iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 2 **Issue:** IX **Month of publication:** September 2014

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

Wheat Yield Prediction Using Artificial Neural Network and Crop Prediction Techniques (A Survey)

Mr. Vikas Lamba[1], Dr.V.S.Dhaka[2]
Computer Science and Engineering Department,
Jaipur National University,
Jaipur, Rajasthan, India.

Abstract: - This survey paper represents the forecasting techniques in the field of Agricultural (Wheat crop). Paper shows all the past research development of forecasting in all areas. The major forecasting models are Statistical, Metrological, Simulation, Agronomic, Remote Satellite Sensed, Synthetic and Mathematical in the field of Agricultural Yield. This paper shows compact combination of all these models and shows why Neural Network Model is important from other models for nonlinear data behavior system like wheat crop yield prediction.

Keywords: - Artificial Neural Network (ANN), Neurons, Back Propagation (BP) Algorithm, Transfer Function, Network Performance, Mean Square Error (MSE).

I. INTRODUCTION

The prediction of product yield in every region in order to planning & policy making future for food providing distribution, pricing and also its import and export is so important since product yield is as result of different processes interaction in plant and these processes are influenced by weather factors, and studying their relationship and product yield are necessary to product-climate models extraction. Product climate analyzing are practical instruments to analyze plant reaction analyzing models are practical instruments analyze plant reaction to climate changes [1]. Common statistical processes based on regression relations are used to evaluate coefficients which relate plant reactions to climate indices. Agricultural region can be considered as a collection of individual fields that vary in environmental conditions and management practices. Day to Day increase in the population demands increase in the agricultural production with available resources. Efficient management of available resources with variable weather conditions is essential to increase productivity of agriculture. In addition to this, the focus of agricultural production is changing from quantity towards quality and sustainability. Solution of these new challenges requires consideration of how numerous components interact to effect plant growth. These transitions force farmers and agricultural advisors to deal with increasing bulks of information. They need to analyze vast and sporadically located information resources. As Information Technology has opened up new challenges to automate data and analysis, computer programs that simulate the crop growth or yield of crops under different management regimes help farmers make technical decisions to manage their crops better [2].

Crop modeling can play a significant part in systems approaches by providing a powerful apability for scenario analyses. Crop modeling has developed extensively over the past 30 years and a diverse range of crop models are now available [2].

Agriculture now-a-days has become highly input and cost intensive area without judicious use of fertilizers and plant protection measures, agriculture no longer remains as profitable as before because of uncertainties of weather, production, policies, prices etc that often lead to losses to the farmers [3]. Under the changed scenario today, forecasting of various aspects relating to agriculture are becoming essential. Wheat yield prediction Reliable and timely forecasts provide important and useful input for proper, foresighted and informed planning.

But in-spite of strong need for reliable and timely forecasts, the current status is far from satisfactory. The impact of weather and climate on food production is of vital importance. Weather variables affect the crop differently during different stages of development. Thus there is need to develop reliable and timely statistically sound objective forecasts of crop yield based on weather variables and another factors so that reliable forecasts can be obtained for efficient production [4].

II. NETWORK AND MODEL

2.1 Neural network:

Work on artificial neural networks, commonly referred to as neural networks, has been motivated right from its inception by the recognition that the brain computes in an entirely different way from the conventional digital computer. The struggle to understand the brain owes much to the pioneering work of Ramóny Cajál (1911), who introduced the idea of neurons as structural constituents of the brain. Typically, neurons are five to six orders of magnitude slower than silicon logic gates; events in a silicon chip happen in the nanosecond (10-9 s) range, whereas neural events happen in the millisecond (10-3 s) range. However, the brain makes up for the relatively slow rate of operation of a neuron by having a truly staggering number of neurons (nerve cells) with massive interconnections between them [5].

The brain is a highly complex, nonlinear, and parallel information-processing system. It has the capability of organizing neurons so as to perform certain computations (e.g. pattern recognition, perception, and motor control) many times faster than the fastest digital computer. It resembles the brain in two respects:

- Knowledge is acquired by the network through a learning process.
- Interneuron connection strengths known as synaptic weights are used to store the Knowledge [5].

2.2 Benefits of neural networks:

It is apparent from the above discussion that a neural network derives its computing power through, first, its massively parallel distributed structure and, second, its ability to learn and, therefore, generalize. The use of neural networks offers the following useful properties and capabilities:

1. Nonlinearity: A neuron is basically a nonlinear device. Consequently, a neural network, made up of an interconnection of neurons, is itself nonlinear. Moreover, the nonlinearity is of a special kind in the sense that it is distributed throughout the network.
2. Input-output mapping: A popular paradigm of learning called supervised learning involves the modification of the synaptic weights of a neural network by applying a set of training samples [5].
3. Adaptively: Neural networks have a built-in capability to adapt their synaptic weights to changes in the surrounding environment.
4. Contextual information: Knowledge is represented by the very structure and activation state of a neural network.
5. Fault tolerance: A neural network, implemented in hardware form, has the potential to be inherently fault tolerant in the sense that its performance is degraded gracefully under adverse operating.
6. VLSI implements ability: The massively parallel nature of a neural network makes it potentially fast for the computation of certain tasks.
7. Uniformity of analysis and design: Basically, neural networks enjoy universality as information processors. This feature manifests itself in different ways:

   a) Neurons, in one form or another, represent an ingredient common to all neural networks.

   b) This commonality makes it possible to share theories and learning algorithms in different applications of neural networks.

   c) Modular networks can be built through a seamless integration of modules.
8. Neurobiological analogy: The design of a neural network is motivated by analogy with the brain, which is a living proof that fault-tolerant parallel processing is not only physically possible but also fast and powerful [5].

2.3 Models of a neuron:

A neuron is an information-processing unit that is fundamental to the operation of a neural network [6].
We may identify three basic elements of the neuron model:

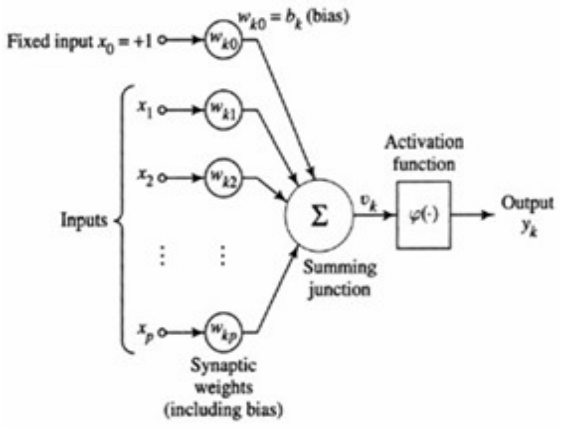


Figure 1-1

nonlinear model of a neuron

- A set of synapses, each of which is characterized by a weight or strength of its own. Specifically, a signal xj at the input of synapse j connected to neuron k is multiplied by the synaptic weight wkj. The weight wkj is positive if the associated synapse is excitatory, it is negative if the synapse is inhibitory.
- An adder for summing the input signals, weighted by the respective synapses of the neuron [7].
- An activation function for limiting the amplitude of the output of a neuron. The activation function is also referred to in the literature as a squashing function in that it squashes (limits) the permissible amplitude range of the output signal to some finite value. Typically, the normalized amplitude range of the output of a neuron is written as the closed unit interval [0, 1] or alternatively [-1, 1].

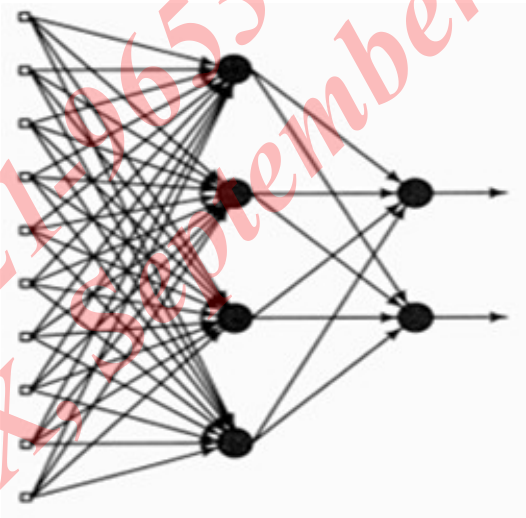
The model of a neuron also includes an externally applied bias (threshold) wk0 = bk that has the effect of lowering or increasing the net input of the activation function [8].

In mathematical terms, we may describe a neuron k by writing the following pair of Equations:

$$v_k = \sum_{j=0}^{p} w_{kj} x_j \quad (1)$$

$$y_k = \phi(v_k) \quad (2)$$

Multilayer feed forward network



Input layer of source nodes | Layer of Hidden Neurons | Layer of Output Neurons

Figure 1-2 fully connected FF network

The source nodes in the input layer of the network supply respective elements of the activation pattern (input vector), which constitute the input signals applied to the neurons (computation nodes) in the second layer (i.e. the first hidden layer). The output signals of the second layer are used as inputs to the third layer, and so on for the rest of the network. Typically, the neurons in each layer of the network have as their inputs the output signals of the preceding layer only. The set of output signals of the neurons in the output layer of the network constitutes the overall response of the network to the activation pattern supplied by the source nodes in the input layer [8].

III. LITERATURE REVIEW

Artificial neural networks, which are nonlinear data-driven approaches as opposed to the above model-based nonlinear methods, are capable of performing nonlinear modeling without a priori knowledge about the relationships between input and output variables. Thus they are a more general and flexible modeling tool for forecasting. The idea of using ANNs for forecasting is not new. The first application dates back to 1964. Hu (1964), in his thesis, uses the Widrow's adaptive linear network to weather forecasting. Due to the lack of a training algorithm for general multi-layer networks at the time, the research was quite limited. It is not until 1986 when the back propagation algorithm was introduced (Rumelhart et al., 1986b) that there had been much development in the use of ANNs for forecasting. Werbos (1974), (1988) first formulates the back propagation and finds that ANNs trained with back propagation outperform the traditional statistical methods such as regression and Box-Jenkins approaches.

Lapedes et al. (1987) conduct a simulated study and conclude that ANNs can be used for modeling and forecasting nonlinear time series. Weigend et al. (1990), (1992); Cottrell et al. (1995) address the issue of network structure for forecasting real-world time series. Tang et al. (1991), Tang and Fishwick (1993), among others, report results of several forecasting comparisons between Box-Jenkins and ANN models.

Christopher Gan et al. 2005, interest in applying artificial neural networks (ANN) to analyze consumer behavior and to model the consumer decision-making process [9]. Heping Pan et al. 2005, presents a computational approach for predicting the Australian stock market index–AORD using multi-layer feed-forward neural networks from the time series data of AORD and various interrelated markets. This effort aims to discover an effective neural network or a set of adaptive neural networks for this prediction purpose [10]. KOŠCAK et al., 2009, have compared common meteorological forecasting method with ANN and he found the performance of ANN with high accuracy. Mahdi Pakdaman Naeini et al. 2010, two kinds of neural networks, a feed forward multilayer Perception (MLP) and an Elman recurrent network, are used to predict a company's stock value based on its stock share value history. The experimental results show that the application of MLP neural network is more promising in predicting stock value changes rather than Elman recurrent network and linear regression method [11].

Nekoukar et al., 2010, have used radial basis function neural network for financial time-series forecasting, and the result of their experiment shows the feasibility and effectiveness. Geetha and Selvaraj, 2011, have predicted Rainfall in Chennai using back propagation neural network model, by their research the mean monthly rainfall is predicted using ANN model [12].

Jyothi Patil et al. 2013, Researchers have attempted to comprehend the pest population dynamics by applying analytical and other techniques on pest surveillance data sets. In this paper, an intelligent system for effectual prediction of pest population dynamics of Thrips Tabaci Linde (Thrips) on cotton (Gossypium Arboreum) crop is presented. The feed forward Multi-Layer Perception (MLP) Neural Network with back propagation training algorithm is employed in the design of the intelligent system. The neural network is trained and tested with the data prepared. The experimental results portray the effectiveness of the proposed system in predicting pest population dynamics of Trips on cotton crop. Moreover, a comparative analysis is performed between the proposed system and two of the existing works [13].

The feed forward Multi-Layer Perception (MLP) Neural Network with back propagation training algorithm is employed in the design of the intelligent system. The neural network is trained and tested with the data prepared. The experimental results portray the effectiveness of the proposed system in predicting pest population dynamics of Trips' on cotton crop. Moreover, a comparative analysis is performed between the proposed system and two of the existing works [13].

Prakash Ramani et al. 2013 proposed a stock price prediction model using multi-layer feed forward Artificial Neural Network (ANN). In this model we have used back propagation algorithm. As the closing price of any stock already covers other attributes of the company, we have used historical stock prices (closing) for training the network [14].

Dase R.K. et al. predicated stock rate because it is a challenging and daunting task to find out which is more effective and accurate method so that a buy or sell signal can be enervated for given stocks. Predicting stock index with traditional time series analysis has proven to be difficult an artificial neural network may be suitable for the task.

Lubomir Macku et al. 2013. Contribution studies prediction of the given semi batch reactor using multilayer feed-forward neural networks. The two prediction approaches are tested – signal prediction approach and system prediction methodology. The most common are multilayer feed-forward neural networks. Fairly wide group of artificial neural networks belongs to recurrent neural networks. Very popular due to their fast training are radial basis function neural networks [15].

Alshayea et al. use generalized regression neural network for prediction of Spanish banks data) develop Neural Network Model for energy consumption and analyze the performance model. Energy consumption prediction which focuses on structures and the parameters used in developing Neural Network models proposed neural network energy prediction model is able to demonstrate an adequate performance with least Root Mean Square Error.

IV. FORECASTING TECHNIQUES

Forecasting techniques in agriculture include, inter alia, forecasting of production/ yield/ area of crops and forewarning of incidence of crop pests and diseases. Crop yield forecasts are extremely useful in formulation of policies regarding stock, distribution and supply of agricultural produce to different areas in the country. However, statistical techniques employed should be able to provide objective crop forecasts with reasonable precisions well in advance before harvests for taking timely decisions. Various approaches have been used for forecasting such agricultural systems. Prominent among the methods of forecasting are based on models that utilize data on crop biometrical characters, weather parameters, farmers' eye estimates, agro meteorological conditions and remotely sensed crop reflectance observations etc., utilized either separately or in an integrated approach [16].

Statistical Forecast Models in Agriculture:

The following statistical forecast models are taken up for discussion for their applications in forecasting agricultural systems. However, the same cannot be claimed to be complete and exhaustive as far as the forecast models and/ or the application areas are concerned. Nevertheless, the essentials are hopefully covered.

4.1 Regression Models:

• Multiple Linear Regressions (MLR) models using plant characters (for forecasting crop yields)

• Weather indices based MLR models (for forecasting crop yields/ crop pest counts)

• Logistic regression models (Models for forecasting/ forewarning qualitative response variable like low or high crop yields, presence or absence of crop pests/ diseases etc.)

4.2 Time Series Models:

• Exponential smoothing models (for forecasting area/ production of crops)

• Auto-Regressive Integrated Moving Average (ARIMA) models (for forecasting area/ production of crops)

4.3 Probabilistic Models:

• Markov chain models (for forecasting crop yields) Forecasting Techniques in Agriculture.

4.1 Regression Models

Let the response variable (variable to be forecasted) be denoted by Y and the set of predictor variables by, X_1, X_2, …, X_p, where p denotes the number of predictor variables. The true relationship between Y and (X_1, X_2, …, X_p) can be approximated by a multiple Linear regression model given by

$$Y= \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon. \quad (3)$$

β_0 and βi (i=1,2,…,p) are parameters to be estimated and ε is random error. Some assumptions are made about this model like the relationship of the response Y to the predictors

$X_1, X_2, \ldots, X_p$ is linear in the regression parameters $\beta_0, \beta_1, \ldots, \beta_p$, the errors are assumed to be independently and identically distributed (iid) normal random variables with mean zero and a common variance σ2, the errors are independent of each other (their pair-wise covariance's are zero) and that the predictor variables $X_1, X_2, \ldots, X_p$ are non-random and measured without error.

4.1.1 Multiple Linear Regression (MLR) Models using Plant Characters:

In the crop yield forecasting context, Y and X_i (i=1,2,…,p) are crop yield and plant biometrical characters respectively. These plant characters may be used in original scale or some suitably transformed variables of these can be used. The plant characters could be average plant height, number of ear heads / panicles etc. depending upon the crop in question. Even though, models using what are called 'growth indices' of plant characters require data at intermediate periods of crop growth, usually these models are developed in a simple manner by utilizing data at one point of time only during the crop growth period (Jain et al. 1985; Aggrawal and Jain, 1996).

4.1.2 Weather Indices based MLR Models:

This is a model similar to the one used in previous subsection. However, instead of using the predictor weather variables as such suitable transformations are made to construct weather indices which in turn are used as predictor variables. The model is given by:

$$y = a_0 + \sum_{i=1}^{p}(a_i * Z_i) + \sum_{i\neq 1}^{p}(a_{ij} * Z_i) + e \quad (4)$$

Where $Z_i = \sum_{w=1}^{m}(r_{iw} * X_{iw})$;

$$Z_{ij} = \sum_{w=1}^{m}(r_{ijw} * X_{iw} * X_{jw});$$

With Y and e denote the crop yield/ pest count (dependent variable) and random error, r_{iw}/r_{ijw} denotes correlation coefficient of Y with ith weather variable/ product of i^{th} and j^{th} weather variables in w^{th} week and m and p denote week of forecast and number of weather variables used respectively. Here Z_i's and Z_{ij}'s are the independent variables which are functions of the basic weather variables like, say, maximum temperature, minimum temperature, rainfall, relative humidity etc. These models utilize data on weather parameters over weeks within years during the crop growth, response variable, if crop yield is available, over years and if pest count is available, over weeks for various years.(Agrawal et al. 1983; Ramasubramanian et al. 2006)

4.1.3 Logistic Regression Models:

When the response variable is quantitative, the usual theory of MLR analysis holds good. However, situations where the response variable is qualitative are also quite common and occur extensively in statistical applications. Logistic regression is widely used when the response variable is qualitative.

Sometimes quantitative information on pests and diseases is not available but is available in qualitative form such as occurrence / non-occurrence, low / high incidence etc. The statistical model preferred for the analysis of such binary (dichotomous) responses is the binary logistic regression model. It can be used to describe the relationship of several independent variables to the binary (say, named 0 & 1) dependent variable. The logistic regression is used for obtaining probabilities of occurrence of the different categories (Misra et al. 2004; Agrawal, et al. 2004) the model is of the form: P (E =1) = 1 / (1+exp (-z)) Where z is a function of weather variables. If P (E =1) ≥ 0.5 then there is more chance of occurrence of disease and if P (E=1) < 0.5 then probability of occurrence of disease is minimum. If the experimenter wants to be more stringent, then the cutoff value of 0.5 could be increased to, say, 0.7.

4.2 Time Series Models

Time series (TS) data refers to observations on a variable that occur in a time sequence. Mostly these observations are collected at equally spaced, discrete time intervals. A basic assumption in any time series analysis/modeling is that some

aspects of the past pattern will continue to remain in the future. Also under this set up, the time series process is based usually on past values of the main variable but not on explanatory variables which may affect the variable/ system. So the system acts as a black box and we may only be able to know about 'what' will happen rather than 'why' it happens. So if time series models are put to use for forecasting purposes, then they are especially applicable in the 'short term'. Ideally, at least 50 observations are necessary for TS forecasting, however practically some 20 observations are needed [16].

4.2.1 Exponential Smoothing Models

An important step in analyzing TS data is to consider the types of data patterns, so that the models most appropriate to those patterns can be utilized (Makridakis et al. 1998). Four types of time series components can be distinguished. They are:

(i) Horizontal − when data values fluctuate around a constant value.

(ii) Trend − when there is long term increase or decrease in the data.

(iii) Seasonal − when series is influenced by seasonal factor/ recurs on regular periods.

(iv) Cyclical − when the data exhibit rises and falls that are not of a fixed period.

Note that many data series include combinations of the preceding patterns. After separating out the existing patterns in any time series data, the pattern that remains unidentifiable form the 'random' or 'error' component. Time plot (data plotted over time) and seasonal plot (data plotted against individual seasons in which the data were observed) help in visualizing these patterns while exploring the data. A crude yet practical way of decomposing the original data (ignoring cyclical pattern) is to go for a seasonal decomposition either by assuming an additive or multiplicative model viz.

$$Y_t = T_t + S_t + E_t \text{ or } Y_t = T_t \,.\, S_t \,.\, E_t, \qquad (5)$$

Where Y_t - Original TS data

T_t - Trend component

S_t – Seasonal component

E_t – Error/ Irregular component

If the magnitude of a TS varies with the level of the series then one has to go for a Multiplicative model else an additive model. This decomposition may enable one to study the TS components separately or will allow workers to de-trend or to do seasonal Adjustments if needed for further analysis.

4.2.1.1 Simple Exponential Smoothing (SES) Models:

Let the time series data be denoted by $Y_1, Y_2,\ldots,Y_t$. Suppose we wish to forecast the next value of our time series Y_{t+1} that is yet to be observed with forecast for Y_t denoted by Ft. Then the forecast F_{t+1} is based on weighting the most recent observation Y_t with a weight value α and weighting the most recent forecast Ft with a weight of (1-α) where α is a smoothing constant/ weight between 0 and 1. Thus the forecast for the period t+1 is given by

$$F_{t+1} = F_t + \alpha (Y_t - F_t) \qquad (6)$$

Note that the choice of α has considerable impact on the forecast. A large value of α (say 0.9) gives very little smoothing in the forecast, whereas a small value of α (say 0.1) gives considerable smoothing. Alternatively, one can choose α from a grid of values (say α = 0.1, 0.2,…, 0.9) and choose the value that yields the smallest MSE value. If the above model is expanded recursively then F_{t+1} will come out to be a function of α, past Y_t values and F_1. So, having known values of α and past values of Y_t, our point of concern relates to initializing the value of F_1. Because the weight attached to this user- defined F_1 is minimal, its effect on F_{t+1} is negligible [16].

4.2.1.2. Double Exponential Smoothing (Holt) Models:

This is to allow forecasting data with trends. The forecast for Holt's linear exponential smoothing is found by having two more equations to SES model to deal with – one for level and one for trend. The smoothing parameters (weights) α and β can be chosen from a grid of values (say, each combination of α = 0.1, 0.2,…,0.9 and β = 0.1,0.2,…,0.9) and then select the

combination of α and β which correspond to the lowest MSE. In SPSS, the weights α and β can be chosen from a grid of values (say, each combination of α = 0.1,0.2,…, 0.9 and β = 0.1,0.2,…,0.9) and then select the combination of α and β which correspond to the lowest MSE.

4.2.1.3 Triple Exponential Smoothing (winters) Models:

This method is recommended when seasonality exists in the time series data. This method is based on three smoothing equations – one for the level, one for trend, and one for seasonality. It is similar to Holt's method, with one additional equation to deal with seasonality. In fact there are two different winters' methods depending on whether seasonality is modeled in an additive or multiplicative way [4] [5].

In SPSS, periodicity can be assigned to data by defining "Dates". In SPSS, choose, Data - > Define Dates ->Cases Are: - defines the time interval used to generate dates; First Case Is: - Defines the starting date value, which is assigned to the first case. Sequential values, based on the time interval, are assigned to subsequent cases; Periodicity at higher level:-Indicates the repetitive cyclical variation, such as the number of months in a year or the number of days in a week. The value displayed indicates the maximum value you can enter. A new numeric variable is created for each component that is used to define the date.

4.2.2 ARIMA Models

A TS is said to be stationary if its underlying generating process is based on a constant mean and constant variance with its autocorrelation function (ACF) essentially constant through time (Pankratz, 1983). Thus, if we consider different subsets of a realization (time series 'sample') the different subsets will typically have means, variances and autocorrelation functions that do not differ significantly. A statistical test for stationary is the most widely used Dickey Fuller test. To carry out the test, estimate by OLS the regression model y't = φyt-1+b1y't-2+…+ bp y't-p where y't denotes the differenced series (yt -yt -1). The number of terms in the regression, p, is usually set to be about 3. Then if φ is nearly zero the original series yt needs differencing. And if φ<0 then yt is already stationary. Autocorrelation refers to the way the observations in a time series are related to each other and is measured by the simple correlation between current observation (Yt) and observation from p periods before the current one (Yt−p). It ranges from −1 to +1. The maximum number of useful autocorrelations are roughly n/4 where n is the number of periods upon which information on y_t is available. Partial autocorrelations are used to measure the degree of association between y_t and y_{t-p} when the y-effects at other time lags 1, 2,…,p-1 are removed. Note that usually up to order 2 for p, d, or q are sufficient for developing a good model in practice. Theoretical ACFs and PACFs (Autocorrelations versus lags) are available for the various models chosen. Thus compare the correlograms (plot of sample ACFs versus lags) with these theoretical ACF/PACFs, to find a reasonably good match and tentatively select one or more ARIMA models [4]. The general characteristics of theoretical ACFs and PACFs are as follows:-

Table-1 Models with ACF and PACF

Model	ACF	PACF
AR	Spikes decay towards zero	Spikes cutoff to zero
MA	Spikes cutoff to zero	Spikes decay to zero
ARMA	Spikes decay to zero	Spikes decay to zero

In general, an ARIMA model is characterized by the notation ARIMA (p,d,q) where, p, d and q denote orders of auto-regression, integration (differencing) and moving average respectively. In ARIMA parlance, TS is a linear function of past actual values and random shocks. For instance, given a time series process $\{y_t\}$, a first order auto-regressive process is denoted by ARIMA (1, 0, 0) or simply AR (1) and is given by:

$$y_t = \mu + {}_{\varphi 1} y_{t-1} + \varepsilon_t \qquad (7)$$

And a first order moving average process is denoted by ARIMA (0, 0, 1) or simply MA (1) and is given by

$$y_t = \mu - \theta_1 \varepsilon_{t-1} + \varepsilon_t \qquad (8)$$

Alternatively, the model ultimately derived, may be a mixture of these processes and of higher orders as well. Thus a stationary ARMA (p, q) process is defined by the equation:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ldots + \varphi_p y_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} + \ldots - \theta q \varepsilon_{t-q} + \varepsilon_t \quad (9)$$

Where ε_t's are independently and normally distributed with zero mean and constant variance σ 2 for t = 1, 2,...n. Note here that the values of p and q, in practice lie between 0 and 3. For development of seasonal ARIMA models, identification of relevant models and inclusion of suitable seasonal variables are necessary for seasonal modeling and their applications, say, forecasting production of crops. Seasonal forecasts of production of principal crops are of greater utility for planners, administrators and researchers alike. Agricultural seasons vary significantly among the states of India. For example, Tamil Nadu has unique three-season cropping pattern for Paddy crop whereas two-season paddy rules elsewhere in the country. Thus seasonal forecasts of crop production can also be made using seasonal ARIMA models. The basic stages involved in developing ARIMA models are now discussed [2].

(1) Identification Stage: The foremost step in the process of modeling is to check for the stationary of the series, as the estimation procedures are available only for stationary series. There are two kinds of stationary, viz., stationary in 'mean' and stationary in 'variance'. A cursory look at the graph of the data and structure of autocorrelation and partial correlation coefficients may provide clues for the presence of stationary. Another way of checking for stationary is to fit a first order autoregressive model for the raw data and test whether the coefficient 'φ1' is less than one. If the model is found to be non-stationary, stationary could be achieved mostly by differencing the series. Stationary in variance could be achieved by some modes of transformation, say, and log transformation. This is applicable for both seasonal and non-seasonal stationary. Thus, if 'X t' denotes the original series, the non-seasonal difference of first order is Y t = X t-X t-1 followed by the seasonal differencing (if needed) $Z_t = Y_t - Y_{t-s} = (X_t - X_{t-1}) - (X_{t-s} - X_{t-s-1})$. The next step in the identification process is to find the initial values for the orders of seasonal and non-seasonal parameters, p, q, and P, Q. They could be obtained by looking for significant autocorrelation and partial autocorrelation coefficients. If second order auto correlation coefficient is significant, then an AR (2), or MA (2) or ARMA (2) model could be tried to start with. This is not a hard and fast rule, as sample autocorrelation coefficients are poor estimates of population autocorrelation coefficients. Still they can be used as initial values while the final models are achieved after going through the stages repeatedly.

(2) Estimation Stage: At the identification stage one or more models are tentatively chosen that seem to provide statistically adequate representations of the available data. Then we attempt to obtained precise estimates of parameters of the model by least squares as advocated by Box and Jenkins. Standard computer packages like SAS, SPSS etc. are available for finding the estimates of relevant parameters using iterative procedures.

(3) Diagnostic Stage: Different models can be obtained for various combinations of AR and MA individually and collectively. The best model is obtained with following diagnostics:

(i) Low Akaike Information Criteria (AIC)/ Schwarz-Bayesian Criteria (SBC): AIC is given by AIC = (-2log L+2 m) where m=p+ q+ P+ Q and L is the likelihood function. Since -2 log L is approximately equal to $\{n (1+\log 2\pi) + n \log \sigma^2\}$ where σ^2 is the model MSE, AIC can be written as AIC= $\{n (1+\log 2\pi) + n \log \sigma^2 + 2 m\}$ and because first term in this equation is a constant, it is usually omitted while comparing between models. As an alternative to AIC, sometimes SBC is also used which is given by SBC = $\log \sigma^2 + (m \log n) /n$.

(ii) Non-significance of auto correlations of residuals via Portmanteau tests (Q-tests based on Chi square statistics)-Box-Pierce or Lunge-Box texts: After tentative model has been fitted to the data, it is important to perform diagnostic checks to test the adequacy of the model and, if need be, to suggest potential improvements. One way to accomplish this is through the analysis of residuals. It has been found that it is effective to measure the overall adequacy of the chosen model by examining a quantity Q known as Box-Pierce statistic (a function of autocorrelations of residuals) whose approximate distribution is chi-square and is computed as follows:

$$Q = n \Sigma r^2 (j) \quad (10)$$

Where summation extends from 1 to k with k as the maximum lag considered, n is the number of observations in the series, r(j) is the estimated autocorrelation at lag j; k can be any positive integer and is usually around 20. Q follows Chi-square with (k-m1) degrees of freedom where m1 is the number of parameters estimated in the model. A modified Q statistic is the Lunge-box statistic which is given by:

$$Q = n (n+2) \Sigma r^2 (j) / (n-j) \quad (11)$$

The Q Statistic is compared to critical values from chi-square distribution. If model is correctly specified, residuals should be uncorrelated and Q should be small (the probability value should be large). A significant value indicates that the chosen model does not fit well. All these stages require considerable care and work and they themselves are not exhaustive.

4.3 Probabilistic Models

4.3.1 Markov Chain Models

A system which takes the form of a chain of finite stages with a limited number of possible states (condition classes) within each stage is called a Markov chain if there exist a case of simple dependence that any state of a particular stage depends directly on any of the states of the preceding stage. In the crop yield forecasting context, a first order Markov chain is constructed to provide a forecast distribution of crop yield for the various crop condition classes (states) at selected phonological (or calendar dates) stages of plant life. The only assumption under the first order Markov chain (or simply a Markov chain) set up is that the past (crop) conditions are statistically uninformative for predicting the future (yield forecasts), after the present (crop) conditions are known (Jain and Agrawal, 1992; Ramasubramanian et al. 2004) For simplicity, the Markov chain forecast model is discussed by taking the particular case of sugarcane (a one-year growth period crop) yield forecasting. Let there be six stages including the final harvest stage, the stages having been determined on the basis of calendar dates i.e. first stage is 3 months after planting, second stage is 4 months after planting etc. and sixth stage at harvest. For development of model using the available (say, for one year) data, firstly construct states within stages on the basis of percentiles upon the data points on biometrical characters, say, average plant height, girth of cane, number of plants per plot etc. at each of the first five stages and upon (only) yield at final stage. For brevity, let us consider, only one biometrical character, say, plant height (X_1). At stage−1, let the percentiles, say, quartiles for the data points upon X_{11} (i.e. X1 at stage−1) be a11, a12, a13. Thus four states viz. X_{11}<=a11; a11<=X_{11}<=a12; a12<=X_{11}<=a13; X_{11}>a13 (classes denoted as a1, a2, a3, a4) can be obtained with distributions say, f11, f12, f13, f14. Similarly with quartiles (a21, a22, a23) upon X12 four states within stage−2 as X_{12}<=a21;a21<=X_{12}<=a22;a22<=X_{12}<=a23; X_{12}>a23 (classes denoted as b1, b2, b3, b4) with distributions say, f21, f22, f23, f24 can be obtained. Likewise four states of other stages can be defined noting that the final stage will have ten states because more finer percentiles, say, deciles (say, a61, a62, a63, a64, a65, a66, a67, a68, a69) can be used for forming states on the basis of variable Y (yield) with distributions say, f61, f62, f63, f64, f65, f66, f67, f68, f69 as this is the main character under study.

Now compute the transition counts of data points moving from any state of one stage to any of the states of the succeeding stage. For example, to find transition probability matrix (TPM) from stage−1 to stage−2, to begin with, consider the f11 data points in state-1 (i.e. a1) of stage-1 and count the corresponding data points of X12 that satisfy the state conditions b1, b2, b3, b4 of stage-2. In this way, the data points which are in state-1 of stage-1 are redistributed into the different states of the next stage-2. Next consider the 36 data points in state-2 (i.e. a2) of stage-1 and count the corresponding data points of X_{22} that satisfy the state conditions b1, b2, b3, b4 of stage-2. The same procedure can be followed for finding other transition counts as well to form a transition frequency matrix say ((ff_{ij})) where i, j=1,2,3,4 for stage-1 to stage-2.

Here, it is noted that the data points of any stage tend to remain in the same states on transiting to the next stage but it can also

be noted that they will be redistributed to other (usually adjacent) states of the next stage as well. However such phenomenon may become less apparent when more than one variable/ transformed data are used. The TPM from stage-1 to stage-2 is finally obtained by dividing each row element of this matrix by its corresponding row sum (these row sums are nothing but distributions at various states of stage-1). This TPM say A12 from stage-1 to stage-2 will be a matrix of order 4. Other TPMs can similarly be obtained. Similarly, TPMs A_{23}, A_{34}, A_{45} will be of order four and A_{56} will be of order (4*10). At stage-6, form a (10*1) vector, say, y_m, on the basis of midpoints of class intervals of yield. Now calculate the predicted yield distributions (PYDs) which will be a (4*1) vector at each stage as given here in the table below [16].

Table-2 The PYD with stages

At stage	PYD
1	A_{12} A_{23} A_{34} A_{45} A_{56} Y_m
2	A_{23} A_{34} A_{45} A_{56} Y_m
3	A_{34} A_{45} A_{56} Y_m
4	A_{45} A_{56} Y_m
5	A_{56} Y_m

In the year for which yield forecast has to be found out, collect observations upon a sample of plots, information upon variable X_1 only for the same set up of stages 1 through 5. Classify these observations as per states of stages of the model developed already. This will result in 'weights' that fall in various states of any particular stage of the developed model. The mean yield forecast at any particular stage can be obtained as the weighted average of the PYDs (four in number) at that stage, weights (four in number) being the number of observations of the forecast year in different states (in the particular stage) of the developed model. Thus the developed model can be used in practice for crop yield forecasting [16].

V. CONCLUSION

Some of the statistical forecast modeling techniques in agriculture has been discussed. Many other models like models based on Group Method of Data Handling (GMDH), growth models, models based on soft-computing techniques such as fuzzy regression, Artificial Neural Networks (ANNs), state space modeling and forecasting using remotely sensed data have not, however, been included. State space modeling and forecasting using remotely sensed data uses large amount of data with many costly tools and technology, these are performed at govt. level research institutes. But ANN tool is easily available tool for us and it is very efficient for nonlinear data also. Many researches in forecasting using ANN tool has been performed till now and they were very accurate.

REFERENCES

[1] Fariba, Ghafoor, Ali "Wheat yield prediction through agro meteorological indices for Arbebil District" World academy of Science, Engineering and technology, 2009.

[2] R. Nagarjuna, B.sailaja, S.R. Voleti "Crop modeling with special reference to rice crop" Directorate of research, Hyderabad, India, 2013.

[3] Vision 2030, India agricultural statistics research institute, Delhi, India, 2013.

[4] Ratna, Amrender "Weather based forecasting model for crops yield using neural network approach" Statistics and Applications, Vol.9, PP. 55-69, 2011.

[5]https://uqu.edu.sa/files2/tiny_mce/plugins/filemanager/files/4230066/reports/Neuralnetworks.pdf.

[6] L.N.M. Tawfig, M.N.M. Tawfig "The effect of number of training samples for artificial neural network" IBN AL-Haitham journal for pure and application science, Vol.23, 2010.

[7] M.Ziarati, O.N. Vcan, R. Ziarati "Optimization of economic order quantity using neural network's approach" Economic Science, Vol.5, 2010.

[8] Guruprasad Bhat, Vidyadevi G. Biradar "Artificial neural network based cancer cell classification (ANN-C3)" Computer engineering and intelligent systems, Vol.3, No.2, 2012.

[9] Christopher, Mike, Jing "An empirical analysis of New Zealand bank customers satisfaction" Banks and Bank systems, Vol.6, Issue.3, 2011.

[10] Hepling Pan, Chandima, John "Predicting Australian stock market Index using neural network exploiting dynamical swinge and intermarket influence" Journal of research and practice in information technology, Vol.37, No.1, February 2005.

[11] Mahdi Pakdaman, Hamidreza, Homa Baradaran "Stock market value prediction using neural networks" IEEE, 2010.

[12] Arti R. Naik, S.K. pathan "Weather classification and forecasting using Back Propagation Feed-Forward neural network" International journal of scientific and research publication" Vol.2, Issue.12, December 2012.

[13] Jyoti Patil, V.D. Mytri "Prediction model for population dynamics of cotton pest (Thrips tabaci Linde) using multilayer-perception neural network" International journal of computer applications, Vol.67, No.4, April 2013.

[14] Prakash Ramani, P.D. Murarka "Stock market prediction using artificial neural network" International journal of advanced research in computer science and software engineering" Vol.3, Issue.4, April 2013.

[15] Lubomir Macku, David Samek "Multilayer Feed-Forward Neural Network in prediction and predictive control of Semi-Batch reactor" International journal of mathematics and computers in simulation, Vol.7, Issue.1, 2013.

[16] Ramasubramanian V. "Forecasting techniques in Agriculture" Indian Agricultural Statistics, New Delhi, India, 2010.

[17] Reza Ghodsi, Ruzbeh, "Predicting wheat production in Iran using an Artificial neural networks approach" International journal of Academic research in business and social sciences, Vol.2, No.2, February 2012.

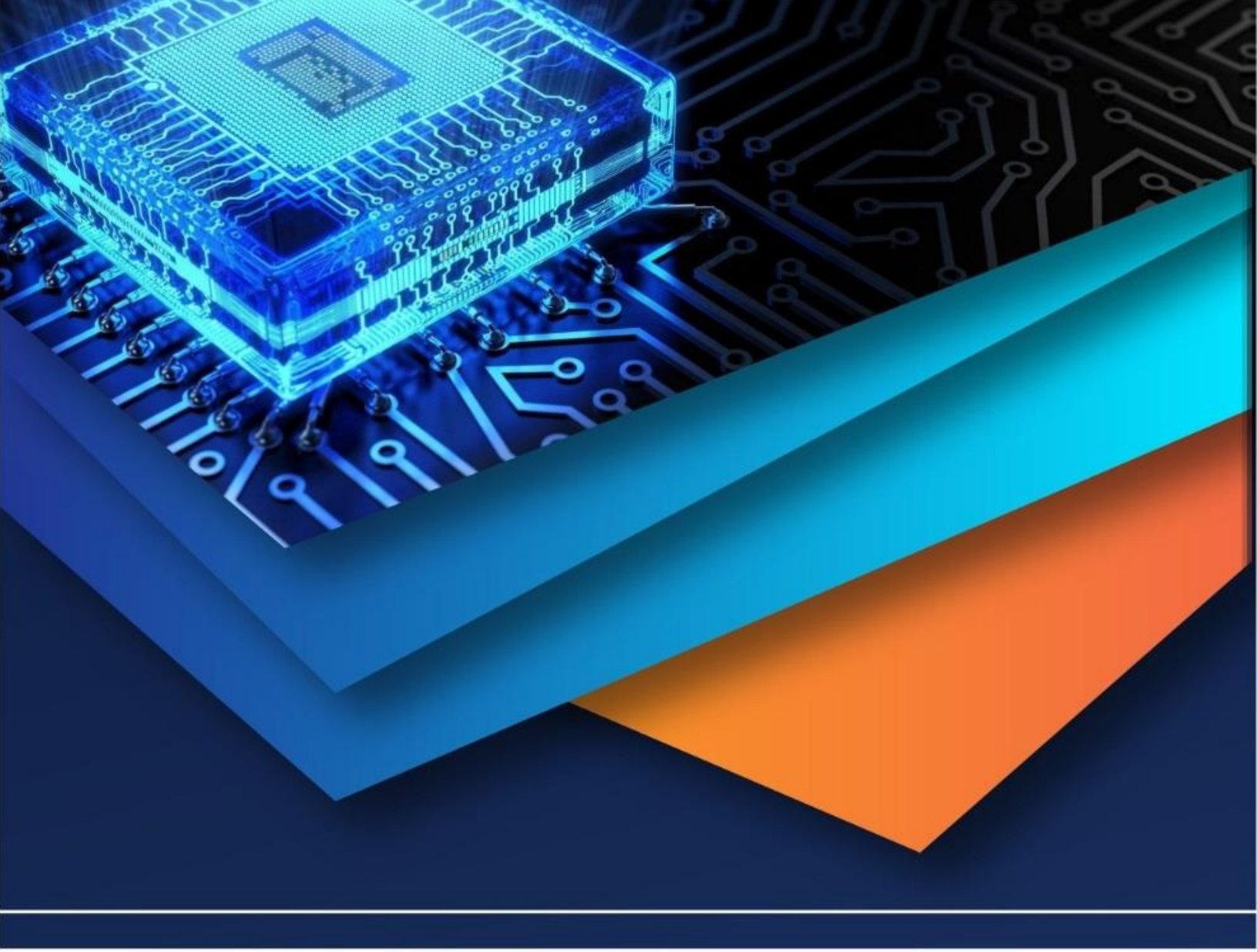
[18] Vida Varahrami "Good Prediction of Gas Price between MLFF and GMDH Neural network" International Journal of Finance and Accounting, Vol.1 2012.

[19] Yue Ma, Yu Chang, Chunyu Xia "Applied Research on Stock Forecasting Model Based on BP neural network" International Conference on Electronic & Mechanical Engineering and Information Technology 2011.

[20]L. Chenl, M. R. Smith, *Y.* Hui "Comparing the Performance of Two Feedforward Neural Network Training Algorithms in MRI Reconstruction" IEEE, 1996.

[21] Mahdi Pakdaman Naeini, Hamidreza Taremian, Homa Baradaran Hashemi "Stock Market Value Prediction Using Neural Networks" IEEE 2010.

[22] Kanna Bhaskar, S. N. Singh "AWNN-Assisted Wind Power Forecasting Using Feed-Forward Neural Network" IEEE Transactions on Sustainable Energy, vol. 3, no. 2, April 2012.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY