Project Dissertation on

DEVELOPING A DIAGNOSTIC TOOL AND TIME SERIES ANALYSIS IN RESPECT OF SOYBEAN ARRIVALS AND PRICES

Submitted By:

Shivesh Dwivedi

(2K15/MBA/49)

Under the Guidance of:

Dr. Pradeep Kumar Suri

Professor

DELHI SCHOOL OF MANAGEMENT

Delhi Technological University

Bawana Road, Delhi 110042

Jan - May 2017

CERTIFICATE

This is to certify that the dissertation report titled **"Developing a Diagnostic Tool and Time Series Analysis in Respect of Soybean Arrivals and Prices"** is a bonafide work carried out by **Mr. Shivesh Dwivedi** of **MBA 2015-17** and submitted to Delhi School of Management, Delhi Technological University, Bawana Road, Delhi-42 in partial fulfilment of the requirement for the award of the Degree of Masters of Business Administration.

Signature of Guide **Signature of Head (DSM)**

Seal of Head

Place:

Date:

DECLARATION

I, **Shivesh Dwivedi**, student of **MBA 2015-17** of Delhi School of Management, Delhi Technological University, Bawana Road, Delhi – 42, hereby declare that the Dissertation report **"Developing a Diagnostic Tool and Time Series Analysis**

in Respect of Soybean Arrivals and Prices" submitted in partial fulfilment of Degree of Masters of Business Administration is the authentic work conducted by me.

The information and data given in the report is original to the best of my knowledge.

This report is not being submitted to any other University, for award of any other Degree, Diploma or Fellowship.

Place: **Shivesh Dwivedi**

Date:

ACKNOWLEDGEMENT

Firstly, I would like to thank Dr. Pradeep Kumar Suri, who was my guide and my mentor during this project, for providing me with the opportunity to do this dissertation under his guidance and expertise.

I would also like to thank Mrs. Varuna, research scholar at Delhi School of Management, for her continuous support throughout the project.

I also extend my gratitude to my Parents, Mr. Ramesh Chandra Dwivedi and Mrs. Brinda Dwivedi, who provided me with their love and support whenever I needed it.

Lastly, I would like to thank God for giving me the skills and the patience because of which this project was made possible.

Shivesh Dwivedi

ABSTRACT

“India is an agriculture based country, so appropriate distribution of marketing related information of agricultural commodities throughout the nation is essential in today’s scenario. The AGMARKNET portal is the leading step taken by the Directorate of Marketing and Inspection, Ministry of Agriculture and Farmer Welfare, Government of India, with the target of collecting, cleaning and distributing marketing related information of agricultural commodities throughout the country. This information is of utmost importance in nature as it can be used by various stakeholders, such as farmers, traders and policy makers among others, to make important decisions at individual as well as national level. It also aims to strengthen the economic position of farmers as well as consumers by providing them with marketing related information of agricultural commodities spanning over all the markets in the country”. Such information will let the farmers get fair returns on their crops. “For consumers, it means that they will be able to obtain agricultural commodities at fair and affordable prices”.

Two main objectives of this study. “First, to develop a diagnostic for major markets where Soybean arrivals are high. This diagnostic tool will help the concerned stakeholders to have various checks on the entire process of data entry. Thereby, improving the data quality by monitoring the data reporting process continuously. Statistical Process Control (SPC) can help the correct reporting of data”. Microsoft Excel 2013 was used for diagnostic tool development.

The second objective of the study is to forecast the modal prices of Soybean for a particular variety using time series modelling. “Data filtering, sorting and cleaning are the essential tasks conducted as part of this process. The time series modelling has been attempted to forecast Soybean prices in 2017. Microsoft Excel 2013 and ETS forecasting in R (language) have been used for time series analysis”.

Table of Contents

Chapter 1

Introduction

1.0 Introduction to the Project

The main emphasis of the study is on the importance of agriculture and agricultural related information for our country India. The importance and the various objectives to be covered in the study are also to be highlighted in this chapter.

1.1 Background

India is mainly an agricultural country. Agriculture is the most important occupation for most of the Indian families. In India, agriculture contributes about sixteen percent (16%) of total GDP and ten percent (10%) of total exports.

Over 60 % of India's land area is arable making it the second largest country in terms of total arable land. Agricultural products of significant economic value includes rice, wheat, potato, tomato, onion, mangoes, sugar-cane, beans, cotton, etc.

Agriculture is the backbone of Indian economy. Though, with the growth of other sectors, the overall share of agriculture on GDP of the country has decreased. Still, Agriculture continues to play a dominant part in the overall economic scenario of India. India exports excess food and agricultural products. A large proportion of India's export trade is based on the agricultural products, such as jute, tea, tobacco, coffee, spices, and sugar. It helps in increasing the foreign exchange. India is ranked seventh in terms of agricultural exports. In 2013, India exported agricultural products valuing around 39 billion dollars.
Agriculture provides raw materials to our leading industries such as cotton textiles and sugar industries. Not only this the workers in industries depend on agriculture for their food. Agriculture also provides the market for a variety of goods.

So, the Central Sector Scheme project of Agricultural Marketing Information Network (AGMARKNET) was started in March of the year 2000 by the Ministry of Agriculture under Government of India. It targeted to network all the agricultural commodities wholesale markets throughout the nation and the State Agricultural Marketing Boards and Directorates. The project is technically supported by the National Informatics Centre (NIC).

The target is to collect, break down and distribute market information to the stakeholders. More than 3200 markets are mapped under the scheme and more than 2700 markets are giving information to AGMARKNET portal. More than 350 commodities and 2000 varieties are under the plan. Every unit is furnished with important IT hardware, alongside web network. "AGMARK," a convenient portal has been made to encourage gathering and dissipation of information at market level. The framework is currently web empowered. It additionally gives week by week information for real markets and important agribusiness products. Online Exchange Portals are connected to the AGMARKNET entryways to give current and future costs.

1.1.1 Directorate of Marketing and Inspection

"The Directorate of Marketing and Inspection (DMI), under Ministry of Agriculture & Farmers Welfare, was set up in the year 1935. The Agricultural Marketing Adviser to the Government of India (AMA) heads the Directorate."

"Its prime aim is to implement agricultural policies and plans so that an integrated development of marketing information pertaining to agriculture can take place. By maintaining a close liaison with Central Government and State Government DMI also thrives to safeguard the interest of both farmers and consumers" (**source**: http://dmi.gov.in/About.aspx).

1.1.2 National Informatics Centre (NIC)

Since its establishment in the year 1976, National Informatics Centre (NIC) has been the major builder applications pertaining to the domain of e-governance having a reach up to the grassroots levels. "NICNET", connects all the Ministries and Departments of the Central Government, links 36 of the State Governments or Union Territories, and spans about 688 district administrations across India. NIC aims to implement Information Communication Technology based applications to derive competitive advantage in social as well as public administration.

NIC also has under its belt various initiatives such as "Government eProcurement System (GePNIC), Office Management Software (eOffice), Hospital Management System (eHospital), Government Financial Accounting Information System (eLekha)."

For the Agricultural Marketing Network Scheme, NIC had provided computer hardware, developed the software, "provided training to market personnel towards the operation of the hardware and software systems and provided internet connectivity. It has also developed the integration between the software packages developed by the various states with AGMARKNET to bring about seamless" uniformity in the database.

1.1.3 State Agricultural Marketing Boards

For AGMARKNET, the State Government/Marketing Boards gave the list of markets to be secured under the AGMARKNET project. The chosen markets were to give site to establishment including offices for computer installation, phone network and computer operating personnel.

Showcase Committees/Controlling specialists of AGMARKNET hub at market level were doled out to gather pertinent information and data, encourage it and transmit it to the State level and AGMARKNET entry. NIC had additionally prepared suitable people from every hub in working on a computer and dealing with programming bundle.

At each market hub, there is a man appointed to gather information and transmit it. A motivation plot has been acquainted with reward information section administrators for keeping up execution guidelines consistently.

1.2 Significance of Study

The "Agricultural Marketing Information System Network (AGMARKNET) which is based on the "NICNET" targets at linking all the important Agricultural Produce Market Committees (APMC), State Marketing Boards and Directorate of Marketing and Inspection regionals offices". So, that there can be seamless networking of information.

The idea is to help AGMARKNET to enhance the data quality of the reporting prices. As data quality of an agricultural commodity is of utmost significance to the concerned party, particularly when it comes to issues such as inflation. Moreover, forecasting the future prices of an agricultural commodity can help the stakeholders to be better prepared for price rise or fall. Through timely awareness, predicted values can provide the much-needed leverage of time to adapt and perform accordingly. Data filtering, sorting and data gap filling can help us to draw meaningful inferences from the agricultural data and perform activities such as forecasting and diagnostic tool development using that data for distinct markets

1.3 Objectives of the Study

The data belonging to both arrivals and prices is available on the AGMARKNET portal, though the reported data has many shortcomings when it comes to the quality. The daily price reporting data also shows irregularities because of lack of daily updated data. Also, the daily reporting makes the data vulnerable to mistakes which might be unintentional or deliberate in nature. The study aims to find out any gaps in the data reporting, and then to bridge those gaps in order to improve upon the data quality. The study will primarily revolve around the agricultural commodity Soybean and will focus on the study of modal prices. After the analysis of data, the study aims to fulfil two major objectives:

- To develop a tool which can report irregularities in price for major Soybean markets
- Predict the future prices of Soybean for markets and mandis

1.4 Remarks

In this chapter, it was seen that how AGMARKNET is emphasizing the need to collect information and is facilitating it. Various agencies involved with the AGMARKNET project were also highlighted. The importance of the project and the key objectives of the study were highlighted in this chapter.

Chapter-2

Review of Literature

2.0 Introduction to Review of Literature

A lot of literature is available on agricultural marketing and its' management. It is neither desirable nor possible to survey the whole literature. Therefore, only those relevant works have been reviewed, which reveal some general idea and provide a rationale for the present study. Such review of literatures always helps the researcher in getting an overview of the problem under study. It, moreover, helps to identify areas where in-depth research has not been carried out. Such as identification of fresh areas facilitate the taking up of new and meaningful research work.

2.1 What Is Agricultural Marketing?

Agricultural marketing is mainly the buying and selling of agricultural products. In earlier days when the village economy was more or less self-sufficient the marketing of agricultural products presented no difficulty as the farmer sold his produce to the consumer on a cash or barter basis.

Today's agricultural marketing has to undergo a series of exchanges or transfers from one person to another before it reaches the consumer. There are three marketing functions involved in this, i.e., assembling, preparation for consumption and distribution. Selling on any agricultural produce depends on some couple of factors like the demand of the product at that time, availability of storage etc. The products may be sold directly in the market or it may be stored locally for the time being. Moreover, it may be sold as it is gathered from the field or it may be cleaned, graded and processed by the farmer or the merchant of the village. Sometime processing is done because consumers want it, or sometimes to conserve the quality of that product. The task of distribution system is to match the supply with the existing demand by whole selling and retailing in various points of different markets like primary, secondary or terminal markets.
Most of the agricultural products in India are sold by farmers in the private sector to moneylenders (to whom the farmer may be indebted) or to village traders. Products are sold in various ways. For example, it might be sold at a weekly village market in the farmer's village or in a neighbouring village. If these outlets are not available, then produce might be sold at irregularly held markets in a nearby village or town, or in the mandi.

In India, there are several central government organisations, who are involved in agricultural marketing like, Commission of Agricultural Costs and Prices, Food Corporation of India, Cotton Corporation of India, Jute Corporation of India, etc. There

are also specialised marketing bodies for rubber, tea, coffee, tobacco, spices and vegetables.
"Under the Agricultural Produce (grading and marketing) Act of 1937, more than forty primary commodities are compulsorily graded for export and voluntarily graded for internal consumption. Although the regulation of commodity markets is a function of state government, the directorate of marketing and inspection provides marketing and inspection services and financial aid down to the village level to help set up commodity grading centres in selected markets".

(**Source:** TN Agricultural University)

2.2 Marketing Information System

"A marketing information system (MIS) is a set of procedures and methods designed to generate, analyse, disseminate, and store anticipated marketing decision information on a regular, continuous basis. An information system can be used operationally, managerially, and strategically for several aspects of marketing".

A marketing information system can be used operationally, managerially, and strategically for several aspects of marketing.

Current interchanges innovations open up the conceivable outcomes for market data administrations to enhance data conveyance through SMS on mobile phones and the fast development of FM radio stations in many creating nations offers the possibilities of more localized information services. Radio projects like Kisanvani on All India Radio and Kisan Suvidha on versatile give famers the genuinely necessary data in regards to farm and farm produce delivery. Also, there has been an increase in penetration of internet over the last 16 years. This is evident from the fact in the year 2000, India had about 0.5% of its population using the internet whereas in July, 2016 34.8% of the total population of India uses internet, be it on desktop, laptop or mobile phones. As a result, there is rapid flow of information. AGMARKNET is one such portal giving agricultural related information of various varieties of crop to its stakeholders. The information needs to be complete and correct.

2.3 Data Reporting on AGMARKNET portal

Data on the AGMARKNET portal is primarily reported as product Arrivals and Prices. The arrival data shows the amount of agricultural produce received in tonnes. Arrival data is reported on yearly, monthly, weekly and even daily basis which is used to compile reports. One can easily see which state, district or market receives maximum arrivals of a particular agricultural commodity. On the other hand, the price data is reporting in the unit of Rs/Quintal. There are three types of prices reported on the portal they are:

- Maximum Price

  The maximum prices paid for an agricultural commodity during several transactions held in a day in a market is termed at maximum price.

- Minimum Price

  The minimum price paid for an agricultural commodity during several transactions held in a day in a market is termed at minimum price.

- Modal Price

  The modal price is the price at which most of the commodity was sold in a market during the day.

Along with these three prices one can also get the details of state name, district name, group to which agricultural commodity belongs, variety and grade.

Thus, the information being disseminated through the AGMARKNET portal is quite comprehensive and vital to all the stakeholders of agriculture.

2.4 Overview of Soybean Production

"Soybean contributes significantly to the Indian edible oil pool. Presently soybean contributes 43 % to the total oilseeds and 25% to the total oil production in the country. Currently, India ranks fourth in respect to production of soybean in the world. The crop helps earn valuable foreign exchange (Rs. 62000 millions in 2012-13) by way of soya meal exports. Soybean has largely been responsible in uplifting farmer's economic status in many pockets of the country". It usually fetches higher income to the farmers owing to the huge export market for soybean de-oiled cake. "While on one hand production of Soybean in India has increased at a CAGR of 9.60 per cent from 6.87 million tonnes in 2004-05 to 15.68 million tonnes in 2012-13. On the other hand Soybean meal consumption has also increased at a CAGR of 10.82 per cent over the last eleven years from 1365 thousand million tonnes in 2004-05 to 4225 thousand million tonnes in 2014-15".

2.4.1 Climatic Conditions

The climatic conditions suitable for growing Soybean are given below:

- Temperature: The temperature appropriate for soybean growth falls in the range of 21°C and 32°C amid the time of sowing. Soybean grain does not develop completely if the temperature falls underneath 21°C. Temperature as high as 40°C hamper the crop production.

- Rainfall: Soybean grows in moderate precipitation areas. Regions having 400 to 500 cm yearly normal rainfall develop Soybean. Lack of rainfall can also be tackled with proper irrigation.

2.4.2 Trade Policy

As per the government of India policy, the export of Soybean is duty free. It generates a lot of foreign exchange for the country.

The import of soybean for public consumption is free, however, import of soybean as an oilseed can be done after paying certain duty. There have not been instances of import of soybean however.

2.4.3 Major Soybean Producing States (2015-16)

The Major Soybean producing states for the year 2015-16 according to Soybean Processor Association of India are:

SNo.	States	Kharif 2016		
		Sowing Area	Expected Yield	Estimated Production
1.	Rajasthan	9.814	968	9.499
2.	Madhya Pradesh	54.010	1059	57.170
3.	Maharashtra	35.809	1102	39.455
4.	Andhra Pradesh	2.993	815	2.439
5.	Chhattisgarh	1.340	975	1.307
6.	Gujarat	1.379	1005	1.386
7.	Karnataka	3.240	785	2.543
8.	Others	1.131	980	1.108
	Grand Total	**109.716**	**1047**	**114.907**

Table 2.2: Major Soybean producing states

In the case of agriculture the quality of reported data is of utmost importance. Daily updating of agricultural data is however a cumbersome procedure. Many discrepancies can arise while reporting the data due to fledgling nature of the task. Thus, the entire data reporting process keeps on running without any checks and controls.

2.5 Statistical Process Control

Statistical Process Control (SPC) is not new to industry. In 1924, a man at Bell Laboratories developed the control chart and the concept that a process could be in statistical control. His name was William A. Shewart. He eventually published a book titled *"Statistical Method from the Viewpoint of Quality Control"* (1939). The SPC process gained wide usage during World War II by the military in the munitions and weapons facilities. The demand for product had forced them to look for a better and more efficient way to monitor product quality without compromising safety. SPC filled that need. The use of SPC techniques in America faded following the war. It was then picked up by the Japanese manufacturing companies where it is still used today. In the 1970s, SPC startedto gain acceptance again due to American industry feeling pressure from high quality products being imported from Japan. Today, SPC is a widely used quality tool throughout many industries. SPC is method of measuring and controlling quality by monitoring the manufacturing process. Quality data is collected in the form of product or process measurements or readings from various machines or instrumentation. The data is collected and used to evaluate, monitor and control a process. SPC is an effective method to drive continuous improvement. By monitoring and controlling a process, we can assure that it operates at its fullest potential. One of the most comprehensive and valuable resources of information regarding SPC is the manual published by the Automotive Industry Action Group (AIAG).

Control Charts

One of the most widely used control charts for variable data is the X-bar and R chart. X-bar represents the average or "mean" value of the variable x. The X-bar chart displays the variation in the sample means or averages.

Analysing the Data

The data points recorded on a control chart should fall between the control limits, provided that only common causes and no special causes have been identified. Common causes will fall between the control limits whereas special causes are generally outliers or are outside of the control limits.

2.6 Time Series Forecasting

Time series forecasting is the use of a model to predict future values based on previously observed values. “While regression analysis is often employed in such a way as to test theories that the current values of one or more independent time series affect the current value of another time series, this type of analysis of time series is not called "time series analysis", which focuses on comparing values of a single time series or multiple dependent time series at different points in time”.

Ramasubramanian V. of Indian Agricultural Statistics Research Institute(IASR) with expertise in Agricultural Economics, Aquaculture, Artificial Neural Network has discussed various time series models for agricultural forecasting in“Forecasting Techniques in Agriculture” namely:

- Exponential Smoothing Models
- Auto Regressive Integrated Moving Average (ARIMA) Models

So, from the above references it is seen that there is evidence that forecasting of time series has been done in agriculture sector to predict the future crop yields.

2.6.1 Exponential Smoothing Models

“Smoothing models represent the evolution of a time series by the model:

$y_t = \mu_t + \beta_t t + s(t) + a_t$ where

μ_t is the time-varying mean term,

β_t is the time-varying slope term,

$s(t)$ is one of the s time-varying seasonal terms,

a_t are the random shocks.

Models without a trend have $\beta_t = 0$ and nonseasonal models have $s(t) = 0$. The estimators for these time-varying terms are

L_t smoothed level that estimates μ_t

T_t is a smoothed trend that estimates β_t

S_{t-j} for $j = 0, 1, \ldots, s-1$ are the estimates of the $s(t)$.

“Each smoothing model defines a set of recursive smoothing equations that describes the evolution of these estimators. The smoothing equations are written in terms of model parameters called *smoothing weights*.” They are

α, the level smoothing weight”

γ, the trend smoothing weight
ϕ, the trend damping weight
δ, the seasonal smoothing weight.
While these parameters enter each model in a different way (or not at all), they have the common property that larger weights give more influence to recent data while smaller weights give less influence to recent data.
Each smoothing model has an ARIMA model equivalent. You may not be able to specify the equivalent ARIMA model using the **ARIMA** command because some smoothing models intrinsically constrain the ARIMA model parameters in ways the ARIMA command will not allow.

Simple Moving Average
A simple moving average model (SMA) produces plotted values that are equal to the average of consecutive observations in a time window. The plotted values can be uncentred or centred in the time window.
If we let $f_t=(y_t+y_{t-1}+y_{t-2}+...y_{t-(k-2)}+y_{t-(k-1)})/k$ be the average of *k* consecutive observations, then:
•for an uncentred SMA(*k*), f_t is the plotted value for time *t*.
•for a centred SMA(*k*) with odd *k*, f_t is the plotted value for time *t*-(*k*-1)/2.
•for a centred SMA(*k*) with even *k*, f_t is the plotted value for time *t*-(*k*-1)/2. When saved to a data table, f_t is at *t*-(*k*-2)/2.
•for a centred SMA(*k*) with double smoothing for even *k*, the smoothing estimates are given by the following:

$$f_{t-\frac{k}{2}} = \frac{y_t + 2\sum_{i=1}^{k-1} y_{t-i} + y_{t-k}}{2k}$$

To fit a simple moving average model, select **Smoothing > Simple Moving Average**. A window appears with the following options:

Enter smoothing window width
Enter the width of the smoothing window.

Centering
Choose between **No Centering**, **Centred**, and **Centred and Double Smoothed for Even Number of Terms**.
The Simple Moving Average report shows a time plot of the data and the fitted model. The red triangle menu has the following options:

Add Model
Select this option to fit another model. When additional models are fit, the model is added to the time plot of the data.

Save to Data Table
Saves the original data, and plotted values of all moving average models.
Show Points
Shows or hides the points on the plot.
Connecting Lines
Shows or hides the lines on the plot.

Smoothing Model Dialog
The Smoothing Model dialog appears in the report window when you select one of the smoothing model commands. The **Confidence Intervals** popup list allows you to set the confidence level for the forecast confidence bands. The dialogs for seasonal smoothing models include a **Periods Per Season** box for setting the number of periods in a season. The **Constraints** popup list lets you to specify what type of constraint you want to enforce on the smoothing weights during the fit. The constraints are:

Zero To One
Keeps the values of the smoothing weights in the range zero to one.

Unconstrained
Allows the parameters to range freely.
Stable Invertible
Constrains the parameters such that the equivalent ARIMA model is stable and invertible.
Custom
Expands the dialog to allow you to set constraints on individual smoothing weights. Each smoothing weight can be **Bounded**, **Fixed**, or **Unconstrained** as determined by the setting of the popup menu next to the weight's name. When entering values for fixed or bounded weights, the values can be positive or negative real numbers.

The example shown here has the Level weight (α) fixed at a value of 0.3 and the Trend weight (γ) bounded by 0.1 and 0.8. In this case, the value of the Trend weight is allowed to move within the range 0.1 to 0.8 while the Level weight is held at 0.3. Note that you can specify all the smoothing weights in advance by using these custom constraints. In that case, none of the weights would be estimated from the data although forecasts and residuals would still be computed. When you click **Estimate**, the results of the fit appear in place of the dialog.

Simple Exponential Smoothing

The model for simple exponential smoothing is $y_t = \mu_t + \alpha_t$.
The smoothing equation, $L_t = \alpha y_t + (1 - \alpha)L_{t-1}$, is defined in terms of a single smoothing weight α. This model is equivalent to an ARIMA(0, 1, 1) model where
$(1 - B)y_t = (1 - \theta B)\alpha_t$ with $\theta = 1 - \alpha$.

The moving average form of the model is

$$y_t = a_t + \sum_{j-1}^{\infty} \alpha a_{t-j}$$

Double (Brown) Exponential Smoothing

"The model for double exponential smoothing is $y_t = \mu_t + \beta_1 t + a_t$.
The smoothing equations, defined in terms of a single smoothing weight α are
$L_t = \alpha y_t + (1-\alpha)L_{t-1}$ and $T_t = \alpha(L_t - L_{t-1}) + (1-\alpha)T_{t-1}$.
This model is equivalent to an ARIMA(0, 1, 1)(0, 1, 1)$_1$ model"
$(1-\mathrm{B})^2 y_t = (1-\theta\mathrm{B})^2 a_t$ where $\theta_{1,1} = \theta_{2,1}$ with $\theta = 1-\alpha$.
The moving average form of the model is

$$y_t = a_t + \sum_{j=1}^{\infty} (2\alpha + (j-1)\alpha^2)a_{t-j}$$

Linear (Holt) Exponential Smoothing

"The model for linear exponential smoothing is $y_t = \mu_t + \beta_t t + a_t$.
The smoothing equations defined in terms of smoothing weights α and γ are"
$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + T_{t-1})$ and $T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1}$
"This model is equivalent to an ARIMA(0, 2, 2) model where
$(1-\mathrm{B})^2 y_t = (1-\theta\mathrm{B} - \theta_2\mathrm{B}^2)a_t$ with $\theta = 2 - \alpha - \alpha\gamma$ and $\theta_2 = \alpha - 1$.
The moving average form of the model is"

$$y_t = a_t + \sum_{j=1}^{\infty} (\alpha + j\alpha\gamma)a_{t-j}$$

Damped-Trend Linear Exponential Smoothing

"The model for damped-trend linear exponential smoothing is $y_t = \mu_t + \beta_t t + a_t$.
The smoothing equations in terms of smoothing weights α, γ, and ϕ are"
$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + \varphi T_{t-1})$ and $T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)\varphi T_{t-1}$
This model is equivalent to an ARIMA(1, 1, 2) model where
$(1-\varphi\mathrm{B})(1-\mathrm{B})y_t = (1-\theta_1\mathrm{B} - \theta_2\mathrm{B}^2)a_t$ with $\theta_1 = 1 + \varphi - \alpha - \alpha\gamma\varphi$ and
$\theta_2 = (\alpha - 1)\varphi$.
The moving average form of the model is

$$y_t = \alpha_t + \sum_{j=1}^{\infty} \left(\frac{\alpha + \alpha\gamma\varphi(\varphi^j - 1)}{\varphi - 1} \right) \alpha_{t-j}$$

2.6.1 Making Time Series Stationary

To check a whether a time series is stationary or not a test known as Breakpoint Unit Root Test is used. This test has a null hypothesis that *Ho: Series has a unit root.*

If a series has unit root then it a non-stationary series. To check the hypothesis, p-value is seen i.e. if p-value is less than 0.05 then hypothesis is rejected and time series is stationary, where as if p-value is greater than 0.05 then hypothesis is accepted and the time series is non-stationary.

One approach to make a time series stationary is to find the differences between the data recorded. This is known as differencing. Changes, for example, logarithms can balance out the variations of a time series. Differencing can help balance out the mean of a time series.

- Random walk model

"The differenced series is formed by subtracting the next observation from the previous one, and can be written as"

$$y'_t = y_t - y_{t-1}$$

"The differenced series will have only $t-1$ values since it is not possible to calculate a difference for the first observation."

2.7 Conclusion of the Literature

As we look into the literature we can say that a lot of content available on the topic which is diverse and encompassing in nature. Also, the reporting technique of AGMARK for timely notation of arrival and prices were also noted. There are evidences regarding the applicability of statistical process control in non-manufacturing sector and how time series forecasting has been done on the agricultural sector as well. Now, the methodology involved in the study shall be discussed.

Chapter-3

Methodology

3.0 Introduction to Methodology

“In organizations, **analytics** enables professionals to convert extensive data and statistical and quantitative analysis into powerful insights that can drive efficient decisions. Therefore with **analytics** organizations can now base their decisions and strategies on data rather than on gut feelings. After the data is gathered and assessed, the sources are utilized to demonstrate a speculation or bolster a thought”. A person can bring out small details to form greater assumptions about the material by using critical thinking skills effectively.

The research in this project is on secondary data obtained from the AGMARKNET.GOV.IN portal. Through this analysis, the study aims to bring to forward the gaps in order to develop an effective diagnostic tool and a forecasting model.

3.1 Project Phases

3.1.1 Understanding Phase

Step 1: To comprehend the broader purpose(s) of the project.

The project is targeted at developing a diagnostic tool for a particular commodity in order to improve the quality of the data reported in future. Developing a diagnostic tool will help farmers and citizens of the country to have correct information regarding the crop produce as it will help address the issues such as modal prices reported as zeros, modal prices reported as averages of maximum and irregularities in data reporting. Thereby reducing the error in reporting.

The second objective of the project is to forecast the prices of the chosen commodity. This activity will help farmers and government to understand the prices in the coming months of 2017.

Step 2: To understand the data reporting fields of AGMARKNET portal.

The AGMARKNET portal reports Arrivals (in tonnes) and Prices (in

Rs/Quintal) along with crop commodities' state, district, variety, grade and date of reporting.

3.1.2 Defining Phase

Step 1: To choose a commodity for analysis.

The commodity chosen for the project was Soybean.

Step 2: To define a time frame of the data reporting period for that commodity.

The time frame for the data so that the diagnostic tool can be developed was 01 January, 2016 to 31 December, 2016. Whereas, the time period for data forecasting for the year 2017 was chosen to be 01 January,2013 to 31 December,2016.

Step 3: To identify the key data term(s) to be worked upon for analysis from amongst maximum price, modal price and minimum price.

The data reporting on the AGMARKNET portal consists of maximum price, minimum price and modal price. Out of them modal prices were chosen for the analysis part as this is the price at which most of the commodity is transacted in a market.

3.1.3 Data Preparation and Analysis Phase

Step 1: Tools chosen for Analysis

The software chosen for development of diagnostic tool was Microsoft Excel 2013, where the softwares chosen for time series forecasting was 'R', an open source programming language.

Step 2: To find the state, district and market with maximum arrivals within the reporting period.

The state with maximum arrivals for Soybean was found out. Then within that state the district and market with maximum arrivals for Soybean were found.

Step3: Check for the reporting frequency of the data fields (modal price and date).

The plot of prices reported per month was done for the whole reporting period to check the consistency in the data.

Step 4: Clean the data for better analysis.

The data gaps were identified and they were filled using the average method in order to have a consistently reported past data which can be worked upon easily to develop diagnostic tool and forecast model.

Step 5: Analyse the data to develop diagnostic tool and forecast model in the next phase.

The data was analysed for data gaps, zeros and reporting of averages of maximum prices and minimum prices in place of modal prices.

3.1.4 Developing Phase

Step 1: To develop a diagnostic tool incorporating statistical process control charts.

To develop a diagnostic tool, control chart was used. Various reports pertaining to control charts were studied and the apt chart was chosen according to the best data fit. The chart helped us to get the upper and lower control limits.

Step 2: Develop a time series forecast model for the modal price of the chosen commodity.
To forecast the time series data into 2017, Error, Trend, Seasonal (ETS) forecasting method was chosen. For data smoothening, Holtwinters framework was chosen from ‘R’.

3.1.5 Testing Phase

Step 1: Test the diagnostic tool on other markets reporting the same commodity.

The diagnostic tool thus developed for one market was applied on other markets as well by using their respective control limits to check the applicability of the tool across markets for the same commodity.

Step 2: Check the accuracy of the forecast model.

To check the accuracy of the forecast various parameters like Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) were used.

3.1.6 Findings and Recommendations Phase

After extensive analysis, one will be able to bring out some findings and recommendations through which the AGMARKNET portal can benefit. The limitations of the diagnostic tool and forecasting model are also highlighted.

3.2 Concluding Remarks

The approach towards the study of the project and the project methodology was discussed in this chapter. Now in the next chapter the data analysis, findings, limitations and recommendations regarding the study shall be discussed.

Chapter-4

Data Analysis

4.0 Introduction to Data Analysis

The commodity chosen for analysis in the understanding phase was Soybean. The commodity data for Soybean obtained from the AGMARKNET portal was usually in .xlsx or .csv format and various operations such as data filtering, data sorting and data cleaning was done in order to draw and make sense out of the compiled data. Since, the project had two major objectives the entire analytics process was classified into two major heading.

4.1 Development of a Diagnostic Tool

The development of diagnostic tool was a multi phased process. The phases were as follows:

4.1.1 Choosing a Tool to Analyse Data

Since, the data obtained from AGMARKNET was in excel format, so, Microsoft Excel 2013 was chosen to analyse the data and develop a diagnostic tool.

4.1.2 Selection of Market for Analysis

The arrival data for Soybean for the past one year i.e. from 01 January, 2016 to 31 December, 2016 was obtained for all the states from the AGMARKNET portal. The arrivals were for each state and the percentage share of total arrivals for each state was found. It was clear to see that Madhya Pradesh was the leader in Soybean arrivals with 20.9% share of the total Soybean arrivals. Hence, the data of the state of Madhya Pradesh was analysed to find the district and market showcasing maximum arrivals in tonnes. The district that gave maximum arrivals was Ujjain with a share of nearly 22%. So, it was decided to develop the diagnostic tool for Ujjain market. However, it was found that the Ujjain market had less no. of arrivals and also the frequency of arrivals or say Data entry was pretty less. So, it was decided to do analysis for the next major market in terms of Soybean arrivals viz. Shajapur. Hence the market of analysis was shifted from Ujjain to Shajapur, which had 8.9% of the total Soybean arrivals during the year 2016. The arrivals were seen throughout the year making the market appropriate for further analysis.

Market with variety	Count of Price Date	Sum of Arrivals (Tonnes)
Shajapur	2007	252126.45
Rajgarh	1512	138335.13
Ujjain	1454	622667.09
Sagar	1195	69752.18
Dhar	1125	163942.26
Ratlam	1063	119504.55
Dewas	1044	230180.52
Vidisha	1028	53415.5
Sehore	941	172033.46
Khargone	908	41359.48
Guna	840	42038.46
Indore	768	153315.26
Mandsaur	734	118286.03
Shivpuri	679	21697.7

Table 4.1: Arrival Data for Soybean (State, District and Market Wise)

(**Source:** Own Analysis on AGMARKNET Data)

Market Name	Percentage of Total Arrival	No. of Arrivals
Shajapur	10.24%	293
Shujalpur	19.96%	257
Kalapipal	6.54%	240
Nalkehda	3.48%	229
Agar	16.50%	224
Akodia	37.41%	202
Badod	1.50%	181
Momanbadodiya	1.55%	148
Susner	1.60%	112
Berachha	0.53%	66
Maksi	0.66%	42
Soyatkalan	0.03%	13

Table 4.2: Arrival Data for Soybean, Madhya Pradesh

(**Source:** Own Analysis on AGMARKNET Data)

4.1.3 Checking for Modal Price Reporting Frequency

Along with arrivals it became necessary to see the frequency with which prices were reported for Shajapur market in Madhya Pradesh. We found the price reporting to be good throughout the year as it reported for 293 days out of 366 days. Considering the fact that soybean is a seasonal crop these entries are somewhat satisfying.

Market with variety	Count of Price Date
Yellow	12872
Other	9907
Soyabeen	1337
Local	182
Black	7
Grand Total	**24305**

Figure 4.1: Soybean reporting with varieties

(**Source:** Own Analysis on AGMARKNET Data)

4.1.4 Data Cleaning of Modal Prices

The price reporting sheet for Shajapur market was analysed on the following points

- Modal Price is Average

  In the 293 entries, it was checked that how many entries had modal price equal to average of maximum and minimum price by using the "IF" statement in MS Excel. It was found that there were 23 entries in all where modal price was actually average.

- Modal Price is Zero

  The reported entries were checked for where the modal price is reported to be zero. No entry was found to be reporting modal price zero. The entries having modal price as average of maximum and minimum prices were deleted. 257 entries were now reported after deletion and by using VLOOKUP command

PIVOT TABLE the daily data gaps were filled by the average modal price of that month. Thereby, having a continuous series.

4.1.5 Developing a Control Limits and Control Charts

Once the data was properly cleaned it was necessary to develop control limits. For this statistical process control charts were referred. Since, the modal prices were of variable nature, measured on a continuous scale of daily dates and the data set had a subgroup size of one (as each day's price data cannot be combined with the next day's data), so, the Control Chart was chosen to get the upper control limit and lower control limit.

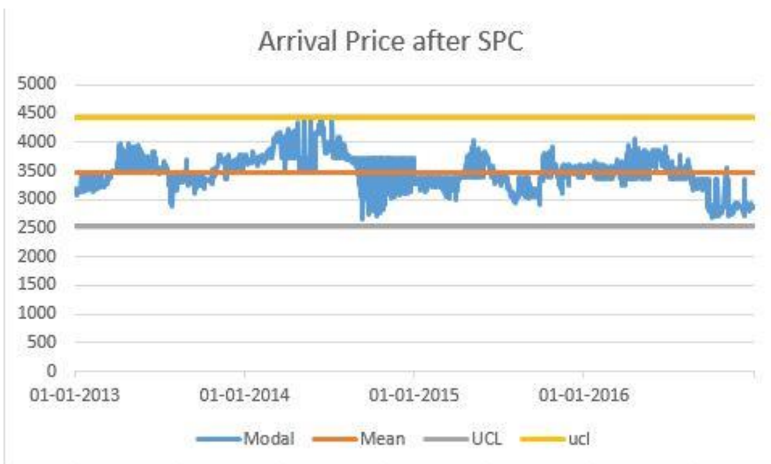

Figure 4.2: Control Chart for Shajapur Market

(**Source:** Own Analysis on AGMARKNET Data)

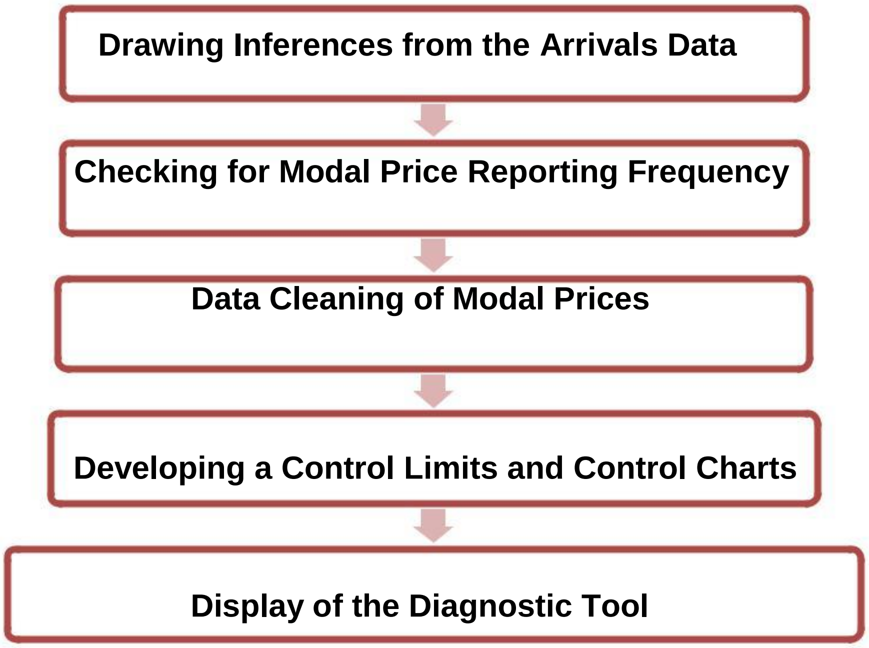

Figure 4.3: Diagnostic Tool Development

Date	Modal	Mean	UCL	ucl	OUT OF BOUND
01-01-2013	3204	3477.494	2529.483	4425.505	
02-01-2013	3195	3477.494	2529.483	4425.505	
03-01-2013	3173	3477.494	2529.483	4425.505	
04-01-2013	3191	3477.494	2529.483	4425.505	
05-01-2013	3114	3477.494	2529.483	4425.505	
06-01-2013	3114	3477.494	2529.483	4425.505	
07-01-2013	3074	3477.494	2529.483	4425.505	
08-01-2013	3125	3477.494	2529.483	4425.505	
09-01-2013	3129	3477.494	2529.483	4425.505	
10-01-2013	3150	3477.494	2529.483	4425.505	
11-01-2013	3150	3477.494	2529.483	4425.505	
12-01-2013	3130	3477.494	2529.483	4425.505	
13-01-2013	3130	3477.494	2529.483	4425.505	
14-01-2013	3463.845	3477.494	2529.483	4425.505	
15-01-2013	3197	3477.494	2529.483	4425.505	
16-01-2013	3156	3477.494	2529.483	4425.505	

Table 4.2: Data before Diagnostic Tool
(**Source:** Own Analysis on AGMARKNET Data)

Date	Modal	Mean	UCL	ucl	OUT OF BOUND
01-01-2013	3204	3478.624	2525.3	4431.947	
02-01-2013	3195	3478.624	2525.3	4431.947	
03-01-2013	3173	3478.624	2525.3	4431.947	
04-01-2013	3191	3478.624	2525.3	4431.947	
05-01-2013	3114	3478.624	2525.3	4431.947	
06-01-2013	3114	3478.624	2525.3	4431.947	
07-01-2013	3074	3478.624	2525.3	4431.947	
08-01-2013	3125	3478.624	2525.3	4431.947	
09-01-2013	3129	3478.624	2525.3	4431.947	
10-01-2013	4800	3478.624	2525.3	4431.947	OUT OF BOUND
11-01-2013	3150	3478.624	2525.3	4431.947	
12-01-2013	3130	3478.624	2525.3	4431.947	
13-01-2013	3130	3478.624	2525.3	4431.947	
14-01-2013	3463.845	3478.624	2525.3	4431.947	
15-01-2013	3197	3478.624	2525.3	4431.947	
16-01-2013	3156	3478.624	2525.3	4431.947	
17-01-2013	3165	3478.624	2525.3	4431.947	
18-01-2013	3147	3478.624	2525.3	4431.947	

Table 4.3: Data of the Diagnostic tool
(**Source:** Own Analysis on AGMARKNET Data)

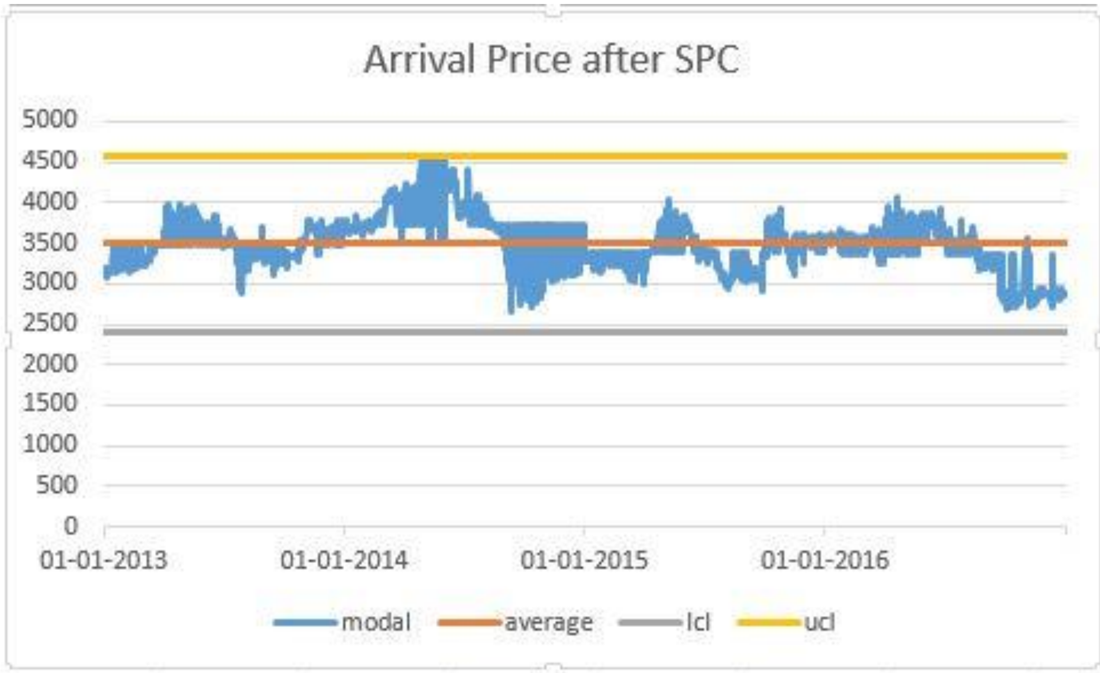

Figure 4.3: Plot After Diagnostic Tool Application
(**Source:** Own Analysis on AGMARKNET Data)

4.2 Forecasting the Modal Prices of Soybean

In order to forecast the Soybean modal prices for the year 2017 based on the modal prices of the previous years following steps were followed:

4.2.1 Choosing the Tools of Analyses

To develop a better forecast model two tools were used

- Microsoft Excel 2013: For sorting, filtering and cleaning data
- R: To help choose a forecast model

4.2.2 Drawing Inferences from the Arrivals Data

It was aimed to develop forecast model for one of the top Soybean producing markets of Madhya Pradesh, and for that it became necessary to look into the arrivals data of Madhya Pradesh's markets.

But after analysing the data of Shajapur Market (leading market in terms of arrivals) belonging to the leading district Shajapur in terms of Soybean arrivals it was found that there were a few data gaps because of price reporting gaps.

There were two major varieties of Soybean reported, namely:

- Yellow
- Other

Yellow was reported the maximum times, quite ahead of the other variety. Other is no variety of Soybean actually but it could be the normal Soybean available in the market.

4.2.3 Data Cleaning of Modal Prices

To forecast the data, it was necessary to fill the data gaps and improve the quality of data. For this it was seen that how many number of times (count) the Soybean variety-Other was reported is month i.e. in days. To bridge the data gaps, the months which had more than 15 days of reporting of modal prices had their remaining non-reported days filled with the averages of that month. Whereas, the months having data reported days' count greater than 10 but less than or equal to 15 were replaced by the average of the modal price values present of that month and the average of annual modal price. Lastly, if the count of the modal price reported days was less than 10, then, the missing values were replaced by annual average.

4.2.4 Choosing a Forecast Model and Technique

R is the perfect programming language for forecasting when a time series is involved. Holtwinters framework was used which utilises the ETS model and smoothens the data through a single algorithm.

First of all, it was checked that was the series stationary or not. Using the Dickey-Fuller test, the p-value came out to be 0.2112 which means that the null hypothesis "Series has a unit root" was not rejected, this acceptance means that the time series is non-stationary in nature.

```
> y2<-diff(y1)
> adf.test(y2, k=0)

        Augmented Dickey-Fuller Test

data:  y2
Dickey-Fuller = -64.564, Lag order = 0, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(y2, k = 0) : p-value smaller than printed p-value
```

Figure 4.4: Dickey Fuller test for stationary H_o

(**Source:** Own Analysis on AGMARKNET Data)

So, in order to make the time series stationary first level difference was taken i.e. Differencen+1=Pricen+1-Pricen. We again ran the unit root test on it based on Dickey-Fuller test to get the p-value. The p-value this time was found to be less than 0.01, hence the null hypothesis was rejected, meaning that it does not have a unit root and thus is stationary in nature.
Thus, a stationary series at first level of difference was found.

Now, Holtwinters framework does an exponential smoothening of the data which helps users to find the best fit model out of the given models. The ETS(Error-Trend-Seasonality) Exponential Smoothing of Holtwinters gives the best fit model according to the select criteria and justifies the chosen model statistically.

```
> # Augmented Dickey-Fuller test
> adf.test(y1)

        Augmented Dickey-Fuller Test

data:  y1
Dickey-Fuller = -3.6034, Lag order = 11, p-value = 0.03235
alternative hypothesis: stationary
```

Figure 4.5: Dickey Fuller test to prove H1

(**Source:** Own Analysis on AGMARKNET Data)

RGui (64-bit) - [D:\downloads\SHIVESH Time Series Analysis.R - R Editor]

File Edit Packages Windows Help

```r
# Time Series HOLTWINTERS Model for forecasting soyabean for year 2016-2017

# install packages
library(tseries)
library(forecast)

mydata<- read.csv("C:/Users/SHAKTI/Desktop/SHIVESH PROJECT/M1.csv")
attach(mydata)

# Descriptive statistics and plotting the data
summary(mydata$PRICE)
plot.ts(mydata$PRICE)

#Converting the data into Training data and Test data
a1<-ts(mydata$PRICE[c(1:1095)], frequency=365) #Training Data
a2<-ts(mydata$PRICE[c(1095:1461)], frequency=365)#Test Data
a3<-ts(mydata$PRICE,frequency=365)

#Applying the Forecast model and test for 2016 data
q1<-HoltWinters(a1)
q2<-predict(q1,n.ahead=366,prediction.interval = TRUE)
q3<-predict(q1,n.ahead=366)

plot(a2) #plotting the Test Data
plot(q2) #plotting the forecasted data
plot(a3)
lines(q3,col="blue") #plot the whole data with forecasted data

#Transfer the data to excel file
write.table(q2,file="forecated1.csv", append=TRUE,sep=",",col.names=FALSE,row.names=FALSE)

#Applying model for 2017 data
f1<-HoltWinters(a3)
f2<-predict(f1,n.ahead=100,prediction.interval = TRUE)
f3<-predict(f1,n.ahead=100)


plot(f2)
plot(f3) #plotting the forecasted data

#Transfer the data to excel file

write.table(f2,file="forecated2.csv", append=TRUE,sep=",",col.names=FALSE,row.names=FALSE)
```

Figure 4.6: R program for forecasting

(**Source:** Own Analysis on AGMARKNET Data)

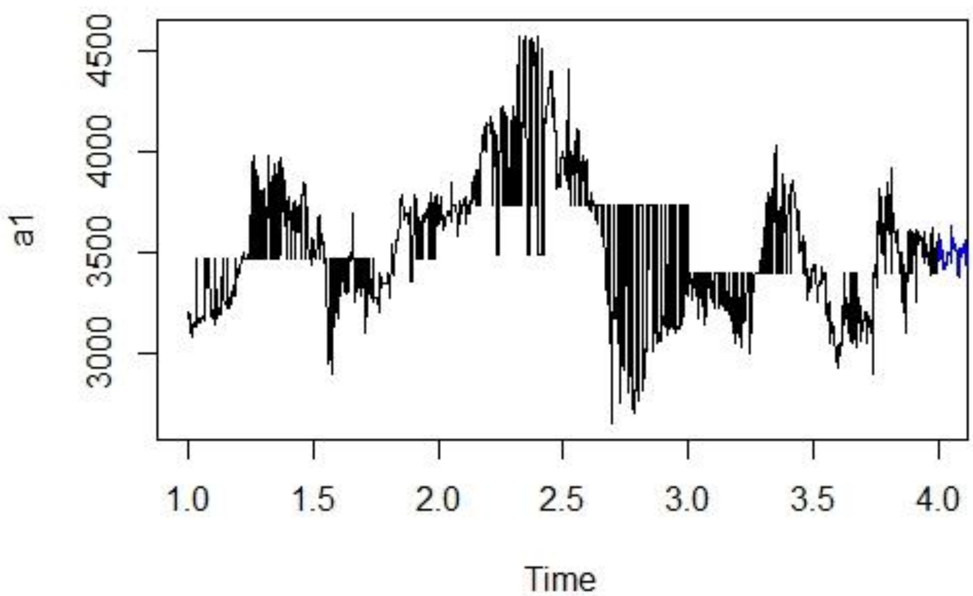

Figure 4.7: R Plot of Forecasted Data (2016)

(**Source:** Own Analysis on AGMARKNET Data)

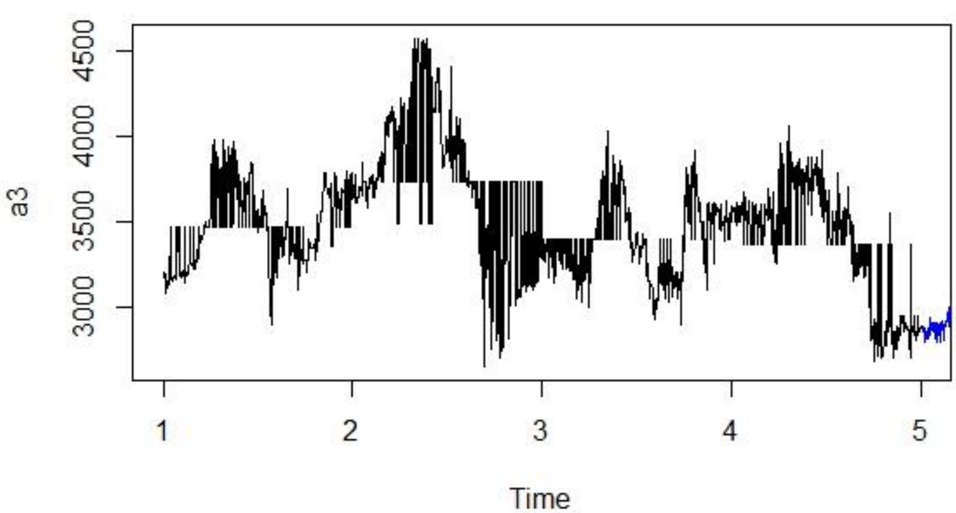

Figure 4.8: R plot of Forecasted Data (2017)

(**Source:** Own Analysis on AGMARKNET Data)

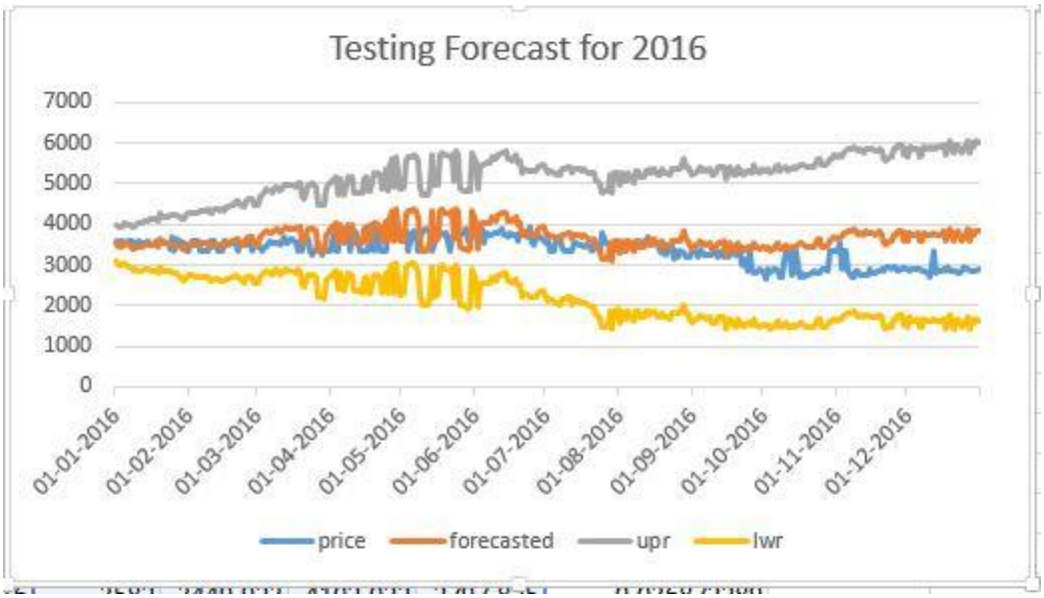

Figure 4.9: Excel Plot for forecasted Data (2016)
(**Source:** Own Analysis on AGMARKNET Data)

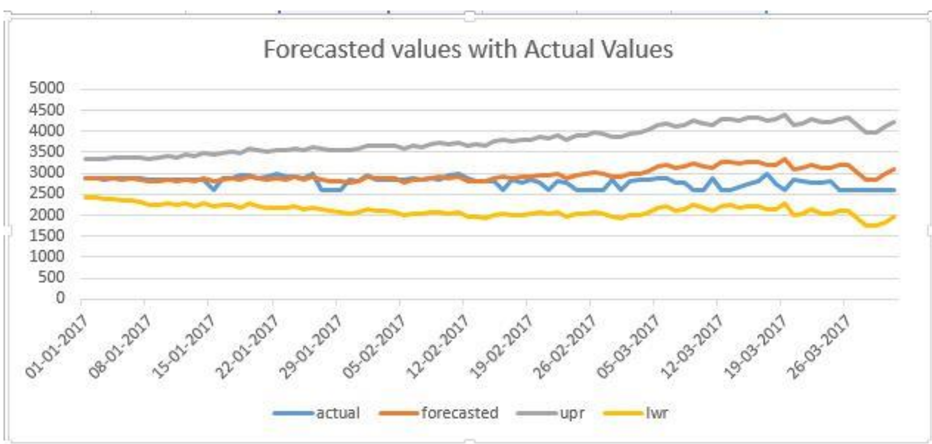

Figure 4.10: Soybean Price Forecast for 2017
(**Source:** Own Analysis on AGMARKNET Data)

The ETS exponential smoothing method was first employed on the period 01 January 2012 to 31 December 2015 to forecast the modal prices for the month January to December 2016. Since the data for 2016 was already present it was easy to check the model's accuracy by finding out various errors.

4.4 Limitations

- The analysis was done only on modal price and not on maximum and minimum prices.
- The data used for the development of the diagnostic tool was only of one year i.e. of 2016. To develop a better tool the data horizon can be widened to include previous year data.
- The diagnostic tool needs to be developed for each market separately which can make the task cumbersome.
- The data used for forecasting the modal prices of 2017 was taken from 01 January 2012 to 31 December 2016, giving a forecast horizon of only five years which can be made large.
- The forecast was done only for one market and one specific variety of Soybean. Thus, it can only be applied to that market and that variety of Soybean.

4.5 Concluding Remarks and Recommendations

After the data analysis, the diagnostic tool was developed for five major Soybean markets (in terms of arrivals) of Madhya Pradesh. Also, the time series forecasting for Soybean for a chosen market (-Madhya Pradesh) was done. To conclude the data analysis of the project, here are some recommendations.

- The diagnostic tool developed as part of the project work can be accommodated alongside the data entry reporting portal of Soybean prices i.e. http://agmarknet.nic.in/market_online/. Integration with the portal will give the checks and raise alerts then and there itself, making the data entry process effective and thereby improving the data quality of reporting.

- The price data reporting frequency should be monitored daily as it can help provide the actual data. This actual data can in turn help in developing robust diagnostic tools and accurate forecast models. For example, Ujjain-the leading Soybean market for 2016 had a poor reporting frequency and the market of analysis was shifted.

- The time series forecast in the case of this study was done for a specific variety of Soybean and for a specific market but instead the forecast model needs to be developed for all varieties of Soybean across various markets.

- The variety terms of Soybean on AGMARKNET portal like “Others” need to classified as well to maintain effective tracking on the Soybean arrivals.

- In the findings of the study it was seen that the data reported historically pertaining to an agricultural commodity can be edited and changed. The editing of already reported data can disturb and hamper the entire data analysis of the project and thus, the changing of the data should be done in a restrictive manner to ensure integrity of reported data.

Chapter 5

Conclusion of the Study

The project began by understanding the agencies associated with AGMARKNET. Also, the importance of the study and objectives of the study were highlighted which gave a direction to develop a well-planned methodological approach to the project study.

Further on in the project, the review of literature had highlighted the importance of the information particularly when it comes to the aspect of agricultural marketing information. The review of literature had also brought to light the need manage and improve the quality of data being reported. Various definitions of marketing information were obtained. The working of AGMARKNET, an agricultural commodity data reporting portal which can help in bringing insightful information to the farmers, government and consumers was understood. It was also seen how the data is being reported on this portal. The agricultural commodity's data worked on in this project was that of Soybean. It was also necessary to understand some facts related to Soybean like major Soybean producing states, temperature and rainfall conditions and trade policies, and the same was done in this study.

The study had also highlighted how statistical process control can be used in non-manufacturing sector to continuously monitor the data being reported. In this project, it was also emphasized that which statistical process control charts can be used to have mechanism of check on the prices being reported on the AGMARKNET portal for a particular commodity such as Soybean. The equations of upper control limit, lower control limit and control limit were highlighted for variable process control chart i.e. individual moving range chart. It was also discussed that how the control limits can be improved by developing seasonal factors for each month so that marginally neglected prices can be accommodated. The diagnostic tool developed on the platform of Microsoft Excel was on the basis of statistical process control charts for one major Soybean arrival market of Madhya Pradesh.

To check whether the approach adopted for developing the diagnostic tool was scalable to other markets or not, the same approach was applied to other four markets of Madhya Pradesh. It was seen that more than 90% of the modal prices reported for all the five markets was within the limits developed by using statistical process control chart. The diagnostic tool also took care of the modal prices that were being reported as average of maximum and minimum prices and reported a warning whenever modal prices was average.

Also, the study of this project emphasized the importance of time series forecasting particularly when it comes to the prices of agricultural commodities. Through the review of the available literature it was seen that what were the different models and techniques to analyse the time series. With a well-defined mechanism for data gap filling and after applying data filtering and data sorting in Microsoft Excel 2013, R was used to develop a suitable model to forecast the modal prices for Soybean for the year 2017. Using the ETS forecasting gave a particular model, this model was applied to a subset of the dataset and was checked for errors. The error percentage came out to be less than 4%. Then the ETS model was applied on the entire data set to get the forecasted values for the year 2017. Till March 2017, the mean absolute error percentage was 7.92%. The forecasting of time series was done for one particular market of Madhya Pradesh and for one specific variety of Soybean.

During the study, it was also seen that AGMARKNET had some data reporting inconsistencies i.e. the data reported was changed after a time period. Screenshot comparison of the data taken on a previous date with that of the reported data values in the future showed enough evidence of data inconsistencies. Along with the findings, the study also highlighted the findings and recommendations of the study.

Thus, we see that by developing a diagnostic tool for various markets can help in improving the quality of data being reported as it will create a mechanism of checks and balances. The need for developing these kinds of diagnostic tools across all markets and for all varieties of agricultural commodities was emphasized in the study. Whereas, the forecasted modal prices will assure the farmers, government agencies and consumers to get a prior information regarding the agricultural commodities. It was also highlighted that with a wider range of data set the forecasted model thus developed can be more accurate.

Bibliography

1) B B Singh,(2013),"Success of Soybean in India: The Early Challenges and Pioneer Promoters",Research Journal, Institute of Tropical Agriculture (IITA)
2) Archer Daniels Midland Company (ADM) and Department of Agriculture, Government of Maharashtra Supported by Public Private Partnership for Integrated Agriculture Development Programme(2014), "Documentation of the project on Improving productivity of Soybean in Maharashtra". Vol 1
3) Gareth James, Daniela Witten, Trevor Hastie Robert Tibshirani(2005); An Introduction to Statistical Learning with Applications in R. Vol 8
4) https://stats.stackexchange.com/questions/63250/forecasting-function-r-holt-winters-hw-approach.
5) https://www.rstudio.com/wp-content/uploads/2016/11/ggplot2-cheatsheet-2.1.pdf, Data visualization with ggplot2 RStudio Cheat Sheet.
6) Rob Hyndman, Mitchell O'Hara-Wild, Christoph Bergmeir, Slava Razbash, Earo Wang, Forecasting Functions for Time Series and Linear Models.R studio cheatsheet.
7) https://stats.stackexchange.com/questions/72243/holts-linear-and-holt-winters-in
-