2019-09-25

DEPARTMENT OF STATISTICS THE IMPACT OF SELECTED CLIMATE VARIABLES AND FERTILIZERS TYPE ON SELECTED CEREAL CROP PRODUCTION IN AMHARA REGION.

FASILEDES, FETENE

BAHIR DAR UNIVERSITY

DEPARTMENT OF STATISTICS

THE IMPACT OF SELECTED CLIMATE VARIABLES AND FERTILIZERS TYPE ON SELECTED CEREAL CROP PRODUCTION IN AMHARA REGION.

BY:-FASILEDES FETENE

ADVISOR:-SALIE AYALEW(PhD)

A THESIS RESEARCH SUBMITTED TO THE DEPARTMENT OF STATISTICS, COLLEGE OF SCIENCE, BAHIR DAR UNIVERSITY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTERS OF SCIENCE IN ECONOMETRICS.

July 18, 2019

BAHIR DAR,ETHIOPIA

Declaration

I, the undersigned, declare that this thesis is my original work, has not been presented for degrees in any other Universities and all sources of materials used for this thesis have been duly acknowledged. The assistance received during the course of these investigations has been duly acknowledged. Therefore, we recommend it to be accepted as fulfilling the thesis requirements.

Name of Student	Signature	Date

This thesis has been submitted for examination with my approval as a University advisor.

Name of Advisor	Signature	Date

Approval Sheet

We, the undersigned, members of the board of examiners of the final open defense by Fasiledes Fetene have been read and evaluated his thesis entitled **"The impact of selected climate variables and fertilizers type on selected cereal crop produced in Amhara region."** This is therefore to certify that the thesis has been accepted in partial fulfillment of the requirement for the degree of Master of Sciences in Statistics with specialization of Econometrics.

Name of Chairperson	Signature	Date
Name of External Examiner	Signature	Date
Name of Internal Examiner	Signature	Date

Acknowledgment

Words are powerless to express my praises and adoration to the Almighty God and his holy mother for their love, comfort, strength, mercies and favor. The strength and guidance of God have enabled me to complete this MSc. program in Econometrics.

I would also like to show my gratitude to my advisor Salie Ayaliew (PhD)for sharing his pearls of wisdom with me during the course of this research, and I am also immensely grateful to my colleagues for their comments.

My heartfelt thanks goes to Woldia University staffs for giving me these priceless chance and statistics department staffs for supporting me through these program. My gratitude also goes to Bahir Dar University statistics staffs for sharing me their knowledge. At last I take this opportunity to express my heartfelt gratitude to all my beloved family members specially Abayneh Moges.

Abstract

Backgrounds:*Food security is an enduring critical challenge in Ethiopia. The 2015 El Niño drought is one of the strongest droughts that have been recorded in Ethiopian history .*

Method: *The study aims to evaluate the impacts of climate change and fertilizers applied on wheat and barley yield per hectare from 1987 to 2017 using an autoregressive distributed lag to cointegration approach.*

Result: *The mean wheat and barley yield was 13.48 and 11.47 quintal per hectare respectively.The bounded F-test for cointegration among the variables show evidence of a long-run relationship with a short run among climate change, fertilizers applied and barley yield per hectare. from the F-statistic of cointegration test there was no evidence that wheat has cointegration with others. Average urea,precipitation and temperature have a positive significant impact but average DAP and rainfall have no significant impact on the amount of wheat yield produced per hectare. On the barley model in the long run, precipitation and rain both had significant positive impacts and average DAP had negative impact on the barley yield per hectare.*

Conclusion:*The results have implications for national and local agriculture policies under climate change and fertilizers used to design well-targeted agriculture adaptation policies for the future and to reduce the adverse effects of climate change on the wheat and barley yield.*

KEYWORDS: Crop yield; Fertilizers; climate variable;autoregressive distributed lag; Cointegration

Acronyms

ADF	Augmented Dickey Fuller
AGRA	Alliance for a Green Revolution in Africa
AIC	Akakie Information Criteria
AR	Auto Regressive
ARDL	Autoregressive Distributed Lag
ARIMA	Auto regressive Integrated Moving Average
CSA	Central Statistical Agency
CUSUM	Cumulative Sum
DAP	Di Ammonium Phosphate
DF	Dickey and Fuller
ECM	Error Correction Model
FAO	Food and Agricultural Organization
GDP	Gross Domestic Product
HQIC	Hannan-Quinn Information Criteria
NASA	National Aeronautics and Space Administration
NMA	National Meteorological Agencie
OLS	Oridnary Least Square
PP	Phillips and Perron
PSS	Pesaran, Shin and Smith
SBIC	Schwarz Bayesian Information Criteria

Contents

List of Tables

List of Figures

CHAPTER ONE

1 INTRODUCTION

1.1 Background of the problem

Agriculture is a proven path to prosperity. No region of the world has developed a diverse, modern economy without first establishing a successful foundation in agriculture. This is going to be critically true for Africa where, today, close to 70 percent of the population is involved in agriculture as small holder farmers working on parcels of land that are, on average, less than 2 hectares. As such, agriculture remains Africa's surest bet for growing inclusive economies and creating decent jobs mainly for the youth [for a green revolution in Africa()].

Global mean temperatures have already risen by 0.8^0C above preindustrial levels. Scientific reviews published in the last few years indicate that recent green house gas emissions and future 21^{st} century emissions trends are higher than previously projected. In the absence of further mitigation there is a 40% probability that global mean temperatures will exceed 4^0C above preindustrial levels and a 10-percentage chance that they will exceed 5^0C [Analytics et al.(2013)].

In the face of global warming, agricultural production systems must become more resilient to long-term changes in temperature and precipitation, as well as to disruptive events. By the year 2100, under different scenarios, climate change is predicted to have an impact on the market (as a percent of GDP) for the entire world but more so for developing countries than for developed ones. Agriculture, as a climate sensitive sector, plays an important role in the economies of developing countries, where the impact is larger and the relationship between crop responses and temperature follows an inverted U-shape relationship [Mendelsohn et al.(2006)Mendelsohn, Dinar, and Williams].

The food crisis trap that threaten African continent is primarily the result of lack of investment in the agricultural sector. Its vulnerability to climate adds to the burden. Since farming in Africa is largely done under rain fed conditions, and Africa's reliance on agriculture and its very low levels of irrigation make it singularly vulnerable to the vagaries of its highly variable and changing climate [for a green revolution in Africa()].

Yield gains associated with high-yielding varieties have been much lower in sub-Saharan Africa than in other regions, partly as a result of the inadequacies of input and output markets and extension services and poor infrastructure. This in turn has resulted in a low use of irrigation, fertilizers, advanced seeds and pesticides [Food and agricultural Organization()].

Crop yield per area (amount of crop harvested per amount of land cultivated) is the most commonly used impact indicator for agricultural productivity activities. Crop yields are inevitably affected by many factors, these are weather, input price, changes in farming practices, amounts of fertilizer used, quality of seed varieties, and use of irrigation [(CSA)()].

Ethiopia enjoyed remarkable growth in agricultural production and overall real incomes (GDP/capita) from 2004/05 to 2008/09, due to a combination of factors, including good weather, increased efforts in agricultural extension, increased usage of fertilizer, and foreign capital inflows that funded major increases in private and public infrastructure investments. In spite of these developments, prices of major cereals (teff, maize, wheat and sorghum) have fluctuated dramatically in both nominal and real terms. International prices of cereals also fluctuated dramatically, particularly between 2006 and 2008. The reality of Ethiopia's agriculture and food security situation is complex because of variations across space within Ethiopia as well as variations over time due to changes in policies, weather shocks, and other factors [Dorosh et al.(2009)Dorosh, Ahmed, et al.].

1.2 Statement of the problem

Food security is an enduring critical challenge in Ethiopia. The 2015 El Niño drought is one of the strongest droughts that have been recorded in Ethiopian history. There were more than 27 million people became food insecure and total population of 18.1 million people require food assistance in 2016. Ethiopia ranked first in having the highest number of people in state of undernourishment which is 32.1 million people in 2014. The number of food insecure people in the country increased from time to time; which was estimated to 2.9 million in 2014 and 4.5 million in August, 2015 and by the end of the same year this figure had more than doubled to 10.2 million food insecure people. [Mohamed(2017)].

In Ethiopia's economy agriculture continues to be the dominant sector, accounting for 51%of the GDP in 2009 [Bank(2013)]. Within agriculture, cereals play a central role accounting for roughly 60% of rural employment, 80% of total cultivated land[Agency)()].

There has been no much research on impact of fertilizers on crop with respect to the Amhara region of Ethiopia. For example,[Matsumoto and Yamano(2011)] found that the fertilizers credit access had a significant impact on teff but not on maize and wheat. The researcher used total amount of fertilizer, as a limitation this doesn't show which fertilizer is more effective on the yield of crops.

There have been limited scientific evidences on impacts of climate change on wheat and barley production in Ethiopia. For instance different scholars addressed the economic impacts of climate change in Ethiopia were based on aggregate agricultural crops produced by farmers[Deressa and Hassan(2009), Ferede et al.(2013)Ferede, Ayenew, Hanjra, and Hanjra]. But in reality; climate change affects different crops differently as long as different crops have different climate requirements and amount of fertilizers required. Therefore, under such findings it is difficult to disaggregate the impact on wheat and barley production. The study analyzed intensification of use of modern inputs like fertilizers and relation between climatic variables and fertilizers with crop yield. In the research paper we looked both climate and fertilizers impact at the same time for each crop type.

1.3 Objectives of the study

1.3.1 General objective

The general objective of the study was assessing the impact of climate variables and type of fertilizers on crop production in Amhara region using time series models.

1.3.2 Specific objectives

The specific objectives of the research were:

- To examine the effect of selected climate variables on selected crop produced.

- To analyze type of fertilizers impact on crop produced.
- To analyze the trend of climatic variables and fertilizers used.

1.4 Significance of the study

A study showed increased temperature and erratic rainfall patterns linked to climate change and variability were the major challenges responsible for the decline in agriculture production [Samuel et al.(2017)Samuel, Shem, Daniel, and Silas]. Agriculture takes the largest share of GDP of Ethiopia. Cereal production and marketing were the means of livelihood for millions of households in Ethiopia, and Amhara region is one of the highest Cereal producer region in the country, so it is mandatory to examine the impacts that climate change has had, and might continue to have on agricultural production.

The fertilizer credit was found to increase input application for crop production [Matsumoto and Yamano(2011)]. The main aim of the study was to provide a meaningful insight and contribute to efforts aimed at ensuring increased food availability through sustainable domestic production and increased income from agricultural production by Identifying and quantifying the effects of climate variability and fertilizers used on crop productivity.

The study also adds Knowledge to the limited but growing literature on impact of climate and fertilizer on agricultural production in Ethiopia. It is also available to give information for designing policies on adaption and mitigation to reduce the effects of fertilizers and climate risk associated with climate change.

1.5 Limitations of the study

In time series and econometrics analysis, the major limitation is availability of data. The accuracy of the data is again another limitation to the study since the inconsistency of data collected on the same variable from different institutions is unbelievable. Since the study used annual data it was difficult to find large number of observation on economical variables. In our country the data recording system is too weak. Shortage of availability of data and inconsistency of data and this makes the study hard.

1.6 Operational definitions

Crop:-A cultivated plant that is grown on a large scale commercially, especially a cereal, fruit, or vegetable.

cereal:-A grain used for food, for example wheat, maize, or rye.

Crop production: - is the process of growing and harvesting of crops for own consumption and/or sale.

Weather: - according to National oceanic and atmospheric administration (noaa) weather is the way in which atmosphere is behaving mainly with respect to its effect up on life and human activities.in most place weather can change minute-to-minute, hour-to hour, day-to-day and month-to-month.

Climate: -is the description of the long-term pattern of weather in a particular area. Climate is also defined as the average weather of a particular region and time period. Usually taken over 30 years.

Climate change: -Climate change is a change of climate which is attributed directly or indirectly to human activity. It alters the composition of the global and/or regional atmosphere and natural climate variability observed over comparable time periods.

Climatic variability:- are the types of changes (temperature, rainfall, occurrence of extremes); magnitude and rate of the climate change that causes the impacts on the area of public health, agriculture, food security, forest hydrology and water resources, coastal area, biodiversity, human settlement, energy, industry, and financial services.

Precipitation:-In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity.it includes rain in the form of ice.

Rainfall:-Rain is liquid water in the form of droplets that have condensed from atmospheric water vapor and then becomes heavy enough to fall under gravity

Fertilizer: – refers to anything that is added to the soil and intended to increase the amount of plant nutrients available for crop growth.

Organic fertilizer:- Contains carbon in its chemical make up. Can be fast or slow release. Can be synthetic or natural in origin.

CHAPTER TWO

2 LITERATURE REVIEW

2.1 Theoretical Review

Agriculture and climate function hand in hand. Today, 32–39% of global crop yield variability is explained by climate; this translates to annual production fluctuations of 2–22 million tones for major crops such as maize, rice, wheat, and soybean. At the same time, agriculture and livestock directly contribute about 11% of global greenhouse gas emissions, and agriculturally-driven land use changes cause additional emissions[Lipper et al.(2014)Lipper, Thornton, Campbell, Baedeker,

By 2050, a growing global population with shifting consumption patterns will require agriculture to deliver 60% more food, yet every 1^0C of warming above historical levels is likely to cause a decrease of approximately 5% in crop productivity. Continuing uneven rural development and inattention to the resource gaps that women and youth are facing will exacerbate inequality. These trends and drivers present a global challenge that requires concerted action [Lipper et al.(2014)Lipper, Thornton, Campbell, Baedeker, Braimoh, Bwalya, Caron, Cat

The success story of the Asian Green Revolution has encouraged governments in African to promote the application of chemical fertilizers, improved seeds and irrigation schemes. However, studies indicated that the chemical fertilizer application of Sub-Saharan countries is lower than the South and East Asian counties. In Sub-Saharan Africa (SSA), the intensity of chemical fertilizer application is 11 kg/ha while in South Asia and East Asia the intensity is 130kg/ha and 271kg/ha respectively [Abrha(2015)].

Agriculture is the major supplier of raw materials to food processing, beverage and textile industries. It accounts for more than 85% of the labor force and 90% of the export earnings [MoFED and economic development()]. Cereal production and marketing are the means of livelihood for millions of households in Ethiopia and is the single largest sub-sector within Ethiopia's agriculture, far exceeding all others in terms of its share in rural employment, agricultural land use, calorie intake, and contribution to national income [Rashid et al.(2019)Rashid, Abate, Lemma, Warner, Kasa, and Minot].

In the country, cereals are also the major stable food crops taking a significant share of area cultivated and volume of production obtained. Out of the total grain crop area, 79.69% (8.7 million hectares) was covered by cereal [Kaso and Guben(2015)].

According to official statistics, Ethiopia's annual rate of economic growth, which averaged 10.3% over 2005/06–2015/16 (compared with the regional average of 5.4%), slowed to 8% in 2016 due to drought-related lower agricultural production. Real GDP growth was estimated to 10.9% in 2017 (July 2016 to June 2017), this was due to a recovery in agricultural production after 2015 drought. The crop harvest was estimated to have increased by 7.9% during the 2017 (compared with a 2.4% increase during 2016) [Bank()].

Ethiopia's crop agriculture is complex, involving substantial variation in crops grown across the country's different regions and agroecology. The Central Statistical Agency (CSA) classified Ethiopian farms into two major groups: smallholder farm (¡25.2ha) and large commercial farms (¿25.2ha). The majority of farmers in Ethiopia are smallholder farms. Smallholders account for 96% of total area cultivated and generate the key share of total production for the main crops. The core crop season is the Meher season, with harvests between September and February. Five major cereals (teff, wheat, maize, sorghum, and barley) are the core of Ethiopia's agriculture and food economy, accounting for about three-quarters of total area cultivated [Alemayehu et al.(2011)Alemayehu, Paul, and Asrat].

In the main agricultural regions in Ethiopia there are two rainy seasons, the Meher and the Belg, and consequently there are two crop seasons. Meher is the main crop season. It encompasses crops harvested between Meskerem (September) and Yeaktit (February). Crops harvested between Megabit (March) and Nehase (August) are considered part of the Belg season crop. The Meher season is overwhelmingly important (96.9% of total crop production and 95.5% of total cereal production in 2007/08). Only smallholders cultivate crops during the Belg season and yields are smaller in the Belg than in the Meher season. In 2007/08 4.5% of national cereal production was produced in the Belg season [Alemayehu et al.(2011)Alemayehu, Paul, and Asrat].

In Ethiopia the total amount of fertilizer applied to area under crops in 2016 G.c was estimated to be more than 12 million quintal and the crop area to which fertilizer was applied estimated about 8.3 million hectares. Of all the quantity of fertilizers used 1.4 million quintal was Urea & DAP, 8 millionquintal was that of NPS & Urea, about 1.9 million quntal was NPS, about 500 thousand quintal was DAP, and about 428 thousand quintals was Urea. Most of the fertilizer used was applied to cereal crops (about 10.4 million quintal) of which teff accounted 3.2 million quintals, Wheat 2.4 million quintals and Maize 3.4 million quintals. The largest area to which fertilizer was applied was that of teff (about 2.4 million hectares) followed by Maize (about 1.7 thousand hectares) and wheat (nearly 1.5 million hectares) crops[(CSA)()].

The world will not be able to meet its food production goals without biotechnology, improved genetics, and without fertilizer. Commercial fertilizer was responsible for 40 to 60% of the world's food production [Roberts et al.(2009)].

The Amhara region had climatic zones ranging from hot dry tropical (800-1830 m above sea level), subtropical (1830-2440 m above sea level), temperate (2440-3000 m above sea level), and alpine (over 3000m above sea level). Highlands above an altitude of 1500m experiences relatively cool temperature conditions in contrast to the lowlands [BoFED and economic development(2011)].

The growth in total output and productivity in the last fifteen years in the Amhara region were not accompanied by significant changes in the marketed surplus of cereal crops. The use of improved seeds and biological and chemical inputs have increased; but not at the rate required to commercialize the agriculture to produce high marketed surplus. For instance, In the case of cereals, the proportion of output marketed has increased marginally from 12.99% to 15.2% [M.(2017)].

2.2 Empirical review

A study done by [Urgessa(2015)] using fixed effect model found, labor per unit of cultivated area of land, the use of pesticides, extension service, number of household member size and the age of the household head found to be the determinants of agricultural land productivity of rural households.

Using multiple regression model done by [Mohammed(2010)] showed that all the cereals mean yields were affected by zone, fertilizer type and crop damage effects. Another studies done by [Bewket(2009)] using correlation analysis showed that Annual rainfall was weakly correlated with production of cereals, and hence it was a poor predictor of yields as well as total outputs.

A study done by [Gedefaw et al.(2018)Gedefaw, Denghua, Hao, Alemu, Chanie, and Agitew] using the binary logistic regression showed that, the reduction of crop productivity since the last two decades was related to drought, flooding, conflict, wind force, ice, insect infestation, inflation, shortage of ploughing land, shortage of grazing land and population growth.

A study done in 2012 by [Tesso et al.(2012)Tesso, Emana, and Ketema] using co-integrated Vector Auto Regressive and Error Correction Models showed food production was significantly affected by improved technology, area under irrigation, manure usage, Meher rain and temperature, while fertilizer application and Belg rain were found to be less significant in the model. The Johannes' approach revealed that 90% of the variation in productivity was explained by area under irrigation, area covered by manure per hectare, the change in usage of improved variety, and the three climate parameters (Meher Rain, Belg rain and Average temperature).

The fertilizer credit was found to increase input application for crop production. As a consequence, it had a substantial impact on the yield of teff. they also found that the impact on net crop income per cultivated area and also on per capita income was marginal because of the low profitability due to the low output price and high input cost of agricultural.[Matsumoto and Yamano(2011)]

A study done in Pakistan using the methods of feasible generalized least square (FGLS) time series data for the period 1989 to 2015 on wheat, rice, maize and sugarcane, reveals maximum temperature adversely affects wheat production, while the effect of minimum temperature was positive and significant for all crops. Rainfall effect towards the yield of a selected crop was negative, except for wheat.[Ali et al.(2017)Ali, Liu, Ishaq, Shah, Ilyas, Din, et al.]

A study done in Nigeria by [Ayinde et al.(2011)Ayinde, Muchie, and Olatunji] using co-integration model, Temperature change was revealed to exert negative effect while rainfall change exerts positive effect on agricultural productivity in Nigeria.

The study done by [Samuel et al.(2017)Samuel, Shem, Daniel, and Silas] showed that an average of 73% in household agricultural production change was due to erratic rainfall, increase sunshine hour, increase in temperature and pest infestation in 2014 and 2015 production years.

From the study done by [You et al.(2009)You, Rosegrant, Wood, and Sun] using OLS and AR(1) model for data from 1979–2000 Chinese crop-specific panel dataset, they found that the climate impact on Chinese wheat yield growth that a 1^0C increase in wheat growing season temperature reduced wheat yields by about 3–10% and also temperature over the past two decades accounts for a 4.5% decline in wheat yields in China.

A study done using multiple regression model on two region in china by [Licker et al.(2013)Licker, Kucharik, Doré, Lindeman, and Makowski] showed winter wheat yields had significant negative responses to increased in minimum summertime temperatures, explaining 11% and 23% of the interannual yield variability on the two region.1^0C increase in minimum summer temperatures contributed to a $0.405 \pm 0.123T/ha$ decrease in winter wheat yields in the first region and $0.193 \pm 0.110T/ha$ decrease in in the second region. A 1mm decreased in precipitation makes the wheat yield increased by 0.009 ± 0.003 T/ha

A study Using different economic models like pooled ordinary least square, fixed effect and random effect model by [Urgessa(2015)] found Fertilizer inputs and the number of household size was found to be the most significant effect on the determinants of land productivity in the rural households of

Ethiopia.

A study done by [Matsumoto and Yamano(2011)] using the fixed effect model found that the fertilizers credit access had a significant positive impact on teff but not on maize and wheat. The estimated coefficient of the credit was 0.37 on teff, suggesting that the teff yield increased by 37 percent if credit was provided.

A study done in 2012 by [Tesso et al.(2012)Tesso, Emana, and Ketema] using co-integrated Vector Auto Regressive and Error Correction Models showed that in Meher rainfall and the risen in average temperature had a productivity reducing effect.

A study done by [Janjua et al.(2014)Janjua, Samad, and Khan] using ARDL model showed that in the short run a 1% increased in fertilizers make the yield to increased by a 0.30% and in the long run it had a positive significance impact but temperature and precipitation had no significance impact on wheat in yield in both long-run and short-run.

A study done by[Amikuzino and Donkoh(2012)], by applying cointegration and Granger causality models, The study suggested that inter-annual yields of the crops had been influenced by the total amounts of rainfall in the planting season a 1% risen in rainfall causes the yield to increased by 1.27%.

There was lack of studies on the determinant impacts of crop production in the Amhara region. Since agriculture had large amount of share on the GDP, study should be done on the factors affected agricultural production.

CHAPTER THREE

3 DATA AND METHODOLOGY

3.1 Source of data

The data for the research were secondary data from the year 1987-2017. Climate variables were taken from national metrological agency and NASA climate data center. Selected cereal yields and amount of fertilizers applied were taken from yearly mehere season of private peasant agriculture and production report of CSA Ethiopia and Amhara national and regional Agricultural office.

3.2 Study area description

The study was conducted in Amhara region. Amhara Region is one of the nine regional state of Ethiopia. The Amhara National Regional State extends from 9^0 to 13^0 45' N and 36^0 to 40^0 30'E. It is bounded by Tigray region and Eritrea in the north, Oromia region in the south, Benishangul-Gumuz region and Sudan in the west, and Afar region in the east.

It covers approximately 161,828.4km^2 in area and is moderately compact in shape. The regional state is made up of 13 administrative zones. This is 21% of Ethiopia's total area.

This land consists of three major geographical zones. Highlands (above 2,300 metres above sea level), semi-highlands (1,500 to 2,300 metres above sea level) and lowlands (below 1,500 metres above sea level) accounting 20%, 44%, and

28% respectively. Its capital city is Bahir Dar. Ethiopia's largest inland body of water, Lake Tana, which is the source of the Blue Nile river, is located within this region.The region also contains the Semien mountain national park,which includes Ras Dashen,the highest point in Ethiopia. In general the region is located in moderate temperate zone[BoFED and economic development(2011)]

3.3 Variables in the study

The variables in the study were Meher season wheat and barley production in quintal/hectare collected from CSA annual report of Ethiopia, selected indicator of climatic variables were average temperature in degree Celsius and precipitation in mm/d collected from the NASA satellite for world climate and rainfall in mm from national meteorology of Ethiopia. Type and amount of fertilizers used (average urea and average dap) in quintal/hectare from CSA annual report.

3.4 Methodology

Time series analysis is a statistical technique that deals with time series data, or trend analysis. Time series data means that data is in a series of particular time periods or intervals. The purpose of time series analysis is generally twofold: to understand or model the stochastic mechanism that gives rise to an observed series and to predict or forecast the future values of a series based on the history of that series and, possibly, other related series or factors.Time series is broadly classified in to two major parts:

1. Univariate time series :- uses only the past history of the time series being forecast plus current and past random error terms. ARIMA

modeling is a specific subset of univariate modeling in which a time series is expressed in terms of past values of itself (the autoregressive component) plus current and lagged values of a 'white noise' error term (the moving average component).

2. Multivariate time series: - consists of multiple single series referred to as components. It involves a vector of time series data that can be modeled simultaneously.

3.4.1 Stationarity

Stationary: -A stationary process has the property that the first and second moments do not change over time. It is an essential property to define a time series process. Stationarity may be weak or strong.

Weak stationarity: - a series is weakly stationary if its first and second moments are time-invariant. In particular, the mean vector and covariance matrix of a weakly stationary series are constant over time[Lütkepohl(2005)]. For a weakly stationary time series y_t, we define its mean vector and covariance matrix as

$$E(y_t) = \mu, \qquad \sigma_h = E(y_t - \mu)(y_{t-h} - \mu) = \sigma(-h) \text{where}, t, h = 0, 1, 2 \cdots \tag{3.1}$$

where the expectation is taken element by element over the joint distribution of y_t.The mean μ is a m-dimensional vector consisting of the unconditional expectations of the components of y_t. The covariance matrix σ_h is a mXm matrix. The diagonal element of σ_h is the variance of y_{it} , whereas the $(i, j)^{th}$ element of σ_h is the covariance between y_{it} and y_{jt}.

Strong Stationary: - a process is strictly stationary if the joint probability distributions of m consecutive variables are time invariant.i.e $(y_1, y_2, ...y_m)$ at any set of time t should have the same joint probability distribution with $(y_{t+1}, y_{t+2}, .., y_{t+m})$. There are several effects of unit root. Among these the followings are some of them

- The process has permanent effects which do not decay as they would if the process were stationary.
- process has a variance that depends on t, and diverges to infinity.
- In multivariate frameworks, one can get spurious regression results.
- statistically, the existence of unit roots can be problematic because OLS estimate of the AR coefficient ϕ is biased.

There are several methods through which we can check the stationary (or unit root) in time series setting. Let's start with simple autoregressive scheme, which is as follows

$$y_t = \alpha + \phi_1 y_{t-1} + trend + \varepsilon_{t-1} \tag{3.2}$$

Where, ϕ_1 is autoregressive (AR) coefficient, where $\varepsilon_{t-1} is N(0, \sigma^2)$. The null and alternative hypothesis are

$$H_0 : \phi_1 = 0 \; vs \; H_1 : |\phi_1| < 1$$

If $\phi_1 = 1$, implies there is unit root and it is non-stationary, meaning the mean and variance are non-constant and violates the normal requirement of time series modeling.

3.4.2 Stationarity test (Unit root test)

Statistics and econometrics use a single-equation or multi-equation regression to model time series economic variables and their interrelations. These models are based on the Box and Jenkins methodology and the fundamental assumption for their use is time series stationarity or their linear combinations stationarity in the case of multi-equation models. In time series models or in econometrics (the application of statistical methods to economics), a unit root is a feature of processes that evolve through time that can cause problems in statistical inference if it is not adequately dealt with.

A non-stationary stochastic process could be trend Stationary (deterministic) Process (TSP) or Difference Stationary Process (DSP). A time series is said to be trend stationary process if the trend is completely predictable and not variable, whereas if it is not predictable, we call it difference or integrated stochastic trend or difference stationary process. In the case of deterministic trend, the divergence from the initial value (represents non-stationary mean) is purely random and they die out quickly. They do not contribute or affect the long run development of the time series.

There are different methods of checking stationarity of a series. Among these test the Augmented Dickey Fuller and Philips-Perron test were used in the research. The difference between these two tests i.e. ADF and PP tests differ mainly in how they treat serial correlation in the test regressions.

3.4.2.1 Augmented Dickey and Fuller test (ADF test)

Prior to run the any time series model the non-stationarity property of the series must be checked. It is important to make sure that all of the variables are stationary, because if they are non-stationary the above listed problem will occur. There are several methods to test stationarity of time series data such as Augmented Dickey-Fuller and Philips-Perron[Dickey and Fuller(1981), Dickey and Fuller(1979) Phillips and Perron(1988)]. In this study only Augmented Dickey-Fuller test and Philips-Perron are considers.

David Dickey and Wayne Fuller in 1979 has proposed the best known and most widely used unit root tests. It is based on the model of the first-order autoregressive process [Box et al.(2015)Box, Jenkins, Reinsel, and Ljung]. The ADF test is used to tests whether a unit root is present in an autoregressive model or not. It is developed by the statisticians David Dickey and Wayne Fuller in 1979. To calculate the test statistic for ADF test, we use an equation that is differenced.

$$\Delta y_t = c + \delta y_{t-1} + \alpha_i \sum_{i=1}^{m} \Delta y_{t-i} + \beta t + \varepsilon_t \tag{3.3}$$

where $\Delta y_t = y_t - y_{t-1}$ and testing the null hypothesis $H_0 : \delta = 0$ is the same as testing $\phi = 1$ because by adding y_{t-1} to both side $(1 + \delta) = \phi$. Under the null hypothesis, we have a non-stationary series, i.e.$\Delta y_t = \varepsilon_t$. The ADF test involves testing for the negativity of δ.

The null hypothesis is $H_0 : \delta = 0$, and the alternative hypothesis is $H_1 : \delta < 0$. A rejection of the null hypothesis then implies $\phi < 1$ and hence y_t is $I(0)$. If we are unable to reject the null, then we conclude that y_t $I(d)$,where $d \geq 1$.The test statistics is given by

$$t = \frac{\hat{\delta}}{s.e(\hat{\delta})} \tag{3.4}$$

Where the numerator is an OLS estimator of δ and the denominator is the standard error of δ. The primary difference between the DF and ADF tests is that the ADF is utilized for a larger and more complicated set of time series models. The augmented Dickey-Fuller statistic used in the ADF test is a negative number, and the more negative it is, the stronger the rejection of the hypothesis that there is a unit root. Of course, this is only at some level of confidence. That is to say that if the ADF test statistic is positive, one can automatically decide not to reject the null hypothesis of unit root.

3.4.2.2 Phillips-Perron (PP) Unit Root Tests

From equation (3.3) the ADF test involves fitting the regression model by ordinary least squares (OLS), serial correlation will present a problem. To account for this, the augmented Dickey–Fuller test's regression includes lags of the first differences of y_t. The Phillips–Perron test involves fitting Δy_t, and the results are used to calculate the test statistics. The error term maybe heteroscedastic. The PP tests correct for any serial correlation and heteroscedasticity in the errors ε_t by non-parametrically modifying the Dickey Fuller test statistics. Phillips and Perron's test statistics can be viewed as Dickey–Fuller statistics that have been made robust to serial correlation by using the heteroscedasticity and autocorrelation consistent covariance matrix estimator[Newey and West(1987)] . The modified Z_t and Z_π are given by:

$$Z_t = \sqrt{\frac{\hat{\delta}^2}{\hat{\lambda}^2}} + \frac{1}{2}\frac{\hat{\delta}^2 - \hat{\lambda}^2}{\hat{\lambda}^2}\frac{T * s.e(\hat{\phi})}{\hat{\lambda}^2} \tag{3.5}$$

$$Z_\pi = T * \hat{\phi} + \frac{1}{2}(T^2 * \frac{s.e(\hat{\phi})}{\hat{\delta}^2})(\hat{\delta}^2 - \hat{\lambda}^2) \tag{3.6}$$

Where $\hat{\delta}^2$ and $\hat{\lambda}^2$ are the variance parameter $\hat{\delta}^2 = \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E(\varepsilon_t^2)$ and $\hat{\lambda}^2 = \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E(s_t^2)$ where $s_t^2 = \sum_{t=1}^{T}\varepsilon_t$ and s_t^2 is the OLS estimate of the error term.Under the null hypothesis $H_0 : \delta = 0$, Z_t and Z_π

have the same asymptotic distributions as the ADF t-statistic and normalized bias statistics. One advantage of the PP tests over the ADF tests is that the PP tests are robust to general forms of heteroscedasticity in the error term and no need of lag determination.

3.4.3 Autoregressive distributed lag (ARDL) Model

Since the data for this research is annual it is better to use $ARDL$ ratherthan ARIMA and GARCH. As non-stationary variables change in time, so OLS estimates show high t values by mistake as they become inflated due to common time component. In econometric it is called spurious results where R^2 value of the model becomes higher than the Durbin-Watson Statistic.To avoid a spurious regression $ARDL$ model was introduced by [Pesaran et al.(2001)Pesaran, Shin, and Smith] in order to incorporate $I(0)$ and $I(1)$ variables in the same estimation. This model is an ordinary least square (OLS) based model which is applicable for both non-stationary time series as well as for times series with mixed order of integration.

If the variables are all stationary $I(0)$ then VAR is suitable and if they are all non-stationary $I(1)$ and there is long run relation then $VECM$ (Johanson Approach) is recommended. Conventional OLS is not appropriate if at least one variable is $I(1)$.

ARDL models are linear time series models in which both the dependent and independent variables are related not only contemporaneously, but across historical (lagged) values as well. In particular, if y_t is the dependent variable and $(x_1, x_2, \cdots x_k)$ are K explanatory variables, a general $ARDL(p, q)$ model is given by:

$$y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^{p} \delta_i y_{t-i} + \sum_{l=1}^{k} \sum_{i=0}^{q} \beta_{ji} x_{j,t-i} + \varepsilon_t \qquad (3.7)$$

Where y_t and x_t are stationary variables,α_0 is a constant term,$\alpha_1, \delta_i and \beta_{ji}$ are respectively the coefficients associated with a linear trend, lags of y_t, and lags of the regressors $x_{j,t}$ for $j = 1, 2, \cdots, k$ and ε_t is a white noise process. Alternatively, let $Ly - y_t - 1$ denote the usual lag operator L and define $\delta(L)$ and $\beta_j(L)$ as the lag polynomials

$$\delta(L) = 1 - \sum_{i=0}^{p} \delta_i L^i \quad and \quad \beta_j(L) = 1 - \sum_{i=0}^{q} \beta_{j,i} L^i$$

then equation (3.7) can be written as

$$\delta(L) y_t = \alpha_0 + \alpha_1 t + \sum_{j=1}^{k} \beta_j(L) x_j, t - i + \varepsilon_t \qquad (3.8)$$

A general $ARDL$ model given in this research as $ARDL(p, q_1, q_2, q_3, q_4, q_5)$ is:

$$\begin{aligned} barley_t = \gamma_{01} + a_1 t + \sum_{i=i}^{p} \alpha_{1j} barley_{t-i} + \sum_{i=0}^{q_1} \alpha_{2j} aveurea_{t-i} + \sum_{i=0}^{q_2} \alpha_{3j} avedap_{t-i} \\ + \sum_{i=0}^{q_3} \alpha_{4j} prec_{t-i} + \sum_{i=0}^{q_4} \alpha_{5j} temp_{t-i} + \sum_{i=0}^{q_4} \alpha_{6j} rain_{t-i} \end{aligned} \qquad (3.9)$$

$$wheat_t = \gamma_{01} + a_{2t} + \sum_{i=i}^{p} \beta_{1j} wheat_{t-i} + \sum_{i=0}^{q_1} \beta_{2j} aveurea_{t-i} + \sum_{i=0}^{q_2} \beta_{3j} avedap_{t-i} + \sum_{i=0}^{q_3} \beta_{4j} prec_{t-i} + \sum_{i=0}^{q_4} \beta_{5j} temp_{t-i} + \sum_{i=0}^{q_4} \beta_{6j} rain_{t-i} \quad (3.10)$$

Where γ^s are intercepts, and α^s and β^s are coefficients(slopes), a_1 and a_2 are coefficients of unrestricted linear trend and ε_{it}^s are a vector of serially uncorrelated white noise process. In the $ARDL$ model the dependent variable is a function its own lagged values, current and lagged values of the exogenous variables.

3.4.4 Lag order selection criteria

A critical element in the specification of $ARDL$ models is the determination of the lag length of the model. The importance of lag length determination is demonstrated by [Braun and Mittnik(1993)] who show that estimates of a VAR whose lag length differs from the true lag length are inconsistent as are the impulse response functions and variance decomposition derived from the estimated VAR.

Over fitting (selecting a higher order lag length than the true lag length) causes an increase in the mean-square forecast errors of the model and that underfitting the lag length often generates autocorrelated errors[Lütkepohl(2005)].

There is no hard-and-fast-rule on the choice of lag length. It is basically an empirical issue. As noted in [Gujarati(2009)], there is no a prior guide as to what the maximum length of the lag should be. Taking in mind that, as one estimates successive lags, there are fewer degrees of freedom left, making statistical inference somewhat unstable. Econometricians are usually not that lucky to have a long series of data so that they can go on estimating numerous lags. More importantly, in economic time series data, successive values (lags) tend to be highly correlated increasing the likelihood of multicollinearity in the model.

One possible approach is to start from a model with some prespecified maximum lag length, p_{max} and apply tests sequentially to determine a suitable model order. For example, the following sequence of null hypotheses may be tested until the test rejects:$H_0 : p_{max} = 0, H_0 : p_{(max-1)} = 0..$, and so on. In this procedure, a decision on p_{max} has to be made. Occasionally this quantity is chosen by some theoretical or institutional argument. For instance, one may want to include lags of at least one year, and thus four lags have to be included for quarterly data and twelve lags may be used for a monthly model.

The general approach is fitting $ARDL(p, q)$ models and choose an estimator of the order p where $p, q = 0, 1, 2, \cdots , p_{max}, q_{max}$ that minimizes the preferred criterion. Most ARDL models are estimated using symmetric lags, i.e. the same lag length is used for all variables in all equations of the model [Klein and Welfe(1983)]. This lag length is frequently selected using an explicit statistical criterion such as the Akakie Information criteria (AIC), Schwarz Bayesian Information Criteria (SBIC), Hannan-Quinn Information criteria (HQIC).

The initial maximal lag in this research has been set equal to two, which is the maximal order recommended by[Pesaran and Shin(1998)] for annual data. Also, from [Wooldridge(2015)]: A Modern Approach with annual data, the number of lags is typically small, 1 or 2 lags in order not to lose degrees of freedom. With quarterly data, 1 to 8 lags are appropriate, and for monthly data, 6, 12 or 24 lags can be used given sufficient data point. Many of the criteria in current use have the general form

$$cr(m) = \log(\det \hat{\Sigma}_{\varepsilon}(p)) + c_T \phi(p) \tag{3.11}$$

where $det(.)$ denotes the determinant, log is the natural logarithm, $\hat{\Sigma}_{\varepsilon} = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_t \varepsilon_t'$ is the residual covariance matrix estimator for a model of order m, C_T is a sequence that depends on the sample size T, and $\phi(p)$ is a function that penalizes large orders. For instance, (p), may represent the number of parameters that have to be estimated in a $ARDL(p, q)$ model. The term $\log(\det(\hat{\Sigma}_{\varepsilon}(p))$ measures the fit of a model with order p. Because there is no correction for degrees of freedom in the covariance matrix estimator, the log determinant decreases (or at least does not increase) when m increases. As in the univariate case, the sample size has to be held constant; hence, the number of presample values set aside for estimation is determined by the maximum order p_{max} [Lütkepohl(2005)].The three most commonly used criteria are

$$AIC(p) = \log(\det(\hat{\Sigma}_{\varepsilon}(p))) + \frac{2}{T} mK^2 \tag{3.12}$$

$$SBIC(p) = \log(\det(\hat{\Sigma}_{\varepsilon}(p))) + \frac{2\ln(\ln(T))}{T} mK^2 \tag{3.13}$$

$$HQIC(p) = \log(\det(\hat{\Sigma}_{\varepsilon}(p))) + \frac{(\ln(T))}{T} mK^2 \tag{3.14}$$

Where K is the number of variables and m is the order.

3.4.5 Parameter Estimation

The first assumptions in $ARDL$ model is the error terms are uncorrelated with both exogenous or endogenous variables. Ordinary least squares, or OLS, in which parameter estimates are chosen to minimize a quantity called the residual sum of squares. So OLS estimation technique is applied because they are efficient and consistent. Let assumed that a m-dimensional multiple time series

$Y = (y_1, y_2, ..., y_T) \qquad (K \times T)$

$B = (c, A_1, A - 2, \cdots A_p) \qquad (K \times (Kp + 1))$

$Z_t = (1, y_t, \cdots y_{t-p+1})' \qquad ((Kp + 1) \times 1)$

$Z = (Z_0, Z_1, \cdots Z_{T-1}) \qquad ((Kp + 1) \times T)$

$U = (u_1, u_2, \cdots u_T) \qquad (K \times T)$

Using this notation, for $t = 1, 2, \cdots, T$, any multivariate regression model can be written compactly as

$$Y = BZ + U \tag{3.15}$$

Thus the OLS estimator of the parameter B is

$$\begin{aligned} \hat{B} &= YZ'(ZZ')^{-1} \\ &= (BZ + U)'Z(Z\prime Z)^{-1} \\ &= B + UZ'(ZZ')^{-1} \end{aligned} \tag{3.16}$$

3.4.6 Model adequacy checking

Model adequacy checking, also known as diagnostic check or residual analysis, plays an important role in model building. Before proceeding to further steps, it is mandatory to diagnose the model that is build. The most useful and informative diagnostic checks deal with determining whether or not the assumptions underlying the innovation series are satisfied by the residuals of the calibrated *ARDL* model.

When fitting a model to time series variables the model should be stable and the estimated innovations or residuals are assumed to be independent, homoscedastic (i.e. have a constant variance) and normally distributed. Estimates for the error terms are automatically calculated at the estimation stage for the model parameters.

Of the three innovation assumptions, independence and whiteness, is by far the most important. A data transformation cannot correct dependence of the residuals because the lack of independence indicates the present model is inadequate. Rather, the identification and estimation stages must be repeated in order to determine a suitable model. If the less important assumptions of normality violated it can often be corrected by a Box-Cox transformation of the data.

3.4.6.1 Stability test

It is important to check the stability of parameters in the model. In the *ARDL* model this is checked by CUSUM and CUSUM SQUARE. The CUSUM and CUSUMSQ tests for parameter stability were first introduced into the statistics and econometrics literatures by [Brown et al.(1975)Brown, Durbin, and Evans]. The main difference between these two tests are depends on the nature of the structural change taking place. If the break is in the intercept of the regression equation then the CUSUM test has higher power. However, if the structural change involves a slope coefficient or the variance of the error term, then the CUSUMSQ test has higher power.

3.4.6.2 Specification test

The Ramsey Regression Equation Specification Error Test (RESET) test is a general specification test for the linear regression model.This test was developed by James B. Ramsey. More specifically, it tests whether non-linear combinations of the fitted values help explain the response variable. The intuition behind the test is that if non-linear combinations of the explanatory variables have any power in explaining the response variable, the model is misspecified in the sense that the data generating process might be better approximated by a polynomial or another non-linear functional form[Ramsey(1969)].Suppose that in the multiple linear regression model

$$y = c + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \varepsilon \tag{3.17}$$

This implies that no nonlinear functions of the independent variables (such as squares and cubes of x_j^s should be significant when added to the model. But, in the White heteroscedasticity test, adding squares, cubes and cross-products uses up many degrees of freedom.Instead of this, we can add squares and cubes of the fitted values, $\hat{y}^2, \hat{y}^3$, into the model and test for the joint significance of added terms using F-statistic or LM test. The auxiliary regression for the RESET test statistic can be written as follows:

$$y = c + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \delta_1 \hat{y}^2 + \delta_2 \hat{y}^3 + \varepsilon \tag{3.18}$$

The null hypothesis of the RESET test says that the model is correctly specified:

$$H_0 : \delta_1 = \delta_2 = 0 \; vs \; H_1 : \delta_1 \neq \delta_2 \neq 0$$

In large samples and under the Gauss-Markov assumptions, the usual F restrictions test follows the $F(2; n-k-3)$ distribution. the rule of thumb is reject H_0 If the F statistic is greater than the critical value at 5% level of significance. This indicates that there is a functional form of misspecification.RESET is a general test for the following types of specification errors:

- Omitted variables; does not include all relevant variables.
- Incorrect functional form; some or all of the variables in and should be transformed to logs, powers, reciprocals, or in some other way.
- Correlation between and , which may be caused, among other things, by measurement error in , simultaneity, or the presence of lagged values and serially correlated disturbances.

3.4.6.3 Residual autocorrelation test

The fitted model is assumed to be stationary, invertible, and identifiable and fitting this model to a series of length m, the residuals, $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t}......, \varepsilon_{m,t})'$, where $t = 1, 2, \cdots$ may be estimated and used to check the model assumption that the innovations are white noise. There are different procedures to test whether the residuals are autocorrelated or not these are

- Breusch–Pagan test (Autocorrelation Lagrange multiplier test)
- Durbin–Watson test
- Ljung–Box test (Multivariate portmanteau tests)

among these tests Lagrange multiplier test of Breusch-Pagan is used .

3.4.6.3.1 Autocorrelation Lagrange multiplier test

This test was developed by Trevor S. Breusch and Leslie G. Godfrey in 1978 to overcome the drawback of the Durbin Watson test. Consider the equation (3.1) with no Trend given above, this is most easily implemented by partialling out lagged residuals from the original regressors $\varepsilon_t = \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + ...\rho_p\varepsilon_{t-p} + u_t$ where u_t denotes a vector of white noise error terms and re estimating the original system using the new regressors. Again, we set missing observations to zero. Therefore, the null and the alternative hypotheses are:

$$H_0 : \rho_1 = \rho_2 = ... = \rho_p \ = 0 \ vs \ H_1 : \text{atleast one } \rho \neq 0$$

The Breusch-Godfrey LM test combines these two equations:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + ...\rho_p\varepsilon_{t-p} + u_t \quad (3.19)$$

Where $\hat{u}_(t)$ denotes the residuals from the auxiliary regression. The auto correlation LM test statistic at lag p is

$$LM_s = (T - d - 0.5)\ln(\frac{\hat{\Sigma}}{\tilde{\Sigma}}) \sim X^2_{m^2} \quad (3.20)$$

Where T is the number of observations in the $ARDL$, d is the number of coefficients estimated in the augmented $ARDL$,$\hat{\Sigma}$ is the MLE of Σ and $\tilde{\Sigma}$ is the MLE of the augmented model.The Lagrange multiplier test statistic is equivalent to a test based on multivariate portmanteau [Hosking(1980), Poskitt et al.(1982)Poskitt, Tremayne, et al., Ljung and Box(1978)]

3.4.6.4 Heteroskedasticity test

The Breusch-Pagan-Godfrey test is a Lagrange multiplier tests whether the variance of the errors from a regression is dependent on the values of the independent variables. In that case, heteroskedasticity is present. The test is performed by completing an auxiliary regression of the squared residuals from the original equation on $(1, x_t)$.Assume our regression model is

$$Y_i = \beta_1 + \beta_2 X_{2i} + \mu_i \tag{3.21}$$

i.e we have simple linear regression model, and $E(\mu_i^2) = \sigma_i^2$, where $\sigma_i^2 = f(\alpha_1 + \alpha_2 Z_{2i})$, σ_i^2 is some function of the non-stochastic variable Z's. The f() allows for both the linear and non-linear forms of the model. The variable Z is the independent variable X or it could represent a group of independent variables other than X.Step to Perform Breusch Pagan tests are:

1. Estimate the model by OLS and obtain the residuals $\hat{\mu}_1, \hat{\mu}_2 + \cdots$
2. Estimate the variance of the residuals i.e. $\hat{\sigma}^2 = \frac{\sum e_i^2}{(n-2)}$
3. Run the regression $\frac{e_i^2}{\hat{\sigma}^2} = \beta_1 + \beta_2 Z_i + \mu_i$ and compute explained sum of squares (ESS) from this regression
4. Test the statistical significance of ESS/2 by χ^2-test with 1 df at appropriate level of significance (α).
5. Reject the hypothesis of homoscedasticity in favour of heteroscedasticity if $\frac{ESS}{2} > \chi^2_{(1)}$ at the appropriate level of α

3.4.6.5 ARCH test

The ARCH test is a Lagrange multiplier (LM) test for autoregressive conditional heteroskedasticity (ARCH) in the residuals.The ARCH LM test statistic is computed from an auxiliary test regression. To test the null hypothesis that there is no ARCH up to order q in the residuals, we run the regression:

$$\varepsilon_t^2 = \beta_0 + \sum_{i=1}^{q} \beta_i \varepsilon_{t-i}^2 + v_t \tag{3.22}$$

where ε is the residual. This is a regression of the squared residuals on a constant and lagged squared residuals up to order q. The null and alternative hypothesis is given by:

$$H_0 : \text{No ARCH up to order q } vs \ H_1 : NotH_0$$

3.4.6.6 Normality test of the residuals

The tests are based on the third and fourth central moments (skewness and kurtosis) of the normal distribution. If x is a univariate random variable with standard normal distribution, i.e., $x \sim N(0,1)$, its third and fourth moments are known to be $E(x^3) = 0$ and $E(x^4) = 3$. Let ε_t be a m-dimensional Gaussian white noise process with $\varepsilon_t \sim N(\mu_\varepsilon, \Sigma_\varepsilon)$ and let P be a matrix satisfying $PP' = \Sigma_\varepsilon$ For example, P may be obtained by a Choleski decomposition of Σ_ε [Lütkepohl(2005)]. Then

$$w_t = (w_{1t}, w_{2t}, \cdots, w_{mt}) = P^{-1}(\varepsilon_t - \mu_\varepsilon) \sim N(0, I_m) \tag{3.23}$$

In other words, the components of w_t are independent standard normal random variables. Hence, the null hypothesis is

$$H_0 : E \begin{Bmatrix} w_{1t}^3 \\ \cdot \\ \cdot \\ w_{mt}^3 \end{Bmatrix} = 0, \quad E \begin{Bmatrix} w_{1t}^4 \\ \cdot \\ \cdot \\ w_{mt}^4 \end{Bmatrix} = \begin{Bmatrix} 3 \\ \cdot \\ \cdot \\ 3 \end{Bmatrix} \quad vs \quad H_1 : not\ H_0$$

The most commonly known for testing this hypotheses is Jarque-Bera test which is proposed in 1987. Let $\bar{\varepsilon}$ be mean of $(\varepsilon_1, \varepsilon_2, \cdots \varepsilon_T)$ and $S_\varepsilon = \frac{1}{T-1}\sum_{t=1}^{T}(\varepsilon_t - \bar{\varepsilon})(\varepsilon_t - \bar{\varepsilon})'$ and P_s is a matrix for which $P_s P_s' = S_\varepsilon$ and such that $p_{lim}(P_s - P) = 0$. Moreover,

$$\begin{aligned} v_t &:= (v_{1t}, v_{2t}, \cdots, v_{mt})' = P_s^{-1}(\varepsilon_t - \bar{\varepsilon}), t = 1, 2, \cdots, T, \\ b_1 &:= (b_{11}, b_{21}, \cdots, b_{m1})' \ with \ b_{m1} = \frac{1}{T}\sum_{t=1}^{T} v_t^3, m = 1, 2, \cdots, m \end{aligned} \tag{3.24}$$

and

$$b_2 := (b_{12}, b_{22}, \cdots, b_{m2})' \ with \ b_{m2} = \frac{1}{T}\sum_{t=1}^{T} v_t^4, m = 1, 2, \cdots, m \tag{3.25}$$

Thus, b1 and b2 are estimators of the vectors in equation the hypothesis. The asymptotic distribution of skewness and kurtosis is

$$\lambda_{sk} = \frac{T b_1' b_1}{m} \quad \xrightarrow{d} \chi^2(m) \tag{3.26}$$

and

$$\lambda_{ku} = \frac{T(b_2 - 3m)'(b_2 - 3m)}{24} \quad \xrightarrow{d} \chi^2(m) \tag{3.27}$$

3.4.7 Error correction model (ECM) and Cointegration

many econometric methods have been proposed for investigation of the long-run equilibrium cointegration approach with time series variables. Among these Engle and Granger (1987), Johnson's Co-integration technique following Johansen (1988) and Johansen and Juselius (1990), but there outcome is not reliable for small sample[Narayan(2005), Udoh and Ogbuagu(2012)]. The concept of cointegration was defined by Engle, Robert F and Granger, Clive WJ in (1987) . A multivariate process y_{it} are $I(1)$ processes, but a nontrivial linear combination $\phi\prime y_{it}$ is an $I(0)$ series, then y_{it} is said to be Cointegrated of order 1. In general, if y_{it} are $I(d)$ non stationary and $\phi\prime y_{it}$ is $I(h)$ with $h < d$, then y_{it} is Cointegrated. Cointegration often means that a linear combination of individually unit-root non stationary time series becomes a stationary and invertible series. The linear combination ϕ vector is called a cointegrating vector. A process consisting of cointegrated variables is called a cointegrated process [Engle and Granger(1987)]. The EC take into account any cointegrating relationships among the variables. The presence of a cointegrating equation is interpreted as a long-run equilibrium relationship among the variables.

3.4.7.1 Cointegration test

Many time series variables are stationary only after differencing. Hence, using differenced variables for regressions imply loss of relevant long run properties or information of the equilibrium relationship between the variables under consideration[Nkoro et al.(2016)Nkoro, Uko, et al.]. We need to find the way to retain the long run relationship, cointegration makes it possible to retrieve the relevant long run information of the relationship between the considered variables that had been lost on differencing. Cointegration is concerned with the analysis of long run relations between integrated variables and reparametrizing the relationship between the considered variables into an Error Correction Model.

There are different methods of testing cointegration among variables. These are

- the Engle and Granger
- Phillips-Ouliaris methods
- Johansen's procedure
- ARDL Bound test

3.4.7.1.1 ARDL Bounds tests for cointegration

In order to empirically analyze the long-run relationships and short run dynamic interactions among the variables of interest (wheat, barley, average dap, average urea, precipitation,temperature and rainfall) we apply the ($ARDL$) Bound test technique. After checking and estimating the long run relationship of the variables, then one can estimate the appropriate short run parameters by using Error Correction model (ECM).

The $ARDL$ cointegration approach developed to investigate the long run relationship. It has many advantages in comparison with other previous and traditional cointegration methods. The first one is that the $ARDL$ bound test does not need that all the variables under study must be integrated of the same order and it can be applied when the under-lying variables are integrated of order one, order zero or mixed integrated. The second advantage is that the $ARDL$ test is relatively more efficient in the case of small and finite sample data sizes. A dummy variable can be included in the co-integration test process, which is not permitted in Johansen's method. ARDL procedure employs only a single reduced form equation, while the other co-integration procedures estimate the long-run relationships within a context of system equations. The last advantage is that by applying the $ARDL$ bound test technique we obtain unbiased estimates of the long-run model [Udoh and Ogbuagu(2012), Rahimi and Shahabadi(2011), Harris and Sollis(2003)].

In the traditional cointegration methodologies of Engel and Granger (1987), phillips and hansen(1990) and Johanson (1988) typically fail since all variables need to have identical orders of integration, usually $I(1)$. This requires pre-testing for the presence of a unit root in each of the variables under consideration, which is clearly subject to misclassification, particularly since unit root tests are known to suffer size and power problems in many cases of interest [Perron and Ng(1996)].

The shortcomings follow from its reliance on the presence of a single cointegrating vector. Secondly, the ARDL estimator may not provide robust results in the presence of I(2) variables. Finally, the value of the F-statistic may be sensitive to the number of lags imposed on the differenced variables [Bahmani-Oskooee and Goswami(2003)].

For the purpose of this study Autoregressive Distributed Lag ($ARDL$) bound test approach developed by Pesaran, Shin and Smith (2001) was used .The general form of the model in the research was:

$$\boldsymbol{\Delta barley_t} = \gamma_{01}+\sum_{i=i}^{p}\alpha_{1j}\Delta barley_{t-i}+\sum_{i=0}^{q_1}\alpha_{2j}\Delta aveurea_{t-i}+\sum_{i=0}^{q_2}\alpha_{3j}\Delta avedap_{t-i}+$$
$$\sum_{i=0}^{q_3}\alpha_{4j}\Delta prec_{t-i}+\sum_{i=0}^{q_4}\alpha_{5j}\Delta temp_{t-i}+\sum_{i=0}^{q_4}\alpha_{6j}\Delta rain_{t-i}+\delta_1 barley_{t-i}$$
$$\delta_2 aveurea_{t-1}+\delta_3 avedap_{t-1}+\delta_4 prec_{t-1}+\delta_5 temp_{t-1}+\delta_6 rain_{t-1} \quad (3.28)$$

$$\boldsymbol{\Delta wheat_t} = \gamma_{01}+\sum_{i=i}^{p}\beta_{1j}\Delta wheat_{t-i}+\sum_{i=0}^{q_1}\Delta\beta_{2j} aveurea_{t-i}+\sum_{i=0}^{q_2}\Delta\beta_{3j} avedap_{t-i}+$$
$$\sum_{i=0}^{q_3}\Delta\beta_{4j} prec_{t-i}+\sum_{i=0}^{q_4}\Delta\beta_{5j} temp_{t-i}+\sum_{i=0}^{q_4}\Delta\beta_{6j} rain_{t-i}+\pi_1 wheat_{t-i}+$$
$$\pi_2 aveurea_{t-1}+\pi_3 avedap_{t-1}+\pi_4 prec_{t-1}+\pi_5 temp_{t-1}+\pi_6 rain_{t-1} \quad (3.29)$$

Where δ^s and π^s are long-run coefficients, Δ is first differences,α^s and β^s are short run coefficients, t is time trend and $\varepsilon_i^s \sim (\mu, \sigma^2 I)$. $ARDL$ procedure is statistically much more significant approach to determine the cointegration relationship in small samples, which allows different optimal lags of variables. To investigate the long run relationship $ARDL$ is applied to compute the long run association of the variables.

The bound test approach for the long-run relationship between the variables of interest and the exogenous variables are based on the Wald test (F-statistic), by imposing restrictions on the long-run estimated coefficients of one period lagged level of the variables of interest and the exogenous variables to be equal to zero, that is,

$$H_0 : \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = \delta_6 = 0 vs H_1 : \delta_1 \neq \delta_2 \neq \delta_3 \neq \delta_4 \neq \delta_5 \neq \delta_6 \neq 0$$

and the same for the second model then the calculated F-statistic is compared to the tabulated critical value in Pesaran ,Shin and Smith (2001). The explanatory variables are assumed to be integrated of order zero, or I(0) for values of the lower bound, while the upper bound values assumed they are integrated of order one, or I(1). Therefore, the decision rule is that if computed F-statistic falls below the lower bound value, I(0), the null hypothesis (no cointegration) cannot be rejected. Contrarily, if the computed F-statistic exceeds the upper bound value, I(1) then it can be concluded that variables of interest and the exogenous variables co-integrated [Pesaran et al.(2001)Pesaran, Shin, and Smith].

3.4.7.2 Error correction of the ARDL model

A dynamic error correction model (ECM) can be derived from $ARDL$ through a simple linear transformation. Likewise, the ECM integrates the short-run dynamics with the long-run equilibrium without losing long-run information and avoids problems such as spurious relationship resulting from non-stationary time series data.

The EC coefficient shows how fast variables restore to their equilibrium value. EC term is one period lagged residual saved from the estimated dynamic long run relationship. The ECM_{t-1}, which measures the adjustment to restore equilibrium in the dynamic model.If the ECM coefficient is significance ensuring the long run equilibrium can be attained [Banerjee et al.(1998)Banerjee, Dolado, and Mestre]. Let's re-parameterize equation (3.7) with only one explanatory cariable and no trend term to the ECM. Equation (3.7) can be written as

$$y_t = (1 + \alpha_1 + {\alpha_1}^2 + ...)c + (1 + \beta_1 L + \beta_2 L^2 + ...)(\beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t) \quad (3.30)$$

the current value of y_t depends on the current and all previous values of x_t and ε_t and

$$\frac{\delta y_t}{\delta x_t} = \beta_0 \quad \text{[an impact multiplier]}$$

then the effect of the first lag (after one period) is,

$$\frac{\delta y_{t+1}}{\delta x_t} = \beta_0 \alpha_1 + \beta_1 \tag{3.31}$$

Effect at second lag is,

$$\frac{\delta y_{t+2}}{\delta x_t} = \alpha_1 \beta_1 + \alpha_1^2 \beta_0 \tag{3.32}$$

Since y_t is equivalent to $y_{t-1} + \Delta y_t$ *and* x_t is equivalent to $x_{t-1} + \Delta x_t$, by substituting y_t and x_t with $y_{t-1} + \Delta y_t$ and $x_{t-1} + \Delta x_t$ in equation (3.9) we have,

$$\Delta y_t = c + \beta_0 \Delta x_t - (1 - \alpha_1) y_{t-1} + (\beta_0 + \beta_1) x_{t-1} + \varepsilon_t \tag{3.33}$$

$$\Delta y_t = \beta_0 \Delta x_t - (1 - \alpha_1)[y_{t-1} - (c/(1 - \alpha_1)) - ((\beta_0 + \beta_1)/1 - \alpha_1) x_{t-1} \tag{3.34}$$

If we let:$\phi_0 = (1 - \alpha_1)$ and $\phi_1 = (\beta_0 + \beta_1)$ then equation (3.33) become

$$\Delta y_t = c + \beta_0 \Delta x_t - \phi_0 [y_{t-1} - \phi_1 x_{t-1}] + \varepsilon_t \tag{3.35}$$

And the total long term effect/long run multiplier (equilibrium), say k_1 is therefore $k_1 = [\frac{\phi_1}{-\phi_0}]$. Y and X will be in their long term equilibrium state when $y = k_0 + k_1 x$, where $k_0 = [\frac{c}{-\phi_0}]$. In summary the ECM concludes that the current change in y is the sum of two components.

- current change in y is proportional to the current change in x
- current change in y is a partial correction for the extent to which the

lag of y (i.e. y_{t-1} deviates from the equilibrium values corresponding to x_{t-1} the equilibrium error).

Hence, by differencing and forming a linear combination of the non-stationary data, all variables in $ARDL$ model are transformed equivalently into an ECM with stationary series only. ECM in the research was given by:

$$\boldsymbol{\Delta barley_t} = \gamma_{01} + \sum_{i=i}^{p} \alpha_{1j} \Delta barley_{t-i} + \sum_{i=0}^{q_1} \alpha_{2j} \Delta aveurea_{t-i} + \sum_{i=0}^{q_2} \alpha_{3j} \Delta avedap_{t-i} + \sum_{i=0}^{q_3} \alpha_{4j} \Delta prec_{t-i} + \sum_{i=0}^{q_4} \alpha_{5j} \Delta temp_{t-i} + \sum_{i=0}^{q_4} \alpha_{6j} \Delta rain_{t-i} + \phi_1 ECTt - 1 + \varepsilon_{1t} \tag{3.36}$$

$$\boldsymbol{\Delta wheat_t} = \gamma_{01} + \sum_{i=i}^{p} \alpha_{1j} \Delta wheat_{t-i} + \sum_{i=0}^{q_1} \alpha_{2j} \Delta aveurea_{t-i} + \sum_{i=0}^{q_2} \alpha_{3j} \Delta avedap_{t-i} + \sum_{i=0}^{q_3} \alpha_{4j} \Delta prec_{t-i} + \sum_{i=0}^{q_4} \alpha_{5j} \Delta temp_{t-i} + \sum_{i=0}^{q_4} \alpha_{6j} \Delta rain_{t-i} + \phi_2 ECTt - 1 + \varepsilon_{2t} \tag{3.37}$$

3.4.8 Causality Analysis

After confirming the long-run relationship by applying the ARDL bounds test and combined cointegration techniques, the Granger causality can be applied to investigate the direction of causality among the variables. The Error Correction Model (ECM) based Granger causality test is applied to investigate the direction of causality between the variables in equation (3.36) and (3.37).

The ECM is an important model that distinguishes the short- and long-run Granger causalities. The lag of the individual coefficients is utilized to test the significance of the short-run relationship. Furthermore, the coefficient of

ECT_{t-1} is statistically significant an indicates long-run causality. Jointly-lagged coefficients and the EC are used to verify joint causality between the variables.

CHAPTER FOUR

4 Result and discussion

4.1 Descriptive Analysis

Table (4.1) shows descriptive statistic of annual wheat and barley production by private peasant in the Amhara region, fertilizers applied and climate variability from 1987 to 2017. The mean annual yields were 13.48 q/ha and 10.05 q/ha for wheat and barley respectively. The amount of urea and DAP fertilizers used for the two crops are 0.65 q/ha and 0.81 q/ha respectively. The minimum (maximum) amount produced from 1987-2017 for the two cereals crops in the region were, 7.11(19.74) q/ha and 6.36(25.33) q/ha for barley and wheat respectively. The standard deviations of the variables indicates no high variability from year to year in the yields of cereal production, amount of fertilizers used and climate for the region in the last 30 years except wheat and rain . Based on Jarque-Bera test statistic, skewness and kurtosis all the variables are normal except rain.

Table 4.1: Descriptive Statistics

	WHEAT	BARLEY	TEMPS	RAIN	PREC	AVEUREA	AVEDAP
Mean	13.48097	11.47129	21.47968	111.6944	3.507264	0.650538	0.812043
Median	11.90000	10.10000	21.40000	113.6822	3.640909	0.650000	0.820000
Maximum	25.33000	19.74000	25.14000	142.1351	4.851818	0.893333	1.133333
Minimum	6.360000	7.110000	18.57000	60.67500	1.963636	0.303333	0.496667
Std. Dev.	5.327074	3.666677	1.372529	16.81541	0.836613	0.134651	0.148652
Normality test							
Skewness	0.7044	0.7579	0.3755	-1.1878	-0.3572	-0.3650	-0.1453
Kurtosis	2.4928	2.4136	3.6330	5.0851	1.9315	2.9380	2.6995
Jarque-Bera	2.896035	3.411030	1.246181	12.90551	2.133996	0.693352	0.225717
Probability	0.235036	0.181679	0.536285	0.001576	0.344040	0.707034	0.893277
Observations	31	31	31	31	31	31	31

The mean annual values for the three selected climatic variables were 21.48^0C, $111.69mm$ and $3.51mm/d$ for temperature, rainfall and precipitation respectively. From Figure (4.1 and 4.2), during the growth stage of wheat and barley, maximum average temperature was 25.14^0C and 18.57^0C in the year 2005 and 1989 respectively. Average temperature doesn't show high variation from 1987 to 2017 ($y = 0.0997x - 177.43$). By comparison, the total precipitation fluctuated and showed a slight downward trend ($y = -0.0188x + 41.053$). The maximum value for precipitation is $4.85mm/d$ in 1987, whereas the minimum value is $1.96mm/d$ in 2005.

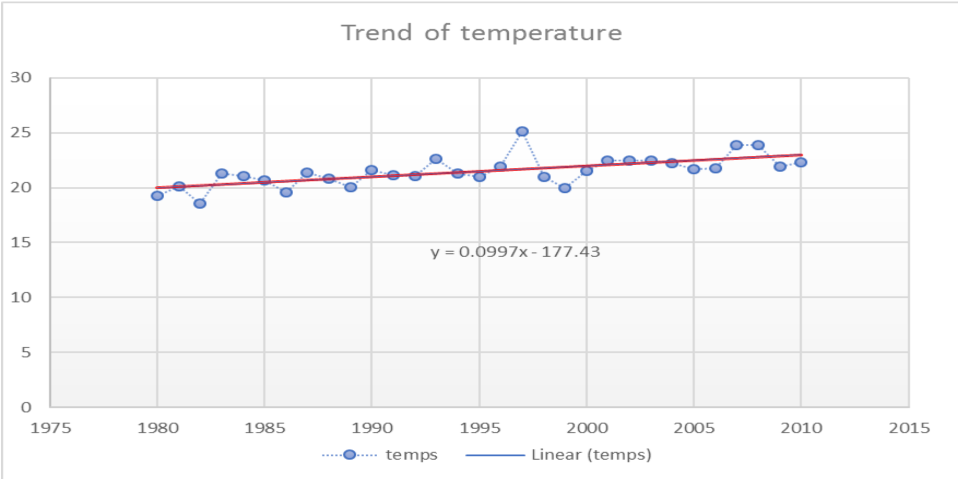

Figure 4.1: Trend of Temperature

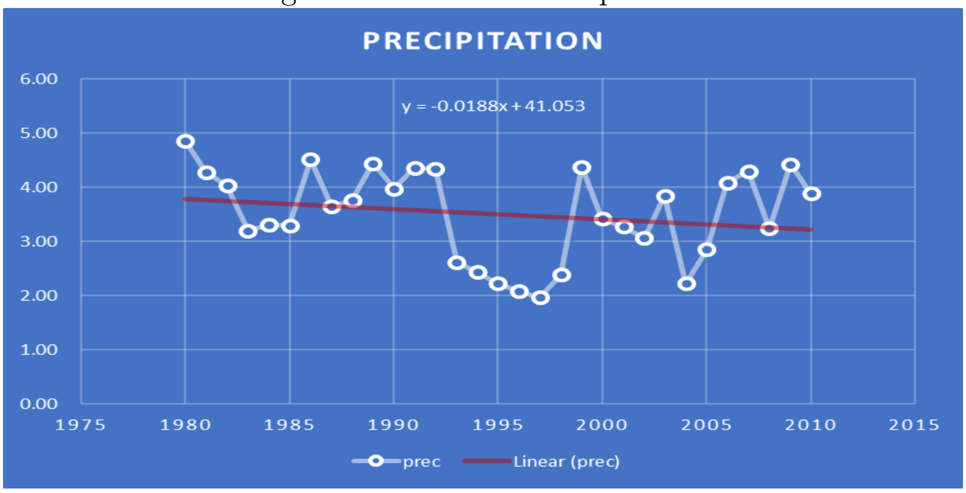

Figure 4.2: Trend of Precipitation

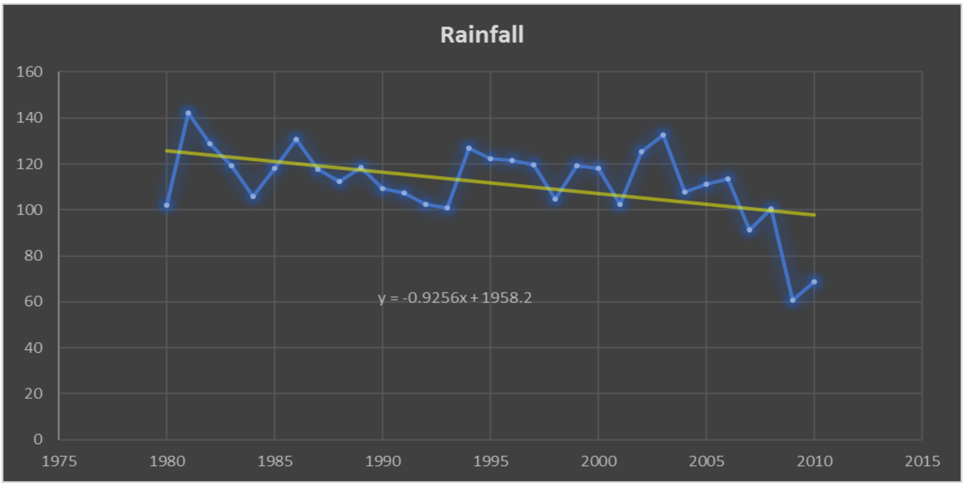

Figure 4.3: Trend of Rainfall

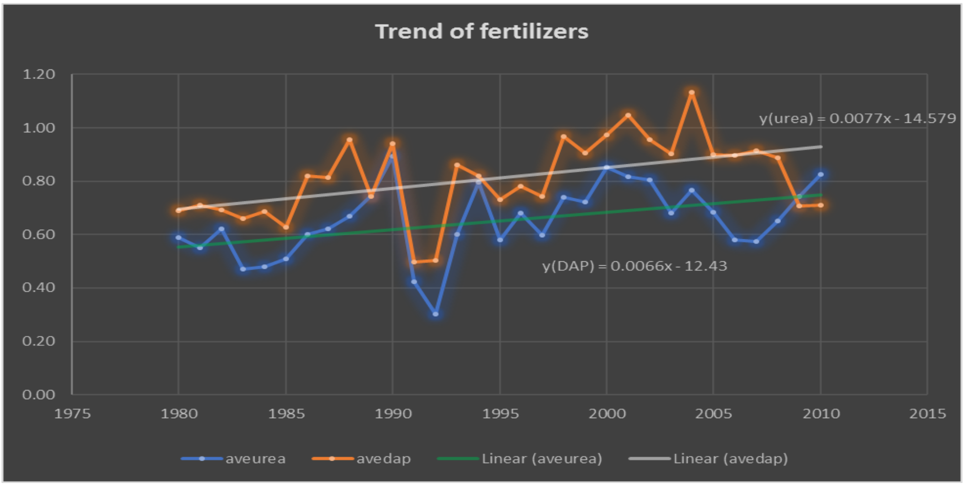

Figure 4.4: Trends of fertilizers applied

From Figure(4.3 and 4.4), the minimum amount of rainfall recorded in 2016 was 60.675mm and maximum amount was 142.14mm in 1988. There is a substantial variation by years and it has a decreasing trend with time ($y = -0.9256x + 1958.2$). The maximum(minimum) average urea and DAP was 0.89q/ha in 1997, 0.3q/ha in 1999 and 1.13q/ha in 2011 and 0.49q/ha in 1998-99 respectively. The average amount of urea and DAP fertilizers doesn't differ substantially by years but there is an increasing trend on the use of both fertilizers ($y_{urea} = 0.0077x - 14.579$) and ($y_{DAP} = 0.0066x - 12.43$).

4.2 Inferential Statistics

4.2.1 Unit root test

The $ARDL$ bounds test is based on the assumption that the variables are $I(0)$, $I(1)$ or mixed order but not $I(2)$. So, before applying the test, we determine the order of integration of all variables using the unit root tests. The objective is to ensure the variables are not $I(2)$ so as to avoid spurious results. In the presence of variables integrated of order two, we cannot interpret the values of F-statistics [Pesaran et al.(2001)Pesaran, Shin, and Smith].

There are two ways of testing stationarity these are graphical and common test. In the research graphical methods included time series plot and common tests include Augmented Dickey-Fuller and Philips-Perron test statistic. First it is recommended to plot the series in order to look the pattern of the data weather to include the trend part or not.

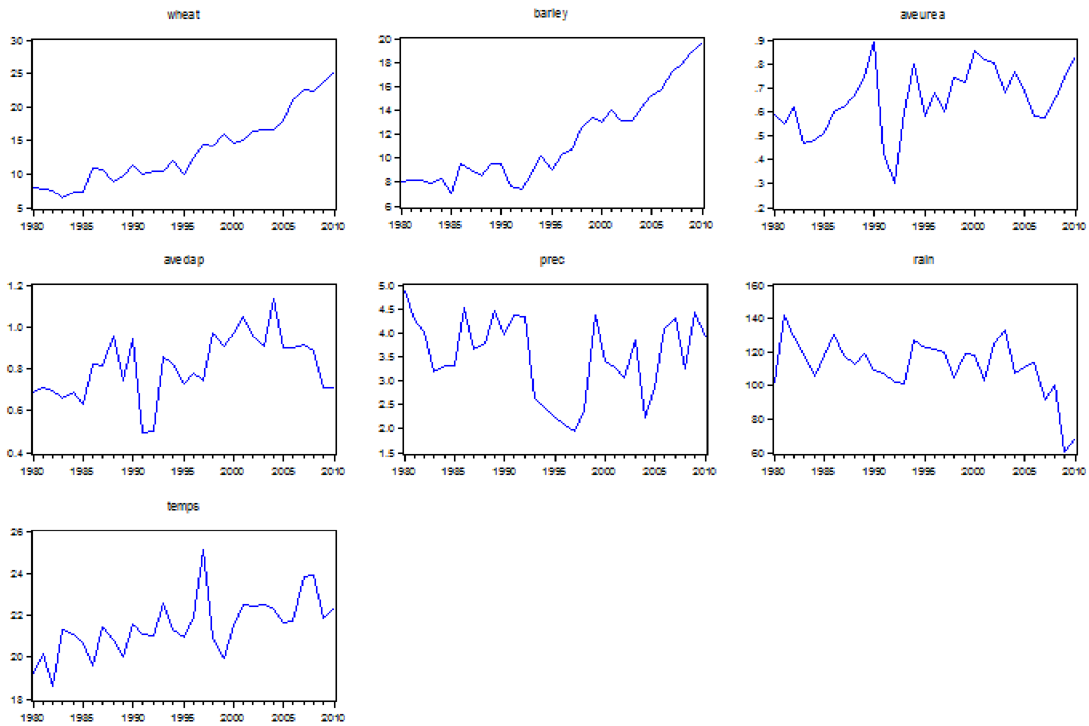

Figure 4.5: Time series plot of variables at level

From figure (4.5) the mean and variances of some variables seems increase with time, implies the series are not stationary. Taking the first difference may result in stationary processes. Since the decision is subjective, the stationary process also cannot be determined from figure (4.6). Therefore, we undertake a formal test for unit root test.

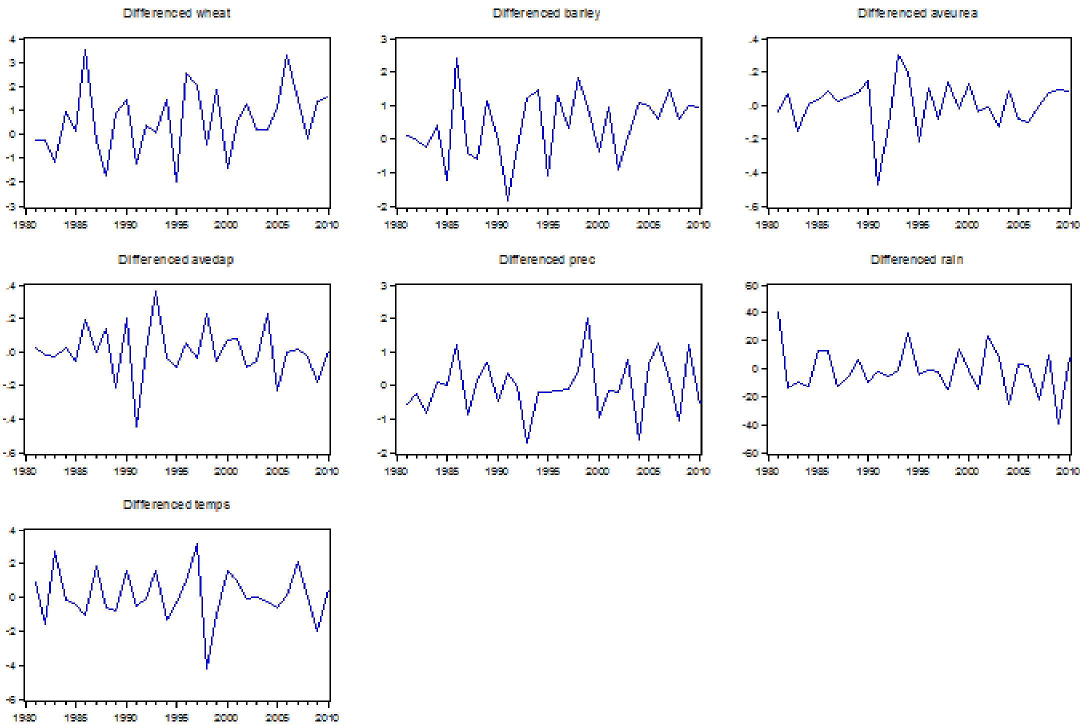

Figure 4.6: Time series plot of variables at first difference

The ADF and PP tests the stationarity of the series. The choice (with or without trend and intercept) to choose depends on economic theory and trending behavior of the data. From the first graph the data had an increasing trend with time. So, the choice for testing unit roots include trend.

$$H_0 : X_t \sim I(1) \ vs \ H_1 X_t \sim I(0)$$

$$H_0 : X_t \sim I(2) vs H_1 X_t \sim I(1)$$

Table 4.2: Unit root test

Variables	ADF level		ADF diff		PP level		PP diff		Order
	Intercpt	Int& Trend	Intercpt	Int & trend	intercept	Int & trend	intercpt	Int & trend	
Barley	0.87 (0.994)	-1.59 (0.772)	-6.06 (0.001)	-6.69 (0.001)	1.72 (0.999)	-1.59 (0.772)	-6.06 (0.001)	-8.15 (0.001)	I(1)
Wheat	1.56 (0.999)	-2.02 (0.566)	-6.13 (0.001)	-5.38 (0.0000)	2.76 (1.0000)	-1.93 (0.6154)	-6.32 (0.001)	-11.50 (0.001)	I(1)
Average urea	-3.16 (0.032)	-4.07 (0.017)	-5.33 (0.001)	-5.23 (0.001)	-2.93 (0.054)	-3.63 (0.0434)	-10.22 (0.001)	-10.16 (0.001)	I(0)
Average dap	-3.24 (0.028)	-3.57 (0.049)	-7.90 (0.001)	-7.85 (0.001)	-3.18 (0.031)	-3.56 (0.050)	-9.21 (0.001)	-9.54 (0.001)	I(0)
Precip itation	-3.36 (0.021)	-3.21 (0.102)	-6.84 (0.001)	-6.82 (0.001)	-3.36 (0.0208)	-3.21 (0.105)	-7.21 (0.001)	-7.3 (0.001)	I(1)
Tempe rature	-3.74 (0.008)	-6.06 (0.001)	-5.38 (0.001)	-5.24 (0.002)	-3.73 (0.009)	-7.9 (0.001)	-15.62 (0.001)	-16.37 (0.001)	I(0)
Rain	-2.33 (0.17)	-3.41 (0.07)	-8.23 (0.001)	-8.13 (0.001)	-2.35 (0.164)	-3.53 (0.054)	-8.45 (0.001)	-8.13 (0.001)	I(1)

From table (4.2) wheat, barley, precipitation and rainfall variables are not stationary at level and becomes stationary at first difference but temperature, average urea and average DAP are stationary at level.The variables are mixed order of integration i.e $I(0)$ and $I(1)$ no variables are integrated of order two, so $ARDL$ bound test cointegration technique for long-run and short-run relationship is appropriate than others cointegration technique.

4.2.2 Lag order selection

Practical problem in the estimation of time series and econometrics models relates to the number of variables to be included in the model and the maximum lag length to be applied. Time series and econometrics analysis depends critically on the lag order selected. Different lag orders can significantly affect the substantive interpretation of the estimates when those differences were large enough [Mukaras(2012)]. From table (4.3 & 4.4) maximum lag order selected by AIC, HQIC and SBIC was two when the dependent variable is barley and one for wheat.

Table 4.3: Lag order selection for barley

Lag	AIC	SBIC	HQ
0	2.875833	3.249486	2.995368
1	2.681208	3.288394	2.875452
2	2.176795*	3.072609*	2.457353*
*indicates lag order selected by the criterion			

Table 4.4: Lag order selection for wheat

Lag	AIC	SBIC	HQ
0	3.799794	4.126741	3.904387
1	3.360358*	3.920837*	3.539660*
2	3.726376	4.575042	3.992168
*indicates lag order selected by the criterion			

4.2.3 ARDL bound test for cointegration

After verifying the unit root properties of the variables, the bounds test of cointegration can be implemented for Equations (3.28), and (3.29) in order to analyze the long-run relationship between the variables. Table (4.5) shows the computed F-statistic and critical value (at the 5% significance level). We have evidence to reject the null hypothesis of no cointegration when the dependent variable is barley because 5.22 is higher than I(1) of PSS(2001) critical value at 5% level of significance. Since 0.43517 is lower than I(0) of

PSS(2001) critical value at 5% level, we have no evidence to reject the null hypothesis of no cointegration among variables when wheat is the dependent variable.

Table 4.5: ARDL Bound test

F-bound test	barley	wheat	signf.	I(0)	I(1)
F-statistic	5.22	0.43517	10%	3.087	4.277
k=5			5%	3.673	5.002
			1%	5.095	6.77

4.2.4 Parameter estimation and interpretation for wheat

Because of the non existence of cointegration among the variables was confirmed, the ARDL short run dynamic regression relationships between wheat yield per unit hectare, climate variables and fertilizers was estimated. Figure (4.7) shows top 20 evaluated models selected by AIC. There are 32 model evaluated among these the selected model due to minimum AIC is $ARDL(1,0,0,0,0,0)$, implies only the dependent variable is lag one and the others are at a current value that is lag(0).

From the selected ARDL model only the current values of the independent variables have significance effect on the current amount of Wheat produced in Amhara region, implies the lags of the independent variables have no impact on the amount per unit hectare of wheat produced in the region.

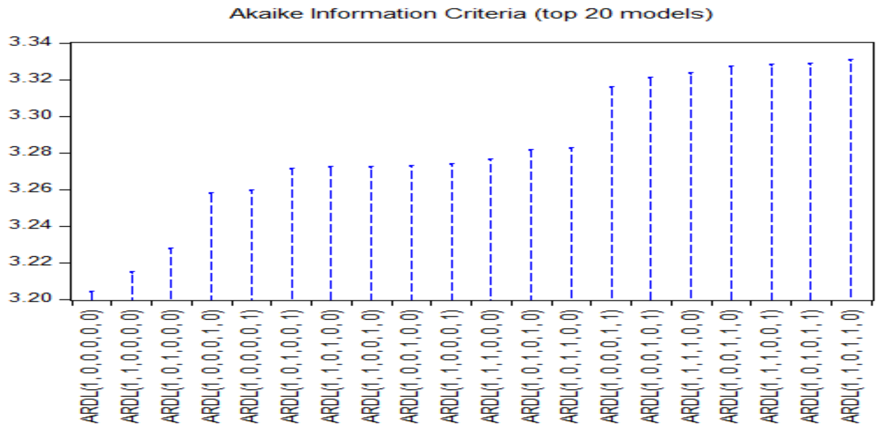

Figure 4.7: Model selection criteria graph for wheat

From Table (4.6) the value of adjusted R^2 is 0.399 implying 40% of the variation in amount of wheat yield is explained by the dynamic regressors. Current values of average urea, precipitation, temperature and the constant have a positive significance impact on the amount of Wheat yield per unit hectare but average DAP and amount of rainfall have no significance impact on the yield of wheat.

Holding other things constant, On average a 1% rise in current precipitation leads to an improvement of wheat yield per unit hectare by 1.317%. The coefficient of average urea is 4.4872 which implies that at citrus paribus a 1% increase in current amount of urea leads to a 4.4872% increase in the wheat yield per unit hectare also 1% rise in current temperature leads to an improvement in wheat yield per unit hectare by 0.4213% keeping the other constant. The substituted Coefficients in the model are:

$$\Delta wheat_t = 0.62 + 4.49\Delta aveurea_t + 1.32\Delta prec_t + 0.42\Delta temp_t$$

Table 4.6: Short run parameter estimate for wheat

Dependent Variable: D(WHEAT) Method: ARDL Model selection method: Akaike info criterion (AIC) Dynamic regressors (1 lag, automatic): D(AVEUREA) D(AVEDAP) D(PREC) D(RAIN) D(TEMPS) Number of models evalulated: 32 Selected Model: ARDL(1, 0, 0, 0, 0, 0)					
Variable	Coefficient	Std. Error	t-Statistic	Prob.	[95% Conf. Interval]
D(WHEAT(-1))	-0.071556	0.153453	-0.466303	0.6456	(-0.3898 0.2467)
D(AVEUREA)	4.487203	2.067295	2.170568	0.0410	(0.1999 8.7745)
D(AVEDAP)	0.526583	1.910526	0.275622	0.7854	(-3.4356 4.4888)
D(PREC)	1.317164	0.306491	4.297564	0.0003	(0.6815 1.9528)
D(RAIN)	0.013143	0.014804	0.887755	0.3848	(-0.0176 0.0438)
D(TEMPS)	0.421314	0.161915	2.602074	0.0163	(0.0855 0.7571
C	0.620697	0.222638	2.787918	0.0107	(0.1590 1.0824)
R-squared	0.528106	Mean dependent var		0.605172	
Adjusted R-squared	0.399408	S.D. dependent var		1.397443	
S.E. of regression	1.082989	Akaike info criterion		3.203831	
Sum squared resid	25.80302	Schwarz criterion		3.533868	
Log likelihood	-39.45555	Hannan-Quinn criter.		3.307195	
F-statistic	4.103445	Durbin-Watson stat		1.781891	
Prob(F-statistic)	0.006516				

4.2.5 Parameter estimation and interpretation for barley

Using model selection criteria, the model with ARDL (1,2,1,0,2,2) model has been selected from the 486 evaluated model because of the minimum values of AIC. This means the dependent variable Barley has lag one, average urea has lag 2, average DAP has lag one, precipitation has lag 0, rain and temperature has lag two each. From the graph below we can see the top 20 evaluated model.

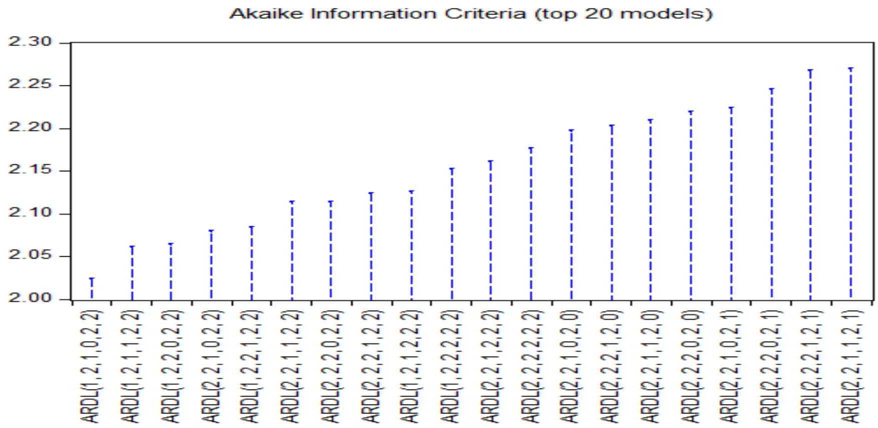

Figure 4.8: Model selection criteria Graph for Barley

4.2.5.1 Long run parameter estimate for barley

As a result of cointegration among variables is confirmed. So, the ARDL bound test for the short-run and long-run relationships between barley yield, climate variables and fertilizer used is estimated.

Table 4.7: Long run parameter estimate for Barley

Long-run estimates				
variables	coefficients	std.error	t-statistic	prob
aveurea	-3.494692	5.029044	-0.694902	0.4985
avedap	-28.76323	11.31615	-2.541787	0.0235
prec	2.768306	0.865427	3.198774	0.0064
rain	0.409652	0.149469	2.740723	0.0159
temp	-1.622681	0.837329	-1.937926	0.0731

The estimated coefficients in Table (4.7) shows, in the long run average DAP have a negative significant impact. Precipitation and rainfall have a positive significance impact but average urea and temperature have no significant impact on the amount of yield per hectare of barley at 5% level of significance.

In the long run estimates every 1% decrease in current average dap increase the yield per hectare of barley by 28.76% at citrus paribus and a 1% rise in current precipitation and rainfall increase the yield of barley by 2.768% and 0.41% respectively keeping the effect of one variable to the other constant.

4.2.5.2 Short run parameter estimation for barley

Once the study identified the presence of long run cointegration through the F-statistics and estimation of the long run coefficients, the next step is estimation of the error correction representation of long run relationship. The estimated ECM has two parts. First part contains the estimated coefficients of short run dynamics and the second part consists of the estimates of the error correction term (ECT) that measures the speed of adjustment whereby short-run dynamics converge to the long-run equilibrium path in the model. The ECM coefficient shows how fast variables restore to their equilibrium value.

From Table (4.8) the short-run coefficients estimates show the dynamic adjustment of all variables.The short run coefficients for D(average urea), D(average urea(-1)), D(average dap), D(prec), D(rain(-1)), D(temperature) and D(temperature (-1)) are statistically significant at the 5% level of significance. The magnitude of the adjusted (R^2) and the F-statistics show the model goodness of fit. Based on the value of adjusted R^2, the explanatory variables explained almost 76% of the variation in the production of barley and the F-statistics shows the model is well fitted .

Table 4.8: Short run coefficients and Error correction

Error Correction Representation				
regressor	coefficients	std.error	t-statistic	prob
c	-0.733735	0.221384	-3.314312	0.0051
t	0.443967	0.084822	5.234129	0.0001
D(AVEUREA)	6.870459	1.121342	6.126997	0.0005
D(AVEUREA(-1))	6.568769	1.202268	5.463646	0.0005
D(AVEDAP)	4.437762	1.279058	3.469553	0.0038
D(PREC)	1.056267	0.257176	4.107170	0.0011
D(RAIN)	0.025255	0.008768	2.880253	0.0121
D(RAIN(-1))	-0.084793	0.013954	-6.076724	0.0005
D(TEMPS)	-0.389559	0.078432	-4.966810	0.0002
D(TEMPS(-1))	0.350777	0.088972	3.942580	0.0015
CointEq(-1)*	-0.381557	0.058519	-6.520171	0.0001
R-squared	0.834689	Mean dependent var	0.398276	
AdjustedR-squared	0.756384	S.D .dependent var	0.993342	
S.E. of regression	0.490289	AIC	1.679155	
Sum squared resid	4.567283	SBIC	2.150636	
Log likelihood	-14.34775	HQIC	1.826817	
F-statistic	10.65942	Durbin-Watson stat	2.765895	
Prob(F-statistic)	0.000010			

In the short run, the current year barley production is affected positively by current and previous year amount of urea used, previous year temperature, current year precipitation and current year rainfall and also affected negatively by previous year rainfall and current year temperature. The change of time itself affects the yield positively.

As a 1% increase in current DAP make the yield of barley to increase by 4.43% quintal per hectare. A 1% increase in current urea improved the yield of barley by 6.87% and when current temperature rise by 1% the yield to decrease by 0.39% in the short run. A 1% rise in current precipitation increase the yield by 1.05%. If the current amount of rainfall rise by 1% the yield increase by 0.025%significance.

The coefficient of the error correction term that captures the speed of adjustment towards the long run equilibrium is found with the correct sign and magnitude. The speed of adjustment is -0.381557 which is highly significant indicating the speed of the adjustment back to the long run equilibrium after a short run shock. The estimated value of the coefficient indicates that about 38.15% of deviations from the long run equilibrium is adjusted in the current year, in the next year another 38.15% will be adjusted and the rest 23.7% will be adjusted in the third year. This shows that once the disequilibrium happens it takes around two and half year to correct any deviation from the long run equilibrium. The substituted parameter for equation (3.36) is

$$\begin{aligned} D(BARLEY) = -0.73 + 0.44t + 6.87D(aveurea) + 6.57D(aveurea(-1)) \\ +4.44 * D(avedap) + 0.0252 * D(rain) - 0.08 * D(rain(-1)) \\ -0.39(barley - (-28.76 * avedap(-1) + 2.77 * prec(-1) + 0.41 * rain(-1)) \end{aligned} \tag{4.1}$$

4.2.6 Model diagnostic

At this stage, an evaluation of the tentative model based on the estimated residual properties is to be performed.

4.2.6.1 Stability test

The stability of the coefficients in the model are checked by CUSUM and CUSUMSQ test conducted based on the recursive regression residuals as suggested by [Brown et al.(1975)Brown, Durbin, and Evans]. The stability of the long-run coefficient is tested by the short-run dynamics. Once the

ECM model has been estimated, the cumulative sum of recursive residuals (CUSUM) and the CUSUM of square tests are applied to assess the parameter stability [Peseran and Peseran(1997)].

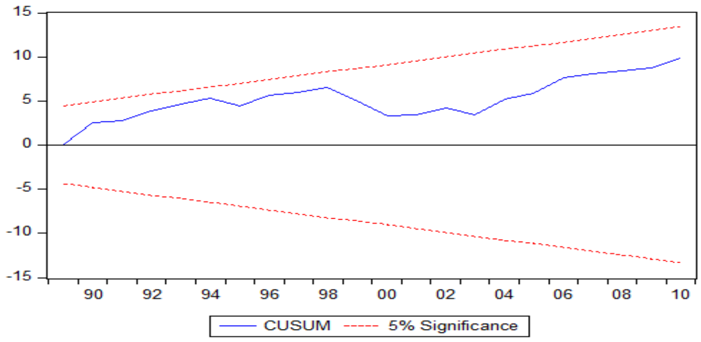

Figure 4.9: CUSUM plot for wheat

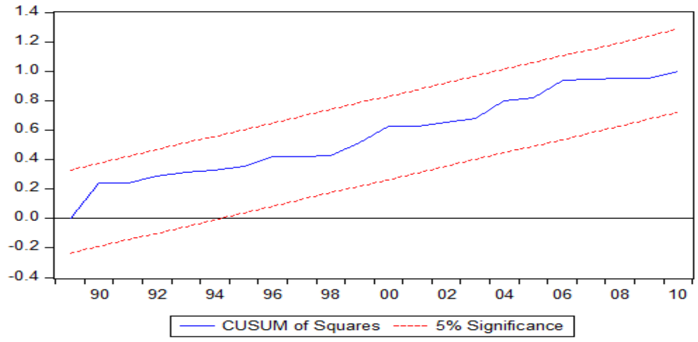

Figure 4.10: CUSUMSQ plot for wheat

Figure (4.9 & 4.10) shows the stability of coefficients in the model. Implies the stability assumption is fulfill for the model when wheat is the dependent variable.

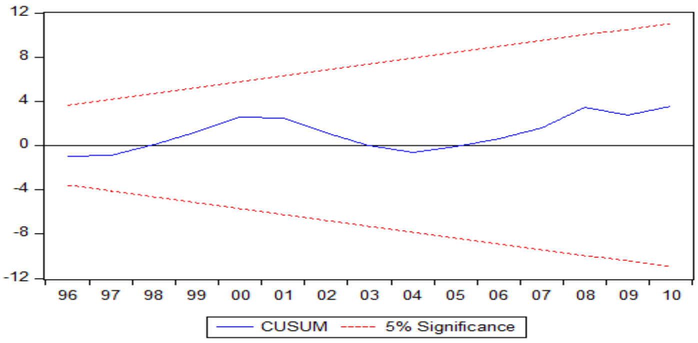

Figure 4.11: CUSUM plot for barley

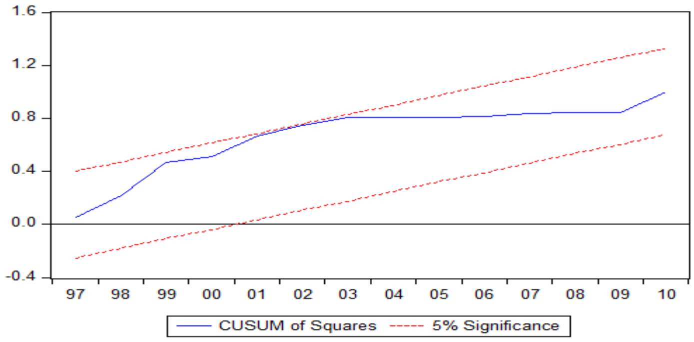

Figure 4.12: CUSUMSQ plot for barley

From figure (4.11 & 4.12) we conclude parameters are stable when the dependent variable is barley.

4.2.6.2 Specification test

Another potential problem may be omitted variable bias where some temperature-related and fertilizers-related variables such as (disease or pests,NPK,NPS,natural fertilizers, sunshine, humidity etc) that affect cereal yield but have been left out of *ARDL* model. Ramsey (1969) regression specification error test (RESET) for omitted variables is used.The purpose of this test is to provide evidence concerning the existence or non-existence of structural breaks in the ARDL model for barley.the null and alternative hypothesis is given by:

$$H_0 : the\ model\ is\ specified\ correctly \ \ vs\ H_1 :$$
$$the\ model\ is\ not\ correctly\ specified$$

Table 4.9: Ramsey Reset test for barley

Ramsey RESET Test Omitted Variables: Powers of fitted values from 2 to 3			
	Value	df	Probability
F-statistic	2.152804	(2, 15)	0.1507

Table 4.10: Ramsey Reset test for wheat

Ramsey RESET Test Omitted Variables: Squares of fitted values			
	Value	df	Probability
F-statistic	0.603378	(1, 21)	0.4460

From table (4.9 & 4.11) there is no evidence to reject null hypothesis of the Ramsey RESET test which is based on that the model is specified correctly because p-value(0.1507 and 0.4460) of both model is greater than the 5% level of significance. Therefore, Ramsey RESET test for functional form specification accepts the regression specification of the dynamic model. Furthermore, failing to reject the null in Ramsey reset test also further confirms that our model did not suffer from omitted variable bias [Tsadkan(2013)]

4.2.6.3 Autocorrelation test

Autocorrelations of the residuals are tested by using the Breusch-Godfrey Serial Correlation LM Test.

From table (4.11) we are in favor of not rejecting the null hypothesis of the residuals are serially uncorrelated because the p-values of the F-statistic is 0.5597 and 0.9183 for wheat and barley models respectively which is much

Table 4.11: Autocorrelation test using Breusch-Godfrey

serial correlation		
statistic	wheat	barley
F-statistic	0.3513	0.085745
p-value	0.5597	0.9183
decision	not correlated	not correlated

higher than the 0.05 level of significance, implying the residuals are serially uncorrelated.

4.2.6.4 Heteroscedasticity test

From table (4.12) we have no evidence to say the variance of the residuals are not constant since the p-value 0.6024 for barley model and 0.8222 for wheat model are greater than the 5% level of significance, implies the homoscedastic of residual assumption is satisfied.

Table 4.12: heteroscedasticity test using Breusch-Pagan-Godfrey

homoscedasticity of variance		
statistic	wheat	barley
F-statistic	0.8549	0.4712
p-value	0.6024	0.8222
decision	not correlated	not correlated

4.2.6.5 ARCH test

The F-statistic is an omitted variable test for the joint significance of all lagged squared residuals. The exact finite sample distribution of the F-statistic under H_0 is not known, but the LM test statistic is asymptotically χ^2 distributed as a under quite general conditions. From Table(4.13 & 4.14) we have evidence to reject the hypothesis of the square residual are autocorrelated.

Table 4.13: Auto regressive conditional heteroscedasticity test using LM test for barley

ARCH		
Variables	Barley	p-value
F-statistic	0.447950	0.8697
Obs*R-squared	4.829152	0.7757
decision	do not reject H_0	

Table 4.14: Auto regressive conditional heteroscedasticity test using LM test for wheat

ARCH		
Variables	wheat	p-value
F-statistic	0.328416	0.9389
Obs*R-squared	3.771970	0.8771
decision	do not reject H_0	

4.2.6.6 Normality of the residual

Jarque-Bera test which is based on the skewness and kurtosis of the residuals is used to know whether the distributions of the residual in the model are normal or not. For the model to provide a good description of the series, no underlying structure might be left in the residuals. So first of all it could be useful to examine the standardized residual plot. Quantile-Quantile plots are used to assess whether the data in a single series are normally distributed or not. If the two distributions are the same, the QQ-plot should lie on a straight line. If the QQ-plot does not lie on a straight line, the two distributions differ along some dimension. The pattern of deviation from linearity provides an indication of the nature of the mismatch. From Figure(4,13 & 4.14) we can see the two distribution doesn't differ along some direction

Table 4.15: normality test of residuals using Jarque-Bera test

Residuals	Jarque-Bera	P-value	Skewness	Kurtosis
From barley model	0.061	0.9700	0.0419	3.2081
From wheat model	1.8423	0.3981	-0.5943	3.3339

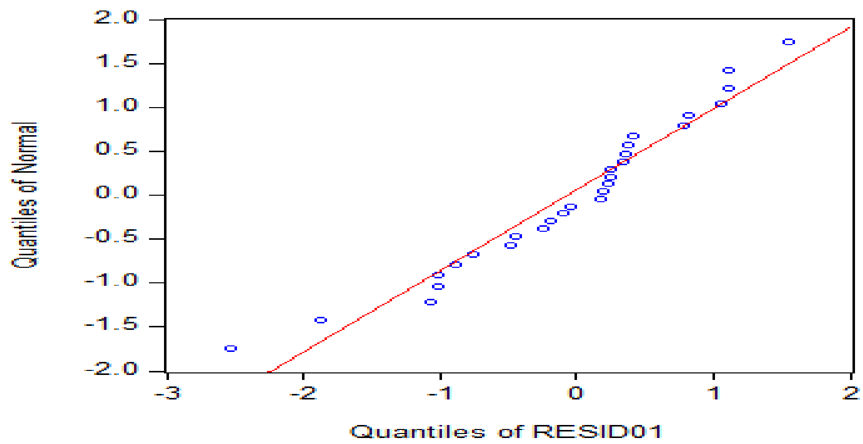

Figure 4.13: Q-Q plot of standardize residuals from wheat model

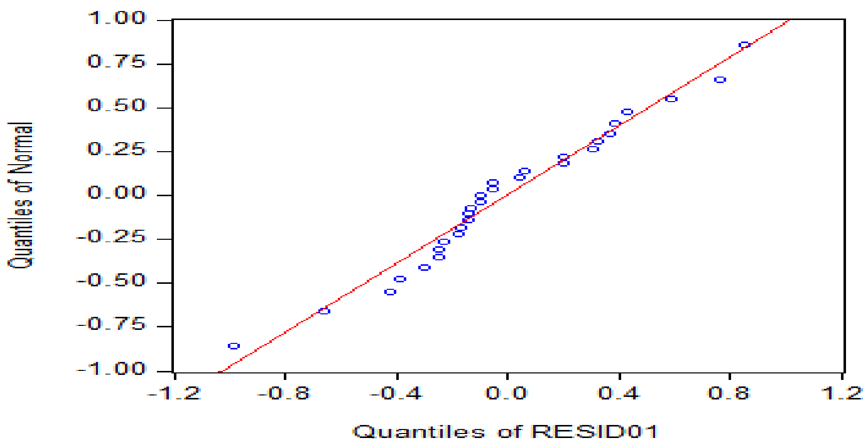

Figure 4.14: Q-Q plot of standardize residuals from barley model

From table (4.15) we have no evidence to reject the null hypothesis of normality because p-values in the Jarque-Bera test for both model i.e 0.97 and 0.398 are greater than $\alpha(0.05)$ level of significance. The residuals of both models are normally distributed.

The diagnostic tests reveal that the equation passed all the tests, i.e., the Breusch – Godfery's LM test rejects the presence of serial correlation, the Jarque-Bera test passed the normality assumption, Breusch-Pagan-Godfrey test does not indicate any evidence of heteroscedasticity of the residual for both model.

4.2.6.7 Granger causality test

Three types of Granger causalities were applied to find the causality: (1) Short-run causality—the Wald test was applied for all the lag independent variables using the joint F test; (2) long-run causality—investigated by verifying the coefficient of the error correction term (it should be between 0 and 1 with a negative sign), which implies convergence of the system back to the long-run equilibrium position, (3) joint (short-run and long-run) causality—the Wald test was applied to both the lagged independent variables and the error correction term using the joint F test[Türsoy(2017)].

From table (4.16) in the short run urea, DAP and precipitation Granger cause barley productivity. In the long run all variables Granger cause wheat yield because the ECT is negative and also between 0 and 1 finally in the joint causationall,all variables Granger cause barley productivity

Table 4.16: Granger causality test

Granger Causality			
variables	F-Statistics(Probability) (Short-Run)	t-Statistics (Long-Run)	Joint (Short- and Long-Run)
Δ urea	11.49207(0.0044)	4.393903(0.0006)	11.31134(0.0012)
Δ DAP	6.240988(0.0256)	"	9.772751(0.0022)
Δ precipitation	10.18507(0.0065)	"	
Δ rainfall	0.993788(0.3357)	"	9.745318(0.0022)
Δtemperature	0.009921(0.9221)	"	9.761519(0.0022)

4.2.7 discussion

In this study the $ARDL$ regression of wheat with other exogenous variable have no long run relationship which was inconsistent with previous studies exemplified by [Tao et al.(2014)Tao, Zhang, Xiao, Zhang, Rötter, Shi, Liu, Wang, Liu, and Zhang, Zhang and Huang(2013)]

In the short run model temperature has a positive impact on wheat yield which is the same as [Zhang and Huang(2013)] who found that the positive effect of temperature on the wheat yield existed in the southern part of Henan Province.

In the short run model precipitation has a significant positive effect on wheat yield which is consistent with [Zhai et al.(2017)Zhai, Song, Qin, Ye, and Lee] but in their study fertilizers effect is negative, which is inconsistent with this research.This may be due to that each type of average fertilizers are used separately in this research and aggregate fertilizers are used in their research.

In both short run and long run rain have a significance impact on the production of barley which is he same as [Fahimifard et al.(2011)Fahimifard, Sabouni, et al.]

CHAPTER FIVE

5 Conclusion and Recommendation

5.1 Conclusion

Increasing cereal crops production is one of the most important determinants of economic of one country. This study focused on determining empirically the impact of major factors on grain crops production in Amhara region. Based on the analytical results of this work, two issues can be resolved. One is whether the threat of climate change impacted the wheat and barley yield per hectare in Amhara region and the other is whether average quantity of fertilizer type used increased the wheat and barley yield per hectare.

In this paper, we examined the relationships among wheat and barley yield per unit hectare with precipitation,temperature,rainfall,average urea and average DAP in Amhara region from 1987 to 2017 by using an Autoregressive Distributed Lag (ARDL) to co-integration model. Employing the bounds testing to cointegration, the results showed that there is no-cointegration between the yield of wheat crops with average urea, average DAP,temperature, precipitation and rain but for the cereal crop barley the result showed that there is cointegration relationship.

In the short run static model,amount of wheat yield is positively affected by amount of average urea used,precipitation and temperature but average DAP and rain has no significance impact on the yield of wheat. This implies the significance variables improved the amount of wheat yield in the region through the period.

In the short run for barley, current amount of average urea, previous year of Average urea, precipitation,previous year temperature,current amount of rainfall and average DAP have positive impacts on the barley yield per hectare but temperature and previous year rainfall affects yield of barley negatively and also the time change have a positive impact. The coefficient of error correction term ecm(-1) which is highly significant indicating that there is a disequilibrium that can be adjusted in the long run. The estimated value of the coefficient indicates that disequilibrium in barley is offset by the short run adjustment in the same year.

In the long run precipitation and rain has a positive impact on the yield of barley but DAP have highly negative impact. Temperature and average urea has no impact in the long run. Average urea, average DAP, precipitation and rainfall increase the barley yield in the short run.

5.2 Recommendation

Depending on the result of this research we have several policy implications that could ensure continuous increases in the wheat and barley production per unit hectare and food security under climate change and fertilizers used in Amhara region. Average urea could be effective measures but average DAP have high negative impact in the yield of barley in the long run and have no significance impact in the short run in wheat yield. Therefore, to increase both wheat and barley yield federal and regional governments have to work lots of research on average DAP replacement and increase on usage of urea fertilizers. The uptake of nitrogen, phosphorus and potassium during the period of wheat and barley growth must change based on the characteristics of particular wheat and barley varieties. Government may need to make urea fertilizer available to farmers at low cost .In addition, agro-technicians could be arranged to guide farmers regarding the reasonable use of fertilizers.

In the presence of climate change the farmers should use different techniques such as irrigation to increase the impacts of precipitation on the yield of wheat and barley. since temperature have a negative impact on barley yield, government has to take initiative to introduce the high temperature resistant crops because temperature have an increasing trend. Rainfall and precipitation have a decreasing pattern across the region; therefore it is necessary for researchers to introduce the drought resistant crops.

References

[Abrha(2015)] Bihon Kassa Abrha. *Factors affecting agricultural production in Tigray region, Northern Ethiopia*. PhD thesis, 2015.

[Agency)()] CSA (Central Statistical Agency). National statistics report abstract.

[Alemayehu et al.(2011)Alemayehu, Paul, and Asrat] Seyoum Taffesse Alemayehu, Dorosh Paul, and Sinafikeh Asrat. Crop production in ethiopia: Regional patterns and trends. *international food policy research institute*, (0016), 2011.

[Ali et al.(2017)Ali, Liu, Ishaq, Shah, Ilyas, Din, et al.] Sajjad Ali, Ying Liu, Muhammad Ishaq, Tariq Shah, Aasir Ilyas, Izhar Din, et al. Climate change and its impact on the yield of major food crops: Evidence from pakistan. *Foods*, 6(6):39, 2017.

[Amikuzino and Donkoh(2012)] J Amikuzino and SA Donkoh. Climate variability and yields of major staple food crops in northern ghana. *African Crop Science Journal*, 20(2):349–360, 2012.

[Analytics et al.(2013)] Climate Analytics et al. Turn down the heat: climate extremes, regional impacts, and the case for resilience. 2013.

[Ayinde et al.(2011)Ayinde, Muchie, and Olatunji] OE Ayinde, Mammo Muchie, and GB Olatunji. Effect of climate change on agricultural productivity in nigeria: A co-integration model approach. *Journal of Human Ecology*, 35(3):189–194, 2011.

[Bahmani-Oskooee and Goswami(2003)] M Mohsen Bahmani-Oskooee and Gour G Goswami. A disaggregated approach to test the j-curve phenomenon: Japan versus her major trading partners. *Journal of Economics and Finance*, 27(1):102–113, 2003.

[Banerjee et al.(1998)Banerjee, Dolado, and Mestre] Anindya Banerjee, Juan Dolado, and Ricardo Mestre. Error-correction mechanism tests for cointegration in a single-equation framework. *Journal of time series analysis*, 19(3):267–283, 1998.

[Bank()] World Bank. "inescapable manufacturing- services nexus exploring the potential of distribution services".

[Bank(2013)] World Bank. World annual country's development report, 2013.

[Bewket(2009)] Woldeamlak Bewket. Rainfall variability and crop production in ethiopia: Case study in the amhara region. In *Proceedings of the 16th International Conference of Ethiopian Studies*, volume 3, pages 823–836. Norwegian University of Science and Technology Trondheim, Norway, 2009.

[BoFED and economic development(2011)] bureau of finance BoFED and economic development, 2011.

[Box et al.(2015)Box, Jenkins, Reinsel, and Ljung] George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

[Braun and Mittnik(1993)] Phillip A Braun and Stefan Mittnik. Misspecifications in vector autoregressions and their effects on impulse responses and variance decompositions. *Journal of Econometrics*, 59(3):319–341, 1993.

[Brown et al.(1975)Brown, Durbin, and Evans] Robert L Brown, James Durbin, and James M Evans. Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(2):149–163, 1975.

[(CSA)()] central statistical agencie (CSA). Agricultural sample survey report on land utilization.

[Deressa and Hassan(2009)] Temesgen Tadesse Deressa and Rashid M Hassan. Economic impact of climate change on crop production in ethiopia: Evidence from cross-section measures. *Journal of African economies*, 18 (4):529–554, 2009.

[Dickey and Fuller(1979)] David A Dickey and Wayne A Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431, 1979.

[Dickey and Fuller(1981)] David A Dickey and Wayne A Fuller. Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica: Journal of the Econometric Society*, pages 1057–1072, 1981.

[Dorosh et al.(2009)Dorosh, Ahmed, et al.] Paul Dorosh, Hashim Ahmed, et al. Foreign exchange rationing, wheat markets and food security in ethiopia. Technical report, Citeseer, 2009.

[Engle and Granger(1987)] Robert F Engle and Clive WJ Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276, 1987.

[Fahimifard et al.(2011)Fahimifard, Sabouni, et al.] SM Fahimifard, MS Sabouni, et al. Supply response of cereals in iran: an autoregressive distributed lag approach. *Journal of Applied Sciences*, 11 (12):2226–2231, 2011.

[Ferede et al.(2013)Ferede, Ayenew, Hanjra, and Hanjra] Tadele Ferede, Ashenafi Belayneh Ayenew, Munir A Hanjra, and M Hanjra. Agroecology matters: impacts of climate change on agriculture and its implications for food security in ethiopia. *Global food security: Emerging issues and economic implications*, pages 71–112, 2013.

[Food and agricultural Organization()] (FAO) Food and agricultural Organization. the special challenge for sub-saharan africa.

[for a green revolution in Africa()] AGRA Alliance for a green revolution in Africa. Africa agricultural status report the business of smallholder agriculture in sub-saharan africa.

[Gedefaw et al.(2018)Gedefaw, Denghua, Hao, Alemu, Chanie, and Agitew] Mohammed Gedefaw, Yan Denghua, Wang Hao, Basaznew Alemu, Mersha Chanie, and Genanew Agitew. Evaluation of adoption behavior of soil and water conservation practices in the simein mountain national park, highlands of ethiopia. *Cogent Food & Agriculture*, 4(1):1513782, 2018.

[Gujarati(2009)] Damodar N Gujarati. *Basic econometrics*. Tata McGraw-Hill Education, 2009.

[Harris and Sollis(2003)] Richard Harris and Robert Sollis. Applied time series modelling and forecasting. 2003.

[Hosking(1980)] John RM Hosking. The multivariate portmanteau statistic. *Journal of the American Statistical Association*, 75(371):602–608, 1980.

[Janjua et al.(2014)Janjua, Samad, and Khan] Pervez Zamurrad Janjua, Ghulam Samad, and Nazakatullah Khan. Climate change and wheat production in pakistan: An autoregressive distributed lag approach. *NJAS-Wageningen Journal of Life Sciences*, 68:13–19, 2014.

[Kaso and Guben(2015)] Tura Kaso and Gashaw Guben. Review of barley value chain management in ethiopia. *J Biol Agric Healthc*, 5:84–97, 2015.

[Klein and Welfe(1983)] Lawrence Robert Klein and Władysław Welfe. *Lectures in econometrics*. North-Holland Amsterdam, 1983.

[Licker et al.(2013)Licker, Kucharik, Doré, Lindeman, and Makowski] Rachel Licker, Christopher J Kucharik, Thierry Doré, Mark J Lindeman, and David Makowski. Climatic impacts on winter wheat yields in picardy, france and rostov, russia: 1973–2010. *Agricultural and Forest Meteorology*, 176:25–37, 2013.

[Lipper et al.(2014)Lipper, Thornton, Campbell, Baedeker, Braimoh, Bwalya, Caron, Cattaneo, Leslie Lipper, Philip Thornton, Bruce M Campbell, Tobias Baedeker, Ademola Braimoh, Martin Bwalya, Patrick Caron, Andrea Cattaneo, Dennis Garrity, Kevin Henry, et al. Climate-smart agriculture for food security. *Nature climate change*, 4(12):1068, 2014.

[Ljung and Box(1978)] Greta M Ljung and George EP Box. On a measure of lack of fit in time series models. *Biometrika*, 65(2):297–303, 1978.

[Lütkepohl(2005)] Helmut Lütkepohl. *New introduction to multiple time series analysis*. Springer Science & Business Media, 2005.

[M.(2017)] Kassie D; Raman M. Process of commercialization of agriculture in amhara region-ethiopia: Prospects and constrains. *Global Journal of*

Science Frontier Research, 2017.

[Matsumoto and Yamano(2011)] Tomoya Matsumoto and Takashi Yamano. The impacts of fertilizer credit on crop production and income in ethiopia. In *Emerging Development of Agriculture in East Africa*, pages 59–72. Springer, 2011.

[Mendelsohn et al.(2006)Mendelsohn, Dinar, and Williams] Robert Mendelsohn, Ariel Dinar, and Larry Williams. The distributional impact of climate change on rich and poor countries. *Environment and Development Economics*, 11(2):159–178, 2006.

[MoFED and economic development()] Ministry of finance MoFED and economic development. "survey of the ethiopian economy ii, reviewof developments (1998/99-2003/04)".

[Mohamed(2017)] Abduselam Abdulahi Mohamed. Food security situation in ethiopia: a review study. *International Journal of Health Economics and Policy*, 2(3):86–96, 2017.

[Mohammed(2010)] Yunus Hussien Mohammed. *Modeling the factors affecting cereal crop yields in the Amhara National Regional State of Ethiopia.* PhD thesis, 2010.

[Mukaras(2012)] Mohamed Mukaras. *Fundamental Principles of Time Series Econometrics , Theory and Applications.* 2012.

[Narayan(2005)] Paresh Kumar Narayan. The saving and investment nexus for china: evidence from cointegration tests. *Applied economics*, 37(17): 1979–1990, 2005.

[Newey and West(1987)] Whitney K Newey and Kenneth D West. r= k. this orthogonality condition can be employed to form a generalized method of moments (gmm, hansen (1982)) estimator of 6* by choosing 6) as the solution to. *Econometrica*, 55(3):703–708, 1987.

[Nkoro et al.(2016)Nkoro, Uko, et al.] Emeka Nkoro, Aham Kelvin Uko, et al. Autoregressive distributed lag (ardl) cointegration technique: application and interpretation. *Journal of Statistical and Econometric Methods*, 5(4):63–91, 2016.

[Perron and Ng(1996)] Pierre Perron and Serena Ng. Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *The Review of Economic Studies*, 63(3):435–463, 1996.

[Pesaran and Shin(1998)] M Hashem Pesaran and Yongcheol Shin. An autoregressive distributed-lag modelling approach to cointegration analysis. *Econometric Society Monographs*, 31:371–413, 1998.

[Pesaran et al.(2001)Pesaran, Shin, and Smith] M Hashem Pesaran, Yongcheol Shin, and Richard J Smith. Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, 16 (3):289–326, 2001.

[Peseran and Peseran(1997)] MH Peseran and B Peseran. Working with microfit 4: Interactive econometric analysis, 1997.

[Phillips and Perron(1988)] Peter CB Phillips and Pierre Perron. Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346, 1988.

[Poskitt et al.(1982)Poskitt, Tremayne, et al.] Don Stephen Poskitt, AR Tremayne, et al. Diagnostic tests for multiple time series models. *The Annals of Statistics*, 10(1):114–120, 1982.

[Rahimi and Shahabadi(2011)] Mohammad Rahimi and Aboulfazl Shahabadi. Trade liberalization and economic growth in iranian economy.

Available at SSRN 1976299, 2011.

[Ramsey(1969)] James Bernard Ramsey. Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society: Series B (Methodological)*, 31(2):350–371, 1969.

[Rashid et al.(2019)Rashid, Abate, Lemma, Warner, Kasa, and Minot] Shahidur Rashid, Gashaw T Abate, Solomon Lemma, James Warner, Leulsegged Kasa, and Nicholas Minot. The barley value chain in ethiopia. *F1000Research*, 3, 2019.

[Roberts et al.(2009)] TL Roberts et al. The role of fertilizer in growing the world's food. *Better crops*, 93(2):12–15, 2009.

[Samuel et al.(2017)Samuel, Shem, Daniel, and Silas] Ogallah Samson Samuel, Wandiga Shem, Olago Daniel, and Oriaso Silas. Impacts of climate variability and climate change on agricultural productivity of smallholder farmers in southwest nigeria. *International Journal of Scientific & Engineering Research*, 8(10), 2017.

[Tao et al.(2014)Tao, Zhang, Xiao, Zhang, Rötter, Shi, Liu, Wang, Liu, and Zhang] Fulu Tao, Zhao Zhang, Dengpan Xiao, Shuai Zhang, Reimund P Rötter, Wenjiao Shi, Yujie Liu, Meng Wang, Fengshan Liu, and He Zhang. Responses of wheat growth and yield to climate change in different climate zones of china, 1981–2009. *Agricultural and Forest Meteorology*, 189:91–104, 2014.

[Tesso et al.(2012)Tesso, Emana, and Ketema] Gutu Tesso, Bezabih Emana, and Mengistu Ketema. A time series analysis of climate variability and its impacts on food production in north shewa zone in ethiopia. *African Crop Science Journal*, 20(2):261–274, 2012.

[Tsadkan(2013)] A Tsadkan. The nexus between public spending and economic growth in ethiopia: Empirical investigation. *Unpublished Master Thesis, Addis Ababa University*, 2013.

[Türsoy(2017)] Turgut Türsoy. Causality between stock prices and exchange rates in turkey: Empirical evidence from the ardl bounds test and a combined cointegration approach. *International Journal of Financial Studies*, 5(1):8, 2017.

[Udoh and Ogbuagu(2012)] Elijah Udoh and Uchechi R Ogbuagu. Financial sector development and industrial production in nigeria (1970-2009): An ardl cointegration approach. *Journal of Applied Finance and Banking*, 2(4):49, 2012.

[Urgessa(2015)] Tessema Urgessa. The determinants of agricultural productivity and rural household income in ethiopia. *Ethiopian Journal of Economics*, 24(2):63–91, 2015.

[Wooldridge(2015)] Jeffrey M Wooldridge. *Introductory econometrics: A modern approach*. Nelson Education, 2015.

[You et al.(2009)You, Rosegrant, Wood, and Sun] Liangzhi You, Mark W Rosegrant, Stanley Wood, and Dongsheng Sun. Impact of growing season temperature on wheat productivity in china. *Agricultural and Forest Meteorology*, 149(6-7):1009–1014, 2009.

[Zhai et al.(2017)Zhai, Song, Qin, Ye, and Lee] Shiyan Zhai, Genxin Song, Yaochen Qin, Xinyue Ye, and Jay Lee. Modeling the impacts of climate change and technical progress on the wheat yield in inland china: An autoregressive distributed lag approach. *PloS one*, 12(9):e0184474, 2017.

[Zhang and Huang(2013)] Tianyi Zhang and Yao Huang. Estimating the impacts of warming trends on wheat and maize in china from 1980 to 2008 based on county level data. *International Journal of Climatology*, 33(3):699–708, 2013.