R. Rojas "Neural Networks- The Backpropagation Algorithm", Springer-Verlag, Berlin, 152, 1996.

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

[6] Holland, J. H., "Adaptation in Natural and Artificial Systems", University of Michigan Press. Ann Arbor, MA, 1975.

[7] Schalkoff, R.J., "Artificial Neural Networks", New York: McGraw-Hill, 1997.

[8] Maqsood I., Khan M.R. and Abraham A., "Weather Forecasting Models Using Ensembles of Neural Networks", University of Regina, Regina, SK S4S 0A2, Canada, 2004.

[9] Rao H. Sudarsana, Ghorpade Vaishali G., Mukherjee A. , "A genetic algorithm based back propagation network for simulation of stress-strain response of ceramic-matrix-composites", Computers and Structures, Pergamon Press, Inc, Jan. 2006 vol 84.

[10] Abraham Ajith et. al., "Soft Computing Models for Weather Forecasting" 2002

[11] Chun Lu, Bingxue Shi, "Hybrid back propagation/ GA for Multilayer Feedforward Neural Networks", Inst. of Microelectronics, Tsinghua University Beijing 100084, China IEEE, 2000.

[12] Ileană Ioan, Corina Rotar, Arpad Incze, "The Optimization of Feed-forward Neural Networks Structure using Genetic Algorithms", Proc. 2004 Int. Conf. on Theory and Applicat. of Math. and Informatics, Thessaloniki, Greece, 2004.

[13] Sarangi1 Pradeepta et. al., "Short Term Load Forecasting using Artificial Neural Network: A Comparison with Genetic Algorithm Implementation", ARPN J. of Eng. and App. Sci., Nov. 2009, vol. 9.