International Journal of Emerging Technologies in Engineering Research (IJETER)
Volume 3, Issue 1, October (2015) www.ijeter.everscience.org

An Efficient Neural Networks based Genetic Algorithm Model for Soil Temperature Prediction

Jasmeen Gill
Research Scholar, IKG PTU, Kapurthala, India.

Shaminder Singh
Research Scholar, IKG PTU, Kapurthala, India.

Abstract –Suitable soil temperature predictions can help the farmers and producers in providing valuable information for deciding the right time for crop cultivation and harvesting. Due to non-linearity in climatic physics, neural networks are suitable to predict these meteorological processes. Back Propagation is the most important algorithm to train a neural network. It is a systematic gradient descent method that suffers from local minima problem, scaling problem, long training times, etc. In this article, an efficient soil temperature prediction model is proposed using neural network based genetic algorithm technique to solve these problems. The results are very encouraging.

Index Terms – Neural Networks, Back propagation Algorithm, Genetic Algorithm, Soil Temperature, Hybrid Model.

1. INTRODUCTION

Soil temperature prediction is important as soil temperature plays an important role in various ecosystems from deserts to forests. Soil temperature affects infiltration of water through soil surface. Soil temperature also influences the biological process of plants, insects and other organisms. Degradation of pesticides and bio-degradation processes in soil are greatly affected by soil temperature. Hence it has become necessary to develop accurate and efficient soil temperature prediction models .Earlier models were based on mathematical equations and conceptual theory to simulate the fluctuations of soil temperature at different depths [4]. They involve complex calculations and were very time consuming. 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 forecasting. ANNs are simplified model of biological nervous system and have taken the motivation from kind of computations performed by human brain. Neural Networks have characteristics like pattern matching, object recognition, mapping capabilities and high speed information processing [6].

ANNs learn by examples so they can be trained with known examples. One of the most popular training algorithms in the domain of neural networks is the back propagation algorithm. It is a gradient descent method. The algorithm suffers from several problems, discussed later in this article [1]. 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 [2], [3], [4] and [12]. But, in the field of soil temperature forecasting, such efforts are still to be put up. So, the motivation of this work is firstly, to develop an efficient soil temperature prediction model using neural network based genetic algorithm approach and secondly to compare the proposed model with the previous one [6].

The remainder of the article is organized as follows: Section 2 introduces the artificial neural networks along with a brief description of the algorithms under analysis. The details of proposed approach for soil temperature prediction model are given in Section 3, followed by results in Section 4. Finally, the conclusions are summarized in Section 5.

2. 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 [12].

Simulating the structure of brain, neural networks are highly interconnected processing elements represented by a graph G consisting of V set of vertices & E set of edges. The vertices represent neurons and edges represent synaptic links i.e. weights. NN architecture is broadly classified as Single-layer feed forward network and Multi-layer feed forward network. Single-layer feed forward network comprises of two layers-Input layer to receive signals and output layer to give response. Multi-layer feed forward network comprises input layer to receive signals, hidden layer for performing intermediary computations and output layer to give response [11].

2.1. 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. As the gradient of the error function is to be calculated, it should be continuous and differentiable. Unlike perceptron where step function is used, the sigmoid function is used as an activation threshold for the network [12].

The merits of BP 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 any network 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 [13].

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, BPN suffers from the scaling problem. It works well on simple training problems. However, as the problem complexity increases, the 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 [9]. 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 [7].

2.2. Genetic Algorithms

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. Fitness is defined as a figure of merit which is to be either maximized or minimized. An initial population of chromosomes (set of strings) is taken to generate offspring (from fit parents) that competes for survival to make up the next generation of population [9].

Three main inheritance operators used are reproduction, crossover and mutation. Reproduction is the first operator applied on a population for selecting good chromo in a population to form the mating pool. Crossover is applied next. Here a better offspring is produced to combine the good substring from either parent. Mutation changes one of the elements of substring with a small probability to keep the diversity of population. Successive generation of chromosome improve in quality provided that the criteria used for survival is appropriate. This process is referred to as Darwinian natural selection or survival of the fittest [3].

As compared with BP, 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 [5]. Secondly, Genetic Algorithms work with a string coding of variables instead of the variables. The advantage of working with a coding of variables is that coding discreteness the search space even though the function may be continuous. Thus, a discrete function can be handled with no extra cost. Also Genetic Algorithm can be applied to a wide variety of problems. We always get the solution & the solution gets better with time [10].

The price one pays for Genetic Algorithm is its slowness. The slowness is mainly due to the slow but crucial exploration mechanisms employed, which has three basic arithmetic operators: reproduction, crossover and mutation. In addition, Genetic Algorithm starts searching from random genes, which will cost a lot of time [5], [7].

3. MATERIALS AND METHOD

The proposed soil temperature prediction model based on BP/GA technique starts with the collection of data, feature extraction, data normalization and finally, formation of training data set (containing inputs and outputs) and testing data set (containing inputs).

3.1. 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 2014. The data collected are carefully analyzed and are required to be processed first to attain useful information from it. It was found during the analysis that the soil temperature parameter cannot be predicted using its features alone. Due to its dependence on other parameters like air temperature and daily rainfall, data regarding these parameters was also collected.

3.2. 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 network may cause convergence problem. All the weather related data sets were, therefore, transformed into values between 0 and 1 using the following formula [8].

$$d_{norm} = \frac{d_t - d_{min}}{d_{max} - d_{min}}$$

Where, dt is the actual data value, dmin is the minimum data value, dmax is the maximum data value and dnorm is the normalized data value.

3.3. Feature Extraction

In certain cases, the weather parameter to be forecast can be estimated based on the features obtained from the past data of same parameter. However, in some cases the parameter to be forecast exhibits a strong dependence on other weather parameters also. In such cases the input to the network should include the features extracted from other weather parameters also. As mean soil temperature depends on its past data and the present day data of air temperature and rainfall, so the features extracted are day number, moving average, oscillator of mean soil temperature and moving average of mean air temperature and daily rainfall parameter of that day.

Parameter	Inputs	Outputs
Features of soil temperature	Day number	Mean Soil Temperature
	Moving Average	
	Oscillator	
Features of depe - ndent parameters	Moving Average (temperature)	
	Moving Average (daily rainfall)	

Table 1 Extracted features for Soil Temperature

3.4. Proposed Approach

The idea is to integrate the two approaches. Chromosomes form the initial population. Chromosomes are coded into real coding system in which every chromosome is a string of some values. Every gene is encoded of a randomly chosen string length. Weights are assigned to each chromosome randomly. The network is trained for first cycle and the error is calculated. The search for selecting an individual is guided by a fitness function meant to evaluate the quality of each individual. The efficiency of a genetic algorithm is connected to the ability of defining a “good” fitness function [7].

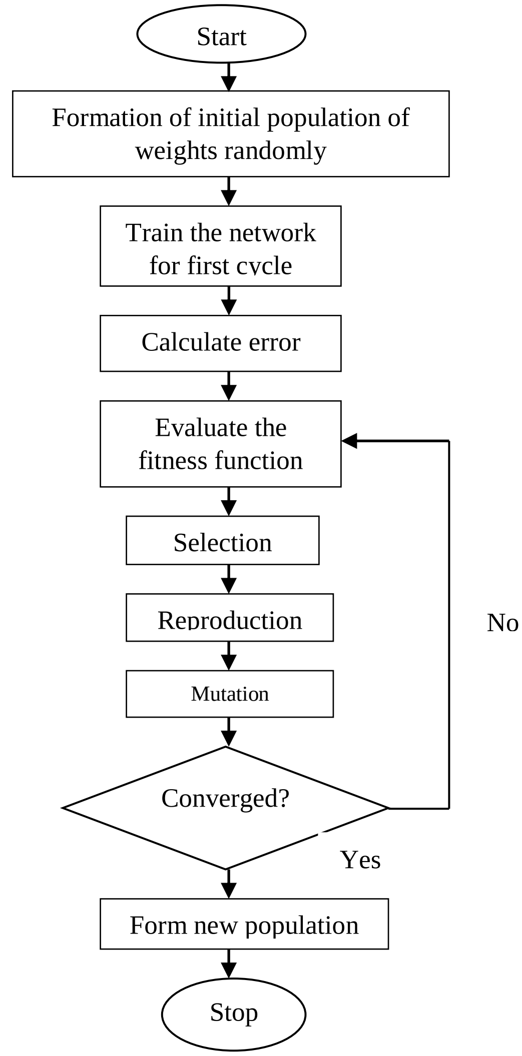

Figure 1 The Proposed Algorithm

More is the fitness value; more are the chances of a chromosome to be selected for reproducing an offspring. So, fitness function is evaluated from the error value. Then the three operators- reproduction, crossover and mutation are applied one after the other. After the calculation of fitness of each individual in the population, individuals to act as parents are randomly selected. Then a mating pool is prepared by replacing the individual with minimum fitness value by individual having maximum fitness value. The population is enriched with better individuals after reproduction phase is over. Afterwards the individuals are mutated (if required). The process of generating populations continues until majority of the chromosomes acquired same fitness value. The final weights calculated are used by the back propagation to train the network. The whole process is described in figure 1. The proposed technique solves the problem of local minima as stochastic optimizing and global searching genetic algorithm is integrated with the back propagation. Furthermore, it is easy to be implemented by hardware.

4. RESULTS AND DISCUSSIONS

The proposed BP/GA technique is implemented using the MATLAB tool (The Mathworks Inc.). The network is trained for soil temperature values obtained at a depth of 5 cm, 10 cm and 30 cm. The error values corresponding to soil temperature at 5 cm depth shown in table 2 along with the desired output and the forecasted output for the proposed BP/GA and gradient-BP technique.

Day	Desired Output	Forecasted Output (BP/GA)	Error Value	Forecasted Output (Gradient-BP)	Error Value
1	20.5	20.9	-0.4	21.4	-0.9
2	22.6	22.0	0.6	20.3	2.3
3	22.1	22.2	-0.1	20.0	2.1
4	24.0	24.2	-0.2	24.2	-0.2
5	26.0	26.0	0.0	25.2	0.8

Table 2 Soil Temperature (at 5 cm depth)

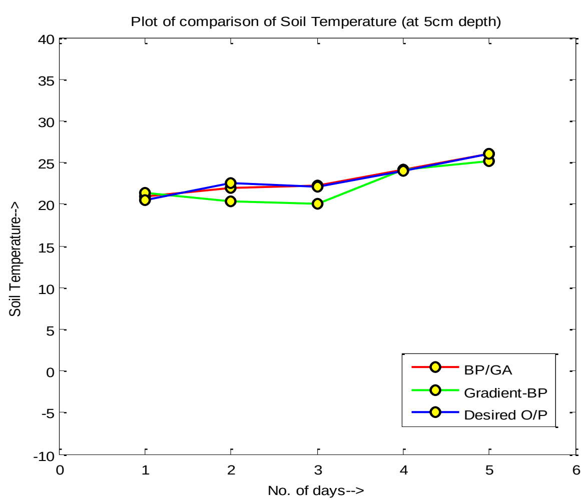

Figure 2 5-day-soil temperature predictions (at 5 cm depth)

These error values are calculated as desired output – actual output and are represented graphically in figure 2. Table 3 shows the prediction of soil temperature at 10 cm depth with its graphical representation in figure 3.

Day	Desired Output	Forecasted Output (BP/GA)	Error Value	Forecasted Output (Gradient-BP)	Error Value
1	20.5	20.1	0.4	18	1.5
2	22.3	22.4	-0.1	23.5	-1.2
3	20.8	21.0	-0.2	22.0	-1.2
4	22.3	22.3	0.0	23.0	-0.7
5	23.5	23.7	-0.2	25.0	-1.3

Table 3 Soil Temperature (at 10 cm depth)

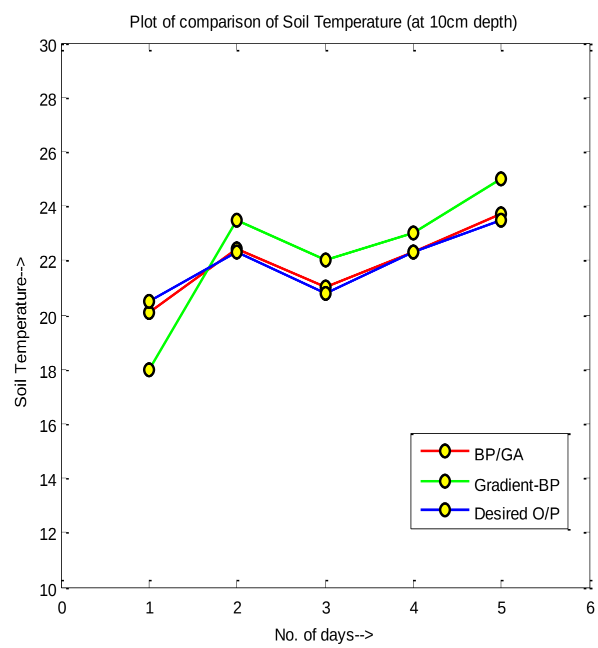

Figure 3 5-day-soil temperature predictions (at 10 cm depth)

Forecasted soil temperature values at 30 cm depth are shown in table IV and the graphical representation for the same is given in figure 4.

The error value at 30 cm depth shown the better results, when we compare these results with other cases discussed earlier. So results will improve when we will increase the domain of the problem.

Day	Desired Output	Forecasted Output (BP/GA)	Error Value	Forecasted Output (Gradient-BP)	Error Value
1	22.0	21.9	0.1	20.0	2.0
2	23.0	23.0	0.0	21.0	2.0
3	21.6	21.4	0.2	22.0	-0.4
4	21.5	21.6	-0.1	22.0	-0.5
5	21.8	21.9	-0.1	21.0	0.8

Table 4 Soil Temperature (at 30 cm depth)

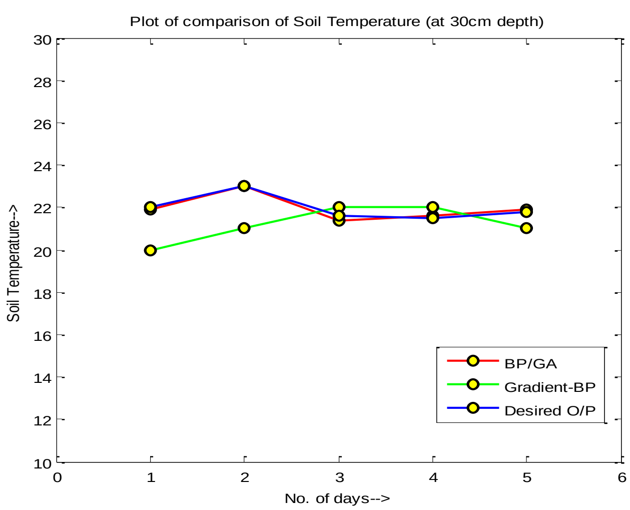

Figure 4 5-day-soil temperature prediction (at 30 cm depth)

5. CONCLUSION

The proposed BP/GA technique can learn efficiently by combining the strengths of GA with BP. 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. By using local gradient information advantageously, the BP/GA is more speed efficient than GA. Hence the use of BP/GA is proposed.

REFERENCES

[1] A Azadeh , S.F. Ghaderi, S.Tarverdian and F.Saberi, "Integration of ANN and GA to predict electrical energy consumption"in Proc. 32nd IEEE Conference on Industrial Electron., Paris, France, vol. 7, no. 4,2006 , pp. 2552–2557.

[2] J.Gill , B.Singh. and S.Singh , "Training Back Propagation Neural Networks with Genetic Algorithm for Weather Forecasting", 8th IEEE SISY, Subotica Serbia, 2010, pp 465–469.

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

[4] A.S.Jebamalar , "Prediction of Annual and Seasonal Soil Temperature Variation using Artificial Neural Network", Indian J. of Radio & Space Physics", vol. 41, 2012, pp. 48–57.

[5] C.Lu. and B.Shi, "Hybrid back propagation/GA for Multilayer Feed forward Neural Networks", Institute of Microelectronics, Tsinghua Univ., China, 2000.

[6] I.Maqsood, M.R.Khan and A. Abraham, "Weather forecasting models using ensembles of neural networks", Int. Conf. on Intel. Sys. Design and Applications, Germany, 2003, pp. 33–42.

[7] D.J.Montana and L.Davis, "Training Feed forward Neural Networks Using Genetic Algorithms", BBN Systs and Tech. Corp. MA.

[8] H.S.Rao, V.G.Ghorpade, A Mukherjee, "A genetic algorithm based back propagation network for simulation of stress-strain response of ceramic-matrix-composites", Computers and Structures, vol. 84, Issue 5-6, 2006.

[9] S.Rajasekaran and P. Vijayalakshmi, "Neural Networks, Fuzzy Logic and Genetic Algorithms", Prentice Hall of India, New Delhi, 2004, pp. 253–265.

[10] R.Rojas, "Neural Networks- The Back propagation Algorithm", Springer-Verlag, Berlin, 1996, pp. 152.

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

[12] S. Singh ,P. Bhambri and J. Gill, "Time Series based Temperature Prediction using Back Propagation with Genetic Algorithm Technique", International Journal of Computer Science Issues, vol. 8, Issue 5, no. 3, pp. 28-32,2011.

[13] S. Singh and J. Gill, "An optimized Neural Network Model for Relative Humidity Prediction", in 'International Journal of Research in Advent Technology', vol. 2, no. 4, pp. 381-385,2014.

Authors

Jasmeen Gill received her B.Tech. and M.Tech. degrees in Computer Science from Punjab Technical University, Jalandhar, in 2006 and 2011, respectively. Presently she is pursuing Ph.D. in Computer Science from IKG Punjab Technical University, Kapurthala.

She has an experience of more than 10 years in teaching and presently working as Assistant Professor at RIMT –IET, Mandi Gobindgarh. She has published more than 20 research articles at national as well as international levels including IEEE, SCI Journals. Her main areas of interest include Artificial Intelligence, Image Processing, Neural networks, Genetic algorithms, Machine Learning and Pattern recognition applications.

Shaminder Singh received his B.Tech. and M.Tech. degrees, from Punjab Technical University, Jalandhar, in 2005 and 2013, respectively. Presently he is pursuing Ph.D. in Computer Science from IKG Punjab Technical University, Kapurthala.

He has an experience of more than 10 years in teaching and presently working as Assistant Professor at GGI, Khanna. He has published more than 17 research articles at national as well as international levels including IEEE, SCI Journals. He is author of book entitled "Temporal weather prediction using Genetic algorithm" by OLAP available at amazon.in His main areas of interest include Artificial Intelligence, Fuzzy Logics, Neural networks, Soft Computing Techniques, Machine Learning and Pattern recognition applications.