STATISTICAL MODELLING TECHNIQUE ON EXPORT OF FRUIT CROPS IN INDIA

Thesis submitted to the
University of Agricultural Sciences, Dharwad
in partial fulfillment of the requirements for the
Degree of

MASTER OF SCIENCE (AGRICULTURE)

In

AGRICULTURAL STATISTICS

By

S. HARI PRASAD

DEPARTMENT OF AGRICULTURAL STATISTICS
COLLEGE OF AGRICULTURE, DHARWAD
UNIVERSITY OF AGRICULTURAL SCIENCES,
DHARWAD - 580 005

MAY, 2010

ADVISORY COMMITTEE

DHARWAD **(ASHALATHA K.V.)**
MAY, 2010 **MAJOR ADVISOR**

Approved by :

Chairman : ______________________________
(ASHALATHA K.V.)

Members : 1. ______________________________
(A.R.S. BHATT)

2. ______________________________
(S.B. HOSAMANI)

3. ______________________________
(VILAS S. KULKARNI)

C O N T E N T S

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Statistical modeling or model building is an activity aimed at learning rules and restrictions in a set of observed data proverbially called “Laws” or “the go of it” as stated by Maxwell.

A statistical model is an equation or set of equations which represents the relationships among variables. Modeling may lead to less adhoc experimentation, as models sometimes make it easier to design experiments to answer particular questions or to discriminate between alternative mechanisms. Modeling provides powerful tools for investigating the dependence and nature of relationship among the variables of interest. The relationship among variables must be determined for the purpose of predicting the values of one or more variables on the basis of observation on other variables. Model building or regression techniques are currently applied in many fields i.e., agriculture, Biometrics, Economics, Education, Meteorology, industry, Sociology, Anthropology etc.The primary objective of these policies is not only to regulate the exports and imports but to spur economic growth.

Models for research and models for Management:

Is there a difference between a management model (representing applied knowledge) and a research model (representing basic knowledge)?

Since applied knowledge can be viewed as a subset or derivative of basic knowledge, in terms of the finished product or finished model, the answer is no, there is no absolute difference between a research model and a management model. However, in terms of what modelers do, that is, the modeling process, there is a very definite difference with a research model, especially with a model that is aiming to improve understanding, as much as if not more progress can be made with a model that is wrong, it fails to describe correctly what happens as with a model that more correctly describes behavior. In fact, part of the enjoyment of modeling is that one can speculate inexpensively and try something out is to be successful (which means being used), it must give predictions or information that are clearly better, in some way, than existing practice.

All models, whether built for research or management purpose, are based on a mix of data, knowledge and conjecture. It is permissible, even desirable, for a research model to have a high proportion of conjecture. The management oriented models must not only have a smaller proportion of guess work, but should also be based on data and knowledge that are relatively secure. As with large models versus small models, unless objectives are clearly defined, there is a danger of doing work that is neither of value to the research nor of use as a management tool.

The role of statistics in research is to function as a tool in designing research and drawing valid conclusions. This is essential for a scientific study ensuring that we have all relevant data for making comparisons and further analysis. Some attempts have been made in the past to explore the trends for Export of fruit crops. The techniques used were;

a) Markov chain Analysis
b) Compound Growth Rate
c) Simple Linear Regression models

Not much work has been carried out in statistical field to forecast the export data. Hence this research aim is to develop a suitable statistical model for describing the trend by employing different statistical modeling. Due to this reason the present study was designed to fit different models to fruit export data. At last the best one is recommended for the further study. As the aspect of this following approaches are conducted.

First approach is Polynomial function fitting approach:

In mathematics, a polynomial is an expression of finite length constructed from variables (also known as indeterminates) and constants, using only the operations of addition, subtraction, multiplication and non-negative, whole number exponents.

Polynomials appear in a wide variety of areas of mathematics and science. For example, they are used to form a polynomial equation, which encodes a wide range of problems, from elementary word problems to complicated problems in the sciences; they are used to define polynomial functions, which appear in settings ranging from basic chemistry and physics to economics and social sciences; they are used in calculus and numerical analysis to appropriate other functions. In advanced mathematics, polynomials are used to construct polynomial rings, a central concept in abstract algebra and algebraic geometry.

Second approach is fitting Non-linear Regression models:

Non-linear regression is a powerful statistical analysis programme that performs non linear regression analysis, surface and curve fitting. Non-linear regression determines the value of parameters for an equation, whose form you specify, that causes the equation to best fit a set of data values. Non-linear regression can handle linear, exponential, logistic, periodic and general non-linear functions. Many "non-linear" regression programmes can handle essentially any function whose form you can specify algebraically.

Third approach is time series ARIMA model and Exponential Smoothing :

For analysis of time series data, a model is essential. The Box-Jenkins procedure is concerned with fitting a mixed Auto Regressive Integrated Moving Average (ARIMA) model to a given set of data. The objectives in fitting this ARIMA model are to identify the stochastic process of the time series and predict the future values accurately. These methods have also been useful in many types of situation which involves the building of models for discrete time series and dynamic systems. But this method was not good for lead times for seasonal series with a large random component. (Granger and Newbold, 1970)

Originally ARIMA models have been studied extensively by George Box and Gwilym Jenkins (1968) and their names have frequently been used synonymously with general ARIMA process applied to time series analysis, forecasting and control.

Exponential Smoothing methods forecasts using weighted averages of past data. They are commonly used in inventory control systems where many items to be forecasted and low cost is a primary concern. The simplest exponential smoothing method is single exponential smoothing (SES), suitable for data with no trend or seasonal patterns. For trend data, Holt's method (Double Exponential Smoothing) is suitable and for seasonal data, Holt-Winters method may be used. (*Cooray, 2008)*

Indian Scenario:

India is the largest producer of fruits in the world and is known as "Fruit basket" of the world. The covered area under fruits in India was 5775.30 thousand hectares with the production 63502.7 thousand million tones in 2007-08. The major fruits grown in India are mangoes, grapes, bananas, apples, oranges, sapota etc. This is due to its potential in different Agro climatic zones. India's Export of fruits was Rs.304.53 crores in 2007-08. The major countries which imports are UK, Netherlands, UAE, Russia, Kuwait, USA etc.

India's Potential Untapped:

Despite vast potential and favourable conditions India lags behind several countries in terms of productivity and quality of Horticultural produce. The country's Horticultural produce has not made any major dent in international trade so far, except for mangoes and of late grapes. This is primarily due to the fact that these crops were treated as one of the several means of land use of secondary importance with food grain crops receiving prime attention and consequently heavy investments. Fortunately, the practices pursued during late sixties and in seventies paid rich dividends in terms of achieving manifold increase in food production making the country self reliant. This has led to a beginning of new awakening among planners and policy makers on the need for improving returns from unit area developing strong Export base and providing nutritional security to the masses.

Despite such large production of fruits and vegetables, exports of these products including the processed one from India are insignificant, while India contributes 10% to world fruit production, its share in exports of these products is less than 1% of total domestic production. Abundant investment opportunities are there in expanding the export market. An increasing acceptance of new products with market development efforts has been witnessed lately given the fact that there is good international demand for certain fruits and vegetables products. The export value of the main exporting fruit crop from India is Mango.

Constraints in Export of Fruits:

The export of fruit and vegetables (fresh, chilled and processed) from the country are subject to a number of constraints, many of which are related to the maintenance of human and plant health in the importing country involved. Exporters not only have to meet the requirements of the importers and consumer of the tropical fruit regarding its quality, but also have to meet the phytosanitary requirements of the concerned Government agencies that deal with human and plant health. Further, lack of proper and efficient transport facilities, poor quality of products, absence of modern and scientific storage facilities as well as packing materials are also considered major hindrance for export growth. India horticultural products are also often out priced in the world market.

In fact, in India few players have invented in developing quality standards. Exporters usually target ethnic population abroad or less discerning markets such as Middle-East to ride over quality issues.

In the present study the analysis was carried on Mango exports, Banana exports, annual fruits export and monthly fruits export respectively. Mango and Banana are taken in broad for the study because of their high quantity of export and also for their production statistics. So, based on the availability of data, the polynomial and non-linear regression models are fitted to all the above mentioned export datas. The ARIMA model which needs atleast 50 data points to run this model was of main constraint and therefore monthly Export data was taken to fit the model. Exponential Smoothing methods are also tried to monthly fruit Exports which was also of the same constraint. According to the availability of the data the following objectives are made

OBJECTIVES

1. To determine the export growth using different degrees of polynomial function fitting approach.
2. To examine the growth in export by using different non-linear regression models.
3. To fit ARIMA model and exponential smoothing technique.
4. To compare different models and forecast values for fruit crops export by using the selected model.

2. REVIEW OF LITERATURE

Review of literature helps the researcher in knowing the previous work done in their field of research and thereby helps them in identifying a proper analytical and methodological issue relevant to the study. It enables the researcher to carry out their research work in a proper direction and enables them to draw conclusions. Therefore, the past studies were reviewed and are presented under the following headings.

2.1 Polynomial Function fitting approach.

2.2 Fitting of Non-linear regression models.

2.3 Fitting ARIMA and Exponential Smoothing.

2.4 Comparison of all the above models and selection of the best model.

2.1 Polynomial Function fitting approach.

Jahne (1984) analysed on Growth curve in female cattle. A polynomial approximation was used to study growth up to 21 months of age in 4617 German Black Pied heifers born at 3 farms over a 5-yr period. It was shown that the accuracy of fit of the growth curve did not differ between weekly or monthly recording of body weight. In the case of daily body weight gain, a 4th-degree polynomial had the greatest accuracy. Growth curves are illustrated for heifers with different ages at 1st conception.

Tennent (1986) examined on Intra-annual growth of young Pinus radiata in New Zealand. A study was made of monthly ht. and diam. growth of 5-yr-old P. radiata at 4 forests throughout New Zealand, with the aim of providing a method of distributing growth within years. A conditioned 7 degree polynomial was fitted to the cumulative growth. The differential was used to calculate the proportion of growth in each month. There were significant differences in growth pattern between the 4 forests. Comparison between ht. and diam. growth supported the hypothesis that ht. growth leads diam. growth.

Rao (1992), attempted a study on the relationship between prices of lac and quantity exported. Data are presented and analysed for lac exports from India over the period 1967-68 to 1986-87, in relation to international prices of shellac/seedlac (in Rs/t) over the same period. The highest export figure was 17 712 t in 1968-69, and the highest prices (Rs 93/t) were in 1985-86 when the export figures were 5624 t (nearly at their lowest over the period). A second degree polynomial regression relation was fitted to international lac prices, using polynomial coefficients as independent variables and export as dependent variable. The relation derived fitted the data with reasonable accuracy.

Singh *et al*. (1993) analysed the behaviour of market arrivals and prices of potato in Punjab. They reported a rising trend and the trend co-efficients of potato arrivals were significant. During the post-harvest period, the prices ruled very low whereas during the lean period, the indices of prices ruled very low whereas during the lean period, the indices of prices remained high and ranged widely mainly due to the seasonal and perishable nature of the crop. They further indicated that the potato prices depicted a cycle of three years

Balasubramanian and Rema (1996) reported that the trend in export was found to be directly correlated to imports. A high correlation was found in that for every tonne of kernel exported, 5 tonnes of raw nuts was imported. The dependence of cashew nut export trends on the international market price for kernels and on the foreign exchange rate of the Indian rupee was examined. They further reported that the export was more the international price was less. Indian raw nuts are fairly priced against the international price for kernels. The pricing trends of raw cashew nuts in different states are evaluated and the effect of a monopoly procurement policy in 1981, 1982 and 1988-93 on the price obtained by farmers was also critically discussed

Vaishnav *et al*., (1998), examined Trends of groundnut productivity in long term experiment. Groundnut yield data were collected from a long-term (1979-92) experiment conducted at Gujarat Agricultural University, Junagadh Campus, Junagadh, Gujarat, India, with different rates and frequencies of application of farmyard manure, N, NP and NPK. The use of moving average values reduced the seasonal/irregular fluctuations in the data effectively and gave a better fit compared with that of the original set of data. A third degree polynomial equation expressed the trend for all the treatments, which showed an initial decline in yield, followed by a rapid increase to the maximum, followed by another decline.

Upadhyay *et al*. (2001) examined on relative importance of reproductive traits on lifetime milk. Relationship of reproductive traits (age and weight at first calving and days open percent) with life time milk yield and profit for 1380 Holstein x Sahiwal crossbreed cattle maintained at six military farms located in Northern India was studied by the third degree polynomial regression analysis and R2. Age at first calving, weight at first calving and day open per cent accounted for 4.99, 2.05 and 19.64 percent, respectively, of variation in life time milk yield, and 5.66, 9.87 and 15.92 percent, respectively, variation in lifetime profit suggesting that the days open percent of the cows was more important than age at first calving and weight at first calving for milk yield as well as profit.

Zhang-HuiLin *et al*. (2006) attempted a study on adjusting coefficient for milk yield of dairy cattle. Based on the milk yield records of 178 Chinese dairy cattle from the Fourth Dairy Cattle Farm of Caotan Farm from 1994 to 2005 in Xi'an. The adjusting coefficients of milk yield with calving numbers, years, milking days, calving months were investigated with the simple statistical polynomial regression analysis method. Four best regression equations were established and their significances were measured by difference significance test.

Munir *et al*. (2006) analysed on growth and flowering response of snapdragons after release from apical dominance. Plants of an early flowering Antirrhinum cultivar 'Chimes White' were pinched at 4, 5, 6, and 7 leaf-pair stage to observe the effects on flowering time and plant quality. Though control plants flowered earlier (81 days) than the pinched ones, they produced less number of flower buds. Flower time and rate of progress to flowering in pinched plants increased linearly and significantly. The quality of pinched plants regarding branch numbers, leaf area, plant height, plant fresh weight etc. was significantly improved in all treatments. Many plant growth parameters were successfully fitted by the second degree polynomial model whereas linear model indicated a good fit in reproductive development.

Therthappa (2005) worked on yield estimation of sunflower with yield contributing character who observed that cubic model was found best in predicting seed yield based on the head diameter and disease at first stage.

Konz *et al*., (2007) worked on Implementation of a process-based catchment model in a poorly gauged, highly glacierized Himalayan headwater. . The distributed catchment model TACD, which is widely based on the HBV model, was further developed for the application in highly glacierized catchments on a daily timestep and applied to the Nepalese Himalayan headwater Langtang Khola (360 km2)Low laying reference stations are taken for temperature extrapolation applying a second order polynomial function.

Yeber *et al*. (2007) analysed on Decolorization of kraft bleaching effluent by advanced oxidation processes using copper as electron acceptor. Two advanced oxidation processes were used to remove colour from a Kraft bleaching effluent. The optimal decoloration rate was determined by multivariate analysis, obtaining a mathematical model to evaluate the effect among variables. Concentrations and the reaction times were optimized. The experimental design resulted in a quadratic matrix of 30 experiments. The experimental design methodology indicated that a quadratic polynomial model may be used to represent the efficiency for degradation of the Kraft bleach pulp effluent by a photocatalytic process.

Swetha (2009) worked on the Impact of rain water harvesting on farming economy. Polynomial model of 5P^{thP} degree proved to be the best with highest RP^{2P} (0.631) followed by 4P^{thP} degree polynomial model (RP^{2P}= 0.592), cubic (0.514), quadratic (0.491) and linear (0.482). In case of relative humidity and wind speed none of the models was significant hence none of the models have been selected.

Teresa Jacobson *et al.* (2009) examined using linear regression of seasonal weather patterns to enhance undergraduate learning based on the students' knowledge of polynomials. It was estimated that a model of at least degree seven would be necessary to satisfactorily describe the pattern. The sixth order polynomial was used to compare the erratic pattern of the maximum standard deviations to the corresponding pattern of the best fitting polynomial for the minimum standard deviations. R-square value including the eighth degree term increased by 0.020975 to 0.885433

Wang-YeCheng *et al*. (2009) analysed on optimization of parameters of blackcurrant harvesting mechanism. The operation parameters of a black currant harvester were evaluated using frequency, amplitude and location as input variables, and the harvesting rate and branch vibration force as output parameters. A quadratic orthogonal rotary regressive experimental design was employed to develop the second order polynomial regression models, which explained the relationship between the input and output parameters.

2.2 Fitting of Non-linear regression models.

Chengappa (1981) made a study in growth rates of area, production and yield of coffee in India. Linear model and Exponential model were used and corresponding growth rates were worked out. The exponential model provided a good fit for Arabica and Robusta coffee production.

Umakapila (1982) calculated the growth rate of groundnut for all India and state wise for the period, 1951-52 to 1974-75 by using exponential growth function. She reported that growth rate was negative and growth crate of output was 0.55 percent. The major source of growth in output in the states of Andhra Pradesh and Karnataka was found to be the area under the crop.

Fialar (1985) analysed the production pattern in marketing of coca in Ghana using Exponential model. The study indicated that the rate of growth in average under coca for the world as a whole has decreased. However the total production has increased per annum. The negative growth rate for exports for the world as a whole reflected that the international trade on cocoa was a decline.

Sadasivan (1989) analysed the pattern of pulse production in India from 1955-56 to 1984-85. The analysis was done at the disaggregate levels of seasons, periods, states and crops. He used exponential form to estimate growth rates. It was concluded that stagnancy in pulse production at all India level was the net result of changing situation at the disaggregate level rather than on overall stagnancy.

Devi *et al*. (1990), used semi-log, exponential and quadratic growth models to analyse the trends in area, production and productivity of Banana in Kerala. The annual average rate was computed from these functions.

Kaushik (1993) studied the growth of Oilseeds production in India for the period of 1968-69 to 1991-92. The period was divided into two periods period 1 (1980-81 to 1991-92) to clearly bring out the trend in the more recent period. The exponential growth model was used. The study revealed that during period 1, most of the growth in oilseeds output was due to growth in area, where as in period 2 it was mainly due to improvement in productivity.

Ali and Singh (1995) studied the growth by using time series data on area, yield and production of wheat for twenty years, from 1970-71 to 1989-90. The total period was divided into two sub-periods: 1970-71 to 1979-80 as period 1; and 1980-81 to 1989-90 as period 2. The linear growth function was, then fitted to the data using OLS method.

Gan-YanTai *et al*. (1996) worked on evaluation of selected nonlinear regression models in quantifying seedling emergence rate of spring wheat. Fast and uniform seedling emergence increases yield potential of spring wheat in short-season areas. In this study the relative effectiveness of the Gompertz, Logistic, and Weibull models in quantifying emergence rate of spring wheat was compared. Wheat cv. Roblin was grown in a growth room at soil water potentials of -0.002, -0.165, -0.41, -1.00 or -1.45 MPa. Daily-recorded emergence data were fitted to each of the models. The Gompertz and Logistic models functioned in a similar way with great stability and accuracy in most cases. The Gompertz predictions most closely fitted the observed set of responses with residual points scattered around zero. For log normally distributed emergence patterns common under field conditions, the Gompertz model provided the most appropriate characterization of emergence.

Coelho *et al*. (1998) examined on Estimating limits of soil water availability as a function of bulk density and soil texture. Regression models are evaluated for predicting the water content at the limits of water availability in soil, based on texture and bulk density for various particle size classes. The nonlinear models presenting the bulk density in clay and clay loam soils and silt in clay and sand clay loam soils as independent variables showed better fitting. The bulk density as an independent variable resulted in more cases of good fits that the texture parameters.

Bergamasco *et al*. (2001) attempted to study Non-linear models fitting to growth data in females of the Holstein heifers. Brody, Gompertz and Logistic models applied to growth data in Holstein heifers were evaluated. Age-weight data from 65 Holstein heifers daughters of 20 bulls were used. The parameters were estimated by the generalized least squares method using nonlinear regression models with autocorrelated errors. Models were compared using some indicators of goodness of fit and biological interpretation of parameters. All of the models adjusted very well to the data. The Gompertz model was used to describe these data.

Mohanty *et al*., (2006) analysed the impact of puddling, tillage and residue management on wheat (*Triticum aestivum* L.) seedling emergence and growth in a rice-wheat system using nonlinear regression models. The nonlinear regression models study indicated that the Logistic model predicted the shoot growth of wheat under different tillage and residue management practices better than the Gompertz model. Whereas, for root growth the Monomolecular model fitted well with the experimental data.

Gaddour *et al*. (2008) analysed on adjustment of the kid's growth curve in pure goat breeds and crosses under southern Tunisian conditions. Five non-linear statistical models, Gompertz, Brody, Richards, Logistic and Exponential, were tested to fit the kid's growth curve parameters and shape of indigenous goat, Alpine, Damascus and their crosses. Data from 16 years periodical weighing was used to adjust growth curve before 5 months age of 1687 suckled kids. Gompertz model was the best to adjust kid's growth evolution.

Chetana (2009) examined prediction models in teak based agroforestry systems in northern transitional zone of Karnataka. Different prediction models namely Linear, Quadratic, Cubic, Exponential, Growth, Sigmoid, vapor pressure, Hassel, MMF compound, Logarithmic, Logistic, Weibull, Gompertz, Power etc., have been tried to predict the diameter and height growth of teak tree. MMF model was found better followed by Gompertz and Weibull for diameter prediction whereas for height growth prediction MMF model followed by logistic and Richards were found to be better.

Sreekanth *et al.* (2009) used a prediction model to forecast ground water level at Maheshwaram watershed, Hyderabad, India. The model efficiency and accuracy were measured based on the root mean square error (RMSE) and coefficient of determination (RP^{2P}). The model provided the best fit and the predicted trend followed the observed data closely (RMSE = 4.50 and RP^{2P} = 0.93).

Swetha (2009) worked on the Impact of rain water harvesting on farming economy who used nine models (Linear, Logarithm, Power, Compound, S-curve, Logistic, Growth, Exponential and Inverse models) and observed that linear model was showing significant and best model. The R-square was used to compare and choose the best fit model. The variation in ground water level was found to increasing over the periods as the temperature increases.

2.3 Fitting ARIMA and Exponential Smoothing

2.3.1 Box-Jenkins Model and its Application

A class of ARIMA (Auto Regressive Integrated Moving Average) model is called Box-Jenkins model. Box and Jenkins popularized it during late sixties. The application of these models for predicting prices of agricultural commodities is very few. Some studies which have used this modular, reviewed below.

Leuthold *et al.* (1970) forecasted daily hog prices and daily quantities supplied by using several alternative techniques. A distinction between econometric and the Box Jenkins models was made. It was stated that the former identified and measured both economic and non-economic variables affecting price and quantity, while the latter identified the stochastic components. The models were tested using Theils 'U' coefficient and the authors concluded that the econometric models yielded slightly superior forecasts. Finally, it was concluded that although better forecasts would be obtained by econometric models yet stochastic models were less prone to error and were less expensive.

Schmity and Walts (1970) forecasted wheat yield changes in four largest wheat exporting countries US, Canada, Australia and Argentina using Box Jenkins models. These forecasts were compared with those obtained by exponential smoothing using Theils 'U' inequality coefficient and concluded that forecasts with parametric modelling gave better results for the US but not for the others.

Chatfield and Protharo (1973) observed that the Box Jenkins procedure was not suitable for the sale forecasts with a multiplicative seasonal component. In this analysis, monthly data on sales of a company was used. The adequacy of the model was tested using Box-Pierce Test.

Govindan (1974) used Box Jenkins model to analyze wholesale price indices of rice, wheat, jowar and gram. The short term forecasts were found to give good results while the same was not true of long term forecasts. Janus quotients of the forecasts showed that the model gave good results.

Newbold and Granger (1974) compared the forecast performance of the Box-Jenkins, Holt-winters and step-wise regression models. The study indicated that each method had its own advantage over the others. It was opined that the Box-Jenkins gave better forecast in the short–run, but the method required time and skill to compute. The results indicated that for time series with less than 30 observations, step wise regression was better. For data between 30 to 50 observations, a combination of Holt-winters and step wise regression was found suitable. For series of 50 and above the Box-Jenkins performed well. For data with strong seasonal and long fluctuations, the Holt-winters model was suggested.

Protharo and Wallis (1976) examined the extent to which variations in a series could be explained first by a dynamic econometric model and then by ARIMA model. Econometric model clearly indicated that they provided a closer estimate of behaviour of the series during the sample periods.

Chatfield (1977) observed that the Box Jenkins approach being a valuable addition in the forecast tool bag which gave a deeper understanding of time series behaviour. Even though it was found to be more expensive yet the accuracy justified the cost.

Makridakis and Hibbon (1979) averred that accuracy of forecasts are negatively associated with the error term. Several tests to arrive at the accuracy of forecasts like mean square error (MSE), Theils 'U' coefficient and mean absolute percentage error (MAPE) were suggested.

Chengappa (1980) applied the Box Jenkins model to forecast poor sale and export auction prices of coffee. Monthly data were used and due to the distinct seasonal variation in prices, the ARIMA seasonal model was applied. The poor sale price forecasts were found to be accurate when compared to forecast of export prices. This was attributed to a possible lack of stationarity of the data. Hence adoption of differencing procedure or a transformation to make the data stationary was found necessary for a better estimate of export prices.

Achoth (1985) analyzed the supply, price and trade of Indian tea by fitting ARIMA models to data on prices and production. The moving average models were found to be most suitable. Among the price series a particular month's price was not related to the price of the immediate previous month but significantly related to the price of same month in previous years. However, the production in a particular month was related both to production of the previous month as well as to the production of same month in previous years. The forecasts yielded reasonably good results as judged from the tests of their efficiency. The forecasts of prices were superior when compared to the forecasts of quantities, which was attributed to the highly structured pattern of price behaviour.

Achoth (1985) fitted the seasonal ARIMA model to price data of tea at Calcutta and cochin auctions to production data of Northern and Southern regions of the country and quantity of tea exports and their prices. He identified that the moving average model was most suitable. The forecasts from these models yielded reasonably good ex-post and ex-ante forecasts judging from the test of their efficiency. By for the forecast of prices were superior to the forecasts of quantities which may be due to the predictable pattern of price behaviour. Further, some of the models fitted to the quantity series did reveal a certain degree of inadequacy which was not considered serious probably because certain cyclic pattern may not have been captured by the model.

Devaiah *et al.* (1988) attempted forecasting the prices of cocoons at Ramnagaram market by using ARIMA models. The forecasts were made for 13 months from April 1987 to April 1988. The forecasted values were observed to be close to he actual prices.

Lanciotti (1990) presented a paper that analysis of time series data of monthly prices for a group of diary products with the aim of obtaining reliable forecasts. The method of analysis employed is ARIMA as put forward by Box-Jenkins. The time series data covers both wholesale and retail prices for butter, Gorgonzola, Provolone, Grana Padano and Pasmigiano Reggiano. To estimate the reliability of the forecast obtained, a comparison is made with those resulting from naïve models do not require any estimates. Indicators on the accuracy of the forecasts show that except for Grana Padana, Le ARIMA forecasts are better.

Yin-Runsheng and Mins-Rs (1999) conducted timber price forecasts were univariate Auto Regression Integrated Moving Average (ARIMA) models employing the standard Box-Jenkins modeling strategy by using quarterly price series Timber Mart South. The results showed that most of the selected pipe pulpwood and saw timber markets in six southern US states can be evaluated using ARIMA models, and that short-term forecasts, especially those of one lead forecast, are fairly accurate. It is suggested that forecasting future prices could aid timber producers and consumers alike in timing harvests reducing uncertaining and enhancing efficiency.

Mastny (2001) used ARIMA models, also called Box and Jenkins models after their developers, is a group of models allowing the analysis of the time series with various features. The article demonstrates the possible usage of the Box-Jenkins methodology for the analysis of time series for agricultural commodities. The paper contains a basic mathematical explanation of ARIMA models together with a practical illustration of a price development forecast for a selected agricultural commodity.

Gangadharappa (2005) fitted ARIMA model to study the variation in arrivals and prices of potato in Bangalore, Belgaum, Kolar, Hassan and Hubli markets of Karnataka during 1996-97 to 2003-04. Box-Jenkins method was applied for precise forecasting of arrivals and prices of potato for the monthly data to all the selected markets. Of all the ten series, he found only two series, which yielded Box –Pierce 'Q' statistic which was significant and AIC was minimum.

Punitha (2007) attempted to fit ARIMA model to forecast the values of arrivals and prices of maize and ground nut for Davengere market and Hubli market. The forecasted values of groundnut arrivals and prices showed an increasing trend in Davangere market, but in Hubli market prices showed decreasing trend. The forecasted values of arrivals and prices of maize showed an increasing trend in both the markets.

Satya *et al.* (2007) made an attempt to forecast milk production using statistical time series modeling techniques such as double exponential smoothing and Auto- Regressive Moving Average (ARIMA) for the study period of twenty five years (1980-81 to 2004-05). On validation of the forecast from these models, ARIMA model performed better than the other one.

Barathi (2009) examined the price behaviour of mulberry silk cocoon in Ramnagar and Siddlaghatta market. Box-Jenkins ARIMA model was used to forecast the monthly arrivals and prices of mulberry silk cocoons in Ramanagara and Siddlagatta markets.

Chandrakala (2009) analysed spatial and temporal behavior of arrivals and prices of groundnut in Karnataka. ARIMA model was employed to forecast the arrivals and prices of groundnut in selected markets. Among five markets (Challakere, Chitradurga, Bellary, Yadagir and Davangere markets) the Bellary market yielded the best results.

2.3.2. Exponential Smoothing Methods and its Application.

Belov *et al.*, (1985) examined on promising chopper mechanisms for forage harvesters. Design features of the chopping mechanisms of forage harvesters are briefly evaluated and future manufacturing trends are analysed by means of time series analysis and exponential smoothing. Harvesters having cylinder choppers and blowers constituted > 50 percent of the existing population and it was predicted that this trend would continue. The proportion of harvesters having disc chopper-forwarders would reach 10-12 percent in the coming 5 yr.

Deluyker *et al*.,(1987) analysed on modeling daily milk yield in Holstein cows using time series analysis. Time series analysis of milk yields of cows milked 3 times daily was carried out on 513 partial or complete lactation yield records. It was found that the exponential smoothing function was most appropriate for the modelling of individual milking and daily yield data. Model parameters were influenced by parity, stage of lactation, occurrence of missed milkings and treatment for diseases. An examination of the residual variances showed that the model to forecast daily total yield performed as well as the model to forecast individual-milking yield.

Sisak (1989) worked on the principle of adaptive models of time series with regard to short term forecasting and the possibilities for application to cost planning. An adaptive model for the exponential smoothing of time series data is used to determine short term forecasts in the development of production costs for a farm forestry enterprise in Czechoslovakia. The results obtained are compared to those derived from an extrapolation of regression estimates. It is argued that the exponential smoothing technique can be more easily used in the computer programs for farm management and provide better quality forecasts for farm planning/budgeting than the regression model forecasts.

Manurung (1991) analysed on forecasting of oil palm hectarage and the need for seed in the second long term development plan. Forecasts of oil palm hectarage in Indonesia over the period of the second long term development plan (1994-2018) are made using the double exponential smoothing method. The average forecast growth rate is 3.24 percent per year. Three seed sources in Indonesia will be able to supply the projected number of seeds needed to support this development in oil palm area.

Sheldon., (1993) worked on forecasting tourism: expenditures versus arrivals. This examines issues relating to the measurement and forecasting of international tourist expenditures and arrivals. It shows that the two series fluctuate differently, and examines the accuracy of six different forecasting techniques (time series and econometric causal models) to forecast tourism expenditures. The results show that the accuracy of the forecasts differs depending on the country being forecast, but that the no-change model and Brown's double exponential smoothing are, overall, the two most accurate methods for forecasting international tourism expenditures

Mohan (1995) worked on forecasting weekly reference crop evapotranspiration series. The time variant characteristics of evapotranepiration (ET) necessitate the need for forecasting ET. In this, two techniques, namely a seasonal ARIMA model and Winter's exponential smoothing model, have been investigated for their applicability for forecasting weekly reference crop ET. A seasonal ARIMA model with one autoregressive and one moving average process and with a seasonality of 52 weeks was found to be an appropriate stochastic model. The ARIMA and Winter's models were compared with a simple ET model to assess their performance in forecasting. The forecast errors produced by these models were very small and the models showed promise of great use in real-time irrigation management.

Lim (2001) worked on forecasting tourist arrivals. Various exponential smoothing models are estimated over the period 1975-99 to forecast quarterly tourist arrivals to Australia from Hong Kong, Malaysia, and Singapore. The root mean squared error criterion is used as a measure of forecast accuracy. Prior to obtaining the one-quarter-ahead forecasts for the period 1998-2000, the individual arrival series are tested for unit roots to distinguish between stationary and non-stationary time series arrivals. The Holt-Winters Additive and Multiplicative Seasonal models outperform the Single, Double, and the Holt-Winters Non-Seasonal Exponential Smoothing models in forecasting. It is also found that forecasting the first differences of tourist arrivals performs worse than forecasting its various levels.

Rancheva (2002) attempted to study trends of average yield changes of cereals in Bulgaria. This estimates the main trend in the productivity/yields of cereals in Bulgaria, using exponential smoothing methods. Time series data for the period 1940-95 are used in the analysis.

Vasanthakumar (2002) worked on statistical evaluation of price variation in tropical Timbers. Exponential smoothing model is preferred to the multiplicative time series model for forecasting purpose. The single parameter exponential smoothing model was used to predict the prices of teak, rosewood and yellow teak.

Kumar *et al*., (2005) analysed on Price forecasting of different classes of teak by the application of exponential smoothing model. A single-parameter exponential smoothing model was used to forecast prices of different classes of teak in the Dandeli timber depot in Karnataka, India. Price data for the period May 1987-May 2001 were used, and both ex-post and ex-ante forecasts were made. The results of the ex-post forecasts reveal that the predicted prices are close to the actual prices.

Gajendra Singh (2006) examined assessment of food security situation for disaster risk management: an analysis for the Gujarat State. This study analyzes and forecasts the food security situation in the state of Gujarat, India. The study is based on the premises that the gap between food supply and food requirement is a better indicator of the food security level than the usually adopted supply-demand gap. The study considers the exponential smoothing and moving averages for making projections of food grains. The estimates of food grain production and requirement indicate that the overall cereals and pulses requirement would continue to be in deficit in both periods.

Huertas *et al*. (2007) attempted to study forecasting international tourist demand using Holt-Winters. This examined forecasts of international tourism arrivals to Spain. A survey was conducted on residents from 10 other major origin countries with respect to their future visits to Spain. The Holt-Winters exponential smoothing model was used to forecast the residents' demand for tourism in Spain by 2007-08.

2.4 Comparison of all the above models and selection of the best model.

Alon *et al*., (2001) have analyzed forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods. Neural networks versus other traditional methods such as winters exponential smoothing, Box-Jenkins ARIMA and multivariate regression, on retail sales data.

Adhikari *et al*., (2007) worked on Modelling swine supply response using a structural time series approach. Pig supply response in Georgia (USA) was analysed using a structural time series model (STSM) with both seasonal and yearly effects introduced stochastically. Pig production is divided into four production phases. Parameters of each phase were then estimated using a structural time series model. Out-of-sample forecasting robustness of the STSM was compared against a commonly used deterministic trend and deterministic seasonal components (DTDS) model using root mean square error (RMSE) and mean absolute percentage error (MAPE) criteria. The STSM model produced lower RMSE and MAPE values than those obtained using a DTDS modelling approach.

Dursun Aydin (2007) worked on comparison of the Non parametric regression models using smoothing spline and kernel regression. The study was about using of nonparametric models for national product data in turkey and Stanford heart transplant data. It is discussed two nonparametric techniques called smoothing spline and kernel regression. The main goal is to compare the techniques used for prediction of the nonparametric regression models. The accuracy measures MAPE and MSE were used. According to the results of numerical studies, it is concluded that smoothing spline regression estimators are better than those of the kernel regression.

TWeisent *et al*., T(2007) attempted to Compare three time-series models for predicting campylobacteriosis risk. Three time-series models (regression, decomposition, and Box–Jenkins autoregressive integrated moving averages) were applied to national surveillance data for campylobacteriosis with the goal of disease forecasting in three US states. Datasets spanned 1998–2007 for Minnesota and Oregon, and 1999–2007 for Georgia. Mean absolute percent error, mean square error and coefficient of determination (T*R*TP^{2P}) were the main evaluation fit statistics. Results showed that decomposition best captured the temporal patterns in disease risk. Training dataset T*R*TP^{2P} values were 72·2 percent, 76·3 percent and 89·9 percent and validation year T*R*TP^{2P} values were 66·2 percent, 52·6 percent and 79·9 percent respectively for Georgia, Oregon and Minnesota. All three techniques could be utilized to predict monthly risk of infection for T*Campylobacter*T sp. However, the decomposition model provided the fastest, most accurate, user-friendly method. Use of this model can assist public health personnel in predicting epidemics and developing disease intervention strategies.

Chen *et al*. (2008) made an attempt to compare forecasting models in tourism. This study uses three major U.S. national parks as applications of statistically selecting appropriate methods to forecast attendance. Forecasting methods assessed include Naive 1, Naive 2, single moving average (SMA), single exponential smoothing (SES), Brown's, Holt's, autoregressive integrated moving average (ARIMA), derived time series cross-section regression (TSCSREG), and time series analysis with explanatory variable models. The mean absolute percentage error (MAPE) is used to measure the accuracy of forecasting methods. Based on the MAPE values, SMA produces the most accurate forecasting, followed closely by ARIMA, Brown's, and Naive 1 models. Holt's and TSCSREG models produce the next most accurate forecasting, followed by SES, time series analysis with explanatory variable model, and Naive 2.

Chu-FongLin (2008) analyzed and forecasted tourism demand with ARAR algorithm. This study investigates the ARAR model and its usefulness as a forecast-generating mechanism for tourism demand for nine major tourist destinations in the Asian-Pacific region. The analysis is conducted on monthly inflow of international visitors covering the period 1975:01-2006:12. In monthly series, we consider 6-, 12-, 18-, and 24-months-ahead forecasting horizons. The accuracy of the forecasts in most cases is robust to the forecasting horizon based on such forecasting performance metric as mean absolute percentage error (MAPE) and root mean square error (RMSE). The study is further expanded by including quarterly series in the forecasting exercise to ensure the reliability of the forecasting evaluation. The 2-, 4-, 6-, and 8-quarters-ahead forecasting horizons are investigated. A comparison between forecasts generated by monthly and quarterly data reveals that the performance is broadly similar. Finally, the forecasts of ARAR are compared with those of seasonal ARIMA, which has been proven reliable in many forecasting contexts. Indeed, ARAR model can be deemed as credible alternatives when modeling and forecasting tourism demand.

Nesreen *et al*. (2009) analyzed on Empirical Comparison of Machine Learning Models for Time Series Forecasting. In this large scale comparison study for the major machine learning models for time series forecasting was done. Specifically, they applied the models on the monthly M3 time series competition data (around a thousand time series). The models considered are multilayer perceptron, Bayesian neural networks, radial basis functions, generalized regression neural networks (also called kernel regression), K-nearest neighbor regression, CART regression trees, support vector regression, and Gaussian processes. The study reveals significant differences between the different methods. The best two methods turned out to be the multilayer perceptron and the Gaussian process regression.

3. METHODOLOGY

To realize the various objectives of the study, an appropriate methodology describing sampling designs, data collection and tools of analysis for the conduct of the study are inevitable. In this chapter the methodology adopted for the present study, including the selection and description of the study area, sampling design, collection of data and analytical tools employed are presented under the following heads

3.1 Description of the study area

3.2 Collection of data

3.3 Statistical tools employed

3.1 Description of the study area:

India is one of the oldest civilizations with Kaleidocopic variety and rich cultural heritage. It covers an area of 3287263 sq.km, extending from the snow covered Himalayan heights to the tropical rain forests of the south. As the seventh largest country in the world, India stands apart from the rest of Asia, marked off by mountains and the sea, which give the country a distinct geographical entity. Bounded by the great Himalayas in the north, it stretches southwards and the Tropic of cancer tapers off into the Indian Ocean between the Bay of Bengal on the east and the Arabian Sea on the west.

Lying entirely in the northern hemisphere, the mainland extends between latitudes 8P^{0P}4P^{IP} and 37P^{0P}6P^{IP} north, longitudes 68P^{0P}7P^{IP} and 92P^{0P}25P^{IP} east and measures about 3214 km from north to south between the extreme latitudes and about 2933km from east to west between the extreme longitudes. It has a land frontier of about 15200km. The total length of the coastline of the mainland, Lakshadweep islands and Andaman and Nicobar islands in 7516.6km (MIB, GOI, 2000)

Based on the estimation of population reference Bureau (2000), India's population is 1002.1 million growing at a natural rate of 1.80% per annum. The second largest populous country, India is the home of 16.5% of the world's population. The country, however, accounts for 2.42% of the total world area.

India has achieved multi-faceted Socio-economic progress during the last few years of Independence. It has become self sufficient in Agricultural production and is now the tenth industrialized country in the world and the sixth nation to have gone into outer space to conquer nature for the benefit of people.

India is one among the few countries endowed ideal and best suited agro-climatic conditions where all types and ranges of tropiuca, subtropical and temperate fruits and vegetables are grown throughout year in one or the other part of the country.

3.1.1 Distribution of major Fruit crops in India

India is bestowed with a varied agro-climate, which is highly favourable for growing a large number of horticultural crops such as fruits, vegetables, root tuber, ornamental, aromatic plants, medicinal herbs, spices and plantation crops like coconut, arecanut, cashew and cocoa. Presently, horticultural crops occupy around 10% of gross cropped area of the country producing 160.75 million tones. India is the second largest producer of fruits and vegetables. Total production of fruits has been estimated at 49.36 million tones from 4.81 million ha. Vegetables occupy an area of 6.3 million ha with a production of 93.00 million tones. Our share in world fruit and vegetable production is 10% and 13.28 % respectively.

India has a large range of varieties of fruit in its basket and accounts for 10% of worlds total fruit production. Mango, Banana, Citrus, Pineapple, Papaya, Guava, Cheeku, Jackfruit, Litchi and Grape among the tropical and sub-tropical fruits Apple, Pear, Peach, Plum, Apricot, Almond and Walnut among the temperate fruits and Aonla, Ber, Pomegranate, Annona, Fig, Phalsa among the arid zone fruits are important. India leads the world in the production of Mango, Banana, Cheeku and Acid lime and in productivity of Grapes per unit land area.

India is the largest producer of Mango, Banana, Cheeku and Lime. About 39% of world's Mango and 23% of world's Banana is produced in the country. In Grapes, India has recorded the highest productivity per unit area in the world. The overall production of horticultural crops registered an increase of 8% during 2004-05 as compared to 2003-04. The percentage increase in fruit production has been to the tune of 1.5% during the period. (*India 2009, Annual Journal*)

3.1.2 Mango

Mango, cultivated in India since times immemorial, is regarded as the National Fruit of the country. Described as the "Food of the Gods", in the sacred Vedas, the fruit is grown almost in all parts of India, except the hilly areas, but is mainly available in the summer season only. There are more than 100 varieties of mangos in India, in a range of colors, sizes, and shapes. The common names used in context of the fruit are, Mangot, Manga, and Mangou. The exact origin of the term 'mango' is not known. It is believed to have come from the Portuguese term 'manga', which is probably from Malayalam 'manga'.

Mango finds a mentioned in the Indian history as well. In fact, the famous poet Kalidasa is known to have sung its praises. Apart from that, ancient Greek King Alexander the Great and Chinese pilgrim Hieun Tsang have been said to have savored its taste. Historical records also mention the instance of Mughal King Akbar planting 100,000 mango trees in Darbhanga, known as Lakhibagh. Mangos, liked for their sweet juice and bright colors all around the world, are known to be rich in vitamin A, C and D.

3.1.2.1 Description

Mangoes are available in different sizes, ranging from 10 to 25 cm in length and 7 to 12 cm in width. In terms of weight, a single mango can be as heavy as 2.5 kg. The fruit come in a wide variety of colors, such as green, yellow, red, and even various combinations of all these colors. Mango has a flat, oblong seed in the center, which is covered by the sweet pulp. Covering the pulp is a thin layer of skin, which is peeled off before eating the fruit. When ripe, the unpeeled fruit gives off a distinct, resinous sweet smell.

3.1.2.2 The important commercial varieties are as :

Sl.no	State	Varieties
1	Andhra Pradesh	Banganapalli, Suvarnarekha, Neelum and Totapuri
2	Bihar	Bombay Green, Chausa, Dashehari, Fazli, Gulabkhas, Kishenbhog, Himsagar, Zardalu and Langra Kishen Bhog, Himsagar, Zardalu and Langra
3	Gujarat	Kesar, Alphonso, Rajapuri, Jamadar, Totapuri, Neelum, Dashehari and Langra
4	Haryana	Chausa, Dashehari, Langra and Fazli
5	Himachal Pradesh	Chausa, Dashehari and Langra
6	Karnataka	Alphonso, Totapuri, Banganapalli, Pairi, Neelum and Mulgoa
7	Madhya Pradesh	Alphonso, Bombay Green, Dashehari, Fazli, Langra and Neelum
8	Maharashtra	Alphonso, Kesar and Pairi
9	Punjab	Chausa, Dashehari and Malda
10	Rajasthan	Bombay Green, Chausa, Dashehari and Langra
11	Tamil Nadu	Alphonso, Totapuri, Banganapalli and Neelum
12	Uttar Pradesh	Bombay Green, Chausa, Dashehari and Langra
13	West Bengal	Fazli, Gulabkhas, Himsagar, Kishenbhog, Langra and Bombay Green

(APEDA Database)

3.1.2.3 Cultivation

Frost-free climate is best for the growth of Mangos. If temperatures drop below 40° F, even for a short period, the flowers and small fruits already grown on the tree can get killed. In other words, warm and dry weather is required for the cultivation of the fruit. This is it available in the summer season only. Mango can grow well in large containers and a greenhouse as well. Mango trees are shady in nature. They grow very fast and can reach a height of as much as 65 ft. The life of mango trees is generally very long and some specimens are known to be over 300 years old and still fruiting.

3.1.3 Banana

Banana is said to be a common man's fruit, because there is large area under banana in India, its per hectare yield is also the highest (30 - 35 MT/ha) of all the fruits. Therefore its production is large and the prices are within the reach of common people. Total area under banana is 3.70 lakh ha, next to Mango. Among the States, Maharashtra contributes the maximum area of about 90,000 ha. In Maharashtra, area under banana is concentrated in Jalgaon district with nearly 60% area in that district.

Other districts growing banana are Parbhani, Nanded and Thane. Banana fruits from Jalgaon area are marketed in most parts of India, particularly in the north. In Tamil Nadu, banana is grown in Cauvery delta area.

Banana is a very perishable fruit and hence its marketing faces many problems such as chain of middlemen, transport, storage, etc. In local or short distance marketing number of intermediaries is small (one or two) but in long distance marketing commission agents and wholesalers are involved in addition to co-operative society.

3.2 Database:

The secondary data on Export of fruit crops was collected from below mentioned sources based on the availability as mentioned below:

- Annual Fruit crops Export data was collected from the CMIE, Foreign Exchange Reviews from 1994 to 2009 (in US million$).
- Annual Mango Export data was collected from the APEDA DATABASE from 1985 to 2009 (in Rs crores).
- Annual Banana Export data was collected from FAO STAT DATABASE from 1985 to 2008(in 1000$US).
- Monthly fruit crop Export data was collected from CMIE, Economic Intelligence Service Monthly reviews from 2000(June) to 2009(March) (in US million$).

3.3 Statistical tools Employed:

Following methods were employed to analyze the data, interpret the results, to draw influences and to design policy options for adoption by farmers, researches and government.

3.3.1 Polynomial function fitting Approach:

Under the first approach polynomial functions of different degrees were fitted to the Export data set using the method of least squares. Over a long period of time, the time series is very likely to show tendency to increase or decrease overtime. There are different types of trend. Some of them are linear and some are non linear in their form. For shorter period of time, in most of the situation, the straight line provided the best description of trend and for longer period of time, the non – linear form generally provides a good description of trend.

The polynomial regression is used to determine the long term behaviour.The polynomials tried are shown below.

1P^{stP} degree (straight line) $Y_x = a_0 + a_1 * x + c$

2P^{ndP} degree polynomial (Quadratic) $Y_x = a_0 + a_1 * x + a_2 * x^2 + c$

3P^{rdP} degree polynomial (Cubic) $Y_x = a_0 + a_1 * x + a_2 * x^2 + a_3 * x^3 + c$

4P^{thP} degree polynomial $Y_x = a_0 + a_1 * x + a_2 * x^2 + a_3 * x^3 + a_4 * x^4 + c$

5P^{thP} degree polynomial $Y_x = a_0 + a_1 * x + a_2 * x^2 + a_3 * x^3 + a_4 * x^4 + a_5 * x^5 + c$

6P^{thP} degree polynomial

$Y_{x=} a_0 + a_1 * x + a_2 * x^2 + a_3 * x^3 + a_4 * x^4 + a_5 * x^5 + a_6 * x^6 + c$

The best fitting of the data with the polynomial function:

$Y(x) = a_0 + a_1 * x + a_2 * x^2 + a_3 * x^3 + ... + a_n * x^n + c$ will be found

Where,

Y_x= trend values at time t

c = error term

$a_0, a_1, a_2, a_3, a_4, a_5, a_6 ...... a_n$ = the coefficients to be estimated.

The suitable model for data i.e., goodness of fit of model will be assessed by computing R^2 (co efficient of determination) value.

3.3.1.2 Coefficient of determination (R^2)

Coefficient of determination (R^2) is used as the measure of explanatory value of the model. It is calculated as,

$$R^2 = \frac{\text{(SS due to regression)}}{\text{(Total sum of squares)}} = \frac{b. \sum (X_i - \bar{X})(Y_i - \bar{Y})}{\sum (Y_i - \bar{Y})^2}$$

R^2 measures the contribution of the linear function of independent variables to the variation in dependent variable. Based on the R^2, model of best fit to the data was selected.

3.3.2 Non-linear Regression Models

3.3.2.1 Logarithmic model

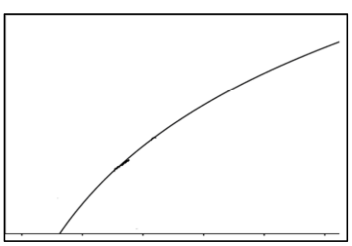
$Y = b_0 + b_1 \ln(t)$ is the linear form of the model.

Y and t_i's are value and time respectively.

b_0 and b_i's are constants to be estimated and

ln is Natural Log.

The curve of the model is given below



3.3.2.2 Power model

Power model has the form,

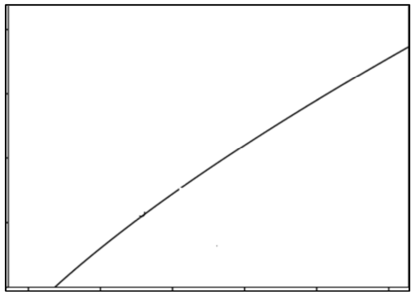
$Y = b_0 t^{b_1}$

on transformation

$\ln(Y) = \ln(b_0) + b_1 \ln(t)$ is obtained which is of the linear form where, Y and t_i's are value and time respectively.

b_0 and b_i's are constants to be estimated and ln is Natural Log

The shape of the curve in this model is,



3.3.2.3 Compound model

Compound model has the form,

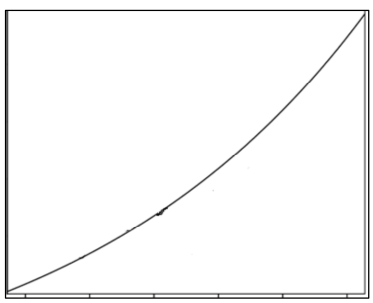
$Y = b_0 \; b_1^t$

On transformation

$\ln(Y) = \ln(b_0) + \ln(b_1)$ t. is obtained which is of the linear form where,

Y and t_i's are value and time respectively.

b_0 and b_i's are constants to be estimated and ln is Natural Log.
The shape of the curve in this model is,



3.3.2.4 S-curve

S-curve has a form

$Y = e^{(b0 + (b1/t)}$ or

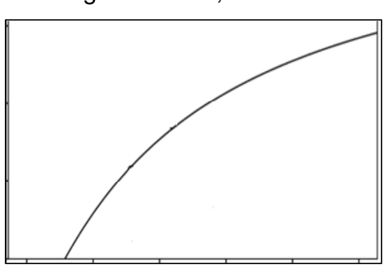
$$\ln(Y) = b_0 + \left(\frac{b_1}{t}\right)$$ which is a linear form of the model.

In this study Y and t_i's are value and time respectively.

b_0 and b_i's are constants to be estimated

ln is Natural Log and "e" is the exponential function.

The S-curve is given below,



3.3.2.5 Logistic model

Logistic model or Verhulst model. The word "logistic" has no particular meaning in this context, except that it is commonly accepted. The second name honors P. F. Verhulst, a Belgian mathematician who studied this idea in the 19th century. Using data from the first five U.S. censuses, he made a prediction in 1840 of the U.S. population in 1940 -- and was off by less than 1%.

It has a form

$$Y = \frac{1}{\left(\frac{1}{u} + b_0 b_1^t\right)}$$

or

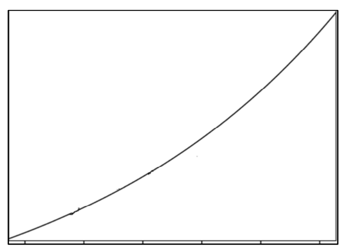
$\ln\left(\frac{1}{y} - \frac{1}{u}\right) = \ln(b_0) + t\ln(b_1)$ is a linear form of the model. In this model,

Y and t_i's are value and time respectively.

b_0 and b_i's are constants to be estimated,

ln is Natural Log and u is the upper boundary value.

The logistic curve is given below



3.3.2.6 Growth model

Growth model has the form,

$Y = e^{(b0 + (b1 * t))}$

On transformation,

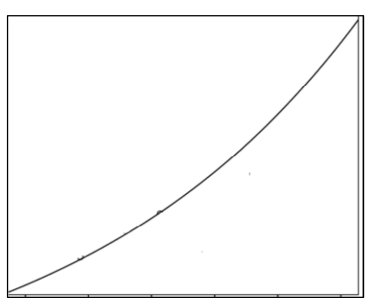
$\ln(Y) = b_0 + (b_1 \times t)$ is obtained which is of the linear form where,

Y and t_i's are value and time respectively.

b_0 and b_i's are constants to be estimated,

ln is Natural Log and “e” is the exponential function.

The shape of the growth curve is,



3.3.2.7 Exponential model

Model under consideration is,

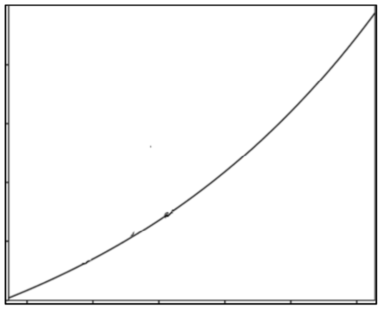
$Y = b_0 \times e^{(b1 * t)}$

or

$\ln(Y) = \ln(b_0) + (b_1 t)$ where,
Y and t_i's are value and time respectively.

b_0 and b_i's are constants to be estimated,

ln is Natural Log and "e" is the exponential function.
The logistic curve is given below



3.3.2.8 Gompertz model:

Gompertz Relation: $Y=b_0 * b_1^{e^t}$

$Y=b_3 \exp[-b_1 \exp\{-b_0 (t-1)\}]$

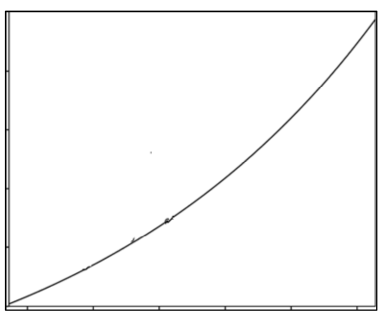
Where, Y and t_i's are value and time respectively

b_0, b_1 and b_2 are the parameters in the model.

The Gompertz reliability growth model is often used when analyzing reliability data. It is most applicable when the data set follows a smooth curve.

The Parameter Estimation for the Gompertz Models can be carried out by Using Least Squares in Nonlinear Regression as well as using linear regression methods.

The gompertz curve is shown below:



3.3.2.9 Simple linear regression analysis

The export value was predicted using linear regression analysis by considering time as independent variable:

$Y = \alpha + \beta x + \varepsilon$ was the model under consideration (NageshwaraT Rao, 2^{nd} Editn, 2007).

Where,

α is the intercept

β is the regression co-efficient.

ε is the random error.

Here the task was to find out the estimates of $\hat{\alpha}$ (α = a) and $\hat{\beta}$ (β = b). They were calculated as

$$\hat{\beta} = b = \frac{n\left(\sum XY\right) - \sum X \sum Y}{n \sum X^2 - \left(\sum X\right)^2}$$

$$\hat{\alpha} = a = \bar{Y} - b\bar{X}$$

3.3.2.10 Significance of the model

The significance of the regression model was tested using F statistic. Here the null hypotheses was set as,

HB_0: β = 0

To test this F – test was used.

$$F = \frac{\text{Regression sum of squares / (n-2)}}{\text{(Error sum of squares) / (n-k)}}$$

Where,

Regression sum of squares = $= \sum (b_i)\left(\sum x_i y\right)$

Error sum of squares = $\sum y^2 - SSR$

This calculated F value was compared with table F for 1 and (n-2) degrees of freedom.

R^2 measures the contribution of the linear function of independent variables to the variation in dependent variable. Based on the R^2 and MSE, model of best fit to the data was selected. (under 3.3.1.2 R^2 measure and 3.3.4 MSE).

3.3.5 Box-Jenkins models

The Box-Jenkins procedure is concerned with fitting a mixed Auto Regressive Integrated Moving Average (ARIMA) model to a given set of data. The main objective in fitting this ARIMA model is to identify the stochastic process of the time series and predict the future values accurately. These methods have also been useful in many types of situation which involve the building of models for discrete time series and dynamic systems. But, this method was not good for lead times or for seasonal series with a large random component (Granger and Newbold, 1970).

Originally ARIMA models have been studied extensively by George Box and Gwilym Jenkins during 1968 and their names have frequently been used synonymously with general ARIMA process applied to time series analysis, forecasting and control. However, the optimal forecast of future values of a time-series are determined by the stochastic model for that series. A stochastic process is either stationary or non-stationary. The first thing to note is that most time series are non-stationary and the ARIMA model refer only to a stationary time series. Therefore, it is necessary to have a distinction between the original non-stationarity time series and its stationarity counterpart.

3.3.5.1 Stationarity and non-stationarity

The term stationarity meaning that the process generating the data is in equilibrium around a constant value and that the variance around the mean remains constant over time. The data must be roughly horizontal along time axis.

If mean changes over time (with some trend cycle pattern) and variance is not reasonably constant then series is non-stationary in both mean and variance.

If a time series is not stationary, then it can be made more nearly stationary by taking the first difference of the series. Conversely a stationary process may be summed or integrated to give a non-stationary process.

Let X_t be a random variable and x_t (where t=1, 2, . . . n) be the observations on X_t with density function f (x_t). If the observations are independent, then

$$f(X_1, X_2 .... X_n) = f_1(x_1) f_2(x_2)....f_n(x_n) \qquad 3.1$$

This implies that joint distribution is independent of historical time.

The assumption of stationarity reduces the number of parameters in the joint probability density function of a random variable x_t in the series.

Since the ARIMA models refer only to a stationary time series, the first stage of Box-Jenkins model is reducing non-stationary series x_t to a stationary series Y_t by taking first differences as follows.

$$Y_t = \Delta X_t$$

$$= X_t - X_{t-1}$$

$$= X_t - BX_t$$

$$= (1 - B)X_t \qquad 3.2$$

Where,

B = Backward shift operator

The backward shift operator is convenient for describing the process of differencing. To define B, such that

$B_i X_t = X_{t-1}$ i= 1, 2,. . .

Suppose the first difference of the series doesn't become stationary then second order differencing is done as follows

$$Y_t = \Delta(\Delta X_t) \qquad 3.3$$

$$= \Delta(X_t - X_{t-1})$$

$$= (X_t - X_{t-1}) - (X_{t-1} - X_{t-2})$$

$$= X_t - 2X_{t-1} + X_{t-2}$$

$$= X_t - 2BX_t + B^2 X_t$$

$$= (1 - 2B + B^2)X_t$$

$$= (1 - B)^2 X_t$$

In general, if it takes a d^{th} order difference to achieve stationarity we will write.

d^{th} order difference $= (1 - B)^d X_t$ 3.4

The general ARIMA (o, d, o) model will be

$$(1 - B)^2 X_t = e_t \qquad 3.5$$

Where e_t is error term distributed normally with

$E(e_t) = 0, V(e_t) = e_t^2$ and

$Cov(e_i, e_j) = \theta$ for all t ($i \neq j$)

In order to test the stationarity, compute the auto-correlation functions (ACF) of difference series (Y_t) up to 16 lags. If the ACF for first and higher differences (after 2-3 lags) drop abruptly to zero then it indicates the series is stationary.

3.3.5.2 Stationary time series model

3.3.5.2.1 Auto regressive process (p, o, o)

If the observation Y_t depends on previous observation and error term e_t is called auto regressive process (AR process)

$$Y_t = \mu + \varnothing_t Y_{t-1} + \varnothing_2 Y_{t-2} + \ldots\ldots + \varnothing_p Y_{t-p} + e_t$$

$$= \varnothing_p (B) (Y_t - \mu) + e_t \qquad 3.6$$

Note the term μ in equation (3.5) is not quite the same as the "Mean" of the Y series. Rather, the development is as follows.

$$(Y_t - \mu) = \phi_p (Y_t - \mu) + e_t$$

$$= \phi_1 (Y_{t-1} - \mu) + \phi_2 (Y_{t-2} - \mu) + \ldots. + \phi_p (Y_{t-p} - \mu) + e_t \qquad 3.7$$

$$= \phi_1 (Y_{t-1} - \phi_1 \mu) + \phi_2 (Y_{t-2} - \mu) + \ldots + \phi_p (Y_{t-p} - \mu) + e_t$$

$$Y_t = (\mu - \phi_1 \mu - \ldots - \phi \mu) + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + e_t$$

$$= \mu_1 + \phi_1 Y_{t-1} + \ldots + \phi_p Y_{t-p} + e_t$$

3.3.5.2.2 Moving average process (o, o, q)

If the observation Y_t depends on the error term e_t and also on one or more previous error terms (e_t 's) then we have moving average (MA) process.

$$Y_t = \mu + e_t - \theta_1 e_{(t-1)} - \theta_2 e_{(t-2)} - \ldots.\ \theta_q e_{(t-q)} \qquad 3.8$$

Where,

$\theta_i = i^{th}$ moving average parameter

$i = 1, 2, \ldots. q$

q = Order moving average

3.3.5.2.3 Mixtures : ARIMA process

If the non-stationarity is added to a mixed ARIMA process, then the general ARIMA (p, d, q) is implied. Here the word integrated is confusing to many and refers to the differencing of the data series.

$$(1-B)^d (1 - \phi_p B^p) Y_t = \mu + (1 - \phi_p B^q) e_t \qquad 3.9$$

Seasonality and ARIMA models

Some time series exhibit perceptible periodic pattern for instance price and arrivals of Agricultural commodities usually have a seasonal pattern process than the general.

The ARIMA notation can be extended readily to handle seasonal aspects and the general shorthand rotation is ARIMA

(p.d.q.)	(P.D.Q.)s
(non-seasonal part of the model)	(Seasonal part of the model)

s = number of periods per season

The mixture of AR and MA seasonal model is

$$\varnothing_p (B)\ \Delta^d \varnothing_p(B^s)\ \Delta^D x_t = \theta_q (B)\ .\ (H)_Q (B^s)\ e_t \qquad 3.10$$

If $Y_t = \Delta^d \Delta^d x_t$ – the model becomes an integrated model.

The main stages in setting up a Box-Jenkins forecasting model are as follows.

1. Identification
2. Estimating the parameters
3. Diagnostic checking and
4. Forecasting

3.3.5.2.4 Identification of models

A good starting point for time series analysis is a graphical plot of the data. It helps to identify the presence of trends.

Before estimating the parameter (p, q) of model, the data are not examined to decide about the model which best explains the data. This is done by examining the sample ACF (Autocorrelation function) and PACF (Partial Autocorrelation function) of differenced series Y_t.

The sample auto correlations for k time lags can be found and denoted by r_k as follows.

$$\hat{\rho}(Y_t) = r_k(Y_t) \qquad 3.11$$

$$= \frac{C_k(Y_t)}{C_0(Y_t)}$$

where,

$$C_k(Y_t) = \frac{1}{n}\sum_{t-1}^{n-k}\left(Y_t - \overline{Y}\right)\left(Y_{t+k} - \overline{Y}\right)$$

K = 0, 1, 2, ... n

t = 1, 2, . . . n-k

$$Y_t = \frac{1}{n}\sum_{t=1}^{n} Y_t$$

n = Length of time period

Both ACF and PACF are used as the aid in the identification of appropriate models. There are several ways of determining the order type of process, but still there was no exact procedure for identifying the model.

3.3.5.2.5 Estimation of parameters

After tentatively identifying the suitable model, next step is to obtain Least Square Estimates of the parameters such that the error sum of squares is minimum.

$$S(\theta, \varnothing) = \sum e_t^2(\theta, \varnothing) \qquad 3.12$$

where,

t = 1, 2, 3 . . . n

There are fundamentally two ways of getting estimates for such parameters.

a) Trail and error: Examine many different values and choose set of values that minimizes the sum of squares residual

b) Interactive method: Choose a preliminary estimate and let a computer programme refine the estimate interactively.

The latter method is used in our analysis for estimating the parameters.

3.3.5.2.6 Diagnostic checking

After having estimated the parameters of a tentatively identified ARIMA model, it is necessary to do diagnostic checking to verify that the model is adequate.

Examining ACF and PACF of residuals may show an adequacy or inadequacy of the model. If it shows random residuals, then it indicates that the tentatively identified model was adequate. When an inadequacy is detected, the checks should give an indication of how the model need be modified, after which further fitting and checking takes place.

One of the procedures for diagnostic checking mentioned by Box-Jenkins is called over fitting i.e. using more parameters than necessary. But the main difficulty in the correct identification is not getting enough clues from the ACF because of inappropriate level of differencing. The residuals of ACF and PACF considered random when all their ACF were within the limits.

$$\mu \pm 1.96\sqrt{1/(n-12)} \qquad 3.13$$

The minimum Akike Information Coefficient (AIC) criterion is used to determine both the differencing order (d, D) required to attain stationarity and the appropriate number of AR and MA parameters, it can be computed as follows.

$$AIC_{(p+q)} = \{(1+\log 2\pi) + n\log\sigma^2 + 2m\} \qquad 3.14$$

Where,

σ^2 = Estimated MSE

n = Number of observations

m = p + q + P + Q

This diagnostic checking helps us to identify the differences in the model, so that the model could be subjected to modification, if need be.

3.3.5.2.7 Forecasting

After satisfying about the adequacy of the fitted model, it can be used for forecasting. Forecasts based on the model.

$$(1-\varnothing B)\ (1-\phi B)^s\ Y_t = (1-\theta B)\ (1-(H)^s B)\ e_t \qquad 3.15$$

were computed for upto 4 months (m) ahead. The above model (3.16) gives the forecasting equation as

$$Y_t = \varnothing Y_{t-1} + \phi Y_{t-12} - \varnothing\phi Y_{t-13} + e_t - \theta e_{t-1} - (H)\ e_{t-12} + \theta(H)\ e_{t-13} \qquad 3.16$$

Given data upto time't' the optional forecast of Y (also called Ex-Ante forecast) model at the t is the conditional expectation of Y_{t+1}.

It follows, in particular, that

$$e_t = Y_t - Y_{t-1} \qquad 3.17$$

The errors e_t in model (3.18) are in fact that forecast errors for unit lead time. That for an optimal forecast these 'one step ahead' forecast errors ought to form an uncorrelated series is otherwise obvious. Suppose, if these forecast errors were autocorrelated, then it could be possible to forecast the next forecast error in which case it could not be optimal.

The required expectations are easily found because

$$E(Y_{t+m}) = Y_t(m), E(e_{t+m}) = 0 \qquad 3.18$$

Where,

m = 1, 2, 3 . . . . . n

$$E(Y_{t-m}) = Y_{t-m} E(e_{t-m}) = a_{t-m} = Y_{t-m} - Y_{t-m-1} \qquad 3.19$$

Where, m = 0, 1, 2 . . . n

For instance, to determine the three month ahead (1-3) forecast for series YB_{tB} (use equation 3.17).

$Y_{t+1} \quad = Y_{t+3}$

$= \varnothing Y_{t+2} + \phi Y_{t-9} - \varnothing\phi Y_{t-10} + e_{t+13} - \theta e_{t-2} - (H)\ e_{t-9} + \theta\ (H)\ e_{t-10}$

taking conditional expectations at time t,

$Y_t\ (1) = Y_t\ (3)$

$= \varnothing\ Y_{t(2)} + \phi y_{t-9} - \varnothing\ \phi Y_{t-10} + 0\text{-}\theta\ (0)\text{-}(H)\ (Y_{t-9}\text{-}Y_{t-10}) + \theta(H)\ (Y_{t-10}-Y_{t-11})$

Because, $E(e_{t+1)} = 0, E(e_{t-1}) = Y_{t-1} - \hat{Y}_{t-1} = e_{t-1}$

i.e. $Y_t\ (3) = 0\ Y_t\ (2)$

The forecast Y_t (2) can be obtained in a similar way in terms of Y_t (1) from E (Y_{t+2}). Similarly Y_t (1) can be obtained from E (Y_{t+1}).

In practice it is very easy to compute the forecast Y_t (1), Y_t (2), Y_t (3) etc. recursively using the forecast function (3.19).

E (Y_{t+1}) = E(Y_{t+1}-1 + Q_{t+1}- e_{t+1}-1) - θ e_{t+1}-1 – (H) e_{t+1}-12 + θ(H) e_{t+1}-13 and using 3.18 and 3.19.

However, using these methods, Ex-post forecasts can also be calculated for comparing with the value actually realized. The accuracy of forecasts for both Ex-ante and Ex-post were tested using the following tests MAPE and MSE mentioned under 3.3.4.

3.3.6 Exponential smoothing method for forecasting

If the weights for exponential smoothing are available, the calculation can be carried out very simply by the formula

New Estimate = W (last measurement) + (1-W) (last estimate)

The term “last estimate” refers to the previous estimated made in the same way and “W’ is called the smoothing factor at the start of series, any reasonable value can be choosen for the “last estimate term”, it is quickly discounted by the method and any effects due to an error in the initial value quickly disappear. This system using a discounted average is referred to an exponential smoothing or adoptive smoothing.

3.3.6.1 Single exponential smoothing method

This method is also called as Method of estimation of forecasts of single weight / parameter. The equal weights have been given to all items, under exponential smoothing, the weights assigned in geometric progression. Greater weights are assigned to recent observation and smaller weights are given distant observation.

Suppose the weights in geometric series are given below.

$$1, (1\text{-}W), (1\text{-}W)^{2P} \ldots (1\text{-}W)^{n-1} \qquad \ldots\ldots\ldots\ldots\ldots\ldots.. (1)$$

Where weights lies between zero and one, and n = number of observations

$$S_t = \frac{(1\times X_{t-1}) + (1-W)X_{t-2} + (1-W)^2 X_{t-3} + ... + (1-W)^{n-1} X_{t-(n-1)}}{1 + (1-W) + ... + (1-W)^{n-1}} \quad ... (2)$$

Similarly,

$$S_{t+1} = \frac{(1\times X_t) + (1-W)X_{t-1} + (1-W)^2 X_{t-2} + ... + (1-W)^{n-1} X_{t-(n-2)}}{1 + (1-W) + ... + (1-W)^{n-1}} \quad ... (3)$$

Where, X_t = current observation values

X_{t-1}, X_{t-2} …$X_{t-(n-1)}$ = previous observed values

S_{t+1} = current forecasting value

S_t = last estimate value

Taking n large value and neglecting higher power of weights (W) and (1-W). Smoothed or new forecast = W (previous observed value) + (1-W) old forecast

OR

$$S_{t+1} = WX_t + (1-W)S_t \quad …………….. (4)$$

In relation of equation 4, $\overline{X}_{t+1}$, the smoothed value at the next period of time is W times the previous period observed values pus (1-W) times the preceding smoothed value, which is itself an average smoothed value. Now to make a forecast for each successive period, these, smoothed value are used to find out a change in each period, W is called the smoothing coefficient. In this particular method only one smoothing W is used assuming timber prices are random. The single parameter exponential smoothing is used in estimating forecast Export value of fruit crops. The first smoothed value is taken as average of first three observed price values.

3.3.6.1.1 Choice of Weight (W)

The value of W can obtained in two ways, one is by trial and error method and another by Pricey time series forecasting software, such as *MINITAB*, *EViews* etc, which use an algorithm to select the value of the W that minimizes mean square error for in-sample forecasts. In the present study MINITAB package is used to the obtain value of appropriate value of W.

3.3.6.2 Double Exponential smoothing method

This method is also called as trend adjusted exponential smoothing. This situation can be improved by the introduction of a second equation with a second smoothing constant/ second weight, W_2; assuming monthly export value is influenced by the trend component. This must be chosen in conjunction with W_1. The two equations associated double exponential smoothing given by

$$S_t = W_1 X_t + (1-W_1)(S_{t-1} + b_{t-1}) \quad ………….. (1)$$

$$b_t = W_2(S_t - S_{t-1}) + (1-W_2)b_{t-1} \quad ……………… (2)$$

Where, S_t = smoothened value at time period t

S_{t-1} = smoothened value at time period t-1

W_1 = level smoothing constant

X_t = actual price at time period t

b_t= trend estimate of the time period t

b_{t-1}= trend estimate of the period t-1

W_2= trend smoothing constant

Taking different value for both W_1 and W_2 by trial and error method both trend factor and level factor is estimated. Hence by using both trend estimate and level estimate, forecast value for the period m is obtained by

$$F_{t+m} = S_t + mb_t \text{ .... (3)}$$

In relation of equation 3, forecast value for the *m* period ahead is *m* times the trend estimate of the time period t plus the level estimate of the time period t. initial value of the trend estimate (b1) and level estimate (S1) or first smoothened value is taken as follows

$S_1 = X_1$

$$b_1 = \frac{(X_2 - X_1) + (X_4 - X_3)}{2}$$

Where, X_1, X_2, X_3 and X_4 are observed values for the initial three season of the timber time series data set. The current value of the series is used to calculate its smoothed value is replacement in double exponential smoothing. The second smoothing equation then updates the trend, which is expressed as the difference between the last two values. The equation is similar to the basic form of single smoothing, but here applied to the updating of the trend.

3.3.6.3 Winter's seasonal exponential smoothing method

To handle seasonality present in the time series data, the third parameter is added. Introduced third equation along with W_3 takes care of seasonality. The resulting set of equations is called the "Holt-Winters" (HW) method after the names of the inventors. There are two main HW models, depending on the type of seasonality

Holt-winters multiplicative model:

$$S_t = W_1 \frac{X_t}{SI_{t-l}} + (1 - W_1)(S_{t-1} + b_{t-1})$$

Where, W_1 = level smoothing constant

SI = seasonality estimate

l = length of seasonal cycle.

Trend estimate is given by:

$$b_t = W_2\ (S_t - S_{t-1}) + (1 - W_2) b_{t-1}$$

Seasonality estimate given by:

$$SI_t = W_3 \frac{X_t}{S_t} + (1 - W_3) SI_{t-l}$$

Where, W_3 = seasonal smoothing constant

Then m – Period ahead forecast is given by:

$$F_{t+m} = (S_t + mb_t)SI_{t-l+m}$$

Holt-winters additive model:

Smoothing equation is given by:

$$S_t = W_1(X_t - SI_{t-l}) + (1 - W_1)(S_{t-1} + b_{t-1})$$

Trend estimate given by:

$$b_t = W_2(S_t - S_{t-1}) + (1 - W_2)b_{t-1}$$

Seasonality estimate given by:

$$SI_t = W_3(X_t - S_t) + (1 - W_3)SI_{t-l}$$

Forecast equation for additive seasonality:

$$F_{t+m} = S_t + mb_t + SI_{t-l+m}$$

This particular model is known as the additive model (seasonality factor is added to forecast).

3.3.6.3.1 Selection of weights (W_1, W_2, W_3)

The value of all three smoothing constant is obtained by trial and error method. Time series are analyzed by giving different weights and the best triple exponential model in each case is selected, based on the minimum MAPE value under different weights.

3.3.4 Comparison of all the above models and Selection of the best model.

The comparison of models was taken using accuracy of forecasts, tested using the following tests (Markidakis and Hibbon, 1979).

1) Mean square error (MSE); the formula for computing MSE is

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(X_t - \hat{X}_t)^2$$

Where,

X_t = Actual values

$\hat{X}_t$ = Predicted values

n = Number of observations

2) Mean average percentage error (MAPE): The formula for this is

$$MAPE = \frac{\frac{1}{n}\sum_{t=1}^{n}(X_t - \hat{X}_t)100}{X_t}$$

Where,

X_t = Actual values

$\hat{X}_t$ = Predicted values

n = Number of observations

The model was considered as best among all the models based on the least MAPE and MSE values.

4. RESULTS

Keeping in view the specific objectives of the present study, the data collected on the export of annual fruit crops, Mango, Banana and monthly export data of fruit crops of India have been subjected to various statistical methods as outlined in the materials and methods. The results of such analysis are presented in this chapter under the following heads.

4.1 Polynomial Function fitting approach.

4.2 Fitting of Non-linear regression models.

4.3 Fitting ARIMA and Exponential Smoothing.

4.4 Comparison of all the above models and Selection of the best model.

4.1 Polynomial function Fitting approach.

The polynomial function fitting was carried out with different degrees and the selection of the best model was based on R-square value. Tables 4.1, 4.2,4.3 and 4.4 showing the polynomials of different degrees were presented. Table 4.1 represents annual fruit export, table 4.2 represents Mango export, table 4.3 represents Banana export and table 4.4 representing the monthly export data of fruit crops.

In the table 4.1 which represents the annual fruit export different degrees were tried. Among all the degrees carried out 5P^{thP} degree polynomial has shown the highest R-sqaure , therefore this was considered as best fit. In table 4.2 which represents the Mango export 6P^{thP} degree polynomial with highest R-square was best fit. In the case of table 4.3 which represents the Banana export also the 6P^{thP} degree polynomial has shown the best fit. And with regard to the monthly export of fruit crops i.e., the table 4.4 showed that the Quadratic polynomial (2P^{ndP} degree) was the best fit with highest R-square.

4.2 Fitting of Non-linear regression models

The results of different non-linear regression models which are tried to fit were shown in the table 4.5, 4.6 and 4.7 for the annual fruit export, Mango and Banana yearly export data.

As a part of it nearly nine models were tried. Among all the models the best model was selected based on the highest R-square value and MSE. All models were found to be significant (at 1% and 5%). Different models were tried to fit for annual fruit crop exports based on this criteria, which is presented in table 4.5. Among all the models the Compound, Growth and Exponential models were best fit with highest R-square of 0.9413 and least MSE of 0.04. In case of Mango exports (table 4.6), the Linear model has shown best fit with the highest R-square value of 0.9876 and least MSE of 0.15, among all the models. And in case of Banana exports (table 4.7), the Compound, Growth and Exponential models were shown best fit with highest R-square value of 0.8010 and least MSE of 0.71, among all the models.

4.3 Fitting ARIMA and Exponential Smoothing.

4.3.1 Fitting ARIMA model

The detailed analysis of fitting of the ARIMA model for monthly export of fruit crops has been presented below

4.3.1.1 Identification of the model

The tentative models are first identified based on the Auto Correlation Function (ACF) and Partial Auto Correlation Function (PACF) for the different series. The computed value of ACF and PACF of the data is shown in Table 4.8 up to 16 lags. An examination of the ACF and PACF revealed seasonality. However, the series is found to be stationary, since the coefficient dropped to zero after the first or second lag. Each individual coefficient of ACF and PACF are tested for their significance using‘t’ test. Further, the absence of peak at first values clearly indicate suitability of the choice of non-seasonal difference d=1, to accomplish stationarity series. Hence, based on ACF and PACF many models are tried, finally model (1,1,1) (1,1,1) is identified for monthly export of fruits.

Table 4.1 Different degrees of Polynomials for Annual Fruits Export.

Degrees	Equation	R^2
$2P^{ndP}$	$y = 2.892xP^{2P} - 20.89x + 83.71$	0.983
$3P^{rdP}$	$y = 0.134xP^{3P} - 0.328xP^{2P} + 0.386x + 50.86$	0.988
$4P^{thP}$	$y = -0.038xP^{4P} + 1.378xP^{3P} - 13.41xP^{2P} + 50.58x - 0.799$	0.989
$5P^{thP}$	$y = -0.002xP^{5P} + 0.076xP^{4P} - 0.288xP^{3P} - 2.872xP^{2P} + 23.10x + 20.43$	0.994
$6P^{thP}$	$y = 0.001xP^{6P} - 0.067xP^{5P} + 1.266xP^{4P} - 10.88xP^{3P} + 43.61xP^{2P} - 67.51x + 77.18$	0.992

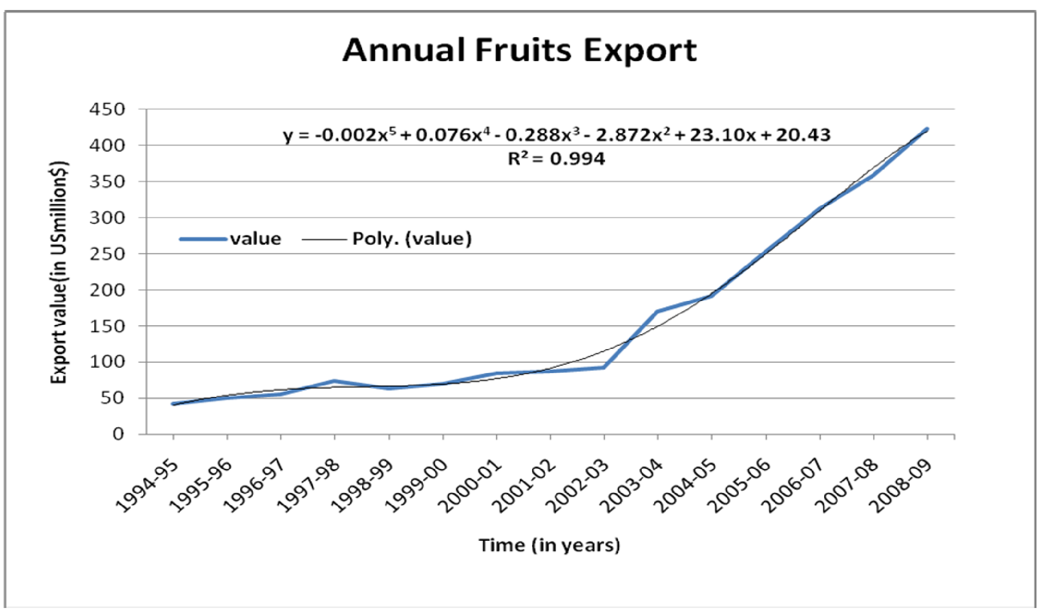


Figure: 4.1 Showing polynomial of Annual Fruits Export.

Figure: 4.1 Showing polynomial of Annual Fruits Export

Table 4.2 Different degrees of Polynomials for Mango Export.

Degrees	Equation	R^2
$2P^{ndP}$	$y = 0.181xP^{2P} + 0.923x + 17.48$	0.924
$3P^{rdP}$	$y = 0.004xP^{3P} + 0.014xP^{2P} + 2.564x + 13.86$	0.925
$4P^{thP}$	$y = 0.001xP^{4P} - 0.032xP^{3P} + 0.591xP^{2P} - 0.671x + 18.48$	0.928
$5P^{thP}$	$y = -0.001xP^{5P} + 0.013xP^{4P} - 0.315xP^{3P} + 3.233xP^{2P} - 10.54x + 28.73$	0.931
$6P^{thP}$	$y = -0.0005xP^{6P} + 0.003xP^{5P} - 0.078xP^{4P} + 0.895xP^{3P} - 4.474xP^{2P} + 10.53x + 11.53$	0.940
$7P^{thP}$	$y = -0.000002xP^{7P} - 0.0003xP^{6P} + 0.001xP^{5P} - 0.06xP^{4P} + 0.78xP^{3P} - 4.54xP^{2P} + 8.53x + 6.53$	0.938

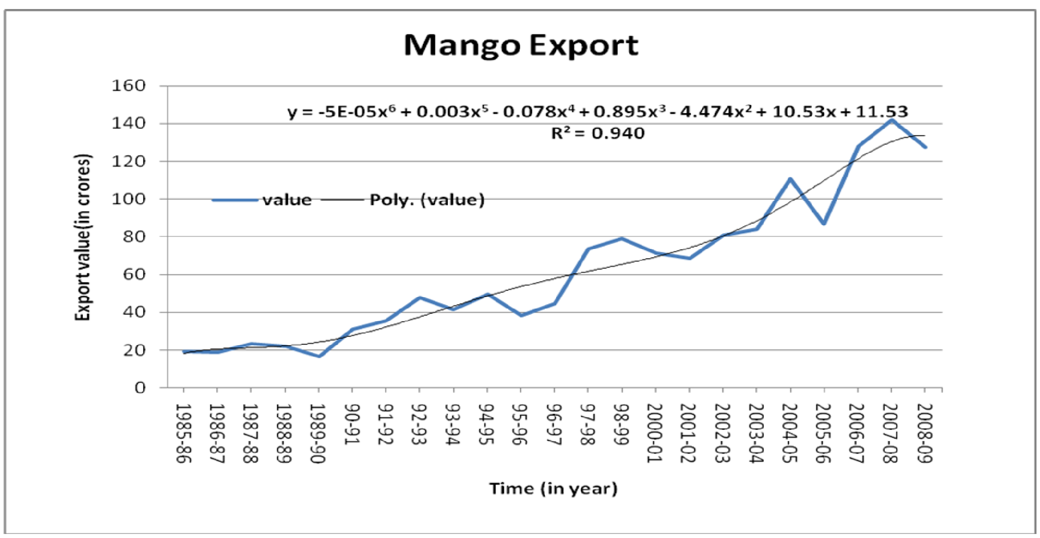


Figure: 4.2 Showing Polynomial of Mango Export.

Figure: 4.2 Showing Polynomial of Mango Export.

Table 4.3 Different degrees of polynomials for Banana Exports.

Degrees	Equation	R^2
$2P^{ndP}$	$y = 7.791xP^{2P} + 64.30x - 338.0$	0.771
$3P^{rdP}$	$y = -0.785xP^{3P} + 36.08xP^{2P} - 213.0x + 274.9$	0.780
$4P^{thP}$	$y = 0.201xP^{4P} - 10.43xP^{3P} + 187.1xP^{2P} - 1060.x + 1484.$	0.798
$5P^{thP}$	$y = 0.064xP^{5P} - 3.684xP^{4P} + 73.43xP^{3P} - 593.8xP^{2P} + 1859.x - 1546.$	0.828
$6P^{thP}$	$y = 0.002xP^{6P} - 0.082xP^{5P} + 0.360xP^{4P} + 20.25xP^{3P} - 255.4xP^{2P} + 933.9x - 791.1$	0.860
$7P^{thP}$	$y = 0.0003xP^{7P}+0.00012xP^{6P} - 0.06xP^{5P} + 0.0260xP^{4P} + 16.25xP^{3P} - 182.4xP^{2P} + 653.9x - 456.1$	0.852

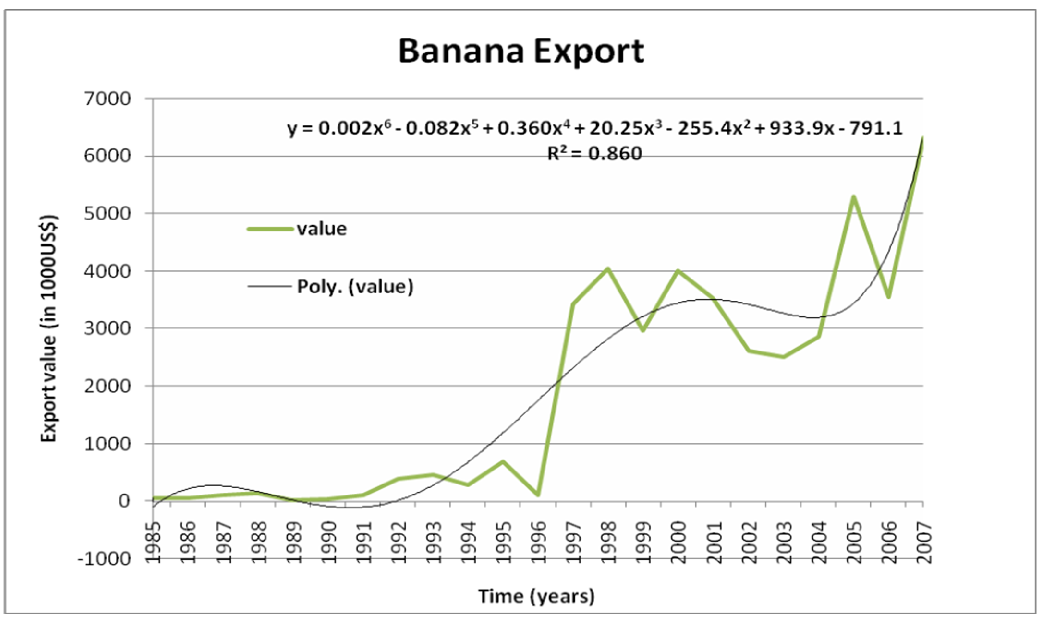


Figure: 4.3 Showing Polynomial of Banana Export

Figure: 4.3 Showing Polynomlal of Banana Export

Table 4.4 Different degrees of Polynomials for Monthly Export of fruit crops.

Degrees	Equation	R^2
$2P^{ndP}$	$y = 3.644+0.166x+0.00154xP^{2P}$	0.6255
$3P^{rdP}$	$y =3.457+0.18642x+0.00106xP^{2P}+2.9688xP^{3P}$	0..6252
$4P^{thP}$	$y = 4.252+0.043x+0.007xP^{2P} – 0.0008xP^{3P}+0.000000xP^{4P}$	0.6163

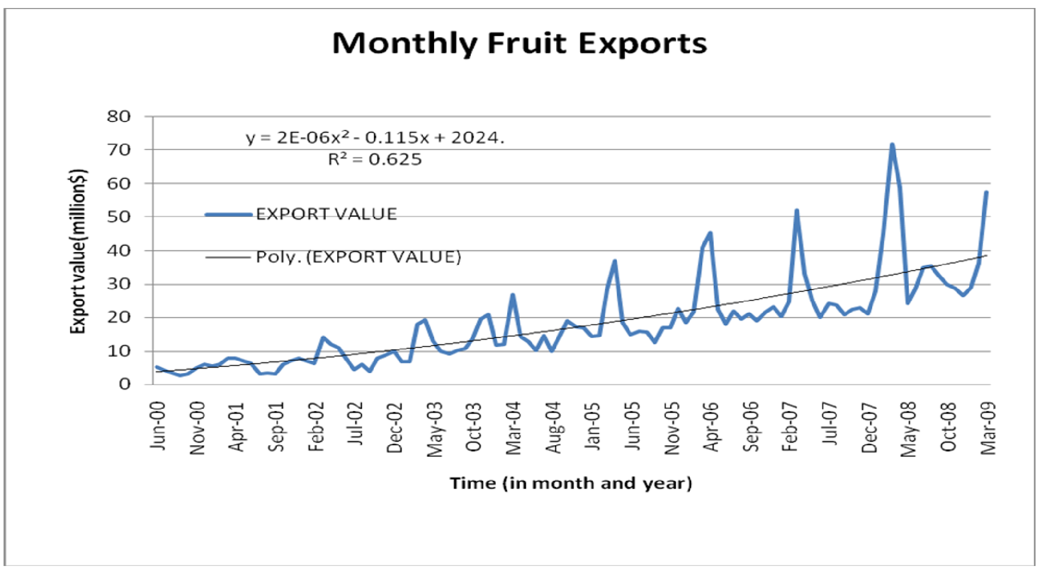


Figure: 4.4 Showing polynomial of Monthly Export of Fruit crops

Figure: 4.4 Showing polynomlal of Monthly Export of Fruit crops

Table 4.5 Different Non-linear regression models including linear model for Annual Fruits Export

Model	Equation	RP^{2P}	MSE
Linear	Y=-47.415+25.3818x	0.8256P^{**P}	29.84
Logarithm	y=-65.285+118.779ln(t)	0.5530 P^{**P}	7515.16
Power	Y=24.5*t P$^{0.85P}$	0.7424P^{**P}	0.16
Compound	y=31.133*1.180P^{tP}	0.9413 P^{**P}	0.04
S-Curve	y=eP$^{(5.215+(-2.03/t))P}$	0.4264P^{**P}	0.36
Logistic	y=1/(1/430+0.637*0.710P^{tP})	0.7662 P^{**P}	0.77
Growth	y=eP$^{(-327.27(0.1658t))P}$	0.9413P^{**P}	0.04
Exponential	y=7.383*eP$^{(0.166t)P}$	0.9413P^{**P}	0.04
Gompertz	Y=367.84*1.864P^{etP}	0.6200P^{**P}	17956.00

Note: ** Significance at 1%
* Significance at 5%

Table 4.6 Different Non-linear regression models including linear model for Mango Exports

Model	Equation	RP^{2P}	MSE
Linear	Y=-0.687115+5.285x	0.9876P^{**P}	0.15
Logarithm	y=-21.162+37.388ln(t)	0.667 0P^{**P}	499.16
Power	Y=10.219*t P$^{0.723P}$	0.8188P^{**P}	0.83
Compound	y=16.7620*1.0985P^{tP}	0.9310 P^{**P}	0.03
S-Curve	y=eP$^{(4.281+(-2.065/t))P}$	0.4432P^{**P}	0.25
Logistic	y=1/(1/145+0.0.904*0.824P^{tP})	0.8220P^{**P}	0.39
Growth	y=eP$^{(2.819(0.094t))P}$	0.9310P^{**P}	0.03
Exponential	y=16.7620*eP$^{(0.094t)P}$	0.9310P^{**P}	0.03
Gompertz	Y=2951.20*0.00698P^{etP}	0.7410P^{**P}	1573.71

Note: ** Significance at 1%
* Significance at 5%

Table 4.7 Different Non-linear regression models including linear model for Banana Exports

Model	Equation	R^2	RMSE
Linear	Y=-112.126+251.28t	0.7467**	118.81
Logarithm	y=-2027.1152+1749.512ln(t)	0.5374 **	24.40
Power	$Y=9.914*t^{1.858}$	0.6881**	1.12
Compound	$y=34.1734*1.276^{t}$	0.8010 **	0.71
S-Curve	$y=e^{(7.297+(-5.14/t))}$	0.3491**	2.34
Logistic	$y=1/(1/6325+0.064*0.693^{t})$	0.7455 **	2.21
Growth	$y=e^{(2.53145(0.2442t))}$	0.8010 **	0.71
Exponential	$y=34.1734*e^{(0.2442t)}$	0.8010 **	0.71
Gompertz	$Y=1.439*0.0025^{et}$	0.4910**	4278857.73

Note: ** Significance at 1%
* Significance at 5%

Table 4.8 ACF and PACF of Monthly Fruits Export.

	Value	
Lags	ACF	PACF
1	-0.446	-0.446
2	0.108	-0.113
3	-0.126	-0.155
4	0.084	-0.038
5	-0.113	-0.119
6	0.003	-0.138
7	.063	0.00
8	-0.003	0.00
9	-0.061	-0.080
10	-0.033	-0.128
11	0.210	0.168
12	-0.330	-0.230
13	0.175	-0.090
14	-0.121	-0.119
15	0.079	-0.132
16	-0.119	-0.161

ACF = Autocorrelation Function
PACF = Partial Autocorrelation Function

4.3.1.2 Estimation of parameters

After identifying the models tentatively the next step is to obtain the estimates by the method of Least Squares Estimates of the parameters ϕ and θ for the monthly export. Such that the error sum of square is to be minimum.

i.e. $S(\phi \,.\, \theta) = et^2 (\phi \,.\, \theta)$

The parameters of the tentatively identified models are estimated by an iterative process and then the residual of each of the models is to be estimated.

4.3.1.3 Diagnostic checking

Residual analysis is carried out to check the adequacy of the models. The residuals of ACF and PACF are obtained from the tentatively identified model. The adequacy of the model is judged based on the values of AIC and SBC. The values of the statistics are shown in Table 4.9. The model (1,1,1) (1,1,1) is found to be the best model for monthly export of fruits for the least statistics of AIC and SBC.

4.3.1.4 Forecasting the Monthly Exports of Fruit crops.

The method of forecasting has been explained in detail in chapter 3. Both Ex-ante and Ex-post forecast are done and it is compared with actual values of observations. The forecast is done up to July 2009. The results of Ex-ante and Ex-post forecast of monthly export of fruits is shown in Tables 4.11. The forecasts are also depicted in the Fig. 4.11. The accuracy of forecasts for both Ex-ante and Ex-post are tested using MSE and MAPE tests. The values MSE and MAPE are presented in Table 4.10, which are found to be least. Forecasted values of fruits showed an increasing trend.

4.3.2 Fitting Exponential Smoothing Methods

Under exponential smoothing, the weights (WB_{1B}, WB_{2B}, and WB_{3B}) assigned in geometric progression. All the three techniques of exponential smoothing methods were taken to fit for monthly export of fruit crops and presented below.

4.3.2.1 Single Exponential Smoothing Method

The equal weight has been given to all the observations, the weight is assigned by trail and error method. The weight (WB_{1B}) i.e., the alpha (level) was found to be 0.9 which is obtained by trail and error method. The accuracy measures were tested using MSE and MAPE tests. The values MSE and MAPE are presented in Table 4.13, which are found to be least. The model was selected based on the minimum MAPE value which was found to be 25.55 percent. The table 4.14 shows the actual and forecasted values of the model using single exponential smoothing constant.

4.3.2.2 Double Exponential Smoothing Method

This method is also called as Trend adjusted exponential smoothing. This method is improved by the introduction of a second smoothing constant i.e., trend,(WB_{2B}); assuming monthly export value is influenced by the trend component. This must be chosen in conjunction with WB_{1B}. So the constants obtained by trail and error method are 0.9 (level) and 0.1(trend) which are presented in table 4.15. The accuracy measures MAPE and MSE were presented in table 4.16 which are found to be least. The model was selected based on the minimum MAPE value which was found to be 27.97 percent. The table 4.17 shows the actual and forecasted values of the model using double exponential smoothing constants.

4.3.2.3 Winters' Seasonal Exponential Smoothing Method

To handle seasonality present in the time series data, the third parameter (WB_{3B}) is added. For the determination of values of the softening constants WB_{1B}, WB_{2B} and WB_{3B}, trail and error method was carried out based on the minimization of the mean absolute percentile error – MAPE of the forecasting.

The tables 4.18 and 4.19 shows the MAPE (26.3678) which was found least (WB_{1B} = 0.9, WB_{2B} = 0.1 and WB_{3B} = 0.1) for the multiplicative model. And for additive model (WB_{1B} = 0.8, WB_{2B} = 0.1 and WB_{3B} = 0.2), MAPE was 36.59 percent. Therefore for the multiplicative model considering the whole period (from june 2000 to march 2009), the found errors were

4.879 and the mean absolute percentile error – MAPE was 26.36 percent. For the additive model the values were 5.2149 and 36.596 percent respectively. Hence multiplicative model was selected as best. The table 4.20 shows the actual and forecasted values of the Winters multiplicative model using triple exponential smoothing constants.

4.4 Comparison of all the above models and Selection of the best model

For the model comparison monthly export of fruits was considered. The comparison of all the models carried out in the process was considered based on the MAPE and MSE values which were considered to be least. So according to the table 4.21 which shows the MAPE and MSE values the ARIMA model with least MAPE value of 15.198 and MSE value of 3.36 was considered as best fit among all the models considered. And the model next to it was found to be single exponential smoothing with MAPE value 25.554 and MSE value 3.36.

Table 4.9 Residual analysis of Monthly Export

EXPORT	AIC	SBC
VALUE	574.3108	586.974

Table 4.10 Selected measures of predictive performance of Box-Jenkins ARIMA model

EXPORTS	MSE	MAPE
VALUE	3.360	15.198

Table 4.11 Actual and forecasted values of Box-Jenkins ARIMA model for Monthly Fruits Export

Sl. No	Time	Actual	Forecasted	Sl. No	Time	Actual	Forecasted
1	Jun-00	5.4	.	44	Jan-04	11.9	15.39
2	Jul-00	4.2	.	45	Feb-04	12.1	11.78
3	Aug-00	3.4	.	46	Mar-04	27	20.83
4	Sep-00	2.7	.	47	Apr-04	14.3	23.58
5	Oct-00	3.3	.	48	May-04	12.8	14.09
6	Nov-00	5.1	.	49	Jun-04	10.3	12.21
7	Dec-00	6	.	50	Jul-04	14.6	10.39
8	Jan-01	5.5	.	51	Aug-04	9.9	14.38
9	Feb-01	6.1	.	52	Sep-04	14.4	11.27
10	Mar-01	7.9	.	53	Oct-04	19	16.88
11	Apr-01	8	.	54	Nov-04	17.4	20.91
12	May-01	7	.	55	Dec-04	17	19.76
13	Jun-01	6.3	.	56	Jan-05	14.3	13.35
14	Jul-01	3.1	5.13	57	Feb-05	14.8	14.73
15	Aug-01	3.5	2.89	58	Mar-05	28.4	27.31
16	Sep-01	3.2	2.87	59	Apr-05	37	21.85
17	Oct-01	6	3.73	60	May-05	18.5	25.71
18	Nov-01	7	6.87	61	Jun-05	15	16.02
19	Dec-01	7.9	7.45	62	Jul-05	16.1	17.03
20	Jan-02	7.2	7.02	63	Aug-05	15.7	15.13
21	Feb-02	6.4	7.56	64	Sep-05	12.6	18.02
22	Mar-02	14.2	8.64	65	Oct-05	16.9	19.44
23	Apr-02	12	11.97	66	Nov-05	17.1	20.99
24	May-02	11	10.03	67	Dec-05	22.8	20.80
25	Jun-02	7.6	9.56	68	Jan-06	18.6	18.31
26	Jul-02	4.5	5.76	69	Feb-06	21.8	18.95
27	Aug-02	6	5.15	70	Mar-06	41	33.83
28	Sep-02	4	5.52	71	Apr-06	45.5	36.32
29	Oct-02	8	6.54	72	May-06	22.4	28.31
30	Nov-02	9	9.02	73	Jun-06	18.2	19.50
31	Dec-02	10	9.78	74	Jul-06	21.9	20.83
32	Jan-03	6.9	9.26	75	Aug-06	19.7	20.01
33	Feb-03	6.8	7.84	76	Sep-06	21.1	19.89
34	Mar-03	18.2	12.96	77	Oct-06	19	24.77
35	Apr-03	19.3	14.93	78	Nov-06	21.8	22.07
36	May-03	13	15.27	79	Dec-06	23.3	26.09
37	Jun-03	10	10.66	80	Jan-07	20.5	21.36
38	Jul-03	9.2	7.38	81	Feb-07	24.8	23.72
39	Aug-03	10.2	9.23	82	Mar-07	52.3	41.21
40	Sep-03	10.9	8.21	83	Apr-07	33.3	50.31
41	Oct-03	14	12.34	84	May-07	25	21.59
42	Nov-03	19.7	13.56	85	Jun-07	20.2	22.41
43	Dec-03	21	16.88	86	Jul-07	24.4	24.52

Table 4.11 Contd…..

Sl. No	Time	Actual	Forecasted	Sl. No	Time	Actual	Forecasted
87	Aug-07	23.9	23.58	105	Feb-09	36.4	41.37
88	Sep-07	20.9	24.21	106	Mar-09	57.6	65.62
89	Oct-07	22.6	23.96	107	Apr- 09	.	51.11
90	Nov-07	23	25.70	108	May-09	.	30.70
91	Dec-07	21.3	27.98	109	Jun-09	.	31.79
92	Jan-08	27.9	22.44	110	Jul- 09	.	37.46
93	Feb-08	45.2	30.33	111	Aug-09	.	37.48
94	Mar-08	72	59.20	112	Sep-09	.	35.25
95	Apr-08	59	54.16	113	Oct-09	.	34.70
96	May-08	24.3	37.32	114	Nov-09	.	34.54
97	Jun-08	28.7	23.79	115	Dec-09	.	33.32
98	Jul-08	35.1	31.38	116	Jan-10	.	33.62
99	Aug-08	35.5	31.50	117	Feb-10	.	36.40
100	Sep-08	32.6	31.13	118	Mar-10	.	46.80
101	Oct-08	30.1	31.16	119	Apr-10	.	70.54
102	Nov-08	28.7	30.95	120	May-10	.	60.96
103	Dec-08	26.6	30.11	121	Jun-10	.	36.85
104	Jan-09	29	30.62	122	Jul- 10	-	38.30

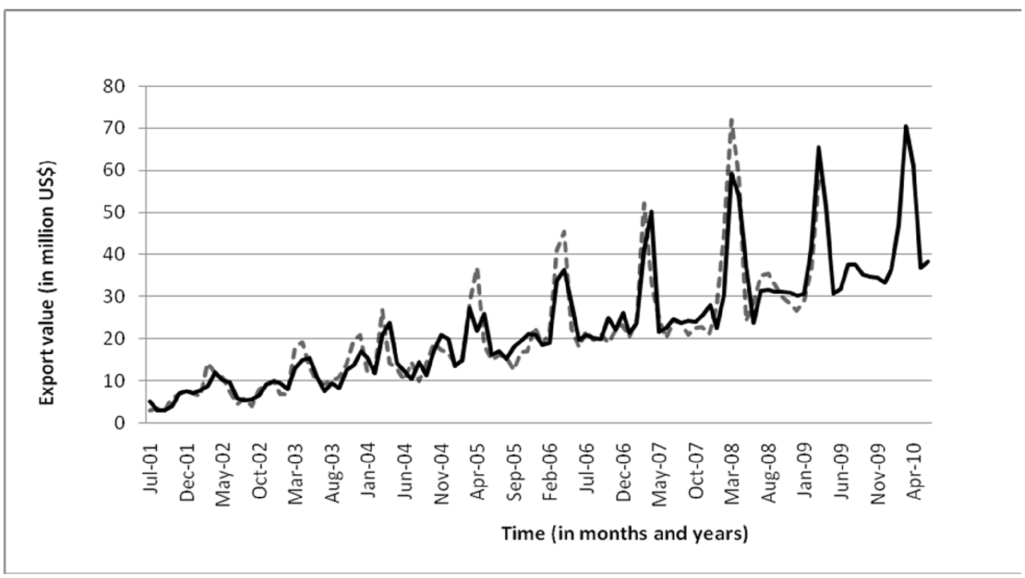


Figure: 4.5 Showing Actual and Forecasted values of ARIMA model

Figure : 4.5 Showing Actual and Forecasted values of ARIMA model

Table 4.12 Single Exponential Smoothing constant Alpha.

SMOOTHING CONSTANT	
Alpha(level)	0.9

Table 4.13 Selected measures of predictive performances of Single Exponential Smoothing

ACCURACY MEASURES	
MAPE	25.554
MSE	4.309

MAPE = Mean Absolute Percentage error

MSE = Mean Square Error

Table 4.14 Actual and Forecasted values of Single Exponential Smoothing for Monthly Fruits Export

Sl.No	Time	Actual	Forecasted	Sl. No	Time	Actual	Forecasted
1	Jun-00	5.4	4.01	54	Nov-04	17.4	19
2	Jul-00	4.2	5.4	55	Dec-04	17	17.4
3	Aug-00	3.4	4.2	56	Jan-05	14.3	17
4	Sep-00	2.7	3.4	57	Feb-05	14.8	14.3
5	Oct-00	3.3	2.7	58	Mar-05	28.4	14.8
6	Nov-00	5.1	3.3	59	Apr-05	37	28.4
7	Dec-00	6	5.1	60	May-05	18.5	37
8	Jan-01	5.5	6	61	Jun-05	15	18.5
9	Feb-01	6.1	5.5	62	Jul-05	16.1	15
10	Mar-01	7.9	6.1	63	Aug-05	15.7	16.1
11	Apr-01	8	7.9	64	Sep-05	12.6	15.7
12	May-01	7	8	65	Oct-05	16.9	12.6
13	Jun-01	6.3	7	66	Nov-05	17.1	16.9
14	Jul-01	3.1	6.3	67	Dec-05	22.8	17.1
15	Aug-01	3.5	3.1	68	Jan-06	18.6	22.8
16	Sep-01	3.2	3.5	69	Feb-06	21.8	18.6
17	Oct-01	6	3.2	70	Mar-06	41	21.8
18	Nov-01	7	6	71	Apr-06	45.5	41
19	Dec-01	7.9	7	72	May-06	22.4	45.5
20	Jan-02	7.2	7.9	73	Jun-06	18.2	22.4
21	Feb-02	6.4	7.2	74	Jul-06	21.9	18.2
22	Mar-02	14.2	6.4	75	Aug-06	19.7	21.9
23	Apr-02	12	14.2	76	Sep-06	21.1	19.7
24	May-02	11	12	77	Oct-06	19	21.1
25	Jun-02	7.6	11	78	Nov-06	21.8	19
26	Jul-02	4.5	7.6	79	Dec-06	23.3	21.8
27	Aug-02	6	4.5	80	Jan-07	20.5	23.3
28	Sep-02	4	6	81	Feb-07	24.8	20.5
29	Oct-02	8	4	82	Mar-07	52.3	24.8
30	Nov-02	9	8	83	Apr-07	33.3	52.3
31	Dec-02	10	9	84	May-07	25	33.3
32	Jan-03	6.9	10	85	Jun-07	20.2	25
33	Feb-03	6.8	6.9	86	Jul-07	24.4	20.2
34	Mar-03	18.2	6.8	87	Aug-07	23.9	24.4
35	Apr-03	19.3	18.2	88	Sep-07	20.9	23.9
36	May-03	13	19.3	89	Oct-07	22.6	20.9
37	Jun-03	10	13	90	Nov-07	23	22.6
38	Jul-03	9.2	10	91	Dec-07	21.3	23
39	Aug-03	10.2	9.2	92	Jan-08	27.9	21.3
40	Sep-03	10.9	10.2	93	Feb-08	45.2	27.9
41	Oct-03	14	10.9	94	Mar-08	72	45.2
42	Nov-03	19.7	14	95	Apr-08	59	72
43	Dec-03	21	19.7	96	May-08	24.3	59
44	Jan-04	11.9	21	97	Jun-08	28.7	24.3
45	Feb-04	12.1	11.9	98	Jul-08	35.1	28.7
46	Mar-04	27	12.1	99	Aug-08	35.5	35.1
47	Apr-04	14.3	27	100	Sep-08	32.6	35.5
48	May-04	12.8	14.3	101	Oct-08	30.1	32.6
49	Jun-04	10.3	12.8	102	Nov-08	28.7	30.1
50	Jul-04	14.6	10.3	103	Dec-08	26.6	28.7
51	Aug-04	9.9	14.6	104	Jan-09	29	26.6
52	Sep-04	14.4	9.9	105	Feb-09	36.4	29
53	Oct-04	19	14.4	106	Mar-09	57.6	36.4

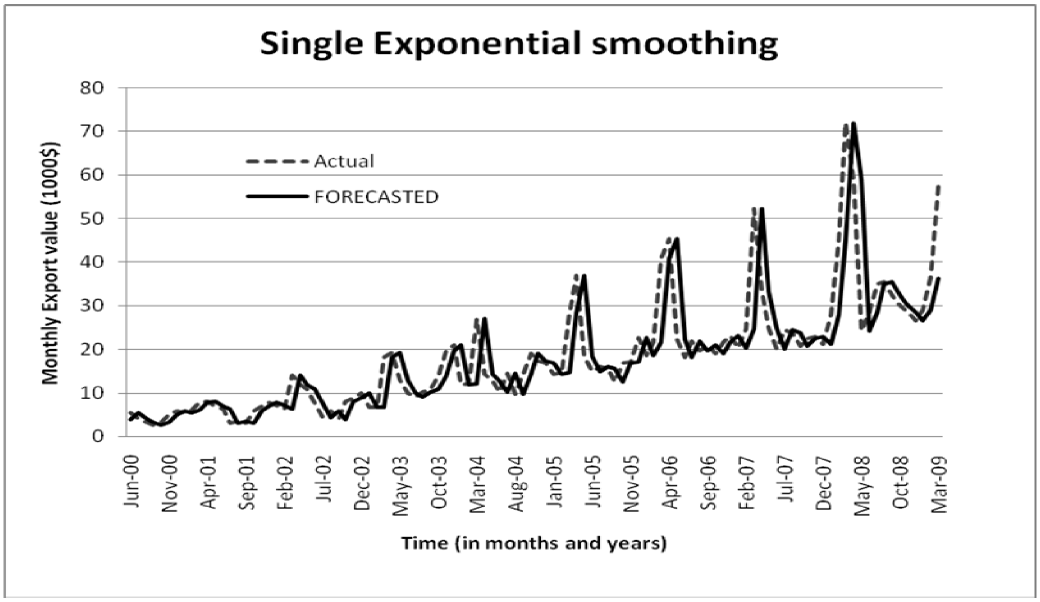


Figure: 4.6 Showing actual and predicted values Single Exponential smoothing

Figure: 4.6 Showing actual and predicted values Single Exponential smoothing

4.3.2.2 DOUBLE EXPONENTIAL SMOOTHING:

Table 4.15 Double Exponential Smoothing constants alpha (level) and Gamma (trend)

SMOOTHING CONSTANT	
Alpha(level)	0.9
Gamma(trend)	0.1

Table 4.16 Selected measures of predictive performance of Double Exponential Smoothing Method

ACCURACY MEASURES	
MAPE	27.9796
MSE	5.085

MAPE = Mean Absolute Percentage error
MSE = Mean Square Error

Table 4.17 Actual and Forecasted values of Double Exponential Smoothing for Monthly Fruits export

Sl.No	Time	Actual	Forecasted	Sl.No	Time	Actual	Forecasted
1	Jun-00	5.4	1.02	54	Nov-04	17.4	19.03
2	Jul-00	4.2	5.69	55	Dec-04	17	17.95
3	Aug-00	3.4	4.94	56	Jan-05	14.3	17.39
4	Sep-00	2.7	4.01	57	Feb-05	14.8	14.63
5	Oct-00	3.3	3.17	58	Mar-05	28.4	14.82
6	Nov-00	5.1	3.63	59	Apr-05	37	28.30
7	Dec-00	6	5.43	60	May-05	18.5	38.17
8	Jan-01	5.5	6.47	61	Jun-05	15	20.74
9	Feb-01	6.1	6.04	62	Jul-05	16.1	15.33
10	Mar-01	7.9	6.54	63	Aug-05	15.7	15.85
11	Apr-01	8	8.33	64	Sep-05	12.6	15.52
12	May-01	7	8.57	65	Oct-05	16.9	12.44
13	Jun-01	6.3	7.56	66	Nov-05	17.1	16.40
14	Jul-01	3.1	6.71	67	Dec-05	22.8	17.04
15	Aug-01	3.5	3.42	68	Jan-06	18.6	22.75
16	Sep-01	3.2	3.46	69	Feb-06	21.8	19.17
17	Oct-01	6	3.17	70	Mar-06	41	21.93
18	Nov-01	7	5.92	71	Apr-06	45.5	41.20
19	Dec-01	7.9	7.19	72	May-06	22.4	47.57
20	Jan-02	7.2	8.19	73	Jun-06	18.2	25.15
21	Feb-02	6.4	7.57	74	Jul-06	21.9	18.50
22	Mar-02	14.2	6.68	75	Aug-06	19.7	21.47
23	Apr-02	12	14.29	76	Sep-06	21.1	19.63
24	May-02	11	12.87	77	Oct-06	19	20.84
25	Jun-02	7.6	11.66	78	Nov-06	21.8	18.90
26	Jul-02	4.5	8.11	79	Dec-06	23.3	21.49
27	Aug-02	6	4.64	80	Jan-07	20.5	23.26
28	Sep-02	4	5.77	81	Feb-07	24.8	20.67
29	Oct-02	8	3.92	82	Mar-07	52.3	24.65
30	Nov-02	9	7.70	83	Apr-07	33.3	52.29
31	Dec-02	10	9.10	84	May-07	25	36.24
32	Jan-03	6.9	10.22	85	Jun-07	20.2	26.16
33	Feb-03	6.8	7.24	86	Jul-07	24.4	20.29
34	Mar-03	18.2	6.81	87	Aug-07	23.9	23.86
35	Apr-03	19.3	18.06	88	Sep-07	20.9	23.77
36	May-03	13	20.28	89	Oct-07	22.6	20.80
37	Jun-03	10	14.18	90	Nov-07	23	22.19
38	Jul-03	9.2	10.49	91	Dec-07	21.3	22.77
39	Aug-03	10.2	9.29	92	Jan-08	27.9	21.16
40	Sep-03	10.9	10.15	93	Feb-08	45.2	27.55
41	Oct-03	14	10.93	94	Mar-08	72	45.35
42	Nov-03	19.7	14.08	95	Apr-08	59	73.64
43	Dec-03	21	20.03	96	May-08	24.3	63.46
44	Jan-04	11.9	21.88	97	Jun-08	28.7	27.68
45	Feb-04	12.1	12.98	98	Jul-08	35.1	28.16
46	Mar-04	27	12.19	99	Aug-08	35.5	34.59
47	Apr-04	14.3	26.85	100	Sep-08	32.6	35.67
48	May-04	12.8	15.76	101	Oct-08	30.1	32.90
49	Jun-04	10.3	13.03	102	Nov-08	28.7	30.12
50	Jul-04	14.6	10.26	103	Dec-08	26.6	28.45
51	Aug-04	9.9	14.25	104	Jan-09	29	26.23
52	Sep-04	14.4	10.03	105	Feb-09	36.4	28.42
53	Oct-04	19	14.05	106	Mar-09	57.6	36.01

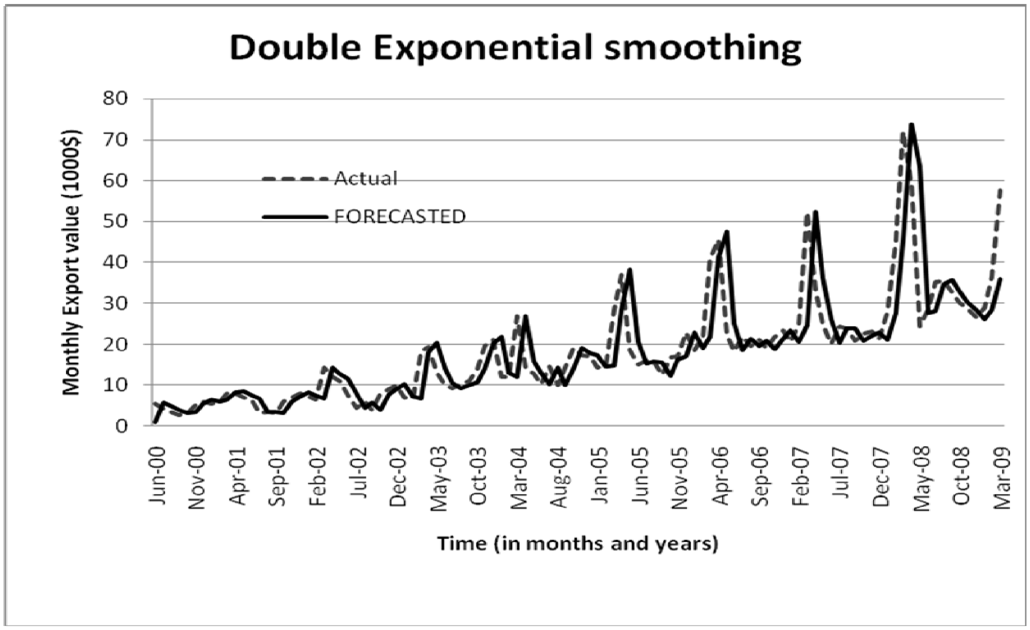


Figure: 4.7 Showing actual and predicted values Double Exponential smoothing

Figure: 4.7 Showing actual and predicted values Double Exponential smoothing

4.3.2.3 WINTERS MULTIPLICATIVE MODEL (TRIPLE EXPONENTIAL SMOOTHING)

Table 4.18 Smoothing constants of Winters' Method

SMOOTHING CONSTANT		
CONSTANTS	MULTIPLICATIVE	ADDITIVE
Alpha(level)	0.9	0.8
Gamma(trend)	0.1	0.1
Delta (season)	0.1	0.2

Table 4.19 Selected measures of predictive performance of Winters method

ACCURACY MEASURES		
	MULTIPLICATIVE	ADDITIVE
MAPE	26.3678	36.596
MSE	4.874642	5.2149

MAPE = Mean Absolute Percentage error
MSE = Mean Square Error

Table 4.20 Actual and Forecasted values of Winters method for Monthly Fruits Export

Sl.No	Time	Actual	Forecasted	Sl.No	Time	Actual	Forecasted
1	Jun-00	5.4	4.62	54	Nov-04	17.4	25.85
2	Jul-00	4.2	6.35	55	Dec-04	17	16.91
3	Aug-00	3.4	2.89	56	Jan-05	14.3	13.39
4	Sep-00	2.7	1.82	57	Feb-05	14.8	14.99
5	Oct-00	3.3	2.01	58	Mar-05	28.4	20.03
6	Nov-00	5.1	3.50	59	Apr-05	37	26.38
7	Dec-00	6	3.80	60	May-05	18.5	29.56
8	Jan-01	5.5	4.16	61	Jun-05	15	20.87
9	Feb-01	6.1	5.42	62	Jul-05	16.1	20.82
10	Mar-01	7.9	8.02	63	Aug-05	15.7	14.87
11	Apr-01	8	6.89	64	Sep-05	12.6	11.83
12	May-01	7	6.04	65	Oct-05	16.9	12.83
13	Jun-01	6.3	7.26	66	Nov-05	17.1	22.30
14	Jul-01	3.1	8.62	67	Dec-05	22.8	16.08
15	Aug-01	3.5	2.73	68	Jan-06	18.6	17.43
16	Sep-01	3.2	2.28	69	Feb-06	21.8	19.68
17	Oct-01	6	2.95	70	Mar-06	41	29.55
18	Nov-01	7	7.63	71	Apr-06	45.5	38.84
19	Dec-01	7.9	6.37	72	May-06	22.4	36.43
20	Jan-02	7.2	6.06	73	Jun-06	18.2	25.37
21	Feb-02	6.4	7.59	74	Jul-06	21.9	25.10
22	Mar-02	14.2	8.78	75	Aug-06	19.7	20.27
23	Apr-02	12	13.02	76	Sep-06	21.1	14.81
24	May-02	11	9.72	77	Oct-06	19	21.63
25	Jun-02	7.6	11.83	78	Nov-06	21.8	25.76
26	Jul-02	4.5	10.89	79	Dec-06	23.3	20.33
27	Aug-02	6	4.25	80	Jan-07	20.5	17.65
28	Sep-02	4	4.32	81	Feb-07	24.8	21.25
29	Oct-02	8	3.93	82	Mar-07	52.3	33.39
30	Nov-02	9	10.29	83	Apr-07	33.3	49.35
31	Dec-02	10	8.49	84	May-07	25	27.49
32	Jan-03	6.9	7.76	85	Jun-07	20.2	26.69
33	Feb-03	6.8	7.36	86	Jul-07	24.4	27.71
34	Mar-03	18.2	9.10	87	Aug-07	23.9	22.42
35	Apr-03	19.3	16.83	88	Sep-07	20.9	18.08
36	May-03	13	15.50	89	Oct-07	22.6	21.39
37	Jun-03	10	14.43	90	Nov-07	23	30.29
38	Jul-03	9.2	14.16	91	Dec-07	21.3	21.53
39	Aug-03	10.2	8.76	92	Jan-08	27.9	16.08
40	Sep-03	10.9	7.69	93	Feb-08	45.2	28.41
41	Oct-03	14	11.28	94	Mar-08	72	61.16
42	Nov-03	19.7	18.96	95	Apr-08	59	69.68
43	Dec-03	21	18.95	96	May-08	24.3	48.74
44	Jan-04	11.9	16.71	97	Jun-08	28.7	27.62
45	Feb-04	12.1	13.14	98	Jul-08	35.1	38.13
46	Mar-04	27	16.44	99	Aug-08	35.5	32.58
47	Apr-04	14.3	25.25	100	Sep-08	32.6	27.13
48	May-04	12.8	11.88	101	Oct-08	30.1	34.06
49	Jun-04	10.3	13.44	102	Nov-08	28.7	40.84
50	Jul-04	14.6	14.02	103	Dec-08	26.6	26.97
51	Aug-04	9.9	13.28	104	Jan-09	29	19.88
52	Sep-04	14.4	7.60	105	Feb-09	36.4	29.48
53	Oct-04	19	14.59	106	Mar-09	57.6	48.50

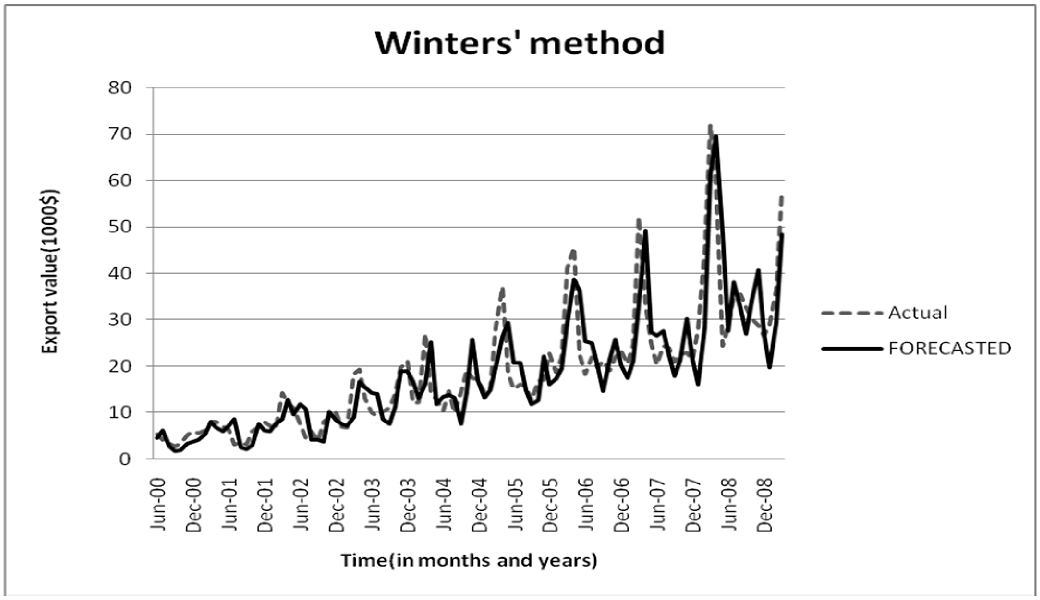


Figure: 4.8 showing actual and predicted values Winters' multiplicative method

Figure: 4.8 Showing actual and predicted values Winters' multiplicative method

Table 4.21 Comparison of models based on MAPE and MSE values and Selection of the best model

Sl. No	Model	MAPE	MSE
1	Linear	31.29	5.321
2	Quadratic polynomial	28.544	5.050
3	Logarithm	56.9	6.655
4	Power	30.467	5.016
5	Compound	25.878	4.922
6	S-Curve	60.47	8.461
7	Logistic	26.77	4.998
8	Growth	25.878	4.920
9	Exponential	25.878	4.920
10	Gompertz	94.204	9.565
11	ARIMA	15.198	3.360 (I)
12	Single Exponential smoothing	25.554	4.309 (II)
13	Double Exponential smoothing	27.98	5.085
14	Winters method(Triple)	26.368	4.875

Note: Abbreviations in the parenthesis
I : First best model
II : Second best model

5. DISCUSSION

The present study is subjected to discussing some of the statistical analysis .The results presented in chapter IV are discussed in this chapter.

5.1 Polynomial Function fitting approach.

5.2 Fitting of Non-linear regression models.

5.3 Fitting ARIMA and Exponential Smoothing.

5.4 Comparison of all the above models.

5.1 Polynomial Function fitting approach.

Over a long period of time, the time series is very likely to show the tendency to increase or decrease overtime. The polynomial function fitting was carried out to know the tendency of the present export data. Polynomials with different degrees were carried out in the process. And the selection of the best model was based on highest R-square value which explains the variation in dependent variable due to independent variable. As a process of this the tables 4.1, 4.2,4.3 and 4.4 showing the polynomials of different degrees were presented where table 4.1 represents annual fruit export, table 4.2 represents Mango export , table 4.3 represents Banana export and table 4.4 representing the monthly export data of fruit crops.

In this process different degrees of polynomial were tried to fit for which the 5P^{thP} degree polynomial has shown stable with the highest R-square of 0.994 (table: 4.1). This was followed by 6P^{thP} and 4P^{thP} degree polynomials with 0.992 and 0.989 R-square values for annual fruit export. The polynomials of these two degrees also showed near stability. Nearly 99.40 percent of the variation was explained in the 5P^{thP} degree polynomial compared to other polynomials with respect to the export data, hence this was considered as best of all the polynomials carried out. In case of Mangos export 6P^{thP} degree polynomial attained stability with highest R-square of 0.940.This was best fit and it is followed by 7P^{thP} and 5P^{thP} degree polynomials with R-square values 0.938 and 0.931for stability. 94 percent of the variation was explained in this polynomial compared to other polynomials with respect to the export data. Hence this was considered as best. And when we look at the Banana export, the 6P^{thP} degree polynomial has shown the best fit with highest R-square value of 0.860 by attaining stability at this point of the polynomial. This was followed by 7P^{thP} and 5P^{thP} degree polynomials with R-square values of 0.852 and 0.828 respectively. 86 percent of the variation was explained in the 6P^{thP} degree polynomial compared to other polynomials with respect to the Banana export data, hence this was considered as best. And in case of monthly fruits export it showed that the 2P^{ndP} degree (Quadratic) polynomial was best fit with an highest R-square value of 0.6255 which was followed by 3P^{rdP} degree with R-square of 0.6252. Nearly 62.30percent of the variation was explained in this polynomial compared to other polynomials with respect to the export data, hence this was considered as best. Hence the polynomial model represented for the export were shown sufficient to describe for the following data pattern, this seemed quite reasonable. This might be because of the erratic behavior of the export data.

Similar work has been done by Teresa Jacobson et., al in 2009 in case of Using Linear Regression of Seasonal Weather Patterns to Enhance Undergraduate Learning Based on the students' knowledge of polynomials. It was estimated that a model of at least degree seven would be necessary to satisfactorily describe the pattern. The sixth order polynomial was used to compare the erratic pattern of the maximum standard deviations to the corresponding pattern of the best fitting polynomial for the minimum standard deviations. The R-square value was found to be increased from 0.0209 to 0.885 including the eighth degree polynomial. Similar work was carried out by Therthappa in 2005 on yield estimation of sunflower with yield contributing character who observed that cubic model was found best in predicting seed yield based on the head diameter and disease at first stage.

5.2 Fitting of Non-linear regression models.

Models were built in order to fit and predict the export of fruits with help of previous year's export data. As a process of this nearly nine models were tried. Among all the models the best model was selected based on the highest R-square value and least MSE value. All models were found to be significant (at 1percent and 5percent).

In case of annual fruit export, among all the models the Compound, Growth and Exponential models were best fit with highest R-square of 0.9413 and least MSE of 0.04. It means that these were best for predicting the trend in the annual fruits export. The three models resulted so close to each other that there is no real choice between them and it always produces a lower residual mean square error. The R-square has explained nearly 95percent of the variation in annual fruits export was due to time factor. Intern this was influenced by the climatic factors of that particular period like rainfall, temperature, relative humidity, wind velocity etc, quality aspects like shelf life, shape, size, colour etc, farming practices, lack of proper and efficient transport facilities, poor quality of products, absence of modern and scientific storage facilities as well as packing materials and many other factors lead to the major hindrance for export growth. Indian horticultural products are also often out priced in the world market. And this was followed by the Linear, Logistic and Power models with R-square values 0.8256, 0.7662, 0.7424 and MSE values 29.84, 0.77,0.16 respectively by showing close stability of the data. And Gompertz (0.620 and 17956), Logarithmic (0.5530 and 7515.16) and S-curve (0.4264 and 0.36) models didn't show its best with fit of models proving the unstable condition for that particular export data.

When we glance through Mango Exports, it was found that the Linear model has shown best fit with the highest R-square value of 0.9876 and least MSE of 0.15 among all the models. By this we can concluded that, the stagnancy in Mango production was almost increasing over the years. This might be because of international market price. The Linear model was best fit which explained nearly 99percent of the variation in Mango export was due to time factor. And this was followed by Growth, Exponential and Compound models with R-square value of 0.9310 and MSE of 0.03. These intern followed by Logistic (0.8220 and 0.39), Power (0.8188 and 0.83), Gompertz (0.7410 and 1573.71), Logarithmic (0.6670 and 499.16) and S-curve (0.4432 and 0.25) models respectively. Linear model was found best to forecast the trend in Mango Export.

The same process was examined for Banana exports also. The Compound, Growth and Exponential models have shown best fit with highest R-square value of 0.8010 and least MSE of 0.71 among all the models. The three models proved that no other models are so close and so best in explaining the accuracy of the Banana export data. This showed 80percent of the variation in Banana export was due to time factor. This was influenced by many constraints in that particular period. These models were followed by Linear (0.7467 and 118.81), Logistic (0.7455 and 2.21), Power (0.6881 and 1.12), Logarithmic (0.5374 and 24.4), Gompertz (0.4910 and 4278.73) and S-curve (0.3491 and 2.34).

This might be because over the years the production has increased, the area under the fruit crops is also increased and the productivity has also shown positive. All these in turn have increased the exports and also the modernization of farming practices has increased the quality of the products.

Similar work was carried by Chetana (2009) on prediction models in teak based agroforestry systems in northern transitional zone of Karnataka**.** Different prediction models namely Linear, Quadratic, Cubic, Exponential, Growth, Sigmoid, vapor pressure, Hassel, MMF compound, Logarithmic, Logistic, Weibull, Gompertz, Power etc., have been tried to predict the diameter and height growth of teak tree. MMF model was found better followed by Gompertz and Weibull for diameter prediction whereas for height growth prediction MMF model followed by logistic and Richards were found to be better. Similar study was carried out by Swetha in 2009 on the Impact of rain water harvesting on farming economy who used above all models and observed that linear model was showing significant and best model with R-square of 0.261 as compared to all other models which indicated that 26.1percent of the variation in ground water level is accounted by temperature.

5.3 Fitting ARIMA and Exponential Smoothing.

5.3.1.1Box-Jenkins ARIMA Model

As explained earlier (Chapter III), fitting Box-Jenkins models, the other name of ARIMA models, involves a four stage procedure. The discussion is presented in the same order.

5.3.1.2 Identification of the model

Identification of the model is the first step which involves a greater deal of skill. It is done based on conjunction of the sample Auto Correlation Function with the Partial Auto Correlation Function (PACF). ACF and PACF are presented in Table 4.8. Since the method of identification does not lay down any hard and fast principles, several possible models are tentatively identified and the best results are presented below.

5.3.1.3 Estimation

Having tentatively identified the model, next the parameters which minimize the sum of squares of errors are estimated. The estimated model for monthly fruit crops export is: (1,1,1) (1,1,1)

5.3.1.4 Diagnostic checking

The residuals of estimated models are examined for testing the randomness of series and analyzed to determine the adequacy of the estimated models.

For all the series of monthly fruit crops export, SBC and AIC is minimum. Seasonality is found and forecast consideration is the best. So this model is chosen.

5.3.1.5 Forecasting

Ex-ante and Ex-post obtained by the Box-Jenkins method is presented in Table 4.11. The forecasts from the various models are checked for their efficacy by comparing them with the actual values.

Similar work was carried out by Venugopalan et.al., in 1996 for Trend analysis in all-India marine products export using statistical modeling techniques. They fitted ARIMA model to the monthly all-India marine export. They showed that ARIMA forecasts are quite near to the actual export figures during 1993-95. The marine products export is finally forecasted for next three years using the selected model. It is estimated that by 1997-98 marine products export in our country would be about 3, 14,451 tonnes.

The similar model (Box-Jenkins) is used by Achoth (1985) analyzed the supply, price and trade of Indian tea by fitting ARIMA model to data on prices and production. The forecasts yielded reasonably good results as judged from the tests of their efficiency. The forecasts of prices are superior when compared to the forecasts of quantities, which is attributed to the highly structured pattern of price behaviour.

5.3.2 Exponential Smoothing Methods

In these methods for updating the forecast in light of changing parameters of a regression model is shown. There are several ways to update the parameters: One is to re-estimate them every time a new data point is obtained another is to use of moving average technique on the series before attempting to compute the parameters. By far the most common and efficient way of handling the problem is to use Exponential Smoothing technique. In this we discuss the methods of obtaining the correct weighting factors and building a prediction for the updated forecast. An Exponential Smoothing model is preferred to the multiplicative time series model for forecasting purposes. The Exponential smoothing is best model for short term forecasting than regression and moving average. The three methods of Exponential smoothing techniques are carried out for monthly fruit export data using MINITAB SOFTWARE.

5.3.2.1 Single Exponential Smoothing

Single Exponential Smoothing method is also called as Method of estimation of forecasts of single weight / parameter. In order to smooth a set of data correctly, we must first obtain the proper weighting factor. In theory this weighting factor alpha can range from 0.01 to 1, but it has been used that any estimated value of alpha that is greater than 0.3 indicates that the error terms are not random. However, recent research has shown that the values greater than 0.3 are acceptable (Gander, 1985)

Once it has established that the time series variable is stationary, and then it is possible to apply the exponential smoothing method. The equal weight has been given to all the observations of the export data; the weight is assigned by trail and error method. The weight (WB_{1B}) i.e., the alpha (level) was found to be 0.9 which is obtained by trail and error method. The accuracy measures considered are MSE and MAPE tests. The values of MSE and MAPE are presented in Table 4.13, which are found to be least and stable. The model was selected based on the minimum MAPE value which was found to be 25.55 percent. The table 4.14 shows the actual and forecasted fit values of the model using single exponential smoothing constant.

Similar research work was carried out by Vasanth Kumar in 2002 on Statistical Evaluation of price variation in tropical timbers. In this Single Exponential Smoothing was used for forecasting the price of timber in different depots.

5.3.2.2 Double Exponential Smoothing Method

In previous section we discussed modeling with single exponential smoothing method, if time series contains linear regression, the correct procedure would be double exponential smoothing method. The argument and techniques for the double exponential smoothing method are similar in nature to that of single exponential smoothing.

Double exponential smoothing method is also called as Trend adjusted exponential smoothing or Brown's double exponential smoothing. This method is improved by the introduction of a deterministic second smoothing constant i.e., trend (WB_{2B}); assuming monthly export value is influenced by the trend component which could be useful. This is chosen in conjunction with W1. Hence during the process constants obtained by trail and error method are 0.9 (level) and 0.1(trend) which are presented in table 4.15. The accuracy measures MAPE and MSE were presented in table 4.16 which are found to be least. The model was selected based on the minimum MAPE value which was found to be 27.97 percent. The table 4.17 shows the actual and forecasted fit values of the model using double exponential smoothing constants.

Similar work was carried out by Weldon in 1995 for application of double exponential smoothing to interlibrary loan forecasting. The difficulties of forecasting interlibrary loan and document delivery acquisition times are briefly discussed using double exponential smoothing technique are explained.

5.3.2.3 Winters' Seasonal Exponential Smoothing Method

. The Holt-Winters exponential smoothing is used when the data exhibits both trend and seasonality. The two main HW models are additive model for time series exhibiting additive seasonality and multiplicative model for time series exhibiting multiplicative seasonality. This method is efficient when the seasonal pattern is constant year after year and the necessary computation for updating the new data are not a problem.

The Holt-Winters method can be extended to accommodate additive seasonality if the magnitudes of the seasonal effects do not change with the series or multiplicative seasonality if the amplitude of the seasonal pattern changes over time. To handle seasonality present in the time series data, the third parameter (WB_{3B}) is added. For the determination of values of the softening constants WB_{1B}, WB_{2B} and WB_{3B}, trail and error method was carried out based on the minimization of the mean absolute percentile error (MAPE) of the forecasting.

The tables 4.18 and 4.19 show the MAPE (26.3678) which is found least (W1 = 0.9, W2 = 0.1 and W3 = 0.1) for the Multiplicative model and for Additive model (WB_{1B} = 0.8, WB_{2B} = 0.1 and WB_{3B} = 0.2), MAPE was 36.59 percent. Hence the multiplicative model outperformed the additive model. Therefore for the multiplicative model considering the whole period (from june 2000 to march 2009), the found MSE is 4.879 and the mean absolute percentile error – MAPE was 26.36 percent. For the additive model the values are 5.2149 and 36.596 percent respectively. Hence multiplicative model was selected as best. The table 4.20 shows the actual and forecasted fit values of the Winters multiplicative model using triple exponential smoothing constants.

Similar work was carried out by Prajakta in 2004 on Time series Forecasting using Holt-Winters Exponential Smoothing. The tests were conducted on the following time series: d1, ebrahim, rose, red-wine and paper. The multiplicative model gave better results as compared to the additive model for all the series.

When we look at the table 4.21 the MAPE and MSE values are close to each other in all the smoothing techniques. In single exponential smoothing the MAPE was found to be (25.56percent and MSE 4.809) very close to double exponential smoothing (MAPE 27.98percent and MSE 5.085). Which shows that after having the second smoothing constant i.e., trend component the accuracy has shown slight downfall. And this was tried with the triple exponential smoothing (Winters' method: MAPE 26.368percent and MSE 4.875) even in this to it didn't show as accurate as single exponential smoothing after having three smoothing constants having season as the third component. This shows that single exponential smoothing method as outperformed well and is the best to forecast the trends to the monthly fruit exports among all the three methods of smoothing.

5.4 Comparison of all the above models and Selection of the best model

In this chapter we give large scale comparison of different models in order to know the best model for the forecasting of the exports. For the model comparison, monthly export of fruits was considered. The comparison of all the 14 models was carried out in the process based on the MAPE and MSE values which were considered to be least. According to the table 4.21 which represents the MAPE and MSE values of the monthly fruit crop export, the ARIMA model with least MAPE value of 15.198 percent and MSE value of 3.360 was considered as best fit among all the models considered. Box-Jenkins ARIMA model outperformed all the models. And the model next to it was found to be Single exponential smoothing with MAPE value 25.554 percent and MSE value 4.309. The other models which showed next best are Exponential, Compound and Growth models with MAPE and MSE values 25.878 percent and 4.920 respectively. The other best were Winters method (MAPE=26.368 percent and MSE=4.875) and Logistic model (MAPE=26.77 percent and MSE=4.998) followed by Double exponential smoothing (MAPE=27.98 percent and MSE=5.085), Quadratic polynomial (MAPE=28.544 percent and MSE=5.050), Power (MAPE=30.467 percent and MSE= 5.016), Linear (MAPE=31.29 percent and MSE=5.321), Logarithmic (MAPE=56.9 percent and MSE=6.655), S-curve (MAPE=60.47 percent and MSE=8.461) and at last the Gompertz model (MAPE=94.204 percent and MSE=9.565).

This in turn proved that the Box-Jenkins ARIMA model is the best model in forecasting the trend for monthly fruit exports which has least MAPE and MSE values. And this model was concentrated on obtaining forecasts from time series methods of monthly fruit crop export from the own current and past values.

Christine Lim *et al.*, worked on forecasting quarterly tourism demand by Hong Kong for Australia. Various forecasting models are estimated over the period 1975(1)-1989(4) for seasonally unadjusted quarterly tourist arrivals to Australia from Hong Kong. As the best fitting ARIMA model out performed all the exponential smoothing methods. The time series models are used to obtain the post-sample forecasts, with RME being used to evaluate their accuracy.

Henry worked on Tforecasting Thailand's rice export: statistical techniques vs. artificial neural networks in 2007T. This paper compares the performance of artificial neural networks (ANNs) with exponential smoothing and ARIMA models in forecasting rice exports from Thailand. To ascertain that the models can reproduce acceptable results on unseen future, we evaluated various aggregate measures of forecast error (MAE, MSE, MAPE, and RMSE) during the validation process of the models. The results reveal that while the Holt-Winters and the Box-Jenkins models showed satisfactory goodness of fit, the models did not perform as well in predicting unseen data during validation. On the other hand, the ANNs performed relatively well as they were able to track the dynamic non-linear trend and seasonality, and the interactions between them.

6. SUMMARY AND CONCLUSION

In this Chapter, the summary and conclusion of the study based on the findings are presented below

The year to year fluctuations in crop production is quite common in Indian Agriculture. These fluctuations adversely affect the production, employment and income, thereby hampering the economic growth of the country. The Export of fruits in India have recorded increase during certain years and also decrease during some other years. The growth achieved is different in different time period and regions. India is bestowed with a varied agro-climate, which is highly favourable for growing a large number of horticultural crops such as fruits, vegetables, root tuber, ornamental, aromatic plants, medicinal herbs, spices and plantation crops. India has a large range of varieties of fruit in its basket and accounts for 10% of worlds total fruit production. Mango, Banana, Citrus, Pineapple, Papaya, Guava, Sapota, Jackfruit, Litchi and Grape, among the tropical and sub-tropical fruits Apple, Pear, Peach, Plum, Apricot, Almond and Walnut among the temperate fruits and Aonla, Ber, Pomegranate, Annona, Fig, Phalsa among the arid zone fruits are important. India is the largest producer of Mango, Banana, Sapota and Lime. About 39% of world's Mango and 23% of world's Banana is produced in the country.

The secondary data on Export of fruit crops was collected from below mentioned sources based on the availability. Annual Fruit crops Export data was collected from the CMIE, Foreign Exchange Reviews from 1994 to 2009 (in 1000$US). Annual Mango Export data was collected from the APEDA DATABASE from 1985 to 2009 (in crores). Annual Banana Export data was collected from FAO STAT DATABASE from 1985 to 2008(in 1000$US). Monthly fruit crop Export data was collected from CMIE, Economic Intelligence Service Monthly reviews from 2000(June) to 2009(March) (in 1000$US).

In the present study an attempt has been made for the study of trends in Export of Fruits.

Not much work has been carried out in statistical field to forecast the export data. Hence this research aim was to develop a suitable statistical model for describing the trend by employing different statistical modeling. Due to this reason the present study was designed to fit different models to fruit export data. At last the best one is recommended for the further study. As the aspect of this following approaches are conducted.

a) Polynomial function fitting
b) Non-linear regression models
c) ARIMA and Exponential smoothing methods

The polynomial function fitting was carried out with different degrees and the selection of the best model was based on R-square value. Different degrees were tried to fit among all the degrees carried out $5P^{thP}$ degree polynomial showed best fit to the Annual fruit Export, $6P^{thP}$ degree polynomial showed best to the Mango Export, $6P^{thP}$ degree polynomial showed best to the Banana export and Quadratic polynomial ($2P^{ndP}$ degree) showed the best to the Monthly Export of fruit crops.

Nine Non-linear models were tried among all the models the best model was selected based on the highest R-square value and least MSE value. All models were found to be significant (at 1% and 5%). These models were fitted to Annual fruit Export, Mango and Banana Export. In the process it showed that the Compound, Growth and Exponential models were best fit with highest R-square and least MSE values. But only in case of Mango Export it showed Linear model was best fit and the next best was found to be the same models. So these three models are best to find the growth in Export or to find trends for Exports.

The Box-Jenkins ARIMA model is fitted to the monthly fruit exports. Seasonality was found in the data so seasonal ARIMA model is used. Before going to application of Box-Jenkins analysis, more the data should be stationary series. If the series is non-stationary, it could be removed by differencing. So, the differencing was done to remove non-stationarity and made stationary. Making use of differenced series (which is stationary), the ACF and PACF are computed because it helps in tentatively identify the model. Then the parameters of all tentatively identified model was estimated by iterative process. The AIC and SBC values of least were considered to identify the model. The model (1,1,1 : 1,1,1) was identified and subjected to diagnosis checking in order to determine the adequacy of model. The residues of estimated model is examined for testing the randomness of series and for its significance. Both Ex-ante and Ex-post forecast is done for the identified model.

Exponential Smoothing methods are fitted to the monthly fruit exports. Trail and error method was used in identifying the smoothing constants. The equal weight was given to all the observations. In single Exponential smoothing, the weight (WB_{1B}) i.e., the alpha (level) was found to be 0.9. The accuracy measures were tested using MSE and MAPE tests. Double Exponential smoothing method is used which is improved by the introduction of a second smoothing constant i.e., trend, (WB_{2B}); assuming monthly export value is influenced by the trend component. This must be chosen in conjunction with W1. So the constants obtained by trail and error method are 0.9 (level) and 0.1(trend). Winters seasonal exponential method was tried to handle seasonality present in the time series data, the third parameter (WB_{3B}) is added. For the determination of values of the softening constants WB_{1B}, WB_{2B} and WB_{3B}, trail and error method was done. Winters Multiplicative model showed best among Additive and Multiplicative model. The smoothing constants are $WB_{1B} = 0.9$, $WB_{2B} = 0.1$ and $WB_{3B} = 0.1$ for the Multiplicative model.

In all the smoothing methods, when we look at the MAPE and MSE values they were found to be close to each other. This showed that even after having trend and season smoothing constants the accuracy measures didn't change. In turn proved that single smoothing constant was enough to forecast the Monthly fruit export. Hence Single Exponential is considered best to forecast the trends among the Exponential smoothing methods.

In selecting the best model to forecast the trend for monthly fruit export, the accuracy measures MAPE and MSE are considered. Among all the models tried, the Box-Jenkins ARIMA model was best fit with least MAPE (15.198 percent) and MSE (3.36) values. And the forecasted values from the ARIMA model were much nearer to the Actual values.

6.1 Special Features of Investigation:

Hitting a 13-year low, the country's total exports dropped by 21.7 per cent in February, 2009 to Rs 58, 685 crore, as compared to Rs 60,476 crore in the corresponding month last year (2008). The drop is being attributed to a continued lack of demand for goods in the West. Having been an agro-based economy, Indian trade has always been devoid of manufactured or industrial goods. Post liberalization, imports dominated the Indian trade scene in the form of heavy machinery and information technology products and, thus, created an imbalance of trade.

Horticultural exports have great potentials for developing countries foreign earnings due to the following reasons. Firstly, horticultural products have a high income elasticity of demand. As income goes up, demand raised rapidly, especially in middle and high- income developing countries, although demand is also rising in developed countries. A growing concern for health and nutrition has caused consumer preferences to shift from high fat, high cholesterol foods, such as meat and livestock products, to low-fat, low –cholesterol foods, such as fish, fruits and vegetables. Horticultural crops provide a better alternative for diversification of agriculture in view of high returns available from them. Horticulture sector helps in improving productivity of land, generating employment, improving economic condition of farmers and entrepreneurs, enhancing exports and foreign exchange earnings and above all, providing nutritional security of the people.

Traditional exports such as agricultural raw materials and beverages face slow growth prospects in world market, and food exports such as cereals, livestock products, and sugar cannot compete with the subsidized exports of developed countries.

Statistical modeling can be applied to forecast the future export values, as there is no much research studies in this area. Such statistical studies will be of immense help in formulating suitable strategies for planning and developing future export of fresh fruit crops.

6.2 Future Line of Work:

The present study was limited to only few models. Therefore, it is suggested as mentioned below:

- Other non linear regression models (MMF, Richards, Wiebull etc) can also be used.
- Other forecasting models like ANN, ARCH and GARCH Models can also be used.
- Statistical modeling can be applied to forecast the export of vegetables, plantation crops, oilseed crops etc.

REFERENCES

Achoth, L., 1985, Supply price and trend of Indian tea-An econometric analysis. *Ph.D. Thesis*, Uni.Agric.Sci. Bangalore (India).

Adhikari,M., Paudel,.K., Paudel, L and Bukenya, J., 2007, Modelling swine supply response using a structural time series approach, Applied Economics Letters, **14(7/9) :**467-472.

Ali.M, A and Singh, A.K., 1995, Growth and fluctuation in area, production and productivity of wheat in Chhattisgarh region of Madhya Pradesh. *Agri. Situ. India*, **59(9)**: 608-614.

Alon, M. Qi., and Sadowski, R. J., 2001, Forecasting aggregate retail sales: a comparison of artificial neural networks and traditional methods, *Journal of Retailing and Consumer Services*, **8**, 147–156.

Arun Katyayan., 2001, *Fundamentals of agriculture. Vol-.1.* Kushal Publications and Distributors, Varanasi.

Avila-Curiel, A., Shamah, T., Barragan, L., Chavez, A., Avila, M. A and Juarez, L., 2004, An epidemiological index to assess the nutritional status of children under five years of age in Mexico, *Archivos Latinoamericanos-de-Nutricion*, **54(1)**: 50-57.

Balasubramanian, P. P and Rema, M., 1996, Pricing and Transaction Trend of Raw cashew nut in India, *The Cashew*, 10(4); 13-19.

Belov, M. I and Chepurnoi, A. I., 1985, Promising chopper mechanisms for forage harvesters, *Traktory-i-Sel'-khozmashiny*, **(5)**: 39-41.

Bergamasco, A. F., Aquino, L. H-de, and Muniz, J. A., 2001, Non-linear models fitting to growth data in females of the Holstein heifers, *Ciencia-e-Agrotecnologia*, **25(2)**: 235-241.

Bharathi, R., 2009, Price behaviour of Mulberry silk cocoon in Ramnagar and Siddlaghatta market – A statistical analysis, *M. Sc Thesis*, Uni. Agril. Sci., Dharwad.

Chandrakala, S. Kallimath., 2009, Spatial temporal behaviour of arrivals and prices of groundnut in Karnataka- An economic analysis, *M. Sc Thesis*, Uni. Agril. Sci., Dharwad.

Chatfield, C and Protharo, D.L., 1973, Box and Jenkins seasonal forecasting problems in a case study. *J.Roy.Statist.Soc.A.* 136(3):295-336.

Chatfield, C., 1977 Recent Developments in Time series Analysis. *J.Roy.Statist.Soc.A*. 140(4):492-510.

Chen, R. J. C., Bloomfield, P and Cubbage, F. W., 2008, Comparing forecasting models in tourism, Journal of Hospitality and Tourism Research, **32(1)**: 3-21.

Chengappa, P. G., 1981, Growth rates of area, production and productivity of coffee, *Indian Jrnl. Of Coffee Research*., **11(2)**: 19-26.

Chengappa, P.G., 1980, An analysis of Indian coffee Prices - A Spectral approach. *Unpublished M.Sc.(Agri) Thesis*, IARI, New Delhi.

Chethana, H. M., 2009, Prediction Models in Teak based Agroforestry systems in Northern transitional zone of Karnataka, *M. Sc Thesis*, Uni. Agril. Sci., Dharwad.

Chu-FongLin., 2008, Analyzing and forecasting tourism demand with ARAR algorithm, *Tourism Management*, **29(6)**: 1185-1196.

Coelho, E. F., Conceicao, M. A. F and Souza, V. A. B., 1998, Estimating limits of soil water availability as a function of bulk density and soil texture, *Revista-Ceres*, **45(258)**: 183-192.

Cooray, T. M. J. A.,2008, Applied Time series Analysis and Forecasting, Narosa publication.

Crevier, N., Pourcelot, P., Denoix, J. M., Geiger, D., Bortolussi, C., Ribot, X and Sanaa, M., 1996, Segmental variations of in vitro mechanical properties in equine superficial digital flexor tendons, *American Journal of Veterinary Research*, **57(8)**: 1111-

1117.

Debertin, D. L., 1985, Developing realistic agricultural production functions for use in undergraduate classes, *Southern Journal of Agricultural Economics*, **17(2)**: 207-214.

Deluyker, H. A., Shumway, R. H and Weaver, L. D., 1987, Modeling of daily milk production in dairy cows using time series analysis, Journal of Dairy Science, USA, **70**(Suppl. 1): 227.

Devaiah, M.C, Venkatagiriyappa and Achoth Lalith, 1988, Spatial integration and price relationship of Ramnagaram over minor silk cocoon markets of Karnataka. *Proceedings of the international congress on tropical sericulture, Central Silk Board*, 77-84.

Devi, P. I., Thomas, E. K and Thomas, J. K., 1990, Growth and Supply response of Banana in Kerala. *Agric. Situ. India*, **45(4)**: 239.

Dursun Aydin., 2007, A Comparison of the Nonparametric Regression Models using Smoothing Spline and Kernel Regression, *World Academy of Science, Engineering and Technology* , **36**:253-257.

Fialar, S. C., 1985, An analysis production pattern and marketing of cocoa in Ghana, Unpuplished *M.Sc. Thesis*, Univ. Agric. Sci., Bangalore.

Gaddour, A and Najari, S., 2008, Adjustment of the kid's growth curve in pure goat breeds and crosses under southern Tunisian conditions, *Journal of Applied Animal Research*, **34(2)**, 117-120.

Gajendra Singh and Bhogal, T. S., 2006, Assessment of food security situation for disaster risk management: an analysis for the Gujarat State, *Anvesak*, **36(1)**: 32-44.

Gangadharappa, H., 2005, A Statistical study of variation in arrivals and prices of potato in the selected markets of Karnataka. *M.Sc. (Agri) Thesis,* Uni. Agric.Sci. Dharwad (India).

Gan-YanTai., Stobbe, E. H and Njue, C., 1996, Evaluation of selected nonlinear regression models in quantifying seedling emergence rate of spring wheat, *Crop-Science*, *Canada*, **36(1)**: 165-168.

Gomez., A. K., and Gomez., A. A., 1958, Statistical procedures for agricultural research. John Wiley and Sons, New York.

Govindan, A., 1974, Analysis of Price Behaviour and retail margins in selected food grain markets in India. An application of spectral analysis. *Unpublished Ph.D.Thesis,* IARI, New Delhi.

Granger, C.W.J. and Newbold, P., 1970, Spectral methods. *Econometrica.* 30: 424-438.

Gwathmey, J. K., Nakao, S., Come, P. C and Abelmann, W. H., 1989, Echocardiographic assessment of cardiac chamber size and functional performance in swine, *American Journal of Veterinary Research*, **50(2)**: 192-197.

Henry, C., Rujirek., and Boosarawongse., 2007, Forecasting Thailand's Rice Export, *Computers and Industrial Engineering*, **53**: 610-627.

Huertas Garcia, R and Rodriguez, B. M., 2007, Forecasting international tourist demand using Holt-Winters, *Papers-de-Turisme*, *Spain*, **41**: 75-90.T

Jahne, M., 1984, Growth curve in female cattle, Wissenschaftliche Zeitschrift der Karl Marx Universitat-Leipzig Mathematisch Naturwissenschaftliche Reihe, **33(3)**: 252-262.

Kaushik, K. K., 1993, Growth and instability of Oilseed production in India, *Indian Jrnl.Agric. Econ*., **48(3)**: 334-338.

Kim, O. W., Kim, B. S., Nahm, G. B., Jeong, J. W., Kim, D. C and Keum, D. H., 1995, The effect of air vent holes and stacking methods of fruit and vegetable boxes on static pressure drop in pressure cooling system, *Journal of the Korean Society for Agricultural Machinery*, **20(4)**: 360-367.

Konz, M., Uhlenbrook, S., Braun, L., Shrestha, A and Demuth, S., 2007, Implementation of a process-based catchment model in a poorly gauged, highly glacierized Himalayan headwater, *Hydrology and Earth System Sciences*, **11(4)**: 1323-1339.

Kumar, T. V., Bhat, A. R. S and Patil, S. J., 2005, Price forecasting of different classes of teak by the application of exponential smoothing model, *Karnataka Journal of Agricultural Sciences*, **18(1)**: 55-59.

Lanciotti, C., 1990, Forecasts of monthly prices for a group of dairy products – An application of the Box-Jenkins techniques. *Rivista-di-Economics-Agraria.* 45(3):465-492.

Leuthold, R.M., Maccomnick, Schmity and D.G.Waltis, 1970, Forecasting daily hog prices and quantities: a study of alternate forecasting technique. *J.Ame. Statist.Assoc.* 65(1):90-107.

Lim, C and McAleer, M., 2001, Forecasting tourist arrivals, *Annals-of-Tourism-Research, Australia*, **28(4)**: 965-977.

M. Sc..Thesis, Univ. Agric. Sci., Dharwad

Makridakis, S and Hibbon,M., 1979, Accuracy of forecasting: An empirical investigation. *J.Roy.Statist.Soc.A*. 41(2):97-145

Manurung, A., Chairani, M and Lubis, S., 1991, Forecasting of oil palm hectarage and the need for seed in the second long term development plan, *Buletin-Perkebunan, Indonesia*, **22(4)**: 221-230.

Mastny, V., 2001, The use of Box-Jenkins methodology to predict the price development of agricultural commodities. *Acta Unierisatatis Agriculturae et Silviculturae-Mendelianae-Brunensis*, 49(2):165-172.

Mohan, S and Arumugam, N., 1995, Forecasting weekly reference crop evapotranspiration series, *Hydrological Sciences Journal*, **40(6)**: 689-702.

Mohanty, M., Painuli, D. K., Misra, A. K., Bandyopadhyaya, K. K and Ghosh, P. K., 2006, Estimating impact of puddling, tillage and residue management on wheat (Triticum aestivum, L.) seedling emergence and growth in a rice-wheat system using nonlinear regression models, *Soil and Tillage Research*, **87(1)**: 119-130.

Munir, M and Farhat-Naz., 2006, Growth and flowering response of snapdragons after release from apical dominance, *Journal of Applied Horticulture Lucknow*, **8(1)**: 25-28.

TNageshwara Rao., G., 2008, *Statistics for agricultural sciences*. BS publications, Hyderabad.

Nesreen, K. Ahmed., Amir, F. Atiya and Neamat El Gayar., An Empirical Comparison of Machine Learning Models for Time Series Forecasting, *World Academy of Science, Engineering and Technology* ,1:31

Newbold, P and Granger, C.W.J., 1974, Experience with forecasting Univariate time series and the combination of forecasts. *J.Roy.Statist.Soc.Australia,* 137:131-165.

Nisha Varghese, 2004, Export of groundnut from India under liberalization regime- An economic analysis, *Ph.D Thesis*, Uni.Agric.Sci. Dharwad (India).

Palomares Salas, J. C., J.J. G. de la Rosa, Ramiro, J.D., Melgar, J., Agüera, A and Moreno, A., Comparison of Models for Wind Speed Forecasting, Research Unit PAIDI-TIC-168, *University of Cádiz*, Electronics Area, Escuela Politécnica Superior.

Prajakta, S. Kalekar.,2004, Time series Forecasting using Holt-Winters Exponential Smoothing, *Kanwal Rekhi School of Information Technology* :1-13.

Protharo, D.L, and K.F.Wallis, 1976, Modelling micro-economic time series. *J.Roy.Statistic.Soc.A.* 139: 468-500.

Punitha, S.B., 2007, A Comparative analysis of market performance of agricultural commodities – An Economic approach, *M.Sc. (Agri) Thesis*, Uni.Agric.Sci. Dharwad (India).

Rancheva, E., 2002, Trends of average yield changes of cereals in Bulgaria, *Agricultural Economics and Management*, 47(4): 30-35.

Rao, A. R., 1992, A study on the relationship between prices of lac and quantity exported, *Indian Forester*, **118(2)**: 122-124.T

Sadasivan, S., 1989, Pattern of pulse production – An analysis of growth trends. *Econ. And Political weekly*, **24(52)**: A167-A179.

Satya Pal, Ramasubramanian,V and Mehta,S.C., 2007, Statistical models for Forecasting Milk Production in India, *J.Ind.Soc.Agril.Statist*., 61(2), 2007: 80-83.

Schmity, Andrew and Walts Donald, C., 1970, Forecasting wheat yields: Application of parametric time series modelling. *Am.J.Agric.Econ*. 52:247-254.

Sheldon, P. J., 1993, Forecasting tourism: expenditures versus arrivals, *Journal of Travel Research, USA*, **32(1)**: 13-20.

Sisak, L., 1989, The principle of adaptive models of time series with regard to short term forecasting and the possibilities for application to cost planning, *Lesnictvi*, **35(6)**: 549-562.

Sreekanth, P. D., Geethanjali, N., Sreedevi, P. D., Shakeel Ahmed, Ravi Kumar, N., and Kamala Jayanthi, P.D., 2009, Forecasting Groundwater level using artificial neural networks. *Curr. Sci.* , **96**(7)**:** 933-939.

Swetha, K. S., 2009, A Statistical study on the impact of rain water harvesting on farming economy, *M. Sc.(Agri.) Thesis,* Univ. Agric. Sci., Dharwad

Tennent, R. B., 1986, Intra-annual growth of young Pinus radiata in New Zealand. *New – Zealand Journal of Forestry Science*, **16(2)**: 166-175.

Teresa Jacobson., Josh James and Neil Schwertman, C., 2009, An Example of Using Linear Regression of Seasonal Weather Patterns to Enhance Undergraduate Learning, *Journal of Statistics Education*, **17**:1- 2.

Therthappa., B. S., 2005, Application of multivariate techniques in evaluation of sunflower genotypes,

Umakapila., 1982, Oilseed economy of India-A case study of groundnut. Agricole Pub Company N.D.P.

Upadhyay, P. K., Yadav, M. C., Dwivedi, H. B and Singh, C., 2001, Relative importance of reproductive traits on lifetime milk, *Progressive-Agriculture*, **1(1)**: 47-51.

Vaishnav, P. R and Dixit, S. K., 1998, Trends of groundnut productivity in long term experiment, Gujarat Agricultural University Research Journal, 24(1): 73-82.

Vasanth Kumar, T., 2002, Statistical Evaluation of price variation in Tropical Timbers, *M. Sc Thesis*, Uni. Agril. Sci., Dharwad.

Wang-YeCheng., Chen-HaiTao and Lin-Qing., 2009, Optimization of parameters of blackcurrant harvesting mechanism, *Transactions of the Chinese Society of Agricultural Engineering*, **25(3)**: 79-83.

Weisent, J., Seaver, W., Odoi, A and Rohrbach, B., 2007, Comparison of three time series models for predicting campylobacteriosis risk, Department of Comparative medicine*, The University of Tennessee*, Knoxville, T. N, USA.

Weldon, J. Horton., 1995, An Application of Double Exponential Smoothing to Interlibrary Loan Forecasting, *Journal of Interlibrary Loan, Document Delivery & Information Supply*, **5**: 97 – 100.

Yeber, M. C., Onate, K. P and Vidal, G., 2007, Decolorization of kraft bleaching effluent by advanced oxidation processes using copper (II) as electron acceptor, *Environmental Science and Technology*, **41(7)**: 2510-2514.

Yin Runsheng and Min, R.S., 1999, Forecasting short-term timber prices with Univariate ARIMA model. *J. Ame.Statis.Assoc.* 25(1):154-158.

Zhang-HuiLin., Fu-AiYing., Kang-ZhanWu., Yu-WenWen., Xu-Hua., Liu-XiaoLin and Song-AiLong., Study on adjusting coefficient for milk yield of dairy cattle, 2006, *Jrnl. of Northwest-Sci-Tech, University of Agriculture and Forestry Natural Science Edition*, **34(11)**: 27-32.

STATISTICAL MODELLING TECHNIQUE ON EXPORT OF FRUIT CROPS IN INDIA

HARI PRASAD.S **2010** **Mrs. ASHALATHA K.V.**
MAJOR ADVISOR

ABSTRACT

The Export of fruits in India have recorded increase during certain years and also decrease during some other years. India has a large range of varieties of fruit in its basket and accounts for 10% of worlds total fruit production. India is the largest producer of Mango, Banana, Sapota and Lime. About 39% of world's Mango and 23% of world's Banana is produced in the country. The secondary data on Export of fruit crops was collected based on the availability. Annual Fruit crops Export data from the CMIE, Foreign Exchange Reviews (1994 to 2009 in 1000$US), Annual Mango Export data from the APEDA DATABASE from (1985 to 2009, in crores). Annual Banana Export data from FAO STAT DATABASE (1985 to 2008 in 1000$US). The research aim was to develop a suitable statistical model for describing the trend by employing different statistical modeling. At last the best one is recommended for the further study. The polynomial function fitting was carried out with different degrees and the selection of the best model was based on R-square value. Different degrees were tried to fit among all the degrees carried out 5th degree polynomial showed best fit to the Annual fruit Export, 6th degree polynomial showed best to the Mango Export, 6th degree polynomial showed best to the Banana export and Quadratic polynomial (2nd degree) showed the best to the Monthly Export of fruit crops. Nine Non-linear models were tried among all the models the best model was selected based on the highest R-square value and least MSE value. All models were found to be significant (at 1% and 5%). These models were fitted to Annual fruit Export, Mango and Banana Export. In the process it showed that the Compound, Growth and Exponential models were best fit with highest R-square and least MSE values. But only in case of Mango Export it showed Linear model was best fit and the next best was found to be the same models. So these three models are best to find the growth in Export or to find trends for Exports.

The Box-Jenkins procedure is concerned with fitting a mixed Auto Regressive Integrated Moving Average (ARIMA) model to a given set of data. The objectives in fitting this ARIMA model are to identify the stochastic process of the time series and predict the future values accurately. These methods have also been useful in many types of situation which involves the building of models for discrete time series and dynamic systems. Keeping in view of the specific objective of the present study, the data on the export of monthly fruit crops of India was collected. The Box-Jenkins ARIMA model is fitted to the monthly fruit exports. Seasonality was found in the data so seasonal ARIMA model is used.. So, the differencing was done to remove non-stationarity and made stationary. Making use of differenced series (which is stationary), the ACF and PACF are computed because it helps in tentatively identify the model. Then the parameters of all tentatively identified model was estimated by iterative process. The AIC and SBC values of least were considered to identify the model. The model (1,1,1 : 1,1,1) was identified and subjected to diagnosis checking in order to determine the adequacy of model. The residues of estimated model is examined for testing the randomness of series and for its significance. Both Ex-ante and Ex-post forecast is done for the identified model. In selecting the best model to forecast the trend for monthly fruit export, the accuracy measures MAPE and MSE are considered. Among all the models tried, the Box-Jenkins ARIMA model (1,1,1 :1,1,1) was best fit with least MAPE (15.198 percent) and MSE (3.36) values. And the forecasted values from the ARIMA model were much nearer to the Actual values.