About the Authors

Dr. Sagar Maitra, a Doctorate in Agronomy, specialization in crop husbandry engaged in research activities on cropping system, nutrient management of field crops, crop physiology, precision farming and sustainable agriculture. Dr. Maitra is presently working as Associate Professor and Head, Department of Agronomy and Agroforestry at M.S. Swaminathan School of Agriculture, Centurion University of Technology and Management. Besides, he has exposure to print and electronic media. He is a dynamic professional with versatile experience in academics and developmental activities including field exposure, project support and monitoring and transfer of technologies to the farming community. Previously he worked as Senior Producer in a National Electronic Media Network for dissemination of farming technologies for about 17 years and as SMS (Agronomy) in KVK for 3 years. He published more than 70 articles in different national and international journals and authored and edited some books.

Dr. Biswajit Pramanick, Ph.D., B.C.K.V., is presently working as an Assistant Professor, Agronomy at Dr. Rajendra Prasad Central Agricultural University, Pusa, Bihar. Previously he served in the same post at the harbinger of green revolution, the G.B. Pant University of Agriculture & Technology, Pantnagar during 2015-2019. Dr. Pramanick is recipient of prestigious '*Best Teacher Award 2017, Pantnagar*'. He also received Young Scientist Award from '*The Society of Tropical Agriculture, New Delhi*'. He has been serving as '*Academic Editor*' and '*Editorial Board Member*' of two peer reviewed esteemed international journals since 2016. Dr. Pramanick was honoured as '*Co-Chair*' of Youth Empowerment session in the National Yuva Conference - 2017, Pantnagar. He has 14 publications in the journal of national and international repute including 3 books.

NEW DELHI PUBLISHERS

Head Office:
90 Sainik Vihar, Mohan Garden,
Uttam Nagar, New Delhi-110059

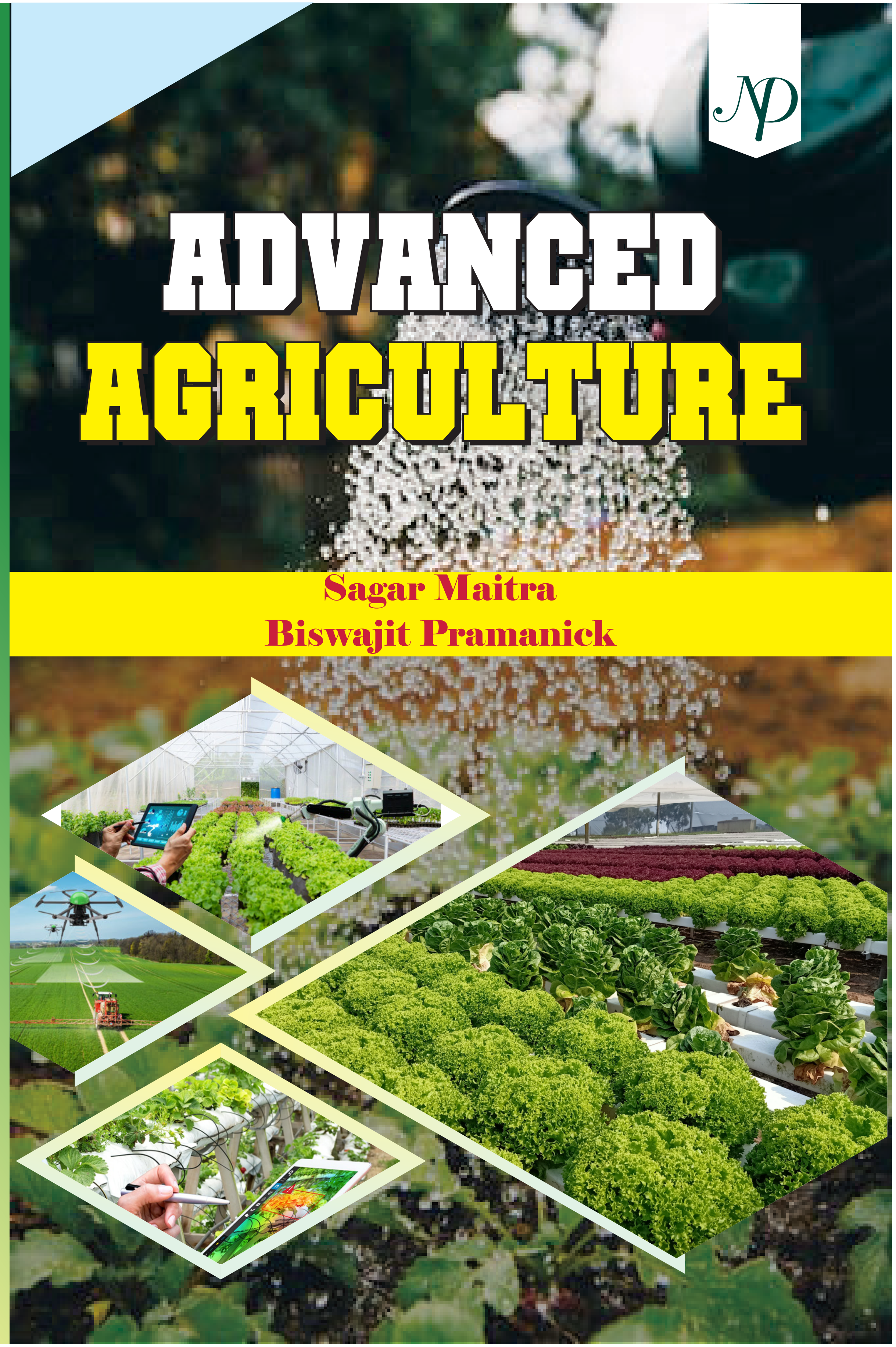
Branch Office:
216, Green Park, Narendrapur,
Kolkata

Corporate Office:
7/28, Room No 208,
Vardaan House, Ansari Road,
Daryaganj, New Delhi-110002

Phone number: 011-23256188, 9953694312, 9582248909
Email: ndpublishers@gmail.com | **Website:** www.ndpublisher.in

ADVANCED AGRICULTURE

Sagar Maitra
Biswajit Pramanick



ADVANCED AGRICULTURE

Sagar Maitra
Biswajit Pramanick

Advanced Agriculture

Sagar Maitra
Biswajit Pramanick

NEW DELHI PUBLISHERS
New Delhi, Kolkata

Advanced Agriculture by Sagar Maitra and Biswajit Pramanick published by New Delhi Publishers, New Delhi.

First Edition: July, 2020

ISBN: 978-93-88879-99-6 (HB)
ISBN: 978-81-947417-6-3 (ebook)

New Delhi Publishers
90, Sainik Vihar, Mohan Garden, New Delhi – 110 059
Tel: 011-23256188, 9971676330
E-mail: ndpublishers@gmail.com
Website: www.ndpublisher.in

Branch Office
216 Flat-GC, Green Park,

Preface

Agriculture,the way of life, has advanced from ploughing soil with a stone fitted on a tree-branch drawn by a horse, to monitoring the entire filed with artificial intelligence and drone technology. Thus, the dimension of it has changed a lot during last couple of decades. Agriculture is not confined managing soils only rather understating what actually the factors going on behind producing the ultimate product, yield. Many new tools like biotechnology, nano-technology, machine learning in agriculture and so on have got their massive importance in the advancement of agriculture. Recently, climate change is one of the burgeoning issues in the agricultural sector. It is a big challenge to the scientists and the policy makers to develop a better resilient agriculture under changing climatic scenario. So, advanced agriculture must consider this aspect as well. Biotic stresses due to weeds, insect and diseases are highly responsible to reduce the crop yield. Better understanding the science behind these stresses and development of their management strategies accordingly is also very much needful. Now a days, horticulture sector has emerged out as an important and highly income generating option to the farmers. This particular sector of agriculture has got its advancement a lot in the recent past creating huge employment generations in the nation. A slogan called 'waste to wealth' is very much trendy now. In the agricultural sector, nothing is waste actually. The so called waste of agriculture mainly consists of crop residues which can be converted into wealth. In the light of advancement one must not overlook the goal of sustainable agricultural development. Any advancement must be nature and biodiversity saving. All these are the milieu of editing a book entitled, 'Advanced Agriculture' which includes the recent advances of agricultural sciences covering the aforementioned topics.

The intended readers of this edited book will mainly consist of agricultural students, researchers and policy makers in the agricultural sector. The book will fulfil the interest of researchers working on various streams of agriculture. On behalf of the editors, we are extremely thankful to the New Delhi Publishers for inviting us to edit such a timely needed book.We would like to convey our appreciation to all contributors including the accepted chapter's authors and many other participants who submitted their chapters that cannot be included in the book due to space limits.

We hope the students, teachers, researchers and extension personnel of the agrarian sector will find this publication informative and useful.

SagarMaitra

BiswajitPramanick

Contents

12

Time Series Modelling and Forecasting in Agriculture: Basic to Advanced In:Advanced Agriculture by *S. Maitra and B. Pramanick* © New Delhi Publishers, New Delhi: 2020, 234-267.
ISBN: 978-93-88879-99-6, DOI: 10.30954/NDP-advagr.2020.12

Time Series Modelling and Forecasting in Agriculture: Basic to Advanced

Amit Saha[1]*, K. N. Singh[2], Ramasubramanan V.[2], Mrinmoy Ray[3], Kanchan Sinha[3], Santosha Rathod[4], D. K. Yadav[5], S. G. Patil[6] and Monica Devi[7]

[1]Scientist, Central Sericultural Research and Training Institute, Mysore, Karnataka, India
[2]Principal Scientist, ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India
[3]Scientist, ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India
[4]Scientist, ICAR-Indian Rice Research Institute, Hyderabad, Telangana, India
[5]Assistant Professor, The National Institute of Health & Family Welfare, New Delhi, India
[6]Assistant Professor, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India
[7]Scientist, CCS Haryana Agricultural University, Hisar, Haryana, India
* Corresponding author: amitsahauni@gmail.com

Abstract: Forecasting in agriculture is a formidable challenge, more so when it needs to be based on objective, systematic methods also having the aspects like timeliness and accuracy. Its need and usefulness for farmers, planners, researchers, agribusiness firms and other stakeholders need hardly be emphasized. Because of the significance of agricultural production in a nation's security, every government utilizes sound methodologies at the institutes established by them and they are not only producers but also main users of agricultural forecasts. The present chapter describes about the various forecasting techniques using time series data in agricultural domain. Along with basics of times series models, a brief overview of some advanced forecasting methodologies are also described. Some possible areas of agriculture are also highlighted in which such forecasting methodologies are useful for effective policy making. To start with, an introduction to the preliminaries of time series and the premise upon which such data rest are discussed. Then the importance of time series in the context of agriculture is outlined by also listing the various related modeling approaches available in the literature. The art of time series model building by considering the very famous and widely used Auto-Regressive Integrated Moving Average (ARIMA) model is then elaborated through its chief stages of identification, estimation, diagnostic checking and forecast performance measures. This is followed by explaining about how indicator variables are used for incorporating the various types like step, pulse and ramp interventions leading to ARIMA

Intervention modelling. When a very few observations are available, how such a data can be fitted using a difference equations based approach called Grey modeling has also been outlined. Moving away from such model based approaches, the data driven technique of Artificial Neural Networks is dealt with the help of feed-forward and recurrent network types also mentioning about the role of activation functions in capturing the non-linearity that exist in the large scale complex data. The treatment of how fuzziness that exists in data is tackled and accounted has also been discussed by giving the steps involved in using Fuzzy time series modeling. Another data driven technique, viz., Support Vector Machine (SVM) has been included that also handles time series in an effective way at times surpassing the usual methods. The long memory time series modeling is also covered. In addition, ARFIMA, where the F stands for Fractional degree of differencing is also given and the situations when these are applicable in time series are also given. Lastly, the subject of technology forecasting has been touched upon and how time series modeling has been used in combination with the former has been highlighted.

1. Introduction

A time series (TS) is a collection of observations made sequentially through time or indexed by time (Chatfield, 2004; Hamilton, 1994). Mostly these observations are collected at equally spaced, discrete time intervals. When TS data refers to observations on a single variable that occur in a (historical) time sequence it will result in an univariate time series (Pankratz, 1983). Put in a different way, a discrete time series is defined as an ordered sequence of random numbers with respect to time (Pfaff, 2008). Thus avoiding any detailed statistical notations and mathematical rigour at this juncture, a TS can be considered as a realization of an unknown stochastic process (which in turn is defined as a sequence of random variables indexed by time) called the data generating process that forms the underlying structure. Often the value of one variable (which constitutes TS) is not only related to its past values in time but, in addition, it depends on past values of other related variables having values at corresponding time periods (Lutkepohl, 2005). A set of such multiple (more than one related variables) series indexed by time is called a multivariate TS. More specifically, if there are k (>1) such variables, then at every time period, say t, there will be a (kx1) vector of main (endogenous) variables. Here, the discussion is restricted to univariate time series data.

A basic assumption in any TS analysis/ modeling is that some aspects of the past pattern will continue to remain in the future (axiom of continuity). Also under this set up, in preliminary TS modeling which were in vogue during the initial years, the TS process is considered to be based on past values of the main variable but not on explanatory variables/ factors which may affect the variable/ system. So the system acts as a black box and if such TS models are put to use, say, for instance, forecasting purposes, it will be possible to know only about 'what' will happen rather than 'why' it happens. Hence they are especially applicable in the 'short term'. Here it is tacitly assumed that information about the past is available in the form of numerical data. Ideally, at least 50 observations are necessary for performing TS analysis/ modeling, as propounded by Box and Jenkins (Box *et al.*, 2016, to mention their recent book edition) who were pioneers in TS modeling. But practically some 15 to 20 observations will suffice for fitting a TS model by keeping in view the statements made by many statisticians over years including George Box that can be loosely stated as "All models are wrong; but some models are useful"!

Any TS data can be resolved, split or decomposed into discernible components such as trend, periodic (say, seasonal; also note that there can be periodic components yet not dependent on seasonality, for instance based on habits), cyclical and irregular variations. One or two of these components may overshadow the others in some series, not to miss a sudden jump or slump in the magnitude of TS observations in an abrupt fashion or a damped trend instead of a gradual pattern in their values. Many techniques such as time plots, multiple seasonal plots, auto-correlation functions, box plots and scatter plots abound for suggesting relationships with possibly influential factors. For long and erratic series, instead of just time plots, alternatives such as smoothing or averaging methods can be tried. These methods could be moving averages, exponential smoothing methods etc. In fact, if the data contains considerable error, then the first step in the process of trend identification is smoothing.

Over the years, various TS analysis literature abound with many improved models due to extensions, modifications and sometimes radical developments in the conventional TS approaches. Few of these methods are briefly mentioned here about which detailed discussion will be dealt with in the subsequent sections. For brevity, the references of the TS tools and techniques mentioned are not included which are anyway given in the respective sections.

In few situations, certain external events called 'interventions' could affect the TS under study. For this, the usual ARIMA methodology can be improved by employing what are called 'transfer function' models to account for the effects of the intervention event on the series by means of including indicator variables about their presence or absence. Such models are referred to as 'ARIMA Intervention models'. ARIMAX models also exist wherein the explanatory variables (usually referred to as X i.e. exogenous variables) apart from the study variables are also included in the ARIMA models. When the underlying relationships among TS variables are so complex that they cannot be described satisfactorily through a linear TS modelling approach like ARIMA, the previously mentioned non-linear TS models like ARCH/ GARCH are preferred. When the functional form is assumed to be known (parametric approach), for such parametric non-linear TS models, non-linear functions relating mean and separately variance of the observed TS and the errors in the underlying system are considered; for e.g. in the function involving variance, the conditional variance is allowed to change over time as a function of squared past errors leaving the unconditional variance constant.

In specific situations, the analysts have to deal with only a few observations, say, if they pertain to some recently developed technologies, only few such values will only be available. In such circumstances, Grey modeling, which works based on difference equations, can be used which can handle as few as four TS observations to construct the model. On the other hand, the conventional TS models will have limitations of requiring a decent number of observations in order to construct the model. Sometimes, long memory TS modelling is resorted to when there is long-range dependence in the sense that the statistical dependence between observations at consecutive time periods does not decay quickly enough, with the event exhibiting a long memory. ARFIMA modelling wherein fractional degree of differencing is also discussed and the situations in TS where they are used are also given.

Machine learning techniques like Artificial Neural Networks (ANNs) are also employed in combination with TS models. Many types of architectures are available in ANN methodology but chiefly the non-linear component in TS is trained using such data driven approaches without imposing any stringent functional form representing the relationships between values of the study variable(s). Of course, to bring in non-linearity in the system, activation functions like logistic, tanh etc. are often used. In certain studies, Support Vector Machines (SVMs) are also coupled with TS models to improve the latter's performance. SVMs are supervised learning models with associated learning algorithms that analyze data used for classification and an SVM model is a representation of the observations as points in space, mapped so that the observations of the separate categories are divided by a clear gap (by means of hyperplanes) that is as wide as possible. SVMs can efficiently perform a non-linear classification using 'kernels', implicitly mapping their inputs into high-dimensional feature spaces. Moreover, data collection in the real world involves a lot of fuzziness and uncertainty. Conventional set theory is based on a crisp boundary between which elements are members and non-members of a particular set. Thus, the answer of the question of whether an element is a member or not of a particular set is always either a strict yes or no. The actual situation can be anywhere between these indicating degrees of membership. Fuzzy set theory assigns membership degrees to each set element and such possibilities have also been integrated in several studies leading to hybrid modeling of Fuzzy logic and TS approach, namely Fuzzy TS modeling.

Technology Forecasting (TF) is the qualitative and/or quantitative prediction with stated level of confidence of feasible and/or desirable characteristics of performance parameters of future technologies given a specific time frame, also with specified level of support (policy, capital, human resource and infrastructural needs). For TF, both quantitative (including TS data) and qualitative information are utilized. One important hybrid approach which is a combination of TS based intervention like modelling and impact analysis is the Trend Impact Analysis (TIA) has also been discussed in a separate section.

After discussing the importance of time series forecasting in agriculture,all the aforementioned TS models will be discussed in detail in the subsequent sections with their applications in agriculture.

2. Importance of Time Series Forecasting in Agriculture

India essentially lives and breathes in her villages which depend almost entirely on agriculture. Green revolution has transformed Indian agriculture from deficit to surplus production which resulted in a paradigm shift in the country's status from being an importer of foodgrains to an exporter. When the time series (TS) on our country's foodgrain production in the past around half century is observed (the period after the implementation of Green revolution in the late 1960s taken), it has nearly rose from 94.01 million tonnes in 1968-69 to a level of 284.83 million tonnes in 2017-18 (4th Advance Estimates; Agristat-GoI, 2018) marking a three-fold increase. Looking into the TS of area under foodgrains, it has marginally increased from 120.43 million hectares in 1968-69 to 127.56 million hectares in 2017-2018 in the same fifty years period. Note that this is obviously due to urbanization and other related factors with the time plot of area figures exhibiting a damped trend pattern with increase at a decreasing rate in the initial years. The surprising fact about this area under foodgrains of our country is that it hovers around the same recent figure even for the past four decades that can be deemed as stagnation with no scope for increase and pointing towards the fact that perhaps it will start decreasing in near future. The TS on productivity during the same period reveals that it has increased almost three-fold (like the aforementioned production figures) from 781 kg/ha to 2233 kg/ha, due to the best technological developments in agriculture over these years inducing more of vertical growth when horizontal expansion in area became increasingly difficult. However, the burgeoning human population seems to undermine this better pace of agricultural progress. Among the individual crops that constitute the foodgrain basket, TS data on paddy and wheat reveal that they together account for around 58% of the area under foodgrains and 75% of the foodgrain production in 2017-18. An assessment and analysis of TS of acreage/ production under various crops is crucial in deciding about shift in production patterns, impact of government policies, farmers' choice, etc. for which TS modeling is best suited leading to forecasting the future scenario for planning purposes. Availability of reliable forecasts of agricultural production/ acreage especially in the short-term has often proved to be a major limitation to planners and TS modeling came to their rescue. So researchers have developed forecasting systems for these statistics employing sound statistical modeling techniques, including, but not limited to TS analysis, modeling and forecasting.

Agricultural forecasts are extremely useful in formulation of policies regarding stock, distribution and supply of agricultural produce to different parts of the country. They are also useful to farmers to decide in advance their future prospects and course of action. Moreover, new emerging forecasting methodologies in agricultural systems may benefit the researchers in enhancing their technical knowledge. However, these statistical techniques should be able to provide objective, consistent and comprehensible forecasts of crop statistics (e.g. area, production and yield) with reasonable precisions well in advance before the harvest for taking timely decisions by the beneficiaries' viz. planners, farmers, agrobased industries and researchers. Various approaches have been depicted in this chapter along with their application in agriculture and allied sectors. Every approach has its own advantages and limitations. TS models have advantages in certain situations. They can be used more easily for forecasting purposes because historical sequences of observations on study variables are readily available from published secondary sources and are being well documented on a regular manner and even available online. The successive observations of TS are statistically dependent and TS modeling is concerned with techniques for the analysis of such dependencies. There are two main reasons for resorting to TS models. First, the system may not be understood, and even if it is understood it may be extremely difficult to measure the cause and effect relationship. Secondly, many a time, collection of information on causal factors affecting the study variable(s) may be cumbersome / impossible and hence availability of long series TS data on explanatory variables is a problem. In such situations, the TS models are a boon to forecasters. Various organisations/ workers in India and abroad have developed forecast models based on TS data using different methodologies with proven applications in the domain of agriculture. As far as utility of TS modeling in agriculture is concerned, its application for statistical forecasting needs hardly any emphasis.

Before venturing into modeling of TS data, the first and foremost step is to study the TS pattern by means of exploratory analysis of the different TS components like trend, seasonality etc. Trend is a 'long-term' behaviour of a TS process usually in relation to the mean level. Such a phenomenon can be seen in many agricultural TS data like crop production and crop yield wherein it can generally be seen that they have an increasing trend due to improvements in technologies and also the management practices that evolved better over the years. Over a period of time, a TS is very likely to show a tendency to increase or to decrease otherwise termed as an upward or downward trend respectively. TS are sometimes encountered with patterns which appear to have had a trend of one type during one part of the period and a different trend of the same, or a different, type during another part of the period. A classic example in agriculture is the pre-and post- green revolution periods with respect to foodgrain scenario of India. Nevertheless, one should not lose sight of the underlying factors that sometimes may cause the trend to change due to factors like growth in population, price changes etc. The trend of a TS may be studied to eliminate it in order to have insight into other components such as periodic variations in the series. Sometimes the TS data are de-seasonalized for the purpose of making the other movements (particularly trend) more readily discernible. The most frequently studied periodic movement is that of seasonal variation which occurs within a year and repeats year after year in a regular fashion. The statistical problem of TS analysis leading to model building consists of deciding the type of components which will fit the data adequately and which is a logical description of the data and thereafter resorting to forecasting and control applications. Not all historical series show upward trends. Some, like plant disease incidence exhibit a generally downward trend. This particular declining trend is attributable to better and more widely available advisory and extension services or due to good government policies. An economic series like agricultural commodity prices may have a downward trend because there could the demand-supply-price dynamics operating on them or may be a better or cheaper substitute became available. Volatility in TS data relating to agricultural commodity prices warrants application of non-linear TS models to capture such phenomenon effectively.

Decomposition models are among the age old approaches to TS analysis although with a number of theoretical pitfalls. These were followed by crude methods called the moving averages method. As an improvement over this equal weighting method is the exponential smoothing methods which gave more weights to recent data thus yielding an unequal weighing procedure. Exponential smoothing models have been shown to be particular cases of the most popular TS model, viz. Auto-Regressive Integrated Moving Average (ARIMA) model. ARIMA models and their variations, such as seasonal ARIMA models, ARIMA models with intervention (introduction of Bt cotton technology formally in the year 2002 is a case in point which can be taken as an intervention), transfer function models, ARMAX models (i.e. ARIMA with explanatory variables) have also been extensively used for forecasting purposes. If the TS show long term dependence property, then long memory TS modeling are quite helpful. Then came many hybrid methods with TS models developed in combination with machine learning techniques such as Artificial Neural Networks (ANNs), Fuzzy Logic, Genetic algorithms, Support Vector Machines (SVMs) etc.

Time-based pattern of one study variable may affect the trend of other variables in a particular location. Such co-variation of TS (of the variables related) that follow similar time-based patterns can be a source of information that may improve forecast accuracy. Various developments in TS modeling under the multivariate framework offer further scope in model improvements. The availability of powerful computers and a variety of readily available software resulted in an impetus in the development of TS based forecast models. Their applications in the field of agriculture for forecasting crop statistics of interest were of immense importance. Not only the agricultural statistics like area/ production/ yield figures, but also the information of weather (particularly rainfall) is available for our country over the past several years and hence TS modeling and forecasting have been used not only for studying, say, the rainfall pattern separately but also utilizing the TS on such variables as exogenous to the variability of crop production by taking the latter as the main or study (endogenous) variables.

As already mentioned in the previous section, when the functional relationships between the various TS data are not known, the plethora of data mining techniques like ANNs, SVMs etc. are resorted to that capture the complex yet non-linear structure present in the data to yield possibly improved forecasts. For this, development of hybrid methodologies by combining both TS models and these machine learning tools have been attempted by several workers. Such soft computing tools can be employed very effectively in this computer era as speed and ease of analysis is no longer a problem nowadays. Even when only very few data points are available in the TS, then also the novel way of Grey modeling can be applied. Even if the TS data is fuzzy, then also methods have been proposed for handling such data and also have the beauty of giving forecasts that tend to provide results that can be likened to interval forecasts instead of obtaining just point forecasts. Moreover, qualitative information have also been utilized along with the quantitative data of TS for better envisioning the future.

The aim of the review is to provide a summary of the main approaches used by agricultural forecasters from basics to advance.

3. Auto Regressive Integrated Moving Average (ARIMA) Model

In modelling time series data, the most important graphical shape is a *time plot* in which the data are plotted over time. A time plot immediately reveals any trends over time, any regular seasonal behavior and other systematic functions of the data. Further, one of the implicit assumptions of time series forecasting is that future will behave like past. This is achieved via the stationarity situation of a series. Therefore, stationarity is a necessary circumstance in building an appropriate model for forecasting. A series is stated to be weakly stationary if its mean, variance and auto-covariance are constant over time i.e., they are time invariant. In other words, stationarity means that there is no growth or decline in the data. The visual plot of a time series is often sufficient to convince a forecaster that the data are stationary or non-stationary. Such an intuitive sense is the starting point of greater formal test of stationarity. One simple test of stationarity is based on the autocorrelation function (ACF). The 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 (y_t) and observation from *p* periods before the current one (y_{t-p}). It levels from -1 to +1. The plot of autocorrelation against lag is known as correlogram or ACF plot. The ACF plot of a stationary data drop to zero relatively quickly, even as for a non-stationary series they are significantly different from zero for several time lags. The autocorrelation function (ACF) plot of a non-stationary data displays a standard pattern with a slow decrease in the size of autocorrelations. Partial autocorrelations are used to measure the degree of association between y_t and y_{t-p} while the *y*-effects at other time lags *1, 2,…, p-1* are not considered.

There are numerous statistical tests to examine the stationary of a series. These are also known as *unit root tests*. The maximum extensively used statistical test for stationarity is Augmented Dickey-Fuller (ADF) test.

The ADF test equation has the following form:

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \sum_{i=1}^{h} \alpha_i \Delta y_{t-i} + \varepsilon_t \qquad (1)$$

where Δy_t denotes the differenced series i.e., $y_t - y_{t-1}$. β_1 and β_2 are the parameters of regression model and *h* is the lag length. The wide variety of lagged difference terms to include is regularly determined empirically, so that the residuals in (1) is serially uncorrelated. In ADF we take a look at whether δ =0 i.e., we have a unit root, which means the time series under consideration is nonstationary. When the observed time series presents trend and heteroscedasticity, differencing and transformation are often applied to the data to cast off the trend and stabilize variance earlier than a model can be fitted.

3.1 Model Specification

Auto Regressive Integrated Moving Average (ARIMA) model is still the most popular choice among many researchers and/or practioners to forecast future values of a series primarily based on past values having linear pattern present among them. Here "AR" means lags of the differenced series appearing in the forecasting equation; "MA" is the lag of the forecast errors and a time series which requires to be differenced for making it stationary is termed as "integrated." Generally, a non-seasonal ARIMA model is denoted as ARIMA *(p, d, q)*; *p* and *q* are the order of autoregressive and moving average order and *d* is the order of differencing.

3.2 Autoregressive (AR) Model

In an auto regressive model, a value of a time series is regressed on previous value from that same time series. Let, the observations of a process at equally spaced time periods *t, t*-1*, t*-2*,…, t-p* are denoted by y_t, y_{t-1}, y_{t-2}, …, y_{t-p} then y_t can be described by way of the following expression and denoted as AR (*p*) model.

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t \quad (2)$$

3.3 Moving Average (MA) Model

In moving average (MA) model an observation is depend on its lagged values of residual error. MA (*q*) model can be written as

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (3)$$

where *q* is the parameter of how many lagged observations are to be added.

3.4 ARIMA Model

In an ARIMA model, the forecast value of a variable is assumed to be a linear function of several lagged values and random errors. That is, the underlying system that generate the time series has the following form:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (4)$$

where, y_t and ε_t are the actual observation and random error at time period *t*, respectively; *(i= 1, 2, …, p)* and θ_j (*j= 1, 2, …, q*) are model parameters, *p* and *q* are integers and often called as orders of the model. Random errors ε_t are assumed to be independently and identically distributed with a mean zero and a constant variance σ^2 .

Equation (4) includes several important special cases of the ARIMA family of models. If *q*= 0, then it becomes an AR model of order *p*. When *p*= 0, the model reduces to an MA model of order *q*. One central challenge of the ARIMA (*p, d, q*) model building is to determine the perfect model order (*p, d, q*) where *d* is the order of differencing. The important steps to analyze and forecast of a time series are presented in the following manner:

Step 1: Stationarity of the time series

A stationary time series has the property that its statistical properties such as the mean and variance are constant over time and for model building the time series must be stationary. The presence of stationarity in the data can be identified via actually plotting the raw information or via plotting the autocorrelation and partial autocorrelation function. Statistical tests like Dickey-Fuller test, augmented Dickey-Fuller test are also to be had to test the Stationarity of a time series.

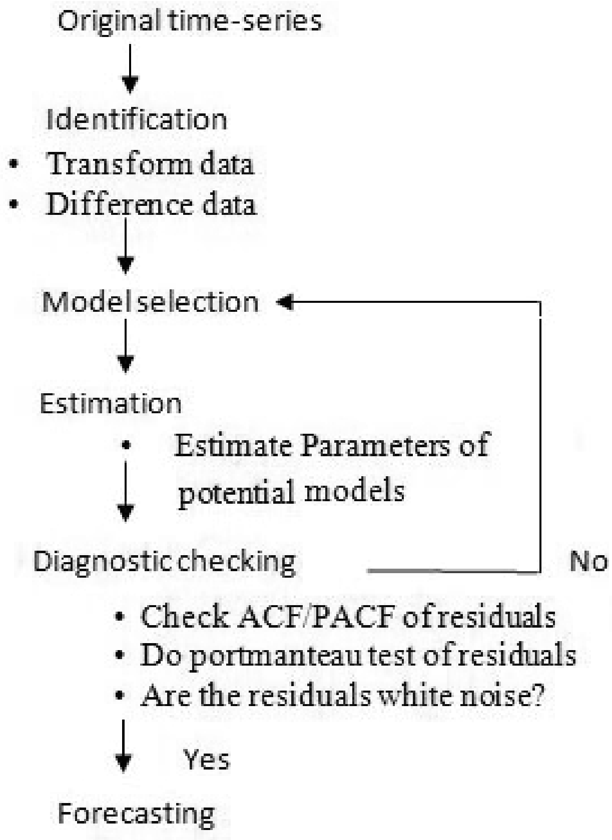


Figure 1: *Box-Jenkins methodology for time series modeling*

Logarithmic, square root or some other modes of transformation are applied to a non-stationary time series to convert a time series stationary in variance. If the series reveals a trend over time or seasonality or some other nonstationary pattern, then the series is differenced repeatedly till the time series becomes stationary.

Step 2: Identification of the model

Candidate ARIMA models are recognized as soon as the time series will become stationary. After obtaining the autocorrelation function (ACF) and partial autocorrelation function (PACF), more than one ARIMA models that closely fit the data can be recognized. In the identification step, the order of tentative models could be received by seeking out significant autocorrelation and partial autocorrelation function.

Step-3: Diagnostic checking

With numerous mixtures of AR and MA model order, different specification of ARIMA model order may be acquired. Some diagnostics measures are used to obtain the best model form. The appropriate ARIMA model is selected using the smallest Akaike Information Criterion (AIC) or Schwarz-Bayesian Criterion (SBC). AIC is given by using the subsequent equation:

$$AIC = (-2\log L + 2m) \tag{5}$$

where, $m= p+q$ *a*nd L is the likelihood function. SBC is likewise used as an alternative to AIC that's written as

$$SBC = \log\sigma^2 + (m\log n)/n \tag{6}$$

If the model is not adequate, a new tentative model has to be diagnosed and the above steps must be repeated. Diagnostic statistics can also help advocate alternative model(s). These steps of model building method are generally repeated numerous times till a satisfactory model is subsequently selected. The final model can then be used for prediction.

The model assumptions about the errors are one of the important steps to consider and are performed by means of employing portmanteau test. The test is utilized to see whether or not the model residuals are white noise. The null hypothesis examined is that the current set of residual is white noise. The Ljung-Box test statistic is utilized to test the residuals that can be expressed as

$$Q = n(n+2)\sum_{k=1}^{h}(n-k)^{-1}r_k^{\,2} \tag{7}$$

wherein, h is the maximum lag, n is the no. of observations, m is the number of parameters in the model. If the data are white noise, the Ljung-Box Q statistic has a chi-square distribution with (h-m) degrees of freedom.

Step-5: Forecast the time series and check accuracy of the model

Finally, the forecasting potential of a method is evaluated based totally on error terms. Root mean square error (RMSE) is one of the most typically used accuracy measures criteria and its scale relies upon on the scale of the data. This is appropriate when comparing unique techniques carried out to the identical set of data, but ought to no longer be used for example, when evaluating across data sets that have specific scales. Root mean square error (RMSE), which measures the overall performance of a model can be expressed as

$$RMSE = \sqrt{\frac{1}{n}\sum_t^n (y_t - \hat{y}_t)^2} \tag{8}$$

where is y_t the actual value for time t, $\hat{y}_t$ is the expected value for time t and n is the number of predictions.

Mean Absolute Percentage Error (MAPE) is some other criteria to check how much a dependent series varies from its model-predicted level. This accuracy measure has the advantage of being scale independent, and so are often used to compare forecast overall performance across different data sets and can be expressed as

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|y_t - \hat{y}_t|}{y_t} \times 100 \qquad (9)$$

4. ARIMA-intervention Model

For time-series data analysis Box-Jenkins Autoregressive integrated moving average (ARIMA) is one the popularly used approach. Three types of processes, viz. Autoregressive (AR) process of order p, Moving average (MA) process of order q and Differencing d times to make the series stationary of degree d are integrated in this model. It is represented as ARIMA (p, d, q) and is employed for analysis of linear time-series data. The principle preferred point of ARIMA is that they provide a general framework for forecasting time-series in which the specification of a model is dependent on data. However in some circumstances, the time series under study could be affected by certain exceptional external events called ‘interventions’. The forecasting performance of ARIMA model may be affected under such circumstances. However, by using suitable approaches it can be improved. To account for the effects of the intervention event on the series but wherein the input series (apart from lags of the main variable and moving average components of errors) will be in the form of a simple indicator variable to indicate the presence or absence of the event ‘transfer function’ models can be employed. In time series literature such models are referred to as ‘Intervention models’. Intervention model had been first applied to study impact of air pollution controls, economic controls on the consumer price index [Box and Tiao (1975)]. An adequate description on intervention modeling can be found in [Box *et al.* (1994)], [Madsen (2008)], [Yaffee and McGee (2000)]etc. Intervention model has been applied in the domain of agriculture for modeling and forecasting cotton yield of India considering the introduction of Bt cotton as unprecedented technology [Ray *et al.* (2014)] and envisioning crop yield scenarios employing time series intervention model based trend impact analysis [Ray *et al.* (2017)]

4.1 Types of Interventions

Three types of intervention are there which is as follows

(i) Step

(ii) Pulse / Point
(iii) Ramp

(i) Step Intervention: At particular point of time it happens and exists in the subsequent time-periods. The effect of step intervention may remain persistent over time or it may enhance or reduce over time.

- Owing to introduction of new variety, pesticide, new economic policy etc. such type of intervention occurs in the domain of agriculture. An instance of this type of intervention is introduction of Bt-cotton in India in 2002.

(ii) Pulse Intervention: Only at particular point of time it happens however the effect of these type of intervention may exists for that particular time period only or it may exists in the subsequent time period.

- These types of intervention is said to occur in specific years with severe drought or flood or severe insect-pest incidence. An instance of this type of intervention is Drought in 2002.

(iii) Ramp Intervention: It happens at particular period of time and exists in the subsequent time-periods with an increasing magnitude. Over time the effect of ramp intervention will always increase.

- The best instance of this type of intervention is the price rise of an agricultural commodity.

An intervention model can be represented as follows:

$$Y_t = \frac{\theta(B)}{\phi(B)}\varepsilon_t + \frac{\omega(B)}{\delta(B)}B^b I_t$$

$$Y_t = \text{ARIMA model} + [\text{Intervention component}]^*I_t \qquad (10)$$

where

y_t = dependent variable

I_t = Indicator variable coded according to the type of intervention (discussed subsequently)

$\delta(B) = 1 + \delta_1 B + ... + \delta_r B^r$ (Slope parameter)

$\omega(B) = \omega_0 + \omega_1 B + ... + \omega_s B^s$ (Impact parameter)

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ (Autoregressive parameter)

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_p B^q$ (Moving average parameter)

ε_t = white noise or error term

b = delay parameter

B = Backshift operator i.e. $B^a Y_t = Y_{t-a}$

4.2 Indicator Variables

Generally, the values an intervention variable can take based on the type of intervention. For step intervention

$$I_t = \begin{cases} 0 & t \neq T' \\ 1 & t \geq T' \end{cases} \qquad \text{(i)}$$

with T' is time of intervention when it first occurred.

For pulse intervention

$$I_t = \begin{cases} 0 & t \neq T' \\ 1 & t = T' \end{cases} \qquad \text{(ii)}$$

For ramp

$$I_t = \begin{cases} 0 & t \neq T' \\ t - T' & t \quad T' \end{cases} \qquad \text{(iii)}$$

The fitting of intervention model consists of the usual three stages i.e. identification, estimation, diagnostic checking alike to ARIMA model. The estimation process and diagnostic checking are similar to ARIMA modeling; however the identification procedure is somewhat distinct. ω, δ and b are the additional three parameters of the intervention model where ω is known as impact parameter which implies magnitude of change (either positive or negative) owing to intervention and δ is known as slope parameter which has divergent meanings in case of different types of intervention. In case of pulse intervention, the effect of the intervention is temporary if δ is near to zero; alternatively the effect of the intervention is permanent if δ is near to one. In case of step intervention, the effect of the intervention remains constant over time if δ is near to zero and the effect of intervention increases over time if δ is near to one. δ has no significant meaning in case of ramp intervention. The delay parameter b generally takes value 0,1 or 2; b=0 implies that the effect of intervention has been observed at the time of intervention itself, b=1 implies that after a delay of one time point the effect of intervention is felt and so on.

Examining the data graphically the order of b can be decided and by comparing estimated impulse response functions with theoretical impulse response function the model form is ascertained. Plotting the residuals which are the absolute difference between the actual values of the post-intervention observations with the forecasted values obtained by fitting ARIMA model on the basis of pre-intervention data the impulse response function can be obtained.

5. Grey Model (GM)

In agricultural systems forecasting is indispensable for the purpose of planning. Various statistical and machine learning approaches are available in the literature for time series analysis, modelling and forecasting in the domain of agriculture. In order to construct the model every time series modelling approach required at least 50 and preferably 100 or more observations. However, sometimes only few time series observations are available owing to rapid innovation in new technologies and policies. For example, Bt-cotton was released in the year 2002-03 hence very few time series are observations are available for modeling Bt-cotton yield. In such circumstances grey model can be deployed. The principle preferred standpoint of this model is that it can be used with as few as four observations in a modelling process. Grey model has been utilized in divergent domain, to cite a few, vehicle fatality risk estimation [Mao and Chirwa (2006)], forecasting of integrated circuit industry output [Wang & Hsu (2007)], time series prediction of time-series [Kayachan *et al.* (2010)], growth trend of internet users, revenue, online game industry [Chang *et al.* (2013)], forecasting hybrid rice yield of India [Ray *et al.* (2017)] etc. GM (1, 1) implies one variable and one order grey model. The general approach, of grey model fitting is as follows:

Step 1: Let the sequence of the original time series is:

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \ldots x^{(0)}(k)\} \tag{11}$$

Step 2: By implementing the first-order accumulated generating operation (1-AGO) on $x^{(0)}$, a new time series can be formed, which is represented as follows:

$$x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \ldots x^{(1)}(k)\}, \tag{12}$$

where $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,\ldots,n.$

Step 3: The background value $z^{(1)}$ built by the method of generations is computed based on average:

$$z^{(1)}(k) = 0.5*[x^{(1)}(k) + x^{(1)}(k-1)], k = 2,3,\ldots,n. \tag{13}$$

Step 4: If $x^{(1)}$ shows an exponential variation, the grey differential equation used in the GM(1,1) model is , and its difference equation is represented as follows:

$$\frac{dx^{(1)}(k)}{dk} + ax^{(1)}(k) = b$$

$$x^{(0)}(k)+\text{a } z^{(1)}(k)=\text{b} \tag{14}$$

where a is the development coefficient, and b is the grey control variable.

Step 5: The OLS method is employed to estimate the parameters a and b

$$\begin{bmatrix} a \\ b \end{bmatrix} = (B^T B)^{-1} B^T Y \tag{15}$$

$$\text{where} \quad Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ . \\ . \\ . \\ x^{(0)}(n) \end{bmatrix}, B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ . & . \\ . & . \\ . & . \\ -z^{(1)}(n) & 1 \end{bmatrix}$$

Step 6: The solution of the grey differential equation under the initial condition $x^{(1)}(1)= x^{(0)}(1)$, yields:

$$\hat{x}^{(1)}(k) = \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-a(k-1)} + \frac{b}{a}, \ k = 2,3,...,n. \tag{16}$$

It is necessary to transfer the data of 1-AGO to actual forecasting value, because the grey forecasting model is formulated using the data of 1-AGO rather than original data. This technique is called the first-order Inverse accumulated generating operation (1-IAGO) and is represented as:

$$\hat{x}^{(1)}(k) = \left(x^{(0)}(1) - \frac{b}{a} \right) e^{-a(k-1)} (1 - e^{a}), \ k = 2,3,...,n. \tag{17}$$

6. Artificial Neural Networks (ANNs)

The terminology of ANNs has evolved from a biological model of the brain. A neural network consists of a set of linked cells (neurons). The neurons obtain impulses from either input cells or other neurons and carry out some sort of transformation of the input and transmit the end results to other neurons or to output cells. The neural networks are built from layers of neurons related in order that one layer receives input from the previous layer of neurons and passes the output on to the subsequent layer. A neuron can be considered as real characteristics of the. The output of a neuron can be obtained as

$$f(x_j) = f(\alpha_j + \textstyle\sum_{i=1}^{k} w_{ij} y_i),$$

wherein f is a function, usually logistic or tangent hyperbolic function and are input vector. The diagrammatic and $y_1, y_2, \ldots, y_k$ nemerical representation of neural network is given in fig. 2 and 3 respectively.

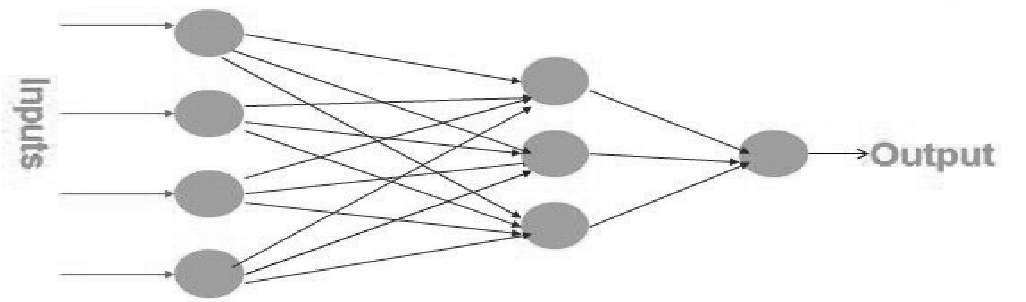


Fig. 2: *Diagrammatic representation of neural network*

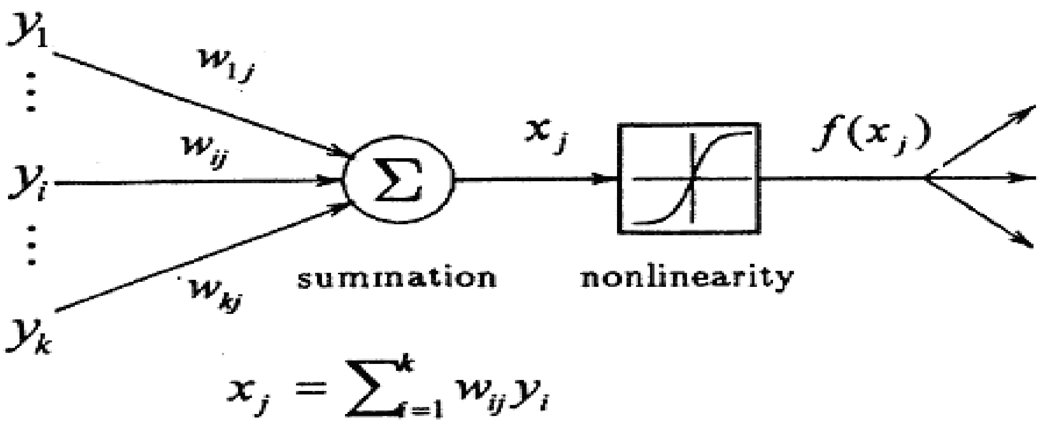


Fig. 3: *The building block of an artificial neuron*

6.1 Neural Networks Architecture

The basic idea of ANN is originated from the anatomy of the cerebral cortex of the brain and is delineated as data processing systems including a big wide variety of simple incredibly inter linked processing elements (artificial neurons). There are several sorts of architecture of ANNs. However, the most broadly used ANNs are discussed below:

6.1.1 Feed Forward Neural Network

In feed forward network, data flows in one route alongside connecting pathways, from the input layer via the hidden layers to the final output layer. There is not any feedback (loops) in this type of neural network architecture. A pictorial representation of feed forward network is given in figure 4.

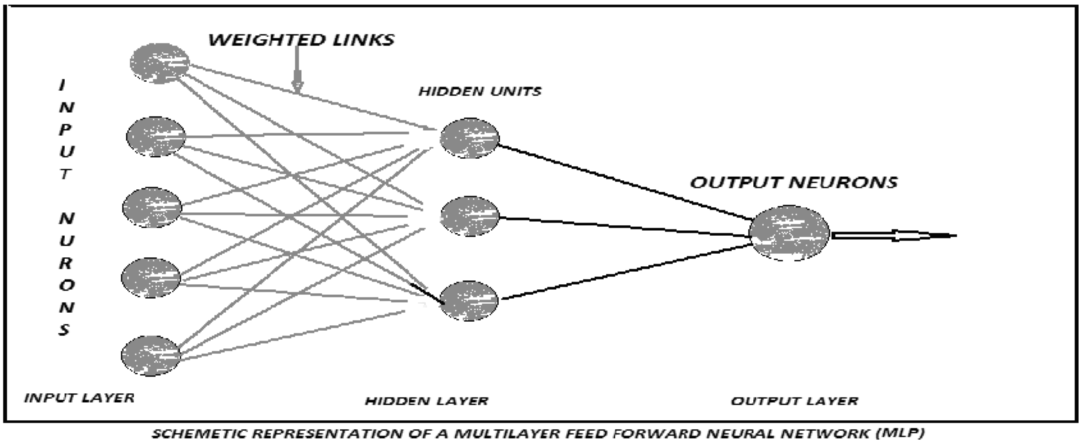


Fig. 4: *Feed forward neural network architecture*

6.1.2 Recurrent Neural Network

These networks fluctuate from feed forward network architectures inside the feel that there is as a minimum one feedback loop. Thus, in those networks, for example, there ought to exist one layer with feedback connections as proven in figure 5. There may also be neurons with self-feedback links, i.e., the output of a neuron is fed back into itself as input. A graphical representation of recurrent neural network is shown in figure 5.

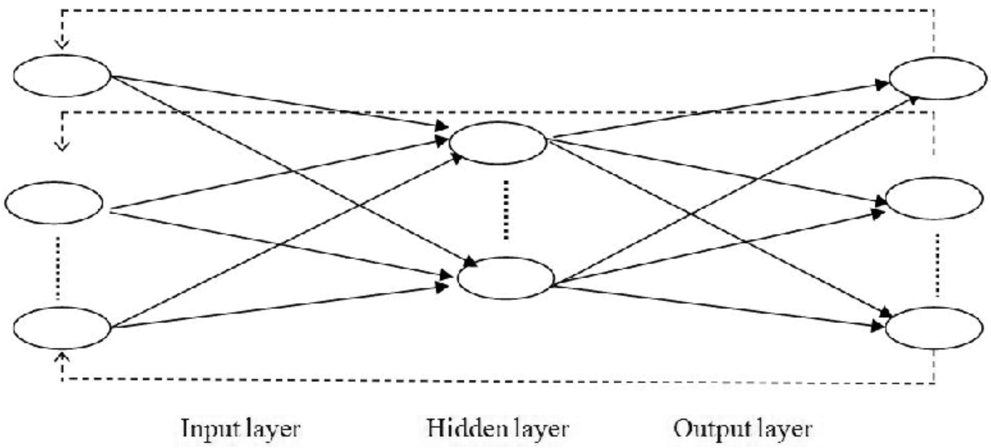


Fig. 5: *Recurrent neural network architecture*

6.2 Development of an ANN model

The overall performance of an ANN model depends on strategic procedure of model development and hence considerable attention is required. One of the critical decisions is to determine an appropriate architecture, that is, the variety of layers, the quantity of nodes in every layer, and the wide variety of arcs which interconnect with the nodes. Other network design choices include selection of activation functions of the hidden and output nodes, the training algorithm, data transformation or normalization methods, training and test sets, and overall performance measures. The most significant steps in developing ANN model are:

6.2.1 Variable selection

Appropriate variable selection techniques as for examples, based on some prior knowledge, effective combination of different input variables, stepwise technique, genetic algorithm, etc., can often apply for selection of significant input variables essential for modelling.

6.2.2 Formation of training and testing set

The data set is divided into two distinct sets called training and testing set. Training set is larger set which use with the aid of ANN to study pattern present inside the data or to adjust the connection weights in a network. The test set is use to assess the performance of the model and also to determine when to stop training i.e., to avoid over fitting.

6.2.3 The network structure

In designing an ANN, one has to decide the ensuing variables:

- The quantity of input nodes.
- The quantity of hidden layers and hidden nodes.
- The variety of output nodes.

Neural networks with larger number of parameters are subjected to poor generalization and over fitting, therefore trial-and-error approach is applied in general for selection of hidden layer nodes. And hence the designing of an ANN is more of an art rather than a science.

6.2.4 The interlinking of the nodes

The network architecture is likewise characterized with the aid of the interlinking of nodes in layers. The connections among nodes in a network fundamentally determine the behavior of the network. For forecasting as well as for different applications, the nodes in all layers are fully interconnected among themselves i.e., from input to hidden and from hidden to output layer.

6.2.5 Activation function

The activation function is also called the transfer function. The degree of nonlinearity in a neural network model is instigated by activation function which is most valuable for ANN applications. Any differentiable function can qualify as an activation function and only a small range of “well behaved” (bounded, monotonically increasing, and differentiable) activation functions are used. The different transfer functions in ANN model building process are:

i. The sigmoid (logistic) function:

$$f(x)=\frac{1}{1+e^{-x}}$$

ii. The hyperbolic tangent (tanh) function;

$$f(x)=\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$$

iii. The sine or cosine function:

$$f(x)=\sin(x) \quad \text{Or} \quad f(x)=\cos(x)$$

iv. The linear function:

$$f(x)=x$$

Among them, logistic transfer function is the most popular preference for all hidden and output nodes. However, for a forecasting issue which involves continuous target values, it's far reasonable to apply a linear activation function for output nodes. It is critical to notice that feed forward neural networks with linear output nodes have the limitation that they cannot model a time series containing a trend. Hence, for this type of neural networks, pre-differencing can be needed to take away the trend effects.

6.3 Performance measure criteria

The most common error function uses in neural networks modelling performance are mean square errors, mean absolute deviations, root mean square error, mean absolute percentage error, etc.

6.4 Applications of neural networks

ANNs have huge applicability to actual world commercial enterprise problems. In fact, they have got already been successfully applied in many domains. Since neural networks are first-class at identifying patterns or trends in data, they're well suited for prediction or forecasting viz., commodity price forecasting, industrial process control, customer research, image processing, risk management, target marketing, etc. The tremendous ability of neural networks to learn by examples makes them highly flexible and powerful tool and thereby exceptionally recommended for real time systems because of their fast response and computational times due to their parallel architecture. Hence, for systematic use of ANNs for different problems; it's miles crucial to understand the potential as well as boundaries of neural networks. With a growing listing of applications, neural networks also make contributions to the areas of studies together with neurology and psychology. They are frequently used to model parts of residing organisms and to analyze the internal mechanisms of the brain. Last but not the least; ANNs have a large capability for prediction as well as for classification when they're combined with computing, artificial intelligence, fuzzy logic and other related techniques.

7. Support Vector Machine (SVM)

SVMs are one of the most used supervised machine learning technique which are the part of artificial intelligence. Machine learning is a technique which allows the machine/computer to learn by itself. Cortex and Vapnik (1995) developed SVM technique for problems of classification. SVM is not only popular for the classification but also for its modelling and prediction performance. SVM is a learning algorithm which is based on kernel. There are different types of kernel which can be used for the classification and prediction purpose. However, there is no such rule to make inference on which kernel should one use all the kernel separately for the given datasets and whichever gives the better result, one should choose that one. SVM is popular in pattern recognition, face and hand writing recognition, datamining, classification of image, categorization of text, chemistry, protein structure prediction, breast cancer diagnosis. SVM has also been applied in rainfall forecasting (Ortiz-Garcia *et al.*, 2014), power load forecasting (Niu *et al.*, 2010) and wind power forecasting (De Giorgi *et al.*, 2014; Mohandes *et al.* 2004).SVMs are now using for prediction of various time series data such as stock market price prediction, weather prediction, agricultural production etc.

Application of SVM in time series is generally utilized when the series shows non stationarity and non-linearity process. In many instances, SVM performed better than the other techniques such as ARIMA, ANN etc. A tremendous advantage of SVM is that it is not model dependent as well as independent of stationarity and linearity. However, it may be computationally expensive during the training. The training of the data driven prediction process SVM is done by a function which is estimated utilizing the observed data. Let, a time series which takes the data at time .

Now, the prediction function for linear regression is defined as:

$$f(y) = (w.y) + c \quad (18)$$

Whereas, for non linear regression, it will be:

$$f(y) = (w.\text{Ø}(y)) + c \quad (19)$$

Where, *w* dentoes the weights,

c represents threshold value,

Ø(y) is known as kernel function.

If the observed data is linear, then equation (18) will be used. But, for non-linear data, the mapping of *y(t)* is done to the hogher dimension "feature" space through some function which is denoted as *Ø(y)* and eventually it is transformed into the linear process. Afer that, a linear regression will carry out in that feature space.

SVM is a learning algorithm which is based on kernel. There are different types of kernel which can be used for the classification and prediction purpose. However, there is no such rule to make inference on which kernel should one use. All the kernels are used separately for the given datasets and whichever gives the better result, one should choose that one. Various types Kernel are listed below:

Polynomial kernel equation: Polynomial kernel is generally used in the image processing. It is useful for nonlinear modelling. This kernel function is very simple yet efficient method.

$k\ (x,y) = (y.x + 1)^p$; p = degree of polynomial (20)

Gaussian kernel function:

$$k(x,y) = exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$$

Or $k(x, y) = exp\ (-\alpha\|x - y\|^2\ )$

Where, shape of hyperplane is controlled by.

Sigmoid kernel function: Sigmoid function is used as the proxy of artificial neural network.

$$k(x,y) = tanh(\theta x^T.y + a) \qquad (22)$$

Linear kernel function: Sometimes, linear kernel gives better results as compared to complex and nonlinear kernels. Linear classifier can be used to test the non-linearity of the datasets.

$$k(x,y) = x.y \qquad (23)$$

8. Fuzzy time series

Time series forecasting is well known method of forecasting in many areas. Time series forecasting is popular because of easiness of evaluation of time series data for getting the forecast values. Another reason of its popularity is that the real world data are mostly time series data. In time series forecasting, time series data are taken as a crisp value. However, data may not be precise and complete in all the cases, e.g. water level data of river, temperature data, agricultural production data etc. Fuzzy techniques are appropriate in those cases when vagueness has been seen in the data. Fuzzy data can be found in artificial intelligence, quality control, biology, psychometry, agriculture, social economy, image recognition etc. Fuzzy time series model can improve the utilization of the data.

Fuzzy time series is the time series with the fuzzy data which is based on fuzzy set theory. Song and Chissom (1993) first proposed fuzzy time series models employing max-min composition operation. They developed a step by step procedure to get the forecast and assessed the proposed model and verified the models' robustness property.

8.1 Definition:

Let, fuzzy sets $f_i(t)$ are defined on $Y(t)$ and $F(t)$ is the collection of $f_i(t)$. Then, $F(t)$ is known as a fuzzy time series on $Y(t)$.

8.2 Fuzzification:

It is the process of conversion from crisp data into fuzzy data.

8.3 Fuzzy logical relationship:

If a fuzzy set A_1 is caused only by A_2, then fuzzy logical relationship is denoted by $A_2 \rightarrow A_1$.

8.4 Fuzzy logical relationship group:

If some fuzzy logical relationship are $A_2 \rightarrow A_1$, $A_2 \rightarrow A_3$, $A_2 \rightarrow A_2$; then fuzzy logical relationship groupis denoted by $A_2 \rightarrow A_1, A_2, A_3$,

8.5 Defuzzification:

Defuzzification is the process of conversion of fuzzified data into the crisp format. Actually, defuzzification is the counterpart of the fuzzification process. In literature, there are many methods for the fuzzification process. Some methods of defuzzification are described in the following:

8.5.1 Centroid Method:

Centroid method is a method of weighted average in which it determines the centroid value of the sets. It is also known as centre of area method. The mathematical formula of this method is-

$$Y = \frac{\sum_{i=1}^{n} \omega_F(x_i) x_i}{\sum_{i=1}^{n} \omega_F(x_i)} \tag{24}$$

Where, Y is the crisp output.

$\omega_F (x_i)$ is the fuzzy output value or value of the membership function of x_i

x_i is the value of the element on the in the fuzzy set F.

8.5.2 Maximum membership method:

This method gives the crisp output value which is equal to the value of associated with the maximum value of membership. It can be expressed as:

$$\omega_F(x_S) \geq \omega_F(x) \text{ for all } x \in F;$$

Where, $x_\$$ is the value associated with the maximum value of membership

8.5.3 Average maximum membership method:

It is like the maximum membership method but the only difference is that it may include more point other than the maximum point. It may include points which belong to some range.

8.6 General Steps Involved in Fuzzy Time Series Forecasting Model

Most of the fuzzy time series forecasting model follow the following steps in forecasting process.

Step 1: Fixing the universe of discourse which is defined as $U = [U_{min}- U_1 , U_{max}- U_2]$, where U_{min} and U_{max} are minimum and maximum value of the data and and are two any two positive values which are selected by the modeler properly. Define the proper universe of discourse to accommodate whole time series data.

Step 2: Division of the universe of discourse or define the intervals.

Step 3: Define fuzzy sets on the universe of discourse.

Step 4: Fuzzify the data which are based on the universe of discourse and corresponding fuzzy set defined in step-2 and step-3.

Step 5: Make the fuzzy logical relationship (FLR).

Step 6: Preparethe fuzzy logical relational groups.

Step 7: Forecast the time series data.

Step 8: Defuzzification of the forecasted fuzzified outputs.

9. Modeling of Long Memory Time Series

Time series modeling is an important area of forecasting in which successive observations recorded over a period of time called as time series, are analyzed to develop a model, which describes the underlying relationship for the study variable. This univariate time series approach is useful when there are no or limited information is available on explanatory variables. Most popular class of model for modeling univariate time series is Box-Jenkins (1970) Autoregressive Integrated Moving Average (ARIMA) model due to its most robust statistical properties and well-known

Box–Jenkins methodology for model building. The ARIMA model is comprises of three steps; identification, estimation and diagnostic checking, once the model is built, forecasting will be done based on the selected model.

The main assumption of the Box and Jenkins methodology, the time series is assumed to be stationary, but in most of the real life practical phenomenon, the stationary assumption does not satisfy. In order to achieve the stationarity, differencing the time series is seems to be good solution, but in many cases researcher face a problem of over-differencing which leads to loss of information and affect the model building. These phenomenon suspects the presence of long memory or long range dependence in time series. In ARIMA model the autocorrelation is expected to decay exponentially as lag increases and in some series the decay occurs very slower hyperbolic rate, such series are called long memory time series. These kind of long memory time series are commonly prevail in stock market prices and in economic time series such as stock price, economic growth rate, inflation rate, oil price and GDP figures, agricultural commodity prices etc, the time series showed a characteristic of "long memory faculty' (Xu 2010., Paul 2014 and Rathod et al, 2017). In the presence of long memory, regular ARIMA model fails to capture the trend and leads poor fitting, a popular class of model for modeling and forecasting long memory time series is Autoregressive Fractionally Integrated Moving Average (ARFIMA) model. ARFIMA is the generalization of ARIMA model allowing the fractional values for the differencing parameter (*d*)

9.1 Long Memory

Long memory time series are the successive observations recorded over a period of time, having long range dependence in the series and have autocorrelation at longer lags. Let us assume that the time series has autocorrelation function and t is the lag number. If satisfies the condition (1) then is called as long memory series;

$$\lim_{T\to\infty} \sum_{t=-\eta}^{n} |\rho_t| \to \infty \qquad (25)$$

In particular, the process is said to be integrated of order *d*, or I(*d*), if ,

$$(1-L)^d X_t = U_t \qquad (26)$$

Where, L is a lag operator, -0.5< d <0.5 and U_t is a stationary process.

Classical R/S method, the modified R/S test (MR/S), KPSS method, logarithmic diagram method (GPH), and Gauss semi-parametric estimation method (GSP) are commonly used methods to detect the presence of long memory in data under consideration (Arathi *et al* 2010).

9.2 R/S Analysis method:

The rescaled range test is most commonly used method in detecting long memory in time series, which is the ratio of the adjusted cumulative mean range to its standard deviation.

$$Q_n = \frac{R(n)}{S(n)} \quad (27)$$

The R/S analysis method is expressed as; $\lim_{n\to\infty} n^{-H} Q_n = C$ is aconstant, and H is Hurst index, and approximation of H can be written as $H = \ln Q_n / \ln n$..

$$(R/S)_n = C.n^H \quad (28)$$

Logarithmic form of eqn. (5), can be obtained as

$$\log(R/S)_n = \log(C) + H\log(n) \quad (29)$$

If H lies between 0.5 to 1 then one can say data under consideration have long memory property in it and autocorrelations take far longer to decay (hyperbolic decay); and if H takes values between 0 to 0.5 then time series depicts long term switching. Value of $d = 0$ indicates short memory corresponding to stationary and invertible ARMA.

9.3. ARFIMA Model:

The most popular statistical model for characterizing long memory series is Autoregressive Fractionally Integrated Moving Average (ARFIMA) introduced by Granger and Joyeux (Granger and Joyeux 1980). ARFIMA model is expressed as follows;

$$\varphi(B)(1-B)^d X_t = \theta(B)e_t \quad (30)$$

Where, X_t represents the time series under consideration, B is the back-shift operator such that $BX_t = X_{t-1}$ and e_t is a white noise process with $E(ei)=0$ and variance σ_e^2 and -0.5<d<0.5.

As like ARIMA model building, ARIMA also involves similar method of modeling steps, namely, identification, estimation and diagnostic checking. Geweke and Porter-Hudak (GPH) method is one of the most commonly used method of parameter estimation in ARFIMA modeling.

10. Technology Forecasting

Technology Forecasting (TF) is nothing but prediction of the future characteristics of useful machines, procedures or techniques (Martino, 1983). TF is distinct from the usual forecasting where technology plays a role but not the central issue (Roper *et* *al.*, 2011). TF can aid in understanding the underlying trends in the key factors of technologies aiding agricultural sector so that they can be influenced to achieve the required needs.

TF indeed leads to better decision making even if it is not going to give accurate forecasts. If one questions as to 'why forecast technology' then Martino (1983) states that it may lead to many false implications such as 'No Forecast' which has an underlying assumption that an unchanging technology is in place and that leads to facing the future blindfolded; 'Anything can happen' which means there is no point in attempting to forecast and thus assumes that nothing can be done to influence the future in the desired direction; 'The Glorious Past' which means that one can just rest on past laurels and achievements which is definitely on a future road to disaster as it falsely assumes that glorious past guarantees glorious future; 'Window-Blind Forecasting' which involves an attitude that technology has a fixed track like 'increasing', 'upward' etc. but may not be true as the future may face a setback or may start moving sideways if another technology supersedes it; 'Crisis Action' which works on the assumption that there will be time to respond once an adverse situation arises which points that the future may be doomed to failure more often than not; 'Genius Forecasting' which again is not preferable as it is impossible to teach or repeat to get the same results, expensive to learn, and allows no opportunity for review by others. All such alternatives thus foolishly point towards the inactions like nothing is going to happen or, something may happen but we should do nothing about it or, we should do something about it but nothing can be done or finally, something we could have done but it is too late! The very purpose of documenting these many undesirable options is to emphasize that there is really no alternative to TF.

Of late, TF is considered as a part of the whole gamut of what is called, Technology Foresight which is more than just forecasting technological trends and needs. There is a whole branch of Futures Studies consisting of the following stages (Hines and Bishop, 2013): Framing - Scoping the domain of study; Scanning - Gathering relevant information; Forecasting- Describing most likely and alternative futures; Visioning - Choosing a preferred future; Planning - Organizing to achieve the vision; and lastly, Acting - Implementing the plan. Such a foresight exercise basically attempts to systematically to look into the future in the long terms with respect to science, technology, economy, environment and society. The two chief objectives of foresight is to identify the areas of strategic research and to list the emerging generic technologies that are likely to yield the greatest economic and social benefits (Georghiou *et al.*, 2008). In addition, the factors influenced by Social, Technological, Economic, Environmental, and Political (STEEP) dimensions should also be considered in foresight exercises; adding Information and Legal, it becomes the mnemonic EPISTLE with again the

Social domain further divided into the Demographic and Cultural domains. TF activity can thus help envision and act upon alternative possibilities of possible through plausible to probable ending with preferable futures.

TF methodologies range from intuitive (e.g. Delphi, Brainstorming) to statistical (e.g. trend extrapolation, growth models) to normative methods (e.g. relevance trees) and also include the monitoring methods like scientometrics. The TF methods can be broadly classified as exploratory and normative forecasting methods. Exploratory technological forecasting has its basis on today's assured basis of knowledge and is projected towards the future, while normative technological forecasting first starts with assessing future goals, needs, desires, mission, etc., and then works in the opposite direction to reach the present. Explorative forecasting is more focused on predicting how a new technology will evolve on a pre-determined curve, which is elongated S-shaped, while the normative forecasting attempts to be more pro-active. In a sense, exploratory forecasting extends the present into the future, whereas normative forecasting recedes from a long-term into an intermediate-term future.

An exploratory forecasting exercise involves forecasting of technologies likely to be available at different times in the future in view of the present trends and momentum. This is based on the fact that there is a definite relationship between the past, present and future and that these days, breakthroughs are more frequently engineered by deliberate and sustained application of finance and manpower, occurring less frequently by accident or through strokes of genius. They also represent examples of successful self-fulfilling forecasts, that is, those which cause certain desirable events to occur mainly because they were forecast. On the other hand, a normative forecasting exercise involves determination of needs for future technologies in order to satisfy the projected needs. It is mission-oriented planning. The needs and goals of a nation, industry, or corporation are identified and the technologies required to achieve these goals are deduced.

The TF methods (both exploratory and normative) are of many types and they can further be classified. The exploratory methods can be subdivided into intuitive methods, extrapolation methods, substitution/ diffusion modelling and technology monitoring methods. The intuitive methods include tools and techniques such as individual forecasting, opinion polls, panel discussion, brainstorming, scenario development, Delphi, cross impact analysis, questionnaire approach and systems dynamics. The extrapolation methods include linear/ exponential/ trend extrapolation and trend correlation. The substitution/ diffusion modelling include Pearl/ Fisher Pry, Gompertz, Lotka-Volterra Competition (Prey-Predator) models and Bass diffusion models. The technology monitoring methods include bibliometrics/ scientometrics and patent analysis methods. The other broad type i.e. normative methods include relevance trees, morphological analysis and mission flow diagrams. It is mentioned here that the aforesaid classification is not exhaustive as some methods such as Analytic Hierarchy Process (AHP), Interpretative Structural Modeling (ISM), Goal Objective Strategy (GOS) technique, Simulation modelling, Technology Roadmapping, Gap analysis, Contextual mapping, etc. which have been used in the literature for TF purposes have not been included in the diagram but are no less important TF techniques.

TF per se is not a time series (TS) based modeling approach. Of course many TF tools like substitution modeling utilize data which are TS based, but these models are called growth curve type of models rather than TS models. Some applications of such modeling approaches have been made in agriculture, to cite one, Ramasubramanian *et al.* (2017), in fisheries, an allied sector of agriculture. However, in one of the TF tools viz., Trend Impact Analysis (TIA) utility of TS modeling can be found. In TIA, in the first step, a baseline model (it could be a TS model as well) based forecast is generated using past data. TIA and also use of TS modeling in TIA are discussed in detail subsequently. In the second step, as qualitative information, a set of future events and their impacts are identified utilizing prior knowledge.

Proceeding further on the steps for implementing TIA after the second step mentioned above, the aforesaid impacts and events are considered along with the baseline (thus TIA is a combination of both qualitative and quantitative approaches) to generate various possible future scenarios via simulation. Then the TIA algorithm involves simulation for combining the impact cum event chance judgments with the results of the baseline model based forecasts to generate a set of alternatives of possible future scenarios. From these very large numbers of alternative future, the median, 5th and 95th percentile scenarios are picked up to act as just three distinct yet useful scenarios.

An improved TIA approach considering low, medium and high degrees of severity was proposedby Agami *et al.* (2008). To overcome some of the limitations of their method, Ray *et al.* (2017a, 2017b) have proposed an enhanced method using TS intervention based Trend Impact Analysis (TIA) for TF. As crop yield in any future year shall depend on many forthcoming technologies and also on certain rare events which can be put considered as unprecedented, to exploit such dependencies such a method was proposed.. Three pre-determined events/ impacts of technologies were considered, which may occur in the long term future. This new TIA method combines the traditional TIA with a TS based intervention model. The steps of this enhanced are as discussed subsequently. Firstly, the base forecasts are computed by using any TS model, say, ARIMA. Secondly, envisioning of all possible alternatives is done using intervention model with certain modifications (for details, refer Ray *et al.* 2017a; 2017b). Thirdly, integration of base forecasts with the results obtained from those of the modified intervention model is done. The study revealed the advantages of improved TIA methodology over the traditional TIA approach by providing alternatives for a wide range of forecast paths from which desirable futures can be picked up. To conclude, it can be seen that TF methods can be improved by employing TS modelling appropriately.

11. Conclusions

To sum up, the various time series based forecasting methodologies applicable in the context of agriculture have been discussed. Both statistical model based and machine learning based data driven approaches have been included in the discussion. The combination of two methods that captures the best of both the methods leading to improved forecasting performance have also been highlighted, with one of the methods not coming under the purview of time series modeling if viewed in isolation. While most of the time series techniques have been outlined and also some hybrid approaches, an exhaustive coverage has been done with useful topics from basic to advanced have been truthfully included. It is hoped that the whole chapter will be useful for the targeted stakeholders who may employ these techniques in their respective areas of work leading to productive forecasts.

References

Aarthi, R, S., Muralidharan, D., and Swaminathan, P. 2012. Double Compression of Test Data Using Huffman Code. *J of Theo. and Appl. Info. Tech.*, **39**(2): 104-113.

Agami, N., Omran, A., Saleh, M., and El-Shishiny, H. 2008. An enhanced approach for trend impact analysis. *Tech. Forecasting and Social Change*, **75**:1439-1450.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. 2016. Time series analysis: Forecasting and Control. 5th Ed., Wiley, Hoboken, New Jersey.

Box, G.E.P., and Jenkins, G.M. 1970. Time series analysis: Forecasting and control. San Francisco: Holden-Day.

Box, G.E.P., and Tiao, G.C.1975. Intervention Analysis with Application to Economic and Environment Problems. *J of the American Statistical Association,* **70**:70-79.

Box, G.E.P., Jenkins, G.M., and Reinsel, G.C. 1994. Time Series Analysis: Forecasting and Control. 3rd Ed., Holden-Day, San Francisco.

Chang, T., C. Ku, and H. Fu. 2013. Grey theory analysis of online population and online game industry revenue in Taiwan. *Tech. Forecasting & Social Change*, **80**:175-185.

Chatfield, C. 2004. The analysis of time series: An introduction. 6th Ed., Chapman Hall, Boca Raton, London.

Cheng, B., and Titterington, D. M. 1994. Neural networks: A review from a statistical perspective. *Statistical Science,* **9**:2-54.

Cortes, C., and Vapnik, V. 1995. Support-vector network. *Machine Learning*, **20**:1-25

David, M. 2017. E1071: Misc Functions of the Department of Statistics. *Prob. Theory Group R package version*, **1**:6–8.

De Giorgi, M.G., Campilongo, S., Ficarella, A., and Congedo, P.M. 2014. Comparison between wind power rediction models based on wavelet decomposition with least-squares support vector machine (LS-SVM) and artificial neural network (ANN). *Energies*, **7**:5251-5272.

Georghiou, L., Harper, J.C., Keenan, M., Miles, I., and Popper, R. 2008. The Handbook of Technology Foresight: Concepts and Practice. *PRIME series on Res. and Innovation Policy,* UK.

Gordon, T. J. 1994. Trend impact analysis, Futures Research Methodology, The Millennium Project, American Council for the United Nations University, http://www.foresight.pl/assets/downloads/publications/Gordon1994-Trendimpact.pdf. [Accessed 2020-04-18]

Granger, C.W.J., and Joyeux, R. 1980. An introduction to long-memory time series models and fractionally differencing. *J of Time Series Analysis*, **1**(1):15-29.

Hamilton, J. D. 1994. Time series analysis. Princeton University Press, Princeton, New Jersey.

Hassoun, M. H. 1995. Fundamentals of Artificial Neural Networks. MIT Press, Cambridge.

Hines, A., and Bishop, P. 2013. Thinking about the future: Guidelines for strategic foresight. Hinesight, Social technologies.

Hyndman, R. J., and Khandakar, Y. 2008. Automatic time series forecasting: the forecast package for R. *J of Statistical Software,* **26**(3):1–22.

Kaastra, I., and Boyd, M. 1996. Designing a neural network for forecasting financial and economic time series. *Neurocomputing,* **10**(3):215-236.

Kayachan, E., Ulutas B., and Kaynak, O. 2010. Grey system theory –based models in time series prediction. *Expert Sys. with Applications*, **37**:1784-1789.

Kumar, T.L.M., and Prajneshu. 2015. Development of Hybrid Models for Forecasting Time-Series Data Using Nonlinear SVR Enhanced by PSO. *J of Stat. Theory and Practice*, **9**(4):699-711.

Lutkepohl, H. 2005. New introduction to multiple time series analysis. Springer, Berlin, Germany.

Madsen, H. 2008. Time Series Analysis. Boca Raton, Chapman and Hall.

Maechler, M., Fraley, C., Leisch, F., Reisen, V., Lemonte, A., and Hyndman, R. 2020. fracdiff: Fractionally Differenced ARIMA aka ARFIMA(P,d,q) Models. R package version **1**.5-1. https://cran.r-project.org/web/packages/fr.

Makridakis, S., Wheelwright, S. C., and R. J., Hyndman 1998. Forecasting methods and applications. 3rd Ed.,Wiley*,* New York.

Mao, M., and E. C. Chirwa 2006. Application of grey model GM(1,1) to vehicle fatality risk estimation. *Tech. Forecasting and Social Change*, **73**:588-605.

Pankratz, A. 1983. Forecasting with univariate Box-Jenkins models: Concepts and cases. John Wiley and Sons, New York.

Martino, J.P. 1983. Technological Forecasting for Decision Making. 2nd Ed., North Holland, New York.

Niu, D., Wang, Y., and Wu, D.D. 2010. Power load forecasting using support vector machine and ant colony optimization. *Expert Syst Appl,* **37**:2531–2539.

Ortiz-Garcia, E.G., Salcedo-Sanz, S., and Casanova-Mateom, C. 2014. Accurate precipitation prediction with support vector classifiers: A study including novel predictive variables and observational data. *Atmos. Res*, **139**:128–136.

Patterson, D. 1996. Artificial Neural Networks. Prentice Hall, Singapore.

Paul, R.K. 2014. Forecasting Wholesale Price of Pigeon Pea Using Long Memory Time-Series Models *Agricultural Economics Research Review*, **27** (2):167-176.

Pfaff, B. 2008. Analysis of integrated and cointegrated time series with R. Springer, New York.

Ramasubramanian, V., Ananthan, P.S., Krishnan, M., and Vinay, A. 2017. Technology Forecasting in Fisheries Sector: Cross Impact Analysis and Substitution Modeling. *J of the Ind Soc. of Agri. Statistics*, **71**(3):231–239.

Rathod, S., Singh, K, N., Paul, R.K., Meher, R.K., Mishra, G.C., Gurung, B., Ray, M., and Sinha, K. 2017. An Improved ARFIMA Model using Maximum Overlap Discrete Wavelet Transform (MODWT) and ANN for Forecasting Agricultural Commodity Price. *J of the Ind Soc. of Agri. Statistics*, **71**(2):103–111.

Ray, M., Rai, A., Ramasubramanian, V., Singh, K. N., and Rathod S. 2017a. Envisioning crop yield scenario using time series intervention based trend impact analysis (TIA), ICAR-IASRI News, 22(1), April-June issue, https://iasri.icar.gov.in/wp-content/uploads/2017/12/April-June20171.pdf. [Accessed 2020-04-18]

Ray, M., Rai, A., Singh, K.N., Ramasubramanian, V., and Kumar, A. 2017b. Technology forecasting using time series intervention based trend impact analysis for wheat yield scenario in India. *Tech. Forecasting and Social Change,***118**: 128–133.

Ray, M., Ramasubramanian, V., Kumar, A., and Rai, A. 2014. Application of time series intervention modelling for modelling and forecasting cotton yield. *Statistics and Application,* **12**:61-70.

Roper, A.T., Cunningham, S.W., Porter, A.L., Mason, T.W., Rossini, F.A., and Banks, J. 2011. Forecasting and Management of Technology. 2nd Ed., John Wiley and Sons, New York.

Schalkoff, R. J. 1997. Artificial Neural Networks. The Mc Graw-Hall.

Song, Q., and Chissom, B.S. 1993a. Forecasting enrollments with fuzzy time series-part I. *Fuzzy Sets and Sys*., **54**:1-10.

Song, Q., and Chissom, B.S. 1993b. Fuzzy time series and its models. *Fuzzy Sets and Sys*., **54**:269-277.

Vapnik, V., Golowich, S., and Smola, A. 1997. Support vector method for function approximation, regression estimation, and signal processing, In Mozer, M., Jordan, M and Petsche, T. (Eds). *Adv. in Neural Info. Proc. Sys.*, **9**:281-287, Cambridge, MA, MIT Press.

Wang, C., and L. Hsu. 2007. Forecasting the output of integrated circuit industry using a grey model improved by the Bayesian analysis. *Techno. Forecasting & Social Change*, **74**:843-853.

Warner, B., and Misra, M. 1996. Understanding neural networks as statistical tools. *American Statistician*, **50**:284-93.

Xu, C. U. 2010. The effectiveness of longmemory of Financial Time series R/S test. *China Manage. Sci.*, **11**(18): 204-208.

Yaffee, R., and McGee, M. 2000. An Introduction to Time Series Analysis and Forecasting: With Applications of SAS and SPSS.1st Ed., Elsevier, New York.

Zadeh, L.A. 1965. Fuzzy sets. Inform. and Control, **8**:338-353.

Zhang, G., Patuwo, B. E., and Hu, M. Y. 1998. Forecasting with artificial neural networks: The state of the art. *Int. J of Forecasting,* **14**:35-62.