International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

Development of Artificial Neural Network Models for Estimation of Yield of Cotton

Pranav Mistry[1], Dr. T. M. V. Suryanarayana[2]

[1]PG Student, Water Resources Engineering and Management Institute (WREMI), Faculty of Technology and Engineering, The Maharaja Sayajirao University of Baroda, Samiala-391410

[2]Associate Professor, Water Resources Engineering and Management Institute (WREMI), Faculty of Technology and Engineering, The Maharaja Sayajirao University of Baroda, Samiala-391410

Abstract: ***In this paper, Artificial Neural Network (ANN) models were developed to estimate the yield of cotton. The climatological parameters considered to predict the yield of cotton were maximum temperature, minimum temperature, wind velocity and relative humidity and sunshine hours. An attempt has been made to develop the ANN model with the best network architecture, to predict the yield of cotton, considering the above given climatological parameters, as input. The models performance were evaluated using performance indices such as coefficient of correlation (r), coefficient of determination (R^2) and discrepancy ratio (D.R). From the results, it can be concluded that the best fit ANN model is generalized regression model, with r, R^2 and D.R. as 0.9770, 0.9545 and 1.003 during training and 0.9287, 0.8624 and 0.9977 during validation respectively.***

Keywords: Artificial Neural Network, MLP Modeling, Performance Indices, Crop Yield, Climatological Data.

1. Introduction

Artificial Neural Networks (ANNs) are computational models inspired by an animal's central nervous systems (in particular the brain) which is capable of machine learning as well as pattern recognition. Artificial neural networks are generally presented as systems of interconnected "neurons" which can compute values from inputs. [1]

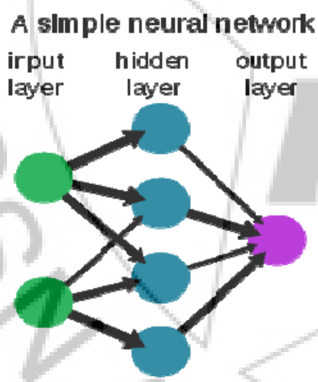

Figure 1: A structure of a three-layer feed-forward perceptron.

Maind and Wankar (2014) [2] gave overview of Artificial Neural Network, working & training of ANN. ANN is mathematical model or computational model, an information processing paradigm i.e. inspired by the way biological nervous system, such as brain information system. As shown in Fig.1, ANN is made up of interconnecting artificial neurones which are programmed like to mimic the properties of m biological neurons. These neurons working in unison to solve specific problems. ANN is configured for solving artificial intelligence problems without creating a model of real biological system. ANN is used for speech recognition, image analysis, adaptive control etc. These applications are done through a learning process, like learning in biological system, which involves the adjustment between neurones through synaptic connection. Same happen in the ANN. Climatic variability is the major factor influencing the agriculture productivity. Global climate change and its impacts on agriculture have becoming an important issue. Agriculture production is highly dependent on climate and it also adversely affected by increasing climatic variability. The aim is to develop the methodology for assessing this component of the total impact of climate variability on agricultural productivity. There is a need to quantify climatic variability to assess its effect on crop productivity. Parekh and Suryanarayana (2012) [3] carried out the study to determine the predominance of various meteorological data on yield of Wheat. Panchal et al. (2013 [4] carried out the Rainfall-Runoff modeling using ANN at Dharoi sub-basin". Suryanarayana and Parekh (2012) [5] carried out the study to develop a weather based model on yield of cotton using Neural Network Fitting Tool for various sizes of data set.

2. Study Area and Data collection

The Area selected for the present study is Vadodara, Gujarat, India. In this study, long term meteorological data from 1981 to 2006 comprising of maximum temperature, minimum temperature, relative humidity, wind speed, sunshine hours and yield of cotton has been collected.

3. Working of ANN

The other parts of the ―art of using neural networks revolve around the myriad of ways these individual neurons can be clustered together. This clustering occurs in the human mind in such a way that information can be processed in a dynamic, interactive, and self-organizing way. Biologically, neural networks are constructed in a three-dimensional world from microscopic components. These neurons seem capable of nearly unrestricted interconnections. That is not true of any proposed, or existing, man-made network. Integrated circuits, using current technology, are two-dimensional devices with a limited number of layers for interconnection.

This physical reality restrains the types, and scope, of artificial neural networks that can be implemented in silicon. Currently, neural networks are the simple clustering of the primitive artificial neurons. This clustering occurs by creating layers which are then connected to one another. How these layers connect is the other part of the "art" of engineering networks to resolve real world problems.

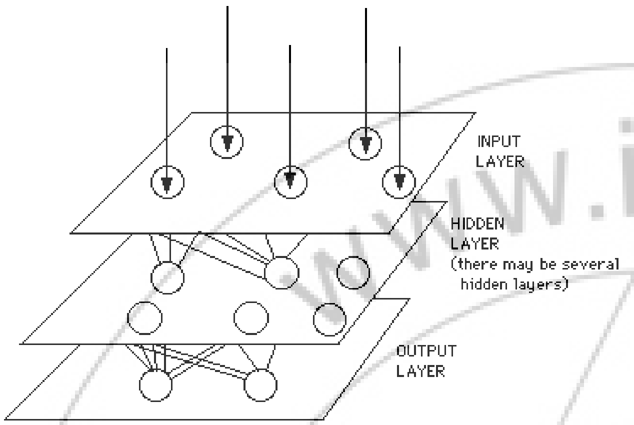

Figure 2:- A Simple Neural Network Diagram.

Basically, all artificial neural networks have a similar structure or topology as shown in Fig. 2. In that structure some of the neurons interface to the real world to receive its inputs. Other neurons provide the real world with the network's outputs. This output might be the particular character that the network thinks that it has scanned or the particular image it thinks is being viewed. All the rest of the neurons are hidden from view. But a neural network is more than a bunch of neurons. Some early researchers tried to simply connect neurons in a random manner, without much success. Now, it is known that even the brains of snails are structured devices. One of the easiest ways to design a structure is to create layers of elements. It is the grouping of these neurons into layers, the connections between these layers, and the summation and transfer functions that comprises a functioning neural network. The general terms used to describe these characteristics are common to all networks. Although there are useful networks which contain only one layer, or even one element, most applications require networks that contain at least the three normal types of layers - input, hidden, and output. The layers of input neurons receive the data either from input files or directly from electronic sensors in real-time applications. The output layer sends information directly to the outside world, to a secondary computer process, or to other devices such as a mechanical control system. Between these two layers can be many hidden layers. These internal layers contain many of the neurons in various interconnected structures. The inputs and outputs of each of these hidden neurons simply go to other neurons. In most networks each neuron in a hidden layer receives the signals from all of the neurons in a layer above it, typically an input layer. After a neuron performs its function it passes its output to all of the neurons in the layer below it, providing a feed forward path to the output.

4. Methodology

In this study, different types of networks are used for training. In which the inputs are presented to the network and the outputs obtained from the network are compared with the real output values (target values) of the system under investigation in order to compute error.

The steps involved in the identification of a nonlinear model of a system are selection of input-output data suitable for calibration and verification; selection of a model structure and estimation of its parameters; and validation of the identified models. Input variables are maximum temperature, minimum temperature; relative humidity, sunshine hours and wind speeds selected to describe the physical phenomena of the climatological parameters, in order to forecast the yield of cotton, and are checked with various performances indices.

The different network types are considered for the study, viz., Cascade forward back propagation, Elman back propagation, Feed forward back propagation, Feed forward distributed time delay, Generalized Regression, Layer recurrent, Linear layer (design) and the values of r, R^2 and D.R. are checked in both the training and validation.

The whole dataset has been divided into 70-30%. That means 70% data has considered for training and remaining 30% data has selected for validation of the model. The performance of the model has been found out by various indices as discussed above.

5. Model Evaluation Criteria

The performances of the developed ANN models were compared using statistical evaluation performance indices, namely the coefficient of correlation (r), coefficient of determination (R^2) and discrepancy ratio (D.R). The values of these performance indices were computed from the observed and model predicted values of the dependent variable. They were calculated for both the training and validation data sets. The values were calculated using the equations of model performance indices given in below Table 1.

Table 1: Performance Indices

Correlation coefficient(r)	$\frac{\sum_{i=1}^{n}(y(i)-\bar{y})(\hat{y}(i)-\hat{y})}{\sqrt{\sum_{i=1}^{n}(y(i)-\bar{y})^2\sum_{i=1}^{n}(\hat{y}(i)-\hat{y})^2}}$
Coefficient of determination(R^2)	$\left(\frac{\sum_{i=1}^{n}(y(i)-\bar{y})(\hat{y}(i)-\hat{y})}{\sqrt{\sum_{i=1}^{n}(y(i)-\bar{y})^2\sum_{i=1}^{n}(\hat{y}(i)-\hat{y})^2}}\right)^2$
Discrepancy Ratio	$\frac{\sum_{i=1}^{n} y(i)}{\sum_{i=1}^{n} \overline{y(i)}}$

Where ŷ (i) is the n estimated crop yield, y (i) is the n observed crop yield, y bar is the mean of the observed crop yield values.

6. Results and Discussion

In this study, a three-layer multilayer perceptron (MLP) model was developed in order to estimate the yield of cotton. As mentioned in the study area and data sections, we have developed the generalized linear regression model using 5 input variable as maximum temperature, minimum temperature, relative humidity, sunshine hours and wind speed data and 1 Output variable as yield of cotton data which were defined based on the problem at hand.

Here, the 70% of the total data has taken for the testing purpose and the remaining 30% data has selected for the validation. Modeling has been done for the all training algorithms to generate best fit ANN model to represent the best model.

Table 2: r, R^2 and D.R. during Training for different Network Types Considered

Model network type	*Training(70% data)*		
	r	*R^2*	*D.R*
Cascade forward back propagation	0.4689	0.2168	0.8970
Elman back propagation	0.7765	0.6029	0.8610
Feed forward back propagation	0.7960	0.6336	0.9468
Feed forward distributed time delay	0.4325	0.1870	1.0751
Generalized Regression	0.9770	0.9545	1.003
Layer recurrent	-0.3773	0.1423	0.7210
Linear layer (design)	0.6600	0.4326	1.000

Table 3: r, R^2 and D.R. during Validation for different Network Types Considered

Model network type	*Validation(30% data)*		
	r	*R^2*	*D.R*
Cascade forward back propagation	-0.2590	0.0670	1.024
Elman back propagation	0.1000	0.010	1.100
Feed forward back propagation	0.9044	0.8179	1.034
Feed forward distributed time delay	-0.1173	0.0137	1.077
Generalized Regression	0.9287	0.8624	0.9977
Layer recurrent	0.1691	0.0285	1.071
Linear layer (design)	0.6757	0.4565	1.000

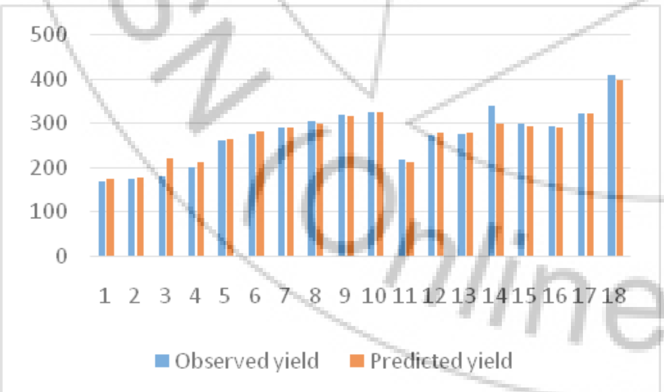

Figure 3: observed yield and predicted yield during training

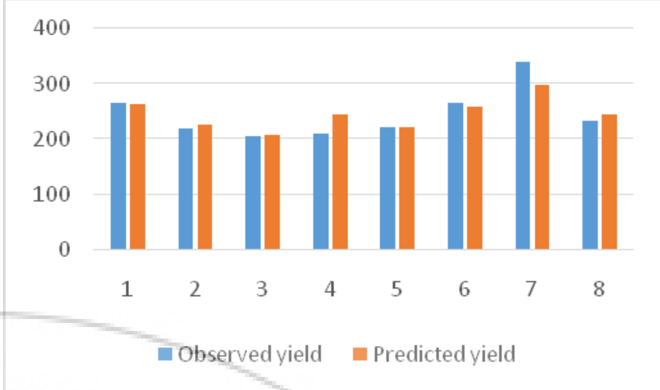

Figure 4: observed yield and predicted yield during Validation

As seen from the Table 2 and Table 3, the network type generalized regression gives the best correlation coefficient (r) value, which is 0.9770, coefficient of determination (R^2) value as 0.9545 and discrepancy ratio (D.R) value as 1.003 for training period and correlation coefficient (r) value as 0.9287, coefficient of determination (R^2) value as 0.8624 and discrepancy ratio (D.R) value as 0.9977 for validation period. The observed yield and predicted yield during training and validation are given in Fig. 3 and Fig. 4 respectively.

Hence, it is observed that the artificial neural network can be used for the estimation of crop yield using climatological data. For the area and data consider for the study, generalized regression proves to be the appropriate network type.

7. Conclusions

The climatic parameters such as maximum temperature, minimum temperature, relative humidity, sunshine hours and wind speed plays an important role for the crop yield estimation. From the results, it can be concluded that the best fit ANN model is generalized regression model, with r, R2 and D.R. as 0.9770, 0.9545 and 1.003 during training and 0.9287, 0.8624 and 0.9977 during validation respectively. In this study, for the area considered, the results obtained shows that the artificial neural networks can be an effective tool for the estimation of crop yield, using climatological data.

8. Future scope of this study

Artificial Neural Network may be efficiently used for development of models for estimation of yield of a crop, with inputs as climatological data, i.e. maximum temperature, minimum temperature, relative humidity, sunshine hours and wind speed, provided one has the data of the same for the previous years.

References

[1] http://en.wikipedia.org/wiki/Artificial_neural_network
[2] Maind S. B.,Wankar PriyankA, (2014) "Research Paper on Basic of Artificial Neural Network". International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 2 Issue: 1, 96-100.

[3] Parekh F. P. and Suryanarayana T. M. V., (2012). "Impact of Climatological Parameters on Yield of Wheat Using Neural Network Fitting.", International Journal of Modern Engineering Research (IJMER), www.ijmer.com Vol.2, Issue.5, Sep-Oct. 2012 pp-3534-3537 ISSN: 2249-6645
[4] Panchal R. A., Suryanarayana T. M. V., and Parekh F. P., (2013). "Rainfall-Runoff modeling using ANN at Dharoi sub-basin". NSCCIWRS, (2013)
[5] Suryanarayana T. M. V., and Parekh F. P. (2013). "Effect of Distribution of Dataset for Model Development using Neural Network Fitting with Emphasis on Yield of Cotton.", International Journal of Agriculture innovations and research, Volume 1 Issue 2, pps. 59-64.

Author Profile

Pranav B. Mistry Completed B.E. (Civil-IWM) in April (2011), P.G. student of M.E. (Civil) in Water Resources Engineering in From Water Resources Engineering and Management Institute (WREMI), Faculty of Technology and Engineering, The Maharaja Sayajirao University of Baroda, Gujarat, India. He is Life Member of various Professional Bodies like Indian Society of Hydraulics, and Indian Society for Geomatics.

Dr. T.M.V. Suryanarayana completed B.E.(Civil-IWM) in May 2001, M.E.(Civil) in Water Resources Engineering in November 2002 and Ph.D. in Civil Engineering in May 2007 from The M.S. University of Baroda, Vadodara, Gujarat, India. He is serving as Associate Professor and recognized Ph.D. Guide in Water Resources Engineering and Management Institute, Faculty of Technology and Engineering, The M. S. University of Baroda. He has 11 years of teaching and Research Experience. His areas of research include Operations Research, Hydrologic Modeling, Conjunctive Use, Hydraulics of Sediment Transport, Soil and Water Conservation, Reservoir Operation, Soft Computing Techniques, Climate Change. Dr. Suryanarayana is Executive Committee Member of Indian Water Resources Society, Secretary and Treasurer of Gujarat Chapter of Association of Hydrologists of India and Joint Secretary of Indian Society of Geomatics_Vadodara Chapter. He is Life Member of various Professional Bodies like Indian Water Resources Society, Indian Society of Hydraulics, Association of Hydrologists of India, Indian Society for Geomatics, Association of Agrometeorologists of India.