ADDIS ABABA UNIVERSITY

GRADUATE STUDIES PROGRAM

COLLEGE OF SCIENCE

DEPARTMENT OF STATISTICS

A MULTIVARIATE TIME SERIES ANALYSISA OF AGROMETEOROLOGICAL DATA TO ASSESS THE EFFECT OF CLIMATE VARIABILITY ON PRODUCTION OF SORGHUM: THE CASE OF MELKASSA, ETHIOPIA

By

TIGIST MIDEKSA

A Thesis submitted to the Office of Graduate Programs of Addis Ababa University in Partial fulfillment of the requirement for the Degree of Master of Science in Statistics (Biostatistics)

June, 2011

ADDIS ABABA UNIVERSITY

GRADUATE STUDIES PROGRAM

COLLEGE OF SCIENCE

DEPARTMENT OF STATISTICS

A MULTIVARIATE TIME SERIES ANALYSISA OF AGROMETEOROLOGICAL DATA TO ASSESS THE EFFECT OF CLIMATE VARIABILITY ON PRODUCTION OF SORGHUM: THE CASE OF MELKASSA, ETHIOPIA

By

TIGIST MIDEKSA

Approve by the Board of Examiners:

-------------------------------------- Signature --------------------

Chairman, Department graduate committee

Signature --------------------

Examiner

Signature --------------------

Examiner

ACKNOWLEDGEMENT

First of all I would like to thank God for His guidance and love and for giving me the opportunity to study MSC. Degree.

My special gratitude goes to my advisor, Dr.Butte Gotu for his continual advice, useful comments and suggestions throughout the preparation of this paper.

I would like to thank the following people for their encouragement, contribution in providing relevant papers, data and their invaluable advice in improving my work with respect to their discipline: Dr.Girma Taye of Ethiopian Institute of Agricultural Reserch, head of department of Biometrics GIS and Agrometeorology Process and Dr.Girma Mamo of Melkassa Agricultural Research Center .

My sincere thanks go to my family for their lovely and tireless support. Last but not the least my appreciation goes to my brother Woldeyesus for his moral and material support.

Table of Contents **Page**

List of tables

List of Figures

List of tables in Appendix

Abbreviations and acronyms

ACF	Auto Correlation Function
ADF	Augmented Dickey – Fuller
AIC	Akaike Information Criteria
ANOVA	Analysis of Variance
AR	Auto Regressive
ARMA	Auto Regressive Moving Average
ARIMA	Auto Regressive Integrated Moving Average
BIC	Bayesian Information Criteria
DF	Degree of Freedom
EIAR	Ethiopian Institute of Agricultural Research
FAO	Food and Agriculture Organization
FEVD	Forecast Error Variance Decomposition
GDP	Gross Domestic Product
GCMs	Global Circulations Models
HQ	Hannan-Quinn Information Criteria
IC	Information Criteria
IRF	Impulse Response Function
JB	Jarque-Bera
LM	Lagrange Multiplier
MAPE	Mean Absolute Percentage Error
MA	Moving Average
MARC	Melkassa Agricultural Research Center
MoARD	Ministry of Agriculture and Rural development

MTS	Multivariate Time Series
NMSA	National Meteorological Services Agency
NRRD / MOA	Natural Resources Regulatory Department of the Ministry of Agriculture
PACF	Partial Autocorrelation Function
PP	Phillips – Perron
STAR	Space Time Autoregressive
U.S	United States of America
VAR	Vector Autoregressive
VARMA	Vector Autoregressive Moving Average
VEC	Vector Error Correction

Abstract

Climate change affects all economic sectors to some degree, but the agricultural sector is perhaps the most sensitive and vulnerable. In the last three decades, Ethiopia has been affected by climate-related hazards. Agriculture, the most dominant sector in the national economy, has been most at risk because of its dependence on seasonal rainfall. Anticipated climate change has negatively impacted the agricultural sector due to increased temperatures and decreased or greater variability in precipitation, leading to increased food insecurity.

The aim of this thesis is to study the effects of rainfall and temperature on sorghum yield over long period of time.

The study consists of thirty two years consecutive summer seasons total summer rainfall in mm, average temperature in degree centigrade and sorghum yield of medium maturing Gambella-1107 variety in quintal per hectare (Q/h) which is used as a control variety in Melkassa Research Center of EIAR. Multivariate time series analysis that is Vector Auto Regressive model is used for this study.

The results obtained show that rainfall and yield are highly variable. Rainfall shock has significant impact on rainfall temperature and yield, temperature shock has a significant impact on temperature rainfall and yield and also yield shock has a significant impact on yield. It was observed that rainfall variation could fully be explained by its own innovations. For temperature, in the first round, 56% variation of temperature has resulted from the shock of its own innovation and 44% variation resulted from the change in rainfall. Yield variation of up to 48.1% is explained by changes in rainfall amounts, the percentage contribution of yield shock for its forecast variance is around 48.6%.

There exist high inter annual variability in summer rainfall total which is an evidence to climate variability in the study area and also yield is highly variable. Past year and following year rainfall temperature and yield are autocorelated and rainwater is a principal component in determining sorghum yield.

Keywords*: VAR, Sorghum(Gambella#1107) variety, Melkassa*

CHAPTER ONE

I. Introduction

1.1 Background of the Study

During the recent decades, global climate change has been at the centre of many scientific studies. Global climate change, caused by natural processes as well as anthropogenic factors, is likely the major and the most important environmental issue that will affect the world during the 21st century.

Climate change is a phenomenon due to emissions of greenhouse gases from fuel combustion, deforestation, urbanization and industrialization (Upreti, 1999) resulting variations in solar energy, temperature and precipitation. It is a real threat to the lives in the world that largely affects water resources, agriculture, coastal regions, freshwater habitats, vegetation and forests, snow cover and melting and geological processes such as landslide, desertification and floods, and has long-term effects on food security as well as in human health.

Climate is the average weather parameters, usually 30 years of a given locality. The measures of climate include mainly the estimates of average values of weather parameters and measures of variability near to the average value. It includes patterns of temperature, precipitation, humidity, wind and seasons. These climate patterns play a fundamental role in shaping natural ecosystems, and the human economies and cultures that depend on them. "Climate change" affects more than just a change in the weather. Because so many systems are tied to climate, a change in climate can affect many related aspects of where and how people, plants and animals live, such as food production, availability and use of water, and health risks.

Climate change will affect all economic sectors to some degree, but the agricultural sector is perhaps the most sensitive and vulnerable. Although the climate is constant in comparison with the weather, it may not be the same from one decade to another.

In the last three decades, Ethiopia has been affected by climate-related hazards. Agriculture, the most dominant sector in the national economy, has been most at risk because of its dependence on seasonal rainfall. Anticipated climate change has impacted negatively on the agricultural sector due to increased temperatures and decreased or greater variability in precipitation, leading to increased food insecurity.

Ethiopia has the most variable rainfall patterns. A number of professionals and organizations have documented scientifically interesting reports on Ethiopian rainfall variability by classifying the country into various and wide temporal and spatial rainfall categories (NMSA,1996a & b; FAO, 1984; FAO, 1989; Degefu, 1987; Gemechu, 1977; Ethiopian Delegation, 1984; Gonfa, 1996; NRRD/MOA, 2000) and many others. According to Haile (1986) drought occurs every 3-4 years in the northern and 6-8 years in other parts of Ethiopia. According to Kidson (1977) a steady downward trend of rainfall since the peaks of the 1950s has expanded more in 1980s covering almost the whole of Africa. In the Ethiopian context, opinions are also divergent regarding the arrival of the rainy season, which used to occur in March but is now gradually shifting to April through July and, therefore, there is a progressive shortening of the growing periods and the corresponding seasonal rainfall totals.

Sorghum is one of the crops mostly grown in wide agro-ecological zones throughout the world. This crop is considered as a potential adaptation option for millions of farmers hit hard by climate change. The crop appeared to have been domesticated in Ethiopia about 5000 years ago (Taylor, 2003). Currently, large part of sorghum production areas in Ethiopia fall under the arid and semi-arid regions of the country that are characterized by high rainfall variability and low soil water storage capacity. The crop is widely grown in low moisture areas due to its high capacity to tolerate soil water deficit and wide range of ecological diversity (Kidane et al., 2006; MoARD, 2007). Despite its significant area coverage, however, the national average sorghum productivity is estimated to be less than one tone per hectare (Mesfine et al., 2005).

Melkassa, which is located in the Central Rift Valley of Ethiopia, is one of the areas where sorghum is widely grown. The localized temporal rainfall and temperature variation during cropping season induces an important challenge to sorghum production and in turn to food security. Therefore, apart from understanding meteorological variability and change on crop production and productivity per se, it is important to study the interaction and comovements among climate variables and agricultural production in order to assess the impact of climate change on agriculture.

1.2 Statement of the Problem

It is a fact that agriculture provides a strong backbone to the overall Ethiopian economic welfare, employing over 80% of the population and accounting for 50 % of the GDP (MoARD, 2007). While the country has about 3.7 million hectare of irrigable land plus 110 billion cubic meter of surface water, the cropping system is mostly carried out under the rainfed condition (Awulachew et al., 2007)

Agricultural economy in Ethiopia has features of subsistence economy, meaning provides sufficient food to last only from one harvest to the next. Therefore, a failure of one harvest means starvation for the ensuing year, shortage of seed for the next cropping season and loss of animal power to plough the fields (Abate, 1994).

Meagre harvests, total crop failures and rapid decline in productivity, particularly in drought prone areas like the Central Rift Valley, are common characteristics. The climate of arid and semi arid region of Ethiopia is characterized by high rainfall variability and unpredictability, strong winds, high temperature and high evapotranspiration (Girma, 2005). It is, therefore, essential to quantify its effects especially on crop yields because it is likely to be most affected by sudden or gradual adverse changes in climatic conditions.

This paper responds to the question of how one can use historic time series data to study the effects of rainfall and temperature on production of sorghum.

1.3 Objectives of the Study

The general objective of this thesis is to study the effects of rainfall and temperature on sorghum yield over long period of time using Multivariate Time Series Techniques.

Specific objectives of this research are:

- To study the dynamic relationships over time among the time series rainfall, temperature and sorghum yield using appropriate statistical model.
- To provide information that could help policy makers in formulating strategies to cope with climate change and its consequences.

1.4 Significance of the Study

Ethiopian rain fed farmers around Melkassa, researchers and extension workers will be the major users of this study by translating the results in to the field level management. In addition officers in the Ministry of Agriculture and Rural Development will use as a supportive tool for disaster management and for agriculture related policy and decision making and the results can also be used for further study in this area.

CHAPTER TWO

LITERATURE REVIEW

2.1 General literature about the model used in the study

2.1.1 Time Series Analysis

Time series data are widely available in different fields including medicine, finance, the physical science, in particularly, the earth science and engineering. Modeling time series data effectively is important for many decision-making activities as the models can be used to forecast future values and to help understand the underlying relationships within the time series.

2.1.2 Time Series

A time series is a series of observations, $\mathbf{Y}_{it}$; i= 1,…,n; t=1, …,T made sequentially over time. Here i indexes the different measurements made at each time point t; n is the number of variables being observed and T is the number of observations made. If n is equal to one then the time series is referred to as univariate (Chatfield, 1989), and if it is greater than one the time series is referred to as multivariate (Hannan, 1970).

2.1.3 Univariate Time series

The majority of the work on time series analysis has been highly concentrated on univariate models, which are the simplest types to work with. Within this area, most work has been concentrated on the family of linear models, which includes the autoregressive (AR), moving average (MA) and autoregressive moving average models (ARMA) (Box, 1970).

2.1.4 Multivariate Time series Model and the Vector Autoregressive Process

Multivariate time series (MTS) analysis is a powerful tool for the analysis of data. It is of considerable interest in a variety of fields such as engineering, the physical sciences- particularly the earth sciences (e.g. metrology and geophysics), and economics and business (Reinsel, 1997).

There are many types of MTS models to choose from. First of all there is a decision about linear and non-linear models (Casdagli, 1992), and then the specific type of model within these two categories. The main linear models are the Vector Autoregressive (VAR) process, the Vector Moving Average (VMA) process and the Vector Auto Regressive Moving Average (VARMA) process (Lütkepohl, 1993).

The application of MTS is wide-spread from, for example the medical field where the relationship between exercise and blood glucose can be modeled (Crabtree et al., 1990), and the engineering field where the process control effectiveness can be evaluated (De Vries & Wu, 1978). Grubb (1992) develop and compare several multivariate models for some multiple time series data to explore the dynamic relationships between the variables and to forecast the system by using a suitable representation. The data are indices of the price of flour at three sites in the USA. The models all come from the vector autoregressive moving average class so that comparisons between them can easily be made using criteria such as Akaike's Information Criterion.

Kenneth et al. (1996) develop a new multivariate, time-series prediction model that employs past values of earnings, short-term accruals and cash-flows as independent variables in a time-series regression. The predictive results indicate that this model clearly outperforms firm-specific and common-structure ARIMA models as well as a multivariate, cross-sectional regression model popularized in the literature employed for prediction of cash flow.

Vector Autoregressive model was firstly introduced by Sims (1980) for macro-economic forecasts. It is found to exhibit greater predictive efficiency and accuracy than the large scale structural economic models. Since then, VAR models have been applied to various problems. Hui (2004) applies VAR model to study the air pollution levels in Hong Kong. In the study the levels of the major air pollutants in Hong Kong are modeled by various vector autoregressive (VAR) models, including the general VAR model, structural VAR model, and the Space-Time autoregressive (STAR) model. Finally the model is improved by using the STAR. The forecasting accuracies of all the final models were found to be quite satisfactory.

Samuel (2007) used VAR model to study the demand of electricity in Ethiopia as related to consumption, population and gross domestic product.

2.2 Specific Literature Related to this Study

Climatic variability exert a strong influence on a variety of economic sectors (Easterling and Mjelde, 1987) including agriculture, forestry, water resource management, road maintenance, construction, tourism and public transport (Taylor, 1972). Agriculture is the most frequently cited human system likely to be affected by climate variations (Smit *et al*., 1996). Climate is the primary driving force of agricultural production (Hollinger, 1994).

Climate affects a wide range of agricultural activities, output and input resources like yields, land quality, on-farm storage, water supplies, labor migration rates of urban and rural communities, population growth, farm income and farmers skills (Washington and Dawning, 1999). Climatic variations have widespread agricultural effects that affect every production management practice from seedbed preparation to harvesting (Hollinger, 1994).

Blignaut et al. (2006) study the impact of changes in climatic conditions on agriculture in South Africa, and indeed throughout much of Africa. To address the broad-scale impact of climate change on agriculture they consider rainfall and temperature data from 1970 to 2006 for South Africa's nine provinces. The relationship between rainfall and temperature was analyzed using ANOVA software and for all areas, the covariance of rainfall and temperature was very significantly negative for most provinces. It was observed that with the exception of the arid areas, the covariance observed between temperature and rainfall has become stronger over the decade. The hotter it gets, the less rainfall there is in all regions.

Naipal (2004) analyzes the trends and variation of rainfall and temperature in Suriname. The analysis in the study was based on monthly time series. Linear regression analyses was used to identify negative or positive trends in monthly rainfall and temperature and five Global Circulations Models (GCMs) were used to predict future changes. The absence of a clear trend in monthly rainfall within the stations as well as between the stations were observed. Results of long-term trends in average monthly and maximum temperature of nearly all months indicate a warmer trend. Results of GCMs have shown that under global climate change by 2100, the monthly rainfall is predicted to decrease from 82 in January to 66 mm in August, and increase from 36 in

September to 47 mm in December. The monthly temperature is predicted to increase with 1.3 to 4.3 ^{o}C by 2100.

Ben Mohamed et al. (2002) investigate the impact of current climate variability and future climate change on millet production for three major millet producing regions in Niger. Statistical models (Multiple regressions) have been used to predict the effects of climate change on future production on the basis of thirteen available predictors. Crop yield (or production) is estimated on the basis of agro-climatic parameters (predictors). Based on the analysis of the past 30-years of rainfall and production data, the most significant predictors of the model were found to be sea surface temperature anomalies, the amount of rainfall in July, August and September, the number of rainy days and the wind erosion factor. In 2025, production of millet is estimated to be about 13% lower as a consequence of climate change, translated into a reduction of the total amount of rainfall for July, August and September, combined with an increase in temperature while maintaining other significant predictors at a constant level.

David et al. (2007) analyzed the relationship between crop yield and three climatic variables (minimum temperature, maximum temperature, and precipitation) for 12 major Californian crops: wine grapes, lettuce, almonds, strawberries, table grapes, hay, oranges, cotton, tomatoes, walnuts, avocados, and pistachios. The months and climatic variables of greatest importance to each crop were used to develop regressions relating yield to climatic conditions. Results show that, most crops, fairly simple equations using only 2–3 variables explained more than two-thirds of observed yield variance. The types of variables and months identified suggest that relatively poorly understood processes such as crop infection, pollination, and dormancy may be important mechanisms by which climate influences crop yield.

Gay et al. (2006) explore the relation between coffee production and climatic and economic variables in Veracruz (Mexico) in order to estimate the potential impacts of climate change. For this purpose, an econometric model is developed (a multiple regression model) that integrates climatic and economic determinants of coffee production in Veracruz. The objective was to construct a production function that could give information on how this activity responds to changes in economic and climatic variables. The model is validated by means of statistical analysis, and then used to project coffee production under different climatic conditions. Climate change scenarios are produced considering that the observed trends of climate variables will continue to prevail until the year 2020. An approach for constructing simple probability scenarios for future climate variability is presented and used to assess possible impacts of climate change beyond what is expected from changes in mean values. Results show that temperature is the most relevant climatic factor for coffee production. The results for the projected climate change conditions for year 2020 indicate that coffee production might not be economically viable for producers, since the model indicates a reduction of 34% of the current production.

Chen et al. (2004) developed quantitative estimates of the impacts of annual average climate conditions on yield variability of major agricultural crops across the U.S. This is accomplished by estimating a Just-Pope stochastic production function using a time series and panel data set of U.S. crop yields for major crops by state. The results show changes in average climate conditions cause alterations in crop yield levels and variability. The effects are found to differ by crop. For instance, temperature has the reverse effect on crop yield. For sorghum, higher temperatures reduce yields and yield variability.

Almaz (2004) analyzed the effect of climate change on agriculture by using different meteorological parameters and different types of crop models and statistical analysis (descriptive) in order to investigate the influence of climatic variability on crop performance in terms of crop production and protection aspect of Ethiopia. The result indicates that climate change induced by increasing greenhouse gases is likely to affect crop yields differently from region to region. Decreases in potential crop yields are likely to be caused by shortening of the crop-growing period, decrease in water availability due to higher rates of evaporation.

Adugna (2005) attempts to show patterns of rainfall and provide insight into the preparation of an early warning system in Ethiopia using time series analysis techniques. Auto-Regressive Moving Average (ARMA) and Vector Auto-Regressive (VAR) models are used to see the pattern of rainfall and response of yield to rainfall as well as to previous yield shocks. Results from estimation of VAR show that current levels of yield respond to previous levels of yield even more than responses to rainfall in most provinces.

The literature mentioned above shows agriculture is the most frequently cited human system likely to be affected by climate variations, and the impact of current climate variability and future climate change on agricultural production can be assessed by using different statistical models and different crop models.

More importantly, what makes this study different from previous ones is that it relates rainfall, temperature and agricultural yield dynamically using a well-established time series technique called VAR model. As far as we know VAR model have never been used in studying climate change impact on agriculture yield in Ethiopia.

CHAPTER THREE
DATA AND METHODOLOGY

3.1 The Data and Study Area

The data for this study were obtained from Melkassa Agricultural Research Center of Ethiopian Institute of Agricultural Research (EIAR). Melkassa was established in 1962 and is about 115 km away from Addis Ababa. It is one of the research centers located in the Central Rift Valley. It is located between 8:24°N and 39:19°E with an altitude of 1540 meters above sea level. It is found in the East Shoa Zone along the Rift valley. Flood, drought, soil erosion and rainfall deficiency are some of the natural hazards that are frequent in the zone.

The study area constitutes the heart and corridor of the Ethiopian Rift that extends from the Afar Triangle in the north to the Chew Bahir in southern Ethiopia (FAO, 1984). Physiographyically, Central Rift Valley is characterized by almost level to gentle slope (reaching up to 1800 m.a.s.l) and a benched rift valley without sedimentary surface features. It has also volcanic lacustrine terraces formed in quaternary lacustrine siltstone, sand stone, inter-bedded pumice and stuffs, with fault topography bordering the major lakes plus parallels and low coastal ridges. It also has quaternary alluvial landforms, mostly bordering the main river valley or located at the foot of the higher plateaus, as alluvial colluvial cones (Markin *et al.*, 1975;FAO, 1989). Despite the variability in rainfall and the prevalence of the long established spiral of land degradation in the region, there is considerable opportunity for raising the level of farmer’s returns through transfer of improved technologies (material and knowledge). To that end, many research and development institutions work in this region. These include Melkassa Agricultural Research Center (MARC), Debre Zeit Research Center, Adami Tulu Research Center, Awasa Research Center, Werer Research Center, Miesso and Arsi Negele sub-stations, Kulumsa Research Center, Wenji Sugar Estate, Metehara Sugar Estate, Upper Awash Agro Industries and Horticultural Crops Farm at Ziway, Adami Tulu Pest Control Plant, as well as many private investors and processing plants.

For this study a total of 32 years summer season (June, July, August and September) total rainfall, average temperature and sorghum yield (Gambella -1107 variety) data obtained from the agrometrology department of the center were used. The time period included is from 1978 to 2009 G.C. The analysis are done using STATA and Eviews software.

3.2 Variables in the Study

Climate variability is partially explained by its key indicators that are rainfall and temperature. Climate change impact on agricultural production can be seen from the historical effects of rainfall and temperature on crop yield, therefore the variables of interest are sorghum yield (Gambella -1107) variety in Quintal/hectare, rainfall in mm and temperature in ^{o}C.

3.3 Methodology

3.3.1 The VAR Model

Multivariate time series analysis is used when one wants to model and explain the interactions and co-movements among a group of time series variables. In analogy with the univariate case, it is one major objective of multiple time series analyses to determine suitable functions that may be used to obtain forecasts with “good” properties. It is also often of interest to learn about the dynamic interrelationships between a number of variables.

The vector autoregressive (VAR) model is one of the most successful, flexible, and easy to use models for the analysis of multivariate time series. It is a natural extension of the univariate autoregressive model to multivariate time series. The model was made famous in Chris Sims’s paper in 1980 for macro-economic forecasts. The term auto regressive is used due to the fact that the variables are regressed on their own past values and the term vector is used due to the fact that we are dealing with a vector of two or more variables.

The VAR model has proven to be especially useful for describing the dynamic behavior of economic and financial time series, and for forecasting. It often provides superior forecasts to those from univariate time series models and elaborate theory-based simultaneous equations models.

The following sections describe the analysis of covariance stationary multivariate time series using VAR models.

Let $\mathbf{Y}_t = (y_{1t}, y_{2t}, \ldots, y_{nt})'$ denote an (n×1) random vector of time series variables.

The basic p-lag vector autoregressive (VAR (p)) model has the form (Hamilton, 1994)

$$\mathbf{Y_t} = \mathbf{c} + \Pi_1\mathbf{Y}_{t-1}+\Pi_2\mathbf{Y}_{t-2}+\cdots+\Pi_p\mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t,\ t = 1, \ldots, T \qquad (3.1)$$

where Π_i’s are (n×n) fixed coefficient matrices, $\mathbf{c} = (c_1, \ldots .c_n)'$ is a fixed (nx1) vector of intercept terms allowing for the possibility of a non zero mean $E(\mathbf{Y}_t)$. Finally $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{nt})'$ is an (n×1) unobservable zero-mean white noise vector process (serially uncorrelated or independent) with time invariant covariance matrix Σ, that is $E(\boldsymbol{\varepsilon}_t)=0$, $E(\boldsymbol{\varepsilon}_t\ \boldsymbol{\varepsilon}_s') = 0$ for $s \neq t$. The covariance matrix Σ is assumed to be non singular.

In lag operator notation, the VAR (p) is written as

$$\Pi(L)\ \mathbf{Y_t} = \mathbf{c} + \boldsymbol{\varepsilon}_t \qquad (3.2)$$

where $\Pi(L) = \mathbf{I}_n - \Pi_{\mathbf{1}} L^1 - \ldots - \Pi_{\mathbf{p}} L^{\mathbf{p}}$ and $L^p \mathbf{Y_t} = \mathbf{Y_{t\text{-}p}}$

The VAR(p) is stable if the roots of

$$\det\left(I_n - \Pi_1 z - \cdot\cdot\cdot - \Pi_p z^p\right) = 0 \qquad (3.3)$$

lie outside the complex unit circle (have modulus greater than one) for complex z, $|z|<1$, or, equivalently, if the eigenvalues of the companion matrix

$$F = \begin{pmatrix} 1 & 2 & \cdots & p \\ I_n & 0 & \cdots & 0 \\ 0 & \ddots & 0 & \vdots \\ 0 & 0 & I_n & 0 \end{pmatrix} \qquad (3.4)$$

have modulus less than one. Assuming that the process has been initialized in the infinite past, then a stable VAR (p) process is stationary with time invariant means, variances, and autocovariances.
If $\mathbf{Y}_t$ in (3.1) is covariance stationary, then the unconditional mean is given by

$$\boldsymbol{\mu} = (I_n - \Pi_1 - \cdot\cdot\cdot - \Pi_p)^{-1} c \qquad (3.5)$$

The mean-adjusted form of the VAR (p) is then

$$\mathbf{Y}_t - \boldsymbol{\mu} = \Pi_1 (\mathbf{Y}_{t-1} - \boldsymbol{\mu}) + \Pi_2 (\mathbf{Y}_{t-2} - \boldsymbol{\mu}) + \cdot\cdot\cdot + \Pi_p (\mathbf{Y}_{t-p} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon}_t \qquad (3.6)$$

3.3.2 Stationary Processes

A stochastic process $\mathbf{Y}_t$ is weak stationary if its first and second moments are time invariant. In other words, a stochastic process is stationary if

$$E(\mathbf{Y}_t) = \boldsymbol{\mu} \text{ for all t} \quad (3.7)$$

and

$$E[(\mathbf{Y}_t - \boldsymbol{\mu})(\mathbf{Y}_{t-h} - \boldsymbol{\mu})'] = \Gamma_y(h) = \Gamma_y(-h)' \quad \text{for all t and h = 0,1,2, …} \quad (3.8)$$

Condition (3.7) means that all $\mathbf{Y}_t$ have the same finite mean vector $\boldsymbol{\mu}$ and (3.8) requires that the autocovariances of the process do not depend on t but just on the time period h the two vectors $\mathbf{Y}_t$ and $\mathbf{Y}_{t-h}$ are apart. Note that all quantities are assumed to be finite. For instance $\boldsymbol{\mu}$, is a vector of finite mean terms and $\Gamma_y(h)$ is a matrix of finite covariances. A stable VAR(p) process $\mathbf{Y}_t$, t=1,2, … , is stationary. Because stability implies stationarity, the stability condition (3.3) is often referred to as stationarity condition in the time series literature. The converse is not true. In other words, an unstable process is not necessarily nonstationary (Lütkepohl, 2005).

3.3.3 Unit Root Tests

To test the presence of stationary formally we can apply the unit root test. There are various kinds of unit root tests. The most popular of these testes are the Augmented Dickey – Fuller (ADF) test and the Phillips – Perron (PP) test.

For testing non-stationarity (i.e., to test for the existence of unit roots) of the variables used in this study Augmented Dickey-Fuller (ADF) (Dickey and Fuller, 1979) and Phillips-Perron (PP) (Hamilton, 1994) tests are used. For the derivation of Dickey-Fuller test for an arbitrary series y_t, consider the model below

$$y_t = \beta_0 + \beta_1 t + u_t \quad (3.9)$$

$$u_t = \alpha u_{t-1} + \varepsilon_t \quad (3.10)$$

where t is time and ε_t is a zero mean covariance stationary process. Using (3.10) the reduced form of (3.9) can be written as:

$$y_t = \delta_0 + \delta_1 t + \alpha y_{t-1} + \varepsilon_t \quad (3.9a)$$

where $\delta_0 = \beta_0(1-\alpha)$ and $\delta_1 = \beta_1(1-\alpha)$.

The Dickey-Fuller test tests the null hypothesis of unit root ($\alpha = 1$) against the alternative hypothesis of $\alpha < 1$ in (3.9a).

Dickey-Fuller test is based on the assumptions that residuals are white noise and the data generating process is autoregressive of order one (AR(1)). It may lead to wrong conclusions if the data generating process is autoregressive of higher order or if the errors are autocorrelated. The Augmented Dickey-Fuller test (Dickey and Fuller, 1981) includes additional higher order lagged differences to the Dickey-Fuller test model. Inclusion of the lagged difference term allows autoregressive moving average (ARMA) errors (Maddala, 1992). ADF test is specified based on the model:

$$y_t = \delta_0 + \delta_1 t + \alpha y_{t-1} + \sum_{i=1}^{n} \gamma_i \Delta y_{t-1} + \varepsilon_{1t} \quad (3.11)$$

where Δ is a first difference operator and *n* is the lag length in the model. The lag length is determined based on Akaike Information Criterion.

ADF test is biased towards accepting the null hypothesis of unit root in the series (Badawi, 2007; Kim, 1990 cited in Maddala, 1992). The Phillips-Perron test is specified based on the model:

$$y_t = \delta_0 + \delta_1\left(t - \frac{T}{2}\right) + \alpha y_{t-1} + \sum_{i=1}^{m} \gamma_i \Delta y_{t-1} + \varepsilon_{2t}$$

where *T* is the number of observations and *m* is lag length. The lag length is determined based on Newey and West (1987) suggestions.

Although differencing may transform non-stationary series into stationary ones, it leads to the loss of important long run information about the variables. To deal with this problem of differencing, Engle and Granger (1987) recommend cointegration.

3.3.4 Computation of Autocovariance and Autocorrelations of Stable VAR Processes

3.3.4.1 Autocovariance of a Stable VAR (p) Processes

Autocovariance is the covariance between two observations separated by k units of time in a time series.

For a higher order stable VAR (p) process,

$\mathbf{Y_t} - \boldsymbol{\mu} = \prod_1(\mathbf{Y_{t-1}} - \boldsymbol{\mu}) +$

$$\mathbf{Y_t} - \boldsymbol{\mu} = \prod_1(\mathbf{Y_{t-1}} - \boldsymbol{\mu}) + \dots + \prod_p(\mathbf{Y}_{t-p} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon}_t \quad (3.12)$$

post multiplying with $(\mathbf{Y}_{t-h} - \boldsymbol{\mu})$ and taking expectations gives

$$E[(\mathbf{Y}_t - \boldsymbol{\mu})(\mathbf{Y}_{t-h} - \boldsymbol{\mu})'] = \prod_1 E[(\mathbf{Y}_{t-1} - \boldsymbol{\mu})(\mathbf{Y}_{t-h} - \boldsymbol{\mu})'] + \dots + \prod_p E[(\mathbf{Y}_{t-p} - \boldsymbol{\mu})(\mathbf{y}_{t-h} - \boldsymbol{\mu})'] + E[\boldsymbol{\varepsilon}_t(\mathbf{Y}_{t-h} - \boldsymbol{\mu})']$$

Thus for h=0

$$\Gamma_y(0) = \prod_1 \Gamma_y(-1) + \dots + \prod_p\Gamma_y(-p) + \Sigma_\varepsilon$$

$$= \prod_1 \Gamma_y(1)' + \dots + \prod_p\Gamma_y(p)' + \Sigma_\varepsilon \quad (3.13)$$

and for h > 0,

$$\Gamma_y(h) = \prod_1 \Gamma_y(h-1) + \dots + \prod_p\Gamma_y(h-p) \quad (3.14)$$

These equations are usually referred to as Yule-Walker equations.

3.3.4.2 Autocorrelations of Stable VAR (p) Processes

Autocorrelation is the correlation of the observations in a time series, usually expressed as a function of the time lag between observations.

Because the autocovariance depend on the unit measurement used for the variables of the system, they are sometimes difficult to interpret. Therefore, the autocorrelations

$$R_y(h) = D^{-1}\Gamma_y(h)D^{-1} \quad (3.15)$$

are usually more convenient to work with as they are scale invariant measures of the linear dependencies among the variables of the system. Here D is a diagonal matrix with the standard deviations of the components of $\mathbf{Y}_t$ on the main diagonal. That is, the diagonal elements of D are the square roots of the diagonal elements of $\Gamma_y(0)$. Denoting the covariance between $y_{i,t}$ and $y_{j,t-h}$ by $\gamma_{ij}(h)$ (i.e., $\gamma_{ij}(h)$ is the ij-th element of $\Gamma_y(h)$) the diagonal elements $\gamma_{11}(0), \dots, \gamma_{nn}(0)$ of $\Gamma_y(0)$ are the variances of $y_{1t}, \dots y_{nt}$. Thus

$$D^{-1} = \begin{pmatrix} 1/\sqrt{\gamma_{11}(0)} & & 0 \\ & \ddots & \\ 0 & & 1/\sqrt{\gamma_{nn}(0)} \end{pmatrix}$$

And the correlation between $y_{i,t}$ and $y_{j,t-h}$ is

$$\rho_{ij}(h) = \frac{\gamma_{ij}(h)}{\sqrt{\gamma_{ii}(0)}\sqrt{\gamma_{jj}(0)}} \tag{3.16}$$

Which is just the ij-th element of Ry(h) given in equation 3.15 above

3.3.5 Structural Vector Autoregressive (SVAR) Measures

The general VAR(p) model has many parameters, and they may be difficult to interpret due to complex interactions and feedback between the variables in the model. As a result, the dynamic properties of a VAR(p) are often summarized using various types of structural analysis. The three main types of structural- analysis summaries are (1) Granger causality tests; (2) impulse response functions; and (3) forecast error variance decompositions. The following sections give brief descriptions of these summary measures, (Lütkepohl, 2005).

3.3.5.1 Granger Causality

The structure of the VAR model provides information about a variable's or a group of variables' forecasting ability for other variables. The following intuitive notion of a variable's forecasting ability is due to Granger (1969). If a variable, or group of variables, $\mathbf{Y}_{1t}$ is found to be helpful for predicting another variable, or group of variables, $\mathbf{Y}_{2t}$ then $\mathbf{Y}_{1t}$ is said to Granger-cause $\mathbf{Y}_{2t}$; otherwise it is said to fail to Granger-cause Y_{2t}. Formally, $\mathbf{Y}_{1t}$ fails to Granger-cause $\mathbf{Y}_{2t}$ if for all $s > 0$ the MSE of a forecast of $\mathbf{Y}_{2,t+s}$ based on ($\mathbf{Y}_{2,t}$, $\mathbf{Y}_{2,t-1}$, . . .) is the same as the MSE of a forecast of $\mathbf{Y}_{2,t+s}$ based on ($\mathbf{Y}_{2,t}$, $\mathbf{Y}_{2,t-1}$, . . .) and ($\mathbf{Y}_{1,t}$, $\mathbf{Y}_{1,t-1}$, . . .). Clearly, the notion of Granger causality does not imply true causality. It only implies forecasting ability. If $\mathbf{Y}_{1t}$ causes $\mathbf{Y}_{2t}$ and $\mathbf{Y}_{2t}$ also causes $\mathbf{Y}_{1t}$ the process $(\mathbf{Y}_{1t}', \mathbf{Y}_{2t}')'$ is called a feedback system.

In a bivariate VAR(p) model for $\mathbf{Y}_t = (\mathbf{Y}_{1t}, \mathbf{Y}_{2t})'$, $\mathbf{Y}_{2t}$ fails to Granger-cause $\mathbf{Y}_{1t}$ if all of the p VAR coefficient matrices $\Pi_1, \ldots, \Pi_p$ are lower triangular.

That is, the VAR(p) model has the form

$$\begin{pmatrix} \mathbf{Y}_{1t} \\ \mathbf{Y}_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \pi_{11}^1 & 0 \\ \pi_{21}^1 & \pi_{22}^1 \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \ldots + \begin{pmatrix} \pi_{11}^p & 0 \\ \pi_{21}^p & \pi_{22}^p \end{pmatrix} \begin{pmatrix} y_{1t-p} \\ y_{2t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

So that all of the coefficients on lagged values of $\mathbf{Y}_{2t}$ are zero in the equation for $\mathbf{Y}_{1t}$. Similarly, $\mathbf{Y}_{1t}$ fails to Granger-cause $\mathbf{Y}_{2t}$ if all of the coefficients on lagged values of $\mathbf{Y}_{1t}$ are zero in the equation for $\mathbf{Y}_{2t}$. Granger non-causality may be tested using the Wald statistic.

3.3.5.2 Impulse Response Functions

Any covariance stationary VAR (p) process has a Wold representation of the form

$$Y_t = \mu + \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \Psi_2 \varepsilon_{t-2} + \ldots \qquad (3.17)$$

where the (n × n) moving average matrices Ψ_s are determined recursively using

$$\Psi_s = \sum_{j=1}^{p-1} \Psi_{s-j} \Pi_j$$

It is tempting to interpret the (i, j)-th element, Ψ_{ij}^s, of the matrix Ψ_s as the dynamic multiplier or impulse response i.e Ψ_{ij}^s represent the effects of unit shocks in the variables of the system.

$$\frac{\partial y_{i,t+s}}{\partial \varepsilon_{j,t}} = \frac{\partial y_{i,t}}{\partial \varepsilon_{j,t-s}} = \psi_{ij}^s \qquad \text{i, j = 1, …, n}$$

However, this interpretation is only possible if $\text{var}(\boldsymbol{\varepsilon}_t) = \Sigma$ is a diagonal matrix so that the elements of $\boldsymbol{\varepsilon}_t$ are uncorrelated.

One way to make the errors uncorrelated is to follow Sims (1980) and estimate the triangular structural VAR(p) model defined by

$$y_{1t} = c_1 + \gamma'_{11} Y_{t-1} + \cdots + \gamma'_{1p} Y_{t-p} + \eta_{1t} \tag{3.18}$$

$$y_{2t} = c_2 + \beta_{21} y_{1t} + \gamma'_{21} Y_{t-1} + \cdots + \gamma'_{2p} Y_{t-p} + \eta_{2t}$$

$$y_{3t} = c_3 + \beta_{31} y_{1t} + \beta_{32} y_{2t} + \gamma'_{31} Y_{t-1} + \cdots + \gamma'_{3} p Y_{t-p} + \eta_{3t}$$

$$\vdots$$

$$y_{nt} = c_n + \beta_{n1} y_{1t} + \cdots + \beta_{n,n-1} y_{n-1,t} + \gamma'_{n1} Y_{t-1} + \cdots + \gamma'_{np} Y_{t-p} + \eta_{nt}$$

In matrix form, the triangular structural VAR(p) model is

$$B\mathbf{Y}_t = \mathbf{C} + \Gamma_1 Y_{t-1} + \Gamma_2 \mathbf{Y}_{t-2} + \cdots + \Gamma_p \mathbf{Y}_{t-p} + \boldsymbol{\eta}_t \tag{3.19}$$

where

$$B = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\beta_{21} & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\beta_{n1} & -\beta_{n2} & \cdots & 1 \end{pmatrix} \tag{3.20}$$

$\mathbf{C} = [c_1\ c_2 \ldots\ldots c_n]'$ and $\boldsymbol{\Gamma}_i = [\gamma'_{1i}\ \gamma'_{2i} \ldots\ldots \gamma'_{ni}]'$ for i= 1,2, …,p

is a lower triangular matrix with 1's along the diagonal. The algebra of least squares will ensure that the estimated covariance matrix of the error vector $\boldsymbol{\eta}_t$ is diagonal. The uncorrelated/orthogonal errors $\boldsymbol{\eta}_t$ are referred to as structural errors.

The triangular structural model (3.18) imposes the recursive causal ordering

$$y_1 \rightarrow y_2 \rightarrow \cdots \rightarrow y_n \tag{3.21}$$

The ordering (3.21) means that the contemporaneous values of the variables to the left of the arrow → affect the contemporaneous values of the variables to the right of the arrow but not vice-versa. These contemporaneous effects are captured by the coefficients β_{ij} in (3.18).

For a VAR(p) with n variables there are n! possible recursive causal orderings. Which ordering to use in practice depends on the context and whether prior theory can be used to justify a particular ordering. Results from alternative orderings can always be compared to determine the sensitivity of results to the imposed ordering.

Once a recursive ordering has been established, the Wold representation of $\mathbf{Y}_t$ based on the orthogonal errors η_t is given by

$$\mathbf{Y}_t = \boldsymbol{\mu} + \Theta_0 \boldsymbol{\eta}_t + \Theta 1 \boldsymbol{\eta}_{t-1} + \Theta_2 \boldsymbol{\eta}_{t-2} + \cdot \cdot \cdot \qquad (3.22)$$

where $\Theta_0 = B^{-1}$ is a lower triangular matrix. The impulse responses to the orthogonal shocks $\boldsymbol{\eta}_{jt}$ are

$$\frac{\partial y_{i,t+s}}{\partial \eta_{j,t}} = \frac{\partial y_{it}}{\partial \eta_{j,t-s}} = \theta_{ij}^{s}, \ \text{i,j=1,…n; s>0} \qquad (3.23)$$

where θ_{ij}^{s} is the (i, j) th element of Θs. A plot of θ_{ij}^{s} against s is called the orthogonal impulse response function (IRF) of y_i with respect to η_j. With n variables there are n^2 possible impulse response functions.

In practice, the orthogonal IRF (3.23) based on the triangular VAR(p) (3.18) may be computed directly from the parameters of the non triangular VAR(p) (3.1) as follows. First, decompose the residual covariance matrix Σ as

$$\Sigma = ADA'$$

where A is an invertible lower triangular matrix with 1's along the diagonal and D is a diagonal matrix with positive diagonal elements. Next, define the structural errors as

$$\boldsymbol{\eta}_t = A^{-1}\boldsymbol{\varepsilon}_t$$

These structural errors are orthogonal by construction since $\text{var}(\boldsymbol{\eta}_t) = A^{-1}\Sigma A^{-1'} = A^{-1}ADA'A^{-1'} = D$.

Finally, re-express the Wold representation(3.17) as

$$\mathbf{Y}_t = \boldsymbol{\mu} + AA^{-1}\boldsymbol{\varepsilon}_t + \Psi_1 AA^{-1}\boldsymbol{\varepsilon}_{t-1} + \Psi_2 AA^{-1}\boldsymbol{\varepsilon}_{t-2} + \cdots$$

$$= \boldsymbol{\mu} + \Theta_0\boldsymbol{\eta}_t + \Theta_1\boldsymbol{\eta}_{t-1} + \Theta_2\boldsymbol{\eta}_{t-2} + \cdots$$

where $\Theta_j = \Psi_j A$. Notice that the structural B matrix in (3.19) is equal to A^{-1}.

3.3.5.3 Forecast Error Variance Decompositions

The forecast error variance decomposition (FEVD) answers the question: what portion of the variance of the forecast error in predicting $y_{i,T+h}$ is due to the structural shock $\boldsymbol{\eta}_j$? Using the orthogonal shocks $\boldsymbol{\eta}_t$ the h-step ahead forecast error vector, with known VAR coefficients, may be expressed as

$$Y_{T+h} - Y_{T+h|T} = \sum_{s=0}^{h-1} \Theta_s \eta_{T+h-s}$$

where $Y_{T+h|T}$ is h-step forecasts based on information available at time T.

For a particular variable $y_{i,T+h}$, this forecast error has the form

$$y_{i,T+h} - y_{i,T+h|T} = \sum_{s=0}^{h-1} \theta_{i1}^{s} \eta_{1,T+h-s} + \ldots + \sum_{s=0}^{h-1} \theta_{in}^{s} \eta_{n,T+h-s}$$

Since the structural errors are orthogonal, the variance of the h-step forecast error is

$$\text{var}(y_{i,T+h} - y_{i,T+h|T}) = \sigma_{\eta 1}^2 \sum_{s=0}^{h-1} \left(\theta_{i1}^{s}\right)^2 + \ldots + \sigma_{\eta n}^2 \sum_{s=0}^{h-1} \left(\theta_{in}^{s}\right)^2$$

where $\sigma_{\eta j}^2 = \text{var}(\eta_{jt})$. The portion of $\text{var}(y_{i,T+h} - y_{i,T+h|T})$ due to shock η_j is then

$$F\,EV\,D_{i,j}(h) = \frac{\sigma_{\eta j}^2 \sum_{s=0}^{h-1} \left(\theta_{ij}^{s}\right)^2}{\sigma_{\eta 1}^2 \sum_{s=0}^{h-1} \left(\theta_{i1}^{s}\right) + \ldots + \sigma_{\eta n}^2 \sum_{s=0}^{h-1} \left(\theta_{in}^{s}\right)^2}, \quad i,j=1,\ldots,n \qquad (3.24)$$

In a VAR with n variables there will be n^2 F $EVD_{i,j}(h)$ values. It must be kept in mind that the FEVD in (3.24) depends on the recursive causal ordering used to identify the structural shocks η_t and is not unique. That is, different causal orderings will produce different FEVD values.

3.3.6 Vector Error Correction and Cointegration Theory

3.3.6.1 VEC Models

The fact that many time series contain a unit root has spurred the development of the theory of non-stationary time series analysis. Engle and Granger (1987) pointed out that a linear combination of two or more non-stationary series may be stationary. If such a stationary, or I(0), linear combination exists, the non-stationary (with a unit root), time series are said to be cointegrated. The linear combination which is stationary is called the cointegrating equation and may be interpreted as a long-run equilibrium relationship between the variables. For example, in income consumption analysis, consumption and income are likely to be cointegrated. If they were not, then in the long-run consumption might drift above or below income, so that consumers were irrationally spending or piling up savings. A vector error correction (VEC) model is a restricted VAR that has cointegration restrictions built in to the specification, so that it is designed for use with non-stationary series that are known to be cointegrated. The VEC specification restricts the long run behavior of the endogenous variables to converge to their cointegrating relationships while allowing a wide range of short-run dynamics. The cointegration term is known as the error correction term since the deviation from long run equilibrium is corrected gradually through a series of partial short run adjustments.

3.3.6.2 Testing for Cointegration

Given a group of non-stationary series, we may be interested in determining whether the series are cointegrated, and if they are, identify the cointegration (long-run equilibrium) relationships. We can interpret the long run paths of cointegrating variables as interdependent. Application of cointegration tests in estimation are analyzed by Johanson and Jusselius (1990). VAR-based cointegration tests using the methodology developed by Johansen (1988) is the most common method. Jhoanson's method is to test the restrictions imposed by cointegration on the unrestricted VAR involving the series. It applies the maximum likelihood method to determine the presence of cointegrating vectors in non-stationary time series. The trace tests and eigen value tests are used to determine the number of cointegrating vectors. This implies a stationary long- run equilibrium relationship between the variables. The maximum lag length of the VAR model which is used in Johanson procedure is determined by the Likelihood Ratio (LR) statistics.

Consider a VAR of order p

$Y_t = A_1\mathbf{Y}_{t-1} + \ldots + Ap\mathbf{Y}_{t-p} + \varepsilon_t$ (3.25)

where $\mathbf{Y}_t$ is an n-vector of non-stationary I(1) variables, and ε_t is a vector of innovations. We can rewrite the VAR as:

$$\Delta\mathbf{Y}_t = \prod\mathbf{Y}_{t-1} + \sum_{i=1}^{p-1}\Gamma_i\Delta Y_{t-1} + \boldsymbol{\varepsilon}_t \quad (3.26)$$

Where $\prod = \sum_{i=1}^{p} A_i - I$ = -(I-A_1-…-A_p)= $\theta\phi'$, $\Gamma_i = -\sum_{j=i+1}^{p} A_j$ = and Δ is the difference operator.

Granger's representation theorem asserts that if the coefficient (parameters) matrix $\prod$ has reduced rank r < n, then there exist nxr matrices Θ and Φ each with rank r such that $\prod = \Theta\Phi'$ and $\Phi'\mathbf{Y}_t$ is stationary. The letter r denotes the number of cointegrating relations (the cointegrating rank) and each column of Φ is the cointegrating vector. The elements of Θ are known as the adjustment parameters in the vector error correction model. Johanson's method is to estimate the $\prod$ matrix in an restricted form, then test whether we can reject the restrictions implied by the reduced rank of $\prod$.

If we have n endogenous variables, each of which has one unit root, there can be from zero to n-1 linearly independent, cointegrating relations. If there are no cointegrating relations, standard time series analysis such as the (unrestricted) VAR may be applied to the first differences of the data. Since there are n separate integrated elements driving the series, levels of the series do not appear in the VAR in this case. Conversely, if there is one cointegrating equation in the system, then a single linear combination of the levels of the endogenous series $\Phi\mathbf{Y}_{t-1}$, should be added to each equation in the VAR.

Each column of the Φ matrix gives an estimate of a cointegrating vector. The cointegrating vector is not identified unless we impose some arbitrary normalization. We can adopt the normalization so that the r cointegrating relations are solved for the first r variables in the Y_t vector as a function of the remaining n-r variables.

When multiplied by a coefficient for an equation, the resulting term $\Theta\Phi' Y_{t-1}$, is referred to as an error correction term. If there are additional cointegrating equations, each will contribute an additional error correction term involving a different linear combination of the levels of the series.

The null hypothesis of at most r cointegrating vectors against a general alternative hypothesis of more than r cointegrating vectors is tested by trace Statistics.

The trace statistic is given by $\lambda_{trace}(r) = -T\sum_{i=r+1}^{n}\ln\left(1-\tilde{\lambda}_i\right)$ (3.27)

where, T is the number of observations and $\tilde{\lambda}_i$ is the eigenvalues.

The null hypothesis of r cointegrating vector against the alternative of r+1 is tested by Maximum Eigen value statistic

Maximum Eigen Value is given by $\lambda_{\max}(r, r+1) = -T\ln\left(1-\tilde{\lambda}_{r+1}\right)$ (3.28)

3.7 VAR Order Selection and Checking Adequacy of the Model

3.7.1 Lag order Selection

The lag length for the VAR (p) model may be determined using model selection criteria. The general approach is to fit VAR (p) models with orders p = 0, ... , p_{max} and choose the value of p which minimizes some model selection criteria. Model selection criteria for VAR(p) models have the form

$$IC(p) = \ln|\tilde{\Sigma}(p)| + c_T \cdot \varphi(n, p) \quad (3.29)$$

where $\tilde{\Sigma}(p) = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t'$ is the residual covariance matrix without a degrees of freedom correction from a VAR(p) model, c_T is a sequence indexed by the sample size T, and φ(n, p) is a penalty function which penalizes large VAR(p) models.

The three most common information criteria are the Akaike (AIC), Schwarz-Bayesian (BIC) and Hannan-Quinn (HQ) defined by

$$AIC(p) = \ln|\tilde{\Sigma}(p)| + \frac{2}{T} pn^2$$

$$BIC(p) = \ln|\tilde{\Sigma}(p)| + \frac{\ln T}{T} pn^2 \text{ and}$$

$$HQ(p) = \ln|\tilde{\Sigma}(p)| + \frac{2\ln\ln T}{T} pn^2 \qquad (3.30)$$

The AIC criterion asymptotically overestimates the order with positive probability, whereas the BIC and HQ criteria estimate the order consistently under fairly general conditions if the true order p is less than or equal to p_{max}.

3.7.2 Checking the Whiteness of the Residuals

It is assumed that ε_t is an n-dimensional white noise process with nonsingular covariance matrix Σ. For instance, ε_t may represent the residuals of a VAR(p) process. The lagrange multiplier tests is a popular statistic for checking the overall significance of the residual autocorrelations.

In Lagrange multiplier tests we wish to test

H_0: $D_1 = \cdot\cdot\cdot = D h = 0$ against

H_1 : $D_j = 0$ for at least one $j \in \{1, . . . , h\}$. (3.31)

Where the error vector, $\varepsilon_t = D_1 \varepsilon_{t-1} + \ldots + D_h \varepsilon_{t-h} + \mathbf{v}_t$, where $\mathbf{v}_t$ is white noise. It is equal to ε_t if there is no residual autocorrelation.

3.7.3 Testing for Nonnormality

A stationary, stable VAR(p) process is Gaussian (normally distributed) if and only if the white noise process ε_t is Gaussian. Therefore, the normality of the y_t' s may be checked via the ε_t's. In practice, the ε_t's are replaced by estimated residuals.

Normality of the underlying data generating process is needed, for instance, in setting up forecast intervals. Nonnormal residuals can also indicate more generally that the model is not a good representation of the data generation process. Therefore, testing this distributional assumption is desirable.

Lütkepohl(1993) suggests using the multivariate generalization of the Jarque-Bera test. Jarque & Bera (1987) established a test statistic to test for the normality of observations. This statistic is based on the skewness and kurtosis properties of the residuals, (3rd & 4th moments). In this study it is used to test the null hypothesis that the disturbances are normally distributed.

The Jarque-Bera test statistic is given by the formula

$$\text{JB} = \frac{T}{6}\left[s^2 + \frac{(k-3)^2}{4}\right]$$

where s is a measure of skewness, k is a measure of kurtosis and n is the sample size. Under H_0 JB has a χ^2 distribution with 2 degrees of freedom asymptotically, and the null hypothesis is rejected if the computed value exceeds a χ^2 critical value (small p-value).

CHAPTER FOUR
RESULTS AND DISCUSSION

4.1 Introduction

The study is based on thirty two years consecutive summer seasons (Jun, July, August and September) total summer rainfall in mm, average temperature in degree centigrade (°C) and sorghum yield of medium maturing Gambella-1107 variety in quintal per hectare(Q/h) which is used as a control variety in Melkassa Research Center of EIAR. The time period included is from 1978 to 2009 G.C. The raw data are given in Table 1 of Appendix B.

4.2 Descriptive analysis

From the correlation matrix presented in Table 4.1 we can observe that there exists high positive correlation between rainfall and yield. The correlation between temperature and rainfall is significant at 0.01 level of significance but it is negative association, and the same is true for temperature and yield. In scientific terms, such a relationship holds true, as for instance, the decline in rainfall could be associated with an increasing air temperature or heat load, while at the same time an increase in temperature beyond a given optimal value could translate into a declining sorghum yield.

Table 4.1 Correlation Matrix

	Rainfall	Yield	Temperature
Rainfall	1	.805**	-.452**
Yield	.805**	1	-.474**
Temperature	-.452**	-.474**	1

**. Correlation is significant at the 0.01 level (2-tailed).

The amount and distribution of annual and seasonal total rainfall is critical rainfall feature that indicates useful information on temporal rainfall variability over an area. The total amount of summer rainfall ranged from 320 to 690 mm (Table 4.2) which indicate a high interannual variability of the summer total rainfall. The standard deviation (s. d) is 93.92 mm indicating higher variability of the summer rainfall total. Also yield exhibits the highest variability from year to year, with a range of 3 Q/h to 70 Q/h and s.d of 20.08 Q/h, where as, s.d of temperature is 0.75°C that is the year to year variability in average summer temperature is not large. Figure.1 of Appendix C shows the plots of each series. From the plots we can observe that all the series are non stationary as they appear to have no fixed level and in addition recent years have more variability than the previous ones.

Table 4.2 Summary Result of the series

Descriptive statistics	Rainfall (mm)	Temperature (°C)	Yield (Q/ha)
Mean	509.59	22.02	31.68
Median	519.55	21.99	28.51
Maximum	690	23	70
Minimum	320	20.6	3
Standard deviation	93.92	0.75	20.08
Skewness	-0.16	-0.25	0.34
Kurtosis	-0.47	-1.05	-0.88

4.3 Nature of the data

4.3.1 Unit Root Tests

Before we apply the different techniques of multivariate time series analysis, we need to check for the stationarity of each variable under study since many of the methods assume that the data is stationary with respect to the mean and variances. The ADF and PP tests differ mainly in how they treat serial correlation in the test regressions. ADF tests use a parametric autoregressive structure to capture serial correlation while PP tests use non-parametric corrections based on estimates of the long-run variance of the differences Δy_t. The hypothesis is that H_0: The series is not stationary (There is a unit root) Versus H_1: The series is stationary. The following table shows the computed values of the two tests for each of the variables.

Table 4.3 ADF and PP Unit Root Test Results for original Series

Variable	Test Statistics		P-Value		Decision
	ADF	PP	ADF	PP	
Rainfall	-3.407	-3.315	0.0505	0.0638	Not stationary
Temperature	-1.641	-1.875	0.7761	0.6678	Not stationary
Yield	-3.603	-3.637	0.0296	0.0269	Not stationary

As we can clearly see from the above table (large p-value indicate non rejection of H_0) all variables do not satisfy the stationarity assumption. But after order one differencing, all the series become stationary.

The next table shows the results of the ADF and PP test results after differencing the series. Since P-Value of all series in column four of Table 4.4 is small (<0.05) we reject the null hypothesis and conclude that the series is stationary. The time plots of the series after differencing are given in Figure.2 of Appendix C, The plots show the difference between the differenced and original series.

Table 4.4 ADF and PP unit Root test results for Differenced series

variable	Order of differencing	Test Statistics		P-Value		Decision
		ADF	PP	ADF	PP	
Rain	1	-5.954	-6.017	0.0000	0.0000	Stationary
Temperature	1	-4.018	-3.887	0.0083	0.0126	Stationary
Yield	1	-7.229	-7.681	0.0000	0.0000	Stationary

4.3.2 Auto Correlation and Partial Autocorrelation Functions

The plot of the autocorrelation versus the lag k is, called the autocorrelation function of the stochastic process. Partial autocorrelation at lag k is the correlation between two observations separated by k time interval with the effects of the intervening observations removed. A plot of the partial autocorrelation coefficient against the lag k is known as the partial autocorrelation function.

Autocorrelation plots are a commonly-used tool for checking randomness/nonstationary in a time series. For our study the ACF and PACF are given on Figure.3 of Appendix C. The plots suggest that our series are all nonrandom/ nonstationary because all are statistically significantly different from zero at lag one and hence we can apply time series technique to analyze the data after differencing with appropriate order.

4.3.3 Cointegration Rank Test

In this section the cointegrating rank (rank of matrix Π) is estimated using Johansen's methodology. If the rank is zero, there is no cointegrating relationship. If the rank is one there is one, if it is two there are two and so on. According to the results of trace statistics, we cannot reject the null hypothesis of no cointegrating equations because the trace statistic at r =0 of 42.3302 is less than its critical value of 42.44, that is there is no cointegrating relationship [$rank(\Pi) = 0$]. In other words, there is no need for VEC model, fitting VAR model to our data will be adequate. Table 4.5 presents the results of Johansen test.

Table 4.5 Johansen test results

Maximum rank	LL	Eigen value	Trace statistics	5% critical value
0	-326.551	.	42.3302*	42.44
1	-315.778	0.50094	20.7844	25.32
2	-307.089	0.42913	3.406	12.25

4.4 VAR Order Selecting and Estimating Model Parameters

4.4.1 Model Selection

The VAR model in this study is specified as a three variable system for a sample period from 1978 to 2009.

The general form of the VAR model is

$$Y_t = \begin{bmatrix} Y_{1t} \\ Y_{2t} \\ Y_{3t} \end{bmatrix} = \begin{bmatrix} R \\ T \\ Y \end{bmatrix} = \begin{bmatrix} Rainfall \\ Temperature \\ Yield \end{bmatrix} = \begin{bmatrix} c_{1t} \\ c_{2t} \\ c_{3t} \end{bmatrix} + \sum_{i=1}^{i=p} \Pi_i \begin{bmatrix} R_{t-i} \\ T_{t-i} \\ Y_{t-i} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix}, \quad t = 1, \ldots, 32$$

In the pre-lag order estimation technique in STATA using four maximum number of lags (i.e $p_{max} = 4$), the suggested model is VAR (1) in all model selection criteria since it has the minimum AIC, BIC and HQ. Table 4.6 reports lag-order selection statistics,

Table 4.6 Pre-lag order selection results from STATA

lag	p	AIC	HQIC	SBIC
0		21.841	21.885	21.9841
1	0.000	20.9892*	21.1638*	21.5602*
2	0.063	21.0531	21.3585	22.0522
3	0.228	21.2761	21.7125	22.7035
4	0.180	21.4676	22.0348	23.3231

The result of the estimated model is given by

$$\hat{Y}_t = \hat{C} + \hat{\Pi}_1 Y_{t-1}$$

The coefficient matrix Π_1 and vector of constants C were successfully estimated by the method of least squares using STATA software.

$$\begin{bmatrix} \Delta R_t \\ \Delta T_t \\ \Delta Y_t \end{bmatrix} = \begin{bmatrix} \Delta Y_{1t} \\ \Delta Y_{2t} \\ \Delta Y\,3t \end{bmatrix} = \begin{bmatrix} -0.0420 & 13.5609 & -0.7774 \\ -0.0011 & 0.2303 & 0.0048 \\ 0.0267 & -1.9887 & -0.4329 \end{bmatrix} \begin{bmatrix} \Delta Y_{1t-1} \\ \Delta Y_{2t-1} \\ \Delta Y_{3t-1} \end{bmatrix} + \begin{bmatrix} -4.5632 \\ 0.0436 \\ 0.3945 \end{bmatrix}$$

This can be explicitly written as:

ΔR_t = - 0.0420 ΔR_{t-1} + 13.5609 ΔT_{t-1} - 0.7774 ΔY_{t-1} - 4.5632

ΔT_t = - 0.0011 ΔR_{t-1} + 0.2303 ΔT_{t-1} + 0.0047 ΔY_{t-1} + 0.0436

ΔY_t = 0.0267 ΔR_{t-1} - 1.9887 ΔT_{t-1} - 0.4329 ΔY_{t-1} + 0.3945

Where Δ represent the first difference operator.

4.4 Model DIAGNOSTIC

In this section the goodness of fit of our VAR model is assessed. If we fit a VAR model and all of the assumptions are not met the inference we make using the model may be erroneous. Diagnostics help us check whether the assumptions of our model are met or not. As in most modeling situations the fit is assessed through the behavior of the residuals. Therefore it is mandatory to study the nature of the residuals before asserting about adequacy of the model.

4. 4.1Whiteness of Residuals

If a model is an adequate representation of the process that generated the time series, the residual should have no significant trend or pattern that means the residuals should be uncorrelated. Thus we test for autocorrelation of the error terms. We run Lagrange-multiplier (LM) test for autocorrelation in the residuals. Table 4.7 shows the results of this test:

Table 4.7 Results of Lagrange multiplier test for autocorrelation

lag	χ^2	df	prob > χ^2
1	13.5649	9	0.13867
2	14.3266	9	0.11117
H_0: no autocorrelation at lag order			

Large p-values (Prob>chi2) indicate that we can not reject H_0, in other words there is no autocorrelation among the residuals in our VAR(1) model. The graph of the ACF of the residuals in Appendix C of Figure.4 also supports this evidence because all values up to 15 lags are random and lie inside the 95% confidence bounds.

4.4.2 Testing for Normality of Residuals

Here, we employ Jarque-Berra test to check if the residuals in the VAR are normally distributed. Normality property is also needed for valid inference when performing hypotheses testing. In Jarque-Berra test the hypothesis is H_0: The residuals are normally distributed versus

H_1: The residuals are non-normally distributed and the null hypothesis is rejected if the computed value exceeds a χ^2 critical value (small p-value). Table 4.8 shows the results of this test:

Table 4.8 Result of Jarque-Bera test for normality

Equation	χ^2-Computed	df	P-value
Rainfall	1.592	2	0.45116
Temperature	0.586	2	0.74612
Yield	1.825	2	0.40159
H0: residuals are normally distributed			

The above result (large p-value) suggests the non rejection of the null hypothesis in every equation at the 5% significance level. Also the normal probability plots in Figure.5 of Appendix C indicate that the disturbances are normally distributed.

Since our model passes all diagnostic tests, we conclude that its results may be trusted, further report and discuss them, perform innovation accounting are possible.

4.5 Structural Analysis

4.5.1 Granger Causality

As presented in chapter three the dynamic properties of a VAR(p) are often summarized using various types of structural analysis. Variable y_1 is said to Granger cause variable y_2, if the lags of y_1 can improve a forecast for variable y_2. The following output is from STATA for pairwise granger causality test among the variables.

Table 4.9 Pairwise granger causality test result

Equation	Excluded	χ^2	df	p-value
Rainfall	Temperature	0.0772	1	0.781
Rainfall	Yield	0.35674	1	0.550
Temperature	Rainfall	0.56709	1	0.451
Temperature	Yield	0.61451	1	0.433
Yield	Rainfall	0.18483	1	0.667
Yield	Temperature	0.04265	1	0.836

A Wald test is commonly used to test for Granger causality. Each row of the above table reports a Wald test that the coefficients on the lags of the variable in the "excluded" column are zero in the equation for the variable in the "equation" column. None of the results shows any causality that is statistically significant at the 5% level (large p-value).

4.5.2 Impulse Response Function

The Impulse Response Function traces the response of an endogenous variable to a change in one of the innovations in the VAR system. Impulse response analysis is a standard tool for investigating the relations between the variables in a VAR model. Usually the response is portrayed graphically with horizon on the horizontal or X- axis and response on the vertical or Y-axis. It traces the effect of a one standard deviation shock to one of the innovations on current and future values of the dependent variables through the dynamic structure of the VAR.

We can use EViews software to compute impulse response functions for each innovation and endogenous variable pair. Tabel.2 of Appendix B is IRF Table. The graphs of these IRF for ten years period are given in Figure.7 of Appendix C. From the graph we can observe the following.

Over the ten years period considered, a shock in rainfall have significant impact on rainfall up to three years into the future and then the impact of the effect dies out quickly**.** Also rainfall shocks (within one standard deviation) have significant impact on temperature and yield up to three and four periods respectively. For example, in the second period a decrease of rainfall by one standard deviation will cause temperature to increase and yield to decrease. In the third period an increase of rainfall will cause temperature to decrease and yield to increase.

This supports the fact that a reduction in rainfall amount (which could be associated with the climate change) resulted in a rising temperature could be explained by the reduced evaporative cooling of the system, while the increase of rainfall could also be explained in terms of increased evaporative cooling of the system. Likewise, the increase in rainfall (particularly under low rainfall condition) could result in significant response of sorghum yield is self explanatory, because with no doubt, sorghum yield is a function of rainfall or water. The fact that rainfall shock has significant impact on rainfall shows the likelihood of past year and current rainfall resembling each other could be high or they are significantly autocorrelated.

Temperature shock have significant impact on temperature, rainfall and yield up to three horizons/years, in the first period if the innovation of temperature is shocked by an increase of one standard deviation, the level of rainfall will be decreased while yield is not affected. In the second period if temperature increases rainfall will decrease by and yield increase. See Figure.7 of Appendix C. This negative association between rainfall and temperature is explained by the level of evaporative cooling in the system, the non-significant impact of temperature shock on sorghum yield could be explained by the fact that, in the first period a shock in temperature within one standard deviation (increment), could mean that still the temperature amount within a range that sorghum crop could use in photosynthesis and for raising water productivity. Also temperature shock has significant impact on temperature means that last year temperature effect has not been totally removed for the next year.

Yield shock has a significant impact on yield up to four horizons. This is because of the fact that as we have seen above last year bad climate condition that may be draught or excess rainfall has impact on the following year and this indirectly affects the yield quality (quantity) as well.

4.5.3 Forecast Error Variance Decomposition

Variance decompositions offer a slightly different method for examining VAR system dynamics. They give the proportion of the movements in the dependent variables that are due to their 'own' shocks, versus shocks to the other variables. A shock to the *i* th variable will directly affect that variable of course, but it will also be transmitted to all of the other variables in the system through the dynamic structure of the VAR. The variance decomposition of a VAR model gives information about the relative importance of each of the random innovations in explaining each endogenous variable in the system. In practice, it is usually observed that own series shocks explain most of the (forecast) error variance of the series in a VAR. To some extent, impulse responses and variance decompositions offer very similar information.

Usually, we plot the decomposition of each forecast error variance as line graphs. The variance decomposition is displayed as separate line graphs with the Y- axis measures the relative importance of each innovation. Tabel.3 of Appendix B is FEVD Table. The graphs of the FEVD for ten years period are given in Figure.8 of Appendix C.

In the first round, 100 % change in rainfall resulted from the shock of its own innovation. In the second round, 98.4% change in rainfall level resulted from the shock to the rainfall innovation. This shock has also caused the levels of yield and temperature to change little; and the effect of the change in yield towards the change in rainfall is about 1.1%. However the change in temperature which resulted from the shock of rainfall is about 0.5%, then it remains the same after the third period: 98.04 % of change is resulted from rainfall, the effects of the change in temperature towards the change in rainfall is about 0.532 % and the remaining 1.419% change is yield.

For temperature, in the first round, 56% variation of temperature has resulted from the shock of its own innovation and 44% variation is resulted from the change in rainfall. However the change in yield resulted from the shock of temperature will not have any impact to the variation of temperature immediately. In the third period, the level of rainfall can explain about 45.8% of the forecast variance of temperature, while the level of yield can only explain about 1.8% temperature forecast variance and 52.4% variation of temperature is resulted from the shock of its own innovation. This remains the same up to the tenth period. (Table 3. of Appendix B and Figure.8 of Appendix C). This result provides similar evidence as the impulse response method. That is the relationship between rainfall and temperature is explained by the level of evaporative cooling in the system.

For yield, in the first round, 48.6% variation of yield resulted from the shock of its own innovation and 3.2% variation is resulted from the change in temperature and the remaining 48.1% from the change rainfall, that is yield and rainfall is similar. In the long term, rainfall and temperature can explain 44.5% and 3% forecast variance of yield, respectively, and 52.5% variation of yield is resulted from the shock of its own innovation. This shows that rainwater is a principal component in determining sorghum yield.

CHAPTER FIVE
CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

The aim of this study is to assess the effect of climate variability and change (rainfall and temperature) on production of sorghum over long period of time in Melkassa area, Eastern Ethiopia. Station level rainfall and temperature data from 1978-2009 and agricultural production data of sorghum (Gambella -1107 variety) from Melkassa Research Center of EIAR were used. Vector Auto-Regressive (VAR) model was employed.

The statistical analyses have provided essential numerical evidences for the existence of variability in total-rainfall and yield in the study area. The long-term annual rainfall and yield showed high variability from year to year with 18.43% and 63.38% coefficient of variation respectively.

Results from IRF measures show that the response of a variable for a one standard deviation of its innovations has a significant impact on its own for all variables included in the analysis. For a one SD change in rainfall, temperature has an opposite response while sorghum yield has direct response to rainfall. Rainfall and yield showed a negative response for a one SD change in temperature.

The decomposition of the variations into the component shocks was made by FEVD analysis. The results of the decomposition give the proportion of the movements in the dependent variables that are due to their 'own' shocks, versus shocks to the other variables. Accordingly, the rainfall movement is almost totally (100%) explained by its own innovations. Yield variation of up to 48.1% is explained by changes in rainfall amounts, the percentage contribution of yield shock for its forecast variance is around 48.6%. Temperature shock provides 52.4 % of the forecast error variance of temperature and rainfall shock has 45.8% contribution for the forecast error variance of temperature.

5.2. Recommendations

The following recommendations are forwarded:

- Researchers should collect and use meteorology data and make detailed analysis of rainfall impacts on sorghum to utilize silent features of the data and to help advice policy makers.
- An early warning of rainfall patterns and distribution should be the priority to mitigate impacts of rainfall variability in order to provide useful information that can easily be used by researchers and farmers.
- Better statistical model of the rainfall variability that can provide precise forecast with high predictive power should be developed.
- Introduction of new technologies that goes in line with the changing climate patterns should be the prior agenda for researcher and development planners in order to arrest declining of sorghum yield of Gambella 1107 variety in the study area due to climate change.

REFERENCES

Abate, K. (1994): The Climatology of Drought over Parts of Ethiopia and their Impacts on Crop Production with Special Reference to the Impact of Drought on the Production of Barley and Maize. PhD thesis (Climatology), University of Nairobi, Kenya.

Adugna, L. (2005): Rainfall Probability and Agricultural Yield in Ethiopia, Eastern Africa Social Science Research Review, 21(1), January 2005, 57-96.

Almaz, D. (2004): Effects of Climate Change on Agriculture Particularly In Semi-Arid Tropics of the World with some examples of Ethiopian condition; BALWOIS.

Awulachew S.B., Aster D.Y., Mekonnen L., Willibald L., Mekonnen A. and Tena A. (2007): Water Resource and Irrigation Development in Ethiopia. Colombo, Sri lanka: International Water Management Institute pp78 (working paper123)

Badawi, A. (2007), “Testing Stationarity in Selected Macroeconomic Series from Sudan,” African Journal of Economic Policy.

Ben M., and Abdoussallam S. (2002): Impact Of Climate Change On Agricultural Production In The Sahel – Part 1. Methodological Approach And Case Study For Millet In Niger., Climatic Change 54: 327–348.

Blignaut U., Ueckermann.L and Aronson J. (2006): Agricultural Production’s Sensitivity to Changes in Climate in South Africa.

Borrell, A., E.V. Oosterom, H. Graeme, J. David and H. Bob (2003): Using science to combat drought: a case study of stay-green in sorghum. In: R. Stone and I. Partridge (Eds). Science for Drought. Proceedings of the National Drought Forum, 120-123.

Box G. E. P., Jenkins G. M. (1970): Time Series Analysis Forecasting and Control, Holden-Day.

Box, George P. and Jenkins, G.M. (1976): Time Series Analysis Forecasting and Control, Revised edition Holden-Day.

Casdagli, M., and Eubank S., (1992): Nonlinear Modeling and Forecasting., Addison-Wesley.

Chatfield C. (1988): What is the ’best’ method of forecasting?, Journal of Applied Statistics, 15(1), 19-38.

Chatfield C. (1989): The Analysis of Time Series . An Introduction.,Chapman and Hall, 4th edition.

Chen C., McCAR B., and Schimmelpfennig D., (2004); Yield Variability As Influenced By Climate: A Statistical Investigation; Climatic Change 66: 239–26. Kluwer Academic Publishers. Printed in the Netherlands.

Crabtree, B.F., Ray, S.C., Schmidt, P.M., O'Connor, P.J. & Schmidt, D.D. (1990): The individual over time: time series application in health care research. Journal of Clinical Epidemiology, 43(3), 241-260.

David B., Kimberly N. and Christopher B. (2007): Historical effects of temperature and precipitation on California crop yields. Climatic Change 81:187–203.

Degefu, W. (1987): Some Aspects of meteorological drought in Ethiopia. In: M.H. Glantz (Ed).

De Vries, W.R. & Wu, S.M. (1978): Evaluation of Process Control Effectiveness and Diagnosis of Variation in Paper Basis Weight via Multivariate Time-Series Analysis. IEEE Transactions on Automatic Control, 23(4), 702-708.

Dickey, D. A., and Fuller, W. A. (1979): "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of American Statistical Association*, 74, 427-431.

Dickey, D. A., and Fuller, W. A. (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with Unit Root," *Econometrica*, 49(4), 1057-1072.

EARO (Ethiopian Agricultural Research Organization) (2000). National Sorghum Strategy Document. pp28.

Earsterling, W.E. and Mjelde, J.W. (1987): The importance of seasonal climatic lead time in agricultural decision-making. Agriculture and Forestry Meteorology 40:37-50.

Engle, R.F., and Granger C.W.J. (1987): "Co-Integration and Error Correction: Representation, Estimation, and Testing." *Económétrica* 55:251-76.

Ethiopian Delegation (1984). Climatic and drought conditions in Ethiopia. Scientific Round Table Discussion on the Climatic Situations and Drought in Ethiopia. Addis Ababa, Ethiopia. 20-23 Feb 1984, Mimograph.

FAO (Food and Agriculture Organization) (1984): Assistance to land use planning in Ethiopia: geomorphology and soil. Report prepared and submitted to the Government of Ethiopia. AGDA. ETH/78/003, Field Document 3.

FAO (1989): Assistance to land use planning of Ethiopia: Physiography and soils of the Hykoch and Butajira and Yerer and Kereyu Awurajas (Shewa). FAO/AG: DP/ET/87/006. Field Document No. 37.

Faraway J., Chatfield C. (1998): Time series forecasting with neural networks: a comparative study using the airline data., Journal of the Royal Statistical Society: Applied Statistics (C), 42(2), 231-250.

Gemechu, D. (1977). Aspects of climate and water budget in Ethiopia. Addis Ababa University Press. pp77.

Gay C., Estrada F., Conde C., Eakin H. and Villers L. (2006): Potential Impacts Of Climate Change On Agriculture: A Case Of Study Of Coffee Production In Veracruz, Mexico; Climatic Change 79: 259–288

Girma M. (2005): Using Seasonal climate outlook to advise on sorghum production in the centeral rift vally of Ethiopia. PhD thesis, Blomefontein, Republic of South Africa.

Glantz, M.H.; R.W. Katz and N. Nicholls (1991): Tele-connections Linking World Climate Anomalies. Cambridge University Press, New York. pp527.

Gonfa, L. (1996): Climatic classification of Ethiopia. NMSA, Addis Ababa, Ethiopia.

Granger, C.W.J. (1969). “Investigating Causal Relations by Econometric Models and Cross Spectral Methods,” Econometrica, 37, 424-438.

Green, G.C. (1966). The Evaluation of Methods of Rainfall Analyses and the Application to the Rainfall Series of Nelspruit. MSc thesis (Agrometeorology) University of the Orange Free State, Bloemfoentein.

Grubb H. (1992): A Multivariate Time Series Analysis of Some Flour Price Data Journal of the Royal Statistical Society. Series C (Applied Statistics), 41(1), 95-107.

Haile T. (1986): Climatic Variability and Support Feedback Mechanisms in Relation to the Sahelo-Ethiopian Droughts. MSc thesis (Meteorology), Reading University, UK.

Hollinger, S.E., (1994). Future direction and needs in agricultural meteorology-climatology and modeling. Agriculture and Forestry Meteorology 69: 1-7.

Hannan E. J. (1970): Multiple Time Series., Wiley

Hoshmand, A.R. (1998): Statistical Methods for Environmental and Agricultural Sciences 2nd ed. CRC Press. pp439

Hui W.L. (2004): Application of Vector Autoregresive models to Hong Kong Air Pollution Data

Johanson, S. (1988): Likilihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford :Oxfored University Press.

Johanson S. and Juselius K. (1990): Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money. Oxford Bulletin of Economics and Statistics, 169-210.

Lorek and Willinger (1996): A Multivariate Time-Series Prediction Model For Cash-Flow Data Source: The Accounting Review, 71(1), 81-102.

Kibrom T.(2008): The Sources of the Recent Inflationary Experience in Ethiopia. MSc thesis, Addis Ababa University, Ethiopia.

Kidane G., Abebe T. and Degefie T. (2006): Estimating crop water use and simulating yield reduction for maize and sorghum in Adama and Miesso districts using the Cropwat model. Ethiopian Agricultural Research Institute, 1-14.

Kidson, J.W. (1977): African rainfall and its relation to the upper air circulation. Quart. J. R. Meteorol. Soc. 103: 441-456.

Landberg, H. (1960). Physical Climatology, 2nd Ed.: Gray Printing Co., Inc., Dubois, Pensylvania.

Lütkepohl H. (1993): Introduction to Multivariate Time Series Analysis., Springer-Verlag, 2nd edition

Maddala, G. S. (1992): Introduction to Econometrics, 2nd eds, MACMILLAN Publishing Company, New York.

Meinke, H. (2003): Options for cropping systems management at a range of time scales. In: R.C. Stone and I. Partridge (Eds). Science for Drought. Proceedings of the National Drought.

Mahmoud E. (1984): Accuracy in Forecasting: a Survey., Journal of Forecasting, 3, 139-159.

Makridakis, S., S.C. Wheelbright and V.E. McGee (1983): Forecasting Methods and Applications, 2nd ed., John Wiley and Sons, New York. pp923.

Mesfine T., G. Abebe and A.M. Al-Tuwaha (2005): Effect of reduced tillage and crop residue ground cover on yield and water use efficiency of sorghum (*Sorghum bicolar L.)* under semi arid conditions of Ethiopia. *World Journal of Agricultural Sciences,* 1(2):152-160.

MoARD, (2007): Ministry of Agriculture and Rural Development crop development crop variety register. Issue No.7. Addis Abeba, Ethiopia. pp104.

Naipal S. (2004): Trends And Variation In Monthly Rainfall And Temperature In Suriname.

Newey, W. K. and Kenneth D. W. (1987): "A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55:703-8.

NMSA (National Meteorological Services Agency) (1996a): Climate and agroclimatic resources of Ethiopia. Meteorological Research Report Series. 1(1), 137, Addis Ababa.

NMSA (National Meteorological Services Agency) (1996b): Assessment of drought in Ethiopia: Meteorological Research Report Series. 1(2), 259, Addis Ababa.

NRRD / MOA (Natural Resources Regulatory Department of the Ministry of Agriculture). (2000) Agroecological Zones of Ethiopia. pp129

Numata M., Sugawara K., Yoshihara I., Abe K., (1998): Time Series Prediction by Genetic Programming., Genetic Programming Conference: Late Breaking Papers, Edited by Koza J., 176-179.

Phillips, P. C. B. and Perron, P. (1988): "Testing for a Unit Root in Time Series Regression", *Biometricka*, 75, 335-346.

Reinsel G. (1997): Elements of Multivariate Time Series Analysis. Springer-Verlag New York, Inc.,

Samuel A. (2007): Demand for Electricity in Ethiopia as Related to consumption, population and Domestic product. MSc thesis, Addis Ababa University, Ethiopia.

Scott M. Robenson.,(1997): Statistical considerations. Appliedclimatology, 22-35.

Sims, C.A. (1980): "Macroeconomics and Reality, "Econometrica, 48, 1-48.

Smit, B., Mcnabb, D., and Smithers, J., (1996): Agricultural adaptation to climatic variation. Climatic Change. 33, 7-29.

Taylor, J.A. (1972): Weather forecasting for agriculture and industry. David & Charles Newton Abbot. 250 pp.

Taylor, J.R.N., (2003): Overview: Importance of sorghum in Africa.www.afripro.org.uk/ papers/paper

Tilahun, K. (2005): Analyses of rainfall climate and evapotranspiration in arid and semi-arid regions of Ethiopia using data over the last half a century. J. Arid Enviro. (article in press).

Upreti, D.C., (2006): Rising Atmospheric CO and crop response. SASCOM Scientific Report, pp1-8. World Climate News.

Washington, R. and Downing, T.E., (1999): Seasonal forecasting of African rainfall, prediction, responses and household food security. The Geographical Journal. 165 (3):255-274.

Weigend A. S. and Garshenfeld N. A. (1994), Time Series Prediction, Addison-Wesley.

Appendix A: Introduction of Key Words in the Thesis

Agrometeorology is abbreviated from agricultural meteorology, puts the science of meteorology to the service of agriculture, in its various forms and facets, to help with the sensible use of land, to accelerate the production of food, and to avoid the irreversible abuse of land resources (Smith,1970). Agrometeorology is also defined as the science investigating the meteorological, climatological, and hydrological conditions that are significant to agriculture owing to their interaction with the objects and processes of agriculture production (Molga, 1962).

Medium Maturing Sorghum: grows in intermediate altitude between 1600-1900m above sea level and with maturity date of 120-130.

Gambella – 1107 Varieties: is a white seeded variety with semi compact semi oval and erect panicle. Its height may be within the range of 120-200cm. Usually part of the head is covered by the flag leaf (not well exscerted).

Appendix B:

Table.1 Rainfall, Temperature and Yield Data (1978-2009) Melkassa, Ethiopia

Year	Rainfall in mm	Yield in Quintal/hectare	Temperature in C^0
1978	445.00	18.90	21.77
1979	520.00	22.35	21.51
1980	516.00	19.30	21.40
1981	511.72	17.76	21.30
1982	386.00	6.11	21.43
1983	393.00	7.93	22.00
1984	385.60	4.12	22.87
1985	519.10	19.48	22.86
1986	405.00	15.60	22.96
1987	440.40	28.30	22.91
1988	560.30	28.72	22.45
1989	544.00	27.01	21.94
1990	459.00	20.96	21.94
1991	535.74	48.01	21.98
1992	344.00	7.46	22.81
1993	520.00	47.00	22.00
1994	535.00	47.56	21.40
1995	490.60	4.62	21.50
1996	540.20	53.00	21.40
1997	575.00	69.00	21.09
1998	690.00	70.00	20.75
1999	682.00	69.50	20.67
2000	599.00	52.00	20.60
2001	500.00	45.00	21.70
2002	320.00	3.00	22.93
2003	498.00	15.20	22.65
2004	572.00	47.00	22.34
2005	590.00	48.00	22.13
2006	632.80	49.00	22.69
2007	615.00	38.00	22.80
2008	602.00	35.00	22.98
2009	380.40	29.00	23.00

Table 2. Impulse Response Functions for the VAR (1) model

Response of RD:

Period	RD	TD	YD
1	97.26466	0	0
2	-18.519	7.283139	-10.4052
3	1.236237	-0.10671	5.806057
4	-0.95052	0.361292	-1.8994
5	0.250569	-0.07801	0.710071
6	-0.10245	0.03549	-0.24895
7	0.034603	-0.01152	0.0896
8	-0.01266	0.004287	-0.03189
9	0.004474	-0.0015	0.011403
10	-0.0016	0.000541	-0.00407

Response of TD:

Period	RD	TD	YD
1	-0.30103	0.339186	0
2	-0.11349	0.06171	0.063674
3	-0.0179	0.010068	-0.0014
4	-0.00148	0.001089	0.003274
5	-0.0007	0.000326	-0.00073
6	6.60E-05	-8.62E-06	0.000327
7	-5.15E-5	1.94E-05	-0.00011
8	1.42E-05	-4.46E-06	3.96E-05
9	-5.70E-6	1.97E-06	-1.39E-05
10	1.94E-06	-6.46E-07	5.01E-06

Response of YD:

Period	RD	TD	YD
1	13.31289	-3.45181	13.3846
2	-2.56896	0.820046	-5.79539
3	0.843632	-0.28336	2.104924
4	-0.29668	0.099819	-0.75361
5	0.106019	-0.03574	0.269085
6	-0.03782	0.012743	-0.0961
7	1.35E-02	-4.55E-3	0.034313
8	-0.00482	0.001625	-0.01225
9	1.72E-03	-5.80E-04	0.004375
10	-0.00062	0.000207	-0.00156

Table 3. Forecast Error Variance Decomposition Function

Proportion of Rainfall explained by innovation of Rainfall, Temperature and Yield

Variable	Steps	Rainfall	Temperature	Yield
Rainfall	1	1	0	0
	2	0.983812	0.005323	0.010865
	3	0.980497	0.005306	0.014198
	4	0.980132	0.005316	0.014552
	5	0.980082	0.005316	0.014601
	6	0.980076	0.005317	0.014608
	7	0.980075	0.005317	0.014608
	8	0.980075	0.005317	0.014608
	9	0.980075	0.005317	0.014608
	10	0.980075	0.005317	0.014608

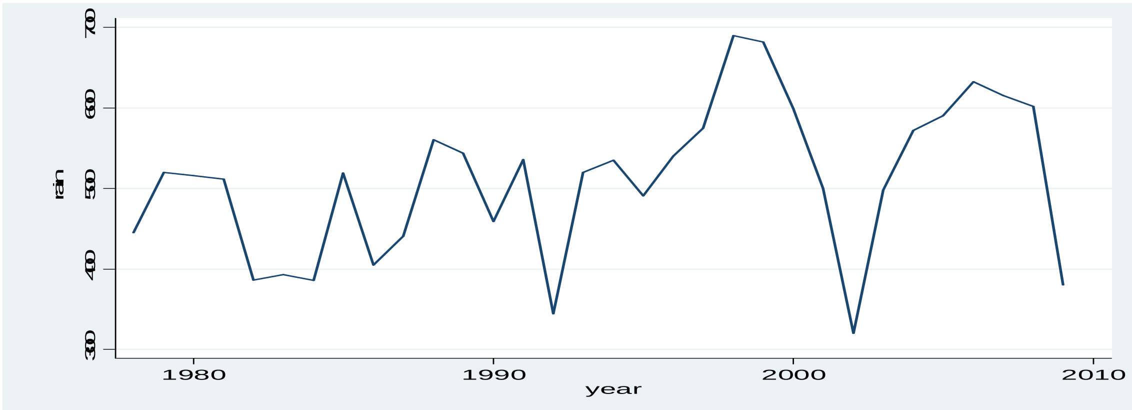
Proportion of Temperature explained by innovation of Rainfall, Temperature and Yield

Variable	Steps	Rainfall	Temperature	Yield
Temperature	1	0.44061	0.55939	0.00000
	2	0.45713	0.52496	0.01791
	3	0.45769	0.52443	0.01788
	4	0.45767	0.5244	0.01793
	5	0.45767	0.5244	0.01793
	6	0.45767	0.5244	0.01793
	7	0.45767	0.5244	0.01793
	8	0.45767	0.5244	0.01793
	9	0.45767	0.5244	0.01793
	10	0.45767	0.5244	0.01793

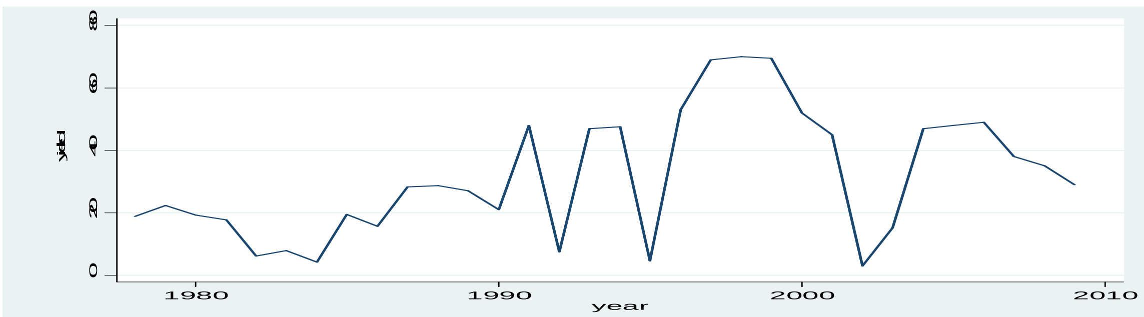
Proportion of Yield explained by innovation of Rainfall, Temperature and Yield

Variable	Steps	Rainfall	Temperature	Yield
Yield	1	0.48123	0.03235	0.48642
	2	0.4493	0.03077	0.51994
	3	0.44535	0.03057	0.52408
	4	0.44485	0.03055	0.5246
	5	0.44479	0.03054	0.52467
	6	0.44478	0.03054	0.52468
	7	0.44478	0.03054	0.52468
	8	0.44478	0.03054	0.52468
	9	0.44478	0.03054	0.52468
	10	0.44478	0.03054	0.52468

Appendix C: **List of Figures**



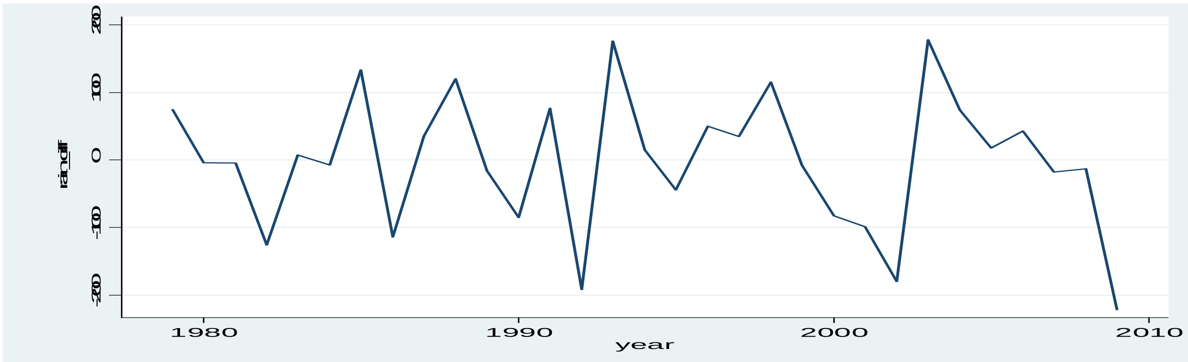
Time Plots of Rainfall



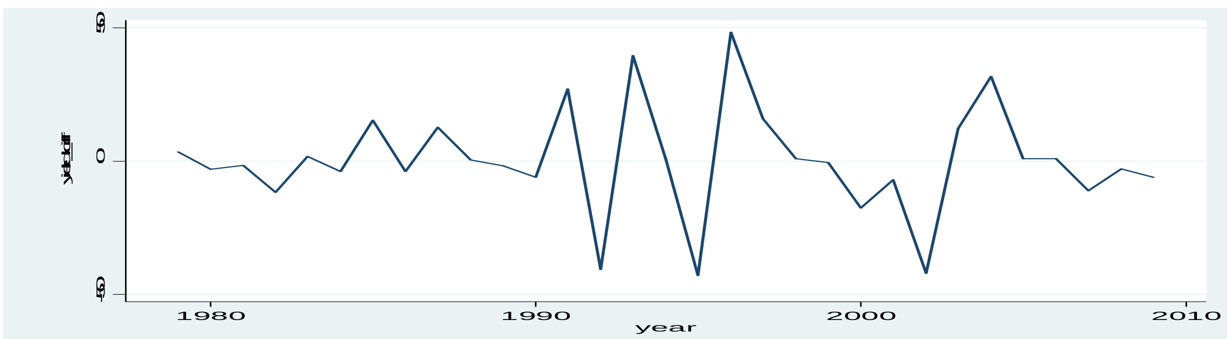
Time Plots of Yield

Time Plots of Temperature

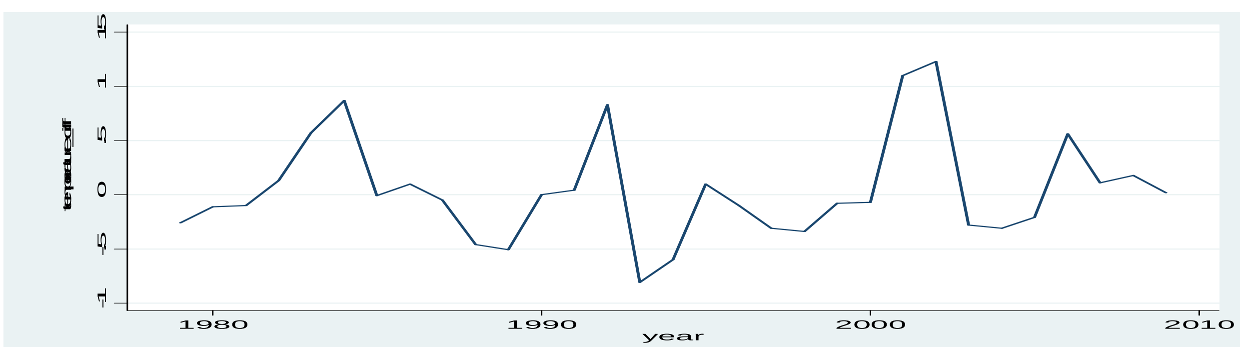
Figure.1 Time plots of the original series



Rainfall differenced

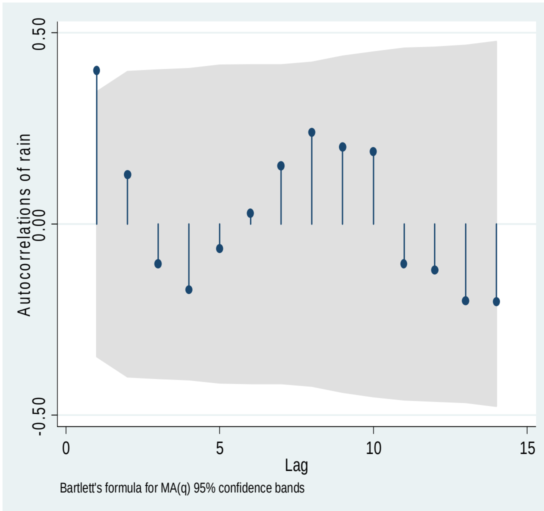


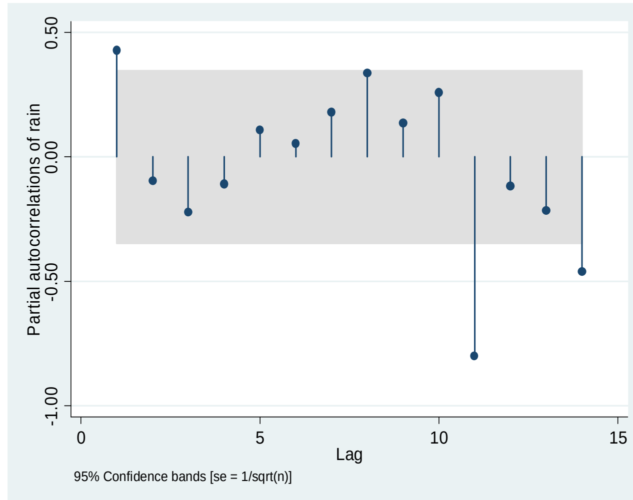
Yield differenced



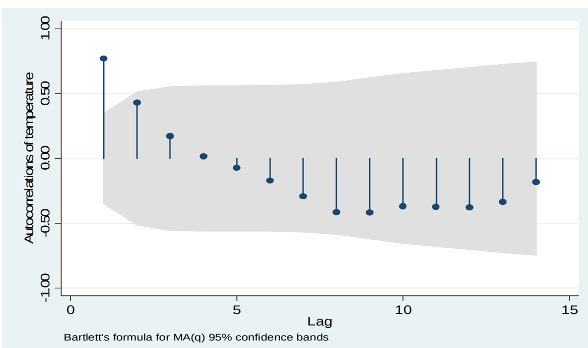
Temperature differenced

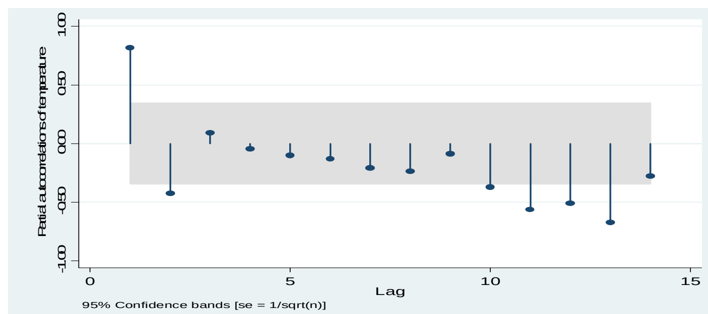
Figure.2 Time plots of the differenced series



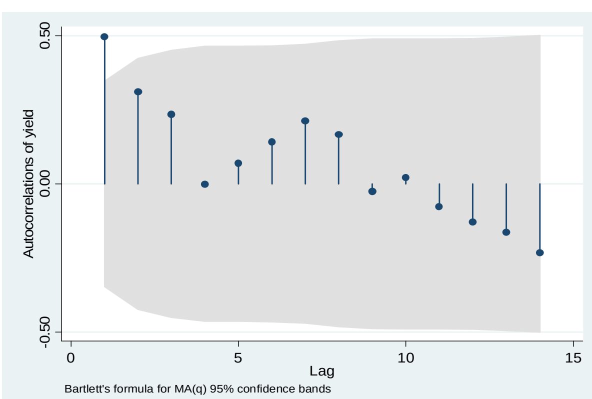


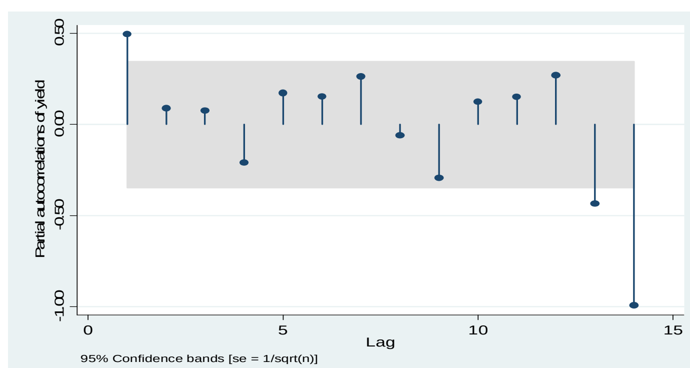
Rainfall





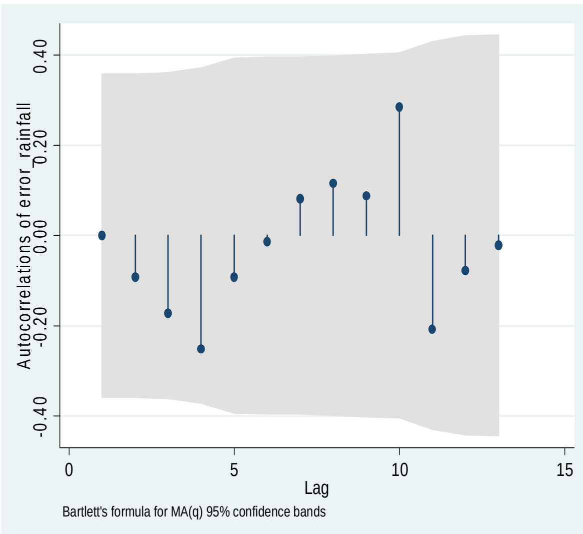
Temperature

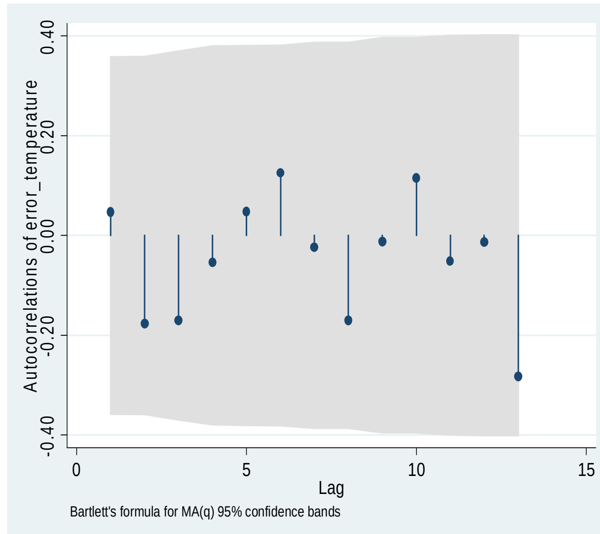




Yield

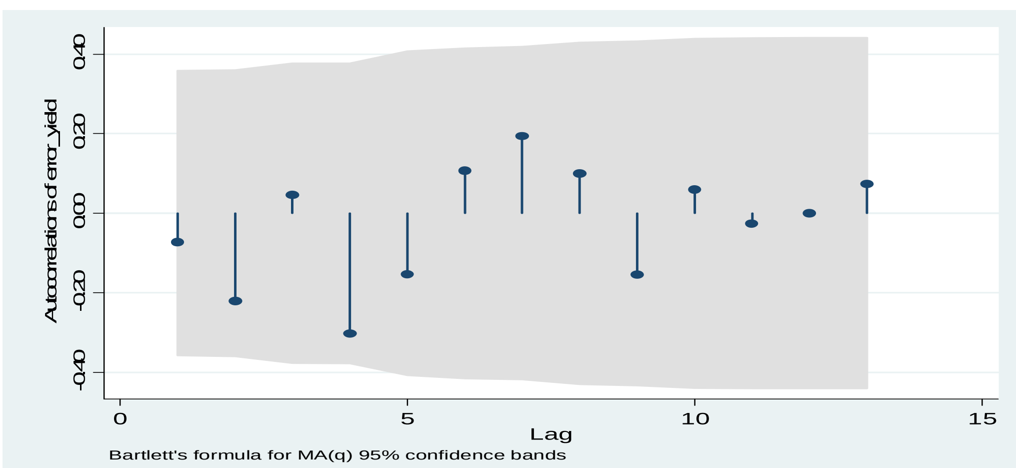
Figure.3 Autocorrelation Function and Partial Autocorrelation Function of our series





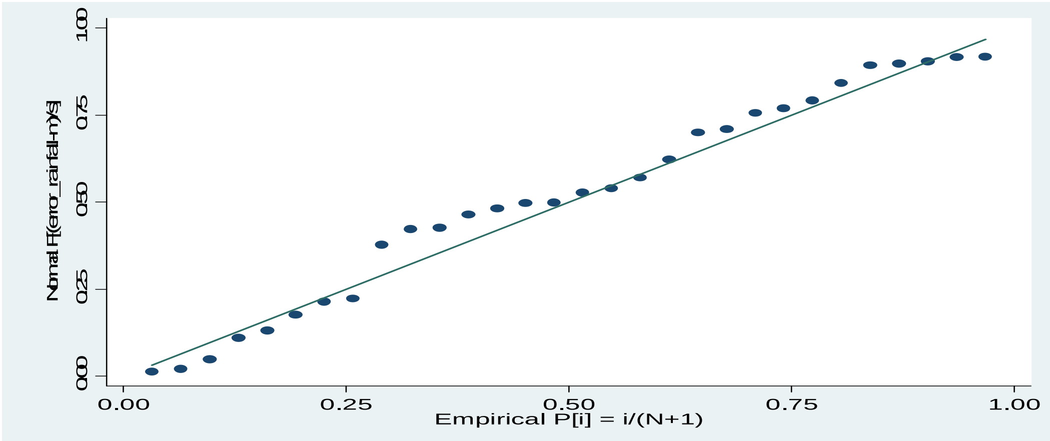
Rainfall differenced

Temperature differenced

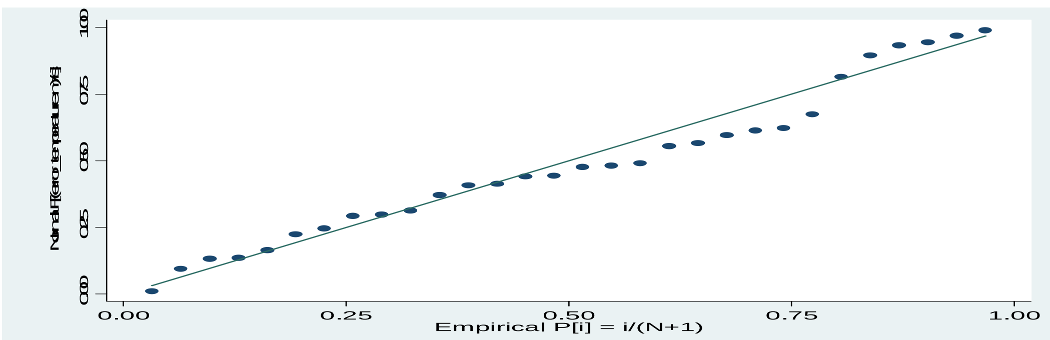


Yield differenced

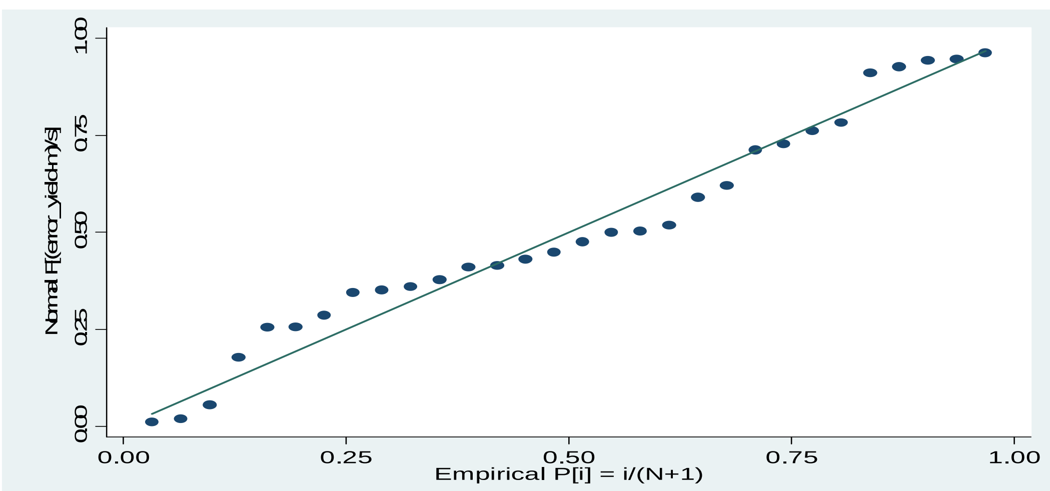
Figure.4 Autocorrelation Functions of the residuals



Rainfall residual



Temperaturel residual



Yield residual

Figure5. Normal probability plots of the residuals

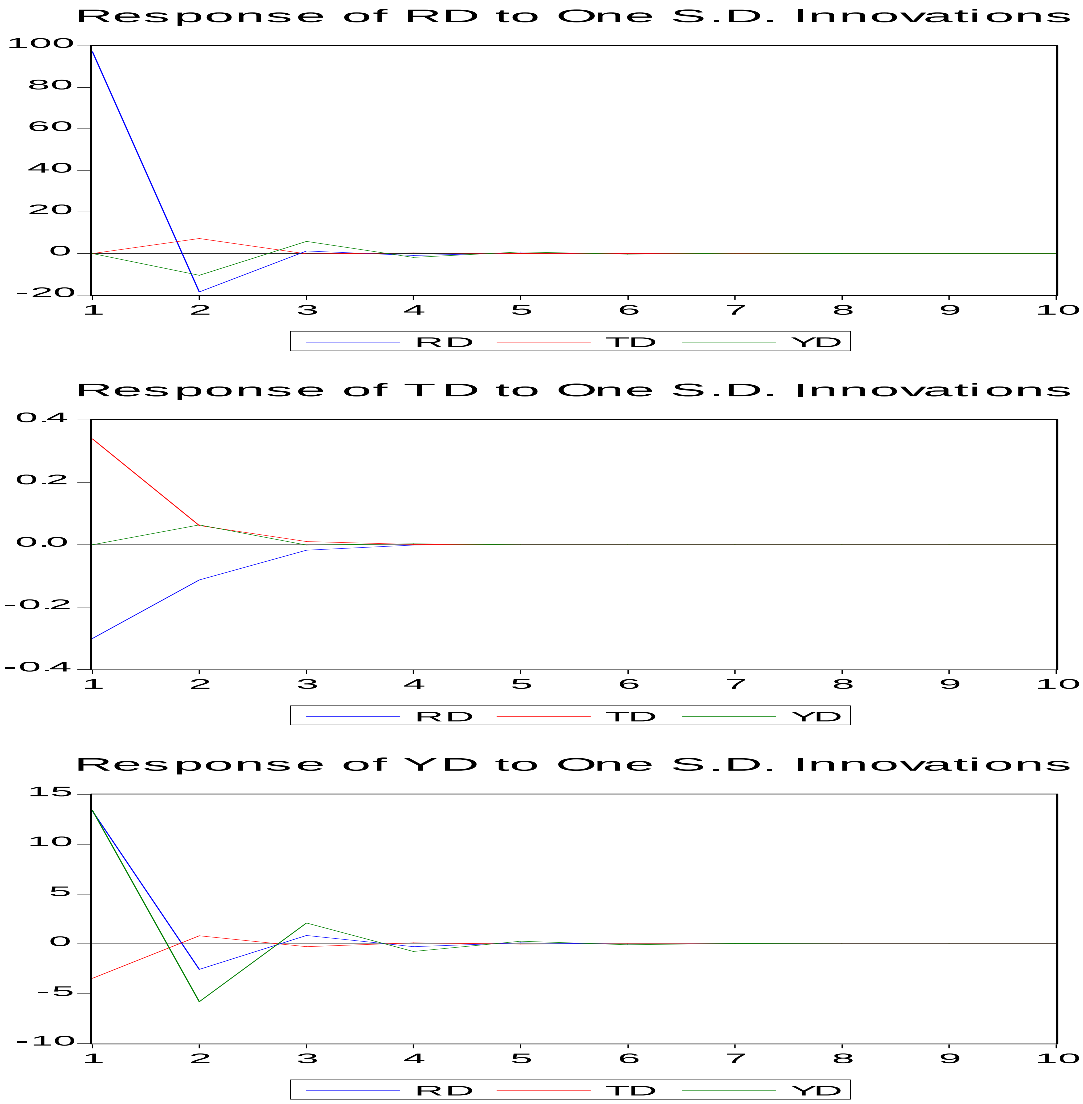


Figure.6 Graphs of Impulse Response Function

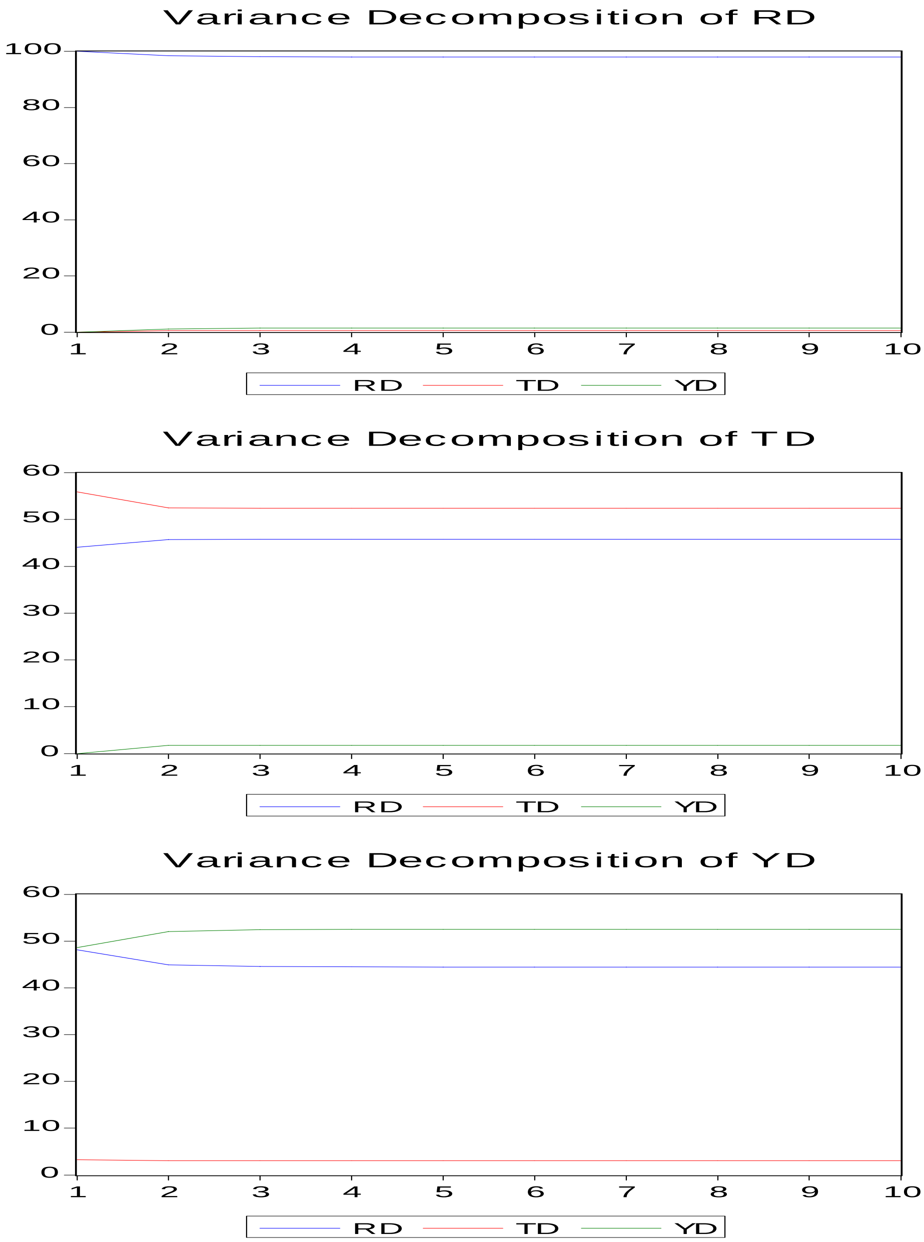


Figure.7 Forecast Error Variance Decomposition

DECLARATION

I, the undersigned, declare that this thesis is my original work, has not been presented for degrees in any other University and all source materials used for the thesis have been duly acknowledged.

Name: Tigist Mideksa

Signature: ___________________________

Date: _________________________________

Place: College of Science, Addis Ababa University

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

Name: Butte Gotu (PhD)

Signature: ___________________________

Date: ________________________________

Place: College of Science, Addis Ababa University