An Optimized Neural Network Model for Relative Humidity Prediction

Shaminder Singh[1], Jasmeen Gill[2]
cse[1], cse[2],Research Scholar PTU[1], Research Scholar PTU [2]
Email: er_shaminder@yahoo.co.in[1], er.jasmeengill@gmail.com[2]

Abstract- Weather forecasting is the application of science and technology to predict the state of the atmosphere for a given location. Weather forecasts are made by collecting quantitative data about the current state of the atmosphere on a given place and using scientific understanding of atmospheric processes to project how the atmosphere will evolve on that place. Accurate forecasting is important in today's world as agricultural is largely dependent on weather. 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 humidity on a particular data series, a humidity 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.

Index Terms- Artificial Neural Networks, Back Propagation Algorithm, Genetic Algorithms.

1. INTRODUCTION

Forecasts based on humidity important to agriculture, and therefore to traders within commodity markets. Humidity forecasts are used by utility companies to estimate demand over coming days. On an everyday basis, people use weather forecasts to determine what to wear on a given day. Since outdoor activities are severely curtailed by heavy rain, snow and the wind chill, forecasts can be used to plan activities around these events, and to plan ahead and survive them. This is done to protect life and property.

Forecasting is a phenomenon of knowing what may happen to a system in the next coming time periods [1]. 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 [4]. As weather is a continuous, data-intensive and dynamic process, the parameters required to predict humidity are enormously complex such that there is uncertainty in prediction even for a short period [3]. These properties make humidity 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 humidity forecasting.

1.1. *Weather Forecasting*

Generally, two methods are used to forecast weather (a) the empirical approach and (b) the dynamical approach [5], [6]. The first approach is based upon the occurrence of analogs and is referred to as analog forecasting. This approach is useful for predicting local-scale weather if recorded cases are plentiful. The second approach is based upon equations and simulations of the atmosphere, and is referred to as computer modeling or Numerical Weather Prediction

that predicts large-scale weather phenomena efficiently. In this technique the atmosphere is considered to be a fluid and the state of the fluid is estimated at a given time using the equations of fluid dynamics and thermodynamics. These calculations were very complex and must be done for every single box in the grid – a huge amount of work. This approach is only useful for modeling large-scale weather phenomena and may not predict short-term weather efficiently. Most weather prediction systems use a combination of empirical and dynamical techniques.

At macro level, weather forecasting is usually done using the data and images taken by remote sensing satellites to assess future trends. These systems are inherently costlier, require complete support system and provide information generalized over a larger geographical area [7]. However, a little attention has been paid to the use of artificial neural networks in weather forecasting.

1.2. *Neural Networks*

Artificial neural network can be defined as: "A computing system made up of a number of simple, highly interconnected processing elements, which process information by their dynamic state response to external inputs". Various learning methods exist to train neural networks. The learning mechanisms are categorized as- Supervised learning and unsupervised learning. In supervised learning, output set along with input set is fed to the network during the learning process. However in unsupervised learning process, no output set is available and the network learns of its own.

Neural networks are suitable for solving non-linear problems that are difficult to be solved by the traditional techniques. Most meteorological processes often exhibit temporal and spatial variability, and are further plagued by issues of nonlinearity of physical processes, conflicting spatial and temporal scale and uncertainty in parameter estimates. Thus, these properties of artificial neural networks are well suited to the problem of weather forecasting under consideration [6].

The major reasons of using artificial neural networks in the field of weather forecasting are: they would reduce the computational power required to accurately predict atmospheric variables from a supercomputer to a single neural network, a large database is available of historical weather data which can be used as training sets, Neural networks find the best generalized network to solve patterns outside their training sets, Neural networks can detect and utilize hidden patterns in data to arrive at a solution, and finally, Neural networks have been shown to accurately predict irregular and complex variables in past work [11].

1.3. *Back Propagation Algorithm*

The merits of back propagation 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 & doesn't 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.

In the field of neural networks weather forecasting is done by training the network through back propagation algorithm. Back propagation is a systematic method of training multilayer artificial neural networks. 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 [10].

1.4. *Genetic Algorithms*

The key feature of such algorithms is characterized by possessing a chromosome. A chromosome is composed of strings of symbols called bits. Each production of GA makes a new population of the existing type. Suppose that the population size is P initially. P individuals are assigned values to their chromosomes, where the assignment can be either random or deterministic. A permutation of such strings can be introduced to construct a population of designs which each design has its own fitness value [2].

The pros and cons of genetic algorithms are as follows: As compared with back propagation, genetic algorithm is more qualified for neural networks as it is good at global searching (not in one direction) and it works with a population of points instead of a single point. Secondly, genetic algorithms work with a string coding of variables instead of the variables which requires only function values at discrete points, a discrete function can be handled with no extra cost [8]. Also inherently parallel nature of genetic algorithms makes the processing faster as compared to the back propagation algorithm [9]. Another merit of genetic algorithm is that it is easy to be implemented by hardware. First of all, the required precision is not high. Second, if binary encoding is adopted, the results can be directly reflected to digital storage. The last, the arithmetic operation is simple, which is quite favorable for hardware implementation [12].

1.5. *Hybrid Techniques*

In order to solve the problems of back propagation algorithm, efforts have been made to integrate it with the genetic algorithms [25].

In the integrated BP/GA technique, the minimization of error through gradient descent method is replaced by the minimization of the error value or maximization of the fitness value through the three basic operators used in the genetic algorithms. These operators are the selection, crossover and mutation operators, which are applied in a sequence one after the other over the population of individuals.

2. PROPOSED TECHNIQUE

It is clear from the problem formulation part that a neural network can be trained with the help of back propagation algorithm based genetic algorithm integrated technique. So, the present work involves training a back propagation network through genetic algorithms so as to blend the merits of both the techniques in weather forecasting . The proposed technique will try to solve the problems incurred in the hybrid BP/GA technique and provide a complete and accurate humidity prediction model.

2.1. *Collection of Data*

The first step in the design of weather forecasting model is to collect weather related data obtained through various instruments like thermometer, barometer etc. The data used in this research are the daily weather data. The data in the un-normalized form have been collected.

2.2. *Parameters used for prediction*

The parameters chosen in this setup for the prediction 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 1. Weather Parameters

Sr. No.	Meteorological Variables	Unit
1.	Mean Air Temperature	ºC
2.	Rainfall	mm
3.	Relative Humidity	%

2.3. *Application of N-Sliding Window*

Work in neural networks has concentrated on forecasting future developments of the time series from values of x up to the current time. 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 [1]. Fig. 1 gives the basic architecture.

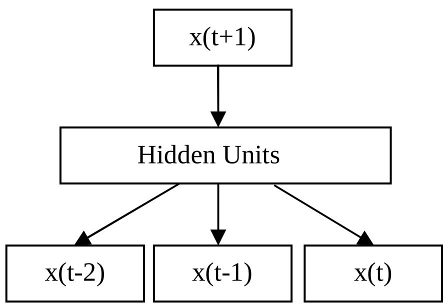


Fig. 1. Time Series Prediction using N-Sliding Window Technique

This method is often called the sliding window technique as the N-tuple input slides over the full training set which is mathematically known as the moving average and is calculated progressively as an average of N number data values over the certain period. For a data set is represented by d_t, d_{t-1}, d_{t-2} ,………………., d_0, where d_t is present and d_0 is the first data value, the moving average with a sliding window of period N is

$$MA_N = \frac{d_t + d_{t-1} + d_{t-2} + \ldots\ldots\ldots + d_0}{N} \qquad \ldots(2.3)$$

2.4. *Feature Selection*

At this step the various features for predicting relative humidity are extracted from data and fed to the neural network model as input. Given below is the table for extracted features from relative humidity data and features extracted from temperature as well as rainfall parameters. Temperature and rainfall are two such parameters over which humidity is highly dependent. Hence the choice is made.

Table 2. Features for estimating relative humidity

Parameter	Inputs	Output
Features of Relative Humidity	• Day number • Moving Average • Oscillator	Relative Humidity
Features of dependent parameters	• Moving Average (temperature) • Moving Average (rainfall)	

2.5. *Data Normalization*

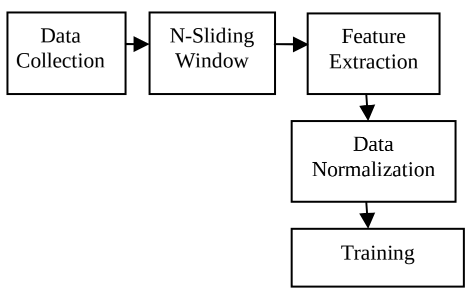
After the collection of data and selection of the weather parameters, next issue is the normalization of data. Neural networks generally provide improved performance with normalized data. The use of original data to network may cause convergence problem. All the weather data sets were, therefore, transformed into values between 0 and 1 by dividing the difference of actual value d_t and minimum value d_{min} by the difference of maximum value d_{max} and minimum value [6] to obtain the normalized data d_{norm} as follows:

$$d_{norm} = \frac{d_t - d_{min}}{d_{max} - d_{min}} \qquad \ldots. (2.4)$$

Thus, the gradient, which is a function of the derivative of the nonlinearity, will always be different from zero. At the end of the algorithm, the outputs are de-normalized into the original data format for achieving the desired result.

2.6. *Methodology*

The main steps in developing relative humidity prediction model are shown in fig. 2. After data normalization step, features are fed to train the network using supervised training method. When the network has learned appropriately, testing is performed where actual prediction is done.



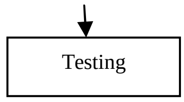


Fig. 2. Steps to predict relative humidity

3. RESULTS

The error values corresponding to relative humidity are shown below in table 3. The error values a[redacted] calculated as the absolute difference of desired outp[redacted] and the forecasted ones for 5-days. Along with this[redacted] comparison of the hybrid and back propagation is al[redacted] shown.

Table 3. Forecasted values of relative humidity

Day No.	DO	Hybrid Technique		Back Propagation	
		FO	Error Value	FO	Error Value
1	65	57.3	7.7	64.8	0.2
2	67	68.5	1.5	71.3	4.3
3	62	62.8	0.8	71.6	9.6
4	72	70.6	1.4	71.6	0.4
5	65	61.4	3.6	71.7	6.7

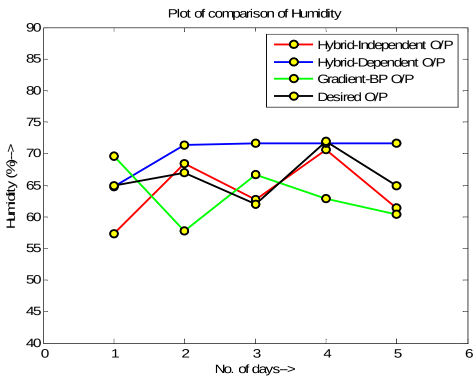


Fig. 3. Humidity prediction for the hybrid technique vs. BP and desired output

The above comparison shows clearly that the hybrid model is more suitable to predict weather than the traditional gradient based back propagation algorithm GA techniques are more close to the desired output than the back propagation algorithm.

In other words, the overall accuracy of the proposed relative humidity prediction model based on back propagation algorithm with genetic algorithm technique is suitable to predict relative humidity than the gradient based back propagation technique.

3.1. *Effects of data series and under training and over training the model*

The error values corresponding to relative humidity are shown in figure 4 along with the series 1, series 2 and series 3. 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.

[redacted]

The above shown figures clearly depict the effects of under training and over training the model as the error values are minimum for the case of 400 epochs. Secondly variation in data series as input too has different effects on output values. Different data series have different error values.

4. CONCLUSION AND FUTURE SCOPE

The author believes that a reasonable prediction accuracy rate could be achieved with this methodology given a larger training set, using faster and better training algorithms, and more known atmospheric values. This objective will obviously be the goal of future work in this area of research. However, it is the author's belief that higher accuracy rates can be achieved by following the below given alternatives:

- The predictions could be done based on the seasonal division i.e. dividing the model into four main seasons- summer, winter, autumn and rainfall.
- Larger training sets of previous years could be used to improve the performance of the system.
- This system could be further extended to include more weather parameters like sunshine, wind speed and evaporation.

REFERENCES

[1] Frank R. J., Davey N. and Hunt S. P. (2008): Time Series Prediction and Neural Networks, Deptt. of Comp. Sci., University of Hertfordshire, Hatfield, UK, vol. 31, pp 91-103.

[2] Goldberg D. E. (2008): Genetic Algorithms in Search, Optimization and Machine learning, ed. 3rd, published by Dorling Kindersley Pvt. Ltd. with Pearson Education Inc., pp. 120-125.

[3] Herrera J. L. (1999): Time series prediction using inductive reasoning techniques, doctoral thesis from Universidad Poliytechnica De Cataluna.

[4] Farahat A.M. and Talaat M. (2012): Short-Term Load Forecasting Using Curve Fitting Prediction Optimized by Genetic Algorithms, International Journal of Energy Engineering , vol. 2, no. 2, pp 23-28 .

[5] Maqsood I., Khan M.R. and Abraham A. (2002): Neuro-computing based Canadian weather analysis, 2nd International Workshop on Intell. Sys. Design and Applications, Atanta, Georgia, pp. 39–44.

[6] Maqsood I., Khan M.R. and Abraham A. (2003): Weather forecasting models using ensembles of neural networks, Proc. 3rd International Conference on Intell. Sys. Design and Applications, Germany, pp. 33-42.

[7] Paras, Mathur S., Kumar A. and Chandra M. (2007): A feature based neural network model for weather forecasting, Journal of World Academy of Sciences, Engg. and Tech., vol. 34,pp. 66-73.

[8] Rajasekaran S and Vijayalakshmi P. (2004), "Neural Networks, Fuzzy Logic and Genetic Algorithms, published by Prentice Hall of India, New Delhi, pp. 253-265.

[9] Rao H. S., Ghorpade V. G. and Mukherjee A. (2006): A genetic algorithm based back propagation network for simulation of stress-strain response of ceramic-matrix-composites, Journal of Computers and Structures, vol. 84, no. 5-6, pp. 330-339.

[10] Rojas R. (1996): Neural Networks- A Systematic Introduction, published by Springer-Verlag, Berlin, no. 3-5, pp. 152-187.

[11] Schwaiger (1996): Evolving topologies of artificial neural networks adapted to image processing tasks, Proc. 26th International Symposium on Remote Sensing of Environment, Vancouver, BC, pp. 71-74.

[12] Srinivas M. and Patnaik L.M., (1994): Genetic algorithms: a survey, IEEE Transactions on Computation, vol. 27, no. 6, pp. 17-26.