IOSR Journal of Mobile Computing & Application (IOSR-JMCA)
e-ISSN: 2394-0050, P-ISSN: 2394-0042.Volume 3, Issue 3. (May. - Jun. 2016), PP 01-06
www.iosrjournals.org

Nonlinear Regression and Artificial Neural Network Based Model for Forecasting Paddy (*Oryza Sativa)* Production in Tamil Nadu

Dr. B.Vinoth[1], Dr. A.Rajarathian[2], Dr. S.K. Manju bargavi[3]
[1]*Assistant Professor, Department of Statistics, Eritrea Institute of Technology, Eritrea, N.E. Africa*
[2]*Manonmaniam Sundaranar University, Tirunelveli, Tamil Nadu.*
[3]*Department of Computer Science, Eritrea Institute of Technology, Eritrea, N.E. Africa*

Abstract: *Artificial neural networks (ANNs) are non-linear mapping structures based on the function of the human brain. ANNs are known to be universal function approximators and are capable of exploiting nonlinear relationships between variables. ANNs can identify and learn correlated patterns between input data sets and corresponding target values. Crop production forecasting is a very important task for researchers in agriculture. Problems exist with multiple factors in the cropland ecosystem. This paper describes the successful application of an artificial neural network in developing a model for paddy production forecasting using back-propagation algorithms. In this paper nonlinear regression models namely Modified Horel and Morgan-Mercer-Flodin models has been used for forecasting paddy production and this approach has been compared with ANN methodology. For the choice between nonlinear regression and ANN models, the error measures namely, Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) have been estimated in the study.*
Keywords: *Artificial neural network, Back propagation, MAE, Nonlinear regression, RMSE.*

I. Introduction

India is one of the world's largest producers of paddy production, accounting for 20% of all world paddy production. Paddy is India's pre-eminent crop, and is the staple food of the people of the eastern and southern parts of the country. India is the second largest producer and consumer of rice in the world. Paddy is one of the chief grains of India. Moreover, this country has the biggest area under rice cultivation, as it is one of the principal food crops. It is in fact the dominant crop of the country. India is one of the leading producers of this crop. Rice is the basic food crop and being a tropical plant, it flourishes comfortably in hot and humid climate. Tamil Nadu one of the leading rice growing states in India, has been cultivating rice from time immemorial as this State is endowed with all favorable climatic conditions suitable for rice growing. The Government is implementing various programs to address the issues and constraints faced by the farmers to achieve the targeted growth in agriculture.

Parametric non-linear regression (NLR) models based on S-shaped curves (e.g. Weibull) using thermal and hydro-thermal indices as explanatory variables have been used widely for weed emergence prediction in [6],[8] and [13]. In many cases, NLR models have demonstrated adequate representation to the observed data; however, they present several major limitations. Specifically, NLR models are sometimes not flexible enough to capture complex features in the explanatory variable, such as abrupt 'jumps' or heavy 'tails' in [4]. Recently, artificial neural network (ANN) non-linear models have been widely used for resolving forecast problems in [1],[7] and [15]. Recent investigations in ANN research have found connections linking ANNs and statistics-based regression modeling. By comparing the two modeling structures, new insight can be gained on the functionality of ANNs.

The objective of the present work was to perform a comparative study between ANNs and NLR approaches to model paddy production using data gathered from locations in the Tamil Nadu.

II. Materials And Methods

To achieve the stipulated objectives, the present study had been carried out on the basis of time-series data on area and production of paddy crop pertaining to the period 1950-51 to 2009-10 had been collected from the office of Statistics and Economics, Teynampet, Chennai-600 006, Tamil Nadu.

2.1 Non-Linear Models

In parametric model, different non-linear models in [2],[5][10-11] and [14] given in Table 1 were employed. Among the non-linear models, the model having highest adjusted R^2 with significant F value was selected, so that it satisfied test for goodness of fit in [10]. In case of more than one model being the good fit for the data, the best model was selected based on lower values of Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) values. Levenberg-Marquardt algorithm in [11] which is widely used was utilized for fitting Logistic, Gompertz Relation, Sinusoidal and Rational Function. Different sets of initial parameter values were tried so as to ensure global convergence.

Table 1. List of non-linear models

Model No.	Model	Name of the Model
1	Y=A*B^(1/X)*X^C + e	Modified Hoerl Model (MH)
2	Y=(A*B+C*X^D)/(B+X^D)+e	Morgan-Mercer-Flodin (MMF)

Here Y is the area/ production and X is the time points; A, B, C and D are the parameters and e is the error term. The parameter A represents carrying capacity; C is the intrinsic growth rate; B represents different functions of the initial value y(0) and D is the added parameter.

2.2 Artificial neural networks

An ANN is a mathematical or computational model based on biological neural networks. It consists of an interconnected group of artificial neurons and processes information using a connectionist approach to computation. In more practical terms, ANNs are nonlinear statistical data modeling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. In Ref. [9]describes for the first time proposed the idea of the artificial neural network but because of the lack of computing facilities they were not in much use until the back propagation algorithm was discovered by [12]. Multilayer feed forward neural network or multi layer perception (MLP) is very popular and is used more than other neural network type for a wide variety of tasks are shown in Fig. 1. Multilayer feed forward neural network learned by back propagation algorithm is based on supervised procedure, i.e., the network constructs a model based on examples of data with known output. It has to build the model up solely from the examples presented, which are together assumed to implicitly contain the information necessary to establish the relation.

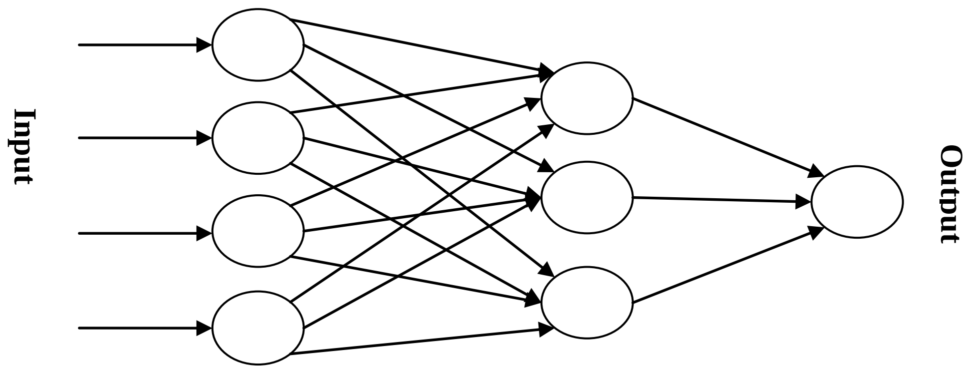

Figure 1. Architecture of neural network(Multi Layer Perception)

An MLP is a powerful system, often capable of modeling complex, relationships between variables. It allows prediction of an output object for a given input object. The architecture of MLP is a layered feed forward neural network in which the non-linear elements (neurons) are arranged in successive layers, and the information flow uni-directionally from input layer to output layer through hidden layers.

2.3 Back propagation Algorithm

Back propagation is a common method of training artificial neural networks so as to minimize the objective function. The MLP network is trained using one of the supervised learning algorithms of which the best known example is back propagation, which uses the data to adjust the network's weights and thresholds so as to minimize the error in its predictions on the training set. First, it computes the total weighted input x_j, using the formula:

$$X_j = \sum_i y_i W_{ij} \tag{1}$$

Typically we use the hyperbolic tangent function:

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{2}$$

The backpropagation algorithm consists of four steps:

(i). Compute how fast the error changes as the activity of an output unit is changed. This error derivative (EA) is the difference between the actual and the desired activity.

$$EA_j = \frac{\partial E}{\partial y_j} = y_j - d_j \qquad (3)$$

(ii). Compute how fast the error changes as the total input received by an output unit is changed. This quantity (EI) is the answer from step (i) multiplied by the rate at which the output of a unit changes as its total input is changed.

$$EI_j = \frac{\partial E}{\partial X_j} = \frac{\partial E}{\partial y_j} \times \frac{\partial y_j}{\partial x_j} = EA_j y_j \left(1 - y_j\right) \qquad (4)$$

(iii). Compute how fast the error changes as the weight on the connection into an output unit is changed. This quantity (EW) is the answer from step (ii) multiplied by the activity level of the unit from which the connection emanates.

$$EW_{ij} = \frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial X_j} \times \frac{\partial X_j}{\partial W_{ij}} = EI_j y_i \qquad (5)$$

(iv). Compute how fast the error changes as the activity of a unit in the previous layer is changed. When the activity of a unit in the previous layer changes, it affects the activities of all the output units to which it is connected. So to compute the overall effect on the error, we add together all these separate effects on output units. It is the answer in step (iii) multiplied by the weight on the connection to that output unit.

$$EA_i = \frac{\partial E}{\partial y_i} = \sum_j \frac{\partial E}{\partial x_j} \times \frac{\partial x_j}{\partial y_i} = \sum_j EI_j W_{ij} \qquad (6)$$

By using steps (ii) and (iv), we can convert the EAs of one layer of units into EAs for the previous layer. This procedure can be repeated to get the EAs for as many previous layers as desired. Once we know the EA of a unit, we can use steps (ii) and (iii) to compute the EWs on its incoming connections.

Different criteria were used to make comparisons between the forecasting ability of the nonlinear regression models and the neural network models. The first criterion is the root mean square error (RMSE). RMSE is given by $RMSE = \left[\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 / n \right]^{1/2}$. The second criterion is the mean absolute error (MAE) $MAE = \sum_{i=1}^{n} \left| (Y_i - \hat{Y}_i) \right| / n$. The lower the values of these statistics, the better are the fitted model.

III. Result And Discussion

A multilayer feed forward neural network (MLP) was fitted to the data. The data was divided into three sets viz. training, testing and holdout. The result shown in Table 2.

Table 2. ANN case processing summary of paddy production in Tamil Nadu.

		N	Percent
Sample	**Training**	33	55.0%
	Testing	19	31.7%
	Holdout	8	13.3%
Valid		60	100.0%
Excluded		0	
Total		60	

The information about the neural network architecture is given in Table 3 which shows that network has an input layer, a single hidden layer and an output layer. In the hidden layer there are 4 units and the activation function used is the hyperbolic tangent.

Table 3.Network information for paddy production

Input Layer	Covariates	Year Area
	Number of Units	2
	Rescaling Method for Covariates	Standardized
Hidden Layers	Number of Hidden layers	1
	Number of Units in Hidden Layers	4
	Activation Function	Hyperbolic tangent
Output Layer	Dependent Variables	Production
	Number of Units	1
	Rescaling Method for Scale Dependents	Standardized
	Activation Function	Identify

	Error Function	Sum of squares

The below drawn architecture of the network has been shown in the Fig. 2, light color lines show weights greater than zero and the dark color lines show weights less than zero.

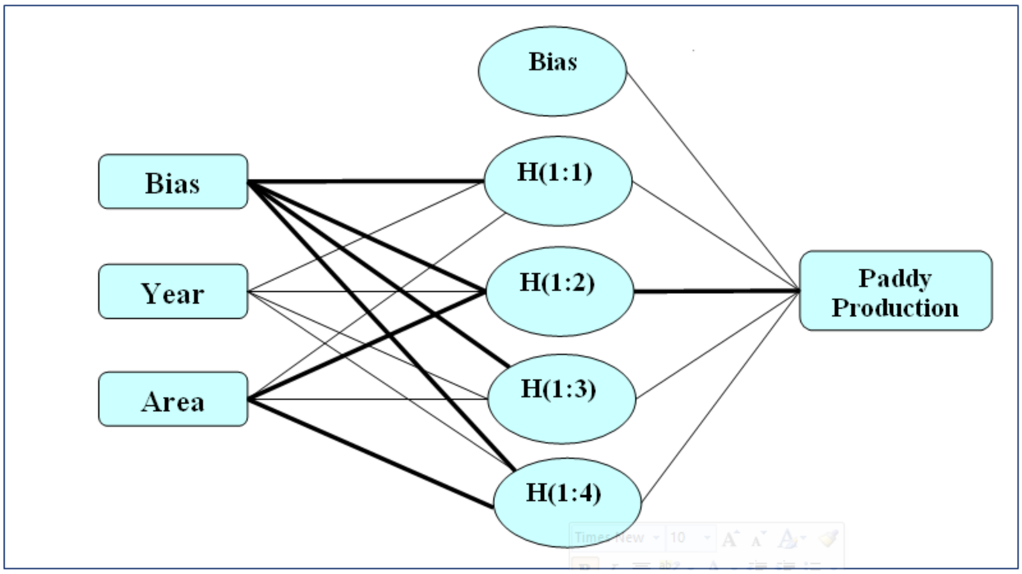

Figure 2. Hidden layer activation function: Hyperbolic tangent; Output layer activation function: Identify.

The training summary and the fit statistics for the training, testing and the holdout sets are given below in Table 4.

Table 4. The training summary and the fit statistics of ANN of paddy production.

Training	Sum of Squares Error	1.977
	Relative Error	0.124
	Stopping Rule Used	1 consecutive step(s) with no decrease in error
Testing	Sum of Squares Error	1.946
	Relative Error	0.310
Holdout	Relative Error	0.171

The estimates of the weights and bias are given below in Table 5. This table shows the value of weights from input to the hidden layer and from the hidden layer to the output layer. H(1:1) means Hidden layer 1 and 1st neuron. The weight attached to the neuron from bias is -0.511, from year is 1.021, from area is 0.875. H(1:2) means Hidden layer 1 and 2nd neuron. The weight attached to the neuron from bias is -0.599, from year is 0.345 from area is -0.823. H(1:3) means Hidden layer 1 and 3rd neuron. The weight attached to the neuron from bias is -0.527, from year is 1.028 from area is 1.073. H(1:4) means Hidden layer 1 and 4th neuron. The weight attached to the neuron from bias is -0.652, from year is 1.520 from area is -1.444.

The weights from the hidden layer to the output layer for bias 0.192 and from 1st neuron in the hidden layer to the output is 0.216, from 2nd neuron in the hidden layer to the output is -0.890, from 3rd neuron in the hidden layer to the output is 0.678 and from 4th neuron in the hidden layer to the output is 0.992.

Table 5. The estimates of the weights and Bias of ANN fitted to paddy production.

Predictor		**Predicted**			
		Hidden Layer 1			
		H(1:1)	H(1:2)	H(1:3)	H(1:4)
Input layer	(Bias)	-0.511	-0.599	-0.527	-0.652
	Year	1.021	0.345	1.028	1.520
	Area	0.845	-0.823	1.073	-1.444
Hidden Layer 1	(Bias)	H(1:1)	H(1:2)	H(1:3)	H(1:4)
Paddy Production Output Layer	0.192	0.216	-0.890	0.678	0.992

The observed and estimated values are depicted in the Fig. 3 which reveals that except for few outliers it is a straight line and there exists a one to one correspondence among the observed and estimated vales. Hence it can be inferred that the performance of ANN is satisfactory.

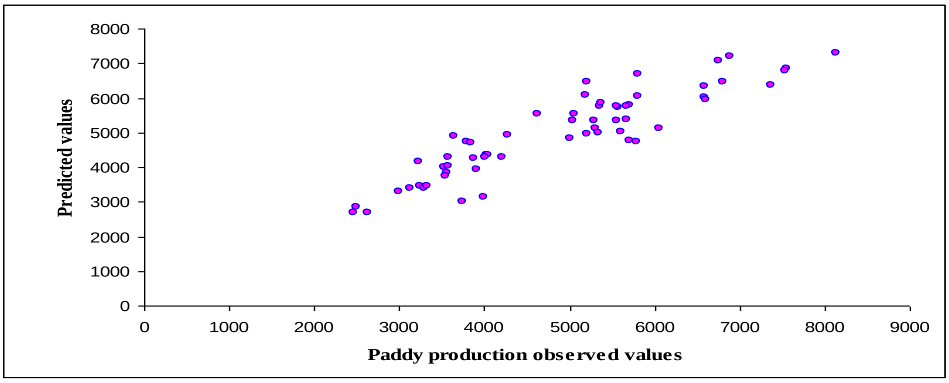

Figure 3. Observed vs Predicted values of ANN

Table 6. Comparison of different forecasting methods

S.No	Observed Values	Estimated Values Based on			S.No	Observed Values	Estimated Values Based on		
		MH	MMF	ANN			MH	MMF	ANN
1	2459	2502	3106	2681	31	4279	5129	5349	4928
2	2635	2600	3108	2679	32	5681	5177	5404	5378
3	2489	2778	3112	2834	33	3642	5225	5456	4898
4	3749	2948	3121	3006	34	4633	5271	5503	5543
5	3995	3103	3136	3141	35	5365	5317	5546	5763
6	3002	3243	3158	3303	36	5370	5362	5586	5841
7	3247	3373	3189	3460	37	5333	5406	5622	4981
8	3288	3492	3230	3386	38	5614	5449	5656	5013
9	3134	3604	3281	3397	39	5704	5491	5687	4783
10	3333	3708	3344	3466	40	6063	5533	5716	5116
11	3559	3806	3417	3846	41	5782	5574	5742	4748
12	3907	3899	3501	3926	42	6596	5614	5766	6007
13	4024	3987	3595	4354	43	6806	5654	5789	6449
14	3876	4071	3697	4259	44	6750	5693	5809	7074
15	4036	4152	3806	4346	45	7559	5732	5828	6845
16	3524	4229	3920	3999	46	5290	5770	5846	5343
17	3791	4303	4037	4733	47	5805	5807	5862	6689
18	3846	4374	4155	4695	48	6894	5844	5878	7191
19	3550	4443	4273	3752	49	8141	5880	5892	7301
20	4012	4510	4389	4287	50	7532	5916	5905	6798
21	5001	4574	4502	4835	51	7366	5951	5917	6379
22	5302	4636	4611	5133	52	6584	5986	5928	6323
23	5569	4697	4716	5717	53	3577	6020	5939	4275
24	5558	4756	4815	5334	54	3223	6054	5949	4150
25	3575	4813	4908	4020	55	5062	6088	5958	5527
26	5203	4869	4995	4957	56	5209	6121	5966	6472
27	4215	4923	5077	4284	57	6611	6154	5974	5949
28	5705	4977	5153	5788	58	5040	6186	5982	5353
29	5559	5028	5224	5776	59	5183	6218	5989	6076
30	5800	5079	5289	6045	60	5665	6250	5995	5755

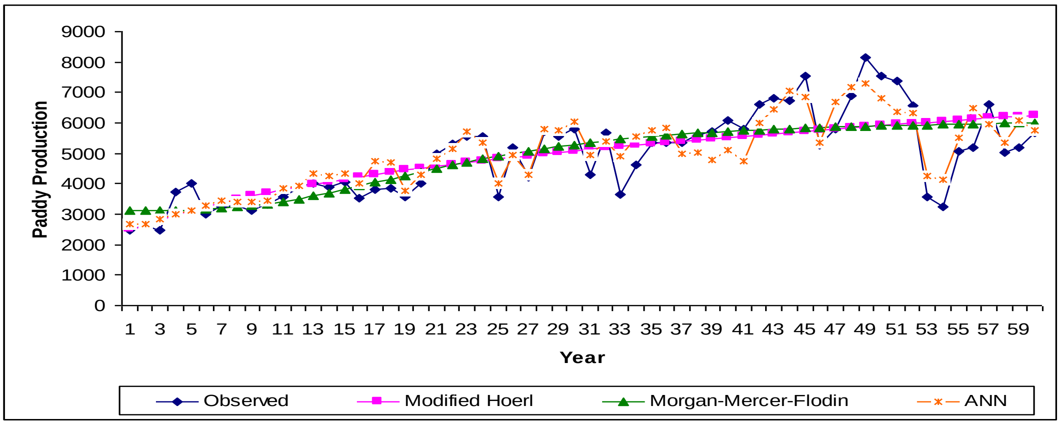

Figure 5. Trends in paddy production based on nonlinear regression models and ANN Method

Table 7. Characteristics of Model Performance parameters

Nonlinear Regression & ANN	RMSE	MAE
Modified Hoerl Model (MH)	921.64	698.65
Morgan-Mercer-Flodin Model (MMF)	895.11	673.87
Artificial Neural Network (ANN)	576.19	474.88

IV. Conclusions

The study demonstrated that the artificial neural network(Multilayer Perception) model was found to be a suitable model to study the trends in production of paddy crop gives the lowest values of RMSE, MAE among compared with the different non-linear models such modified Hoerl and Morgan-Mercer-Flodin. Also a Decreasing trends in production of paddy crops have been observed. For further development and improvement of different statistical methods for the assessment of crops productivity it is necessary to improve and extend the theoretical basis of agricultural statistics.

References

[1]. H. Altun, A. Bilgil, and B.C. Fidan, Treatment of multidimensional data to enhance neural network estimators in regression problems, *Expert Systems with Applications, 32(2)*, 2007, 599–605.

[2]. Y. Bard, *Nonlinear Parameter Estimation*, (Academic Press: New York, 1974).

[3]. W. J. Bullied, R.C. Van Acker and P.R. Bullock, Hydrothermal modelling of seedling emergence timing across topography and soil depth. *Agronomy Journal, 104*, 2012,pp.423–436.

[4]. R. Cao, M. Francisco-Fernández, A. Anand, F. Bastida and J.L.González-Andújar, Computing statistical indices for hydrothermal times using weed emergence data. *Journal of Agricultural Science*, Cambridge, *149*, 2011, pp.701–712.

[5]. N.R. Draper and H. Smith, *Applied Regression Analysis*,(3rd Edition, John and Wiley & Sons, USA, 1998).

[6]. F. Forcella, Real-time assessment of seed dormancy and seedling growth for weed management, *Journal of Seed Science Research, 8*, 1998, 201–209.

[7]. T. Hill, M. O'Connor and W. Remus, Neural network models for time series forecasts, *Management Science, 42(7)*, 1996, 1082–1092.

[8]. J. Izquierdo, J.L. González-Andújar, F. Bastida, J.A. Lezaún and M.J. Sánchez Del Arco, A thermal time model to predict corn poppy (Papaver rhoeas) emergence in cereal fields, *Weed Science, 57*, 2009, pp. 660–664.

[9]. W.S. McCulloch and W.Pitts, A Logical Calculus of Ideas Immanent in Nervous Activity, Bulletin of Mathematical Biophysics, 1943, 115 - 133.

[10]. D.C. Montgomery, E.A. Peck and G.G. Vining , *Introduction to Linear Regression Analysis*, (John Wiley & sons, Inc., 2003)

[11]. D. A. Ratkowsky, *Handbook of Non-linear Regression Models*, (Marcel Dekker, 1990, New York).

[12]. D. E. Rumelhart, G.E. Hinton and R.J. Williams, Learning representations by back-propagating error. Nature, 323, 1986, 533–536. Reprinted in Anderson and Rosenfeld [1988], 696–699.

[13]. .J. Schutte, E.E. Regnier, S.K. Harrison, J.T. Schmoll, K. Spokas and F. Forcella , A hydrothermal seedling emergence model for giant ragweed (Ambrosia trifida), *Weed Science, 56*, 2008, 555–560.

[14]. G.A.F. Seber and C.J. Wild, *Non-Linear Regression* (John Wiley and Sons, New York, 1989).

[15]. F.M. Tseng, H.E. Yu and G.H. Tzenf, Combining neural network model with seasonal time series ARIMA model. *Technological Forecasting and Social Change, 69*, 2002, 71–87.