TIME SERIES MODELING AND FORECASTING OF TEA PRICES IN INDIA

By

Deenamol Joy

(2019-19-003)

Department of Agricultural Statistics

COLLEGE OF AGRICULTURE

VELLANIKKARA, THRISSUR – 680656

KERALA, INDIA

2021

TIME SERIES MODELING AND FORECASTING OF TEA PRICES IN INDIA

By
DEENAMOL JOY
(2019-19-003)

THESIS

Submitted in partial fulfillment of the requirement for the degree of

MASTER OF SCIENCE IN AGRICULTURAL STAISTICS

Faculty of Agriculture

Kerala Agricultural University

Department of Agricultural Statistics

COLLEGE OF AGRICULTURE

VELLANIKKARA, THRISSUR – 680656

KERALA, INDIA

2021

DECLARATION

I hereby declare that this thesis entitled **"Time series modeling and forecasting of tea prices in India"** is a bonafide record of research work done by me during the course of research and that the thesis has not previously formed the basis for the award of any degree, diploma, fellowship or other similar title, of any other University or Society.

Vellanikkara

Date: 13/04/2021

DEENAMOL JOY
(2019-19-003)

CERTIFICATE

Certified that this thesis entitled **"Time series modeling and forecasting of tea prices in India"** is a bonafide record of research work done independently by **Ms. Deenamol Joy (2019-19-003)** under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, fellowship or associateship to her.

Vellanikkara

Date: 13-09-2021

Dr. Laly John C.

(Chairperson, Advisory Committee)

Professor

Department of Agricultural Statistics

College of Agriculture

Vellanikkara

CERTIFICATE

We, the undersigned members of the advisory committee of **Ms. Deenamol Joy (2019-19-003)**, a candidate for the degree of **Master of Science in Agricultural Statistics** with major field in Agricultural Statistics, agree that this thesis entitled **"Time series modeling and forecasting of tea prices in India"** may be submitted by **Ms. Deenamol Joy** in partial fulfillment of the requirement for the degree.

Dr. Laly John C. 13/09/2021
(Chairperson, Advisory Committee)
Professor
Department of Agricultural Statistics
College of Agriculture, Vellanikkara

13/09/2021
Dr. Ajitha T. K.
((Member, Advisory Committee)
Professor and Head
Department of Agricultural Statistics
College of Agriculture
Vellanikkara

13/09/21
Dr. Anil Kuruvila
(Member, Advisory Committee)
Professor
Department of Agricultural Economics
College of Agriculture
Vellanikkara

13/09/2021
Mr. Ayyoob K. C.
(Member, Advisory Committee)
Assistant Professor
Department of Agricultural Statistics
College of Agriculture, Vellanikkara

13/09/21.
Dr. B. Suma
(Member, Advisory Committee)
Professor and Head
Cocoa Research Centre, KAU
Vellanikkara

ACKNOWLEDGEMENT

The preparation of this thesis has been one of the most significant academic challenges I have ever had to face. It marks the end of my journey in gaining my degree in M.Sc. Agricultural Statistics, fulfilling my passion towards the subject. I owe my deepest gratitude to all the personages without whose support, patience and guidance, the study would not have been completed.

First and foremost, I bow my head before the God Almighty for giving me the strength, knowledge, ability and opportunity to undertake this research study and to persevere and complete it successfully.

I would primarily like to express deep gratitude to my esteemed guide ***Dr. Laly John C****. (Professor, Department of Agricultural Statistics) for her expert advice, valuable guidance, practical suggestions, constant patience, inspiring encouragement, friendly approach, kind advice and timely help at various stages of my research work and thesis preparation and will be remembered forever. It was a great privilege and honor to work and study under her guidance.*

I am really fortunate to get exemplary guidance, constant encouragement and compassion from ***Dr. Ajitha T.K.*** *(Professor and Head, Department of Agricultural Statistics)**,** so great that even my most profound gratitude is not enough.*

I extend my special thanks to ***Dr. Anil Kuruvila*** *(Professor, Department of Agricultural Economics), who is a member of my advisory committee, for his advices and suggestions that have served as the major contributions towards the completion of the project.*

I also take this opportunity to extend sincere thanks to ***Mr. Ayyoob K. C.*** *(Assistant Professor, Department of Agricultural Statistics) who is also an advisory committee member, for his guidance and useful critiques that helped me throughout my research.*

I wish to express my deep sagacity of gratitude to advisory committee member, ***Dr. B Suma.*** *(Professor and Head, Cocoa Research Centre) for the valuable suggestions and constant encouragement given throughout the research period.*

I owe a great deal of appreciation and gratitude to thank ***Mrs. Dayana David*** *(Assistant Professor, Department of Agricultural Statistics)* ***Mrs. Sajitha Vijayan M.*** *(Assistant Professor, Department of Agricultural Statistics) and* ***Mr. Vishnu B. R.*** *for their immense support and technical help lent during urgent and needy times.*

I also thank my dearest friends, ***Sisira, Anjana, Pooja*** *and beloved senior* ***Pooja chechi,*** *and staff of the department* ***Bindhu chechi*** *for their valuable support.*

I would also like to greatly thank some people towards the final days of my research tenure who immensely supported me both physically and mentally and to whom I am greatly indebted, my dear friends, ***Vaysagh, Prasad, Praseeda, Praise and Shifin***

I am extremely grateful to my beloved father ***Mr. Joy*** *and mother* ***Mrs. Lissy*** *for their love, prayers, caring and sacrifices for educating and preparing me for my future. Profound affection and love to my sister,* ***Mrs. Chily*** *who serves as a great support to me, who keeps on encouraging me to always make the best out of everything. I also thank my aunty* ***Ms. Shiny****, my dear brother* ***Mr. Johnson*** *for their practical and emotional support and my dearest* ***Hazel*** *who always make me smile during my tough days.*

I acknowledge the whole-hearted co-operation and gracious help rendered by each and every member of the ***College of Agriculture*** *during the period of study.*

For the whole journey, my head bows to Kerala Agricultural University for supporting me financially by provision of scholarship and letting my dreams come true…

Deenamol Joy.

CONTENTS

Abbreviations

A-Pr-Pd	Area, Production and Productivity
NI	North India
SI	South India
ACF	Autocorrelation Function
AIC	Akaike Information Criterion
ARIMA	Autoregressive Integrated Moving Average
CAGR	Compound Annual Growth Rate
CMA	Centred Moving Average
DES	Double Exponential Smoothing
HWMS	Holt-Winters' Multiplicative Seasonal
LCL	Lower Confidence Limit
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSE	Mean Square Error
PACF	Partial Autocorrelation Function
R^2	R-Square
RMSE	Root Mean Square Error
SARIMA	Seasonal Autoregressive Integrated Moving Average
SIC	Schwartz Information Criteria
SES	Single Exponential Smoothing
UCL	Upper Confidence Limit

LIST OF TABLES

LIST OF TABLES (Contd.)

LIST OF TABLES (Contd.)

LIST OF FIGURES

Introduction

1. INTRODUCTION

Tea, the "Queen of beverages" is one of the most popular and widely consumed hot beverage worldwide. It is produced in more than forty-five countries in the world. China, India, Kenya and Sri Lanka are the world's largest tea producers. Among them, India is the largest consumer and second largest producer of tea in the world with a production of around 1390 million kg in the year 2020. Among the sixteen tea growing Indian states, Assam, West Bengal, Tamil Nadu and Kerala constitutes about 96 percent of total tea production. Production in North India (NI) was around 1171 million kg in 2020 from 535629 ha in area under tea, whereas, South India (SI) had a production of 219 million kg and an area of 100928 ha. Assam is the largest tea producing state in India with a production of 507 million kg (Tea Board of India, 2020).

Area under tea (in percent) during 2020 in major states of NI and SI are provided in Fig. 1.1 and 1.2 respectively. From figures it could be observed that, Assam is the largest tea-growing state in NI, accounting for 63 percent of the overall area, while Tamil Nadu accounts for 62 percent of the total tea-growing area in SI (Tea Board of India, 2020).

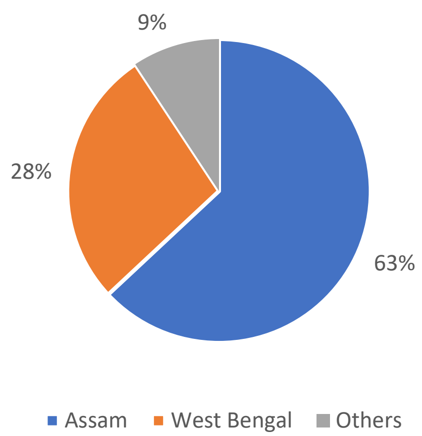

Figure1.1 Statewise area (%) under tea in NI

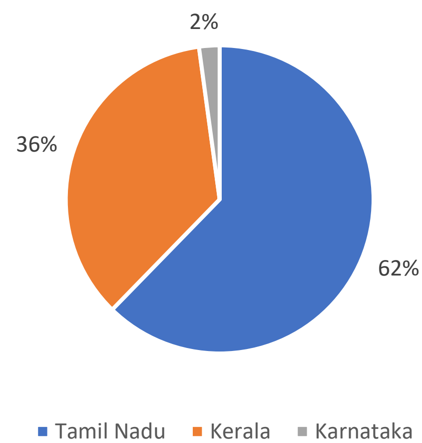

Figure 1.2 Statewise area (%) under tea in SI

Both black and green teas are manufactured in India. CTC (Crush, Tear, Curl) and orthodox are black tea which constitute almost 98 percent of the total production and the rest two percent is green tea. However, CTC has dominance (more than 90%) over orthodox black tea. Total production of CTC was 52.13 million kg in NI and 16.34 million kg in SI, while, orthodox tea had a production of 0.94 million kg in NI and 3.4 million kg in SI (Tea Board of India, 2020). Around 80 percent of the total tea produced in India is consumed by the domestic population.

Tea industry plays an important role in national economy. Price fluctuation is a serious problem for tea producers and has its impact on the economy of the nation. Time series analysis of tea prices will help to reveal the price pattern within and between the years and price forecast will help in formulating proper price policy. A time series is an ordered sequence of values of a variable at equally spaced discrete time intervals. Time intervals can be weeks, months, quarters, years etc. Time series analysis has mainly two objectives viz., to forecast the future values based on past and to study different components like, trend, seasonal variation, cyclic variation and irregular variation.

Forecast models for time series data are fitted on the assumption that some aspects of the pattern of past observations will persist in the future. Autoregressive Integrated Moving Average (ARIMA) models (Box and Jenkins,1970) and Exponential smoothing models are widely used time series forecast models. Exponential smoothing method is an extension of moving average method in which forecasting is done by weighted moving average. As the number of observations becomes older, exponentially decreasing weights are used. They are of different types viz., Single Exponential Smoothing (SES), Double Exponential Smoothing (DES) and Holt-Winters' Exponential Smoothing models. Based on the nature of time series data, an appropriate forecast model may be chosen from among these models. Forecasting of tea price help tea producers to sell or retain their produce and the government for making proper price policy.

Price fluctuation of a commodity is described as price volatility. Price volatility in tea price is a major problem and it can be studied using the statistical model, Auto Regressive Conditional Heteroskedasticity (ARCH), developed by Engle in 1988.

The study of integration between international and Indian tea markets is very important for making price policy. India is mainly in competition with Kenyan and Sri Lankan tea markets. The degree to which consumers and producers would benefit depends on how Indian markets are integrated with international markets and how different Indian domestic markets are integrated with each other. As markets become integrated, it is expected that each market employs information from other markets while forming its own price expectation. The price transmission from the international to domestic markets as well as between domestic markets can be studied by using cointegration analysis.

High cost of production and price fluctuation are the significant issues concerning the Indian tea industry and can cause the country's tea plantation prospects in risk. In this background, the study entitled “Time series modeling and forecasting of tea prices in India” was conducted with the following objectives:

- To analyse the components of time series data on prices of tea in India
- To develop time series forecast models for the tea prices
- To develop statistical models for price volatility
- To study the integration between international and Indian tea markets

Review of Literature

2. REVIEW OF LITERATURE

A comprehensive examination of previous literature is helpful in developing concepts, methodologies, and analysis tools for any research project. The reports of time series analysis of tea price are scanty in literature. Reviews pertaining to other commodities are made and are presented in this section under various headings.

2.1 Trend analysis

2.2 Decomposition of time series components

2.3 Times series forecast models

2.4 Comparison of time series forecast models

2.5 Price volatility

2.6 Cointegration analysis

2.1 Trend analysis

Darvishi and Indira (2013) analysed the changing pattern in A- Pr- Pd of coffee and tea in India for the periods, 1970-71 to 1980-90 and 1990-91 to 2009-10. The results showed that area under coffee and tea showed an increase during the post liberalization period. Similarly, production and productivity also increased during the post liberalization period and the variability has declined.

Kumar et al. (2017) overviewed the growth and instability in A- Pr- Pd of Cassava in Kerala for a 26 years period from 1991 to 2016. The trend estimated using semi log function revealed that there is a decline in area and production.

From the study on Indian tea based on the secondary data from 1971 to 2016, Talukdar and Hazarika (2017) found that the area of tea increased from 3.57 lakh ha in 1971 to 5.67 lakh ha in 2016 and the productivity increased from 1221.56 kg ha^{-1} in 1971 to 2186 kg ha^{-1} in 2016.

Shana (2018) studied the A- Pr- Pd of banana in India and of Nendran banana in Kerala from 1980 to 2018. An increase was identified for area and production in Kerala but the productivity was found decreasing because the increase in area was more than the increase in production in the latter half.

A study was conducted by Nain et al. (2019) on instability and trend in A- Pr- Pd of rice crop in Haryana and India for a period of 47 years from 1966 to 2013 and observed a positive trend in area and production. An increase in area mainly contributed to the increase in production.

Trends in A- Pr- Pd of coffee across the major coffee growing states in India like Kerala, Karnataka and Tamil Nadu for 18 years period from 2000-01 to 2017-18 were studied by Babu et al. (2019). The analysis revealed that the planted area under coffee has shown an upward trend from 3.47 lakh ha in 2000-01 to 4.55 lakh ha in 2017-18 with the compound growth rate of 1.57 percent.

2.2 Decomposition of time series components

Adanaliglu and Yercanm (2012) fitted seasonal indices for tomato prices in Antalya and Turkey. The analysis based on 10 years data from January 2000 to December 2010 revealed a high price in September and October, and, a low price in January.

Kumaraswamy and Sekar (2014) studied the seasonal changes of potato price in major markets in India and Tamil Nadu based on 23 years data on prices of potato. A high price was noticed during May to November.

Indraji (2016) studied seasonality in coconut oil, coconut and copra price at Kochi, Alappuzha and Kozhikode markets of Kerala based on monthly data from January, 1980 to December, 2015. Price of coconut oil followed similar pattern in all the three markets. The study revealed that the price of copra increases from June to August, whereas the price of coconut increased from August to September.

Shana (2018) studied the price behaviour of Nendran bananas in three major Kerala markets: Kozhikode, Ernakulam, and Thiruvananthapuram, over a 16-year period (2003-2018) by decomposing monthly price data into four components: secular trend, seasonal variation, cyclical variation, and irregular variation, using a multiplicative model. Prices in all the three markets showed increasing trend.

Based on monthly price data of coconut in major markets of Kerala for two periods viz., from 1980-81 to 1995-96 and from 1996-97 to 2015-16, Preethi (2018) segregated secular trend, seasonal, cyclical and irregular components. Study showed that despite the high price fluctuations, coconut price in the long run exhibited an increasing trend. The seasonal variations analysed using ratio to moving average method revealed that behaviour of coconut prices in Alappuzha and Kozhikode markets was distinctly different, presumably due to the dissimilarity in the pattern of market arrival of coconut. More than two cycles were observed with varying lengths over time and irregular variations were found highly unpredictable and not following any uniform pattern over the period.

Sutradhar et al. (2019) studied the seasonal and cyclical variations in domestic and international prices of natural rubber in India for a period of 18 years from 2000-01 to 2017-18. In both domestic and international markets, it was shown that prices were lowest in January and November and highest in July and June. It was also discovered that the domestic and international markets for natural rubber had an irregular cycle.

2.3 Times series forecast models

2.3.1 Exponential smoothing models

Rangoda et al. (2006) used monthly data from January 1974 to December 2014 to examine the price behaviour of dried coconut, coconut oil, and fresh coconut. They compared the accuracy of several models like moving average model, Winters model, SES, DES and ARIMA model to forecast coconut oil price and related products in Sri

Lanka. ARIMA and exponential smoothing models were found better than other models.

Jyothi (2011) examined the price behaviour of turmeric in Nizamabad and Erode markets from 1980 to 2009 using various models such as SES, DES, HWMS, moving average and ARIMA models. Among all the models, DES and HWMS model was found to be suitable for Nizamabad and Erode markets respectively.

Ishaque and Ziblim (2013) forecasted the prices of some important food crops in the upper east region of Ghana using the data from January 1992 to December 2000. Results from the study revealed that DES model performed better in cereal crops in which trend was present, than HWMS model.

Vasanthakumar et al. (2014) used exponential smoothing model for forecasting the price of different types of teak based on monthly price data from May 1996 to May 2011 and SES model was identified for forecasting. Ex-post and ex-ante predictions of certain teak classes were made and compared to actual prices. From May 2011, forecasts were given for the next four months.

Booranawong and Booranawong (2017) forecasted monthly lime prices in Thailand using the data from January 2011 to December 2015. Results indicated that double exponential smoothing model shows better forecasting performances and gave the smallest forecasting error measured by Mean Absolute Percentage Error (MAPE).

2.3.2. ARIMA model

Kumar et al. (2011) forecasted potato price in Bangalore market using the data from April 1999 to March 2008. ARIMA $(0,1,0)(0,1,1)_{12}$ with MAPE of 18.28% was found to be the best fitted model.

Barathi et al. (2011) forecasted cocoon price in Sidlaghatta and Ramanagaram markets of Karnataka using ARIMA model. ARIMA $(0,1,0)(1,0,1)_{12}$ was shown to be suitable for both markets for a 10-year period from 1998-99 to 2007-08.

Adanaliglu and Yercanm (2012) developed seasonal ARIMA model for tomato price forecasting using monthly price from 2000 to 2010 and SARIMA (1,0,0) $(1,1,1)_{12}$ was found to be the suitable fit.

Krishnarani (2013) fitted seasonal ARIMA model for tea price during the period from January 2006 to July 2011. ARIMA (1,0,1) $(1,0,1)_{12}$, ARIMA (1,0,0) and ARIMA (1,1,1) $(1,0,0)_{12}$ were selected for tea price in NI, SI and All India tea prices respectively.

Sharma et al. (2014) identified ARIMA (1,1,1) model as the appropriate model to forecast soyabean prices in Kota market of Rajasthan, based on monthly price data for 12 years period from 2000-01 to 2011-12.

In a study carried out by Chaudhari and Iingre (2014), the ARIMA model was used to forecast green gram prices in Maharashtra using time series data from the Akola market's monthly average prices from January 2001 to December 2012. The best-fitting model was found to be ARIMA (0,1,0). The reliability of the model was tested using R^2, MAPE and Bayesian Information Criteria (BIC).

Naidu et al. (2014) obtained ARIMA (1,1,1)$(1,1,1)_{12}$ as the best model for forecasting price of red chilli in Khammam and Warangal markets of Andhra Pradesh based on the monthly price data from April 2002 to September 2013.

Vinayak et al. (2015) conducted a study to forecast the prices of onion at Hubli market of Northern Karnataka. The time series data on monthly prices of onion was collected from 1996-97 to 2010-11. Based on the Akaike Information Criteria and Swarz Bayesian Criteria, ARIMA (1,1,1) $(2,1,1)_{12}$ was found to fit the series suitably.

Nyantakyi et al. (2015) examined the individual behaviour of the prices of tea, rubber and coconut in Sri Lanka using ARIMA. They concluded that ARIMA (0,1,3), ARIMA (0,1,0) and ARIMA (3,1,3) were the best fitted models for the annual prices of coconut, tea and rubber from 1996 to 2009 respectively.

Guha and Bandyopadhyay (2016) used ARIMA model to forecast the future gold price in India. Based on monthly data form November 2003 to January 2014, ARIMA (1,1,1,) was selected in predicting the future gold prices.

Verma et al. (2016) forecasted the coriander prices of Rajasthan using monthly wholesale prices for the period from May 2003 to June 2015 of Ramganj Mandi Rajasthan. They used ARIMA models for forecasting and ARIMA (0,1,1) found to fit the series suitably.

Darekar et al. (2016) forecasted the price of onion in Kolhapur market of Western Maharashtra using monthly data from year 2004 to 2013. ARIMA model (1,1,1) was found fit the data suitably.

Liu and Shao (2016) used ARMA model to analyse weekly tea auction price of India from 2013 to 2014. ARMA (1,1) was identified as the best model as it had minimum AIC and BIC.

Senaviratna (2016) forecasted auction prices of tea in Sri Lanka over the period 1996 to 2016 using Box-Jenkins modeling approach and SARIMA(1,0,0)$(0,1,0)_{12}$ was selected.

Shukla et al. (2017) analysed the monthly auction price of tea for three years from January 2012 to December 2014 for NI, SI and All India and got ARIMA (1,0,1) as the best forecast model.

Jadhav et al. (2017) fitted ARIMA model for forecasting prices of major crops like paddy, ragi and maize in Karnataka using the monthly price data from 2007 to 2016. ARIMA (1,1,1), ARIMA (1,1,2) and ARIMA (1,2,1) were adequate for paddy, ragi and maize respectively.

Darekar and Reddy (2017) conducted a study to forecast the paddy prices in India using the data from January 2006 to December 2016. It was found that seasonal ARIMA (1,1,1) $(0,0,2)_{12}$, ARIMA (0,1,1), seasonal ARIMA (1,1,1) $(0,0,1)_{12}$, ARIMA (2,1,1), seasonal ARIMA (0,1,0) $(0,0,2)_{12}$, seasonal ARIMA (0,1,0) $(0,0,1)_{12}$ model were suitable fit for paddy price data of Punjab, West Bengal, Uttar Pradesh, Andhra Pradesh, Tamil Nadu and India respectively.

Darekar and Reddy (2017) fitted forecast model for cotton price of major cotton producing states of India, using monthly data for a 10 years period from 2006 to 2016. Among ARIMA models tried, ARIMA (1,1,3), ARIMA (1,1,1), ARIMA (1,1,2), ARIMA (0,0,2), ARIMA (2,1,2), ARIMA (1,1,2) and ARIMA (0,1,0) model was selected for Gujarat, Maharashtra, Karnataka, Andhra Pradesh, Haryana and India respectively.

A study was conducted by Reddy (2018) to forecast the tomato price in India using the monthly data from January 2006 to December 2016. The best fit model was selected based on BIC, MAPE, RMSE and MAE. The model ARIMA (2,0,0) $(1,1,0)_{12}$ was found to fit the series suitably.

KumarMahto et al. (2019) applied ARIMA model to forecast sunflower seed price, based on time series data from January 2011 to December 2016. ARIMA (1,1,2) was identified as the most suitable model compared to all other models.

Rotich et al. (2020) carried out a study to estimate the behaviour of Kenyan tea auction prices by including rainfall patterns into the series. The study was conducted using secondary data from the Mombasa Tea Auction Centre from 2009 to 2018. In comparison to VAR (2) model, the Seasonal ARIMA (2,0,1) (0,0,1)12 model was found to fit the series suitably

2.4 Comparison of time series forecast models

Reeja (2011) conducted a study to forecast the price of natural rubber (RSS-4) at Kottayam and Bangkok markets. Various ARIMA models and Artificial Neural Network model were tried and SARIMA (0,1,0) $(1,0,1)_{12}$ was found to be the best fit in Bangkok market.

Sharma and Burark (2015) forecasted sorghum price in the Ajmer market in Rajasthan and ARIMA (1,1,2) model was preferred over other models due to minimum MAPE and Mean Square Error (MSE).

Indraji (2016) used ARIMA models, exponential smoothing models and ANN models to forecast prices of coconut, coconut oil and copra for three markets of Kerala. The study used monthly time series data from 1980 to 2015. It was found that for price of coconut oil at Alappuzha and Kochi markets, Holts-Winters Multiplicative Seasonal model was good. At Kozhikode market, both SARIMA $(1,1,1)(1,0,1)_{12}$ and Holts-Winters Multiplicative Seasonal model were suitable. For all markets, Holt Winters Multiplicative Seasonal model was identified as the best for copra price. At Alappuzha market, ARIMA (0,1,1) was identified as the suitable fit for forecasting coconut price.

Mgale et al. (2021) compared ARIMA and Holt Winters exponential smoothing models for rice price forecasting in Tanzania. When compared to the ARIMA model, Holt Winters exponential smoothing model had the best results for forecasting rice prices.

2.5 Price volatility

Sundaramoorthy et al. (2013) studied volatility in edible oil sector in India from January 2001 to December 2010. For the price series of groundnut and mustard oil in the marketplaces of Hyderabad, Rajkot, and Kanpur, the GARCH (1,1) model was chosen as the most appropriate model. The results revealed that a persistent volatility is prevalent in the markets of Hyderabad for groundnut, while the volatility is less pronounced in the centres like Rajkot (groundnut) and Kanpur (mustard oil).

The price volatility of cereals and pulses in Amhara National Regional State over the period from December 2001 to June 2012 was studied using GARCH models by Kelkay and Yohannes (2014). Among several models tried, ARMA (1,0)-EGARCH (3,4) for wheat, ARMA (4,4)-EGARCH (2,3) for bean and ARMA (1,0)-EGARCH (1, 2) for pea were chosen to be the best fit models.

Sabu (2013) fitted GARCH (1,1) model for the nominal and real prices of black pepper in domestic and international markets by dividing the monthly price of black pepper from January 1980 to December 2018 into two periods, pre-WTO (1980-95) and post WTO (1996-2018). The estimates of the GARCH (1,1) model for nominal pepper prices in the domestic market were found to be significant in the post-WTO period, but volatility in international markets has declined. For real prices of black pepper, volatility was found to be very high in domestic market and low or medium in international markets in the post WTO period.

Bhavani et al. (2015) analysed the price fluctuation in the domestic markets of chilli, using ARCH and GARCH models for the period from 1997 to 2011. A persistent fluctuation was observed in all the markets and it was found to be intense in Nagpur market.

Bodade et al. (2017) evaluated soyabean price movement across major markets of Madhya Pradesh using ARCH and GARCH model. The results revealed a high price fluctuation in Dewas market, whereas, fluctuation was absent in all other markets.

Dudhat (2017) analysed the price volatility of groundnut in major domestic markets of Andhra Pradesh, Gujarat, and Tamil Nadu based on the data from 1996 to 2016. Price volatility was examined by using GARCH model and the Gaussian GARCH model was found to be the best. Higher persistence of volatility was ascertained in Rajkot market (Gujarat) compared to other two markets.

Cermak et al. (2017) carried out a study to determine and forecast the volatility of wheat prices based on weekly time series data from 2005 to 2015 using stochastic models of conditional heteroskedasticity. GARCH (1,1) model was found to be the appropriate model for volatility of wheat price.

Rahmawati et al. (2019) analysed the Value at Risk (VaR) based on the GARCH family volatility model for harvested dry grain of Pemalang district based on weekly price from August 2015 to July 2018. The result showed MA (1)-GARCH (1,1) model as the appropriate one.

2.6 Cointegration analysis

Soe and Fukuda (2010) analysed the market integration of oil seed market in Myanmar from January 2002 to 2007. The markets in the producing area were highly integrated in the long run.

Reddy (2012) assessed the market integration of chickpea in India based on monthly prices from 2003 to 2010 in twelve markets in NI. According to the study, only three markets out of twelve are cointegrated, suggesting a weak integration of chickpea markets in India

To analyse the market integration of apple in India, Wani et al. (2015) carried out a cointegration analysis on three commercial varieties of apple based on weekly price from September 2005 to February 2015. The result revealed that markets were perfectly integrated and Delhi markets are the dominant ones. They also conducted Granger casualty test and found that there are 39 and 18 bidirectional and unidirectional causations respectively under different market situations.

Sabu (2013) conducted a cointegration analysis of domestic and international markets of black pepper in pre- WTO, post-WTO and overall period using monthly data from 1980-2014. Results of pairwise cointegration using domestic prices viz., Cochin Malabar garbles, Cochin ungarbles and Calicut nadan showed that each pair has a single cointegrating relationship. The study also showed that there was price co-movement between domestic and international markets of black pepper even in the pre-WTO.

Price transmission and integration of major pulses in India was studied by Paul et al. (2016) using Johansens cointegration and Vector Error Correction Model (VECM). It was observed that for the major pulses, both wholesale and retail prices showed a strong cointegration and VECM exhibited that disequilibrium got corrected thus restoring the equilibrium situation

A cointegration analysis was carried out by Rani et al. (2017) to analyse the integration of major maize markets based on secondary data of monthly prices of maize from January 2005 to May 2016. The Johannsen test showed two cointegrating equations, indicating that all of the markets were well integrated.

Lavanya et al. (2018) performed the analysis for the market integration between domestic and international market prices of beverage crops like tea, coffee arabica, coffee robusta and cocoa. The study used monthly average price from January 2011 to December 2017 for coffee arabica and cocoa whereas for tea and coffee robusta the data from January 2013 to December 2017 were used. Results revealed that tea and coffee had linkage between domestic and international markets, whereas, cocoa remained independent.

Materials and Methods

3. MATERIALS AND METHODS

The materials and statistical methods used in the study "Time series modeling and forecasting of tea prices in India" have been discussed under various titles in the following sections.

3.1. Data used for the study

3.2. Trend analysis

3.3. Compound Annual Growth Rate (CAGR)

3.4. Decomposition of time series on tea price

3.5. Forecast models for tea prices

3.6. Volatility of tea prices

3.7. Cointegration Analysis

3.1 Data used for the study

Data used for the present study are provided in Table 3.1.

Table 3.1 Data used for the study

Sl. No.	Database	Period	Source
1	Monthly auction price of tea for NI, SI and All India	January 1980-December 2020	Tea Board, India
2	International tea prices for Colombo (Sri Lanka) and Mombasa (Kenya)	January 1980-December 2020	World Bank Commodity Price Data (Pink Sheet)
3	A- Pr- Pd of tea for NI, SI and All India	1970-2019	Tea Board, India

Time series data of monthly auction prices of tea for NI, SI and All India for the period from January 1980 to December 2020 were decomposed to analyse the four components. Time series data on monthly auction prices of tea for NI and SI were used to develop price forecast models and price volatility models. International prices of tea for Colombo (Sri Lanka) and Mombasa (Kenya) for the period from January 1980 to December 2020 were used to analyse the cointegration between domestic and international tea markets. Annual data on A- Pr- Pd of tea from 1970 to 2019 in North India, South India and All India were analysed to have an idea about their trend and compound annual growth rate.

3.2 Trend analysis

The trend in A- Pr- Pd of tea in the long run (1970 to 2019) in NI, SI and All India were studied by fitting suitable model as per Draper and Smith (1998). The following models were fitted:

Linear trend: $Y_t = a + bt + e_t$

Quadratic trend: $Y_t = a + bt + ct^2 + e_t$

Cubic trend: $Y_t = a + bt + ct^2 + dt^3 + e_t$

Exponential trend: $Y_t = ab^t$

t=1970 to 2019 and $e_t \sim N(0, \sigma^2)$

From the different models, suitable one was selected based on MAPE and Adjusted R^2.

MAPE is given by,

$$\text{MAPE} = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - \widehat{Y_t}|}{Y_t} * 100$$

where,

Y_t-Actual area/production/productivity at time t

$\widehat{Y}_t$-Estimated area/production/productivity

n- Number of observations (50 years from 1970 to 2019)

Adjusted R^2 is given by,

$$\text{Adjusted } R^2 = 1-(1-R^2)\left(\frac{n-1}{n-k-1}\right)$$

where,

$$R^2 = \frac{SS_R}{SS_T},\ 0 < R^2 < 1$$

SS_R - Regression sum of squares

SS_T - Total sum of squares

k -Number of independent variables

The model having least MAPE value and high Adjusted R^2 was selected.

3.3 Compound Annual Growth Rate (CAGR)

Compound Annual Growth Rate of a variable is defined as the rate of change per unit time, usually a year. The growth in A- Pr- Pd of tea in NI and SI from 1970 to 2019 was calculated using the following formula,

$$Y_t = ab^t$$

where,

Y_t : Area/production/productivity of tea

a : Intercept

b : Regression coefficient

t : Year, 1970 to 2019

By taking logarithm on both sides of the equation, it is reduced to the following linear form:

$$\ln Y_t = \ln a + t \ln b$$

$$y_t = A + Bt$$

where,

y_t : $\ln Y_t$ where ln is the natural logarithm

A : ln a

B : ln b

The regression coefficient (b) can be calculated using the Ordinary Least Squares (OLS) method. Percentage rate of compound growth per annum was calculated as,

Compound Annual Growth Rate (CAGR) = (b – 1) * 100

where, b=Antilog(B)

3.4 Decomposition of time series on tea price

The various factors at work that affect the variables in a time series can be divided into four categories called time series components. The four components are trend, seasonal variation, cyclic variation and irregular variation. The four components are explained below:

Trend

The general tendency of a time series to increase or decrease over a long period of time is known as trend. The increase or decrease does not have to be in the same direction throughout the time period. In various time periods, distinct patterns of rise, decline, or stability can be observed. However, the general trend might be upward, downward, or stable.

Seasonal variation

Seasonal variations in a time series are caused by rhythmic factors that work in a regular and periodic manner for a period of less than one year. It can be studied only if data are recorded quarterly, daily, weekly or monthly etc. For monthly tea price data, a month is a season. Thus, the months from January to December are the12 seasons.

Cyclic variation

Cyclic variations are the oscillatory movements in a time series with period of oscillation of more than one year. Cyclic fluctuation, though more or less regular, are not uniformly periodic. A cycle may be in the range of 2-11 years.

Irregular variation

Irregular variations in a time series are random or irregular fluctuations that are not explained by trend, seasonal variation and cyclic variation. These changes are completely random, erratic, and unexpected, and are caused by non-recurring and irregular situations beyond human control, such as floods, wars, and earthquakes.

Time series data on monthly tea prices in NI, SI and All India, from January 1980 to December 2020 (492 months) were decomposed into different components using a multiplicative model.

3.4.1 Decomposition model

For decomposition of the tea price data into four components, a multiplicative model was assumed. Let X_t be the price data at time t. Then the multiplicative model is given by,

$$X_t = T_t * S_t * C_t * I_t \quad \rightarrow (1)$$

T_t, S_t, C_t and I_t are the components of time series, where,

T_t: Trend at time t

S_t: Seasonal variation at time t

C_t: Cyclic variation at time t

I_t: Irregular variation at time t

t: January 1980 to December 2020 (492 months)

All the four components were estimated for tea prices in NI, SI and All India using R software and represented graphically.

3.4.2 Estimation of seasonal indices

Seasonal indices were calculated for 12 months (January to December) from the monthly price data of tea in NI and SI from January,1980 to December,2020 to understand the seasonal behaviour of tea prices and to make a comparison among the two regions.

Assuming multiplicative model, $X_t = T_t * S_t * C_t * I_t$ (Equation (1)), seasonal indices were estimated using ratio to moving average method.

The steps involved in ratio to moving average method are given below:

- Calculate the centred 12 month moving average (CMA) of time series data, X_t. CMA values will give the estimates of combined effect of trend and cyclic variation.

  i.e., $CMA = T_t * C_t$
- Express the original data as the percentage of the CMA values.
- This percentage will represent the seasonal (S_t) and irregular components (I_t).

  i.e., $\frac{X_t}{CMA} * 100 = S_t * I_t$
- By averaging the above percentages over years, the irregular components will get eliminated. The resultant value will be preliminary seasonal indices, S.
- The sum of preliminary seasonal indices (S) may not be equal to 1200. Adjust the value by multiplying throughout by the factor $\frac{1200}{S}$. The result will give the seasonal indices for 12 months from January to December.

Seasonal indices of the tea price in NI and SI were plotted against corresponding months. The resultant graph is called seasonal plot.

3.5 Forecast models for tea prices

Time series forecasting is done to estimate how the sequence of observations will continue in the future. A model describing the underlying pattern of the time series is developed by analysing the past observations of the same variable. The test data used for the present study was the monthly tea price of NI and SI from January 1980 to December 2020. Single, Holt's linear (Double), Holt-Winters' additive and multiplicative exponential smoothing models, Auto Regressive Integrated Moving Average (ARIMA) model and Seasonal ARIMA model were fitted to forecast the tea prices. Prices were forecasted for next four months from January 2021 to April 2021 for both NI and SI.

3.5.1 Exponential Smoothing Models

Forecasting using weighted moving average is an obvious extension of the moving average method. In Exponential Smoothing models, exponentially decreasing weights are attached as the observations gets older. The different exponential smoothing models used were,

1. Single Exponential Smoothing model (SES)

2. Double Exponential Smoothing model (DES)

3. Holt-Winters' exponential smoothing model

All the models mentioned above share the property that recent values are given a higher weighting in forecasting than older observations. The weights allocated to the observations in moving averages are a by-product of the particular moving average method used. One or more smoothing parameters must be determined explicitly in exponential smoothing, and those parameters provide the weights assigned to the observations.

3.5.1.1 Single Exponential Smoothing model (SES)

SES is a forecasting method to make short-term forecasts for a time series without trend or seasonality. It constantly repeats enumeration using the most recent data. Let X_t be the actual observation and F_t denote the forecast of the time series at time t. The forecast error is found to be (X_t - F_t). The method of single exponential smoothing takes the forecast for the previous period and adjusts it using the forecast error. Thus, forecast F_{t+1} for the next period is,

$$F_{t+1} = F_t + \alpha\,(X_t - F_t)$$

$$= \alpha\, X_t + (1 - \alpha)\, F_t$$

where, α is the smoothing constant with value ranges from 0 to 1. A high α value (say 0.9) provide a very little smoothing in the forecast, whereas, a small α value (say 0.1) gives considerable smoothing. When α =1, exponential smoothing is equivalent to using the last observation as a forecast. From a grid of values for α (e.g.; α = 0.1, 0.2, …, 0.9), the values that gives least MAPE was selected.

3.5.1.2 Double Exponential Smoothing model (DES)

DES approach may be used to make short-term forecasts for a time series with an increasing or decreasing trend but no seasonality. The level and slope at the current time point can be estimated using DES model. Smoothing is controlled by two parameters, α, for the estimate of the level at the current time point, and β for the estimate of trend component at the current time point. The equations are given as,

$$\text{Level: } L_t = \alpha\, X_t + (1-\alpha)(L_{t-1} + b_{t-1})$$

$$\text{Trend: } b_t = \beta\,(L_t - L_{t-1}) + (1 - \beta) b_{t-1}$$

$$\text{Forecast: } F_{t+m} = L_t + b_t m$$

where,

L_t : level of the time series at time t,

b_t : estimate of trend (slope) of the time series at time t

α and β : smoothing constants, 0< α <1; 0< β <1

F_{t+m} : forecast for m periods ahead of t

The combination of α and β which provides the least MAPE was selected.

3.5.1.3 Holt-Winters' Exponential Smoothing model

Holt's method was extended by Winters (1960) to capture seasonality directly. The Holt-Winters' method is based on three smoothing equations, one for level, one for trend and one for seasonality. There are two models under Holt-Winters' method, depending on whether seasonality is modelled in an additive or multiplicative way: Holt-Winters' Additive Seasonal (HWAS) model and Holt-Winters' Multiplicative Seasonal (HWMS) model. For the present study, model used was HWMS.

The equations for HWMS model:

$$\text{Level: } L_t = \alpha \frac{X_t}{S_{t-12}} + (1-\alpha)(L_{t-1} + b_{t-1})$$

$$\text{Trend: } b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1}$$

$$\text{Seasonality: } S_t = \gamma \frac{X_t}{L_t} + (1-\gamma)S_{t-s}$$

$$\text{Forecast: } F_{t+m} = (L_t + b_t m)S_{t-s+m}$$

where, L_t is the level of the time series at time t, b_t is the estimate of trend (slope) of the time series at time t. S_t is the seasonal component at time t. For monthly data, s=12. F_{t+m} is the forecast for m periods ahead of t. α, β and γ are smoothing constants, each taking values between 0 and 1. The combination of α, β and γ which yields minimum value for MAPE was chosen.

From among different exponential smoothing models fitted for the prices, the best model was selected based on closeness of actual and fitted price plots and MAPE. Exponential smoothing models for the tea prices in NI and SI was fitted using the Minitab package.

The computational formula for MAPE:

$$\text{MAPE} = \frac{1}{492} \sum_{i=1}^{492} \frac{|X_t - \hat{X}_t|}{X_t} * 100$$

where,

X_t = Price at time t

$\hat{X}_t$ = Estimated price from the model at time t

3.5.2 Auto Regressive Integrated Moving Average (ARIMA) model

In early 1970's, George Box and Gwilym Jenkins developed and popularized ARIMA methodology for time series forecasting. Hence it is also called Box-Jenkins ARIMA methodology. The model is based on the assumption that the time series is stationary.

ARIMA models are of two types, Non-seasonal ARIMA and Seasonal ARIMA (SARIMA) models.

Non seasonal ARIMA model

Non seasonal ARIMA model is denoted by ARIMA (p, d, q), where,

p: order of auto-regression

d: order of integration (differencing)

q: order of moving average

AR (p) model is given by

$$X_t = \mu + \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} + \varepsilon_t$$

MA (q) model is given by

$$X_t = \mu - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

A stationary ARMA (p, q) process is defined by the equation

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \cdots + \varphi_p X_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} + \varepsilon_t$$

i.e., $$\left(\mathbf{1} - \sum_{i=1}^{p} \boldsymbol{\varphi}_i \boldsymbol{B}^i\right) X_t = \left(\mathbf{1} - \sum_{j=1}^{q} \boldsymbol{\theta}_j \boldsymbol{B}^j\right) \boldsymbol{\varepsilon}_t$$

$X_t, X_{t-1} \ldots X_{t-p}$ are the values of the time series at times t, t-1, t-2…t-p ;

B is the backshift operator such that $B^i X_t = X_{t-i}$ and $B^j \varepsilon_t = \varepsilon_{t-j}$.

$\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-p}$'s are random errors at times t, t-1, t-2…, t-q; independently and normally distributed with zero mean and constant variance σ^2.

When the time series is non-stationary the ARIMA (p, d, q) model is obtained as

$$\text{i.e.,} \quad \left(\mathbf{1} - \sum_{i=1}^{p} \boldsymbol{\varphi_i B^i}\right)(\mathbf{1} - \boldsymbol{B})^{\boldsymbol{d}} \boldsymbol{X_t} = \left(\mathbf{1} - \sum_{j=1}^{q} \boldsymbol{\theta_j B^j}\right) \boldsymbol{\varepsilon_t}$$

where,

ϕ_i, i=1, 2…p are Auto Regressive (AR) parameters

θ_j, j= 1, 2…q are Moving Average (MA) parameters

$(1-B)^d X_t$ is the non-seasonal difference of order d on X_t

Seasonal ARIMA Model

Seasonal ARIMA model is used when there is seasonality in the time series data. That is, a pattern repeats itself over a fixed interval of time. If seasonality is present, it can be identified by large autocorrelation and partial autocorrelation at lags 12, 24 etc. for monthly data. A seasonal data which is non-stationary can be made stationary by seasonal differencing. That is, difference between an observation and corresponding observation from the previous year.

SARIMA (p, d, q) (P, D, Q)$_s$ model is defined by,

$$\left(1 - \sum_{i=1}^{p} \varphi_i B^i\right)\left(1 - \sum_{i=1}^{P} \Phi_i B^{is}\right)(1-B)^d (1-B^s)^D X_t = \left(1 - \sum_{j=1}^{q} \theta_j B^{js}\right)\left(1 - \sum_{j=1}^{Q} \Theta_j B^s\right)\varepsilon_t$$

(Non seasonal AR(p)) (Seasonal AR(P)) (Non - seasonal difference) (Seasonal difference) (Non - seasonal MA(q)) (Seasonal MA(Q))

where,

Φ_i, i=1,2…p are the seasonal autoregressive parameters

Θ_j , j=1,2…Q are the seasonal moving average

d, D =order of non-seasonal, seasonal differencing respectively

s= no. of seasons (Since the data was collected on a monthly basis, s =12)

The three main steps in developing an ARIMA model are,

a. Model identification

b. Estimation and testing

c. Forecasting

a. Model identification

The first stage in the identification of ARIMA modeling is to check for the stationarity of the time series. A time series is said to be stationary, if it is stationary in both mean and variance. The presence of non-stationarity can be identified by examining time series plots, ACF and PACF plots. A non-stationary TS can be made stationary in mean by the process of differencing. Logarithmic or power transformation can be applied to make the data stationary in variance. For seasonal data, seasonal difference of the data is required.

Statistical tests to determine the stationarity of a time series are known as unit root tests. The most commonly used unit root test is Augmented Dickey Fuller (ADF) test. ADF test consists of estimating the following regression equation.

$$X_t' = \varphi X_{t-1} + b_1 X_{t-1}{}' + b_2 X_{t-2}{}' + \cdots + b_p X_{t-p}{}'$$

where, $X_t{}' = (X_t - X_{t-1})$

The value of ϕ is estimated from the above regression equation using method of least squares and tested for deviation from unity.

$$H_0: \phi = 1$$

$$H_1: \phi < 1$$

$$t_{\phi=1} = \frac{\hat{\phi}-1}{SE(\hat{\phi})}$$

where, $\hat{\phi}$ is the least square estimate of ϕ.

If $\hat{\phi}$ is negative and significant, the time series is considered stationary, if ϕ =1, time series is non- stationary.

The next step is to find the initial values for the orders of non-seasonal parameters p and q, and seasonal parameters P and Q, which can be obtained by observing the significance of autocorrelation and partial correlation coefficients. One or more models are tentatively chosen that seem to provide statistically adequate representation of the time series data.

b. Estimation and Testing

This step involves the estimation of the parameters p, q, P and Q by least squares as given by Box and Jenkins (1970). Standard statistical package SPSS was used for the estimation, after selecting the tentative model. From among the different models, best model could be selected by using criteria like Low Akaike Information Criteria (AIC) or Schwarz-Bayesian Information Criteria (SBC).

AIC is given by

$$AIC = -2\log L + 2m$$

where,

m=p+ q+ P+ Q is the likelihood function.

-2 log L is approximately equal to $\{n(1+\log 2\pi\} + n\log\sigma^2$; σ^2 is the model Mean Squared Error (MSE),

AIC can be written as

$$AIC = \{n(1+\log 2\pi\} + n\log\sigma^2 + 2m$$

$$SBC = \log\sigma^2 + (m\log n)/n$$

First term is usually omitted, as, it is a constant. Model with smallest AIC/SBC value will be selected and it will have residuals resembling white noise. If most of the sample autocorrelation coefficients of the residuals lies within the limits $\pm \frac{1.96}{\sqrt{n}}$ (n = number of observations upon which the model is based), then the residuals are white noise indicating that the model is a good fit.

Portmanteau test: Ljung-Box test

Non-significance of autocorrelations of residuals can be tested by using Portmanteau test. Ljung-Box Q* statistic is the commonly used Portmanteau test and is given by,

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

where,

h = maximum lag for autocorrelations

n = number of observations in the time series

Ljung-Box Q* statistic has a chi-square distribution with (h-m) degrees of freedom, where m is the number of parameters in the model. If Q* is less than the table value of chi-square at (h-m) degrees of freedom, the errors are white noise and if Q* is greater than the table value, the data are not white noise.

c. Forecasting

Forecasting means predicting the future values of the time series based on the fitted model. Prices of tea in NI and SI up to four months were forecasted based on the selected ARIMA or SARIMA model.

Validation of forecasted price was done by comparing the forecast based on Exponential smoothing models and ARIMA model with the actual price. The model which is in more agreement with the actual price and low MAPE vale was selected as the best forecast model.

3.5.3 Forecast Accuracy Measures

Forecast accuracy measures are used to measure the reliability of a forecasting model. If X_t is the actual observation at time period t and F_t is the forecast for the same period, then forecast error, e_t is defined as

$$e_t = X_t - F_t$$

The forecast F_t is calculated using all the observations X_1, X_2, …, X_{t-1} and X_t is not included in the model. F_t is called one step ahead forecast. Suppose there are n observations and n forecasts, then there will be n error terms, and the following statistical forecast accuracy measures can be used.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |e_t|$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {e_t}^2}$$

Mean Absolute Percentage Error (MAPE):

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \frac{|X_t - F_t|}{X_t} * 100$$

where, n is the number of observations in the time series.

MAPE is a unit free measure and is more reliable for comparison of forecast models. Agreement between actual and fitted price plots, residual ACF and PACF plots and MAPE were used for selecting the best model from among the alternate ARIMA/SARIMA model tried, for the tea prices. SPSS package was used for fitting ARIMA/SARIMA model.

3.6 Volatility of tea prices

The fluctuation or variability of commodity prices over a period of time is termed as price volatility. Volatility of tea prices for NI and SI were estimated using intra and inter annual volatility and its significance was tested by fitting suitable ARCH model.

3.6.1 Intra-annual volatility

Using intra-annual standard deviation of changes in log prices (Gilbert, 2006) and scaled inter-annual range as suggested by Parkinson (1980), Garman and Klass (1980), and Kunitomo (1992), the extent of volatility and temporal shifts in volatility of tea prices in NI and SI were analysed by using a series of annual observations from monthly data.

The intra-annual volatility in monthly prices was estimated as the intra-annual standard deviation of changes in logarithm of prices (lnP), which is defined as

$$S_{YM} = \sqrt{\frac{1}{11}\sum_{m=1}^{12}(\ln P_{y,m} - \ln P_{y,m-1} - \delta y)^2} \text{ for year y,}$$

where,

$\delta y = \frac{1}{12}(\ln P_{y,12} - \ln P_{y,0})$ is the yth year drift

$P_{y,0} = P_{y-1,12}$

This estimate was scaled into annual basis using the factor $\sqrt{12}$.

3.6.2 Inter-annual volatility

Estimation of inter-annual volatility of monthly prices was done by using inter-annual volatility measure or the scaled inter-annual range called as the Parkinson's measure as suggested by Parkinson (1980) and modified by Garman and Klass (1980) and Kunitomo (1992).

Parkinson's measure is defined as,

$$S_y^p = \frac{\ln P_y^H - \ln P_y^L}{2\sqrt{\ln 2}}$$

where,

$P_y^H = \text{Max}_{m-1}^{12} P_{y.m}$, the highest monthly average price in the year, y

$P_y^L = \text{Min}_{m-1}^{12} P_{y,m}$, the lowest monthly average price in the year, y

It is an unbiased estimate of the annual price volatility based on the assumption that the price process follows a random walk.

3.6.3 Significance of price volatility using ARCH model

Volatility indices provide only the magnitude of volatility. It will not provide information on whether or not the estimated volatility is statistically significant. To test the statistical significance of price volatility, ARCH model was fitted for tea prices of NI and SI using monthly time series data from January 1980 to December 2020. In 1982, Engle developed Autoregressive Conditional Heteroscedasticity (ARCH) model.

ARCH (q) model can be written as,

$$h_t = b_0 + \sum_{i=1}^{q} b_i \, u_{t-i}^2$$

where,

h_t : conditional variance at time *t*

b_i : ARCH coefficient; $b_0 > 0$, $0 \leq b_i < 1$, i >0

ARCH (q) model simultaneously examines the mean and variance of a variable. It shows that the variance or volatility in a given period depends on the magnitude of the squared error in the past q periods (u_{t-i}^2). The number of lagged periods of the squared error denotes the structure of ARCH model.

The coefficient b_i should be greater than zero to ensure that the conditional variance is always positive. b_i gives the degree of persistence of volatility in the price series. If b_i is close to 1, volatility is more likely to continue for a longer period of time. If b_i exceeds 1, it indicates an explosive series with a tendency to drift away from the mean value.

3.7 Cointegration analysis

Cointegration analysis is used to study the price co-movements and identifies stable, long run relationships between a set of markets. Engle and Granger (1987) pointed out that a linear combination of two or more non-stationary series may be stationary. If such stationary linear combination exists, the non-stationary time series are said to be cointegrated. In the present study, cointegration between tea prices in the markets of NI, SI, Mombasa (Kenya) and Colombo (Sri Lanka) were studied using the concept of market integration based on monthly tea price data from January 1980 to December 2020. Maximum Likelihood method of cointegration was developed by Johansen (1998) and later modified by Johansen and Juselius (1990). Procedure involved for the test of stationarity and cointegration are outlined in section 3.7.1 and 3.7.2 respectively.

3.7.1 Test for stationarity

Stationarity of a time series can be tested statistically by using unit root test. The most commonly used unit root test is Augmented Dickey Fuller (ADF) test. The null hypothesis for of the test is that time series has a unit root or it is non-stationary.

The ADF test was run with the equation,

$$\Delta X_t = \beta_1 + \delta X_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta X_{t-i} + \in_t \qquad \text{(a)}$$

$$\Delta X_t = \beta_1 + \beta_2 t + \delta X_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta X_{t-i} + \in_t \qquad \text{(b)}$$

where,

X_t is the tea price at time t (January 1980 to December 2020)

ΔX_t (first difference) $= X_t - X_{t-1}$

ΔX_{t-i} (ith difference) $= X_t - X_{t-i}$

$\in_t$, t = 1, ... , N, assumed to be Gaussian white noise i.e., $\in_t \sim N(0, \sigma^2)$

p is the number of lagged terms

The equation (a) is with constant term (β_1) and no trend, whereas, the equation (b) is with constant (β_1) and trend term ($\beta_2 t$). ADF test was done for tea prices of NI, SI, Mombasa (Kenya) and Colombo (Sri Lanka).

3.7.2 Test for cointegration

If two or more time series are associated to form a long-run equilibrium relationship, they are said to be cointegrated. When two time series are cointegrated, the series themselves may be non-stationary but will move closely together over time and the difference between them will be stationary. Cointegration test involves the following steps,

3.7.2.1 Determination of lag length using Vector Auto Regression (VAR)

Before applying cointegration, the lag length or order of Vector Auto Regression (VAR) has to be determined. The most commonly used criterion for determining the lag length are the Akaike Information Criteria (AIC) and Schwarz Information Criterion (SIC) and are given below:

$$AIC = \ln \left| \Omega(\hat{r}, p) \right| + (2/T)m$$

$$SIC = \ln \left| \Omega(\hat{r}, p) \right| + (\ln T/T)m$$

Where, $\Omega(\hat{r}, p) = \in_t{}' \in_t / T$

m = the number of freely estimated parameter in a VAR model of lag “p” and cointegration rank “r”

$\in_t{}'$ = a residual vector in the restricted rank VAR

ln = natural log

T = number of observations

Order of VAR for which AIC or SIC is lowest will be chosen. In the present study, SIC was used to determine the lag length.

3.7.2.2 Johansen's cointegration test

The integration between two markets that are spatially separated can be studied using pairwise or multiple cointegration analysis. Johansens Cointegration test is an efficient test procedure for determining the number of cointegrating vectors between the non-stationary time series in the context of a Vector Error Correction Model (VECM).

Error Correction Model (ECM)

Error Correction Model (ECM) for testing cointegration is given by,

$$\Delta X_t = \sum_{i=1}^{k-1} \Gamma i \Delta X_{t-i} + \Pi i X_{t-i} + \varepsilon_t$$

where,

X_t = Vector of endogenous variables (time series of tea price)

Γi = the matrix of short run coefficients

Πi = the matrix of long run coefficients

ε_t =the vector of independently normally distributed errors.

The rank r of matrix Π determines the number of cointegrating vectors, as it identifies how many linear combinations of X_t are stationary. The Johansen technique gives two likelihood ratio tests, the trace statistic and the maximum eigenvalue test, to identify the number of cointegrating relationships r. The trace statistic tests the null hypothesis of r cointegrating relations against the alternative k cointegrating relations and is given by,

$$\text{Trace } (r/k) = T \sum_{i=r+1}^{k} \ln(1-\lambda_i)$$

where,

k= number of endogenous variables, for r = 0, 1,…, k–1.

λ_i= ith largest eigenvalue in the Π matrix

T= number of observations.

The null hypothesis of r cointegrating vectors against the alternative of r + 1 is tested by the maximum Eigen value statistic and is given by,

$$\text{Max } (r/r+1) = -T \ln(1-\lambda_{(r+1)})$$

An increase in the number of cointegration vectors implies an increase in the strength and stability of price linkages.

Vector Error Correction Model

Trace statistic given in the preceding section gives the number of cointegrating equations indicating whether there is a long run relationship between the markets under study. But deviations from this equilibrium can occur in the short run. Hence, it is necessary to check whether such disequilibrium converges on the long run equilibrium or not. As vector error correction model provides a means whereby a proportion of the disequilibrium is corrected in the next period, it was done to generate such short run dynamics. In a system with two or more variables, a VECM, like the VAR model, treats each variable as potentially endogenous and relates the change in one variable to past equilibrium errors and to past changes in all variables in the system.

3.7.2.3 Granger causality Test

Cointegration gives only the existence of causality between two markets in at least one direction but, does not provide information regarding the direction of flow of information on prices (Granger, 1980). The Granger causality tests provide additional evidence about the direction of price transmission. When two markets are cointegrated, the price in one market X_t, would be found to Granger-Cause the price in the other market, Y_t and/or vice versa. Then the Granger causality test involves estimating the following pair of regressions.

$$X_t = \sum_{i=1}^{n} \alpha_i X_{t-i} + \sum_{j=1}^{n} \beta_j Y_{t-j} + \varepsilon_{xt} \qquad \text{(a)}$$

$$Y_t = \sum_{i=1}^{n} \gamma_i Y_{t-i} + \sum_{j=1}^{n} \lambda_j X_{t-j} + \varepsilon_{yt} \qquad \text{(b)}$$

Null hypothesis for equation (a): X_t does not granger-cause Y_t

Null hypothesis for equation (b): Y_t does not granger-cause X_t

Unidirectional causality is indicated when the estimated coefficients in one of the regression equations is statistically nonsignificant. Bilateral causality is suggested when all the regression coefficients in both the equations (a) and (b) are significantly different from zero. If in both regression equations, when all the regression coefficients are not significantly different from zero, the two markets are independent. Otherwise, they are cointegrated. For cointegration analysis, EViews package was used.

Results and Discussion

4. RESULTS AND DISCUSSION

The results of the study "Time series modeling and forecasting of tea prices in India" are discussed in this section under the following titles:

4.1. Trend analysis

4.2. Compound Annual Growth Rate (CAGR)

4.3. Time series decomposition of tea prices

4.4. Forecast models for tea prices

4.5. Price volatility of tea prices

4.6. Cointegration analysis

4.1 TREND ANALYSIS

Trend is defined as the general tendency of a time series data to increase or decrease during a long period of time. Trend analysis was carried out for A- Pr- Pd of tea in NI, SI and All India, by the method of least squares. Different functional forms like linear, quadratic, cubic, exponential, etc were tried and the suitable model in each case was chosen based on the criteria like MAPE and adjusted R^2 value.

4.1.1 Trend analysis of A-Pr-Pd of tea in NI

The area under tea in NI increased from 2.8 lakh ha in 1970 to 5.35 lakh ha in 2019. For trend in area under tea in NI, quadratic model was selected. The production of tea in NI increased from 3.7 lakh tonnes in 1970 to 11.7 lakh tonnes in 2019. Cubic model was observed to be suitable for trend in production of tea in NI. The productivity of tea in NI increased from1108 kg/ha in 1970 to 1804.63 kg/ha in 1998, further a declining trend was noticed till 2010, where the productivity was 1573.13 Kg/ha (Figure 4.3). After 2010, the productivity showed an increasing trend and attained highest productivity of 2334 kg/ha in 2017. Cubic model was found to be the best fit for trend in productivity of tea in NI. The trend equation fitted for A- Pr- Pd along with the accuracy measures are provided in Table 4.1.

Table 4.1 Trend equations for A- Pr- Pd of tea in NI

Trend equation	**Accuracy measures**		
	MAPE	**R^2**	**Adjusted R^2**
Area: $Y_t = 2.79 + 0.015\,t + 0.00064t^2$	2.39	0.965	0.963
Production: $Y_t = 2.75 + 0.245t - 0.0089\,t^2 + 0.00015\,t^3$	3.54	0.98	0.979
Productivity: $Y_t = 1.05 + 0.0605t - 0.0023t^2 + 0.000033t^3$	4.58	0.86	0.86

From the table, it could be observed that, a very low MAPE value of 2.39 and adjusted R^2 value of 96.3 percent show the adequacy of the quadratic model in explaining the trend in area under tea in NI. Cubic model was found to be adequate for trend in production as indicated by the very low MAPE value (3.54) and adjusted R^2 value of 97.9 percent. It could also be observed that, for trend in productivity of tea in NI, cubic model was adequate with a low MAPE value of 4.58 and adjusted R^2 value of 86 percent. The actual and estimated A- Pr- Pd of tea in NI based on the fitted models are depicted in Figure 4.1, 4.2 and 4.3, respectively.

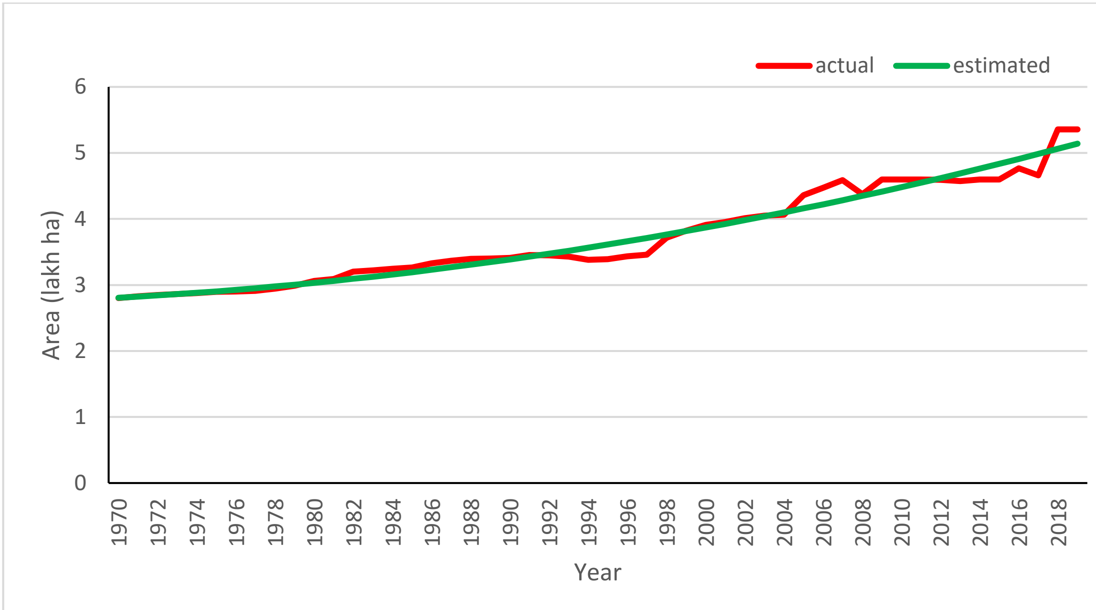

Figure 4.1 Actual and estimated plots of area under tea in NI

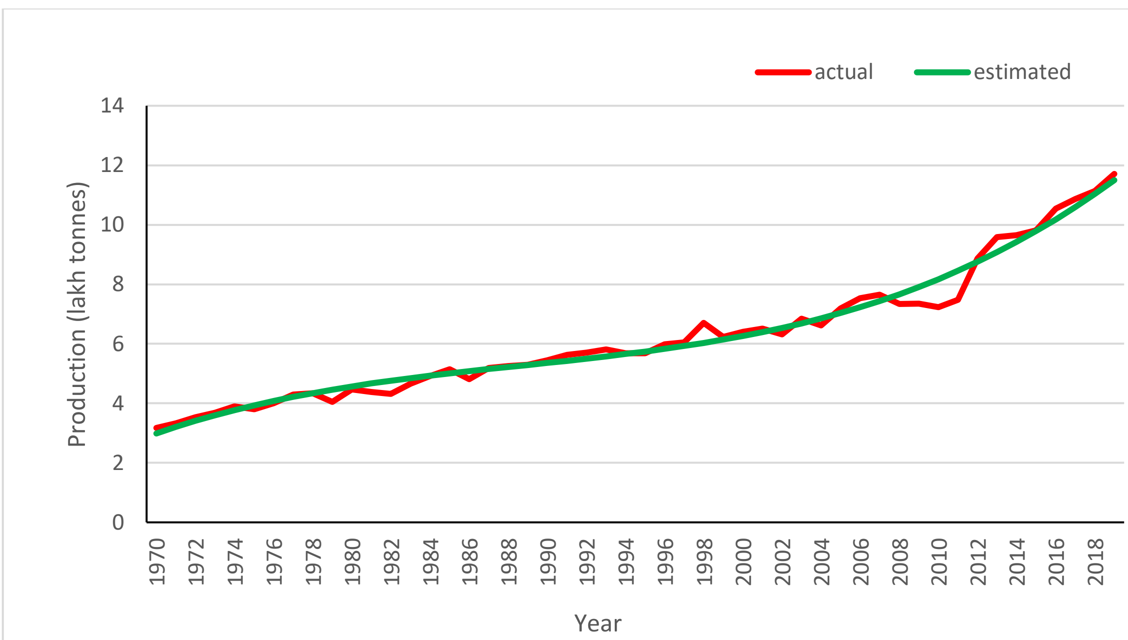

Figure 4.2 Actual and estimated plots of production of tea in NI

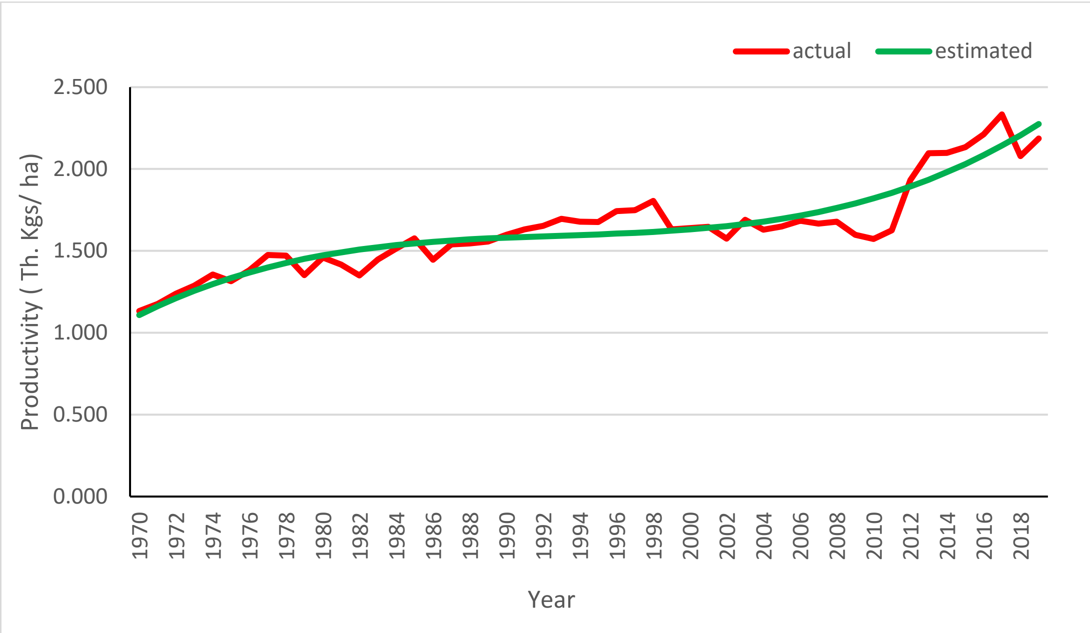

Figure 4.3 Actual and estimated plots of productivity of tea in NI

4.1.2 Trend analysis A-Pr-Pd of tea in SI

For SI, during 1970-1993, there was no visible change in area under tea and an increasing trend was observed from 1993 (0.75 lakh ha) to 2012 (1.197 lakh ha), as can be seen from Figure 4.4. But the area declined to 1.009 lakh ha in 2019. Cubic model was found to be the best fit for area under tea in SI. For production of tea, a general increase could be observed from 1993 (1.79 lakh tonnes) to 2014 (2.42 lakh tonnes) with highest production of 2.469 lakh tonnes in the year 2008 (Figure 4.5). Cubic model was observed to be the best fit for production of tea. For trend in productivity of tea in SI, cubic model was identified to be the best fit. The trend equation fitted for A- Pr- Pd along with the accuracy measures are provided in Table 4.2.

Table 4.2 Trend equation for A- Pr- Pd of tea in SI

Trend equation	Accuracy measures		
	MAPE	**R^2**	**Adjusted R^2**
Area: $Y_t = 0.88 - 0.044t + 0.0027\, t^2 - 0.0000359\, t^3$	5.98	0.903	0.897
Production: $Y_t = 1.038 - 0.0028t + 0.0023\, t^2 - 0.0000356\, t^3$	4.25	0.965	0.963
Productivity: $Y_t = 1.05 + 0.107t - 0.0036t^2 + 0.0000408t^3$	6.81	0.676	0.655

From the table it could be observed that, a very low MAPE value of 5.98 and adjusted R^2 value of 89.7 percent show the adequacy of the cubic model in explaining the trend in area under tea in SI. Cubic model was found to be adequate for trend in production as indicated by the very low MAPE value of 4.25 and adjusted R^2 value of 96.3 percent. It could be also observed that, for trend in productivity of tea in SI, cubic model was adequate with a low MAPE value of 6.81 and adjusted R^2 value of 65.5 percent. The actual and estimated A- Pr- Pd of tea in SI based on the fitted model are depicted in Figure 4.4, 4.5 and 4.6 respectively.

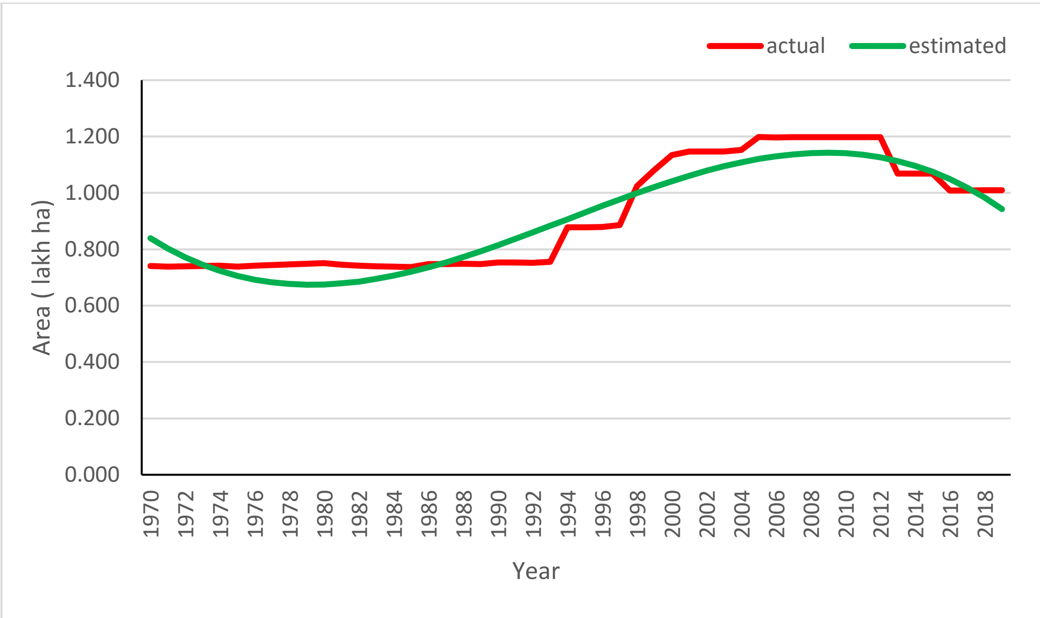

Figure 4.4 Actual and estimated plots of area under tea in SI

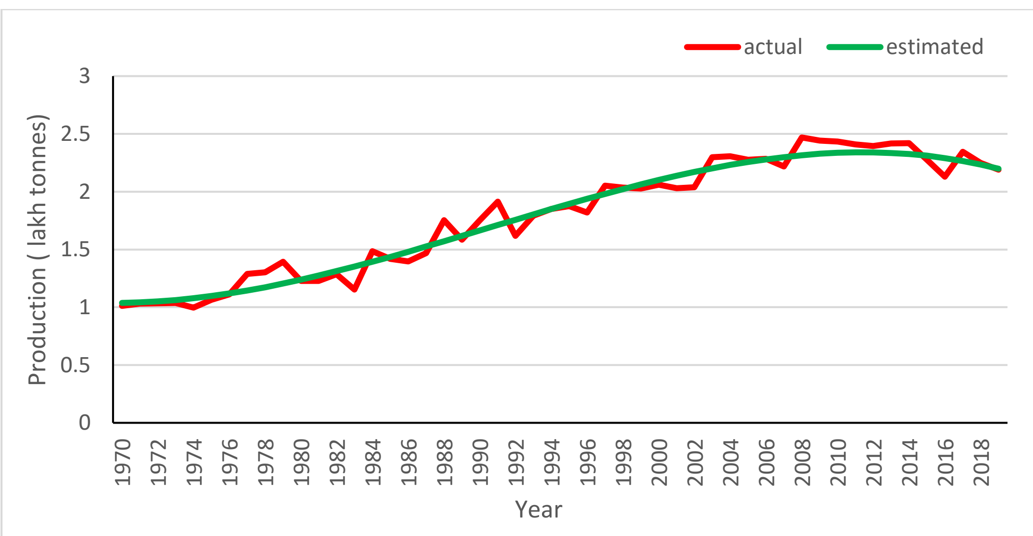

Figure 4.5 Actual and estimated plots of production of tea in SI

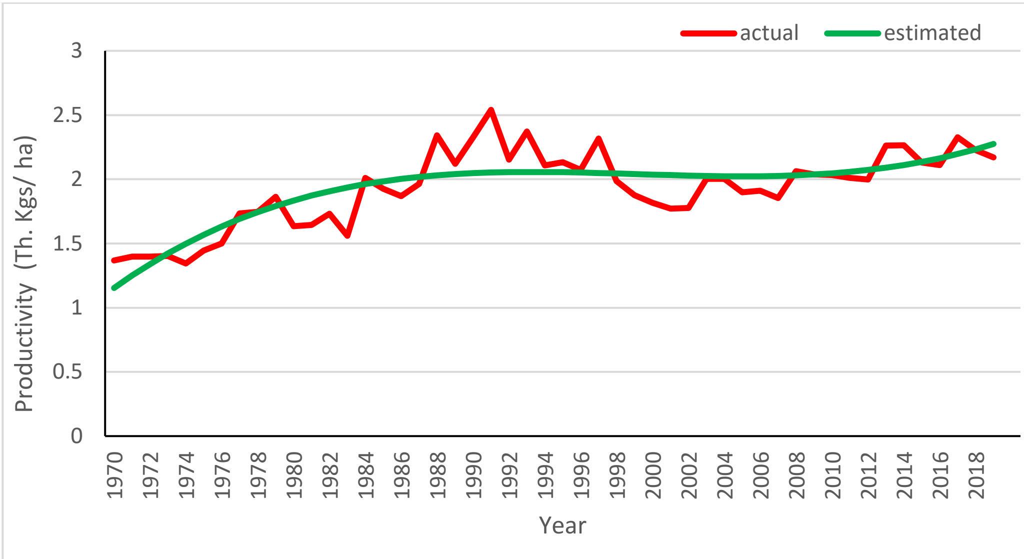

Figure 4.6 Actual and estimated plots of productivity of tea in SI

4.1.3 Trend analysis of tea in All India

The trend equations fitted for A- Pr- Pd along with the accuracy measures are provided in Table 4.3.

Table 4.3 Trend equation for A- Pr- Pd of tea in All India

Trend equation	Accuracy measures		
	MAPE	**R^2**	**Adjusted R^2**
Area: $Y_t = 3.68 - 0.0296t + 0.0033t^2 - 0.0000365t^3$	3.06	0.963	0.957
Production: $Y_t = 3.78 + 0.2451t - 0.0068t^2 + 0.000119t^3$	2.72	0.987	0.986
Productivity: $Y_t = 1.058 + 0.068t - 0.00259t^2 + 0.0000342t^3$	4.59	0.862	0.853

From the table it could be observed that, a very low MAPE value of 3.06 and adjusted R^2 value of 95.7 percent show the adequacy of the cubic model in explaining the trend in area under tea in All India. Cubic model was found to be adequate for trend in production as indicated by the very low MAPE value of 2.72 and adjusted R^2 value of 98.7 percent. It could be also observed that, for trend in productivity of tea in All India, a cubic model was adequate with a low MAPE value of 4.59 and adjusted R^2 value of 85.3 percent.

The actual and estimated A- Pr- Pd of tea in All India based on the fitted models are shown in Figure 4.7, 4.8 and 4.9 respectively. For area of tea in All India, a general increasing trend was observed from 1970 (3.54 lakh ha) to 2019 (6.36 lakh ha). For production of tea in All India, a general increasing trend could be observed from 1970 (4.18 lakh tonnes) to 2019 (13.9 lakh tonnes). Productivity of tea increased from 1180 kg/ha in 1970 to 1870 kg/ha in 1997, then declined to 1670 kg/ha in 2010 (Figure 4.9). It further increased to 2330 kg/ha in 2017.

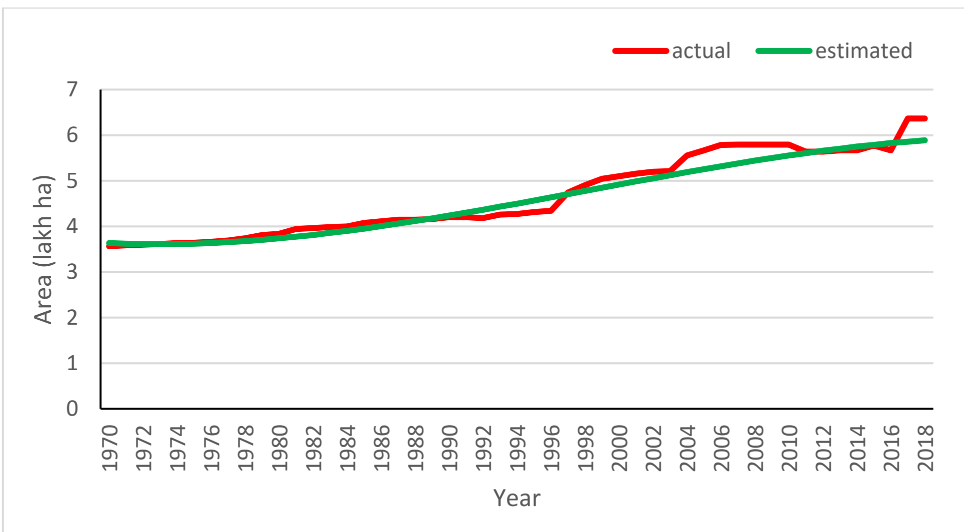

Figure 4.7 Actual and estimated plots of area under tea in All India

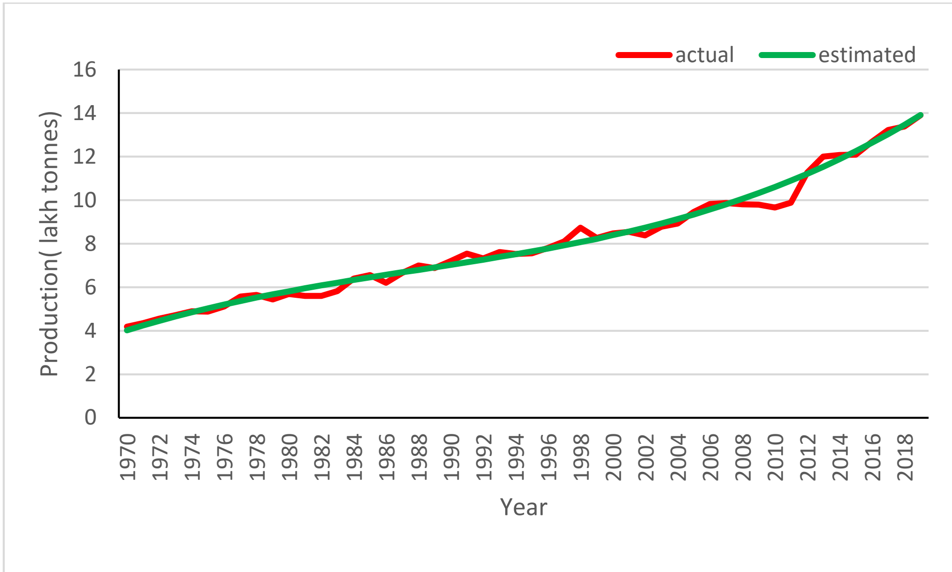

Figure 4.8 Actual and estimated plots of production of tea in All India

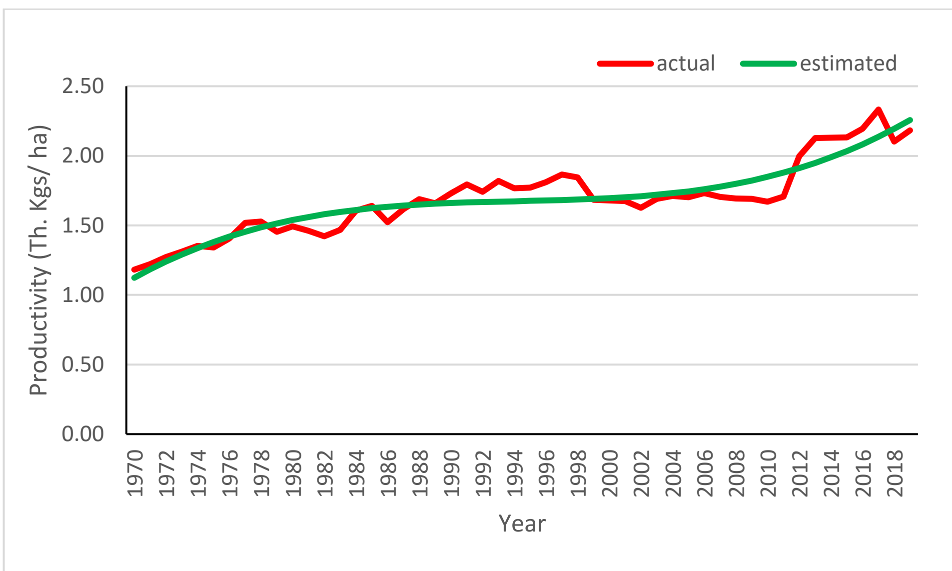

Figure 4.9 Actual and estimated plots of productivity of tea in All India

From the trend analysis carried out for A- Pr- Pd of tea in NI, SI and All India, it was observed that quadratic model was found to the best fit for area under tea in NI while, cubic model was found to be the appropriate fit for production and productivity of tea in NI and, A- Pr- Pd of tea in SI and All India.

4.2 Compound Annual Growth Rate (CAGR)

Compound Annual Growth Rate (CAGR) was used as a tool to study the proportion and extent of changes that have taken place in A-Pr-Pd of tea in NI and SI over a period of 50 years from 1970 to 2019. For detailed analysis, the whole period was divided into two sub period i.e., period I (1970-1995) and period II (1996-2019). The results are provided in Table 4.4.

Table 4.4 CAGR of A- Pr- Pd of tea in NI and SI

		Compound Annual Growth Rate (%)		
		Area	**Production**	**Productivity**
Period I (1970-1995)	NI	0.97	2.32	1.34
	SI	0.84	2.69	2.35
Period II (1996-2019)	NI	1.56	2.89	1.29
	SI	1.27	0.73	0.62
Overall Period (1970-2019)	NI	1.27	2.27	0.98
	SI	1.19	1.96	0.76

The compound annual growth rates of A- Pr- Pd of tea for both NI and SI were positive. It could be observed from the table that during the overall period, the A- Pr- Pd of tea in NI showed growth rate of 1.27, 2.27 and 0.98 percent respectively and corresponding figures were 1.19, 1.96 and 0.76 percent for SI respectively.

From Figure 4.10 and Figure 4.11, it could be observed that the growth rate of area under tea in NI increased from 0.97 percent (1970-1995) to 1.56 percent (1996-2019) and from 0.84 percent (1970-1995) to 1.27 percent (1996-2019) in SI. In the recent years, tea has been introduced into more states which has contributed more towards area (Tea Board, 2020). For NI, an increase in growth rate was noticed in tea production in period II

(2.89 percent) compared to period I (2.27 percent), corresponding to the growth in area (Figure 4.10).

For SI, a decline from 2.69 percent to 0.73 percent was noticed for growth rate in production from period I to period II (Figure 4.11). This is because, after 2014, a steady decline in production was observed due to low number of rainy days, heavy rainfall and drought (UPASI, 2021). Growth rate in productivity was found to be high during period I in both NI (1.34 percent) and SI (2.35 percent) compared to period II. During period II, although growth rate in area under tea was found to be increasing, growth rate in production was low compared to period I. This might be the reason for decline in growth rate of productivity in period II.

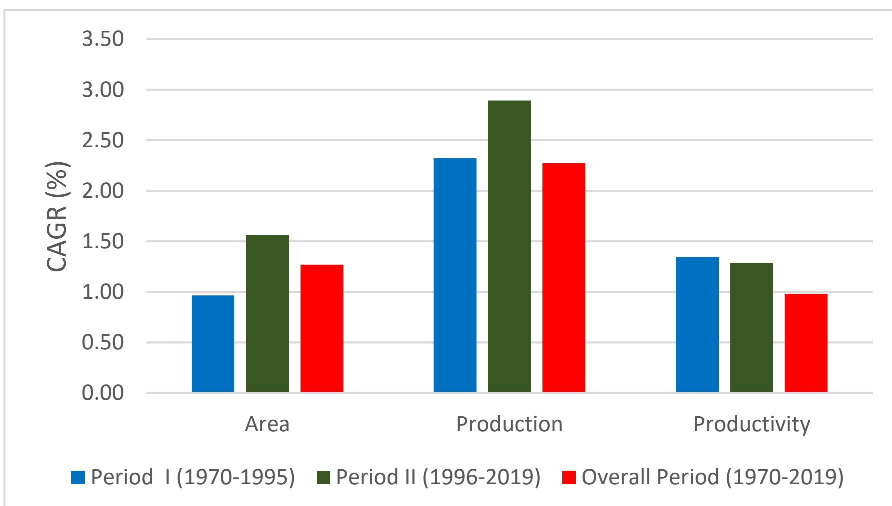

Fig 4.10 CAGR of A- Pr- Pd of tea in NI

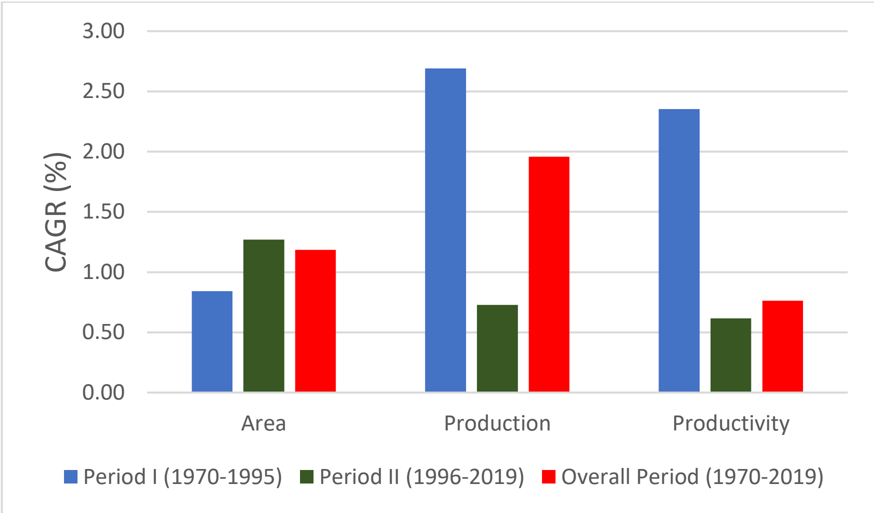

Fig 4.11 CAGR of A- Pr- Pd of tea in SI

4.3 Time series decomposition of tea prices

A time series is a sequence of observations ordered in time (Anderson, 1971). Various forces at work affecting the values of a variable in a time series can be broadly classified into four categories known as components of time series namely, trend, seasonal variation, cyclic variation and irregular variation (Croxton et al., 1979). The time series of tea prices in NI, SI and All India were decomposed into the four components, as per Croxton et al. (1979) and Spiegel (1992) assuming a multiplicative model and are depicted in Figures 4.12, 4.13 and 4.14 respectively. In all figures, four panels are shown. In the first panel, the observed price is plotted, second panel is the plot of trend, third and fourth panels give the plots of seasonal and random variations. For both NI and SI, an overall increasing trend and a prominent seasonal variation could be observed.

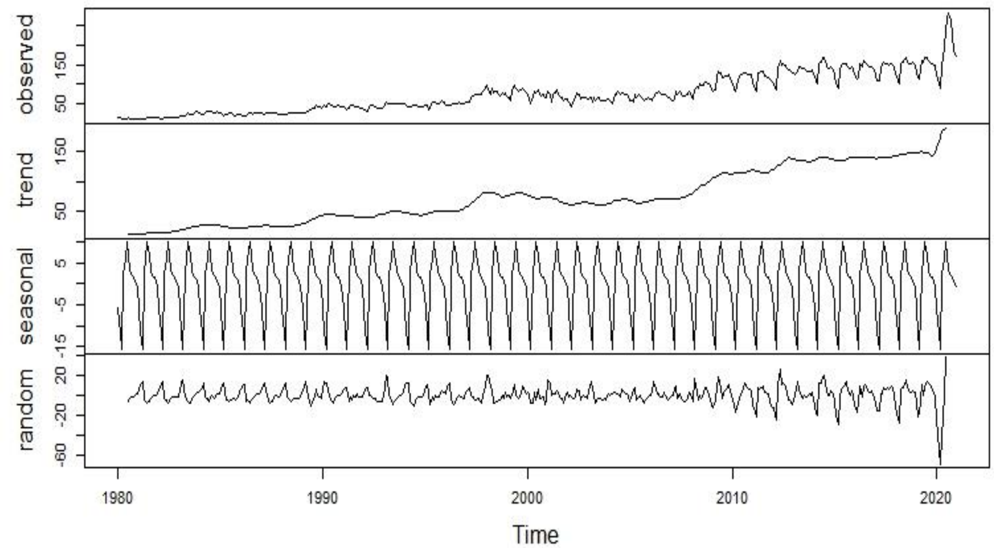

Figure 4.12 Time series decomposition of NI tea price

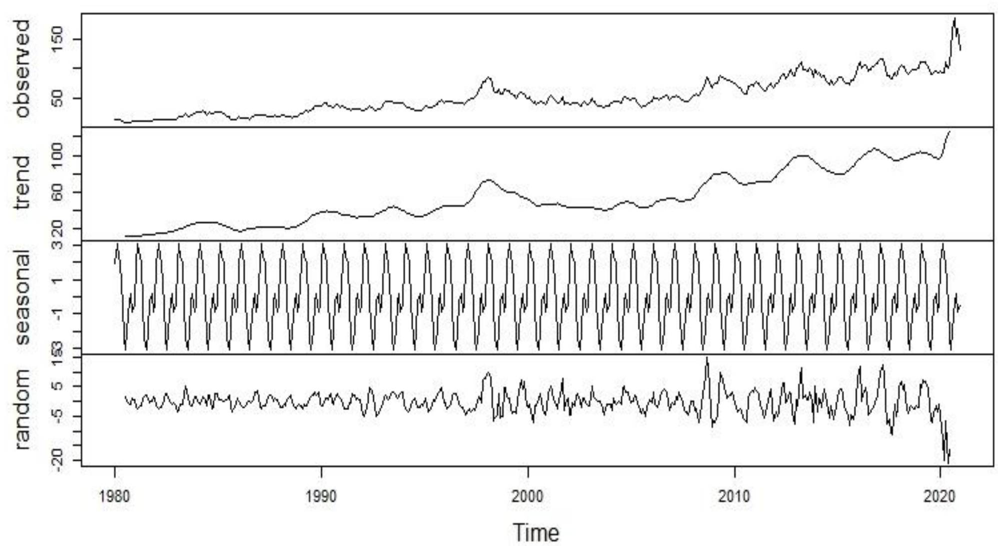

Figure 4.13 Time series decomposition of SI tea price

All India price is the simple average of tea prices in NI and SI. Hence seasonality observed for All India tea price is the combined seasonality of both NI and SI.

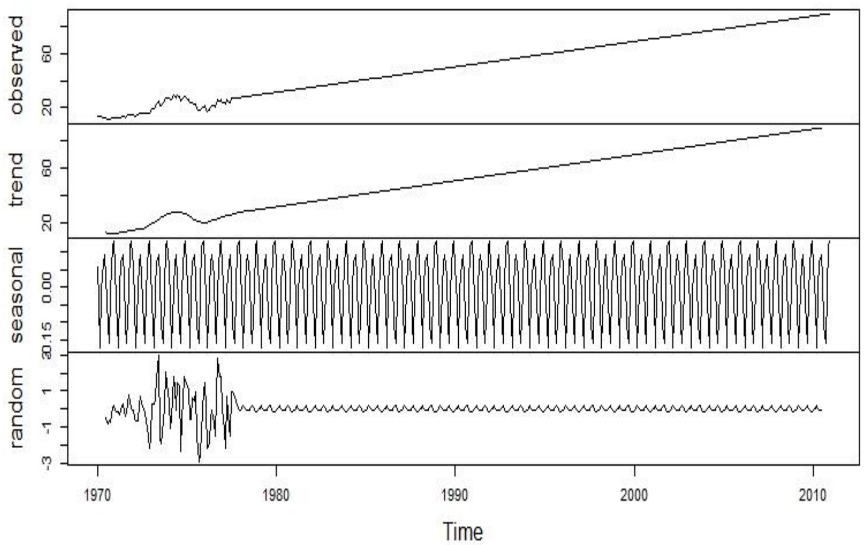

Figure 4.14 Time series decomposition of All India tea price

4.3.1 Seasonal indices of tea price

Decomposition of the time series data of tea prices during January 1980-December 2020 indicated the same pattern in every year for both NI and SI. Hence, seasonal indices were computed for tea prices for NI and SI, to get a clear picture of price pattern in the 12 months from January to December. Seasonal indices are provided in Table 4.5 and seasonal plots are shown in Figure 4.15.

Table 4.5 Seasonal indices of tea price in NI and SI

Month	Seasonal Indices	
	NI	SI
January	94.3	103.4
February	88.1	105.9
March	82.6	105.1
April	102	103.8
May	106.8	99.7
June	111.6	95.7
July	108.9	94.2
August	102.9	95.5
September	101	98.7
October	101.6	100.0
November	100.8	98.8
December	99.3	99.2

From the above table and Figure 4.15, it could be observed that for NI, April to November are the high price months with highest price in June followed by July whereas, December to March are the low-price months with lowest price in March. A hike in price can be observed from March to June. This is because, in NI, first flush starts in late March and second flush occurs during end of May to June. This season produces high quality tea leaves and has got the highest price. Third or autumn flush occurs from October to November which is of low quality compared to first and second flush. In comparison to the first two, this product has the lowest market price. Hence a decline in the price could be observed during these months due to deterioration in quality of autumn flush (Anonymous, 2019).

For SI, high price could be observed from January to April with highest price in February followed by March. Compared to NI, SI doesn't show much seasonal variation. This is because, in SI, tea growers harvest the plants throughout the year. Hence the quality of tea leaves will be same throughout the year. Low price months are from June to September which coincide with southwest monsoon season and lowest price is observed in July.

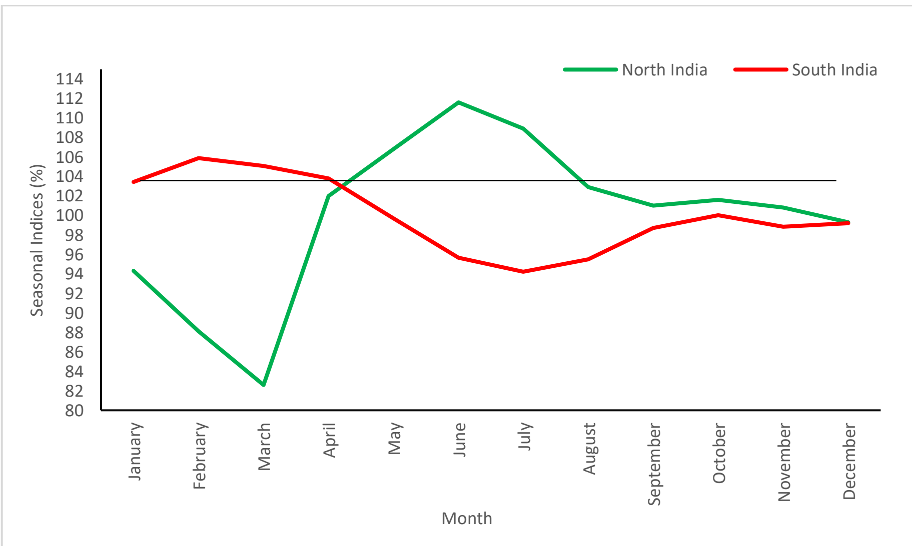

Figure 4.15 Seasonal plots of tea prices for NI and SI

4.3.2 Cyclic variation

Cyclic indices were calculated by eliminating trend, seasonal and irregular variations from the observed tea price data for the period from 1980 to 2020. The cyclic indices of tea price for NI are given in Figure 4.16. A prominent cycle of length 10 to 11 years could be observed from 1998 to 2009. Small cycles of length 2 to 3 years with a distinct cyclic pattern could be observed during the initial few years.

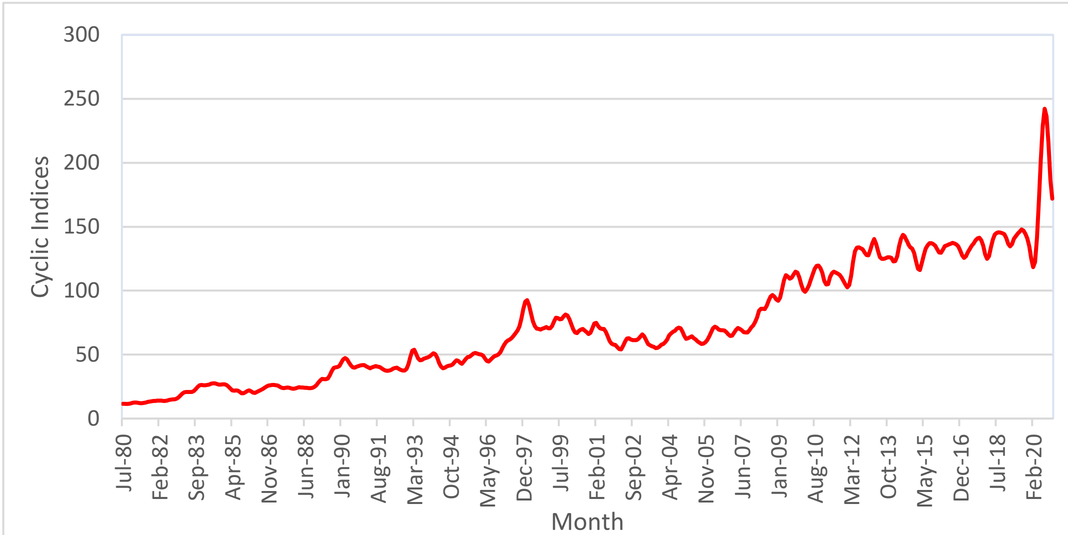

Figure 4.16 Cyclic variation of tea price in NI

Cyclic indices of tea price for SI are depicted in Figure 4.17. For SI, a prominent cycle could be observed from January 1998 to February 2010. A small cycle of length 5 to 6 years could be observed during the period from 1989 to 1994. Compared to NI, SI exhibited more cycles of variation in tea price.

Figure 4.17 Cyclic variation of tea price for SI

4.4 Forecast Models for tea prices

Exponential Smoothing models and ARIMA models were fitted for the tea prices in NI and SI to forecast the price. For fitting the forecast models, price data from January 1980 to December 2020 were used and forecasts were made from January 2021 to April 2021. The tea prices reported by tea board of India were used for validation of forecast. The results for NI and SI tea prices are presented from section 4.4.1 to 4.4.4.

4.4.1 Forecast model for NI tea price

The plot of NI tea price from January 1980 to December 2020 are provided in figure 4.18.

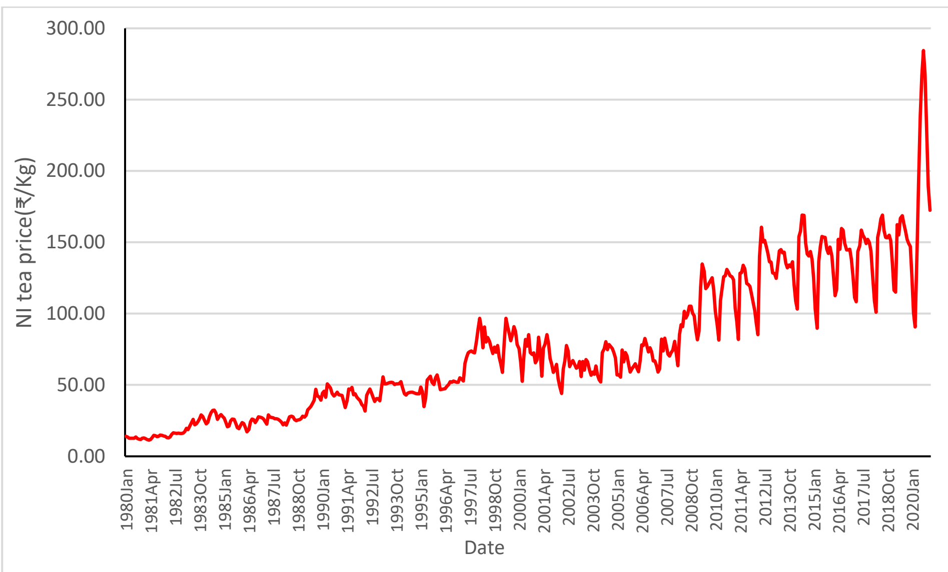

Figure 4.18. Price pattern of tea in NI

4.4.1.1 Exponential Smoothing model for tea price in NI

Different exponential smoothing models like SES, DES, HWAS and HWMS models were compared for the NI tea price based on criterias like agreement between observed and fitted price plots and MAPE. From among the three models tried, Holt-Winters' Multiplicative Seasonal (HWMS) model was selected as the best. The fit of the

HWMS model for tea price in NI is provided in Figure 4.19. It could be observed that the actual and model fit values are in close agreement.

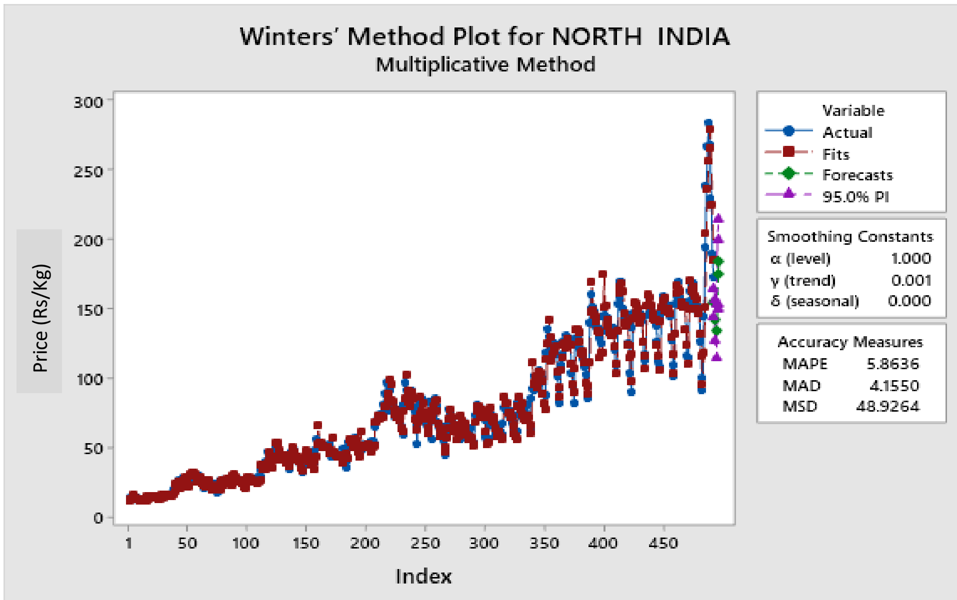

Figure 4.19 HWMS model for tea prices in NI

The estimates of parameters of HWMS model are given in Table 4.6.

Table 4.6 Estimates of parameters of the HWMS model for tea prices in NI

Parameter	$\boldsymbol{\alpha}$	$\boldsymbol{\beta}$	$\boldsymbol{\gamma}$
Estimate	1.00	0.001	0

With these values for the parameters, HWMS model for tea price in NI are given below,

Level: $L_t = \frac{X_t}{S_{t-12}}$

Trend: $b_t = 0.001(L_t - L_{t-1}) + 0.999b_{t-1}$

Seasonality: $S_t = S_{t-12}$

Forecast: $F_{t+m} = (L_t + b_t m)S_{t-12+m}$

The values of L_t, b_t and S_t were substituted in the forecast equation F_{t+m} to obtain the price.

The forecasts of tea price in NI from January to April 2021 based on HWMS model along with forecast accuracy measure MAPE are provided in Table 4.7. The MAPE was 5.86 for the model.

Table 4.7 Price forecasts from HWMS model for tea price in NI

Month	Forecast price (₹/ Kg)	MAPE (%)
Jan-2021	154.54	5.86
Feb-2021	142.13	
Mar-2021	134.46	
Apr-2021	175.14	

4.4.1.2 ARIMA model for tea price in NI

For fitting ARIMA model, the data should be stationary. Autocorrelations and partial autocorrelations were computed for tea price in NI and tested for significance using Ljung- Box Statistic (Table 4.8). It could be observed from the table that autocorrelations declined very slowly from 0.972 to 0.762 and were all significant.

Table 4.8 Autocorrelations and partial autocorrelations of tea price in NI

Lag	Auto-correlation	SE	Ljung- Box Statistic			Partial Auto-correlation	Std. Error
			Value	df	Probability (p)		
1	0.972	0.045	467.58	1	.000	0.972	0.045
2	0.932	0.045	898.09	2	.000	-0.235	0.045
3	0.889	0.045	1290.91	3	.000	-0.016	0.045
4	0.853	0.045	1653.23	4	.000	0.110	0.045
5	0.826	0.045	1993.74	5	.000	0.102	0.045
6	0.808	0.045	2320.27	6	.000	0.089	0.045
7	0.795	0.045	2637.24	7	.000	0.039	0.045
8	0.79	0.045	2950.71	8	.000	0.133	0.045
9	0.795	0.045	3268.60	9	.000	0.184	0.045
10	0.810	0.045	3599.56	10	.000	0.196	0.045
11	0.827	0.044	3945.10	11	.000	0.042	0.045
12	0.838	0.044	4300.36	12	.000	-0.022	0.045
13	0.821	0.044	4642.64	13	.000	-0.392	0.045
14	0.799	0.044	4967.54	14	.000	0.161	0.045
15	0.777	0.044	5275.39	15	.000	0.089	0.045
16	0.762	0.044	5571.76	16	.000	0.097	0.045

The ACF and PACF plots given in Figure 4.20 shows that all spikes of ACF fell outside the 95% confidence limit and PACF also showed spikes at certain lags (1,2,4,8,9,10,13,14) which are outside the confidence limits. It indicates non stationarity of NI tea price series.

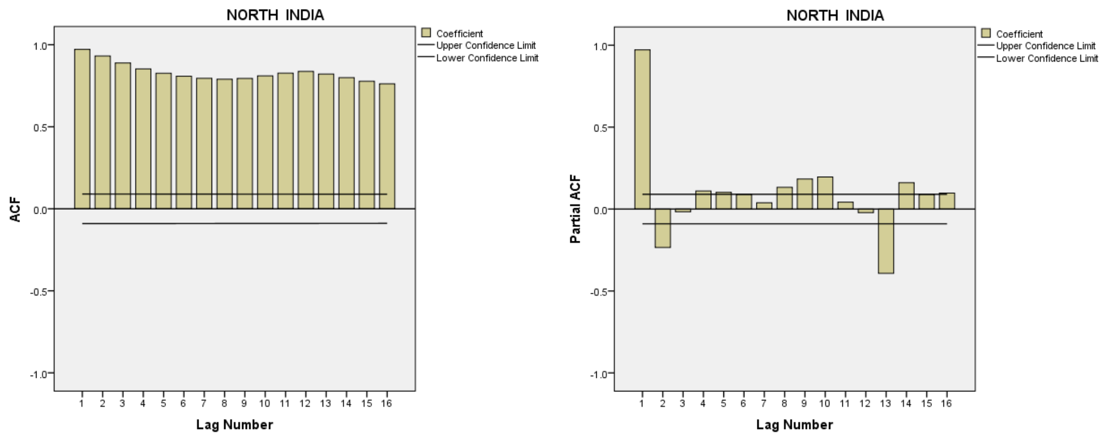

Figure 4.20 ACF and PACF plots of tea price in NI

Augmented Dickey Fuller (ADF) test was conducted for tea price in NI to confirm the non-stationarity of the data and the results are presented in Table 4.9. It could be observed that after first differencing, the data became stationary.

Table 4.9 ADF test for tea prices in NI

NI tea price	ADF test statistic	Probability (p)	Critical values
Actual	0.78	0.99	-3.45
first difference	-6.62**	0.00	-2.87

** indicates significant at 1%level (p<0.01)

Log transformation was done to attain stationarity in variance. From among several ARIMA models tried, SARIMA (0,1,3) $(0,1,1)_{12}$ was chosen as the best forecast model for NI tea price, based on agreement between actual and fitted price plots, MAPE and residual ACF and PACF plots.

The plots of actual, fit, Lower Confidence Limit (LCL), Upper Confidence Limit (UCL) and forecast values for NI tea price for SARIMA $(0,1,3)(0,1,1)_{12}$ are provided in Figure 4.21. Close agreement was obtained between actual and fitted values.

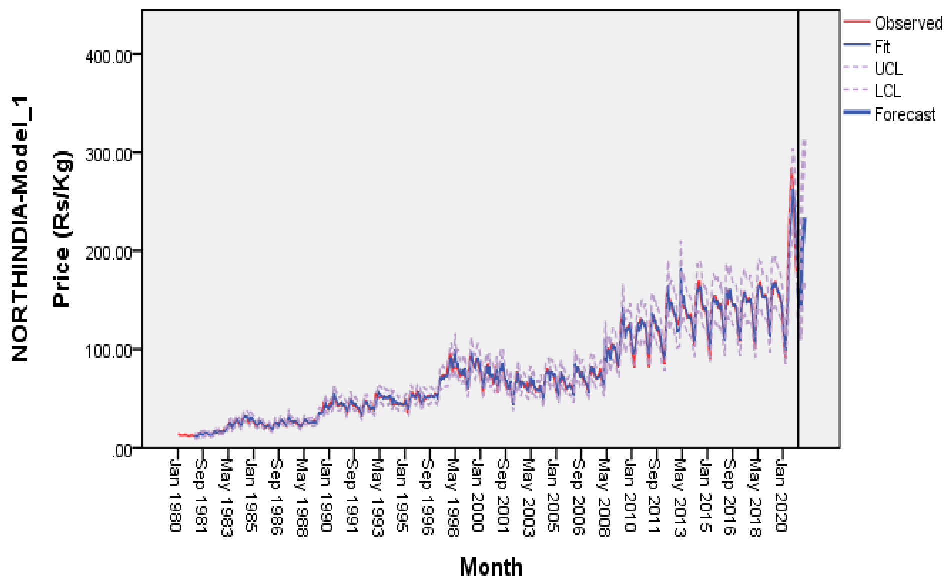

Figure 4.21 Observed, Fit, LCL, UCL, and forecast values with SARIMA(0,1,3)(0,1,1)$_{12}$ for tea price in NI

Residual ACF and PACF plots are provided in Figure 4.22. It could be observed that all ACF and PACF values lie within the confidence limits.

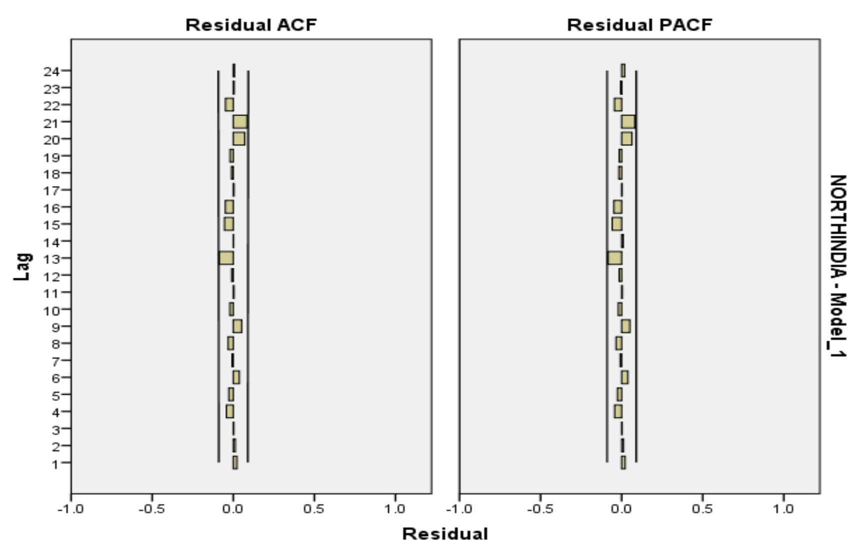

Figure 4.22 Residual ACF and PACF plots for SARIMA(0,1,3)(0,1,1)$_{12}$

The parameters of the model SARIMA(0,1,3)$(0,1,1)_{12}$ along with their tests of significance are provided in Table 4.10.

Table 4.10 SARIMA (0,1,3) $(0,1,1)_{12}$ model parameters for tea price in NI

Transformation	Model Parameters	Estimate	SE	t	Probability(p)
x_t=log$_e$ x_t	Non-seasonal difference	1			
	MA Lag 2 (θ_2)	0.181	0.045	4.007**	0.00
	MA Lag 3 (θ_3)	0.13	0.045	2.834**	0.005
	Seasonal difference	1			
	MA, Seasonal Lag 1(Θ_1)	0.82	0.031	26.375**	0.00

** indicates significance at 1 percent level (p<0.01)

Thus, SARIMA (0,1,3) $(0,1,1)_{12}$ model for NI tea price is

$$(1-B)(1-B^{12})x_t = (1-\theta_2 B^2 - \theta_3 B^3)(1-\Theta_1 B^{12})\varepsilon_t$$

Where, θ_2= 0.181, θ_3 = 0.13 and Θ_1= 0.82

With these parameters, the SARIMA(0,1,3)$(0,1,1)_{12}$ model is

$$(1-B)(1-B^{12})x_t = (1-0.181B^2 - 0.13B^3)(1-0.82B^{12})\varepsilon_t$$

Different accuracy measures of the above model are given in Table 4.11. It could be observed that R^2 was very high (97.8%) and MAPE was very small (5.65) for the model.

Table 4.11 Model fit statistics of SARIMA (0,1,3)$(0,1,1)_{12}$ for tea price in NI

R^2	RMSE	MAPE (%)	MAE	SBC
0.978	7.18	5.65	4.29	3.98

Based on the fitted model, tea prices for NI were forecasted from January 2021 to April 2021 and provided in Table 4.12 along with LCL and UCL.

Table 4.12 Forecasts of tea price at NI

Month	Forecast Price (₹/Kg)		
	Forecast price	LCL	UCL
Jan-2021	161.69	138.94	187.07
Feb-2021	146.16	117.76	179.35
Mar-2021	141.38	110.03	178.91
Apr-2021	201.25	153.18	259.69

4.4.1.3 Comparison of forecast models for tea price in NI

For the tea price of NI, HWMS and SARIMA $(0,1,3)(0,1,1)_{12}$ had low MAPE values. Price forecasted using these models were compared with actual price and is provided in the Table 4.13

Table 4.13 Comparison of forecast models for tea price in NI

Month	Actual Price (₹/ Kg)	Forecast Price (₹/ Kg)	
		SARIMA$(0,1,3)(0,1,1)_{12}$	HWMS
Jan-21	**169.22**	161.69	154.54
Feb-21	**157.45**	146.16	142.13
Mar-21	**150.11**	141.38	134.46
Apr-21	**210.10**	201.25	175.14
MAPE		5.65	5.86

From the Table it could be observed that SARIMA$(0,1,3)(0,1,1)_{12}$ was in more agreement with the actual price and MAPE was low compared to HWMS model. Thus for tea prices of NI, SARIMA$(0,1,3)(0,1,1)_{12}$ was selected as the best forecast model.

4.4.3 Forecast model for tea price in SI

The plot of SI tea price from January 1980 to December 2020 are provided in Figure 4.23.

Figure 4.23. Price pattern of tea in SI

4.4.3.1 Exponential Smoothing model for tea price in SI

Different exponential smoothing models like Single Exponential smoothing (SES), Double Exponential smoothing (DES), Holt-Winters' Additive Seasonal (HWAS) and Holt-Winters' Multiplicative Seasonal (HWMS) models were compared for the SI tea price based on criteria like agreement between observed and fitted price plots and MAPE. From among several models tried, Holt-Winters' Multiplicative Seasonal (HWMS) model was selected as the best. The fit of the HWMS model for tea price in SI is provided in Figure 4.24. It could be observed that the actual and model fit values are in close agreement.

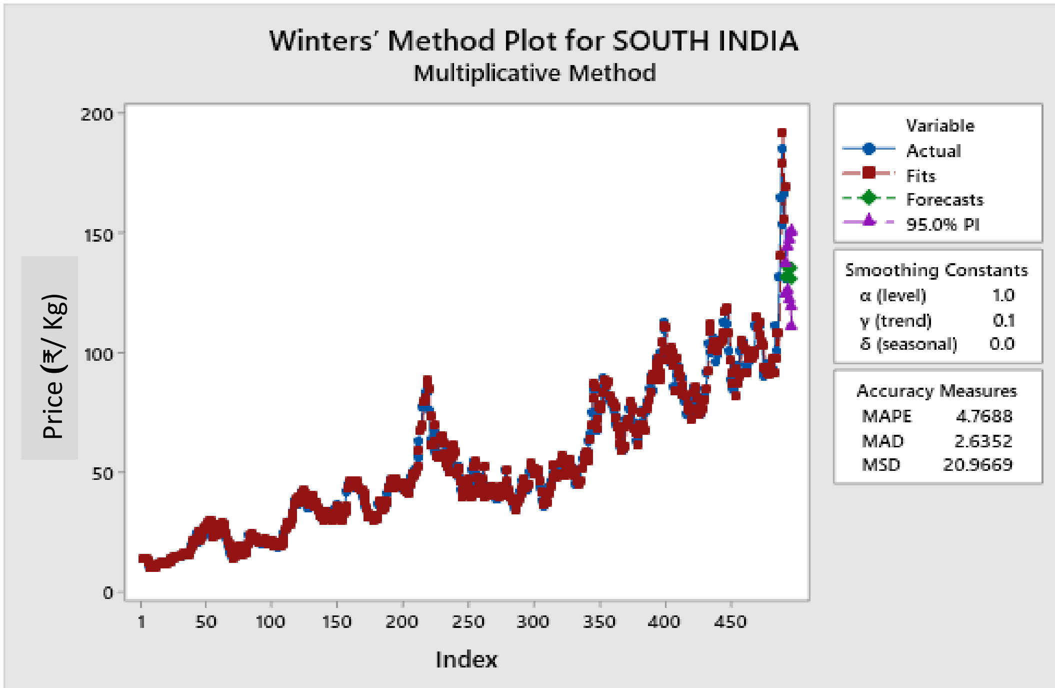

Figure 4.24 HWMS model for tea price in SI

The estimates of parameters of HWMS model are given in the Table 4.14.

Table 4.14 Estimates of parameters of the HWMS model for tea prices in SI

Parameter	α	β	γ
Estimate	1.00	0.1	0

With these values for the parameters, the equation for HWMS model for tea price in SI are given below,

Level: $L_t = \frac{X_t}{S_{t-12}}$

Trend: $b_t = 0.1(L_t - L_{t-1}) + 0.9b_{t-1}$

Seasonality: $S_t = S_{t-12}$

Forecast: $F_{t+m} = (L_t + b_t m)S_{t-12+m}$

The estimated values of L_t, b_t and S_t were substituted in the forecast equation F_{t+m} to obtain one step ahead forecast.

The forecasts of tea price in SI from January to April 2021 based on HWMS model along with forecast accuracy measure MAPE are provided in Table 4.15. The MAPE was 4.76 for the model.

Table 4.15 Price forecasts from HWMS model for tea price in SI

Month	Forecast price (₹/ Kg)	MAPE (%)
Jan-2021	131.68	4.76
Feb-2021	135.65	
Mar-2021	135.28	
Apr-2021	135.62	

4.4.3.2 ARIMA model for tea price in SI

For fitting ARIMA model, the data should be stationary. Autocorrelations and partial autocorrelations were computed for tea price in SI and tested for significance using Ljung- Box Statistic (Table 4.16). It could be observed from the table that autocorrelations declined very slowly from 0.979 to 0.752 and are all significant.

Table 4.16 ACF and PACF values of tea price in SI

Lag	Auto-correlation	SE	Ljung- Box Statistic			Partial Autocorrelation	Std. Error
			Value	df	Probability (p)		
1	0.979	0.045	474.39	1	0.000	0.979	0.045
2	0.950	0.045	921.89	2	0.000	-0.205	0.045
3	0.921	0.045	1343.91	3	0.000	0.042	0.045
4	0.889	0.045	1737.89	4	0.000	-0.122	0.045
5	0.864	0.045	2110.14	5	0.000	0.185	0.045
6	0.844	0.045	2466.69	6	0.000	0.066	0.045
7	0.832	0.045	2813.69	7	0.000	0.153	0.045
8	0.820	0.045	3151.74	8	0.000	-0.092	0.045
9	0.809	0.045	3481.05	9	0.000	0.035	0.045
10	0.802	0.045	3805.02	10	0.000	0.083	0.045
11	0.795	0.044	4124.08	11	0.000	0.049	0.045
12	0.785	0.044	4436.44	12	0.000	-0.046	0.045
13	0.775	0.044	4741.38	13	0.000	-0.011	0.045
14	0.766	0.044	5039.72	14	0.000	0.038	0.045
15	0.758	0.044	5332.69	15	0.000	0.072	0.045
16	0.752	0.044	5621.29	16	0.000	0.038	0.045

The ACF and PACF plots given in Figure 4.25 shows that all spikes of ACF fell outside the 95% confidence limit and PACF also showed spikes at certain lags (1,2,4,5,7) which are outside the confidence limit. It indicates non stationarity of SI tea price series.

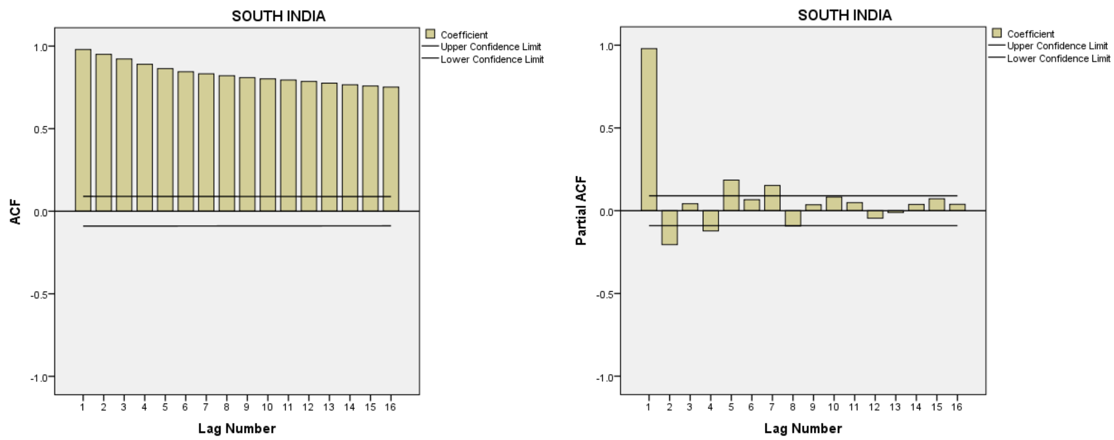

Figure 4.25 ACF and PACF plots of tea price in SI

Augmented Dickey Fuller (ADF) test was conducted for tea price in SI to confirm the non-stationarity of the data and the results are presented in the Table 4.1. It could be observed that after first differencing the data became stationary.

Table 4.17 ADF test for tea prices in SI

	ADF test statistic	Probability (p)	Critical values
Level	-1.03	0.74	-3.45
First difference	-19.08**	0.00	-2.87

** indicates significance at 1%level (p<0.01)

Log transformation was done to attain stationarity in variance. From among several ARIMA models tried, SARIMA (0,1,1) $(1,0,1)_{12}$ was chosen as the best forecast model for SI tea price, based on agreement between actual and fitted price plots, MAPE and residual ACF and PACF plots.

The plots of actual, fit, Lower Confidence Limit (LCL), Upper Confidence Limit (UCL) and forecast values for SI tea price for SARIMA (0,1,1) $(1,0,1)_{12}$ are provided in figure 4.26. From figure close agreement could be observed between actual and fitted values.

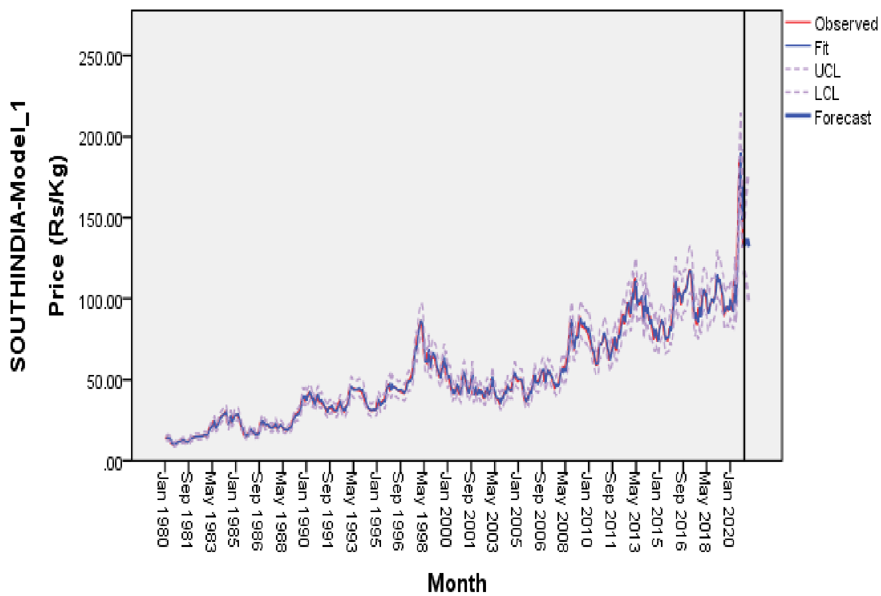

Figure 4.26 Observed, Fit, LCL, UCL, and forecast values with SARIMA(0,1,1)(1,0,1)$_{12}$ for tea price in SI

Residual ACF and PACF plots are provided in Figure 4.27. It could be observed that all ACF and PACF values lie within the confidence limits.

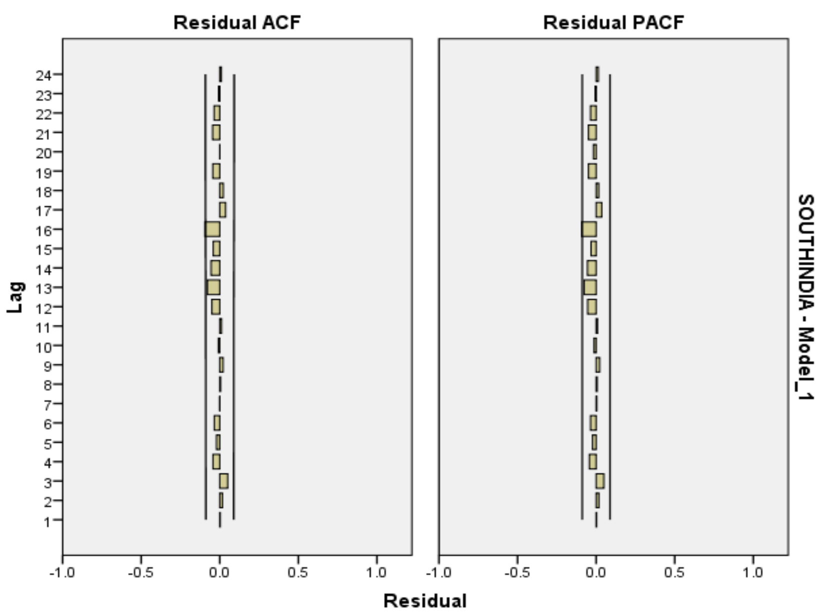

Figure 4.27 Residual ACF and PACF plots for SARIMA(0,1,1)(1,0,1)$_{12}$

The parameters of the model SARIMA (0,1,1) $(1,0,1)_{12}$ along with their tests of significance are provided in Table 4.18.

Table 4.18 SARIMA (0,1,1)$(1,0,1)_{12}$ model parameters for tea price in SI

Transformation	**Model Parameters**	**Estimate**	**SE**	**t**	**Probability(p)**
x_t=log$_e$ x_t	**Non-seasonal difference**	1			
	MA Lag 1 (θ_1)	-0.12	0.46	-2.67**	0.008
	AR, Seasonal Lag 1(Φ_1)	0.99	0.01	95.31**	0.00
	MA, Seasonal Lag 1(Θ_1)	0.97	0.053	18.27**	0.00

** indicates significance at 1% percentage level (p<0.01)

Thus, SARIMA(0,1,1)$(1,0,1)_{12}$ model for tea price in SI is

$$(1-\Phi_1 B^{12})(1-B)x_t = (1-\theta_1 B)(1-\Theta_1 B^{12})\varepsilon_t$$

Where, θ_1= -0.12, Φ_1= 0.99 and Θ_1= 0.97

With these parameters, the SARIMA(0,1,1)$(1,0,1)_{12}$ model is

$$(1-0.99B^{12})(1-B)x_t = (1+0.12B)(1-0.97B^{12})\varepsilon_t$$

Different accuracy measures of the above model are given in Table 4.19. It could be observed that R^2 was very high (97%) and MAPE was very small (4.62) for the model.

Table 4.19 Model fit statistics of SARIMA (0,1,1)$(1,0,1)_{12}$ for tea price in SI

R^2	**RMSE**	**MAPE (%)**	**MAE**	**SBC**
0.97	4.59	4.62	2.58	3.08

Based on the fitted model, tea price for SI were forecasted from January 2021 to April 2021 and provided in Table 4. 20 along with LCL and UCL.

Table 4.20 Forecasts of tea price in SI

Month	Forecast Price (₹/ Kg)		
	Forecast price	LCL	UCL
Jan-2021	132.51	117.02	149.46
Feb-2021	135.84	112.52	162.57
Mar-2021	135.74	107.15	169.62
Apr-2021	135.73	102.89	175.74

4.4.3 Comparison of forecast models for tea price in SI

For the tea price of SI, HWMS and SARIMA (0,1,1) $(1,0,1)_{12}$ had low MAPE values. Price forecasted using these models were compared with actual price and is provided in Table 4.21.

Table 4.21 Comparison of forecast models for tea price in SI

Month	Actual Price (₹/ Kg)	Forecast Price (₹/ Kg)	
		SARIMA(0,1,1)$(1,0,1)_{12}$	HWMS
Jan-21	139.14	132.51	131.68
Feb-21	144.34	135.84	135.65
Mar-21	139.74	135.74	135.28
Apr-21	124.94	135.73	135.62
MAPE		4.62	4.76

From the table it could be observed that SARIMA(0,1,1)$(1,0,1)_{12}$ was in more agreement with the actual price and MAPE was low compared to HWMS model. Thus for tea prices of SI SARIMA(0,1,1)$(1,0,1)_{12}$ was selected as the best forecast model.

4.5 Volatility of tea prices

The intra and inter annual volatility in tea prices in both North and SI are discussed in the following sections.

4.5.1 Intra annual volatility

Intra annual volatility indices of monthly tea prices in NI and SI from 1980 to 2020 are presented in Table 4.22. It measures the dispersion of prices within a year.

Table 4.22 Intra annual volatility indices of tea prices in NI and SI

Year	Intra-annual volatility indices (%) of tea prices	
	NI	SI
1981	5.79	4.18
1982	5.11	1.69
1983	8.53	7.83
1984	9.52	9.02
1985	10.13	9.16
1986	11.12	8.03
1987	8.56	4.68
1988	6.58	4.24
1989	7.56	5.09
1990	8.39	5.77
1991	9.42	3.96
1992	10.84	5.51
1993	6.2	5.39
1994	3.9	2.91
1995	13.66	6.73
1996	3.98	2.11
1997	6.93	3.98

1998	9.64	8.77
1999	13.19	6.62
2000	14.38	6.05
2001	15.87	11.12
2002	14.21	6.08
2003	10.04	7.3
2004	14.05	6.41
2005	12.29	6.7
2006	7.68	6.26
2007	11.06	4.68
2008	11.25	8.13
2009	11.4	4.59
2010	11.29	6.22
2011	15.92	5.54
2012	16.35	5.3
2013	3.76	7.96
2014	13.98	5.21
2015	15.28	4.81
2016	10.08	5.9
2017	9.75	6.78
2018	14.49	4.79
2019	12.24	5.31
2020	21.68	14.89

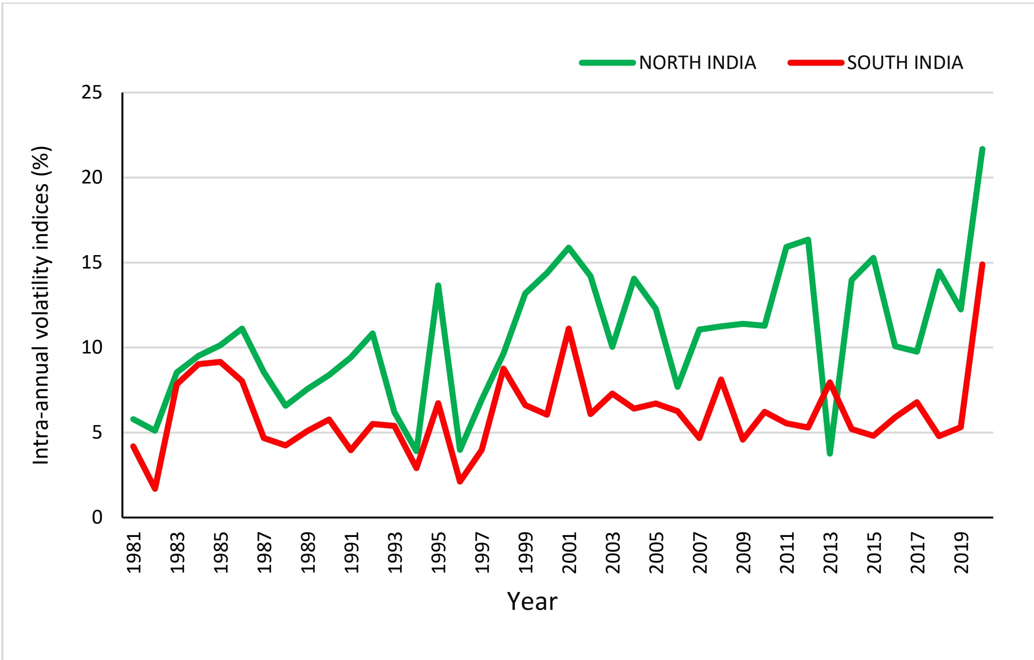

Figure 4.28 Intra annual volatility indices of monthly tea prices (%) in NI and SI

From Figure 4.28 it could be observed that the intra-annual volatility of monthly tea prices in both markets was varying irregularly. As compared to NI, SI showed only narrow variations. Both NI and SI showed almost same pattern from 1981 to 1990. Intra annual volatility in the tea prices shows the instability of returns gained by the producers within a year. High intra annual volatility shows that price is more deviated from the average, indicating an unsteady pattern in the price. A hike in tea price could be observed in 2020 in both NI and SI due to covid-19 pandemic lock down.

4.5.2 Inter annual volatility

Inter annual volatility measures the dispersion of tea prices between two successive years. Inter annual volatility indices of monthly prices of tea in NI and SI from 1980 to 2020 are presented in the Table 4.23.

Table 4.23 Inter annual volatility indices of tea prices in NI and SI

Year	Inter annual volatility indices (%) of tea prices	
	NI	SI
1980	10.88	21.16
1981	16.48	11.84
1982	14.86	5.98
1983	31.13	27.95
1984	20.84	16.76
1985	19.20	38.42
1986	28.20	23.75
1987	14.95	7.90
1988	15.21	9.62
1989	32.17	27.73
1990	12.15	10.64
1991	20.82	11.40
1992	23.85	11.72
1993	32.51	4.80
1994	12.00	15.45
1995	29.58	18.56
1996	7.12	5.73
1997	31.54	36.19
1998	17.59	23.97
1999	29.88	13.99
2000	29.12	14.52
2001	25.07	16.95
2002	33.94	8.55

2003	11.69	21.66
2004	26.01	14.77
2005	17.69	20.61
2006	20.08	10.48
2007	20.14	12.40
2008	30.25	25.93
2009	30.17	11.33
2010	28.45	15.13
2011	29.51	12.37
2012	38.05	15.74
2013	8.99	16.26
2014	29.68	14.59
2015	32.44	12.56
2016	20.91	8.36
2017	22.95	19.98
2018	30.99	7.99
2019	22.96	13.28
2020	68.59	41.75

It is evident from the inter annual volatility indices plotted in Figure 4.29 that the volatility indices of both NI and SI shows the same pattern from 1980 to 1989. Large variation could be observed from 2019 to 2020 in both NI and SI. According to tea growers, the increase in tea prices in 2020 is due to a reduction in production for about a month (from the end of March to the beginning of May) during the Covid -19-induced national lock down apart from the prolonged rainfall.

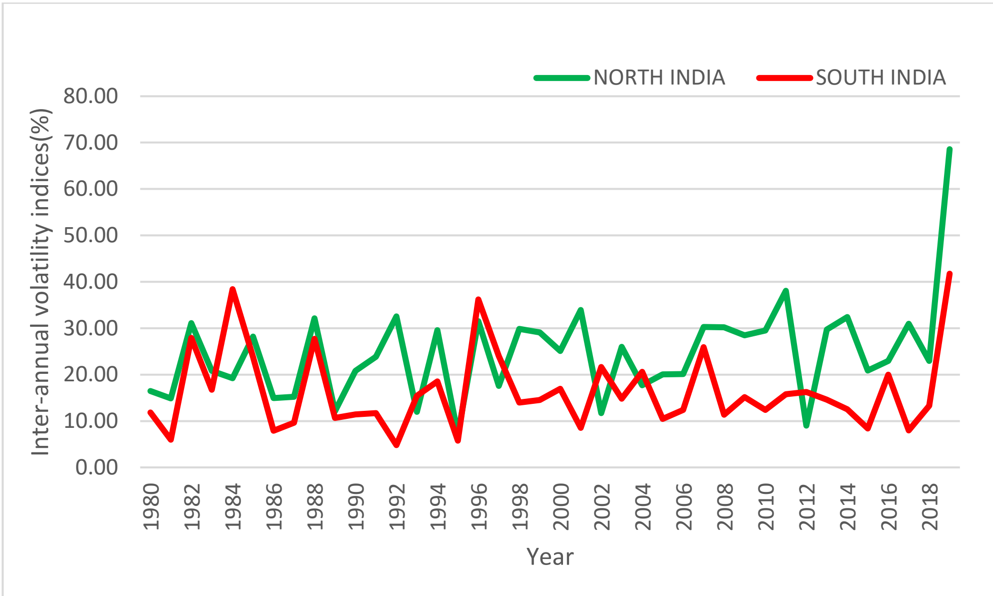

Figure 4.29 Inter annual volatility indices of monthly tea prices (%) in NI and SI

4.5.3 Significance of volatility - ARCH Model

Intra and inter-annual volatility indices only provide information on the magnitude of volatility; it does not indicate whether the estimated volatility is statistically significant. To check the significance, ARCH model was fitted for tea prices of NI and SI.

4.5.3.1 ARCH Model for tea price in NI

The estimates of the ARCH (1) model fitted for the NI tea price is given in the Table 4.24. It could be observed that the constant term and the ARCH parameter are positive and highly significant.

Table 4.24 Estimates of ARCH (1) model for tea price in NI

Parameter	Coefficient	Std.Error	Z-statistic
Constant term (b_0)	0.011	0.003	4.03**
ARCH term (b_1)	0.980	0.163	6.04**

**significant at 1% level

The ARCH (1) model is given by,

$$h_t = 0.011 + 0.98\, u_{t-1}^2$$

The time varying volatility includes a constant (0.011) and a component which depends on past errors ($0.98u_{t-1}^2$). The Z-statistic of the first order coefficient (6.04) suggests a significant ARCH (1) coefficient. Also, this model fits the assumptions $b_0>0$ and $0\leq b_1<1$. The volatility for NI tea price was found to be significant as indicated by the ARCH term.

4.5.3.2 ARCH Model for SI tea price

The estimates of the ARCH (1) model fitted for the SI tea price is given in the Table 4.25. It could be observed from the table that the constant term is positive and highly significant.

Table 4.25 Estimates of ARCH (1) model for tea price in SI

Parameter	Coefficient	Std.Error	Z-statistic
Constant term (b_0)	0.0028	0.0005	4.87**
ARCH term (b_1)	0.99	0.27	3.60**

**significant at 1% level

The ARCH (1) model is given by,

$$h_t = 0.0028 + 0.99\, u_{t-1}^2$$

The time varying volatility includes a constant (0.0028) and a component which depends on past errors (0.99 u_{t-1}^2). The Z-statistic of the first order coefficient (3.60) suggests a significant ARCH (1) coefficient. Also, this model fits the assumptions $b_0>0$ and $0\leq b_1<1$.The estimate of the ARCH parameter showed a high volatility for tea prices in SI.

Thus, ARCH (1) model was identified to be the best fit for volatility of tea prices in both NI and SI and high significant volatility was observed for prices in both markets.

4.6 Cointegration analysis

The concept of market integration can be used to describe the relationship between prices in two or more markets that are spatially separated. Johansen's cointegration method is the most widely used tool to study market integration. In the present study, the integration between the domestic markets, NI, SI and All India, and international markets Mombasa and Colombo were analysed using monthly tea price data from 1980 to 2020.

Before attempting cointegration tests, Augmented Dickey Fuller (ADF) test was performed to confirm that the different price series were non-stationary. A time series exhibits stationarity if the underlying generating process is based on a constant mean or a constant variance. The estimated test statistics from ADF tests for all markets at levels and first difference are presented in Table 4.26. The null hypothesis of non-stationarity for tea prices in the different markets was rejected after first differencing. Thus, all the price series became stationary by taking first difference.

Table 4.26 ADF test for monthly tea prices in the different markets

Market		t-statistic	Probability(p)
Levels	**NI**	0.78	0.99
	SI	-1.03	0.74
	Mombasa	-0.81	0.81
	Colombo	-0.15	0.94
	All India	-0.39	0.90
First difference	**NI**	-6.62**	0.00
	SI	-19.08**	0.00
	Mombasa	-10.98**	0.00
	Colombo	-17.09**	0.00
	All India	-6.61**	0.00

** Denotes significant at one percent level ($p<0.01$)

4.6.1 Johansen's cointegration test

Before estimating cointegration, optimal lag length of the markets has to be determined. Schwarz Information Criterion (SIC) was used to determine the lag length for testing the cointegration between the market pairs, NI- Mombasa, SI- Colombo, NI- SI, All India- Mombasa, All India- Colombo and the results of lag order selection are provided in Table 4.27.

Table 4.27 Lag order selection for different markets pairs

Market pairs	Lags	SIC
NI and Mombasa	0	19.75
	1	13.96
	2	13.85[s]
SI and Colombo	0	19.11
	1	12.58[s]
NI and SI	0	18.17
	1	13.37
	2	13.24[s]
All India and Mombasa	0	17.86
	1	8.26
	2	8.25
	3	8.21[s]
All India and Colombo	0	18.5
	1	8.38
	2	8.41
	3	8.35[s]

s-Smallest SIC value

Optimum lag length was obtained as:

Optimum lag length = Lag corresponding to the smallest SC value - 1

If the lag corresponding to the smallest SIC value is one, it was taken as such.

From Table 4.27, it could be concluded that the optimum lag length to be taken into consideration was, one, for the pairs NI - Mombasa, SI - Colombo and NI - SI. But for All India - Mombasa and All India - Colombo, optimum lag length selected was two. After selection of lag length, cointegration analysis was performed for the different market pairs. The results of pairwise cointegration tests are presented in Table 4.28. The following null hypothesis were used for determining the rank (r) of cointegration.

$H_0 : r = 0 \Rightarrow$ No cointegration between two markets

$H_0 : r \leq 1 \Rightarrow$ At most one cointegration between two markets

The null hypothesis of no cointegration ($r = 0$) was rejected for the market pairs viz, NI - Mombasa, SI - Colombo and NI - SI, as the trace statistics were significant (Table 4.28). The null hypothesis of presence of cointegration ($r \leq 1$) was confirmed in these market pairs since the trace statistics were nonsignificant. Thus, it could be observed that one cointegration relationship exists between all these three market pairs. It indicated that the variation in tea price in one market influences the price in the other market. But for the market pairs, All India - Mombasa and All India - Colombo, the null hypothesis of no cointegration ($r = 0$) was accepted indicating absence of cointegration between them. This is because of the reason that All India prices are the simple average of both SI and NI tea prices.

Table 4.28 Pairwise cointegration tests between different markets for tea price

Markets	Eigen value	Null	Trace statistics	Probability(p)
NI - Mombasa	0.134	r =0	72.17**	0.00
	0.003	r ≤1	1.58	0.20
SI - Colombo	0.073	r =0	30.89**	0.00
	0.002	r ≤1	0.03	0.85
NI - SI	0.220	r =0	123.69**	0.00
	0.002	r ≤1	1.24	0.26
All India - Mombasa	0.028	r= 0	14.14	0.08
	<0.0001	r ≤1	0.01	0.90
All India - Colombo	0.015	r= 0	7.47	0.52
	<0.0001	r ≤1	0.00	0.99

** Denotes significance at one percent level (p<0.01)

From the trace statistics in Table 4.28, one cointegrating equation indicates a long run relationship between the market pairs viz. NI - Mombasa, SI - Colombo and NI - SI. But deviations from this equilibrium can occur in the short run. Hence, it is necessary check whether such disequilibrium converges to equilibrium in the long run or not. So, further analysis was performed for the market pairs which are cointegrated and are given in the following sections.

4.6.2. Cointegration between NI and Mombasa tea markets

The plot for the market pair NI and Mombasa is provided in the Figure 4.30.

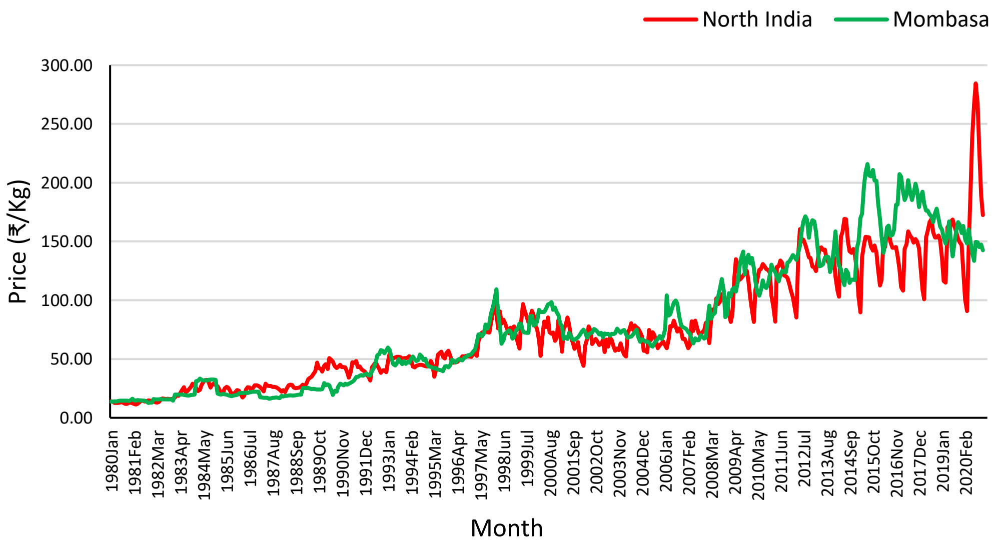

Figure 4.30 Tea price pattern of NI and Mombasa

4.6.2.1 Vector Error Correction Model for NI and Mombasa

Vector Error Correction Model (VECM) was done for NI and Mombasa tea markets to generate short run dynamics. Before performing VECM, normalized cointegrating coefficients were estimated and are provided in the Table 4.29.

Table 4.29 Normalized cointegrating coefficients of NI and Mombasa

<table>
<tr><th>MOMBASA</th><th colspan="2">NI</th></tr>
<tr><td rowspan="2">1.00</td><td>Estimate</td><td>Standard Error</td></tr>
<tr><td>-1.17</td><td>0.05</td></tr>
</table>

The sign of coefficients should be reversed in normalized cointegrating equation which is representing the long run. Mombasa is the dependent variable and from the table it could be observed that NI has a positive impact on Mombasa in the long run. The results of VECM for NI and Mombasa are presented in the Table 4.30.

Table 4.30 Vector Error Correction Model (VECM)- NI and Mombasa

	Coefficient	**Standard Error**	**t- statistic**
Constant term	0.211	0.27	0.77
Error Correction Term (ECT)	-0.047*	0.01	4.15
D(Mombasa(-1))	0.173*	0.04	3.92
D(NI(-1))	0.011	0.03	0.41

Critical value at 5% level of significance: 1.96

* Indicates significance at 5% level

D denotes first difference and (-1) denotes first lag

The cointegrating equation with Mombasa as dependent variable is given by,

$$\Delta \text{ Mombasa}_t = 0.211 - 0.047 \text{ EC}_{t-1} + 0.173 \ \Delta \text{ Mombasa}_{t-1} + 0.011 \ \Delta \text{ NI}_{t-1}$$

From the above equation, it could be observed that the coefficient of Error Correction Term (ECT) of Mombasa is negative and significant, indicating that there is a convergence from short run dynamics towards long run equilibrium. The size of the coefficient indicated that the speed of adjustment to equilibrium is 4.7 percent. A percentage increase in tea price for Mombasa will lead to an increase in Mombasa by 17.3 percent and a percentage increase in NI will lead to an increase in Mombasa by 1.1 percent.

4.6.2.2 Granger causality test

Cointegration analysis does not provide information regarding the direction of flow of information on prices, *i.e.*, whether it is in one direction or in both directions. The Granger causality tests was done to check in which direction price transmission is occurring. The results are presented in Table 4.31.

Table 4.31 Granger causality test for monthly tea prices of NI and Mombasa

Null hypothesis	F statistic	Probability(p)
NI does not Granger Cause Mombasa	15.36**	0.00
Mombasa does not Granger Cause NI	21.35**	0.00

* *Denotes significant at one percent level($p<0.01$)

From Table 4.31, it could be observed that, the F statistic used to test causality is highly significant for both the null hypothesis viz., NI does not Granger cause Mombasa and Mombasa does not Granger cause NI. Thus, it indicates the existence of bidirectional causality between NI and Mombasa tea markets.

4.6.3. Cointegration between SI and Colombo tea markets

The tea price pattern for the market pair SI and Colombo is provided in the Figure 4.31.

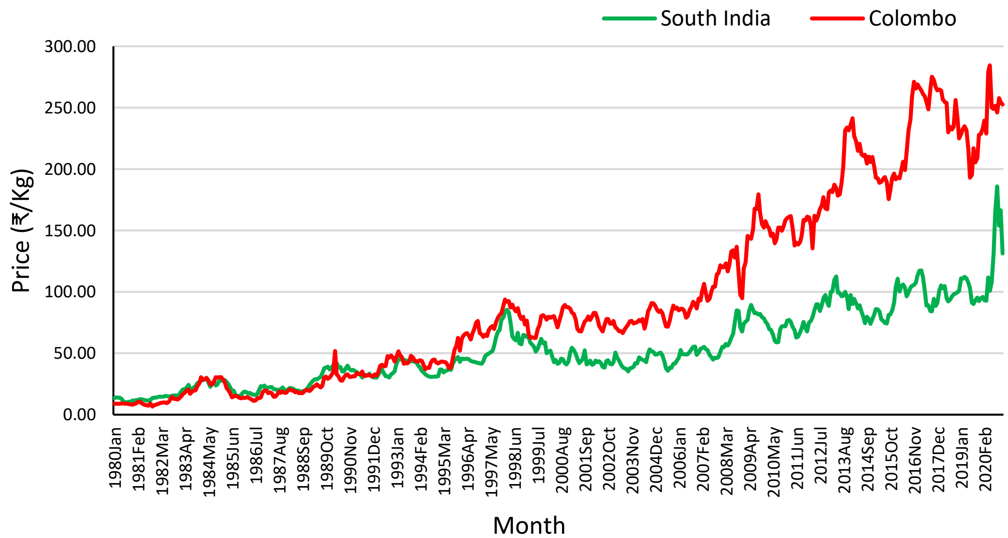

Figure 4.31 Tea price pattern of SI and Colombo tea markets

4.6.3.1 Vector Error Correction Model for SI and Colombo

Vector Error Correction Model (VECM) was done for SI and Colombo tea markets to generate short run dynamics. Before performing VECM, normalized cointegrating coefficients were estimated and are provided in the Table 4.32.

Table 4.32 Normalized co integrating coefficients of SI and Colombo

COLOMBO	SI	
	Estimate	Standard Error
1.00	-2.68	0.16

From Table 4.32, it could be observed that SI has a positive impact on Colombo in long run. The results of VECM for SI and Colombo are presented in Table 4.33.

Table 4.33 Vector Error Correction Model (VECM)- SI and Colombo

	Coefficient	Standard Error	t- statistic
Constant term	0.458	0.295	1.55
Error Correction Term (ECT)	-0.017	0.012	1.63
D(Colombo(-1))	0.048	0.045	1.04
D(SI(-1))	0.046	0.069	0.68

Critical value at 5% level:1.96

D denotes first difference and (-1) denotes first lag

The cointegrating equation with Colombo as dependent variable is given by,

$$\Delta \text{ Colombo}_{t} = 0.458-0.017 \text{ EC}_{t-1} + 0.048 \, \Delta \text{ Colombo}_{t-1} + 0.046 \, \Delta \text{ SI}_{t-1}$$

The estimated coefficient of error correction term in the Colombo equation is negative and nonsignificant (-0.017) which indicates the lack of significant adjustments towards long run equilibrium in any disequilibrium or volatile situation. Thus, it could be concluded that the previous period deviation from long run equilibrium is corrected in the current period with an adjustment speed of 1.7 percent. A percentage increase in itself (Colombo) will lead to an increase in Colombo by 4.8 percent and a percentage increase in SI will lead to an increase in Colombo by 4.7 percent.

4.6.3.2 Granger causality test

Granger causality test was performed for SI and Colombo to know the direction of price transmission and the results are provided in the Table 4.34.

Table 4.34 Granger causality tests for monthly tea prices of SI and Colombo

Null hypothesis	F statistic	Probability(p)
SI does not Granger Cause Colombo	3.24	0.07
Colombo does not Granger Cause SI	19.37**	0.00

**Denotes significant at one percent level(p<0.01)

From Table 4.34 it could be observed that the null hypothesis that Colombo does not Granger cause SI was rejected at one per cent level of significance. The results of the analyses proved the existence of unidirectional causality between SI and Colombo.

4.6.4 Cointegration between NI and SI tea markets

The tea price pattern for the market pair NI and SI is provided in the Figure 4.32.

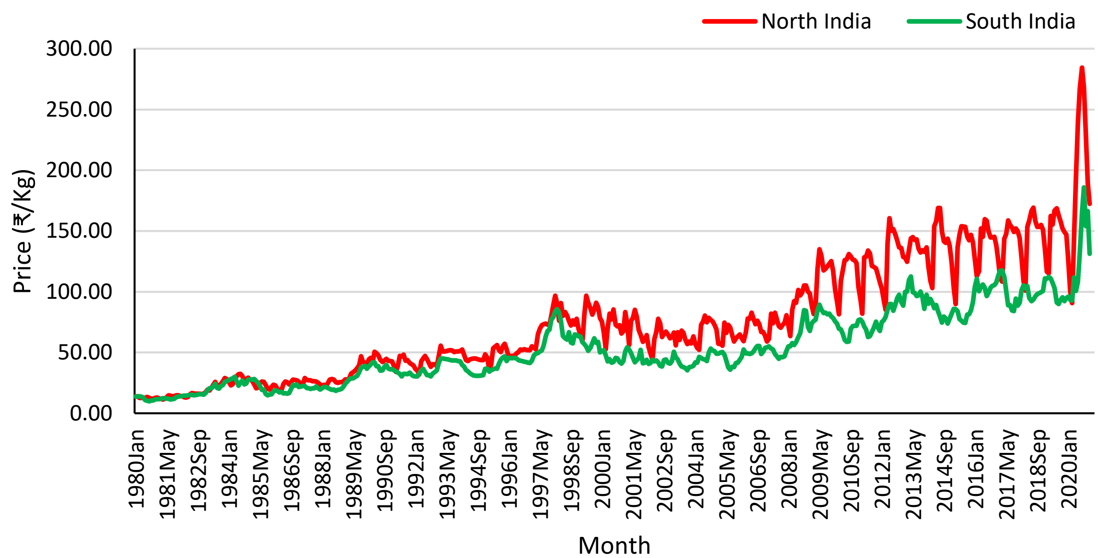

Figure 4.32 Tea price pattern of NI and SI

4.6.4.1 Vector Error Correction Model for NI and SI

Normalized cointegrating coefficients of NI and SI are provided in Table 4.35. NI is the dependent variable. From the table, it could be observed that SI has a positive but nonsignificant impact on NI in long run.

Table 4.35 Normalized cointegrating coefficients of NI and SI

NI	SI	
	Estimate	Standard Error
1.00	-1.56	0.04

Vector Error Correction Model (VECM) was done for NI and SI tea markets to generate short run dynamics. The results are presented in Table 4.36.

Table 4.36 Vector Error Correction Model (VECM)- NI and SI

	Coefficient	Standard Error	t- statistic
Constant term	0.282	0.42	0.66
Error Correction Term (ECT)	-0.216*	0.03	8.08
D(NI(-1))	0.372*	0.04	8.40
D(SI(-1))	-0.293*	0.09	3.03

critical value at 5% level:1.96, * indicates significance at 5% level

D denotes first difference and (-1) denotes lag one

The cointegrating equation (NI as dependent variable) is given by,

$$\Delta NI_t = 0.282 - 0.216\ EC_{t-1} + 0.372\ \Delta NI_{t-1} - 0.293\ \Delta SI_{t-1}$$

The coefficient of Error Correction Term (ECT) of NI is negative and statistically significant indicating that there is a convergence from short run dynamics towards long run equilibrium. The previous period deviation from long run equilibrium is corrected in the current period with an adjustment speed of 21.6 percent. A percentage increase in itself (NI) will lead to an increase in NI by 37.2 percent and a percentage increase in SI will lead to a decline in NI by 29.3 percent. 0.282 is the constant or intercept.

4.6.4.2 Granger causality test for NI and SI

Granger causality test was performed for NI and SI to know the direction of price transmission and the results are provided in the Table 4.37.

Table 4.37 Granger causality tests for monthly tea prices of NI and SI

Null hypothesis	F statistic	Probability
SI does not Granger Cause NI	20.35**	0.00
NI does not Granger Cause SI	36.70**	0.00

**Denotes significant at one per cent level ($p<0.01$)

As can be seen in Table 4.37, the F statistic used to test causality is highly significant in the case of the null hypothesis that, SI does not Granger cause NI. Similarly, the null hypothesis NI does not Granger cause SI was also rejected indicating a bidirectional causality between NI and SI markets.

From the cointegration analysis carried out for tea prices it was observed that a cointegrating relationship exist between the market pairs – NI and SI, NI and Mombasa, SI and Colombo and no cointegration exist between the market pairs-All India and Mombasa and All India and Colombo. Unidirectional causality was observed between SI and Colombo whereas bidirectional causality was observed between market pairs NI and Mombasa and NI and SI.

Summary

5. SUMMARY

The summary of the study entitled "Time series modeling and forecasting of tea prices in India" is provided in this section. The study was conducted to analyse the components of time series data on prices of tea in India, to develop time series forecast models for the prices, to develop statistical models for price volatility and to study the integration between international and Indian tea prices.

Time series data on monthly auction prices of tea for NI, SI and All India for the period from January 1980 to December 2020 collected from Tea Board of India formed the main database for the study. International price of tea for Colombo (Sri Lanka) and Mombasa (Kenya) for the period from January 1980 to December 2020 were also collected to analyse the cointegration between domestic and international markets. To have an idea about the trend in A- Pr- Pd of tea in India, annual data on A- Pr- Pd of tea from 1970 to 2019 in NI, SI and All India were also collected.

Trend equations were fitted for A- Pr- Pd of tea in NI, SI and All India. From among several models tried, best model was selected based on the criteria like MAPE and Adjusted R^2. It was observed that quadratic model was found to be the best fit for area under tea in NI, while, cubic model was found to be the appropriate fit for production and productivity of tea in NI and, A- Pr- Pd of tea in SI and All India. To study the extent and proportion of growth in A- Pr- Pd of tea in India, Compound Annual Growth Rate (CAGR) was estimated for whole period (1970-2019) and by dividing the whole period into two- period I (1970-1995) and Period II (1996-2019). For overall period, growth rate was positive for A- Pr- Pd. For NI, growth rate in tea production was more in period II compared to period I, corresponding to the growth in area. For SI, a decline in growth rate in production was noticed from period I to period II. Growth rate in productivity was found to be more during period I in both NI and SI compared to period II.

NI and SI tea price data were decomposed into the time series components viz., trend, seasonal variation, cyclic variation and irregular variation. Tea prices in both NI and SI showed an overall increasing trend and prominent seasonal variation. All India tea price was found to be the simple average of NI and SI tea prices. Seasonality observed in All India tea price was the combined seasonality of both NI and SI tea prices. Seasonal indices were computed using ratio to moving average method for tea prices for NI and SI, to get a clear picture of price pattern in the 12 months from January to December. For NI, April to November are the high price months with highest price in June whereas, December to March are the low-price months with lowest price in March. For SI, high price was observed from January to April with highest price in February and low-price months were from June to September with lowest price in July. Cyclic variations were plotted for NI and SI tea prices. For NI, a prominent cycle of length 10 to 11 years was observed from 1998 to 2009. For SI, a small cycle of length 5 to 6 years was observed from 1984 to 1990 and a prominent cycle was observed from 1998 to 2009. Tea price in SI exhibited more cycles of variation compared to NI.

Different exponential smoothing models (Single exponential, double exponential and Holt-Winters' multiplicative seasonal (HWMS) models) and ARIMA models were fitted for the tea prices in NI and SI. For fitting the forecast models, price data from January 1980 to December 2020 were used and forecasts were made from January 2021 to April 2021. The model for which actual and forecasted price were in close agreement and having least value for MAPE was selected as the best forecast model. For tea prices in NI, SARIMA$(0,1,3)(0,1,1)_{12}$ was selected as the best forecast model. In the case of SI, SARIMA$(0,1,1)(1,0,1)_{12}$ was identified as the best model to forecast tea price.

Volatility of tea prices for NI and SI were estimated using intra and inter annual volatility indices and its significance was tested by fitting suitable ARCH model. Estimates of intra annual volatility for monthly tea prices showed that, tea prices in both regions were varying irregularly. As compared to NI, SI showed only narrow variations. A hike in tea price was observed in 2020 in both NI and SI due to the effect of national lock down following the covid-19 pandemic. This was established from the study of inter annual volatility, where, large variation was observed from 2019 to 2020 in both NI and SI. Intra and inter annual volatility indices give only the magnitude of volatility, it will not provide information on whether the estimated volatility is statistically significant or not. To check the significance, ARCH (1) model was fitted for tea prices of NI and SI. The estimate of the ARCH parameter showed high volatility for tea prices in both NI (0.98) and SI (0.99).

Cointegration analysis for tea prices indicated that one cointegrating relationship exists between the market pairs, NI - SI, NI – Mombasa and SI -Colombo. No cointegration exist between the market pairs, All India -Mombasa and All India - Colombo. Unidirectional causality was observed between SI and Colombo indicating that SI tea price is influenced by Colombo tea price, whereas, bidirectional causality was observed between market pairs, NI - Mombasa and NI - SI.

Following conclusions were made from the study:

- Quadratic model was the best fit for trend in area under tea in NI
- Cubic model was the best fit for trend in production and productivity of tea in NI
- For trend in A- Pr- Pd of tea in SI, cubic model was the best fit
- Best fit for trend in A- Pr- Pd of tea in All India is cubic model
- For NI and SI, growth rate in area under tea was more during 1996-2019 compared to the period 1970-1995
- In NI, annual growth rate for production of tea was more during 1996-2019 compared to the period 1970-1995, while, for SI, a decline in production was observed during the period 1996-2019
- Overall increasing trend and a prominent seasonal variation were observed for tea prices in NI and SI
- Cyclic variation was found to be high for SI tea price compared to NI tea price
- For NI, highest price for tea was observed in June and lowest in March
- For SI, tea price was highest in February and lowest in July

- SARIMA(0,1,3)(0,1,1)$_{12}$ is the best forecast model for tea price in NI, while, for SI SARIMA(0,1,1)(1,0,1)$_{12}$ is the suitable model for tea price forecasting

- Intra and inter annual volatility indices were used to estimate the price volatility of tea prices in NI and SI and it was observed that, in most of the years, NI showed more volatility in tea price compared to SI

- A hike in tea price could be observed in the year 2020 in NI and SI due to national lockdown following the Covid-19 pandemic

- Significance of price volatility was tested by fitting ARCH (1) model and price volatility was found to be high for both NI and SI tea prices

- Market pairs, NI – Mombasa, SI - Colombo and NI - SI are cointegrated

- SI tea price is influenced by Colombo tea price

- Tea prices in NI and Mombasa influence each other

- Tea prices in the domestic markets NI and SI influence each other

References

REFERENCES

Adanaliglu, H. and Yercanm, M. 2012. An analysis of tomato prices at wholesale level in Turkey: An application of SARIMA model [on-line]. Available: http://www.custoseagronegocioonline.com.br. pdf (8 June 2020).

Anonymous. 2019. Tea planting variety and harvesting season [on-line]. Available: http://www.myassamtea.com/tea-plantation/tea-harvesting. [04 Aug. 2021].

Anderson, T. W. 1971. The statistical analysis of time series. John Wiley and Sons, New York, 704p.

Assis, K. Amran, A. and Remali, Y. 2010. Forecasting cocoa bean prices using univariate time series models. J. Arts Sci. and Commerce. 1(1): 71-80.

Babu, P. B. N., Rudragouda, C.S and Reddy, Y. B. V. 2019. Trends in A- Pr- Pd of coffee across the major coffee growing states in India. Int. J. Agric. Sci. 4(11): 7896-7900.

Barathi, R., Hanaldar, Y. N., Megeri, S. N., Patil, G. M., and Patil, B. L. 2011. A study on market integration of Ramanagara and Shidlaghatta markets and forecasting their prices and arrivals. Karnataka J. Agric. Sci. 24(3): 347-349.

Bhavani, D., Srikala, T. and Ananda. 2015. Price volatility in major chilli markets of India. Indian J. Econ. Dev. 3(3): 194-198.

Bodade, V. M., Deshmukh, R. G. and Prema, B. 2017. Soybean price movement across major markets of Madhya Pradesh. Int. Res. J. Agric. Eco. Stat. 8(1): 162-169.

Booranawong, T. and Booranawong, A. 2017. Simple and double exponential smoothing methods with designed input data for forecasting a seasonal time series: In an application for lime prices in Thailand. Suranaree J. Sci. Technol. 24(3): 301-310.

Box, G. P. E. and Jenkins, G. M. 1970. Time series analysis: Forecasting and control. John Wiley and Sons, New York, 746p.

Chaudhari, D. J. and Iingre, A. S. 2014. Use of ARIMA modeling for forecasting green gram prices for Maharashtra. J. Food Legumes. 27(2):136-139.

Cermak, M., Malec, K. and Maitah, M. 2017. Price Volatility Modelling – Wheat: GARCH Model Application [on-line]. Agris on-line Papers in Eco. Informatics. 9:15-24. [16 May. 2021].

Croxton, F. E., Cowden D. J., Klein, S. Applied general statistics. Prentice Hall of India, New Delhi, India, 1979, 754p.

Darekar, A. S., Pokharkar, V. G. and Datarkar, S. B. 2016. Onion price forecasting in Kolhapur market of Western Maharashtra using ARIMA technique. Int. J. Inf. Res. and Rev. 3(12): 3364-3368.

Darekar, A. and Reddy, A. A. 2017. Forecasting of common paddy prices in India. J. Rice Res. 10(1): 71-75.

Darekar, A. and Reddy, A. A. 2017. Cotton price forecasting in major producing states. Econ. Affairs. 62(3): 1-6.

Darvishi, G. A. and Indira, M. 2013. An analysis of changing pattern in A- Pr- Pd of coffee and tea in India. Int. J Marketing, Financial Services and Manag. Res. 2(9) 46-60.

Draper, N. R. and Smith, H. 1998. Applied regression analysis. John Wiley and Sons, New York, 736p.

Dudhat, A. S. 2017. A statistical analysis on price volatility and market integration in major domestic groundnut markets. PhD (Ag) thesis, Junagadh Agricultural University.

Engle, R. F. 1982. Autoregressive conditional heteroskedasticity with estimates of variance of U.K inflation. Econometrica. 50: 987-1008.

Engle, R. F. and Granger, C.W. J. 1987. Cointegration and error correction: representations, estimation and testing. Econometric. 55: 251-276.

EViews Package. Student Version 11.

Garman, M. B. and Klass, M. J. 1980. On the estimation of price volatility from historical data. J. Business. 53: 67-78.

Granger, C. W. J. 1980. Testing for causality: A personal viewpoint. J. Economic. Dyn. Control. 2: 329–352.

Gilbert, C. L. 2006. Trends and Volatility in Commodity Prices. In: Sarris, A. and Hallam, D. (eds.), Agricultural commodity markets - New Approaches to Analysing Market Structure and Instability. FAO and Edward Elgar.

Gogulan, G. Velavan, C. Sivakumar, S. D. and Selvam, S. 2018. Market Integration of Selected Tomato Markets in Southern India. Indian J. Agric. Marketing. 33(2):24-34.

Guha, B. and Bandyopadhyay, G. 2016. Gold Price Forecasting Using ARIMA Model. J. Adv. Manag., Sci. 4(2): 117-121.

Indraji, K. N. 2016. Price forecast models for coconut and coconut oil. M.Sc. (Ag. Stat.) thesis, Kerala Agricultural University, Thrissur, 123p.

Ishaque, M. and Ziblim, S. 2013. Use of some exponential smoothing models in forecasting some food crop prices in the upper east region of Ghana. Mathematical Theory and Modeling. 3(7):16-27.

Jadhav, V., Reddy, B. V. C. and Gadd. G. M. 2017. Application of ARIMA Model for Forecasting Agricultural Prices. J. Agric. Sci. Tech. 19: 981-992.

Johansen, S. 1998. Statistical analysis of cointegrating vectors. J. Econ. Dyn. Control. 12: 231-254.

Johansen, S. and Juselius, K. 1990. Maximum Likelihood estimation and inference on cointegration with application tom demand for money. Oxford Bull. Econ. Statist. 52(2): 169-210.

Jyothi, T. 2011. Price behaviour of Turmeric in India. M.Sc. (Ag) thesis, Kerala Agricultural University, Thrissur. 152p.

Kelkay, B. D., and Yohannes, E. 2014. The Application of Garch Family Models to Some Agricultural Crop Products in Amhara National Regional State. J. Econ. Sustainable Development. 5(11): 24-35.

Krishnarani, S. D. 2013. Time series outlier analysis of tea price data. Am. J. Theoret. and Appl. Stat. 2(1): 1-6.

Kumar, A. V. S. P. and Bhramaramba. 2017. Time series data analysis on agriculture food production. Adv. Sci. Tech. Letters. 147:520-525.

Kumar, N. S., Joseph, B. and Jaslam, P. K. M. 2017. Growth and instability in A- Pr- Pd of Cassava (*Manihot esculenta*) in Kerala. Int. J. Adv. Res. Ideas and Innovation in Technol. 4(1):446-448.

Kumar, M. T. L., Munirajappa, R., and Suresh, H. S. 2011. Application of Seasonal ARIMA model for forecasting monthly prices of potato in Bangalore market. Mysore J. Agric. Sci. 45(4): 778-782.

Kumaraswamy, N. and Sekar, C. 2014. Growth Performance and Price Forecasting of Potato - An Economic Analysis. Trends Biosci. 7(20): 3315-3320.

KumarMahto, A., Biswas, R. and Alam, M. A. 2019. Short term forecasting of agricultural commodity price by using ARIMA: Based in Indian market. In: International Conference on Advances in Computing and Data Sciences; Springer, Singapore, pp. 452-461.

Kunitomo, N. 1992. Improving the Parkinson methods of estimating security price volatilities. J. Business. 65: 295-302.

Lavanya, S., Arulanandu, U. and Selvam, S. 2018. Cointegration between domestic and international market prices of beverage crops in India. Madras Agric. J. 105(1-3):99-108.

Liu, H. and Shao, S. 2016. India’s tea price analysis based on ARMA model. Mod. Economy. 7: 118-123.

Mgale, Y. J., Yan, Y. X. and Timothy, S. 2021. A comparative study of ARIMA and Holts- Winters exponential smoothing models for rice price forecasting in Tanzania. Open Access Library J. 8:1-9.

Naidu, G. M., Sumathi, P., Reddy, B. R. and Kumari, V. M. 2014. Forecasting price behaviour of red chilli by using seasonal ARIMA model. Bioinfolet. 11(1B): 268-270.

Nain, M., Nisha, Sanjeev and Aneja, R. 2019. Instability and trend in A- Pr- Pd of rice crop in Haryana and India. Current J. Applied Sci. Technol. 37(5): 1-9.

Nyantakyi, K. A., Peiris, B., and Gunaratna, L. 2015. Analysis of interrelationships between prices of Sri Lanka rubber, tea and coconut production using multivariate time series. Adv. Econ. and Business. 3(2): 50-56.

Parkinson, M. 1980. The extreme value method for estimating the variance of the rate of return. J. Business. 53: 61-65.

Paul, R. K., Saxena, R. and Bhatt, S. A. 2016. How price signals are transmitted across regions and vale chain? Examining horizontal and vertical market price integration for major pulses in India. Agric. Econ. Res. Review. 29: 75-86.

Preethi, V. P. 2017. Economic analysis of production, marketing and price behaviour of coconut. M.Sc. (Ag) thesis, Kerala Agricultural University, Thrissur. 145p.

Rahmawati, R., Rusgiyono, A., Hoyyi, A. and Maruddani, D. A. I. 2019. GARCH-family for measuring price fluctuation risk of harvested dry grain in Pemalang district. J. Physics. 11p.

Rangoda, B. D. P., Abeywickrama, L. M. and Fernando, M. T. N. 2006. An Analysis of different forecasting models for prices of coconut products in Sri Lanka. Proceedings of the Third Academic Sessions, Mapalana, Sri Lanka. Coconut Research Institute, Sri Lanka, 8p.

Rani, R., Singh, R., Tewari, H., Singh, S K. and Singh, P. K. 2017. Integration of major Indian maize markets: A cointegration analysis. Int. J. Agric. Stat. Sci. 13(2):601-606.

Reddy, A. 2012. Market integration of grain legumes in India. The case of the chickpea market. SAARC J. Agri. 10(2):11-29.

Reddy, A. A. 2018. Price forecasting of Tomatoes. Int. J. Vegetable Sci. 25(2): 176-184.

Reeja, V. 2011. Price behaviour of natural rubber in India. M.Sc. (Ag) thesis, Kerala Agricultural University, Thrissur, 109p.

Rotich, H. C., Chelule, J. C. and Imboga, H. 2020. Analysis of tea auction prices using non- cointegration based techniques. Am. J. Theoretical and Applied Statistics. 9(6): 296-305.

Sabu, S. S. 2013. Price volatility of black pepper and its implications in Kerala. M.Sc. (Ag) thesis, Kerala Agricultural University, Thrissur, 168p.

Senaviratna, N. A. M. R. 2016. Forecasting tea auction prices in Sri Lanka: Box-Jenkins modeling approach. Ind. J. Appl. Res. 6(11):722-724.

Shana, K. K. 2018. Price behaviour of Nendran banana in Kerala: An economic analysis. M.Sc. (Ag) thesis, Kerala Agricultural University, Thrissur, 96p.

Sharma, H. and Burark, S. S. 2015. Forecasting price of sorghum in Ajmer market of Rajasthan: An empirical study. Ann. Agric. Res. New Series. 36(2):212-218.

Sharma, H., Burark, S. S. and Singh, H. 2014. Soyabean price forecasting in Rajasthan: Application of ARIMA model. Ann. Agri-Bio. Res. **19**(4):763-768.

Shukla, R. K., Negi, Y. and Chaudhry, S. 2017. A time series analysis of auction price of Indian tea. Int. J. Res. in Econ. and Social Sci. 7(6): 100-111.

Soe, W. P. P. and Fukuda, S. 2010. Analysis of the spatial market integration of oil seed crop markets in Myanmar. J. Fac. Agric. Kyushu Univ. 55(1): 159-166.

Spiegel, R. M. (1992). Theory and problems of statistics (2nd Ed.). Mc Graw-Hill International, Singapore. 504p.

Sutradhar, T., Hazari, S. and Nagpure, S. C. 2019. Effect of seasonal and cyclical variation in prices of natural rubber in India. Int. J. Sci. Environ. 8(3):696-701.

Sundaramoorthy, C., Jha, G. K., Pal, S. and Mathur, V. C. 2013. Market integration and volatility in edible oil sector in India. J. Indian Society Agric. Stat. 68(1):67-76.

Talukdar, U. and Hazarika, C. 2017. Production and export of value-added tea in India and its global competitiveness. Econ. Affairs. 62(4): 705-710.

Tea Board, 2020. [on-line]. Available: http://www. teaboard. gov.in. [04 June. 2020].

UPASI, 2021. [on-line]. Available: http://www. upasi.in. [04 June. 2021].

Vasanthakumar, T., Sudheendra, M. and Mallikarjuna, H. B. 2014. Forecasting of price of teak (Tectona grandis Linn.) – By exponential smoothing analysis. Environ. Ecol. 32(2): 566-570.

Verma, V. K., Kumar, P., Singh, H. and Singh, H. P. 2016. Use of ARIMA modelling in forecasting coriander prices of Rajasthan. Int. J. Seed Spices. 6(2): 40-45.

Vinayak., Jalikatti, N. and Patil., B. L. 2015. Onion price forecasting in Hubli markets of Northern Karnataka using ARIMA technique. Karnataka J. Agric. Sci. 28(2): 228-231.

Wani, M. H., Paul, R. K., Bazaz, N. H. and Manzoor, M. 2015. Market integration and price forecasting of apple in India. Indian J. Agric. Econ. 70(2):169-181.

Winters, P. R. 1960. Forecasting sales by exponentially weighted moving averages. Management Sci. 6(3): 324-342.

TIME SERIES MODELING AND FORECASTING OF TEA PRICES IN INDIA

By

DEENAMOL JOY

(2019-19-003)

ABSTRACT OF THE THESIS

Submitted in partial fulfilment of the requirement for the degree of

Master of Science in Agricultural Statistics

Faculty of Agriculture

Kerala Agricultural University, Thrissur

Department of Agricultural Statistics

COLLEGE OF AGRICULTURE

VELLANIKKARA, THRISSUR – 680 656

KERALA, INDIA

2021

ABSTRACT

The study entitled "Time series modeling and forecasting of tea price in India" was conducted to study the components of time series data on prices of tea in India, to develop time series forecast models for the prices, to develop statistical models for price volatility and to study the integration between international and Indian tea prices. Monthly auction prices of tea for North India, South India and All India for the period from January 1980 to December 2020 collected from the Tea Board formed the main database for the present study. International price of tea for Colombo (Sri Lanka) and Mombasa (Kenya) for the period from January 1980 to December 2020 were collected. To have an idea about the trend in A- Pr- Pd of tea in India, annual data on A- Pr- Pd of tea from 1970 to 2019 in North India, South India and All India were also collected.

To have a general idea about trend in A- Pr- Pd of tea in North India, South India and All India, models like exponential, quadratic, cubic etc were fitted. From among several models tried, quadratic model was found to be the best fit for area under tea in North India, while, cubic model was found to be the appropriate fit for production and productivity of tea in North India and, A- Pr- Pd of tea in South India as well as All India. North India and South India tea price data was decomposed to time series components like trend, seasonal variation, cyclic variation and irregular variation. North India and South India showed an overall increasing trend and a prominent seasonal variation. Cyclic variations showed that South India exhibited more cycle of price volatility compared to North India. All India tea price was found to be the simple average of North India and South India tea prices. Compound Annual Growth Rate (CAGR) was estimated for A- Pr- Pd of tea in North India and South India for the period from 1970 to 2019. For North India, growth rate in production was more during 1996-2019 compared to period 1970-1995. For South India, a decline in production was observed during 1970 to 1995.

Price forecast models like exponential Smoothing models and ARIMA models were fitted to forecast the tea prices in North India and South India from January 2021 to April 2021. For North India tea price, SARIMA $(0,1,3)(0,1,1)_{12}$ was identified as the

best forecast model whereas for tea price of South India SARIMA $(0,1,1)(1,0,1)_{12}$ was selected to forecast tea prices.

For tea prices in North India and South India, volatility in prices were estimated using intra and inter annual volatility and its significance was tested by fitting suitable ARCH model. Intra annual volatility indices of monthly tea prices in both regions were varying irregularly. In most of the years, North India showed large variation in tea price compared to South India. ARCH (1) model was fitted to check the significance of tea prices and the estimate of ARCH parameter showed high volatility for tea prices for North India and South India.

Cointegration analysis was carried out for tea prices to study the integration between international and domestic Indian tea markets. One cointegrating relationship exists between the market pairs, North India - South India, North India – Mombasa and South India – Colombo. No cointegration exist between the market pairs, All India -Mombasa and All India - Colombo. Unidirectional causality was observed between South India and Colombo whereas, bidirectional causality was observed between market pairs, North India - Mombasa and North India - South India.