DEVELOPMENT OF FEED FORWARD BACK PROPAGATION NEURAL NETWORK WITH BEST FITTING MODELS TO PREDICT SEASONAL RICE PRODUCTION IN TAMILNADU

S. Arun Balaji[1] and P. Manimegalai[2]

Department of CSE, Karpagam University, Coimbatore – 641021, Tamil Nadu, India;
[1]arunbalaji1983@yahoo.co.in, [2] manimegalai.vairavan@gmail.com

ABSTRACT

The study reported the development of FFBPNN architecture and its corresponding software to predict the rice production data for three seasons in 31 districts of Tamilnadu. It was found that the training and testing data were exactly matching with the predicted data. It was also found that the Absolute Relative Error (ARE) was found to be zero at the 9th iteration itself. The FFBPNN system was improved by integrating it with the best fitting models using the curve expert software. The improved FFBPNN with best fitting model was used to predict the area of rice and its production. The predicted data was compared with the observed data. The paired t-test was conducted between the observed and predicted data. It was found that there is 67% of fittings are showing insignificant difference between the observed area of rice and predicted area of rice cultivation. Similar test was also conducted for the rice production data; it was found that there is 73.3% of fittings showing insignificant difference between the observed and predicted data.

Keywords: FFBPNN, Best fitting models, Prediction, Rice production, Tamilnadu

1. INTRODUCTION

Rice is the main food consumed in south India including the state of Tamilnadu. Rice production is a complex process involving the different types of soil, varieties of seeds, weather conditions, seasons in a year, varied land and water management practices, pest and disease management techniques, manure and fertilizer management methods, weed management and timeliness management of different unit operations like sowing the seeds, growing of rice and harvesting practices. Hence, rice production is a non-linear, parallel and interconnected process. Many mathematical and statistical methods have been developed to predict rice production based on different parameters. All these methods do not involve a non-linear modeling approach. The current research is based on non-linear, highly parallel and interconnected networking approach of using Feed Forward Back Propagation Neural Network (FFBPNN) with sigmoid activation function. The FFBPNN system was integrated with Curve Expert Software to develop the best fitting models and to predict the data pertaining to area of rice cultivation and rice production in different districts of Tamilnadu, which are useful to policy makers of Government of Tamilnadu. The input data like area of rice cultivation in hectare and rice production in tonnes for three seasons namely Kuruvai, Samba and Kodai seasons for the 31 districts of Tamilnadu for five years from 2005-06 to 2009-10 were collected from the Seasons and Crop report published by the Government of Tamilnadu[1] for carrying out the research.

The overall objective of this paper is the "Development of Feed Forward Back Propagation Neural Network (FFBPNN) with best fitting models to predict seasonal rice production in all the districts of Tamilnadu". The specific objectives are given below:

1. To develop the FFBPNN architecture and its corresponding software to predict the area of rice cultivation and rice production in different districts of Tamilnadu for three seasons.
2. To study of the effect of different statistical measures in error reduction pattern of FFBPNN prediction
3. To compare the observed data and the FFBPNN predicted data to measure the level of prediction
4. To develop the best fitting models for area of rice and rice production in different districts during three seasons using the Curve Expert Software and to find out the percentage determination of the best fitted models
5. To integrate the best fitting models into the FFBPNN system and to study the predicted data by means of statistical testing between the observed and predicted data.

2. RELATED WORKS

In agricultural practices, crop production is influenced by a great variety of interrelated factors and it is difficult to describe their relationships by conventional methods. Thus, artificial neural network (ANN) is highly suggested to present the complicated relations and strong nonlinearity between different parameters and crop production. It is considered to be one of the best techniques for extracting information from imprecise and non-linear data[2]. Hence, neural networks (NNs) methods have become a very important tool for a wide variety of applications across many disciplines including prediction of crop production where traditional statistical techniques were used. This has led to a number of studies comparing the traditional statistical techniques with neural networks in a variety of applications. It has been recognized in the literature

that regression and neural network methods have become competing model-building met-hods[3].
Nowadays, NNs methods have been largely used in the areas of prediction and classification[4]. NNs models are also preferred in the area of pattern recognition[5]. Many researchers have shown the relationship between neural networks and statistical models[6],[7],[8],[9],[10] showed a comp-lete analysis and comparison of different network techniques with traditional statistical techniques.
The strong association of the feed forward neural networks with discriminant analysis was also shown by the authors. Schumacher et al.,[11] have shown a compa-rison between feed forward neural networks and logistic regression. The similarities and dissimilarities were also analyzed. Sarle[8] presented a neural network terminology into statistical terminology and showed the relationship between neural networks and statistical techniques. Warner et al.,[4] compared the performances of regression analysis and neural networks using simulated data from known functions and also using real world data. The authors discussed the situations where it would be advantageous to use NNs models in place of parametric regression models. Ripley[7] presen-ted the statistical aspects of neural networks and classi-fied neural networks as one of the flexible non-linear regression methods. This background stimulated the researcher to take up the present research.
Serious efforts to create a model of a neuron have been underway during the last 100 years, and remarkable progress has been made. Neural connections are signify-cantly fewer and similar than the connections in the brains. The basic model of ANN is as shown in Figure 1. This is the combination of perceptron, which forms the Artificial Neural Network which is actually practiced.
Neuron: A neuron (or cell or a unit) is an autonomous processing element. Neuron can be thought of a very sim-ple computer. The purpose of each neuron is to receive information from other neurons, perform relati-vely simple processing of the combined information, and send the results to one or more other neurons. In many of the illustrations shown in various literatures, the neurons are generally indicated by circles or squares.
Layers: A layer is a collection of neurons that can be thought of as performing some type of common funct-ions. These neurons are usually numbered by placing the numbers or letters by each neuron and are generally assu-med that no neuron is connected to another neuron in the same layer. All neurons nets have an input layer and have outputs to interface with the external environ-ment. Each input layer and each output layer has at least one neuron. Any neuron that is not in an input layer or in an output layer is said to be in a hidden layer (shown in Fig. 1) because neither its input nor its output can be observed from outside. Sometimes the neurons in a hidden layer are called feature detectors because they respond to particular feature detectors because they respond to particular features in the previous layer.
Synapses (arcs / links): An arc or a synapse can be a one -way or a two-way communication link between two cells. A feed-forward network is one in which the information flows from the input cells through hidden layers to the output neurons without any paths whereby a neuron in a lower numbered layer receives input from the neurons in a high numbered layer.

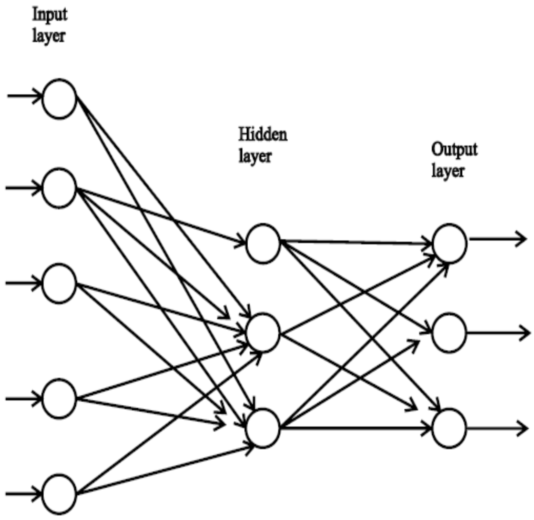

Figure 1. Basic Model of ANN

Weights: A weight w_{ii} is a real number that indicates the influence that neuron n_i has on neuron n_i. For example, positive weights indicate reinforcement, negative weights indicate inhibition and a weight zero (or the absence of weight) indicates that no direct influence or connection exists. The weights are often combined into a matrix *w*. These weights may be initialized as given and predefined values, to initialized as random numbers, but they can be altered by experience. It is in this way that the system learns. Weights may be used to modify the input from any neutron. However, the neurons in the input layer have no weights, which are the external inputs are not modified before going into the input layer.
Propagation Rule: A propagation rule is a network rule that applies to all the neurons and specifies how outputs from cells are combined into an overall net input to neuron *n*. The term 'net', indicates this combination. The most common rule is the weighed sum rule wherein, adding the products of the inputs and their corresponding weights forms the sum,

$$net i = b_i + \sum w_{ij} \, p_{ij} \qquad \text{eq. (1)}$$

where, *j* takes on the appropriate indices corresponding to the numbers of the neurons that send information to the neurons $n_{i.}$ The term in p_j represents the inputs to neurons $n_{i,}$ from neuron n_j. If the weights are both positive and negative, then this sum can be computed in two parts: excitory and inhibitory. The term b_i represents a bias associated with neurons $n_{i.}$ Adding one or more special neuron(s) having a constant input of unity often simulates these neuron biases.
An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way

biological nervous systems, such as the brain process information. The key element of this paradigm is the novel structure of the information processing system. It is composed of a large number of highly interconnected processing elements (neurons) working in unison to solve specific problems. ANNs does its predictions like people learn by experience. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. Learning in biological systems involves adjustments to the synaptic connections that exist between the neurons. This is true of ANNs as well. Neural networks are clusters of neurons that are interconnected to process information. Hence, the researcher has taken up the problem of developing the Feed Forward back Propagation Neural Network (FFBPNN) to predict area of cultivation of rice and rice production from different districts of Tamilnadu. The brief results of the research published by the authors are given below:

Arun Balaji et al.,[12] stated that prediction of annual rice production in all the 31 districts of Tamilnadu is an imp-ortant decision for the Government of Tamilnadu. Rice production is a complex process and non-linear problem involving soil, crop, weather, pest, disease, capital, labour and management parameters. ANN software was designed and developed with Feed Forward Back Propagation (FFBP) network to predict rice production. The input layer has six independent variables like area of cultivation and rice production in three seasons like Kuruvai, Samba and Kodai. The sigmoid activation function was adopted to convert input data into sigmoid values. The hidden layer computes the summation of six sigmoid values with six sets of weights. The final output was converted into sigmoid values using a sigmoid transfer function. ANN outputs are the predicted results. The error between original data and ANN output values were computed. A threshold value of 10^{-9} was used to test whether the error is greater than the threshold level. If the error is greater than threshold then updating of weights was done all summations were done by back propagation. This process was repeated until error equal to zero. The predicted results were printed and it was found to be exactly matching with the expected values. It shows that the ANN prediction was 100% accurate[12].

Arun Balaji et al.,[13] stated the development of Multiple linear regression (MLR) equations between the years of rice cultivation and FFBPNN method of predicted area of rice cultivation/rice production for different districts pertaining to Kuruvai, Samba and Kodai seasons in Tamil-nadu. The average r^2 value in area of cultivation is 0.40 in Kuruvai season, 0.42 in Samba season and 0.46 in Kodai season, where as the r^2 value in rice production is 0.31 in Kuruvai season, 0.23 in Samba season and 0.42 in Kodai season. The Rice Data Simulator (RDS) predicted the area of rice cultivation and rice production using the MLR equations developed in this research. The range of average predicted area for Kuruvai, Samba and Kodai seasons varies from 12052.52 ha to 13595.32 ha, 4899 8.96 ha to 53324.54 ha and 4241.23 ha to 6449.88 ha respectively whereas the range of average predicted rice production varies from 45132.88 tonnes to 46074.48 tonnes in Kuruvai, 128619 tonnes to 139693.29 tonnes in Samba and 15446.07 to 20573.50 tonnes in Kodai seas-ons. The mean absolute relative error (ARE) between the FFBPNN and multiple regression methods of prediction of area of rice cultivation was found to be 15.58%, 8.04% and 26.34% for the Kuruvai, Samba and the Kodai seasons respectively. ARE for the rice production was found to be 17%, 11.80% and 24.60% for the Kuruvai, Samba and the Kodai seasons respectively. The paired t test between the FFBPNN and MLR methods of predicted area of cultivation in Kuruvai shows that there is no significant difference between the two types of prediction for certain districts however it has significant variations for some years[13].

Arun Balaji et al.,[14] stated that to get high accuracy of prediction, the curve expert software was integrated into the FFBPNN software. The curve fitting software developed the best fitting models among the 30 different linear and non linear models for Kuruvai, Samba and Kodai seasons of different districts of Tamilnadu. The test data and training data was fed as input to the FFBPNN soft-ware, it was found that the there was zero error between the observed data and the predicted data. The RMSE is zero and the ARE is also zero at 18th iteration. The curve expert produced the best fitting model to different districts during the three seasons. The curve expert produced the best fitting model to different districts during the three seasons. These developed models were used to simulate the best predicted area of rice cultivation and rice production. The type of fits are found to be 1) Quadratic Fit, 2) Linear Fit 3) User-Defined Model 4) Saturation Growth-Rate Model 5) Logarithm Fit 6) Hyperbolic Fit 7) Exponential Fit and 8) Gaussian Model.

Arun Balaji and Manimegalai Vairavan[15] reported that statistical error analyses have been used to assess the performance of the error reduction pattern of the FFBPNN model. Some of such error analyses methods used are Coefficient of Determination (R^2), Mean Square Error (MSE), Root Mean Square Error (RMSE) and Absolute Relative Error (ARE). It was found that R^2 is a poor statistical measure in the reduction of error for the prediction of rice production. It was also found that RMSE is a much better statistical measure compared to MSE because more data sets get zero error compared to MSE. It was established that ARE is zero for all the data items for the three seasons at the 9th iteration itself. Hence, ARE is the best statistical measure used in FFB-PNN system to predict rice production. The predicted results were printed and it was found to be matching with the expected values.

3. METHODOLOGY

The methodology adopted is based on the work of Makinde, et al.,[16]. There are generally four steps in the modeling process. They are: 1) Assembling the training data, 2) Creating the hybrid network object with integration of Curve Expert Software, 3) Training the network, and 4) Studying the network response to inputs.

3.1 Assembling training and testing data: The training data consist of the area (ha) and rice production (tonnes) for three seasons of Kuruvai, Samba and Kodai for the year 2009-10 pertaining to 31 districts. The test data comprises of the 4 years data from 2005-06 to 2008-09 containing the area (ha) and rice production (tonnes) for three seasons of Kuruvai, Samba and Kodai seasons. The initial weights of 42 data within the range of 0 to 1 were assumed. The training data along with initial weights were stored in a input file. The training data sets were used to train the FFBPNN model. There are 31 districts; each district consists of 7 columns of data items like name of the district, three area of cultivation data and three rice production data for Kuruvai, Samba and Kodai seasons. It causes 217 data items. The initially assumed weights in the data file are 42. Hence, the total data item for training is 259. Testing set for each year also contains similar number of data items like training set. Four years testing data were stored in four different data files for processing by a computer program.

3.2 Normalization of input data: The training data were normalized using sigmoid activation equation before being presented to the network for training. This step was taken to ensure that input data with different ranges were transformed into one similar range of 0 to 1 and allows for easier and faster model training. These training data are converted into sigmoid data between 0 to 1 using a sigmoid activation function and the corresponding converted data between 0 to 1 are represented as x1(i), x2(i)…….x6(i). The general format of the sigmoid action functional equation is given below:

$$S(x) = \frac{1}{1 + e^{(-x)}} \qquad \text{eq. (2)}$$

Where

S(x) is the sigmoid value. It varies from 0 to 1.

x is the independent input values like Area in Hectare and Rice Production in tonnes.

3.3 Creating the FFBPNN system: The multi layered FFBPNN system in this study begins with an input layer. The input layer is connected to a hidden layer; the hidden layer is then connected to output layer. In this study, the architecture used for the neural network consists of one input layer, one hidden layer and lastly, the output layer. The input layer converts the area of rice cultivation in hectare and the rice production in tonnes for three seasons into sigmoid values between 0 to 1 using the sigmoid activation function. The input layer is made with six neurons which are area of rice for Kuruvai, rice production for Kuruvai, area of rice for Samba, rice production for Samba, area of rice for Kodai and rice production for Kodai seasons. The output layer has only one neuron containing the predicted value from 0 to 1. The content of the output neuron will be multiplied with the six input observed values to get the targeted values.

3.3.1 Input layer: The input layer receives the area of rice cultivation in hectare and the rice production in tonnes for three seasons into sigmoid values between 0 to 1 using the sigmoid activation function. The sigmoid values were fed to the hidden layer where the data is multiplied by the assumed initial weights, which is varying from 0 to 1. The input layer is the beginning through which the external environment presents a pattern to the neural network. Input in to a neuron is the weighted sum of outputs from the nodes connected to it. The output from a neuron is the weighed sum of inputs into the neuron. Once a pattern is presented to the input layer, the output layer will produce another pattern. The input layer should represent the condition for which we are training the neural network. Every input neuron should represent some independent variable like area of rice cultivation and rice production for Kuruvai, Samba and Kodai seasons that has an influence over the output of the neural network. The input layer is made with six neurons which are containing the sigmoid values of area of rice for Kuruvai, rice production for Kuruvai, area of rice for Samba, rice production for Samba, area of rice for Kodai and rice production for Kodai seasons.

3.3.2 Hidden layer: The function of the hidden layer is to compute the summation of the sigmoid values and the weights coming from each nodes of the input layer and produces something that the output layer can use. An assumed bias is also added to the summation to prevent negative values. The equation for summation is given below:

$$Yi = \sum Xi\ Wij + Bias\ (+1) \qquad \text{eq. (3)}$$

Where Y_i is the summation of each node X_{ij} with corres-ponding initially assumed weights W_i plus bias. Bias used is +1. Bias is added to the summation to make the summation a number other than 0. It is essential to avoid 0 so that subsequent computations may not face division by zero (infinity).

Different computations performed in the hidden layer are given below:

Steps: Computations performed in the hidden layer of FFBPNN

1. Z_i is processed by multiplication of output Y_i from the hidden layer with initially assumed weight W_k. There are totally six values Z(1), Z(2)….Z(6). $Z_i = Y_i * W_k$
2. Calculation of total of all six Z values. $ZT(i) = \sum Z_i$
3. Convert the ZT(i) into ZS(i), that is sigmoid value from 0 to 1 with same activation function as given:

$$ZS(i) = \frac{1}{1 + e^{-ZT(i)}}$$

4. Convert the ZT(i) into ZS(i), that is sigmoid value from 0 to 1 with same activation function as given:

$$ZS(i) = \frac{1}{1 + e^{-ZT(i)}} \text{FBPNN Output}(i) = ZS(i) * \text{Observed Data}(i)$$

5. Compute the Error between actual input data and the FFBPNN output data using the formula given below:

$$\text{Error } (i) = \frac{[\text{Observed Data}(i) - \text{ FFBPNN oupu}(i)]}{\text{Observed Data}(i)}$$

6. Compare the error with threshold value of 10^{-9}. If the error is greater than threshold value then calculates the updated weights using the formula given below:
   New updated weight=Old weight + Increment of 0.01
7. Back propagation technique was adopted by computing the summation in the hidden layer using updated weights. It computes six sets of Yi and six sets of Zi. The total of six sets Zi is called one set of ZTi. The conversion of ZTi using activation function gave ZSi. The multiplication of ZSi and actual value of variab les gave the FFBPNN output. Repeat the process of feed forward and back propagation techniques until error is 0. When error is 0, print the original input and the corresponding FFBPNN output
8. If original independent input data and the corresponding FFBPNN output are same, the research has predicted correctly without any deviations. This happens only when error is 0.
9. If there is difference between the original input data and the corresponding FFBPNN output, there are some deviations. The errors were worked out to understand how far the FFBPNN output was away from the original data. Then update the weights and follow the back-propagation technique of computing summations.

3.3.3 Output layer: The layer where the processed output information is computed and presented is called the output layer. The hidden layer computes the summa -tion by multiplying each input with its corresponding weights and adds them together with the bias and applies the sigmoid action function. The output layer gets its in-put from the hidden layer. The output layer of the neural network is what actually presents a pattern to the external environment. The pattern presented by the output layer can be directly traced back to the input layer. It is designed that every input neuron provides one predicted value to the corresponding output neuron. The output layer has only one neuron containing the predicted value from 0 to 1. The content of the output neuron will be multiplied with the six input observed values to get the targeted values, which are predicted area of rice for three seasons and predicted rice production for three seasons. The predicted output from the FFBPNN system was compared with the observed values.

3.3.4 FFBPNN Architecture: After careful selection of input layer, number of layers and neurons in the hidden layer, and the output layer; the neural network architecture was achieved. The FFBPNN has one hidden layer with sigmoid neurons followed by an output layer of six sigmoid neurons. Figure 2 shows the FFBPNN architecture.

After careful consideration of the input, hidden, output layer and the neural network architecture, the network was implemented using a C Programming language. Figure 3 shows the programming flow chart to implement FFBPNN system. The Program reads the 259 training data items from a sequential file and computes the Absolute Relative Error (ARE). The program check the condition whether ARE is ≤ threshold value of 10^{-9}. If the condition is not satisfied then the back propagation is started by updating the initial weight with the adding of incre-mental increase in weight. The incremental increase in weight adopted is 0.01. It completes one iteration. This process is repeated until the condition of ARE ≤ threshold is satisfied. The output layer has only one neuron containing the predicted value from 0 to 1. When the condition of ARE ≤ threshold is satisfied then the content of the output neuron will be multiplied with the six input observed values to get the targeted predicted values, which are predicted area of rice for Kuruvai, predicted rice production for Kuruvai, predicted area of rice for Samba, predicted rice production for Samba, predicted area of rice for Kodai and predicted rice production for Kodai seasons. The predicted output from the FFBPNN system was compared with the observed values.

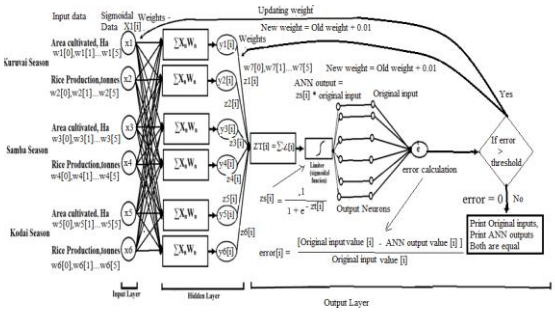

Figure 2: Architecture of FFBPNN designed for predicting the rice production in Tamilnadu

3.3.5 Training of the network: Training set of data consisted of the name of the district, observed area of rice and observed rice production for three seasons of 31 districts for the year 2009-10. The training data was used to build the FFBPNN system. The initial weights of 42 data within the range of 0 to 1 were assumed. The train-ing set of data along with initial weights was stored in an input file. The training data sets were used to train the FFBPNN model. There are 31 districts; each district consists of 7 data items like name of the district, three areas of cultivation data and three rice production data for Kuruvai, Samba and Kodai seasons. It causes 217 data items. The initially assumed weights in the data file are 42. Hence, the total data item for training is 259. Training of the FFBPNN is nothing but updating network architecture and connection weights so that the network can efficiently predict the output.

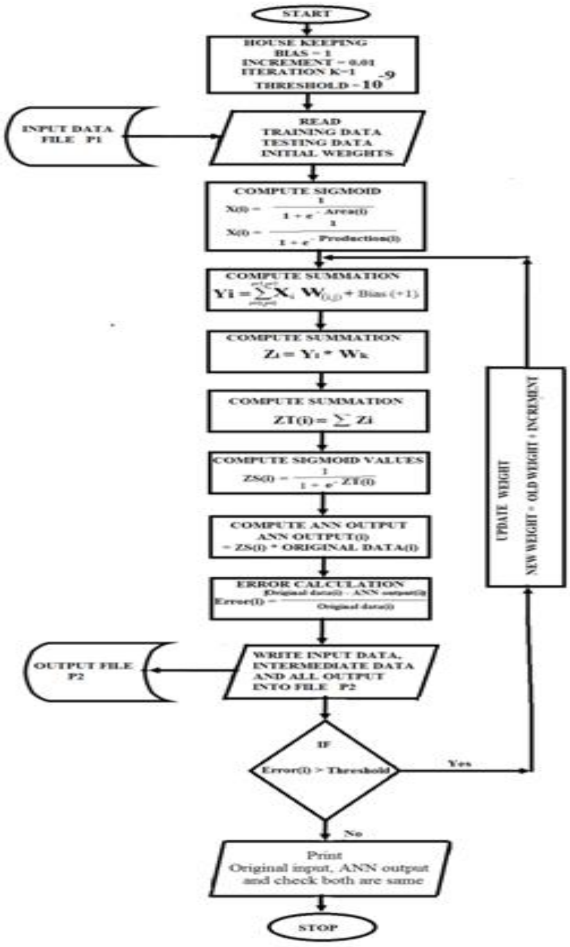

Figure 3. Programming Flow chart to implement FFBPNN system

Performance is improved over time by iteratively updating the weights in the network. In supervised training of the network, the FFBPNN should compute ARE. The C program to implement the FFBPNN check the condition whether ARE is ≤ threshold value of 10^{-9}. If the condition is not satisfied then the back propagation is started by updating the initial weight with the adding of incremental increase in weight. This process is repeated until the condition of ARE ≤ threshold is satisfied. When the condition of ARE ≤ threshold is satisfied then the content of the output neuron will be multiplied with the six input observed values separately to get the six targeted predicted values, which are predicted area of rice for Kuruvai, predicted rice production for Kuruvai, predicted area of rice for Samba, predicted rice production for Samba, predicted area of rice for Kodai and predicted rice production for Kodai seasons.

The C program written to implement the FFBPNN system read the training set of data from input file and computes the various intermediate and final results into the output file. The output file will be examined for the error reduction pattern and the final FFBPNN predicted output.

3.3.6 Validation and testing of FFBPNN: The area of rice cultivation and rice production for 31 districts for four years from 2005-06 to 2008-09 was used as validation and testing set of data. The accuracy of the model built by the training set is compared with the validation and test set of data. Validation and testing set also contains similar number of data items like training set. Testing set consisted of 4 years data at the rate of 259 data items per year. The total data items in the testing set is 1036 data items.

Since, the same C program written to implement the FFBPNN system is able to read the one year input file at a time and produce one year output file. Hence, the same C program is executed 4 times separately with input data for 2005-06, 2006-07, 2007-08 and 2008-09 and four different output files are produced for examination.

3.4 Statistical measures in error reduction: Statistical error analyses have been used to assess the performance of the FFBPNN model developed. Four types of error analyses methods were used are: Coefficient of determination, Mean Squared Error, Root Mean Squared Error and Absolute Relative Error.

3.4.1 Coefficient of Determination: The coefficient of determination (R^2) is used to understand the type of fitting between the predicted value and the known value. The value of R^2 varies from 0 to 1. If R^2=1 then the regression line fits very correctly with the data. If R^2=0 then the regression line does not fits well with the data. It provides a measure of how well predicted data fit with the known data. In order to measure the performance of FFBPNN model on a data set, four measures of error were adopted. According to Sanjay R. Bhatikar et al (2000), the formula adopted for computing R^2 and RMSE are given below.

Coefficient of determination $= 1 - \frac{\sum_{i=1}^{N}(t_i - y_i)^2}{\sum_{i=1}^{N}(t_i - t_m)^2}$ eq. (4)

Where

N is the total number of data points, t_i is the observed data

t_m is the mean data and y_{ii} is the model's predicted data.

3.4.2 **Mean Squared Error:** The performance of FFBPNN is studied by updating the weights during back propagation. The effect of iterations on error reduction between the observed data and the predicted data is measured using the Mean Squared Error (MSE). The formula for the MSE is given below:

Mean Squared Error (MSE) $= \frac{1}{N}\sum_{i=1}^{N}(t_i - y_i)^2$ eq. (5)

Where

N is the total number of data points,

t_i is the observed data and

y_{ii} is the model's predicted data

Root Mean Squared Error: The effect of iterations on error reduction pattern between the observed data and the predicted data is measured using the Root Mean Squared Error (RMSE). The formula for the RMSE is given below:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(t_i - y_i)^2} \quad \text{eq. (6)}$$

Where
N is the total number of data points, t_i is the observed data and y_{ii} is the model's predicted data.

3.4.4 Absolute Relative Error: ARE is used with supervised learning. It is the technique of comparing the FFBPNN's output to the observed data. The error is used to train the system for better performance. The error values can be used to directly adjust the weights, using an algorithm such as the back-propagation algorithm. If the system output is FFBPNN output and the desired system output is known observed data supplied. According to Zabir Haider Khan[17] et al. 2011, the Absolute Relative Error (ARE) can be computed using the formula given below:

$$ARE = \frac{1}{N}\sum_{i=1}^{N}\frac{|(t_i - y_i)|}{t_i} \times 100 \qquad \text{eq. (7)}$$

Where
N is the total number of data points, t_i is the observed data and
y_{ii} is the model's predicted data.
Compare the ARE with threshold value of 10^{-9}. If the error was greater than threshold value then calculates the updated weights and compute the summation using back propagation. This process is repeated until error is zero.

3.4.5 **Percentage determination of model:** The best fitting model is obtained by connecting years (x) as independent variable and the area of rice cultivation or rice production (y) as dependent variable for different districts of rice cultivation during Kuruvai, Samba and Kodai seasons. The coefficient of correlation r value will vary from -1 to +1. The coefficient of determination (r^2) will vary from 0 to +1. The nature of fit depends upon the percentage determination. The percentage determination can be obtained by the multiplication r^2 with 100. The percentage of determination $100r^2$ represents the percent of the data that is the closest to the line of best fit.

Percentage of determination (%) = 100 r^2 eq. (8)

3.4.6 Testing the statistical significance: The t-test is used to test the significance between two sets of paired data. The pair consists of FFBPNN predicted data and the best fitted model's predicted area of cultivation / rice production. The calculated t value for N observations was computed as follows:

Let X is the array of values of FFBPNN method of predicted area of cultivation / rice production

Let Y is the array of values of the best fitted model's predicted area of cultivation / rice production

D = X – Y, Compute $\sum D$, $\sum D^2$ and $\bar{D} = \frac{\sum D}{N}$

where N is the number of observations

$$\text{Standard Deviation (SD)} = \sqrt{\frac{\sum D^2 - \frac{(\sum D)^2}{N}}{(N-1)}}$$

$$\text{Standard Error SE} = \frac{SD}{\sqrt{N}}$$

Calculated Paired t value = $\frac{\bar{D}}{SE}$ It is a positive value.

If negative value then omit the minus sign. The Degrees of Freedom = N-1. Refer the statistical t table for (N-1) degree of freedom at 5% level of significance to get the table t value. If calculated t value is less than table t value then there is no significant difference between the FFB-PNN method of predicted area of cultivation and the multiple regression method of predicted area of cultivation. If calculated t value is greater than table t value then there is significant difference between the FFBPNN method of predicted area of cultivation and the multiple regression method of predicted area of cultivation. The same procedure is used to compute the t value between the pair of FFBPNN method of rice production and multi-ple regression method of rice production.

3.5 **Development of the best fitting models:** It is planned in this research work that more realistic prediction can be obtained by integrating the efficient curve expert software into the FFBPNN predicted data so as to get the best fitting models for area of rice cultivation and rice production for different districts during three seasons. The percentage determination of the best fitted models developed both for area of rice cultivation and rice production for different districts will be computed and analyzed. The predicted area of cultivation and rice production from the best fitted models will be examined.

3.5.1 Development of best fitting models for area of rice cultivation: The best fitting model connecting years (x) as independent variable and the area of rice cultivation (y) as dependent variable for different districts of rice cultivation during Kuruvai, Samba and Kodai seasons will be done using curve expert software along with their coefficient of correlation.

3.5.2 Development of best fitting models for rice production: The best fitting model connecting years (x) as independent variable and the rice production (y) as dependent variable for different districts of rice production during Kuruvai, Samba and Kodai seasons will be done using curve expert software along with their coefficient of correlation.

3.5.3 Percentage determination of the best fitted models: The best fitting model is obtained by connecting years (x) as independent variable and the area of rice cultivation or rice production (y) as dependent variable for different districts of rice cultivation during Kuruvai, Samba and Kodai seasons. The coefficient of correlation r value will vary from -1 to +1. The coefficient of determination (r^2) will vary from 0 to +1. The nature of fit depends upon the coefficient of determination (r^2). The r^2 value can be multiplied by 100 and can be expressed as the percentage of variation "explained" by the best fitted model. The percentage of determination $100r^2$ represents the percent of the data that is the closest to the line of best fit.

3.5.3.1 Percent of data closest to the line of best fit for the area of cultivation: The percentage of the data that is the closest to the line of best fit for the area of rice cultivation of different districts for three seasons will be made. The number of districts coming under the class intervals of > 90%, 80 to 90%, 70 to 80%, 60 to 70%, 50 to 60% and < 50% of data closest to the best fitted models developed for area of rice cultivation will be made available.

3.5.3.2 Percent of data closest to the line of best fit for the rice production: The percentage of the data that is the closest to the line of best fit for the rice production of different districts for three seasons will be made. The number of districts coming under the class intervals of > 90%, 80 to 90%, 70 to 80%, 60 to 70%, 50 to 60% and < 50% of data closest to the best fitted models developed for rice production will be made available.

3.6 Development of FFBPNN with integration of curve expert software for prediction: The hybrid FFBPNN system is created by integrating the Curve Expert Soft-ware into the FFBPNN system. Figure 4 shows the hybrid FFBPNN system integrated with Curve Expert system.

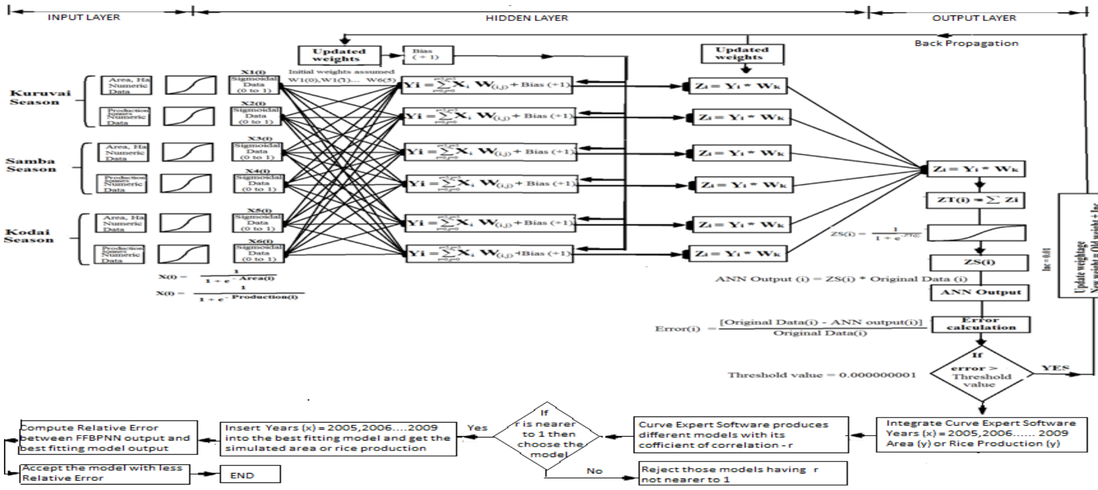

Figure 4. Hybrid FFBPNN system integrated with Curve Expert Software

The predicted output from the FFBPNN system is given as input into the Curve Expert software so as to develop the best fitting linear and non-linear model among the 30 models based on the coefficient of correlation. Further predictions from the hybrid FFBPNN integrated with best fitting models were compared with the FFBPNN outputs also. This makes the system not only a FFBPNN system but also a hybrid FFBPNN system integrated with Curve Expert software for more realistic predictions.

The best fitted models developed and integrated with the hybrid system both for area of rice cultivation and rice production for different districts during the three seasons, the best predicted area of rice cultivation in hectare and rice production in tonnes were obtained by inserting x = (2005, 2006, 2007, 2008 and 2009). The predicted data from the hybrid system will be analyzed. ARE will be worked out between predicted data from the hybrid FFBPNN system and the predicted data from the FFB-PNN system and will be analyzed. Testing of the statistical significance between the hybrid FFBPNN system and the FFBPNN system using paired t-test will be done.

3.6.1 **Prediction of data from the best fitting models of the hybrid system:** The predicted area of cultivation from the hybrid FFBPNN best fitted model will be examined. Some statistics like minimum, maximum mean and standard deviations of the predicted area of rice cultivation for different years during three seasons will be tabulated and discussed.

The predicted rice production from the hybrid FFBPNN best fitted models will be examined. Some statistics like minimum, maximum mean and standard deviations of the predicted rice production for different years during three seasons will be tabulated and discussed.

3.6.2 ARE between FFBPNN method and hybrid FFBPNN best fitted model of prediction: ARE percent was calculated between the FFBPNN method of predicted area of rice cultivation and the hybrid best fitted models predicted area of rice cultivation. Some statistics like minimum, maximum mean and standard deviations of the ARE for area of rice cultivation for three seasons will be tabulated and discussed.

ARE percent was also calculated between the FFBPNN method of predicted rice production and the hybrid FFB-PNN method of best fitted model's predicted area of rice cultivation. Some statistics like minimum, maximum mean and standard deviations of the ARE for rice production for three seasons will be tabulated and discussed.

3.6.3 Testing the statistical significance using paired t-test: The paired t-test was conducted to check the significance between two sets of paired data. The first data field is the FFBPNN method of predicted area of cultivation and the second data field is the hybrid FFBPNN best fitted model's predicted area of cultivation. The two data items for a year form a pair. The calculated t value will be compared with the table t value at 5% level of significance for difference years in three seasons. If the calculated t value is less than the table t value then there is no significant difference between the two samples else there is significant difference exists. This procedure is repeated for different years of rice production for Kuruvai, Samba and Kodai seasons also.

4. RESULTS AND DISCUSSIONS

The data were collected as per the methodology descri-bed in the previous chapter. The data were properly processed in this chapter. The results of data processed are well discussed with suitable data models to arrive at pertinent conclusions. The following are the aspects discussed in this chapter:

1. Development of FFBPNN system to predict the area of rice cultivation and rice production in different districts during three seasons
2. Effect of different statistical measures in error reduction pattern in FFBPNN prediction system.
3. Comparison of the observed data and the FFBPNN predicted data pertaining to the area of rice cultiva-tion and rice production
4. Development of the best fitting models for area of rice and rice production in different districts during three seasons using the Curve Expert Software and to find out the percentage determination of the best fitted models
5. Development of the hybrid FFBPNN system by integrating the best fitting models from Curve Expert Software and to study the predicted data from the hybrid system by means of statistical testing between the observed and predicted data

4.1 Development of FFBPNN system to predict rice data in different districts during three seasons: The FFBPNN architecture was developed as per the methodology (section 3.3) described and a C language application program was also developed to carry out the different unit operations specified in the architecture.

4.2 Statistical error analysis between the observed and predicted data: Statistical error analyses have been used to assess the performance of the error reduction pattern of the FFBPNN model developed in this work. Readers of this paper are advised to refer the published work of Arun Balaji and Manimegalai Vairavan[15], for full information. Some of such error analyses methods used are: coefficient of determination (R^2), Mean Square Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) or also called Absolute Relative Error (ARE). The effect of different statistical measures on the rate of reduction of errors between observed and predicted data in the FFBPNN system is summarized in Table 1

Table 1: Different statistical measures and their rate of reduction of errors

S.No	Statistical techniques	Explanations
1	Coefficient of determination, R^2	R^2 = 1 for all iterations. R^2 does not properly convey at which iteration error between observed and predicted data is 0. Hence R^2 is a poor statistical measure for FFBPNN rice production prediction system
2	Mean Squared Error, MSE	MSE = 0 from 15th iteration onwards to 18th iteration for the different data sets considered. As the iterations increases, there is decrease in MSE. Hence, MSE is a good statistical measure for the error reduction pattern in the FFBPNN for predicting the rice production in Tamilnadu.
3	Root Mean Squared Error, RMSE	RMSE = 0 from 15th iteration onwards to 18th iteration for the different data sets considered. RMSE is a much better statistical measure compared to MSE because more data sets get 0 error compared to MSE.
4	Absolute Relative Error, ARE	ARE =0 for all the data items of area of cultivation and rice productions for the three seasons at the 9th iteration only. Hence, ARE is the best statistical measure used in FFBPNN system to predict rice production.

From table 1, it is clear that the rank of statistical measu-res in the rate of reduction of errors between the observed data and FFBPNN predicted data are in the order of 1) Absolute Relative Error (ARE) 2) Root Mean Squared Error (RMSE) and 3) Mean Squared Error (MSE). It is found that the coefficient of determi-nation, R^2 is the poor statistical measure in the rate of reduction of errors between the observed data and FFBPNN predicted data.

4.2.1 Training and testing data on the FFBPNN system: The C program written to implement the FFBPNN system is able to read the one-year data from an input file at a time and write the results into an output file for that year. Hence, the same C program is executed 6 times separately with input data for 2005-06, 2006-07, 2007-08, 2008-09, 2009-10 and the aver-age of the five years data and six different output files are produced for examination. The output files were examined, it was found that training data and the testing data for the five years also had completed 18 iterations to match the observed data exactly equal to FFBPNN output. Because of completing this research paper within the stipulated pages, the outputs for the testing data were not included in this report. However, it was checked that both training set and testing set took only 18 iterations to make the prediction completely error free with 100% accuracy of observed and predi-cted data. Readers of this paper are requested to refer the published work of Arun Balaji et al.,[12] for full information.

4.3 Comparison of observed and FFBPNN predicted data in different iterations: The predicted data in the 1, 9 and 18th iterations were compared with the observed data and it is shown in Table 2 in the annexure. It shows that the mean district error between the observed training data and FFBPNN predicted data for Kuruvai area is 0.34 ha and Kuruvai production is 1.18 tonnes. The error of Samba season is 2.36 ha and 5.77 tonnes between the observed and predicted data. Similarly, for the Kodai season, the mean district error between the observed training data and FFBPNN predicted data 0.34 ha for area and 1.18tonnes for rice production.

Table 2 in the annexure also shows that the mean district error between the observed training data and FFBPNN predicted data of ninth iteration is 0.01 tonnes for Kuru-vai rice production, 0.03 ha for Samba area of rice cultivation and 0.07tonnes for Samba rice produc-tion. There is zero error found among the Kuruvai area of cultivation and also Kodai area of cultivation and Kodai rice production. This shows that the error is substantially reduced while comparing the first iteration and the ninth iteration.

Table 2 in the annexure again shows the mean district error between the observed training data and FFBPNN predicted data of 18th iteration. It was found that the observed training data is exactly matching with the FFBPNN predicted data since there are zero errors at the end of the 18th iterations. The present research result is similar to the results of the work carried out by Geetha et al.,[18], where the ANN based rainfall predic-tion in Chennai using back propagation neural network, performed well both in train-ing and testing periods. Also, the result of the present research is like the result obtained by Solaiman Karim[19], where the ANN based rainfall-runoff prediction model worked efficiently to predict the river runoff. Hence, the researcher found that there is 100% accuracy in predict-ion both in training and testing phase of the software. The software developed in the present research can be used for Tamilnadu Government's rice prediction studies.

Table 2. Comparison of observed and FFBPNN predicted data in different iterations

No	District	First iteation						9th (middle) iteration						18th (Last) iteration					
		Kuruvai Season		Samba Season		Kodai Season		Kuruvai Season		Samba Season		Kodai Season		Kuruvai Season		Samba Season		Kodai Season	
		Area Ha	Production tonnes	Area Ha	Production tonnes	Area Ha	Production tonnes	Area Ha	Production tonnes	Area Ha	Production tonnes	Area Ha	Productio tonnes	Area Ha	Production onnes	Area Ha	Production tonnes	Area Ha	Production tonnes
1	Kancheepuram	0.70	2.48	2.28	8.22	0.45	1.76	0.01	0.02	0.02	0.08	0.00	0.02	0	0	0	0	0	0
2	Thiruvallur	1.40	4.46	1.19	3.59	0.46	1.59	0.01	0.04	0.01	0.03	0.00	0.02	0	0	0	0	0	0
3	Cuddalore	0.71	2.49	3.31	9.09	0.18	0.56	0.01	0.02	0.03	0.09	0.00	0.00	0	0	0	0	0	0
4	Villupuram	0.94	2.80	4.29	13.69	0.36	1.12	0.01	0.02	0.04	0.12	0.00	0.01	0	0	0	0	0	0
5	Vellore	0.37	1.32	0.61	2.01	0.52	1.78	0.00	0.01	0.01	0.02	0.00	0.02	0	0	0	0	0	0
6	Thiruvannamalai	0.83	2.39	2.17	6.73	0.96	3.09	0.01	0.02	0.02	0.06	0.01	0.03	0	0	0	0	0	0
7	Salem	0.30	1.14	0.53	2.46	0.10	0.40	0.00	0.01	0.00	0.02	0.00	0.00	0	0	0	0	0	0
8	Namakkal	0.11	0.46	0.28	1.23	0.01	0.03	0.00	0.00	0.00	0.01	0.00	0.00	0	0	0	0	0	0
9	Dharmapuri	0.28	1.11	0.40	1.57	0.09	0.31	0.00	0.01	0.00	0.02	0.00	0.00	0	0	0	0	0	0
10	Krishnagiri	0.25	1.05	0.32	1.23	0.05	0.17	0.00	0.01	0.00	0.01	0.00	0.00	0	0	0	0	0	0
11	Coimbatore	0.05	0.17	0.06	0.20	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0	0	0	0	0	0
12	Thiruppur	0.02	0.06	0.36	1.57	0.07	0.25	0.00	0.00	0.00	0.02	0.00	0.00	0	0	0	0	0	0
13	Erode	0.35	1.52	1.03	4.16	0.05	0.20	0.00	0.02	0.01	0.04	0.00	0.00	0	0	0	0	0	0
14	Tiruchirapalli	0.26	0.92	2.08	8.75	0.08	0.35	0.00	0.01	0.02	0.08	0.00	0.00	0	0	0	0	0	0
15	Karur	0.00	0.00	0.77	3.30	0.02	0.08	0.00	0.00	0.01	0.04	0.00	0.00	0	0	0	0	0	0
16	Perambalur	0.03	0.09	0.38	1.35	0.04	0.15	0.00	0.00	0.00	0.01	0.00	0.00	0	0	0	0	0	0
17	Ariyalur	0.05	0.17	0.92	2.81	0.01	0.05	0.00	0.00	0.01	0.02	0.00	0.00	0	0	0	0	0	0
18	Pudukottai	0.02	0.05	3.55	7.00	0.01	0.04	0.00	0.00	0.03	0.06	0.00	0.00	0	0	0	0	0	0
19	Thanjavur	0.83	2.50	5.09	12.91	0.20	0.65	0.01	0.02	0.05	0.12	0.00	0.01	0	0	0	0	0	0
20	Thiruvarur	0.50	1.56	5.47	15.59	0.28	1.02	0.00	0.02	0.05	0.16	0.00	0.01	0	0	0	0	0	0
21	Nagapattinam	1.03	3.08	4.88	13.06	0.03	0.10	0.01	0.03	0.05	0.12	0.00	0.00	0	0	0	0	0	0
22	Madurai	0.15	0.58	1.62	6.55	0.15	0.54	0.00	0.01	0.02	0.06	0.00	0.00	0	0	0	0	0	0
23	Theni	0.21	1.14	0.34	1.69	0.03	0.09	0.00	0.01	0.00	0.02	0.00	0.00	0	0	0	0	0	0
24	Dindigul	0.06	0.28	0.50	2.40	0.06	0.29	0.00	0.00	0.00	0.02	0.00	0.00	0	0	0	0	0	0
25	Ramanathapuram	0.00	0.00	9.78	1.70	0.11	0.37	0.00	0.00	0.14	0.02	0.00	0.00	0	0	0	0	0	0
26	Virudhunagar	0.00	0.00	1.84	4.40	0.12	0.37	0.00	0.00	0.03	0.06	0.00	0.01	0	0	0	0	0	0
27	Sivagangai	0.00	0.00	16.30	29.12	0.00	0.00	0.00	0.00	0.34	0.61	0.00	0.00	0	0	0	0	0	0
28	Tirunelveli	0.70	2.84	2.47	10.28	0.07	0.28	0.01	0.02	0.02	0.09	0.00	0.00	0	0	0	0	0	0
29	Thoothukudi	0.26	1.30	0.43	2.12	0.05	0.22	0.00	0.01	0.00	0.02	0.00	0.00	0	0	0	0	0	0
30	The-Nilgiris	0.12	0.43	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0	0	0	0	0	0
31	Kanyakumari	0.01	0.06	0.01	0.06	0.00	0.00	0.01	0.06	0.01	0.06	0.00	0.00	0	0	0	0	0	0
	Mean Error	**0.34**	**1.18**	**2.36**	**5.77**	**0.15**	**0.51**	**0.00**	**0.01**	**0.03**	**0.07**	**0.00**	**0.00**	0	0	0	0	0	0

4.4 **Development of best fitting models to connect to output of FFBPNN system:** More realistic prediction can be obtained by integrating the efficient curve expert software into the FFBPNN. The curve expert software comprises of 30 different linear and non-linear models. Models were developed for area of rice cultivation and rice production for 31 districts in 3 seasons as per Arun Balaji et al.,[14].

From table 3 in the annexure, it was found that the best fitting equation various from seasons to seasons and from the district to district depending upon the nature of area of rice cultivation in hectare with respect to years of rice cultivation. The more number of fits are found in the order of 1) Quadratic Fit, 2) Linear Fit 3) User-Defined Model 4) Saturation Growth-Rate Model 5) Logarithm Fit 6) Hyperbolic Fit 7) Exponential Fit and 8) Gaussian Model. The nature of fit depends upon the coefficient of correlation (r). The value of r is 0.9962374 for Samba season of Kanyakumari district, which is very close to 1, it means the nature of fit is a perfect with highest reliability of 99.62% for prediction and simulation using the hyperbolic model developed. If the value of r is close to zero, then the model is not the best model and its reliability for prediction is poor.

Table 3. Best fitting model for area of rice cultivation

	District	Kuruvai season -area of rice	Samba season- area of rice	Kodai season- area of rice
1	Kancheepuram	Linear Fit:	Quadratic Fit:	Quadratic Fit:

		$y = a + bx$ Where: a = 1743686.1 b = -859.1 SE=2589.6788153 r=0.5180583	$y = a + bx + cx^2$ Where: a =2.584097e+009 b =-2573281.5, c =640.64286 SE=1387.7771642 r=0.9505217	$y = a + bx + cx^2$ Where: a =7.996972e+009 b =-7965705.5, c =1983.6429 SE=2694.2101729 r=0.9594494
2	Thiruvallur	Linear Fit: $y = a + bx$ Where: a = 1819791.2 b = -886.8 SE=2800.7836284 r=0.5004716	Quadratic Fit: $y = a + bx + cx^2$ Where: a = 6.1498e+009 b = -6126986.2 c = 1526.0714 SE=2346.3276227 r=0.9059931	Quadratic Fit: $y = a + bx + cx^2$ Where: a =4.3037151e+009 b = -4286962.5 c = 1067.5714 SE=1325.0531737 r=0.9638211
3	Cuddalore	Quadratic Fit: $y = a + bx + cx^2$ Where: a =2.8493559e+009 b = -2838075.3 c = 706.71429 SE= 784.3524899 r =0.9758233	Quadratic Fit: $y = a + bx + cx^2$ Where: a =6.4814269e+009 b = -6459923.7 c = 1609.6429 SE=2360.9653233 r =0.9046873	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 68.805182 b = -1979.2248 SE= 171.0517925 r =0.8834125
4	Villupuram	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 157.01568 b = -1995.8502 SE=1848.8717797 r =0.9285353	Quadratic Fit: $y = a + bx + cx^2$ Where: a = 8.5909598e+009 b = -8560802.7 c = 2132.7143 SE=1392.5897565 r=0.9709031	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 39.275129 b = -1999.7409 SE=2803.5182283 r=0.6824962
5	Vellore	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-2.1056772e+009 b = 2097829.3 c = -522.5 SE= 503.0429405 r=0.9630275	Quadratic Fit: $y = a + bx + cx^2$ Where: a =2.7606796e+009 b = -2751172.1 c = 685.42857 SE= 828.8159713 r=0.9117379	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 44.279487 b = -2001.8562 SE=4982.8750880 r=0.7877935
6	Thiruvannamalai	Quadratic Fit: $y = a + bx + cx^2$ Where: a =4.7570186e+009 b = -4738750.2 c = 1180.1429 SE= 992.4629248 r=0.9796578	Quadratic Fit: $y = a + bx + cx^2$ Where: a =1.3046268e+010 b = -12994865 c = 3235.9286 SE=8826.9221320 r=0.8707363	Linear Fit: $y = a + bx$ Where: a = 1971864.3 b = -965.99987 SE=9115.0129493 r=0.1899669
7	Salem	Quadratic Fit: $y = a + bx + cx^2$ Where: a =2.1471009e+009 b = -2139103.6 c = 532.78571 SE= 918.9074103 r=0.8908108	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 89.042528 b = -1996.771 SE=3034.1732766 r=0.7136509	Quadratic Fit: $y = a + bx + cx^2$ Where: a =2.2547174e+009 b = -2246629.3 c = 559.64286 SE= 867.6229927 r=0.8742354
8	Namakkal	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-3.389562e+009 b = 3378194.3 c = -841.71429 SE=1087.3815468 r=0.9134530	Quadratic Fit: $y = a + bx + cx^2$ Where: a =4.3700042e+009 b = -4353683.3 c = 1084.3571 SE= 538.0465195 r=0.9898150	Quadratic Fit: $y = a + bx + cx^2$ Where: a =6.0526458e+008 b = -602913.43 c = 150.14286 SE= 83.2613784 r=0.9923142
9	Dharmapuri	Quadratic Fit: $y = a + bx + cx^2$ Where: a =3.2453323e+009 b = -3233498.9 c = 805.42857 SE=2049.8040186 r=0.7624569	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 102.2878 b = -1989.9121 SE=1427.7431696 r=0.6705804	Quadratic Fit: $y = a + bx + cx^2$ Where: a =2.3771725e+009 b = -2368283.4 c = 589.85714 SE= 141.2503349 r=0.9976424
10	Krishnagiri	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 48.74027 b = -1992.9109 SE= 564.9581785 r=0.8452760.	Quadratic Fit: $y = a + bx + cx^2$ Where: a =2.5389341e+009 b = -2529302.2 c = 629.92857 SE= 315.1166679 r=0.9914517	Quadratic Fit: $y = a + bx + cx^2$ Where: a =9.4119911e+008 b = -937735.71 c = 233.57143 SE= 75.8004900 r=0.9947594

11	Coimbatore	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-3.611354e+008 b = 359851.93 c = -89.642857 SE= 78.8850882 r=0.9514829	Linear Fit: $y = a + bx$ Where: a = 1792512.4 b = -891.6 SE=1140.2577486 r=0.8190485	Quadratic Fit: $y = a + bx + cx^2$ Where: a =4.7931096e+008 b = -477222.36 c = 118.78571 SE= 59.8867980 r=0.9981491
12	Thiruppur	Newly formed district. It has no data to process and hence omitted		
13	Erode	Quadratic Fit: $y = a + bx + cx^2$ Where: a =1.2303838e+009 b = -1225176.4 c = 305 SE= 224.7536429 r=0.9947346	Quadratic Fit: $y = a + bx + cx^2$ Where: a =1.8457344e+009 b = -1838841.2 c = 458 SE= 654.9238124 r=0.9207655	Logarithm Fit: $y = a + b * log(x)$ Where: a = -3067687.1 b = 403605.58 SE= 579.8141688 r=0.5350237
14	Tiruchirapalli	Exponential Fit: $y = ae^{bx}$ Where: a = 9.2118133e-017 b = 0.022763071 SE= 888.3654415 r=0.2898737	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =755.8135, b = -1981.8883 SE=3358.8309585 r=0.7895731	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =4.6025674, b = -2003.5755 SE=1499.9498967 r=0.7727203
15	Karur	Rice not cultivated in this season	Quadratic Fit: $y = a + bx + cx^2$ Where: a =3.3985579e+009 b = -3386250.3 c = 843.5 SE=1710.0211694 r=0.8189261	Quadratic Fit: $y = a + bx + cx^2$ Where: a =2.1588863e+008 b = -215085.51 c = 53.571429 SE= 81.5666248 r=0.9110592
16	Perambalur	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a = 3155.3325 b = 2006.2877 c = 1.223665 SE= 803.1227162 r=0.8776631	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a = 43055.881 b = 2005.3882 c = 1.8182356 SE=6782.4260734 r=0.9545845	Hyperbolic Fit: $y = a + \frac{b}{x}$ Where: a = 366562.7 b = -7.3282097e+008 SE=1011.8276980 r=0.3118982
17	Ariyalur	Newly formed district. It has no data to process and hence omitted		
18	Pudukottai	User-Defined Model: $y = a + b * x$ Where: a = 544601.52 b = -270.90001 SE= 105.6609357 r=0.9779332	Quadratic Fit: $y = a + bx + cx^2$ Where: a =4.4404662e+009 b = -4425305.1 c = 1102.5714 SE=3448.2304612 r=0.6640203	Quadratic Fit: $y = a + bx + cx^2$ Where: a = -47800181 b = 47614.171 c = -11.857143 SE= 37.0898523 r=0.8237899
19	Thanjavur	Hyperbolic Fit: $y = a + \frac{b}{x}$ Where: a = 452075.03 b =-8.5278799e+008 SE=8265.1625434 r=0.0467154	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =-4290.354, b = -2072.7895 SE=3179.7561814 r=0.7603522	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =-8.1583351, b = -2012.0108 SE= 602.0275783 r=0.9229136
20	Thiruvarur	Logarithm Fit: $y = a + b * log(x)$ Where: a = -6801594.8 b = 896950.64 SE=13174.5365883 r=0.0613353	Quadratic Fit: $y = a + bx + cx^2$ Where: a =1.744785e+009 b = -1739456.9 c = 433.57143 SE=1382.8651829 r=0.8584080	Quadratic Fit: $y = a + bx + cx^2$ Where: a =4.105273e+009 b = -4091756.3 c = 1019.5714 SE= 207.5579782 r=0.9979556
21	Nagapattinam	Hyperbolic Fit: $y = a + \frac{b}{x}$ Where: a = 1975008 b = -3.9069138e+009 SE=4518.4469613 r=0.3648908	Linear Fit: $y = a + bx$ Where: a = 909925.1 b = -387.5 SE=3259.3085514 r=0.2121232	Quadratic Fit: $y = a + bx + cx^2$ Where: a =7.2240924e+008 b =-719771.06, c = 179.28571 SE= 345.8650439 r=0.8434410
22	Madurai	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-6.3849087e+009 b = 6362707 c = -1585.1429 SE=1795.3686133 r=0.9193692	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 310.53959 b = -1994.8174 SE=3181.0254633 r=0.924904	Hyperbolic Fit: $y = a + \frac{b}{x}$ Where: a = 353497.84 b =-7.0222335e+008 SE= 357.0785864 r=0.6653966

23	Theni	Quadratic Fit: $y = a + bx + cx^2$ Where: a =2.8209436e+008 b = -281186.01 c = 70.071429 SE= 37.0069492 r=0.9899238	Quadratic Fit: $y = a + bx + cx^2$ Where: a =4.5825371e+008 b = -456404.33 c = 113.64286 SE= 229.1105784 r=0.9377955	User-Defined Model: $y = a + b * x$ Where: a = -326720.79 b = 162.99999 SE= 219.9733317 r=0.8041618
24	Dindigul	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 5.4841938 b = -2001.5997 SE= 155.9981397 r=0.9811653.	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 319.68103 b = -1961.5876 SE=1683.4326684 r=0.3027815	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 6.3306584 b = -2002.0926 SE= 349.1344564 r=0.9572692
25	Ramanathapuram	Rice not cultivated in this season	Quadratic Fit: $y = a + bx + cx^2$ Where: a =5.7880025e+009 b = -5768621.3 c = 1437.3571 SE=2787.2685660 r=0.8411164	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 6.7862166 b = -1999.2383 SE= 119.2505911 r=0.9648408
26	Virudhunagar	Rice not cultivated in this season	Linear Fit: $y = a + bx$ Where: a = 1377377.5 b = -672.5 SE=2447.3137314 r=0.4484271	Linear Fit: $y = a + bx$ Where: a = 420703.8 b = -208.4 SE= 469.5595099 r=0.6295630
27	Sivagangai	Rice not cultivated in this season	Quadratic Fit: $y = a + bx + cx^2$ Where: a =5.5541552e+009 b = -5532138.5 c = 1377.5714 SE=3143.8364961 r=0.9076624	Rice not cultivated in this season
28	Tirunelveli	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a = 26376.013 b = 2007.136 c = 2.4711389 SE=2200.3048107 r=0.9106846	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = -1068.1103 b = -2043.1922 SE=5652.0152824 r=0.4697381	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 5.9480122 b = -2004.0868 SE=2315.1486255 r=0.8938372
29	Thoothukudi	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 151.58641 b = -1965.9526 SE= 990.7883960 r=0.3041219	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = -100.4256 b = -2026.1233 SE=1593.2949601 r=0.5357672	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a = 3603.8834 b = 2006.4299 c = 1.2971632 SE= 637.4731849 r=0.9171349
30	The-Nilgiris	Quadratic Fit: $y = a + bx + cx^2$ where: a =-1.0483882e+008 b =104705.83, c =-26.142857 SE= 42.9378288, r=0.9966401	Rice not cultivated in this season	Rice not cultivated in this season
31	Kanyakumari	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a =10771.465, b =2004.4443 c =6.4430537, SE= 284.0898809 r=0.9782834	Hyperbolic Fit: $y = a + \frac{b}{x}$ Where: a =-1433906.2 b =2.8981968e+009 SE= 114.2776838 r=0.9962374	Rice not cultivated in this season

Where SE – Standard Error and r – correlation coefficient

Similarly, the best fitting models developed for rice production is given in Table 4 in the annexure. From table 4 in the annexure, it was found that the nature of fit depends upon the coefficient of correlation (r). The value of r is 0.9934423 for Samba season of Kanyakumari district, which is very close to 1, it means the nature of fit is a perfect with highest reliability of 99.34% for prediction and simulation using the Saturation Growth-Rate Model developed.

Table 4. Best fitting model for rice production

	District	Kuruvai – rice production	Samba– rice production	Kodai – rice production
1	Kancheepuram	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-1.3871188e+010 b = 13821687 c = -3443.0714 SE=3464.7662421 r=0.9394151	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-5.1022184e+010 b = 50836157 c = -12662.643 SE=41198.7937160 r=0.6809555	Quadratic Fit: $y = a + bx + cx^2$ Where: a =1.7715959e+010 b = -17645532 c = 4393.8571 SE=2914.8146511 r=0.9916769
2	Thiruvallur	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-3.4400411e+010 b =34281651 c =-8540.7857 SE=10180.2940032 r=0.9125863	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-1.8002776e+010 b = 17937323 c = -4468 SE=13217.5046363 r=0.7105419	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 197.42837 b = -1998.2679 SE=8301.4791891 r=0.7588055
3	Cuddalore	Quadratic Fit: $y = a + bx + cx^2$ Where: a =6.5116541e+009 b =-6485888.2 c =1615.0714 SE=5163.3082530 r=0.8382550	Linear Fit: $y = a + bx$ Where: a = -15238047 b = 7700 SE=82719.8343873 r=0.1675473	Linear Fit: $y = a + bx$ Where: a = 1291200.9 b = -635.3 SE=1820.9775671 r=0.5372348
4	Villupuram	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =386.34995, b =-998.3419 SE=9987.8920141 r=0.8878347	Linear Fit: $y = a + bx$ Where: a = -2129409.4, b =1237 SE=22359.6548214 r=0.1004939	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =119.69224, b = -2000.0286 SE=8133.8383235 r=0.7363289
5	Vellore	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a = 37502.553 b = 2008.003 c = 2.8587991 SE= 483.8331549 r=0.9985916.	Linear Fit: $y = a + bx$ Where: a = -3182795.2 b = 1613.9001 SE=6984.3258109 r=0.3887064	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 219.76119 b = -1999.6333 SE=15111.6080050 r=0.6922441.
6	Thiruvannamalai	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 715.87884 b = -1986.2656 SE=8549.1752757 r=0.5775305	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 1118.2917 b = -1994.4858 SE=34863.8974163 r=0.6085882	Linear Fit: $y = a + bx$ Where: a = 1621076.5 b = -754.7 SE=24807.0350795 r=0.0554587
7	Salem	Quadratic Fit: $y = a + bx + cx^2$ Where: a =1.250018e+010 b = -12454424, c =3102.2143 SE=4315.1006977 r=0.9104068	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =1658.7837, b = -1958.9415 SE=16703.7801640 r=0.1412975	Quadratic Fit: $y = a + bx + cx^2$ Where: a =9.5186637e+009 b =-9484697.9, c =2362.7143 SE=2428.7502611 r=0.9363828
8	Namakkal	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-1.2402831e+010 b =12359497, c =-3079.0714 SE=10734.9313600 r=0.6046721	Quadratic Fit: $y = a + bx + cx^2$ Where: a =1.4715317e+010 b =-14660975, c =3651.7143 SE=4008.6395488 r=0.9464677	Quadratic Fit: $y = a + bx + cx^2$ Where: a =1.7260448e+009 b =-1719294.8, c =428.14286 SE= 393.0412919 r=0.9809694
9	Dharmapuri	Quadratic Fit: $y = a + bx + cx^2$ Where: a =9.1154397e+009 b=-9084513.9, c=2263.4286 SE=7664.0810610 r=0.6355731	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a =46655.287, b=2006.8182 c = 3.9666421 SE=4848.4441875 r=0.6094944	Quadratic Fit: $y = a + bx + cx^2$ Where: a =6.4592346e+009 b =-6435285, c=1602.8571 SE= 345.1985763 r=0.9978800
10	Krishnagiri	Quadratic Fit: $y = a + bx + cx^2$ Where: a =6.0730068e+009 b =-6052458.6, c=1508 SE= 972.2065622 r=0.9748071	Quadratic Fit: $y = a + bx + cx^2$ Where: a =5.1730376e+009 b =-5156050, c=1284.7857 SE=1480.4048869 r=0.9423660	Quadratic Fit: $y = a + bx + cx^2$ Where: a =2.2338987e+009 b =-2225933.9, c =554.5 SE= 255.3617042 r=0.9861086
11	Coimbatore	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where:	Quadratic Fit: $y = a + bx + cx^2$ Where:	Hyperbolic Fit: $y = a + \frac{b}{x}$

		a =6468.7172, b =2007.3562 c = 1.9595669 SE= 264.8881852 r=0.9897040	a =-8.5114829e+009 b = 8483131.3 c = -2113.7143 SE=1321.7468093 r=0.9788308	Where: a =-1974558.6 b =3.9688497e+009 SE=1658.7338162 r=0.7351699
12	Thiruppur	Newly formed district. It has no data to process and hence omitted		
13	Erode	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =247.58621, b=-1996.2721 SE=2574.3466085 r=0.9519910	User-Defined Model: $y = a + b * x$ Where: a = -5146607.1 b = 2623.9004 SE=17124.3418355 r=0.2694082	User-Defined Model: $y = a + b * x$ Where: a = -1024356.7 b = 513.0999 SE=2121.7677143 r=0.4038979
14	Tiruchirapalli	Logarithm Fit: $y = a + b * log(x)$ Where: a = -14430152 b = 1900777.6 SE=3019.5668927 r=0.4969286	User-Defined Model: $y = a + b * x$ Where: a = -9006419.3 b = 4596.5 SE=27175.7867982 r=0.2950570	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =26.145095, b = -2002.1593 SE=5427.8558798 r=0.5784679.
15	Karur	Rice not cultivated in this season	Quadratic Fit: $y = a + bx + cx^2$ Where: a =9.667602e+009 b = -9638306.3 c = 2402.2857 SE=6765.6708925 r=0.8682850	Quadratic Fit: $y = a + bx + cx^2$ Where: a =5.5265221e+008 b = -550607.43 c = 137.14286 SE= 217.5804613 r=0.8987622
16	Perambalur	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a = 8585.5609 b = 2006.7458 c = 1.4754603 SE= 540.9664491 r=0.9872764	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-3.1176081e+010 b = 31083675 c = -7747.8571 SE=21005.0044628 r=0.8929604	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = -13.866651 b = -2014.7557 SE=2085.3351772 r=0.4087852
17	Ariyalur	Newly formed district. It has no data to process and hence omitted		
18	Pudukottai	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a = 3898.5473 b = 2005.7047 c = 2.3179513 SE= 520.9259325 r=0.9422405	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 1097.9956 b = -1996.0634 SE=45308.9709600 r=0.6078088	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-2.2745933e+008 b = 226585.69 c = -56.428571 SE= 176.4180101 r=0.8002307
19	Thanajavur	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-3.6123823e+010 b = 35995165 c = -8966.7143 SE=32878.4360089 r=0.5978485P	Linear Fit: $y = a + bx$ Where: a = 48308990 b = -23895.6 SE=108909.2751140 r=0.371857	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = -26.301516 b = -2011.9805 SE=1336.9649901 r=0.9626932
20	Thiruvarur	Hyperbolic Fit: $y = a + \frac{b}{x}$ Where: a = 14741116 b =-2.9458977e+010 SE=48251.8201462 r=0.2667021	Linear Fit: $y = a + bx$ Where: a = 2661561.4 b = -1181.8 SE=197553.5226417 r=0.0109213	Quadratic Fit: $y = a + bx + cx^2$ Where: a =1.227815e+010 b =-12239153, c =3050.0714 SE= 222.3801128 r=0.9998217
21	Nagapattinam	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a = 111755.2 b = 2007.6056 c = 2.1870287 SE=18839.4140761 r=0.8416800	Linear Fit: $y = a + bx$ Where: a = -20435712 b = 10306.9 SE=232688.8242229 r=0.0806077	Quadratic Fit: $y = a + bx + cx^2$ Where: a =2.0430858e+009 b = -2035671.6 c = 507.07143 SE=1351.6446702 r=0.7400011
22	Madurai	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-2.752215e+010 b = 27425929 c = -6832.5 SE=7683.2588924 r=0.9204029	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 2354.2097 b = -1982.455 SE=23528.8094029 r=0.5154196	Hyperbolic Fit: $y = a + \frac{b}{x}$ Where: a =2452383.5 b =-4.897007e+009 SE=2557.6749244 r=0.6554285

23	Theni	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = -161.04523 b = -2019.236 SE=2594.7649256 r=0.8442289	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = -621.45162 b = -2037.0949 SE=3582.5315113 r=0.5705672	User-Defined Model: $y = a + b * x$ Where: a = -1099325.8 b = 548.50004 SE= 678.7312428 r=0.8277842
24	Dindigul	Linear Fit: $y = a + bx$ Where: a = 1220998.9 b = -604.3 SE=1759.8986619 r=0.5311614	Quadratic Fit: $y = a + bx + cx^2$ Where: a =6.6052998e+009 b = -6584592.2 c = 1641 SE=2858.3139088 r=0.9233027	User-Defined Model: $y = a + b * x$ Where: a = 3879971.4 b = -1927.2002 SE=2644.9761688 r=0.7993399
25	Ramanathapuram	Rice not cultivated in this season	Quadratic Fit: $y = a + bx + cx^2$ Where: a =1.4186619e+011 b =-1.4136616e+008 c = 35217 SE=86562.8849508 r=0.7351633	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = 19.827254 b = -2000.5996 SE= 620.4433723 r=0.9527623
26	Virudhunagar	Rice not cultivated in this season	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a = 119166.47 b = 2006.3098 c = 2.1996714 SE=31888.6369435 r=0.7283510	Hyperbolic Fit: $y = a + \frac{b}{x}$ Where: a =-3087383.7 b =6.2128575e+009 SE= 330.8032185 r=0.9931708
27	Sivagangai	Rice not cultivated in this season	Quadratic Fit: $y = a + bx + cx^2$ Where: a =7.5903579e+010 b =-75630109, c =18839.429 SE=26240.3193779 r=0.8976911	Rice not cultivated in this season
28	Tirunelveli	Gaussian Model: $y = ae^{\frac{-(b-x)^2}{2c^2}}$ Where: a =108359.8, b=2007.2511 c = 2.2487998 SE=8329.4578372 r=0.9440887	Exponential Fit: $y = ae^{bx}$ Where: a =1.0707657e-028 b = 0.038263333 SE=36748.0517464 r=0.4036699	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =37.508174, b = -2002.5215 SE=9139.9342904 r=0.6549262
29	Thoothukudi	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = -384.50902 b = -2030.2731 SE=6031.8592010 r=0.3963505	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a = -243.63815 b = -2017.4643 SE=9166.7550685 r=0.6730428	Linear Fit: $y = a + bx$ Where: a = 2873553.3 b = -1427.7 SE=4794.9353941 r=0.4776079
30	The-Nilgiris	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-7.5158574e+008 b=749649.79, c =-186.92857 SE=413.6753730, r=0.9682446	Rice not cultivated in this season	Rice not cultivated in this season
31	Kanyakumari	Quadratic Fit: $y = a + bx + cx^2$ Where: a =-4.6177021e+009 b =4602849.8, c = -1147 SE=1372.4398710, r=0.9474092	Saturation Growth-Rate Model: $y = \frac{ax}{(b+x)}$ Where: a =247.19303, b = -1995.8921 SE= 869.4616473 r=0.9934423	Rice not cultivated in this season

Where SE – Standard Error and r – correlation coefficient

4.5 Statistical t-test between observed data and best fitting models: The paired t-test as explained in methodology was used to test the significance between two sets of paired data. The first data field is the observed area of cultivation and the second data field is the predicted data from the FFBPNN system with best fitted model. The two data items for a year form a pair. The same procedure is used to compute the t value between the pair of observed data and the FFBPNN system integrated with best fitted models predicted rice production.

4.5.3.1 Paired t test between the observed and FFBPNN with best fitting models predicted area of rice: The t-test was conducted between the observed and the FFBPNN system integrated with best fitted model's predicted area of rice cultivation. There are 25 observations for Kuruvai season. The table t value for the degrees of freedom of 24 at 5% level of significance was taken up

from the t table. If the calculated t value is less than the table t value then there is no significant difference between the two samples. This procedure is repeated for different years of rice cultivation for Kuruvai, Samba and Kodai seasons. The summary of the t test is given in Table 5 in the annexure.

Table 5 in the annexure shows the fact that the calculated t value is less than the table t value at 5% level of significance for 2005, 2006 and 2009 in Kuruvai season, i-e there is 95% confidence level between the observed and the FFBPNN with best fitted model's predicted area of rice cultivation for 2005, 2006 and 2009. It is interpreted that there is statistically no significant difference between observed and the FFBPNN system integrated with best fitted models predicted area of rice cultivation for 2005, 2006 and 2009. There is significant difference exist for 2007 and 2008 because calculated t is more than table t value.

Table 5. Result of t test for the area of rice cultivation

Season	No. of districts	DF	Table t at 5% level	Calculated t value for area of cultivation for different years					Remarks
				2005	2006	2007	2008	2009	
Kuruvai	25	24	2.064	1.586	1.856	2.392	2.09 4	1.837	Significant difference 2007 & 2008.
Samba	28	27	2.052	0.423	0.082	1.904	1.23 4	0.565	No Significant difference for all years
Kodai	26	25	2.060	1.571	2.298	2.073	2.29 4	0.784	Significant difference 2006,2007,2008

Note: DF: Degrees of Freedom

With regard to the Samba season, the calculated t value is less than the table t value at 5% level of significance for all the years. The predicted area of rice cultivation by observed data and the FFBPNN system integrated with best fitted model are insignificant for all the years. It means that there is statistically no significant difference between observed and the FFBPNN system integrated with best fitted model of predicted area of rice cultivation for all the years.

Regarding the Kodai season, the calculated t value is less than the table t value at 5% level of significance for the years 2005 and 2009. There is statistically no significant difference between the observed area of rice and the FFBPNN system integrated with best fitted model of predicted area for the years 2005 and 2009. For the year 2006, 2007 and 2008, it was found that the calculated t value is greater than the table t value, it means that there is statistically significant difference between the observed data and the FFBPNN system integrated with best fitted model of predicted area of cultivation for the year 2006, 2007 and 2008.

It is found that totally there are 15 t-test conducted beca-use of three seasons and five years. Among the 15 tests, 2 years in Kuruvai season and 3 years in Kodai seasons are showing significant difference and the rest 10 tests out of 15 tests showed insignificant difference between the observed data and the FFBPNN system integrated with best fitted model of predicted area of cultivation. That is 10 out of 15 t –tests (67% of fitting) are very well fitted. If more number of input data for our analysis is taken then there is possibility to get more than 67% of better fittings.

4.5.3.2 Paired t test between observed and best fitting models predicted rice production: The t-test was conducted between the observed rice production and the FFBPNN system integrated with best fitted models predicted rice production. There are 25 observations for Kuru-vai season. The table t value for the degrees of freedom of 24 at 5% level of significance was taken up from the t table. If the calculated t value is less than the table t value then there is no significant difference between the two samples. This procedure is repeated for different years of rice cultivation for Kuruvai, Samba and Kodai seasons. The summary of the t test is given in Table 6 in the annexure. Table 6 in the annexure shows that for the Kuruvai season, the calculated t value is less than the table t value at 5% level of significance for all the years. It means that there is statistically no significant difference between observed rice production and FFBP-NN system integrated with best fitted models predicted rice production for all the years.

About Samba season, the calculated t value is less than the table t value at 5% level of significance for 2008 and 2009. It is interpreted that there is statistically no significant difference between observed rice production and FFBPNN system integrated with best fitted models predicted rice production for 2008 and 2009. There is significant difference exist for 2005, 2006 and 2007 because calculated t is more than table t value.

Table 6. Result of t test for the rice production

Season	No. of districts	DF	Table t at 5% level	Calculated t value for rice production for different years					Remarks
				2005	2006	2007	2008	2009	
Kuruvai	25	24	2.064	0.949	0.343	1.859	2.449	6.601	No Significant difference for all years
Samba	28	27	2.052	2.523	2.852	2.608	1.521	1.839	Significant difference for 2005,2006 and 2007
Kodai	26	25	2.060	1.347	1.313	1.961	2.324	0.944	Significant difference for 2008 only

Note: DF: Degrees of Freedom

Regarding the Kodai season, the calculated t value is less than the table t value at 5% level of significance for all the years excepting 2008. There is statistically no significant difference between the observed rice production and the hybrid FFBPNN system with best fitted model of predicted rice production for all the years excepting 2008. For the year 2008, it was found that the calculated t value is greater than the table t value; it means that there is statistically significant difference between the observed rice production and FFBPNN system integrated with best fitted models predicted rice production.

It is found that there are 15 t-test conducted because of three seasons and five years. Among the 15 tests, 3 years in Samba season and 1 year in Kodai seasons are showing significant difference and the rest 11 tests out of 15 tests showed insignificant difference between the observed rice production and the FFBPNN system integrated with best fitted models predicted rice production. That is 11 out of 15 t tests (73.3% of our fitting) are very well fitted. If more number of year data for our analysis is taken then there is possibility to get more than 73.3% of better fittings.

5. CONCLUSIONS

During this research, FFBPNN architecture was designed and developed. A program in C++ was developed to implement the FFBPNN system developed. The training data and the testing data were used as input and the output data was recorded in a sequential file. Initial weights were assumed from 0 to 1. Weights were updated with the increment of 0.01 when the observed data was not equal to the predicted data by feed forward back propagation system. Each iteration predicted the data and it was com-pared with the observed data. The different statistical measures like R^2, MSE, RMSE and ARE were computed until the error between observed and predicted data becomes zero. It was found that ARE is zero for all the data items of area of cultivation and rice productions for the three seasons at the 9th iteration itself and hence ARE is the best statistical measure for the FFBPNN system of prediction.

It was found that the targeted data is exactly matching with predicted data at 18th iteration. It means the training of FFBPNN software is perfectly done. The software under-stood the complexities, non-linearity and structure of training data with 100% accuracy. Similarly, the test data also was exactly matched with its predicted data. It was found that the software developed in the present research works with 100% accuracy for both training and testing data and hence it is used for Tamilnadu Government's rice prediction studies.

The FFBPNN system was integrated with best fitting models developed from the curve expert software for three seasons of all districts. These developed models were used to simulate the best predicted area of rice cultivation and rice production. The type of fits is found to be 1) Quadratic Fit, 2) Linear Fit 3) User-Defined Model 4) Saturation Growth-Rate Model 5) Logarithm Fit 6) Hyperbolic Fit 7) Exponential Fit and 8) Gaussian Model.

The predicted data for area of cultivation of rice and rice production is compared with the observed area of cultivation and rice production for three seasons of 31 districts using paired t test. About the area of cultivation of rice, there is 67% of fittings are showing insignificant difference between the observed area of rice and predicted area of rice cultivation. Regarding the rice production data, there is 73.3% of fittings are showing insignificant differ-ence between the observed rice production and predicted rice production. If more number of input data for our analysis is taken then there is possibility to get more than 67% of better fittings.

ACKNOWLEDGEMENTS

The PhD scholar first and foremost thank the almighty God for his abundant blessings and protection during my studies and research work. I am highly indebted to Dr. P. Manimegalai**,** my research guide and Professor, Dept. of Electronics and Communication Engineering, Karpagam University, Coimbatore for her involvement, encouragement, suggestions and constructive criticism to make the research a grand success. I also thank the members of my research committee Dr. M. Hemalatha, Head of the Department of Software systems, Karpagam University Coimbatore. Finally, I thank Dr. K.Baskaran, Professor of Computer Science and Engineering, Government College of Technology, Coimbatore for his advises and suggest-ions to carry out the task more efficiently.

6. REFERENCES

1. Govt. of Tamil Nadu, Department of Economics and Statistics, Chennai. Season and Crop Report, 2005-06 to 2009-10 (Fasli 1419), Report Number R.No. 6 / (2011).
2. Caselli, M., Trizio, L., de Gennaro, G. and Ielpo, P., A simple feedforward neural network for the PM10 forecasting: comparison with a radial basis function network and a multivariate linear regression model, Water Air Soil Pollution 201: 365–377 (2009).
3. Smith, A.E. and Mason, K., Cost estimation predictive modelling: Regression versus neural network, The Engineering Economist 42(2): 137–161 (1997).
4. Warner, B. and Misra, M., Understanding neural networks as statistical tools, The American Statistician 50(4): 284–293 (1996).
5. Setyawati, B.R., Sahirman, S. and Creese, R.C., Neural networks for cost estimation, AACE International Transactions EST 13: 131–138 (2002).
6. Buntine, W.L. and Weigend, A.S., Bayesian Backpropagation, Complex Systems 5: 603–643 (1991).
7. Ripley, B.D. Neural networks and related methods of classification. Journal of the Royal Statistical Society 56(3): 409–456 (1994).
8. Sarle, W.S., Neural networks and statistical models, In: Proceedings of the Nineteenth Annual SAS

Users Group International Conference Pp 1538–1550 (1994)

9. Werbos, P.J., Beyond regression: New tools for prediction and analysis in the behavioral sciences, Ph.D. Thesis, Harvard University, Cambridge, MA (1991).
10. Cheng, B. and Titterington, D.M., Neural Networks: A Review from Statistical Perspective. Statistical Science 9(1): 2-54 (1994).
11. Schumacher, M., Robner, R. and Vach,W., Neural networks and logistic regression: Part I, Computational Statistics and Data Analysis 21: 661–682 (1999).
12. Arun Balaji, S., Baskaran, K., Design and development of Artificial Neural Networking (ANN) system using sigmoid activation function to predict annual rice production in Tamilnadu. International Journal of Computer Science, Engineering and Information Technology (IJCSEIT) 3(1): (2013).
13. Arun Balaji, S., Baskaran, K., Feed Forward Back Propagation Neural Network coupled with Rice Data Simulator for prediction of rice production in Tamil-nadu. International Journal of Computer Science, Engineering and Information Technology (IJCSEIT) 4(5): (2014).
14. Arun Balaji, S., Baskaran, K., Integration of Feed For-ward Back Propagation Neural Network (FFBPNN) and Curve Expert software to predict rice production in Tamilnadu, India. International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE) 3(10): (2015).
15. Arun Balaji, S., Manimegalai Vairavan, P., Effect of different statistical measures in error reduction in Feed Forward Back Propagation Neural Network (FFBPNN) to predict rice production. International Journal of Applied Engineering Research 10(24): 44570-44578 (2015).
16. Makinde, F.A., Ako, C.T. et al., Prediction of crude oil viscosity using Feed-Forward Back-Propagation Neural Network (FFBPNN). Journal of Petroleum & Coal 54(2): 120-131 (2012).
17. Zabir Haider Khan, Tasnim Sharmin Alin, Md. Akter Hussain, Price Prediction of Share Market using Artificial Neural Network (ANN). International Journal of Computer Applications 22(2): (2011).
18. Geetha G., Selvaraj R Samuel, Prediction of monthly rainfall in Chennai using back propagation neural network model. International Journal of Engineering Science and Technology 3(1): 211-213 (2011).
19. Solaiman Karim, Rainfall-runoff Prediction Based on Artificial Neural Network (A Case Study: Jarahi Watershed), American-Eurasian J. Agric. & Environ. Sci. 5(6): 856-865 (2009).