

**APPLICATION OF BAYESIAN VECTOR AUTOREGRESSIVE MODEL IN  
FORECASTING RAINFALL PATTERN IN KENYA.**

**BY**

**GITONGA HARUN MWANGI**

**A RESEARCH THESIS SUBMITTED TO THE SCHOOL OF SCIENCES AND  
AEROSPACE STUDIES, DEPARTMENT OF MATHEMATICS, PHYSICS  
AND COMPUTING IN FULFILMENT OF THE REQUIREMENTS FOR THE  
AWARD OF THE DEGREE OF DOCTOR OF PHILOSOPHY IN  
BIOSTATISTICS**

**MOI UNIVERSITY**

**2022**

## DECLARATION

### **Student Declaration**

I hereby declare that this thesis is my original work and has not been presented for a degree in any other university. no part of this research thesis should be produced without the consent of the author and/or Moi University.

Signature..... Date.....

**Harun Mwangi Gitonga**

**DPS/PhD/07/14**

### **Supervisors Declaration**

This thesis has been submitted with our approval as the university supervisors.

Signature..... Date.....

**Prof. Joseph K. Koske**

Department of Mathematics, Physics and Computing

Moi University

Signature..... Date.....

**Dr. Matthew Kosgei**

Department of Mathematics, Physics and Computing

Moi University

## **DEDICATION**

I would like to dedicate this thesis to my wife; Mary, my Children; Eugene and Elizabeth, for their unwavering support they accorded and displayed to me. To my parent, they encouraged me to always work hard and never to give up.

## **ACKNOWLEDGEMENT**

Firstly, I would like to acknowledge and give my heartfelt thanks to my supervisors Prof. Joseph Koske and Dr. Mathew Kosgei who made this work possible. Their valuable guidance, advice and feedback carried me through all the stages of writing my thesis. I am extremely grateful to you for giving me such a wealth of experience.

I would also like to give special thanks to my wife, Mary and my family as a whole for their continuous support and understanding when undertaking my research and writing my thesis. Your prayers for me were what sustained me this far.

I would also wish to acknowledge my dear friends and colleagues and in particular Dr. Robert Kasisi for their support and encouragement. May our dear almighty always bless you.

Finally, I would like to thank God, for letting me through all the difficulties. I have experienced your guidance day by day. You are the one who let me finish my PhD and I will keep on trusting you for my future.

## ABSTRACT

Time series modelling is of fundamental importance in forecasting weather that is basically one of the most technologically and scientifically challenging problems around the world currently. To make an accurate prediction is certainly one of the key challenges that meteorologists are facing all over the world. One of the most affected areas is the rainfall patterns, which is being influenced by global warming, causing drastic changes in its patterns that are characterized by either very high or low precipitation and temperature. These extreme changes have been identified as major global challenges of recent times. Meteorological scientist always tries to find means to understand the atmosphere of the Earth, and to develop accurate weather prediction models. Several methods have been used in weather prediction, which includes, Classical vector Autoregressive (VAR) models which perform only polynomial-time computation to compute the probability of the next fixed model parameters. While this is attractive, it means they cannot model distributions with a time varying data. They also have a problem with the curse of dimensionality. Recently, machine learning methods are assumed to be accurate techniques and have been widely used as an alternative to classical methods for weather prediction. With all these powerfulness and popularity machine learning methods are not perfect. They have several limitations where they require; massive datasets, enough time and resources, does not work well with high dimensional data and have high error vulnerability among others. Despite the availability of different models that are used by the meteorologists and other departments to make predictions, the same devastating scenarios of unpredictable weather changes are still being experienced. Therefore, robust models reliable for accurate predictions are needed on short- and long-term time scales to reduce potential risks and damages that may occur due to unexpected weather changes. These shortcomings are well addressed by the Bayesian Vector Autoregressive (BVAR) models. The purpose of this study was to develop a BVAR model for predicting rainfall patterns in Kenya. The specific objectives were to; perform diagnostic analysis of the weather variables; develop Bayesian Vector Autoregressive predictive model; conduct model performance analysis and apply the model to forecast the rainfall patterns in Kenya. The Augmented Dicker Fuller and Granger Causality tests were used for diagnostic analysis. The research adopted secondary data for a period of four years (2014-2018), which was sourced from Trans-African Hydro-Meteorological Observatory (TAHMO) and Kenya Meteorological stations. Bayesian Vector Autoregressive model was developed using multiple regression analysis in a system of equations. The model imposes structures through information prior beliefs on the parameters which were obtained from VAR models, likelihood models between the true parameters and the measured variables and the posterior distribution which is the conditional distribution of the parameter given the measurements. The model sensitivity was performed using the confusion matrix. The F-test was used to compare the variances of the actual and the predicted rainfall values. The data was analyzed using R-Statistical Software. The study found that; the data variables were stationary after at least the first differencing. Temperature, atmospheric pressure, wind speed and relative humidity were statistically significant ( $p < 0.05$ ) determinants of rainfall in all five zones, while wind gust and radiation were significant in two zones, coast and arid areas. The BVAR model developed was statistically significant ( $R^2 = 0.9896$ ). The performance of the model was adequate (RMSE= 86.81%) and its sensitivity was 82.52%, making it appropriate for forecasting. There was no significant difference between the variances of the actual and predicted values of rainfall ( $p = 0.3893$ ) at the 5% level of significance in zone five. The study made the following conclusion, after at least the one differencing the weather variables were found to be stable, the developed model coefficients were found to be statistically significant, the model performance was good and its forecasting ability was termed as high. In conclusion, the Bayesian Vector Autoregressive model developed is suitable for forecasting rainfall patterns in Kenya. The study recommends adoption of the BVAR model by relevant authorities to predict rainfall. For further research the study recommends for use of more weather variables to make more accurate prediction. The study also recommends for development of dynamic weather model, which tests for the impulse response of weather variable in respect to change in other endogenous variables.

## TABLE OF CONTENTS

DECLARATION .....	ii
DEDICATION .....	iii
ACKNOWLEDGEMENT .....	iv
ABSTRACT.....	v
TABLE OF CONTENTS.....	vi
LIST OF TABLES .....	xi
LIST OF FIGURES .....	xii
ABBREVIATIONS AND ACRONYMS .....	xiii
<b>CHAPTER ONE .....</b>	<b>1</b>
<b>INTRODUCTION.....</b>	<b>1</b>
1.1 Background of the Study .....	1
1.2 Statement of the Problem.....	4
1.3 Significance of the Study .....	4
1.4 Motivation of the Study .....	6
1.5 Objectives of the Study.....	7
1.5.1 General objectives .....	7
1.5.2 Specific Objectives.....	7
1.6 Scope of the Study .....	8
1.7 Limitation of the Study .....	9
<b>CHAPTER TWO .....</b>	<b>10</b>
<b>LITERATURE REVIEW .....</b>	<b>10</b>
2.0 Introduction.....	10
2.1 Weather Background Information .....	10
2.2 Time Series models.....	12
2.2.1 Augmented Dickey-Fuller test .....	13
2.2.2 Granger Causality test .....	14
2.3 Model Estimation.....	14
2.3.1 Lag order selection .....	15
2.3.2 Vector Autoregressive.....	15
2.3.3 Bayesian Vector Autoregressive Model.....	17
2.3.4 Priors for Model Selection .....	22

2.3.5 Informative and noninformative priors .....	24
2.3.6 Time Varying Parameters.....	25
2.3.7 Breusch Pagan and Breusch Godfrey Test.....	26
2.4 Model Forecasting .....	27
2.5 Critique of the Literature Review .....	28
<b>CHAPTER THREE .....</b>	<b>32</b>
<b>METHODOLOGY .....</b>	<b>32</b>
3.0 Introduction.....	32
3.1 Data Preprocessing.....	32
3.2 Diagnostic Test of the Weather Variables .....	33
3.2.1 Stationarity Tests.....	33
3.2.2 Augmented Dickey-Fuller test .....	34
3.2.3 Granger Causality.....	37
3.3 Developing Specific Model .....	39
3.3.1 Vector Autoregressive Model Specification .....	40
3.3.2 Vector Autoregressive Order .....	43
3.3.3 Estimation of a Vector Autoregressive (p) model.....	44
3.3.4 The Global Vector Autoregressive Model .....	44
3.3.5 The Global Vector Autoregressive framework .....	46
3.3.6 Bayesian models.....	47
3.3.7 BVAR Model Development.....	48
3.3.8 Bayesian model in GVAR.....	51
3.4 Model performance Analysis .....	52
3.4.1 Confusion Matrix .....	52
3.4.2 Training and test data sets .....	53
3.5 Forecasting Analysis.....	53
3.5.1 Evaluation of predictive model .....	54
<b>CHAPTER FOUR.....</b>	<b>56</b>
<b>RESULTS AND DISCUSIONS .....</b>	<b>56</b>
4.0 Introduction.....	56
4.1 Diagnostic analysis of the variables.....	56
4.1.1 Stationary test.....	57
4.1.2 Augmented Dickey fuller (ADF) and Phillips-Perron (PP) test.....	58
4.1.3 ADF and PP test for zone one .....	58

4.1.4 Zone two stationarity test. ....	59
4.1.5 Zone three.....	60
4.1.6 Zone four.....	61
4.1.7 Zone five .....	62
4.1.8 Global vector.....	63
4.1.9 Granger causality test. ....	64
4.1.10 Zone one.....	64
4.1.11 Zone two.....	65
4.1.12 Zone three.....	65
4.1.13 Zone four.....	66
4.1.14 Zone five .....	67
4.1.15 Global vector granger causality test.....	67
4.2 Development of the model.....	68
4.2.1 Lag order selection.....	68
4.2.2 Zone one lag order selection .....	68
4.2.3 The model system output for zone one .....	69
4.2.4 Model diagnostic test .....	70
4.2.5 Zone two.....	71
4.2.6 The model system output for zone two .....	71
4.2.7 Model diagnostic test .....	72
4.2.8 Zone three.....	73
4.2.9 Zone three model.....	73
4.2.10 Model three diagnostic test .....	75
4.2.11 Zone four model.....	75
4.2.12 The model system output for zone four .....	76
4.2.13 Model four diagnostic test.....	77
4.2.14 Zone five model .....	77
4.2.15 The model system output for zone five.....	78
4.2.16 Model diagnostic test .....	78
4.2.17 Global vector.....	79
4.2.18 The model system output for global vector.....	79
4.2.19 Model diagnostic test for global vector.....	80
4.2.20 Models Performance Accuracy Measure .....	81
4.2.21 Performance Measures .....	81



4.2.22 Sensitivity Analysis.....	81
4.3 Forecasting the rainfall pattern .....	82
4.3.1 Zone one: Residual analysis.....	82
4.3.2 Zone one forecasting analysis .....	84
4.3.3 Zone one model comparison results of actual and predictive .....	84
4.3.4 Zone Two .....	85
4.3.5 Zone Three Forecasting.....	86
4.3.6 Zone Four .....	87
4.3.7 Zone Five.....	88
4.3.8 Global Region forecasting analysis.....	89
4.3.9 Global region model comparison results of actual and predictive .....	90
<b>CHAPTER FIVE .....</b>	<b>91</b>
<b>SUMMARY, CONCLUSION AND RECOMMENDATION.....</b>	<b>91</b>
5.0 Introduction.....	91
5.1 Summary of the Findings.....	91
5.1.1 Diagnostic analysis of the Variables .....	91
5.1.2 Model Development.....	91
5.1.3 The Model Performance Measures .....	91
5.1.4 Forecasting Model.....	92
5.2 Conclusion of the Study.....	92
5.2.1 Diagnostic analysis.....	92
5.2.2 Model Development for specific zones.....	93
5.3 Performance and Sensitivity of the Model.....	94
5.4 Forecasting .....	94
5.5 Contribution of the Study.....	94
5.6 Recommendations.....	95
5.6.1 Recommendations of the study .....	95
5.6.2 Recommendations for further study.....	96
REFERENCES .....	97
APPENDICES .....	103
Appendix I: Analysis Tables .....	103
Stationarity Test.....	103
Granger causality test .....	106
Appendix II: Figures .....	109

Appendix III: Output.....	114
Appendix IV: R-Codes.....	132

## LIST OF TABLES

Table 4.1: Lag order and criteria for zone one.....	69
Table 4.2: Zone one model coefficient .....	70
Table 4.3: Lag order and criteria for zone two .....	71
Table 4.4: Zone two model coefficient.....	72
Table 4.5: Zone three Lag order and criteria .....	73
Table 4.6: Zone three model coefficient.....	74
Table 4.7: Lag order and criteria for zone four.....	75
Table 4.8: Zone four model coefficient .....	76
Table 4.9: Lag order and criteria for zone five .....	77
Table 4.10: Zone five model coefficient.....	78
Table 4.11: Lag order and criteria for Global vector .....	79
Table 4.12: Global model coefficient .....	80
Table 4.13: Performance Analysis for the developed models. ....	81
Table 4.14: Zone one forecasting analysis.....	84
Table 4.15: The Prediction Accuracy Test .....	84
Table 4.16: Zone Two Forecasting Analysis .....	85
Table 4.17: Zone Two Prediction Error Analysis.....	86
Table 4.18: Forecasting Intervals Analysis.....	86
Table 4.19: Prediction Error Analysis .....	87
Table 4.20: Forecasting Analysis.....	87
Table 4.21: Zone four Prediction Error Analysis.....	88
Table 4.22: Forecasting Analysis for Zone five.....	88
Table 4.23: Prediction Error Analysis for Zone five .....	89
Table 4.24: Forecasting Analysis for Global Vector .....	89
Table 4.25: Prediction Error for Global Vector .....	90

**LIST OF FIGURES**

Figure 4.1: Time series graph for zone one .....	57
Figure 4.2: Time series graph after differencing .....	58
Figure 4.3: Ljung-Box test.....	82
Figure 4.4: Stability graph .....	83
Figure 4.5: HoltWinters forecast for zone one.....	85

## ABBREVIATIONS AND ACRONYMS

<b>ACM</b>	Atmospheric Circulation Models
<b>ADF</b>	Augmented Dickey Fuller
<b>AIC</b>	Akaike Information Criterion
<b>BMA</b>	Bayesian Model Averaging
<b>BVAR</b>	Bayesian Vector Autoregression
<b>FEVD</b>	Forecast Error Variance Decompositions
<b>FPE</b>	Final Prediction Error Criterion
<b>GVAR</b>	Global Vector Autoregression
<b>HQ</b>	Hannan Quinn
<b>IRF</b>	Impulse Response Functions
<b>ITCZ</b>	Inter-Tropical Convergence Zone
<b>MAE</b>	Mean Absolute Error
<b>MC<sup>3</sup></b>	Monte Carlo Markov-Chain model composition
<b>MCMC</b>	Markov Chain Monte Carlo Composition
<b>MLR</b>	Multi Linear Regression
<b>MPE</b>	Mean Percentage Error
<b>MSE</b>	Mean Squared Error
<b>NWP</b>	Numerical Weather Prediction
<b>RMSE</b>	Root Mean Squared Error
<b>SC</b>	Schwarz - Criterion
<b>SPSS</b>	Statistical Package for Social Scientist
<b>TAHMO</b>	Trans-African Hydro-Meteorological Observatory
<b>TVP-VAR</b>	Time-Varying Parameters in Vector Autoregressive
<b>VAR</b>	Vector Autoregression
<b>VARMA</b>	Vector-valued Autoregressive Moving Average
<b>VCOV</b>	Variance Covariance Vector

## CHAPTER ONE

### INTRODUCTION

This chapter discusses the background of the study, statement of the problem and significance of the study. The chapter also give some insight about the motivation, objectives, scope and the limitations of the study.

#### 1.1 Background of the Study

Bayesian inference has experienced a boost in recent years due to important advances in computational statistics and it has grown over the years from its introduction by Thomas Bayes and its expansion by Richard Price. It uses the prior probability distribution of an uncertain parameter ( $\theta$ ). This prior probability distribution expresses the uncertainty about  $\theta$  before taking into account the data to be used, *Statisticat L.L.C (2016)*. The Bayesian vector Autoregressive (BVAR) is constructed from the vector Autoregressive (VAR) model by treating its parameters as random variables and assigning prior probabilities to them. According to Sims (2007), the key objective aspect of Bayesian inference is the set of rules for transforming an initial distribution into an updated distribution conditioned on observations. One weakness of VAR models is that they require time series to have equal lengths in the estimation processes. This requirement induces a loss of potentially valuable information coming from time series that are longer than others. This problem has a remedy through a Bayesian approach such that information in longer time series is aggregated into a prior, which is then used in the estimation of parameters for the VAR processes. The main advantage of BVAR models is that it avoids the problem of collinearity and over-parameterization that regularly occurs with the use of VAR models since BVAR imposes prior on the VAR parameters.

Weather refers to the condition of air on the earth at a given place and time. Weather forecasting is the application of science and technology in predicting the state of the atmosphere for a future time at a given location. It is carried out by collecting quantitative data about the current and past state of the weather conditions. Weather prediction problem is being experienced when the world is currently generating large datasets which are produced and recorded enormously in virtually all fields including; weather recording, biomedical, social network, mobile network data, digital archives and electronic trading among others. This unanticipated amount of data provides unprecedented opportunities for data-driven decision making and knowledge discovery. Weather basically produces continuous data, large and multidimensional data which is dynamic and has unordered process of recording it. These properties make weather forecasting a formidable challenge. Forecasting is the process of estimating the unknown situation from historical data. However, the massive data and high dimensionality of big data introduces unique computational and statistical challenges, including scalability, storage bottle neck, noise accumulation, spurious correlation, incidental endogeneity, and measurement errors. The task of analyzing such large-scale data sets poses significant challenges and calls for innovative statistical methods specifically designed for faster speed, higher efficiency, and accurate predictions. These challenges are eminent and require a new computational and statistical paradigm shift. In spite of the explosion of this big data, specific tools are required for modelling, mining, visualizing, forecasting to understand these large data sets and make data-driven decisions. In many situations, it is easy to predict the outcome given the cause. However, in science, more often than not, we are faced with the question; when given the outcome of an experiment, what are the causes or the probability of the causes compared to other outcomes? Bayesian theory provides a framework for plausible

reasoning and a concept which is a more powerful and general tool for handling this problem. To apply Bayesian, the data is usually partitioned into training and testing sets, where the training set is used to develop a model and the testing set is for testing the developed model. This idea of Bayesian theory was championed by Jaynes (2013).

There has been a growing interest in applying immense data to many analytical areas, particularly in time series prediction. The primary model in multivariate time series analysis is the VAR. It is the mechanism that is used to link multiple time series variables together. In VAR models, each variable is a linear function of the past values of itself and the past values of all other endogenous variables. It is usually used in simultaneous prediction and structural analysis of a number of temporal observed variables. It is applied when each variable in the system does not only depend on its own lag alone, but also on the lags of other variables. The high-dimensional data sets in time series have become common in many areas like in geo-physics, biomedical, econometrics, and finance, among others. Most of the variables used are codependent and need to be interconnected to give information about a response variable. This cross-sectional dependency of variables brings a sharp focus to the problem of how to uniquely understand the interactions among the components of a large dynamic system from the data set. VAR is commonly used for studying high-dimensional interrelationships among the components of a multivariate time series in a system of equations. This study helps to integrate interdependent variables to develop a computational efficient model for VAR prediction using Bayesian model. One of the actively researched areas is the weather distribution pattern, about which the understanding is still in its early stages of inference. Numerous studies have been conducted to further the knowledge; but Bayesian methodology finds its place to aid in obtaining scientific inferences about certain facts from available data. This study



provides an account of VAR in Bayesian model data analysis applied to weather distribution with a particular focus on rainfall distribution patterns in Kenya. The GVAR model is obtained by integrating the regional model through link matrices.

### **1.2 Statement of the Problem**

Global warming has become a major challenge in the world. It has brought about unpredictable weather patterns which have been changing irregularly from the normal seasons. These, in most cases have caused extreme weather changes that are identified as major global challenges of recent times. In Kenya, unstable rainfall patterns, which are associated with global warming, have been experienced to a greater extent. Accurate rainfall prediction has become very complicated in recent times due to climate variability. Despite the availability of models that are used by the meteorology departments to make predictions, the same devastating scenarios of unpredictable weather changes are still being experienced. Therefore, robust models reliable for accurate predictions are needed on short and long term. Models for precise prediction of weather changes in Kenya are identified as the bases of concern that this study sought to address.

### **1.3 Significance of the Study**

Accurate weather forecast models are important in the developing countries, where most of the agricultural activities depends on weather. It is a major concern to identify any trends of weather parameters that deviate from its periodicity, which would disrupt the economic and non-economic activities of the country. This fear has been aggravated due to threats by the global warming effect. The effect of extreme weather phenomena in our country has led to increased cost of living, infrastructure damages, human and animal injuries as well as loss of life. Therefore, a great understanding of weather

patterns would help to avert this problem. Rainfall is an essential phenomenon in the weather system, which has a direct influence on the agriculture and biological sectors and it is virtually influenced by other weather components. The amount of rainfall in a given region is affected by several factors which include; temperature, atmospheric pressure, wind speed, relative humidity, radiation, and altitude, among others. Different prediction methods have been developed by researchers in different countries. However, robust models that gives accurate predictions are yet to be achieved. Weather variables provides inter dependency where one variable influences the result of other variables. Nevertheless, how do we deal with such a problem of variable interaction? The interdependence of variable that constitute response variables are generated in many areas. VAR method are used to handle the variable inter-linkages which results from large data sets. It represents the correlations among a set of variables, which are used to analyze certain aspects of the relationship between the variables of interest. Understanding the nature and scale of rainfall patterns in Kenya is a key factor to the policy makers because it gives them a chance to prepare for better mitigation and adaptation measures during adverse weather conditions. There have been several forecasting methods that different researchers have been using. However, progress in the operational forecasting of the rainfall at the National environmental prediction and other Numerical weather prediction (NWP) centers has been slow due to the complexity of predicting rainfall and limited computing resources available for the task. For now, rainfall is still not a direct model guidance product produced by NWP centers but is diagnosed by local forecasters based either on statistical methods such as model output statistics. For these purposes, Bayesian statistical analysis of time series of rainfall data is a very valuable tool for investigation its variability pattern and predicting short- and long-term changes in the rainfall patterns. The study also uses global vector

Autoregressive (GVAR) which played a key role. The main idea behind the GVAR framework is to incorporate inter-linkages between cross-sectional weather zones in a viable way. The integration of regional VAR models is used to formulate GVAR. The study was based on developing a model that improved the accuracy of predicting rainfall patterns in Kenya. This was done through the introduction of Bayesian approaches. Bayesian analysis was a useful technique which was used to detect the rainfall patterns and changes using the past data to give the present information about what would happen in future. Thus, a need to develop a better robust prediction model was necessary to overcome this global challenge. Most of the traditional methods employed were probabilistic models, they had challenges of clearly identifying the part of the weather signals that were due to change, making it complex to unravel. Consequently, the probabilistic models had weakness, where, they used fixed parameters along the forecasting duration. One way to overcome this weakness was through the use of Bayesian Models which treated the parameters of the models as random variables. This study used BVAR models to solve the problem of predicting rainfall patterns from the current weather variables. Accurate forecast of rainfall patterns would save lives, support emergency management teams, mitigate the impacts of damages and prevent economic losses, hence the significance of this study.

#### **1.4 Motivation of the Study**

The current knowledge discovery systems are nowadays supposed to handle very large data sets store, process and make deductions from them. When working with vast time series, multivariate prediction becomes more and more complicated because the use of all the variables does not allow to have the most accurate predictions and poses certain problems for classical prediction models. One of the main motivations behind this study is that, using all the available predictor variables does not necessarily yield to the best

forecast accuracy and sometimes renders some statistical prediction models non-solvable due to the high number of variables compared to the number of observations. Therefore, a robust model is required to solve this problem, filter the non-valuable parameters and improves the prediction accuracy of many time series of the big data sets. Bayesian method, due to its shrinkage ability would try and overcome this set of problems. Most of the discussions around climate change have focused on how much the earth would warm up over the coming century. However, climate changes are not limited just to temperature, but precipitation which is both rain and snow changes, would also have a great impact on the global activities. In the recent past Kenyans have experienced erratic weather conditions where the rainfall patterns no longer follow the normal cycles. This study considered a number of variables, they included; rainfall, which was the response variable, and the explanatory variables which were temperature, humidity, atmospheric pressure, wind speed, radiation, and wind gust. The response variables used in this study were considered to be the ones influencing rainfall patterns in Kenya. The effect of adverse weather changes in Kenya have caused huge losses in agricultural sector due to failed rainfall and even death of human and animals.

## **1.5 Objectives of the Study**

### **1.5.1 General objectives**

The general objective of this study was to develop a Bayesian vector Autoregressive model that would be suitable for predicting rainfall patterns in Kenya.

### **1.5.2 Specific Objectives**

The specific objectives were to:

- i. Perform diagnostic analysis of the weather variables.
- ii. Develop Bayesian Vector Autoregressive predictive weather model.

- iii. Conduct model performance Analysis.
- iv. Apply the model to forecast the Rainfall pattern in Kenya.

Accurate forecasts save lives, support emergency management and mitigation of impacts and prevent economic losses from high-impact weather, and they create substantial financial revenue—for example, in energy, agriculture, transport and recreational sectors. Their substantial benefits far outweigh the costs of investing in the essential scientific research, super-computing facilities and satellite and other observational programmes that are needed to produce such forecasts

Accurate forecasts save lives, support emergency management and mitigation of impacts and prevent economic losses from high-impact weather, and they create substantial financial revenue—for example, in energy, agriculture, transport and recreational sectors. Their substantial benefits far outweigh the costs of investing in the essential scientific research, super-computing facilities and satellite and other observational programmes that are needed to produce such forecasts

### **1.6 Scope of the Study**

Kenya lies within the Inter-Tropical Convergence Zone (ITCZ), a narrow belt of very high rainfall that forms near the equator. The ITCZ brings rain southwards through Kenya from October to December where it experiences ‘short rains’, passing again northwards in March, April and May where there is ‘long rains’. Sharon (2018). The timing of these two rainy seasons, how long they last, and their intensity varies from year to year. As is the case for the climate across the rest of sub-Saharan Africa, the surface temperature of the water in the Indian Ocean, atmospheric pressure, humidity, wind speed, among other weather factors greatly influences the rainfall pattern in Kenya. This surface temperature varies from year to year, Sharon (2018). The study

used secondary data, which was sourced from Trans- African Hydro-Meteorological Observatory (TAHMO) and Kenya Meteorological Stations. The data captured over a period of four years from 2014 June to June 2018. The data was collected in Kenya across five regions namely; Coastal, Arid, Semi - arid, Highlands and Lake regions. The data was in the form of daily recordings for at least five evenly distributed weather stations in each of their respective regions. The study considered the data for seven variables, which included; Rainfall, Temperature, Atmospheric pressure, Wind speed, Wind gust, Radiation and Relative humidity.

### **1.7 Limitation of the Study**

One of the limitation was use of secondary data, the time the data was collected and when data was analyzed and decisions made was a bit delayed.

## **CHAPTER TWO**

### **LITERATURE REVIEW**

#### **2.0 Introduction**

This chapter highlights what was done by other researchers in specified areas. It introduces the theories and models other researchers had based their research on. It deals with open issues, Autoregressive, Bayesian vector Autoregressive model, and specific areas that were used in the methodology and in data analyses. It also cites the lacuna that motivated this study.

#### **2.1 Weather Background Information**

Kenya has experienced equally prolonged droughts and intense flooding every year, Mary et al. (2018). Due to the increase in such extreme weather occurrences, the glaciers around Mount Kenya are disappearing, leading to the drying up of rivers and streams. These weather changes have led to harvest losses and food shortages, landslides due to floods, soil degradation, and loss of biodiversity, Otiende and Brian (2009). The diminishing water sources and erratic rainfall have also reduced the availability of water. Climate variability and seasonal changes have adversely affected economic and noneconomic sectors. This situation is expected to deteriorate in the near future. Presently, the weather forecast is solved through the help of Atmospheric Circulation Models (ACM). These are integrated by different weather services on a daily basis, normally on coarse-grained resolution grids which covers a wide geographical coverage. The ACMs describes several meteorological variables such as humidity, temperature, wind component, geo-potential, among others. However, meteorological phenomena like rainfall and temperature, normally vary more on local scales. Linacre and Geerts (2017), stated that Numerical Weather Prediction (NWP) is a simplified set of equations called the primitive equation, used to calculate the changes

of conditions. According to Lutgens and TarBuck (2019), the word “numerical” is misleading, all types of weather forecasting are based on some quantitative data and therefore could fit under this area. NWP is based on the fact that the gases of the atmosphere obey a number of known physical principles. Ideally, these physical laws could not be used to predict the future state of the atmosphere, given the current conditions. The large number of variables that is included when considering the dynamic atmosphere makes this task extremely difficult. Manipulating the large data sets and performing the complex calculations necessary to predict weather and make a resolution good enough to make the result useful requires the use of some of the most powerful tools. These tools need to be supported by enhancements in computing capabilities, steady advances in weather, and climate prediction which will occur at major operational centers across the world, Bauer et al. (2015). Complementing these advances in weather and climate prediction, there have been important milestones in advancing the science and operational infrastructure for prediction using big data source, David Donoho (2020). The first generations of dynamical seasonal forecast systems are implemented at operational centers in the mid-1990s, Ji et al., (2018); Routine weather and climate forecasts at the global and regional levels now provide information critical for the economic welfare of society and for mitigating losses of life and property. According to the State of the Climate in 2020, Blunden et al., (2018), since 1901, the mean annual global that is land and ocean, surface air temperature has warmed by 0.7–0.9° Celsius per century, and the rate of warming has nearly doubled since 1975 to 1.5–1.8° Celsius per century. A steady rise in temperature has triggered important changes in the frequency and intensity of extreme weather and climate events such as heat and cold waves, droughts, floods, hurricanes, and other effects over various parts of the globe according to Intergovernmental Panel on Climate Change, (IPCC),



(2019). These unprecedented long-term climatic changes have influenced sub-seasonal and seasonal-to-interannual variability and have a profound impact on the natural environment as well as on the life, health and well-being of human society, animals and the entire biodiversity, Coumou and Rahmstorf, (2012).

## **2.2 Time Series models**

Time series are sequential data that are mostly applied to model dynamic systems and processes, Stock & Watson (2020). In various real live fields, the capability to make accurate predictions is essential because they provide possible future information about the system under study. One of the main challenges in time series prediction is improving the forecasting model to attain accurate future values. According to Brockwell and Davis (2010), the first prediction models were univariate, which predicted a single time series based on its own lagged information which was historical. They stated that a univariate time series consists of a set of observations on a single variable, which makes a prediction based on the past observations. Later, multivariate models were introduced, which considered multidimensional time series and predicted each variable using its previous values and the previous values of the other predictive variables, Youssef, Hmamouche et al. (2020). This is considered as a way of improving accuracy in developing the new prediction models by varying the structure of the existing models on how to apply and analyze the past data. In the basis of time series prediction with many variables, a common goal is to detect the most independent and relevant predictors with regard to a given target time series. In the literature, several approaches were proposed to handle this problem. Despite the advantages of existing methods in the literature, there are still some problems when dealing with large time series models, where the distribution of the large data storage and processing is also a big challenge, Stock and Watson (2017). When testing, if a variable could be used to

forecast another variable, it is plausible to use an Autoregressive model by taking the past value of a time series and construct it as a regression model, Stock & Watson (2015). Weather changes refers to the effect of atmospheric alteration with time, which would create analytical models and an important field for climate change studies. A possible view of weather changes is a time-dependent random variable that is composed of trend, extremes and variability, Mudelsee (2019). The major task of the analysis is to use the data for estimating the parameters describing the trend, seasonal variability, and other components. Trend estimation quantifying weather changes, is of high priority and of essence in the world. It gives room for the use of a simple parametric trend models to be employed. Smoothing is also used in nonparametric regression; its idea is to cancel out high frequency variability. For time series data to be analyzed, it must be stationary or stable. Stationarity is achieved when a time series variable is independent of time, the variance and the autocovariance are not infinite, Verbeek (2018).

### **2.2.1 Augmented Dickey-Fuller test**

The test for stationarity of the variables in the model is done using the Augmented Dickey-Fuller-test (ADF-test). Moffatt (2020) said that Dickey-Fuller-test was named after American statisticians' David Dickey and Wayne Fuller, who developed the test. The test is used to determine whether a unit root is present in an Autoregressive model or not. When time series variables are independent of time, the autocovariance and variances are not infinite, then the time series variables are said to be stationary, Verbeek (2018). Furthermore, Stock & Watson (2015) state that when the probability distribution has no fluctuation over time, the time series variables are stationary and therefore it follows a random walk. Asteriou & Hall (2017), mentioned stationarity as an important criterion when using a VAR-model. If the time series is not stationary, the

results from the test would not be reliable. According to Stock & Watson (2015), differentiation is a good working technique to overcome the problem of non-stationarity. The ADF test is based on hypothesis testing where the null hypothesis states that the time-series variable is nonstationary. In such circumstance, the variables are differenced a number of times until the ADF test shows a 5% level of significant.

### **2.2.2 Granger Causality test**

When VAR-tests are conducted, Granger causality tests are required to check if there is a significant association between variables. Lütkepohl (2005) states that there is Granger causality if information from one endogenous time series gives a more accurate prediction of another endogenous time series, even though all other possible information is taken into account. Subsequently, Lütkepohl (2005) meant that the idea behind the Granger causality test is that the effect is generated by the cause and not the reverse. However, it is important to note that the test also identifies the direction of the association between variables and not only causality.

### **2.3 Model Estimation**

Time series models are used to forecast events based on verified past data. Not all models will yield the same results for the same data set, so it is critical to determine which model works best based on the individual time series. Once we have chosen the best model, we fit it on the entire training set and evaluate its performance on a separate testing set subsequent in time. Before the best model is attained, there are a number of things that need to be determined, this includes lag order, the model type and the performance measure of the model.

### **2.3.1 Lag order selection**

The dependence of a variable  $y$  on another variable  $x$  is hardly instantaneous, they respond to each other with a lapse of time called lag. When using a VAR-model, it is important to use the correct number of lags, simply because too many lags inflate the standard errors of coefficient estimates and this implies an increase in the forecast error. While omitting lags that should be included in the model may result in an estimation bias. The lag order selection model is used to investigate which number of lags are optimal for the study, Stock & Watson (2015). The importance of lag length determination is demonstrated by Braun and Mittnik (2013) who show that estimates of a VAR whose lag length differs from the true lag length are inconsistent as are the impulse response functions and variance decompositions derived from the estimated VAR. Lütkepohl (2017), indicated that overfitting that is selecting a higher order lag length than the true lag length, causes an increase in the mean-square forecast error of the VAR and that underfitting lag length often generates autocorrelated errors. Hafer and Sheehan (2020) found out that the accuracy of forecasts from VAR models varied substantially for alternative lag lengths. The number of lag length is frequently selected using an explicit statistical criterion such as AIC, SIC, BIC, or HQ, where the criterion with the least number is termed to produce the best model which makes a reliable estimation of the optimal amount of lag in the model, Stock & Watson, (2015). According to Liew et al. (2014), AIC is a good criterion to use when having a sample size of more than 60 observations. Consequently, this study compared the lags from the four criteria and used AIC to measure the optimal number of lags, since it had the least.

### **2.3.2 Vector Autoregressive**

Vector Autoregressive are stochastic processes which are linear multivariate time series models and are a generalization of univariate Autoregressive (AR) models. They are

based on the notion of interdependencies between lagged values of all variables in a given model. They are able to capture the joint dynamic interdependencies among multiple time series. According to Hannan and Deistler (2012), the VAR are commonly resorted to as tools for investigating dynamic effects of shocks. They perform very well in forecasting exercise, where in the canonical tools for the identification of linear dynamic systems are Vector-valued Autoregressive Moving Average (VARMA) and state-space representations. Lütkepohl (2015), advocated for the use of higher-order VAR over more general VARMA models due to numerous identification issues of the latter model class. Andrews (2016), gave a strong theoretical justification of such a modeling strategy which came from the famous Wald decomposition theorem. Which ensured that a large class of stationary processes could be represented as potentially infinite order VAR processes. A major function and common application of VAR models are predictions, Koop & Korobilis (2013). VAR-based forecasts have proven to be superior to many other methods, Bańbura et al. (2010). The GVAR models were proposed by Pesaran et al. (2014), and further developed by D'ees, et al. (2017), for analyzing the transmission of domestic and international shocks to business cycle fluctuations. In essence, the idea of the GVAR modeling approach is that each region could be modeled individually and estimated as a VAR of the dynamics zone variables which are then linked to each other through link matrix. However, unlike standard VAR models, each regional models are linked to the others by including zonal-specific variables. Hjort and Claeskens (2013), used the frequentist model averaging approach in forecasting combination. They used the frequentist model approach in model averaging to address the issue of model uncertainty. Their work was an extension of the traditional approach to model uncertainty which focused on model selection rather than model averaging. Pesaran and Smith. (2018), stated that the conditional specific models

in the GVAR models were structurally stable but the unconditional model used to generate forecasts could be subject to structural breaks. For this reason, they considered a situation whereby some or all models under consideration were subjected to structural breaks and different choices of estimation samples were acceptable. With that in mind, they also talked about, averaging each model over different sampling windows allowing for both model and estimation window uncertainty. In many situations, analyzing a time series in isolation is realistic; while in other cases univariate analysis are limited, as demonstrated by Campbell & Diebold (2018). They linked financial interest variables, including stock returns and default premium in a multivariate system that allowed shocks in one variable to be propagated to the others.

### **2.3.3 Bayesian Vector Autoregressive Model**

One of the major problems in VAR is the number of coefficients to be estimated, they increase quadratically with the number of included variables and linearly in the lag order. Such a dense parameterization often leads to inaccuracies with regards to out of sample forecasting and structural inferences, especially for higher-dimensional models. This phenomenon is commonly referred to as the curse of dimensionality. The cause behind reducing the number of predictors in a multivariate prediction model is generally based on two reasons; for some linear models, if the number of variables is large, the estimation of the parameters may not be achievable. Secondly, applying a prediction model with all available variables is not necessarily the best choice in terms of prediction accuracy, because of the existence of redundant information within the set of variables, Stock and Watson (2017). The Bayesian approach to estimating VAR models tackles this limitation by imposing additional structure on the model. Informative conjugate priors have been shown to be effective in mitigating the curse of dimensionality and allows for large models to be estimated, Bańbura et al. (2010). They

push the model parameters towards a parsimonious benchmark, reducing estimated error and improving out of sample prediction accuracy Koop and Korobilis (2013). Bayesian inference treats the VAR parameters as random variables, and it provides a framework to estimate “posterior” probability distribution of the location of the model parameters. This is done by combining information provided by a sample of observed data and prior information derived from a variety of source according to Silvia and Giovanni (2018). By providing such a framework, the Bayesian approach allows to incorporate prior information about the model parameters into post sample probability statements. The ‘prior’ distributions about the location of the model parameters summarizes pre sample information available from a variety of datasets sources, theoretical models, and or introspection. In the absence of pre sample information, Bayesian VAR inference can be thought of as adopting ‘non informative’ or ‘diffuse’ or ‘flat’ priors, that express complete ignorance about the model parameters, in the light of the sample evidence summarized by the likelihood function. Often, in such a case, Bayesian probability statements about the unknown parameters conditional on the data are very similar to classical confidence statements about the probability of random intervals around the true parameters value. For example, for a VAR with Gaussian errors and a flat prior on the model coefficients, the posterior distribution is centered at the maximum likelihood estimator (MLE), with variance given by the variance-covariance matrix of the residuals. In scientific data analysis, priors on the model coefficients do not incorporate the investigator’s ‘subjective’ beliefs, instead, they summarize stylized representations of the data generating process. Conditional on a model, these widely held standardized priors aim at making the likelihood-based description of the data useful to investigators with potentially diverse prior beliefs, *Sims* (2010). The most commonly adopted priors for VARs are the so called ‘Minnesota’

priors, Litterman, (2019). They express the belief that an independent random walk model for each variable in the system is a reasonable 'centre' for the beliefs about their time series behaviour. BVAR borrows heavily from vector Autoregressive model (VAR). The only difference is that VARs usually assumes linearity and parametric nature of any time series data. The VAR method have a tendency of overfitting and over-parameterization. But the Bayesian model are different from the classical estimation method, the basic idea of the Bayesian estimation method is to treat the parameters of the model to be estimated as random variables which follow a certain distribution. It is also required to give a prior distribution of the parameters to be estimated based on experience and combine it with the sample information. Bayes' theorem is used to calculate the posterior distribution of the parameters to be estimated, thereby obtaining the estimated values of the estimable parameter. Bayesian methods are currently experiencing an increased popularity in the sciences as a means of probabilistic inference, Malakoff, D. (2019). Among their advantages are the ability to include prior information, the ease of incorporation into a formal decision analytic context, the explicit handling of uncertainty, and the straight forward ability to assimilate new information. The Bayesian approaches has shown to be particularly useful for ecological models with poor parameter identifiability, Reichert and Omlin, (2017). The most general time series model is Box – Jenkins model which assumes that the time series is stationary. There are three stages in developing Box – Jenkins time series model; these are model identification, model estimation, and model validation. The problem with Box – Jenkins model is that for effective fitting of the model it requires at least a moderately long time series. Yang D, et al. (2015), recommended at least fifty observations, while many others recommended at least hundred observations. This problem was sorted out by the use of Bayesian inference. Bayesian vector



autoregressions are usually used for forecasting and structural analysis. Until recently, although most empirical work had considered only small systems with a few variables due to parameter propagation concern and computational restrictions, Chan and Joshua C. (2019). The parameters within Bayesian models are stochastic and are assigned appropriate probability distributions, Bernadinelli et al. (2016). Parameters are treated as random variables and probabilities are assigned to these parameters. Bayesian analysis has three components, namely, the prior distribution, likelihood and the posterior distribution. It improves on classical estimation in terms of the precision of estimators. The posterior distribution describes the behavior of the parameters after the data is observed and prior assumptions are made. There is a wide variety of extensions of VAR that takes into account important features of the data, Cogley and Sargent, (2018); Primiceri, (2015). Some recent papers have considered similar extensions for large BVARs. For example, Koop and Korobilis (2013), proposed an approximate method for forecasting using large time-varying parameter BVARs. Chan, J. and Eisenstat, E. (2019). estimated a Bayesian VARMA containing 12 variables. Carriero et al. (2016), suggested that BVAR with informative priors have often proved to be superior tools compared to standard frequentist. Giovanni Ricco (2018), VARs are highly parametrized Autoregressive models, whose number of parameters grows with the square of the number of variables times the number of lags included. These extensions are all found to be out performed by BVARs with homoscedastic and independent innovations. VARs tend to have a lot of parameters, and Bayesian methods that formally incorporate prior information to provide shrinkage are often found to greatly improve the forecast performance, Doan et al. (2014). Until recently, most empirical work has considered only small systems that rarely included more than five dependent variables. This has changed since the seminal work of Bańbura et al. (2010),

who found that large BVARs with more than 10 dependent variables forecast better than small VAR. This has generated a rapidly expanding literature on using large BVARs for forecasting; recent papers include Kapetanios et al. (2019), Large BVARs thus provide an alternative to factor models that are traditionally used to handle large data sets, Forni et al., (2018). According to Rue et al. (2015) the Bayesian Model Averaging approach was not considered an important aspect in economic and other application areas until the late 2000s. The reason BMA approach is gaining popularity is because of the presence of more powerful computers and dramatic increases in numerical methods such as Monte Carlo Markov-Chain model composition (MCMC) which enables researchers to overcome the troubles encountered while implementing BMA. The work of Hoeting et al. (2019), gave a tutorial on Bayesian Model Averaging. In their work, they presented several difficulties encountered while implementing Bayesian Model Averaging and discussed several solutions to these implementation difficulties. Raftery et al. (2017), offered two alternative approaches to Bayesian Model Averaging for linear regression models and described the “Occam’s window” which indicated a small set of models over which a model average could be computed. Secondly, they also described a Markov chain Monte-Carlo approach that directly approximated the exact solution. The result of these two approaches in the presence of model uncertainty are that they both provide better predictive performance than any single model, that is reasonably having been selected. Their work only concentrated on two procedures that accounted for model uncertainty in variable selection for linear regression models but did not address the uncertainty involved in the identification of outliers and in the choice of transformations in regression. To broaden the flexibility of the procedures addressed in Raftery et al. (2017), Hoeting et al. (2019), have extended Bayesian Model Averaging to include transformation, selection, and outlier

identification. The main advantage of incorporating BVAR is because Bayesian statistics, within a solid decision theoretical framework, incorporates a natural and principled way of combining prior information with data. Wasserman, (2014), this means that in our data analysis, we can incorporate past information about a parameter and form a prior distribution for current and future analysis and prediction. This usually follows from Bayes' theorem. In fact, Robert (2020) stated that "The important feature of a Bayesian approach is, thus, that Bayes estimators are derived by an eminently logical process: starting from the requested properties, summarized in the loss function and the prior distribution, the Bayesian approach derives the best solution satisfying these properties".

#### **2.3.4 Priors for Model Selection**

Proper informing prior beliefs is critical and hence the subject of much research. Typical in VAR models the first step is to pre-select the relevant variables to be included in the system and with how many lags. This procedure may be thought of as having assertive priors about which variables have non-zero coefficients in the system. The challenge is in selecting among an expansive set of potential models. Indeed, for a VAR with  $n$  endogenous variables,  $q$  additional potentially exogenous variables including a constant, and  $p$  lags, there are  $2^{(q+pn)n + n(n-1)/2}$  possible models. Jarociński and Maćkowiak (2017) proposed that, to select the variables to be included in the system by systematically assessing the posterior probability of 'Granger causal priority' in a BVAR with conjugate priors. Granger causal priority answers questions of the form "is variable 'y' relevant for variable 'x', after controlling other variables in the system?" The study provided a closed form expression for the posterior probability of Granger causal priority, and suggested that variables associated with high Granger causal priority probabilities can be omitted from a VAR with the variables of

interest. Alternatively, one can adopt priors that support model selection and enforce sparsity. A variety of techniques, including double exponential or Laplace prior, spike-and-slab prior, etc., have been adopted to handle this issue. Some recent theoretical and empirical contributions on this topic are in Bhattacharya et al. (2015), Griffin and Brown (2017), and in Huber and Feldkircher (2017). In multivariate setting, flat priors, which attempt not to impose a certain belief, yield inadmissible estimators and poor inferences Bańbura et al. (2010). Other uninformative or informative priors are necessary. Early contributions Litterman (1980) set priors and their hyperparameters in a way that maximizes out of sample forecasting performance over a pre sample. Del Negro and Schorfheide (2007) choose values that maximize the marginal likelihood. Bańbura et al. (2010) used the in-sample fit as decision criterion and control for overfitting. The contribution of Giannone et al. (2015) focuses on conjugate prior distributions, specifically of the Normal-inverse-Wishart (NIW) family. Conjugacy implies that the Maximum Likelihood is available in closed form, enabling efficient computation. The NIW family includes many of the most commonly used priors, Koop and Korobilis (2010); Karlsson (2013), with some notable exceptions. These include, amongst others, the steady-state prior Villani (2009), the Normal-Gamma prior, Griffin and Brown (2010); Huber and Feldkircher (2019), and the Dirichlet-Laplace prior, Bhattacharya et al. (2015). Many recent contributions focus on accounting for heteroskedastic error structures, Clark (2011); Kastner and Frühwirth-Schnatter (2014); Carriero et al. (2016). This may improve model performance, but is not possible within the conjugate set up and would complicate inference. There are differing opinions and philosophies on the best practices for choosing priors, Wolf et al. (2017), Banner et al. (2020). In ecology, a common practice is to assign so-called non-informative priors that effectively assign equal probability to all possible values using either

uniform or diffuse normal priors with large variances, Lemoine (2019). These priors allow Bayesian inference to proceed that is produce a posterior distribution, but with presumably limited influence of the priors, Lemoine (2019). An attractive feature of the Bayesian approaches is that the models are generative. This means that we can simulate potential data from the model so long as the parameters are assigned a proper probability distribution, Gelman et al. (2013). This feature is routinely used to check models and prior influence after fitting the data using the posterior predictive distribution, Lemoine (2019), Gelman et al. (2020), but also is used before seeing the data using the prior predictive distribution, Gabry et al. (2019)

### **2.3.5 Informative and noninformative priors**

Informative prior probability distributions incorporate information about the VAR parameters that is available before some sample is observed. Such prior information can be contained in samples of past data from the same or a related system, or can be formed from introspection, casual observation, and theoretical models. The first case is sometimes referred to as a ‘data-based’ prior, while the second as a ‘non data based’ prior. An important case arises when the prior probability distribution yields a posterior distribution for the parameters in the same family as the prior probability distribution function. In this case the prior is called a natural conjugate prior for the likelihood function, Nicenboim et al. (2021). In general, it has been shown that exponential distributions are the only class of distributions that admit a natural conjugate prior, due to these, having a fixed number of sufficient statistics that does not increase as the sample size increases according to, Gelman et al. (2020). Because the data is incorporated into the posterior distribution only through the sufficient statistics, formulas for updating the prior into the posterior are in these cases conveniently simple. Prior distributions can be expressed in terms of coefficients, known as hyperparameters,

whose functions are sufficient statistics for the model parameters. It is often useful to think of the hyperparameters of a conjugate prior distribution as corresponding to having observed a certain number of pseudo-observations with properties specified by the priors on the parameters. In general, for nearly all conjugate prior distributions, the hyperparameters can be interpreted in terms of ‘dummy’ or pseudo-observations. The basic idea is to add to the observed sample extra ‘data’ that express prior beliefs about the hyperparameters. The prior then takes the form of the likelihood function of these dummy observations. Hyperparameters can be either fixed using prior information, or associated to hyperprior distributions that express beliefs about their values. A Bayesian model with more than one level of priors is called a hierarchical Bayes model.

### **2.3.6 Time Varying Parameters**

Models that allow parameters to change over time are becoming increasingly popular in empirical research, in recognition of the fact that they can capture structural changes in those time series according to Belmonte, et al. (2014). In fact, it seems to be a common belief that the properties of many if not most time series have changed over time, and can change across phases of the time cycle. Model parameters either change frequently and gradually over time according to a multivariate Autoregressive process as in time-varying parameters VARs or they change abruptly and irregularly as in Markov-switching or structural-break models. Time-varying parameters in vector Autoregressive (TVP-VAR) differ from fixed-coefficient VARs in that they allow the parameters of the model to vary over time, according to some specified rules. TVP-VARs often include also stochastic volatility, which allows for time variation in the variance of the stochastic disturbances. Doan et al. (2014) were the first to show how estimation of a TVP-VAR with Litterman priors could be conducted by casting the VAR in state space form and using Kalman filtering techniques. This same specification

is in Sims (2010). Bayesian time varying parameter VARs have become popular in empirical situations like in macroeconomics following the work of Primiceri (2015) who provided the foundations for Bayesian inference in these models, and used then innovations in MCMC algorithms to improve on their computational feasibility.

### **2.3.7 Breusch Pagan and Breusch Godfrey Test**

The Breusch Pagan test is a diagnostics test of a regression model, where the goal is to understand if there is presence of heteroscedasticity. Heteroscedasticity is defined as a non-constant variance over a period of time. The test assumes that the model has homoscedastic, so if we fail to reject  $H_0$ , the test provides evidence supporting this hypothesis. If we end up rejecting the null hypothesis, we obtain evidence suggesting that there are heteroskedasticity in the regression model, Wooldridge (2015). The Breusch-Godfrey test to detect presence of higher order serial correlation in Autoregressive model. Autocorrelation is present in the data series if the error terms in the regression are serially correlated across time. The test assumes that there is no serial correlation, and the result of the test has similar properties as the Breusch-Pagan Test, Wooldridge (2015). White (1980), proposed a method for correcting the standard errors of the coefficients in the regression model, to produce heteroskedasticity consistent standard errors. This Theory states that a regression model suffering from heteroskedasticity may produce incorrect significance level for the different variables, through a misleading estimate of the included variables' standard errors. These standard errors have a tendency to be under predicted, resulting in increased chance of getting significant values, when this is not the case. Thus, the White correction produces robust standard errors, enabling hypothesis testing, Wooldridge, (2015).

## 2.4 Model Forecasting

Bayesian Vector Autoregressive (BVAR) is used to conduct both classical unconditional as well as conditional forecasts. Forecast for horizon  $h \geq 0$  of an empirical VAR(p) process are generated recursively according to Box and Jenkins (2008). Unconditional forecasts challenging those obtained from factor models in accuracy Giannone et al. (2015) and are used for a variety of analyses. Conditional forecasts allow for elaborate scenario analyses, where the future path of one or more variables is assumed to be known. They are a handy tool for analyzing possible realizations of policy-relevant variables. BVAR features a framework for identification schemes, with two of the most popular schemes currently available; namely short-term zero restriction and sign restriction. The former is also known as recursive identification and is achieved via Cholesky decomposition of the variance covariance vector (VCOV) matrix by Kilian and Lutkepohl (2017). Another method is the use of Impulse Response Functions (IRF) which are central tool for structural analysis. They provided insights into the behavior of weather systems and are another cornerstone of inference in VAR models. IRFs served as a representation of shocks hitting the system and are used to analyze the reactions of individual variables. The exact propagation of these shocks is of great interest, but a meaningful interpretation relies on proper identification. Additionally, identification via sign restrictions comes at the cost of increased uncertainty and a loss of precision for the resulting IRF. Another related tool for structural analysis is Forecast Error Variance Decomposition (FEVD) as stated by Yuriy & Byoungchan (2020). These methods call for the identification of the structural shocks by imposing a sufficient number of identification restrictions on a reduced-form linear vector Autoregressive (VAR) model. However, in many cases it is difficult to come up with adequate credible identification restrictions and thus the use of BVAR.



## 2.5 Critique of the Literature Review

Despite the advantages of prevailing methods in the literature, there are still some problems when dealing with large time series, where the distribution of the data storage and processing is also a challenge. A way of improving the forecast accuracy consists in developing new prediction models by changing the structure of existing models and how they analyze the history of data in order to make predictions. Another way seeks to focus on the other factors that influence the predictions, by considering this problem as a process where the application of the prediction models is just a step. As such, the forecast accuracy can be improved by many ways, for instance, through determining the most optimized structures of the prediction models with respect to the underlying set of predictors. Again, by improving the quality of the input data and lastly by adopting model matching techniques. The problem we are addressing in this study is the prediction of multivariate time series that contain many predictors, that is predictive variables. The main question we are dealing with is how to interrelate interdependent variables and develop an accurate model for forecasting and that can select the subset from the predictors set that allows to obtain the best forecasts for a given target variable. Climate models can only predict a range of possible future scenarios of an individual weather variable. However, Kenya's climate keeps on changing, the most affected is the rainfall pattern which this study addressed. Dezfuli et al. (2015), while the model that have been used by climate scientists generally tried to agree on how different parts of the Earth were changing in temperature, there is much less agreement about where and how precipitation would change. The effects of high temperature bring about fast rate of evaporation, high humidity, and dry surface in the long run, which potentially contribute to the intensity and prolonged duration of drought. However, this increased moisture would not fall evenly across the planet. Some areas would receive increased

rainfall, while other areas are expected to receive less due to shifting weather patterns and other factors. Traditionally, statistical inferences were typically based only on available poorly recorded data sets, and little systematic structure existed to facilitate the combination of results that formed previous studies. To our best knowledge, virtually no direct solution has been offered in the academic literature for the problem of estimation of VAR models lacking equal sized time series. However, in general, modeling of scarce data has benefitted a lot from a rigorous interest of researchers. For a detailed reflective, readers are referred to Kapetanios, et al. (2019), Stock and Watson (2017). Based on these contributions, Bayesian approach to overcome data issues such as noise, errors, and uncertainty has been widely recognized. One obvious reason is that Bayesian methods allow for subjective probabilistic judgments to be included in deriving inferences from data. This stands in unambiguous contrast to the traditional frequentist school of statistical inference which relies on conclusions largely drawn from pure data observations. Quiet often there are a number of events or scenarios where current situational is usually determined by previous events or previous data. As such, prior information is usually carried in the current data and or future information data. This implies that, most Bayesian Vector Autoregressive Model (BVAR) are included essentially to capture prior distributions and also improve out-of-sample performance. Vector autoregression (VAR) models are broadly used to model time series. The main difficulty experienced with these models is the issue of handling a large number of parameters which results to overfitting problems, as stated in most literature. They also require that the time series to have equal lengths in the estimation process, which would induce a loss of potentially valuable information coming from time series that are longer than others. Traditional VARs are a fairly restrictive number of variables versus the number of observations in available samples. With increasing

number of variables and lags, the number of coefficients estimated in VARs quickly increased. This brings another problem of overfitting the data. However, restricting just to a few variables means that the researchers can neglect the information embodied in many other time series that are available because many thousands candidate series are being developed. Most of them bring a limited additional gain in terms of forecasting performance. Several methods are developed to "shrink" the model which allows estimation, but still it has been a big challenge to handle a larger amount of time series. To overcome this difficulty, a Bayesian VAR approach is employed. For unequal time series, the remedy is through Bayesian approach; the information in longer time series is aggregated into prior, which is then used in the estimation of parameters for the VAR process. The BVAR model is a vector autoregression model using the Bayes Theorem based on prior and posterior distribution, therefore, it is simply a VAR model with priors introduced to control coefficients of the variables. As opposed to the point estimators, means and variances used by classical statistics, Bayesian statistics is concerned with generating the posterior distribution of the unknown parameters, given both the data and some prior density for these parameters. As such, Bayesian statistics provides a much more complete picture of the uncertainty in the estimation of the unknown parameters. The central idea of the Bayesian method is the use of study data to update the state of knowledge about the quantity of interest that has been studied. This idea in the Bayesian approach is a very intuitive one, namely, that of updating knowledge. The state of knowledge about the quantities of interest before or prior to a study is updated by the current study data, which yielded the state of knowledge after or posterior to the study. The transformation from prior to posterior is achieved by Bayes Theorem, an explicit mathematical expression for the updating process. Since from the review of the previous literatures, there is still a need for robust models to

handle the challenges in VARs and other classical models for weather prediction. Thus, these motivated the need to explore the idea of the prior-to-posterior transformation by considering the weather data set in Kenya and using BVAR model, hence the study.

## CHAPTER THREE

### METHODOLOGY

#### 3.0 Introduction

This chapter highlights the various methods that were used to achieve the stated objectives in chapter one. The objectives were formulated and their methods of achieving them were developed. The chapter also states some assumptions and theories that were considered when developing the methodologies. In this section, the researcher briefly discussed: sources of data, diagnostic analysis, vector Autoregressive, and Bayesian model. Bayesian Vector Autoregressive (BVAR) constitutes a special type of multivariate time series regression model in a system of equations for estimation and predictability of the data. The data set consisted of daily observations for a period of four years, where the variable of interest was the rainfall.

#### 3.1 Data Preprocessing

The data was stored in the form of excel format, which was captured on a daily basis for a period of four years, starting from June 2014 to June 2018. The data was arranged in column format for each variable and averaged on daily basis. The data was cleaned and the missing values were filled by cluster means. The data was converted into command delimited .csv files so that to import it into R-statistical software. Normalization was done through linear scaling technique to remove scaling. It was essential because all variables used different units of measurements, and also a variable may have a large impact on the prediction value, only because of its numerical size. The technique of linear scaling, which is also referred to as min-max normalization estimations was defined as:

$$x_N = \frac{x - \text{Min}(x)}{\text{Max}(x) - \text{Min}(x)}$$

Normalization transformed the data into a common range of between 0 and 1. Thus, removing the scaling effects from all variables.

### **3.2 Diagnostic Test of the Weather Variables**

This part deals with some preparatory work to ensure that the data are ready for further analysis. This was done through several tests, which are introduced below. Before analyzing any time series data, the data should attain some level of stability. Stationarity was checked using Augmented Dickey fuller (ADF) test. Most of the multivariate time series move simultaneously and they have interdependent with one another which need to be tested. The variables were tested for their causal effect relationship by use of Granger Causality test.

#### **3.2.1 Stationarity Tests**

Stationarity means that the variables included in the analysis have means, variances, and covariance that are constant over time. This implies that each of these variables are equal and independent of what period they represent. For any, time series analysis to be conducted and the data be used to develop a model, the stationarity test is important. If this condition is violated and nonstationary data are used, then the outcome would result in an unpredictable model outcome. The results obtained when using nonstationary data is likely to be unreliable and the outcome may indicate that there is a relationship between variables whereas it does not exist. Many time series data in reality are not stationary and they require to be stationary to be analyzed. Non-stationarity can be detected by visual examining of the time series graph or, by looking at the series correlogram, or by conducting a unit root statistical test. To remove non-stationarity, a time series is transformed by differencing once or several times, until it

becomes stationary. In this study, the unit roots statistical test was employed under Augmented Dickey- Fuller (ADF) test

### 3.2.2 Augmented Dickey-Fuller test

The model variables were tested for stationarity by use of ADF test. The ADF is an extension of Dickey-Fuller test which basically assumes that the time series in question is an Autoregressive AR(1) process.

$$\text{AR}(1) \text{ is of the form } x_t = \alpha + \beta x_{t-1} + u_t$$

That is, the time series is a function of itself lag one period past.

$$H_0: \beta = 1 \text{ not stationary}$$

$$H_1: \beta < 1 \text{ stationary}$$

To make the series stationary we subtract  $x_{t-1}$  on both sides to have

$$x_t - x_{t-1} = \alpha + \beta x_{t-1} - x_{t-1} + u_t$$

$$\Delta x_t = \alpha + \rho x_{t-1} + u_t \text{ where } \rho = \beta - 1$$

$\Delta$  delta notation for differencing,

The hypothesis changes to

$$H_0: \rho = 0$$

$$H_1: \rho < 0$$

Note that  $\Delta x_t$  is converted to be stationary. This is tested by comparing the Dickey-Fuller critical value against the t-statistic.

$$t_{\hat{\rho}} = \frac{\hat{\rho}}{se(\hat{\rho})} \quad \text{we compare the t-value with DF- value which is obtained}$$

from DF distribution

$$t_{\hat{\rho}} < DF_{\text{critical}} \quad \text{we reject } H_0 \text{ (we reject that it has a unit root)}$$

$$t_{\hat{\rho}} > DF_{\text{critical}} \quad \text{we do not reject } H_0$$

If the absolute value of the t-statistic is larger than the critical value, we can say that the time series is stationary. The Augmented Dickey-Fuller test allows for higher-order Autoregressive processes by including  $\Delta x_{t-p}$  in the model.

$$x_t = \alpha + \sum_{i=1}^p \phi_i x_{t-i} + u_t$$

$$\Delta x_t = \alpha + \rho x_{t-1} + \sum_{i=1}^p \phi_i x_{t-i} + u_t$$

Taking  $p = 1$  then VAR(1) process is stable if all eigen values of  $\phi_i$  have a modulus less than one, this stability condition is equivalent to

$$\det(\mathbf{I}K - \phi_i z^i) \neq 0 \text{ for } |z| \leq 1$$

this can be generalized as

$$\det(\mathbf{I}K - \phi_i z^i) = \det(\mathbf{I}K - \phi_1 z^1 - \phi_2 z^2 \dots - \phi_p z^p) = 0$$

It gives the definition of the characteristic polynomial of a matrix. The polynomial is called the reverse characteristic polynomial of the VAR (p) process. The characteristic roots are the values of  $z$  that are obtained after solving this equation. There are  $p$ 's of them, although some of them may be equal.  $x_t$  is stationary if all of the roots "lie outside the unit circle". Since stability implies stationarity, the process is stationary when proved to be stable. The test follows AR (1) process



$$x_t = \rho x_{t-1} + u_t,$$

where,  $u_t$  is identically independently distributed series of random variables.  $x_t$  is non-stationary under the null hypothesis, and is stationary under the alternative hypothesis. The standard t-statistics would not follow t-distribution because of the non-stationarity of  $x_t$  under the null hypothesis. To test the null hypothesis, the test statistics equation below was used

$$ADF = \frac{\rho - 1}{S.e(\rho)}$$

The ADF test of this equation follows the assumption that the error terms are independent and identically distributed, without a drift in the model. The procedure for the ADF test is similar to the Dickey–Fuller test procedure, the only difference is the model where it is applied. The models where ADF was applied is as shown below.

$$\Delta X_t = \alpha + \beta t + \phi_1 x_{t-1} + \delta_1 \Delta x_{t-1} + \dots + \delta_{p-1} \Delta x_{t-1} + \varepsilon_t$$

where,  $\alpha$  denotes a constant,  $\beta$  is the coefficient on a time trend, and  $p$  represents the lag order of the Autoregressive process. Putting the constraints  $\alpha = 0$  and  $\beta = 0$ , this resembles a model with a random walk, and using the constraint  $\beta = 0$  resembles a model of random walk with a drift. The ADF test is performed under the hypothesis

$$H_0: \phi = 1 \text{ against } H_1: \phi < 1$$

The test statistic is computed as:

$$ADF = \frac{\hat{\phi}}{SE(\hat{\phi})}$$

If the  $ADF$  ( $\tau$ ) test statistic is less than the critical value, then the null hypothesis of  $\phi = 0$  is rejected and no unit root is present. When the null hypothesis is not rejected, it means that the time series is not stationary and requires at least differencing once or until when the augmented dickey fuller (ADF) test shows a 5% level significant result.

### 3.2.3 Granger Causality

The basis behind vector autoregression is that each of the time series in the system influences each other. The future prediction for the univariate time series is done by means of its past values. In multiple time series, some of the time series may be influencing each other in their prediction, such situation is referred to as granger causality. Granger causality test is done to a certain, if there is a significant association between different time series model variables. The idea behind the granger causality test is the flow of the information and the effects from one time series to the other time series. A time series  $x_{(t)}$  granger causes  $y_{(t)}$  if the past values of  $x_{(t)}$  help to improve the prediction of the future values of  $y_{(t)}$ . That is  $y_{(t)}$  is a function of its lagged values and the lagged values of  $x_{(t)}$ . This implies that  $y_t = f(y_{t-p}, x_{t-p})$

The following conditions were to be satisfied for the granger causality to be applied;

1. Cause happens prior to effect

$y_t = f(y_{t-p}, x_{t-p})$  where  $y_t$  is the effect and  $x_{t-1}$  is the cause and not  $x_t$  in the same period.

2. Cause has unique information about the future values of its effects

$$y_t = a_1 y_{t-1} + a_2 x_{t-1} + \mu_t$$

therefore  $x_{t-1}$  has an extra effect other than that of  $y_{t-1}$

This relationship needs to be tested before building the model using granger's causality test. What does granger's causality really test? granger's causality tests the null hypothesis that the coefficient of past values in the regression equation are zero. In simple terms, the past values of the time series  $x_{(t)}$  do not cause the other series  $y_{(t)}$ .

$$y_t = a_0 + a_1 y_{t-1} \quad \text{is the first model}$$

$$y_t = a_0 + a_1 y_{t-1} + a_2 x_{t-1} + \mu_t \quad \text{the second model}$$

$$H_0: a_2 = 0 \quad \text{against} \quad H_1: a_2 \neq 0$$

If  $a_2$  is significant then  $x_{t-1}$  adds extra value to the first model. That means, in the presence of  $y_{t-1}$  then  $x_{t-1}$  granger causes  $y_t$ . The t-test is used to test for the significant of the individual coefficients and the f-test to test them jointly. Consequently, if the p-value obtained from the test is smaller than the significance level of 0.05, then the null hypothesis is rejected and we conclude that  $x_{(t)}$  granger causes  $y_{(t)}$ . The test for correlation in most cases does not necessarily imply causality, which makes it difficult to determine whether one variable causes the other. Generally, this difficult in correlation is solved by the use of granger causality where the present or future event is believed to have been caused by the past event. This is impetuous of the granger causality test on time series data which gives evidence that one variable causes the other.

The test is constructed on the following OLS regression model

$$y_i = a_0 + \sum_{j=1}^m \alpha_j y_{i-j} + \sum_{j=1}^m \beta_j x_{i-j} + \varepsilon_i$$

Here, the  $\alpha_j$  and  $\beta_j$  are the regression coefficient and  $\varepsilon_i$  is the error term.

The test is based on the null hypothesis:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_M = 0 \text{ against } H_1: \text{At least one } \beta_i \text{ is not zero}$$

If the p-value for the test is less than the designed value of alpha, then reject the null hypothesis and conclude that granger causality exists. It is assumed that the data is stationary before using the granger causality test.

### 3.3 Developing Specific Model

The model was developed for each specific zone and for the global vector. The zones included;

- zone one – Coast region, zone two – Arid region, zone three – Semi Arid, zone four – Highland
- zone five – Lake region

Each model had seven endogenous variables where six of them were independent and one was a predictor variable. They include

- $x_0$ – Rainfall (Predictor)
- $x_1$ – Atmospheric pressure
- $x_2$ – Radiation
- $x_3$ – Relative humidity
- $x_4$ – Temperature
- $x_5$ – Wind gust

$x_6$ – Wind speed

VAR model was used to determine the relationship between different endogenous variables. The model was a multi-equation system where all the variables were treated as dependent variables. Whereby there was one equation for each variable as dependent variable. The equation included the lagged values of all the weather variables in the system. Primarily,  $x_0$  – Rainfall was set as dependent in the model and each of the other variables were tested separately after which Bayesian vector Autoregressive was applied.

### 3.3.1 Vector Autoregressive Model Specification

Vector Autoregressive (VAR) is a linear multivariate time-series model, able to capture the joint dynamics of multiple time series. They constitute a type of multivariate time series which is applied to examine the active interrelationship between stationary time series variables. VAR model is an extension of univariate to multivariate time series data. It is a multisystem of equations where all variables are treated as endogenous. Model selection is an important integral part of the statistical analysis of VAR. Model selection is made up of two parts; determination of the lag orders, also known as the lag length, and the determination of the structure of the VAR coefficient matrices. The study considered a column vector of k-different variables  $x_t = [x_{1t}, x_{2t}, \dots, x_{kt}]'$  and modeled them, in terms of past values of the vector. The results were a vector Autoregressive of order p or a VAR (p) process which was of the form.

$$x_t = \alpha + B_1x_{t-1} + \dots + B_px_{t-p} + e_t$$

Where:

$x_t$  is a  $k \times 1$  vector of dependent variables

$x_{t-i}$  is a  $k \times N$  matrix which is the number of observations times ( $\times$ ) the number of endogenous variables in the stochastic process

$\alpha$  is a  $k \times 1$  vector of intercept parameters,

$B_1$  through  $B_p$  is a  $k \times k$  matrices of coefficients,

$e_t$  is a  $k \times 1$  vector of white noise process,

and  $P$  is the lag order.

The study considered the following assumptions

i)  $E(e_t) = 0$  for all  $t$

ii)  $E(e_t e_s) = \begin{cases} \sum_e & \text{for } s = t \\ 0 & \text{for } s \neq t \end{cases}$

The covariance matrix  $\sum_e$  was assumed to be a finite positive definite matrix. The lag operators ( $\Delta$ ) were used to develop coefficient matrix. The  $VAR(p)$  process was written in lag operator notation form, which was defined as  $\Delta x_t = x_{t-1}$ . This means that it lags (shifts back) by one period. Using this operator, the equation above was written as

$$x_t = \alpha + (B_1\Delta + B_2\Delta^2 + \dots + B_p\Delta^p)x_t + e_t$$

Or

$$B(\Delta)x_t = \alpha + e_t$$

Where  $B(\Delta) = I_k - B_1\Delta - B_2\Delta^2 - \dots - B_p\Delta^p$

This study considered the model of the form

$$y_t = \alpha + B_1 y_{t-1} + \dots + B_p y_{t-p} + a_1 x_{1t-1} + \dots + a_p x_{1t-p} + \dots + k_p x_{6t-1} + \dots + k_p x_{1t-p} + e_t$$

In this study the matrix formation equation was of the form.

$$\begin{bmatrix} x_{0t} \\ x_{1t} \\ x_{2t} \\ x_{3t} \\ x_{4t} \\ x_{5t} \\ x_{6t} \end{bmatrix} = \begin{bmatrix} \alpha_{0t} \\ \alpha_{1t} \\ \alpha_{2t} \\ \alpha_{3t} \\ \alpha_{4t} \\ \alpha_{5t} \\ \alpha_{6t} \end{bmatrix} + \begin{bmatrix} B_{00} & B_{01} & B_{02} & B_{03} & B_{04} & B_{05} & B_{06} \\ B_{10} & B_{11} & B_{12} & B_{13} & B_{14} & B_{15} & B_{16} \\ B_{20} & B_{21} & B_{22} & B_{23} & B_{24} & B_{25} & B_{26} \\ B_{30} & B_{31} & B_{32} & B_{33} & B_{34} & B_{35} & B_{36} \\ B_{40} & B_{41} & B_{42} & B_{43} & B_{44} & B_{45} & B_{46} \\ B_{50} & B_{51} & B_{52} & B_{53} & B_{54} & B_{55} & B_{56} \\ B_{60} & B_{61} & B_{62} & B_{63} & B_{64} & B_{65} & B_{66} \end{bmatrix} \begin{bmatrix} x_{0t-1} \\ x_{1t-1} \\ x_{2t-1} \\ x_{3t-1} \\ x_{4t-1} \\ x_{5t-1} \\ x_{6t-1} \end{bmatrix} + \begin{bmatrix} e_{0t} \\ e_{1t} \\ e_{2t} \\ e_{3t} \\ e_{4t} \\ e_{5t} \\ e_{6t} \end{bmatrix}$$

The generalized form is as equation below

$$x_{it} = \alpha_{it} + B_{ij} x_{it-1} + e_{it}$$

Where “ $x_{it}$ ” represented the endogenous variables which are independent of a time period “ $t$ ” for a specific zone “ $i$ ”,  $\alpha_{it}$  is the constant term of the models,  $e_{it}$  represented the white noise error terms and  $B_{ij}$  are the vector matrix of the coefficient variables.

where  $i, j = 0, 1, \dots, 6$ .

There is one equation for each variable as the dependent variable. In the reduced form, the right-hand side of each equation included the lagged values of all dependent variables in the system, with no contemporaneous variables. In the application of VAR model, the study employed the following assumptions;

- i) All variables were endogenous
- ii) Time series were stationary which were contained stochastic processes
- iii) The lag order was determined was sufficient for the suitable model.

### 3.3.2 Vector Autoregressive Order

If  $x_t = \alpha + B_1x_{t-1} + B_2x_{t-2} + \dots + B_px_{t-p} + e_t$  is a  $VAR(p)$  process, it is useful to fit the model to the available multiple time series with  $p + 1$  coefficient. If  $x_t$  is a  $VAR(p)$  process, in this sense it is also a  $VAR(p + 1)$  process, therefore, we call  $x_t$  a  $VAR(p)$  process if  $Bp \neq 0$  for  $1 \leq p$  and  $Bp = 0$  for  $l > p$ , so that  $p$  is the smallest possible order. The most popular method to choose the lag order is to use information criteria. An information criterion is designed to consistently find the model that fits better the data from a group of models. As for how many lag orders to be included, this is determined by Akaike Information Criteria (AIC), which is generally given as

$$AIC = 2k - 2\ln(\hat{l})$$

Where  $k$  is the number of the parameters in the model and  $\hat{l}$  is the maximum value of the likelihood function for the model. AIC is generally used to estimate the information loss and the best model is the one that loses the lowest information.

The BIC is calculated by

$$BIC = \log(n) * k - 2 * L$$

where  $n$  is the number of observations of the model variables,  $k$  is the number of covariates used in the model, and  $L$  is the average of the log likelihood chain returned by the function. The main advantage of BIC is its consistent as  $n \rightarrow \infty$ . Alternatively, this is where lags are dropped until the last lag is statistically significant.



### 3.3.3 Estimation of a Vector Autoregressive (p) model

In the equation of  $VAR(p)$ , where  $e_t \sim IID(0, \Omega)$ , if  $x_{it}$  denotes the  $i^{th}$  element of  $x_t$  and  $B_{j,ki}$  denoted the  $k_i^{th}$  element of  $B_j$ , in the  $i^{th}$  row of the equation which is written as

$$x_{it} = m_i + \sum_{k=1}^m \sum_{j=1}^p B_{ki,j} x_{k,t-j} + e_{it}$$

This is just a linear regression in which  $x_{it}$  depended on a constant term and *lag* 1 through  $p$  of all  $m$  variables in the system. Because the same variable appears on the right-hand side of the above equation for all  $i$ , the Ordinary Least Square Estimator for each equation is identical to Generalized Likelihood Squares Estimators.

### 3.3.4 The Global Vector Autoregressive Model

The GVAR model is a multi-zonal framework which allows the investigation of interlinkages among the zones in the country, which are modeled by the corresponding VAR models. The basic building blocks of the global model are the local zonal models. In a GVAR the model comprises two layers via which enables the specification to capture cross-zone links. The first layer separates multivariate time series models, one per climatical zone. In the second layer, the zone models are stacked to yield a global model that is able to trace the longitudinal spread as well as its sequential dynamics. Each of this zonal model are  $VARX(p_i, q_i)$  model, meaning that it is a VAR model with  $p_i$  lags of the endogenous variables and  $q_i$  lags of the set of exogenous variables. The zonal model for the unit  $i$  is to examine the endogeneity of the external variables  $x_{it}^*$ , to solve the entire global model. Stacking over the zonal models is written as

$$x_t = b_0 + b_1 t + \Phi_1 x_{t-1} + \dots + \Phi_p x_{t-p} + \Lambda_0 w x_t + \Lambda_1 w x_{t-1} + \dots + \Lambda_p w x_{t-p} + e_t$$

Where;

$$x_t \text{ is } Nk \times 1$$

$$b_0 \text{ is } Nk \times 1$$

$$b_1 \text{ is } Nk \times 1$$

$$\Phi_1 \dots \Phi_p \text{ is } Nk \times Nk$$

$$x_{t-1} \dots x_{t-p} \text{ is } Nk \times 1$$

$$\Lambda_0, \Lambda_1, \dots, \Lambda_p \text{ is } Nk \times Nk$$

The solution of the stacked model is obtained as

$$x_t = (I_{kN} - \Lambda_0 w)^{-1} (b_0 + b_1 t + \Phi_1 x_{t-1} + \dots + \Phi_p x_{t-p} + \Lambda_1 w x_{t-1} + \dots + \Lambda_p w x_{t-p} + e_t)$$

provided the  $e_t$  are independent in the time dimension, then the endogeneity of the regressors  $w x_t$  follows the from

$$E(w x_t e_t) = w (I_{kN} - \Lambda_0 w)^{-1} E(e_t e_t')$$

Pesaran et al. (2004) noted that the weight matrices  $w_{ij}$  are diagonal with

$$w_{ij} = \text{diag}(w_{ij}^1, \dots, w_{ij}^k)$$

and that

$$\sum_{j=0}^N (w_{ij}^m)^2 \rightarrow 0, \quad \text{as } N \rightarrow \infty, \text{ for all } i \text{ and } m$$

However, this implies that asymptotically external variables have no explanatory power in the model. Asymptotic properties of such models are not used as sample guidance

for our estimators if some degree of cross-sectional dependences were expected in the model.

The assumption

$$\sum_{j=0}^N |w_{ij}^m| \leq c < \infty, \text{ for all } i \text{ and } m,$$

where the constant  $c$  does not depend on the sample size  $N$ . This is clearly a weaker assumption, but it turns out to be strong enough to allow us to derive asymptotic properties of our model.

### 3.3.5 The Global Vector Autoregressive framework

To build a simple version of our Global Vector Autoregression model from each zonal model.

All variables are collected from all zones which creates the global vector.

$$x_t = \begin{pmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{Nt} \end{pmatrix}$$

Which is a  $k \times 1$  vector containing all endogenous variables, where  $k = \sum_{i=1}^N k_i$ .

Following the step that give rise to the zones single model above, the identity below is obtained

$$Z_{it} = w_i x_t$$

For  $i = 1, \dots, N$ , where  $w_i$  is a zonal-specific link matrix of dimensions  $(k_i + k_i^*) \times k$  constructed on the basis of weather variable weights. This identity allows writing each region model in terms of the global vector. By substituting the above equation in the zones single model, the model below is obtained

$$A_i w_i x_{it} = a_{i0} + a_{i1}t + B_{i1} w_i x_{i,t-1} + \dots + B_{ip} w_i x_{i,t-p} + e_t$$

The individual zone model is then stacked, yielding the model for all variables in the global model  $x_t$  to have

$$Gx_t = a_0 + a_1t + \sum_{j=1}^p H_j x_{t-j} + e_t$$

where

$$G = \begin{pmatrix} A_{1,0}w_1 \\ \vdots \\ A_{N,0}w_N \end{pmatrix}, H_j = \begin{pmatrix} B_{1,j}w_1 \\ \vdots \\ B_{N,j}w_N \end{pmatrix}, a_0 = \begin{pmatrix} a_{1,0} \\ \vdots \\ a_{N,0} \end{pmatrix}, a_1 = \begin{pmatrix} a_{1,1} \\ \vdots \\ a_{N,1} \end{pmatrix}, e_t = \begin{pmatrix} e_{1,t} \\ \vdots \\ e_{N,t} \end{pmatrix}$$

Pre-multiplying the global model equation by  $G^{-1}$  yields an Autoregressive representation of the Global Vector Autoregression(p) model shown below

$$x_t = b_0 + b_1t + \sum_{j=1}^p F_j x_{t-j} + \varepsilon_t$$

where

$$F_j = G^{-1}H_j, b_0 = G^{-1}a_0, b_1 = G^{-1}a_1 \text{ and } \varepsilon_t = G^{-1}e_t$$

This equation is treated like any other VAR equation of order p.

### 3.3.6 Bayesian models

Bayesian inference treats the VAR parameters as random variables, and provides a framework to update probability distributions about the unobserved parameters conditional on the observed data. This brings out all its different from frequentist statistics. Classical estimation may yield imprecisely estimated relations that fit the data well only because of the large number of variables included. This results to a problem

known as overfitting which has been greatly discussed in the literature. In fact, the number of parameters to be estimated,  $n(np + d)$ , grows geometrically with the number of variables ( $n$ ) and proportionally with the number of lags ( $p$ ), while ( $d$ ) is the number of exogenous variables. The concept behind the Bayesian modeling frame is Bayes' theorem, which states that any prior beliefs regarding an uncertain quantity are updated, based on new information, to yield a posterior density of the unknown quantity. This process is also referred to as the principle of inverse probability. The concept of Bayesian statistics is that the posterior distribution of a parameter  $\theta$  given data that is proportional to the likelihood of the data given the parameter multiplied by the prior distribution for  $\theta$ , that is,

$$p(\theta/\text{data}) \propto L(\text{data}/\theta)p(\theta)$$

### 3.3.7 BVAR Model Development

Let  $x_t$  be an  $n \times 1$  random vector that takes values in the domain of real numbers. The evolution of  $x_t$  the endogenous variable is described by a system of  $p - th$  order difference equations in the VAR(p):

$$x_t = \alpha + B_1x_{t-1} + \dots + B_px_{t-p} + e_t$$

The vector of  $e_t$ , are independent and identically distributed random variable for each  $t$ . the distribution from which  $e_t$  is drawn which determined the distribution of  $x_t$ , conditional on its past

$$x_{1-p:t} = \{x_{t-p}, \dots, x_0, \dots, x_{t-2}, x_t\}.$$

The standard assumption is that the errors are Gaussian.

$$e_t \sim iid. N(0, \Sigma).$$

This implies that the conditional distribution of  $x_t$  is also Normal. Bayesian inference on the  $x_t$  model amount to updating prior beliefs about the VAR parameters, that are seen as stochastic variables, after having observed a sample

$$x_{1-p:t} = \{x_{t-p}, \dots, x_0, \dots, x_{t-2}, x_t\}.$$

The prior beliefs about the VAR coefficients are summarized by a probability density function and updated using Bayes' Law.

$$p(A, \Sigma/x_{1-p:t}) = \frac{p(A, \Sigma)p(x_{1-p:t}/A, \Sigma)}{p(x_{1-p:t})} \propto p(x_{1-p:t}/A, \Sigma)$$

Define  $A = [A_1 \dots \dots A_p]'$  as a  $k \times n$  matrix of the parameters, with  $k = np + 1$ .  $p(A, \Sigma/x_{1-p:t})$  is the joint posterior distribution of the  $VAR(p)$ .  $P(A, \Sigma)$  is the prior distribution of the parameter. Assuming that the model coefficients are random variables. The prior distribution is specified based on prior information and combined with objective information from the observed data to obtain the posterior distribution using Bayes theorem. The prior distribution specification provides shrinkages preventing the estimated parameters from depicting what are only spurious correlations. The initial information about the model parameters and the sample information is the likelihood function  $p(x_{1-p:t}/A, \Sigma)$ . The posterior distribution summarizes the entire information available and is used to conduct inference on the VAR parameters. Under the assumption of Gaussian error, the conditional likelihood of VAR is

$$p(x_{1-T}/A, \Sigma, x_{1-p:0}) = \prod_{t=1}^T \frac{1}{(2\pi)^{1/2}} |\Sigma|^{-1} \exp \left\{ -\frac{1}{2} (x_t - A'x'_t)' \Sigma^{-1} (x_t - A'x_t) \right\}$$

where

$$x'_t = [x'_{t-1} \dots \dots \dots x'_{t-p}]$$

The likelihood in this equation is written in compact form, by using the apparently unrelated regression representation of the VAR.

$$x_t = Ax + e_t$$

Using this notation and standard properties of the trace operator, the conditional likelihood function is equivalently expressed as

$$p(x_{1:T}/A, \Sigma, x_{1-p:0}) = \frac{1}{(2\pi)^{1/2}} |\Sigma|^{-1} \exp\left\{-\frac{1}{2} \text{tr}[\Sigma^{-1} \widehat{S}]\right\} X \exp\left\{-\frac{1}{2} \text{tr}[\Sigma^{-1}(A - \widehat{A})' x' x (A - \widehat{A})]\right\}$$

Where  $\widehat{A}$  is the maximum likelihood estimator (MLE) of  $A$ , and  $\widehat{S}$  is the matrix of sums of squared residuals that is  $\widehat{A} = (x'x)^{-1}x'x_t$ ,  $\widehat{S} = (x_t - x\widehat{A})'(x_t - x\widehat{A})$

The likelihood is written in terms of the vectorized representation of the VAR

$$x_t = (I_n \otimes x) \alpha + e, \quad e \sim (0, \Sigma \otimes I_T)$$

Where  $x_t = \text{vec}(x)$  and  $e = \text{vec}(e)$  are  $Tn \times 1$  vectors, and  $\alpha = \text{vec}(A)$  is  $nk \times 1$ .

In this vectorized notation, the likelihood function is written as

$$\begin{aligned} p(x_{1:T}/A, \Sigma, x_{1-p:0}) \\ = \frac{1}{(2\pi)^{Tn/2}} |\Sigma|^{-T/2} \exp\left\{-\frac{1}{2} \text{tr}[\Sigma^{-1} \widehat{S}]\right\} X \exp\left\{-\frac{1}{2} (\alpha - \widehat{\alpha})' \Sigma^{-1} \otimes (x'x) (\alpha - \widehat{\alpha})\right\} \end{aligned}$$

Where, consistently,  $\hat{\alpha} = \text{vec}(\hat{A})$  is  $nk \times 1$ . The likelihood function is used to update the prior information regarding the VAR parameters. An interesting case arises when we assume the absence of any information on the location of the model parameters, which can be formalized by assuming that  $A$  and  $\Sigma$  are independently distributed i.e.

$$P(A, \Sigma) = P(A) P(\Sigma) \text{ with prior p.d.f as}$$

$$P(A) \propto \text{constant}$$

$$P(\Sigma) \propto |\Sigma|^{-(n+1)/2}$$

Given this prior we can derive the posterior distribution of the VAR parameters as

$$p(A, \Sigma / x_{1-p:0}, \dots) = |\Sigma|^{-(T+n+1)/2} \exp\left\{-\frac{1}{2} \text{tr}[\Sigma^{-1} \otimes I_T] (x - (I_n \otimes x) \alpha) - (I_n \otimes x) \alpha\right\} \\ \times \exp\left\{-\frac{1}{2} (\alpha - \hat{\alpha})' \Sigma^{-1} \otimes (x'x) (\alpha - \hat{\alpha})\right\}$$

The model was used to forecast the data  $h > 0$  steps ahead.

### 3.3.8 Bayesian model in GVAR

Bayesian approach is employed where information believed to be derived from background knowledge which is used to select a prior probability distribution for the model parameters. The predictions of future observations are made by integrating the models' predictions with respect to the posterior parameter distribution obtained by updating the prior to take account of the data. Bayesian analysis of the GVAR model requires the elicitation of the prior distributions for all parameters of the model. The study uses Minnesota prior structure that were developed for VAR specification for individual Zonal-specific models together with standard prior settings for the parameters corresponding to weakly exogeneous variables and combines the posterior results to obtain the Bayesian Global Vector Autoregressive (B-GVAR) model.



Although other key prior distributions are proposed in literature, the choice of the Minnesota prior in this study was due to its flexibility.

### 3.4 Model performance Analysis

The sensitivity analysis was done using confusion matrix.

#### 3.4.1 Confusion Matrix

A confusion matrix is a summarized table of the number of correct and incorrect predictions or actual and predicted values yielded by a classification model for binary classification tasks. The Confusion Matrix is used to evaluate the sensitivity, accuracy, and specificity of the model.

		ACTUAL VALUES	
		Positive	Negative
PREDICTED VALUES	Positive	TP	FP
	Negative	FN	TN

where: TP = True Positive, TN = True Negative,

FP = False Positive, FN = False Negative

From the confusion matrix, Accuracy, Sensitivity, and Specificity were evaluated using the following equations.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}$$

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

$$\text{Specificity} = \frac{TN}{TN+FP}$$

### **3.4.2 Training and test data sets**

The data was divided into two sets, training data set, which was the material through which the models were built. It was obtained under simple random sampling 80% of the data set, which was used to build the statistical model. The second set was the testing, and this set was used to predict the results. Data points in the training set are excluded from the testing set. The model is ultimately being trained to predict results for which the study does not have the answer. The size of the testing set is typically about 20% of the total sample, although this value depends on how long the sample is and how far ahead you want to forecast. Because the testing data is not used in determining the forecasts, it provides a reliable indication of how well the model is likely to forecast with new data. It is also important to evaluate the forecasted accuracy using the unaffected developed forecasting model. The accuracy of forecasts is only determined by considering how well a model performs on new data that are not used when fitting the model. The following points are noted: a model which fits the training data well will also tend to forecast well. A perfect fit is always obtained by using a model with enough parameters. Over-fitting a model to data is just as bad as failing to identify a systematic pattern in the data. Some references describe the testing set as the “hold-out set” because these data are “held out” of the data used for fitting. Other references call the training set “in-sample data” and the testing set “out-of-sample data.” This study prefers to use “training data” and “testing data”.

### **3.5 Forecasting Analysis**

Sequential analysis is more useful through Bayesian, where the updated information is used to generate further information, and the processes keep on and on. In Bayesian, the prior gives the scope to obtain posterior information. The obtained posterior is used

as the prior for the next posterior generation. The sequential analysis merged with Bayesian gives more consistent results in comparison to the classical methods. In contrast, Bayesian assumes the random effect of the parameter of interest to generate the conclusion about the trial. Whereas, the chances of getting the biased result due to the consideration of the random effect of the parameter of interest are certainly low in Bayesian. This ability in Bayesian assists in the predictability of the future or the current situation. If time series observations are available for a variable of interest and the data from the past contains information about the future development of the available, it is plausible to use a forecast of some function of the data collected in the past. As forecasting is one of the main objectives of multiple time series analysis. Forecast for horizon  $h \geq 0$  of an empirical VAR(p) process are generated recursively.

$$Y_{T+h/T} = A_1 Y_{1+h-1/T} + \dots + A_p Y_{1+h-p/T}$$

$$Y_{T+j/T} = Y_{T+j} \quad \text{for } j < 0$$

Where;

$$\text{Cov} \left( \begin{bmatrix} Y_{T+1} - Y_{T+1/T} \\ \vdots \\ Y_{T+h} - Y_{T+h/T} \end{bmatrix} \right) = \begin{bmatrix} I & 0 & \dots & \dots & 0 \\ \Phi_1 & I & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \Phi_{p-1} & \Phi_{p-2} & \dots & \dots & I \end{bmatrix} (\Sigma_U \otimes I_h) \begin{bmatrix} I & 0 & \dots & \dots & 0 \\ \Phi_1 & I & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \Phi_{p-1} & \Phi_{p-2} & \dots & \dots & I \end{bmatrix}$$

The matrices  $\Phi_i$  are the empirical coefficient matrices of the Wald moving average representation of a stable VAR(p) - process and the operator  $\otimes$  is the Kronecker product.

### 3.5.1 Evaluation of predictive model

The predictive performance is measured on the basis of Root Mean Square Error (RMSE) as defined below. The Root Mean Square Error (RMSE) is defined by the

square root of the mean squared distance between the predicted target values and the observed target value.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{l=1}^n (x_t - \hat{x}_t)^2}$$

Where,  $\hat{x}_t$ , is the forecasted value in the period  $t$ ,  $x_t$  is the actual value in the period  $t$ , and  $n$  is the size of the sample. RMSE measures the difference between the values predicted by a Bayesian model and the observed true values. In other words, it measures the quality of fit between the actual data and the predicted model. For the two measures above, the smaller the value, the better the fit of the model.

In evaluating the accuracy of forecasts, frequently used measures of forecast accuracy are employed to assess the performance of the models. These measures are independent of the scale of the data and are Mean Percentage Error (MPE), Mean Standard Error (MSE), and Mean Absolute Error (MAE). The formulas for these measures of accuracy are given as:

$$\text{MPE} = \frac{1}{n} \sum_{l=1}^n |x_t - \hat{x}_t|$$

$$\text{MSE} = \frac{1}{n} \sum_{l=1}^n (x_t - \hat{x}_t)^2$$

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

Where  $\hat{x}_t$ , is the forecasted value in the period  $t$ ,  $x_t$  is the actual value in the period  $t$ , and  $n$  is the size of the sample and  $e_t$  is the error term.

## **CHAPTER FOUR**

### **RESULTS AND DISCUSSIONS**

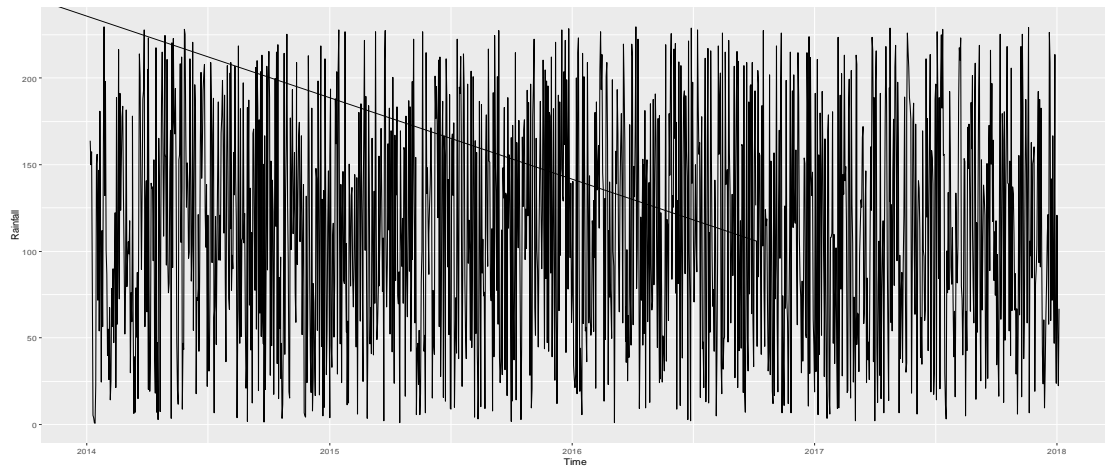
#### **4.0 Introduction**

This chapter deals with data analysis and development of the specific model for each zone and the global vector. The statistical methods employed in the analysis of the data were discussed in Chapter 3. The focus of this chapter was on the application of the methodologies discussed earlier and the analysis and interpretation of the results that generated forecasts of the stated specific and global zones. It shows that BVAR gives a robust model which resulted to a better forecasting.

#### **4.1 Diagnostic analysis of the variables**

To ensure that the time series data contained no flaws, was stable, and not affected by serial correlation, diagnostic analysis was put into use. To achieve these, unit root test was carried out where Augmented Dickey Fuller and Phillips Perron were conducted to ascertain the stability of the data. The Granger causality test was also conducted to find the influence of other time series to the dependent variable time series. The first step was to obtain the time plot graph for the endogenous variable to have a clear visualization of the data.

## Time plot graph for the rainfall data

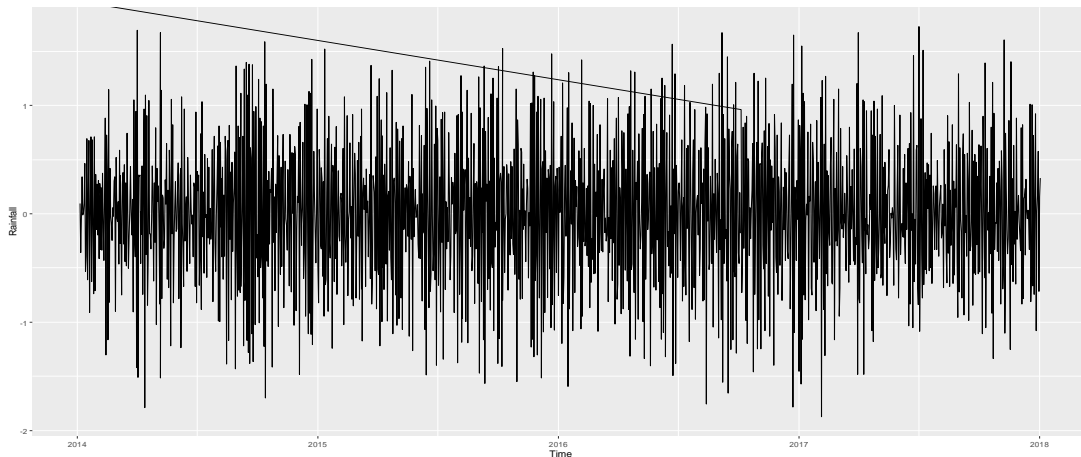


**Figure 4.1: Time series graph for zone one**

It is clear from the above plot that exhibits a time series in nature which contains seasonal variation and shows nonstationary. We difference each of the series appropriate number of times till stationarity is achieved. We carry out the Augmented Dickey Fuller ADF which is the unit root test and Phillips perron test on the differenced series to be certain that they are converted to stationary series.

### 4.1.1 Stationary test

The study adopted two methods of stationarity test the Augmented Dickey fuller (ADF) and Phillips perron test. The output from the test revealed that the data were stationary after at least first differencing for all regions. The obtained results were presented for each region, tabulated and placed in Appendix I. The graph obtained after differencing is displayed in figure 4.2 below.



**Figure 4.2: Time series graph after differencing**

This graph exhibits stationarity where mean and the variance are constant.

#### **4.1.2 Augmented Dickey fuller (ADF) and Phillips-Perron (PP) test**

The Augmented Dickey-Fuller (ADF) test was implemented to check whether the variables were stationary or not. The test assumed that the variables were affected by the unit root, which implied that the variables were nonstationary. The test is used to reject or accept the null hypothesis of the ADF tests. If the alternative hypothesis is accepted then the data is stationary, which was the desired result of the test. In this case, if the absolute value of the test statistics is less than the critical value and  $p\text{-value} < 0.05$  the null hypothesis is rejected which means that a time series does not have a unit root, thus stationary. That means it does not have a time dependent structure. The Phillips-Perron (PP) test was used to confirm the results of ADF test.

#### **4.1.3 ADF and PP test for zone one**

The result of the stationarity test in zone one was listed below and presented in Table 1 in appendix I.

The variable  $x_0$  had an ADF Test Statistics of -28.164, with Phillips-Perron of -273.58, Truncation lag parameter was 3, P-Value ADF of 0.01 and the P-Value P.P of 0.0127 thus showing that it was stationary. Variable  $x_1$  showed ADF Test Statistics of -27.153

with Phillips-Perron of -272.51, Truncation lag parameter of 3, P-Value ADF of 0.0121 while the P-Value P.P was 0.0103 hence it was stationary. Variable  $x_2$  showed ADF Test Statistics of -27.879, Phillips-Perron of -79.12, Truncation lag parameter of 3, P-Value ADF of 0.0213 and the P-Value P.P was 0.0182 thus showing that it was stationary. Variable  $x_3$  showed that ADF Test Statistics was -28.808 with Phillips-Perron of -283.58, Truncation lag parameter was 3, P-Value ADF was 0.0031 while the P-Value P.P was 0.0323 thus showing that unit roots exist. Variable  $x_4$  showed ADF Test Statistics of -29.758 with Phillips-Perron of -297.84, Truncation lag parameter of 3, P-Value ADF of 0.0274 and the P-Value P.P of 0.0145 thus showing that it was unit roots exist. Variable  $x_5$  showed ADF Test Statistics of -28.762 with Phillips-Perron of -286.58, Truncation lag parameter of 3, P-Value ADF of 0.0015 and the P-Value P.P of 0.0113 and hence Heteroscedasticity is impacting the results. Variable  $x_6$  showed ADF Test Statistics of -28.415 with Phillips-Perron of -290.42, Truncation lag parameter of 3, P-Value ADF of 0.0093 and the P-Value P.P of 0.0393.

The result show that after the first differencing the data was stable as reflected by ADF test and confirmed by PP test where their unit root had the p-values  $< 0.05$  which shows that they were significant.

#### **4.1.4 Zone two stationarity test.**

The result is as stated below and presented in Table 2 in appendix I

The Variable  $x_0$  had an ADF test Statistics of -11.019, with Phillips-Perron of -36.716, Truncation lag parameter of 3, P-Value ADF of 0.0156 and the P-Value P.P of 0.0124 thus showed that it was stationary. Variable  $x_1$  showed ADF Test Statistics of -12.984, with Phillips-Perron of -36.762, Truncation lag parameter of 3, P-Value ADF of 0.0126 and the P-Value P.P of 0.0173 thus showed that it was stationary. Variable  $x_2$  showed



ADF Test Statistics of -10.066 with Phillips-Perron of -37.444, Truncation lag parameter of 3, P-Value ADF of 0.0524 the P-Value P.P of 0.0473 hence it was stationary. Variable  $x_3$  showed ADF Test Statistics of -10.249 with Phillips-Perron of -35.759, Truncation lag parameter of 3, P-Value ADF of 0.0026 while the P-Value P.P was 0.0149 thus showing that it was stationary. Variable  $x_4$  showed ADF Test Statistics of -11.108 with Phillips-Perron of --11.108, Truncation lag parameter of 3, P-Value ADF of 0.0138 and the P-Value P.P of 0.0231 hence it was stationary. Variable  $x_5$  showed ADF Test Statistics of -37.055 with Phillips-Perron of -10.715, Truncation lag parameter was 3, P-Value ADF of 0.0361 and the P-Value P.P was 0.0302 thus showing that it was stationary. Variable  $x_6$  showed ADF Test Statistics of -10.715 with Phillips-Perron of -35.215, Truncation lag parameter was 3, P-Value ADF of 0.0054 while the P-Value P.P was 0.0215 thus showing that it was stationary.

The test shows that  $x_2$  is weakly stationary according to ADF but it has been confirmed by PP, thus it was treated to be stable, hence the analysis proceeded as per the study.

#### **4.1.5 Zone three**

The results for zone three were as displayed below and was tabulated in table 3 indicates in appendix I. The variable  $x_0$  had an ADF test Statistics of Test Statistics of -11.173, with Phillips-Perron of -36.636, Truncation lag parameter of 3, P-Value ADF of 0.0155 and the P-Value P.P of 0.0352 thus showing that it was stationary. Variable  $x_1$  showed ADF Test Statistics of -12.238 with Phillips-Perron of -36.714, Truncation lag parameter of 3, P-Value ADF of 0.0268 while the P-Value P.P was 0.01 hence it was stationary. Variable  $x_2$  showed ADF Test Statistics of -10.472 with Phillips-Perron of -37.385, Truncation lag parameter was 3, P-Value ADF of 0.0246 and the P-Value P.P of 0.0165 thus showing that it was stationary. Variable  $x_3$  showed ADF Test Statistics

of -10.907 with Phillips-Perron of -36.66, Truncation lag parameter was 3, P-Value ADF of 0.0276 and the P-Value P.P of 0.0103 thus showing that it was stationary. Variable  $x_4$  showed ADF Test Statistics of -11.555 with Phillips-Perron of -36.763, Truncation lag parameter was 3, P-Value ADF was 0.0371 while the P-Value P.P was 0.0124 hence it was stationary. Variable  $x_5$  showed ADF Test Statistics of -11.633 with Phillips-Perron of -36.884, Truncation lag parameter was 3, P-Value ADF of 0.0188 and the P-Value P.P of 0.01 hence it was stationary. Variable  $x_6$  showed ADF Test Statistics of -11.207 with Phillips-Perron of -35.193, Truncation lag parameter of 3, P-Value ADF of 0.0101 and the P-Value P.P of 0.0297 thus showing that it was stationary. The test result shows stationarity in all the variables.

#### **4.1.6 Zone four**

The results in zone four was as represented below and also in Table 4 Appendix I; Variable  $x_0$  showed ADF Test Statistics of -28.245 with Phillips-Perron of -273.61, Truncation lag parameter of 3, P-Value ADF of 0.01 and the P-Value P.P of 0.0242 thus showing that it was stationary. Variable  $x_1$  showed ADF Test Statistics of -28.327 with Phillips-Perron of -274.36, Truncation lag parameter of 3, P-Value ADF of 0.0260 while the P-Value P.P was 0.0173 thus showing that it was stationary. Variable  $x_2$  showed ADF Test Statistics of -27.892, Phillips-Perron of -277.82, Truncation lag parameter was 3, P-Value ADF of 0.0124 and the P-Value P.P was 0.0732 thus showing that it was stationary. Variable  $x_3$  showed ADF Test Statistics of -28.507, Phillips-Perron of -28.507, Truncation lag parameter was 3, P-Value ADF was 0.0121 and P-Value P.P was 0.0493, this showed stationarity. Variable  $x_4$  showed ADF Test Statistics of -29.518 with Phillips-Perron of -301.02, Truncation lag parameter was 3, P-Value ADF was 0.0386 and the P-Value P.P was 0.0231 thus showing that it was stationary. Variable  $x_5$  showed ADF Test Statistics of -29.222 with Phillips-Perron of -287.64,

Truncation lag parameter was 3, P-Value ADF of 0.0361 and the P-Value P.P was 0.0302 thus showed stationarity. Variable  $x_6$  showed ADF Test Statistics of -28.453, with Phillips-Perron of -291.18, Truncation lag parameter was 3, P-Value ADF of 0.0142 and the P-Value P.P was 0.0215 thus showing that it was stationary.

The stationarity of zone four was obtained after second differencing. The decision was reached after most of the variables turned out to be nonstationary.

#### **4.1.7 Zone five**

The test result is as given below and is presented in appendix I in Table 5 variable  $x_0$  showed ADF Test Statistics of -11.508 with Phillips-Perron of -37.064, Truncation lag parameter was 3, P-Value ADF of 0.01 and the P-Value P.P was 0.0173 thus showed that it was stationary. Variable  $x_1$  showed ADF Test Statistics of -12.403 with Phillips-Perron of -36.616, Truncation lag parameter was 3, P-Value ADF of 0.0387 and the P-Value P.P was 0.0158 thus stationary. Variable  $x_2$  showed ADF Test Statistics of -10.33 with Phillips-Perron of -37.838, Truncation lag parameter was 3, P-Value ADF of 0.0237 and the P-Value P.P was 0.01 thus showing that it was stationary. Variable  $x_3$  showed ADF Test Statistics of -10.901 with Phillips-Perron of -36.79, Truncation lag parameter was 3, P-Value ADF of 0.0275 and the P-Value P.P was 0.0253 thus showing that it was stationary. Variable  $x_4$  showed ADF Test Statistics of -11.438 with Phillips-Perron of -36.762, Truncation lag parameter was 3, P-Value ADF of 0.0421 and the P-Value P.P was 0.0149 thus showing that it was stationary. Variable  $x_5$  showed ADF Test Statistics of -10.702 with Phillips-Perron of -37.208, Truncation lag parameter was 3, P-Value ADF of 0.0173 and the P-Value P.P was 0.0334 thus showing that it was stationary. Variable  $x_6$  showed ADF Test Statistics of -10.586 with Phillips-

Perron of -35.248, Truncation lag parameter was 3, P-Value ADF of 0.0192 and the P-Value P.P was 0.0146 thus showing that it was stationary.

All the variables were found to be stationary after first differencing

#### **4.1.8 Global vector**

Table 6 in appendix I contains the tests in tabulated form Variable  $x_0$  showed ADF Test Statistics of -11.606 with Phillips-Perron of -37.085, Truncation lag parameter was 3, P-Value ADF of 0.01 and the P-Value P.P was 0.01 thus showing that it was stationary. Variable  $x_1$  showed ADF Test Statistics of -11.54 with Phillips-Perron of -37.063, Truncation lag parameter was 3, P-Value ADF of 0.0379 and the P-Value P.P was 0.01 thus showing that it was stationary. Variable  $x_2$  showed ADF Test Statistics of -10.315 with Phillips-Perron of -37.456, Truncation lag parameter was 3, P-Value ADF of 0.0242 and the P-Value P.P was 0.0186 thus showing that it was stationary. Variable  $x_3$  showed ADF Test Statistics of -10.888 with Phillips-Perron of -36.3, Truncation lag parameter was 3, P-Value ADF of 0.0274 and the P-Value P.P was 0.0101 thus showing that it was stationary. Variable  $x_4$  showed ADF Test Statistics of -11.35 with Phillips-Perron of -36.638, Truncation lag parameter was 3, P-Value ADF of 0.0298 and the P-Value P.P was 0.0199 thus showing that it was stationary. Variable  $x_5$  showed ADF Test Statistics of -11.57 with Phillips-Perron of -36.931, Truncation lag parameter was 3, P-Value ADF of 0.0151 and the P-Value P.P was 0.0437 thus stationary. Variable  $x_6$  showed ADF Test Statistics of -11.088 with Phillips-Perron of -35.423, Truncation lag parameter was 3, P-Value ADF of 0.0228 while the P-Value P.P was 0.0382 thus showing that it was stationary.

The global vector variables were stationary as show in the R output above.

#### 4.1.9 Granger causality test.

This was a statistical test for determining whether one time series was useful in forecasting another model. It was conducted for each specific zone and for the global vector. The data was also tested for Granger causality between endogenous and exogenous variables. This analysis was important in that it helped to know whether changes in one variable was influenced by changes in the other variables. This study tested whether exogenous variables had a causality effect onto the endogenous variable in each of the zones.

#### 4.1.10 Zone one

The Granger causality test findings were as represented below and presented in table 7 appendix I model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_1, 1:6)$  and Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.0491 \*. For  $x_2$  against  $x_0$ , Granger causality test shown that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_2, 1:6)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.21703. For  $x_3$  against  $x_0$ , Granger causality test shown that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_3, 1:6)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.006851 \*\*. For  $x_4$  against  $x_0$ , Granger causality test shown that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_4, 1:6)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.02144 \*. For  $x_5$  against  $x_0$ , Granger causality test shown that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_5, 1:6)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.31223. For  $x_6$  against  $x_0$ , Granger causality test shown that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_6, 1:6)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.01279 \*.

Therefore, the Granger causality test for Zone one shows that  $x_2$  and  $x_5$  had no significant influence on the endogenous variable  $x_0$  but the other variable had a significant influence.

#### 4.1.11 Zone two

The test is also presented in Table 8 appendix I

Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_1, 1:6)$  Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.01012 \*. For  $x_2$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_2, 1:6)$  Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.8279. For  $x_3$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_3, 1:6)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.003519 \*\*. For  $x_4$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_4, 1:6)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.04426 \*\*\*. For  $x_5$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_5, 1:6)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.4767. For  $x_6$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:6) + \text{Lags}(x_6, 1:6)$  Model 2:  $x_0 \sim \text{Lags}(x_0, 1:6)$  while the p value was 0.03723 \*.

Except  $x_2$  and  $x_5$ , all the other variables were having a strong significant influence on the causality of the variable  $x_0$ . Their level of significant was less than 0.05, but  $x_2$  had a p-value of 0.8279 and  $x_5$  had a p-value of 0.4767.

#### 4.1.12 Zone three

Table 9 have the tabulated results in appendix I.

Model 1:  $x_0 \sim \text{Lags}(x_0, 1:9) + \text{Lags}(x_1, 1:9)$  Model 2:  $x_0 \sim \text{Lags}(x_0, 1:9)$ , while the p value was 0.03396 \*. For  $x_2$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:9) + \text{Lags}(x_2, 1:9)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:9)$ , while the p value was 0.1548. For  $x_3$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:9) + \text{Lags}(x_3, 1:9)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:9)$  while the p value was 0.004577 \*\*. For  $x_4$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:9) +$

Lags( $x_4$ , 1:9), Model 2:  $x_0 \sim \text{Lags}(x_0, 1:9)$ , while the p value was 0.04553 \*. For  $x_5$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:9) + \text{Lags}(x_5, 1:9)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:9)$  while the p value was 0.04184 \*. For  $x_6$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:9) + \text{Lags}(x_6, 1:9)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:9)$ , while the p value was 0.0007054 \*\*\*.

The test shows that all exogenous variables had a strong causality on the variable  $x_0$ .

#### 4.1.13 Zone four

The test is as shown below and is also tabulated in Table 10 appendix I

For  $x_1$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:5) + \text{Lags}(x_1, 1:5)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:5)$  and the p value was 0.03119 \*. For  $x_2$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:5) + \text{Lags}(x_2, 1:5)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:5)$  and the p value was 1.032e-05 \*\*\*. For  $x_3$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:7) + \text{Lags}(x_3, 1:7)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:7)$  and the p value was 0.006269 \*\*. For  $x_4$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:7) + \text{Lags}(x_4, 1:7)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:7)$  and the p value was 0.04352 \*. For  $x_5$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:7) + \text{Lags}(x_5, 1:7)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:7)$  and the p value was 0.026 \*. For  $x_6$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:7) + \text{Lags}(x_6, 1:7)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:7)$  and the p value was 0.002428 \*\*.

Accordingly, the Granger causality test for Zone four shows that all variables had a p-value  $< 0.05$ , hence causality was positive.

#### 4.1.14 Zone five

The result is as stated below and in Table 11 appendix I.

Model 1:  $x_0 \sim \text{Lags}(x_0, 1:5) + \text{Lags}(x_1, 1:5)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:5)$  and the p value was 0.02655 \*. For  $x_2$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:5) + \text{Lags}(x_2, 1:5)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:5)$  and the p value was 0.0005519 \*\*\*. For  $x_3$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:5) + \text{Lags}(x_3, 1:5)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:5)$  and the p value was 0.04434 \*. For  $x_4$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:5) + \text{Lags}(x_4, 1:5)$  Model 2:  $x_0 \sim \text{Lags}(x_0, 1:5)$  and the p value was 0.000196 \*\*\*. For  $x_5$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:5) + \text{Lags}(x_5, 1:5)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:5)$  and the p value was 0.003104 \*\*. For  $x_6$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:5) + \text{Lags}(x_6, 1:5)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:5)$  and the p value was 0.001708 \*\*.

The variables had some influence to the dependent time series.

#### 4.1.15 Global vector granger causality test

The test result is tabulated in table 12, appendix I.

For  $x_1$  against  $x_0$ , Granger causality test showed Model 1:  $x_0 \sim \text{Lags}(x_0, 1:3) + \text{Lags}(x_1, 1:3)$

Model 2:  $x_0 \sim \text{Lags}(x_0, 1:3)$  and the p value was 0.01859 \*. The test for  $x_2$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:3) + \text{Lags}(x_2, 1:3)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:3)$  and the p value was 0.7194. For  $x_3$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:3) + \text{Lags}(x_3, 1:3)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:3)$  and the p value was 0.01464 \*. For  $x_4$  against  $x_0$ , Granger causality test showed



that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:3) + \text{Lags}(x_4, 1:3)$  Model 2:  $x_0 \sim \text{Lags}(x_0, 1:3)$  while the p value was 0.006 \*\*\*. For  $x_5$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:3) + \text{Lags}(x_5, 1:3)$ . Model 2:  $x_0 \sim \text{Lags}(x_0, 1:3)$  and the p value was 0.04715 \*. For  $x_6$  against  $x_0$ , Granger causality test showed that Model 1:  $x_0 \sim \text{Lags}(x_0, 1:3) + \text{Lags}(x_6, 1:3)$ , Model 2:  $x_0 \sim \text{Lags}(x_0, 1:8)$  and the p value was 0.0274 \*.

The test shows that there is causality among the variables.

## **4.2 Development of the model**

Having confirmed the stationarity of the data, as well as the positive granger causality test for all the regions, the model development was readily formulated. This involved lag setting, obtaining the model coefficient and testing their significance. The estimations of VAR equations were analyzed by use of multiple least square. The endogenous variable rainfall was the determinant, and the target for this study was to obtain equations in terms of  $x_0$  in all the stated zones.

### **4.2.1 Lag order selection**

When using a VAR-model it is important to use the correct number of lags. Lag determination was an important step and optimal lag for each zone was obtained through lag selection criterion of AIC, HQ, SC and FPE.

### **4.2.2 Zone one lag order selection**

In zone one the lag order was 3 as shown below

**Table 4.1: Lag order and criteria for zone one**

\$selection

AIC(n)	HQ(n)	SC(n)	FPE(n)
3	3	3	3

\$criteria

	1	2	3	4	5	6	7	8	9	10
AIC(n)	-6.441	-9.470	-16.989	-16.828	-13.960	-14.725	-15.662	-16.096	-16.459	-11.615
HQ(n)	-6.364	-11.406	-16.312	-15.551	-13.617	-14.315	-15.186	-15.553	-15.849	-9.327
SC(n)	-6.236	-9.087	-15.176	-14.087	-13.041	-13.627	-14.386	-14.641	-14.825	-11.054
FPE(n)	0.002	0.0001	0.00000	0.00001	0.0000	0.0000	0.0000	0.0000	0.0000	0.00000

The lag selection is obtained from AIC where its criteria had the minimal value.

#### 4.2.3 The model system output for zone one

To ensure the model used is stable the Roots of the characteristic polynomial are obtained and confirmed that they are inside the unit circle.

Roots of the characteristic polynomial are as listed below.

0.4781 0.4782 0.4649 0.4289 0.4288 0.4287 0.3997 0.3936 0.3935 0.3659 0.3658  
0.3526 0.3528 0.3085 0.3185 0.2918 0.2915 0.2899 0.2555 0.2434 0.2432

This implies that the Endogenous variables:  $x_0, x_1, x_2, x_3, x_4, x_5, x_6$

are stationary and the model can be developed.

The zone one model

$$\begin{aligned}
 x_0 = & 0.056 + 0.087x_{0.l1} + 0.142x_{1.l1} + 0.187x_{3.l1} + 1.08x_{4.l1} + \\
 & 0.072x_{6.l1} + 0.108x_{0.l2} - 0.101x_{1.l2} + 0.37x_{3.l2} + 0.0165x_{4.l2} + \\
 & 0.39x_{6.l2} + 0.44x_{0.l3} + 0.14x_{1.l3} + 0.45x_{3.l3} + 0.39x_{4.l3} - 0.0415x_{6.l3} .
 \end{aligned}$$

## Zone one model coefficient testing

**Table 4.2: Zone one model coefficient**

Variables	Estimate	Std. Error	t-value	Pr(> t )
$x_{0.l1}$	0.087072	0.009360	-9.303	3.07e-06 ***
$x_{1.l1}$	0.142341	0.031443	-4.527	0.001097 **
$x_{3.l1}$	0.187338	0.044428	-4.217	0.001781 **
$x_{4.l1}$	1.082147	0.084548	-12.79	1.59e-07 ***
$x_{6.l1}$	0.071938	0.008197	-5.368	4.00e-04 ***
$x_{0.l2}$	0.10808	0.040665	-2.479	0.032598 *
$x_{1.l2}$	-0.101519	0.026880	-3.777	0.036216 *
$x_{3.l2}$	0.370303	0.076375	-4.848	0.000673 ***
$x_{4.l2}$	0.016500	0.260521	0.462	0.007417 ***
$x_{6.l2}$	0.392382	0.91500	1.839	0.012372 *
$x_{0.l3}$	0.43523	0.05175	4.905	0.001755 **
$x_{1.l3}$	0.141775	0.08784	-4.792	0.000215 ***
$x_{3.l3}$	0.44951	0.08137	3.535	4.31e-06 ***
$x_{4.l3}$	0.38672	0.07908	2.263	0.002771 **
$x_{6.l3}$	-0.0415	0.0165	0.036	0.095243 *
<i>Const</i>	0.056100	0.212506	0.264	0.007147 ***

The model parameters are all statistically significant at different levels. From the analysis  $x_2$  and  $x_5$  were dropped from the model since they had no granger causality to the endogenous variable  $x$  and their parameters were not significant.

The coefficients were highly significant including the constant. The model had also a great fit, as the adjusted  $R^2$  was 0.9895, meaning that 98.95% of a change in the response variable were explained by the model. This result is presented in the appendix III output of the analysis.

### 4.2.4 Model diagnostic test

The following test were tested to check model suitability; serial correlation, arch, normality test and stability test to a certain that the model was stable. The test found that for serial correlation the p-value = 0.152 was not significant which implied that there was no serial correlation of the residual inside the VAR system and thus the test was approved. The test for heteroscedasticity was done using ARCH test that gave a p-value = 1.2e-16 therefore the test was good and the model does not suffer from

heteroscedasticity. The normality test failed, but it did not rule out the development of the model and its application. The stability test was represented in the graphs which displayed no structural breaks, since no point were outside the confidence interval lines. Thus, the system was stable and suitable for the forecasting. The tests are well presented in appendix III.

#### 4.2.5 Zone two

In zone two the lag order was 3 as shown below

**Table 4.3: Lag order and criteria for zone two**

\$selection										
AIC(n)		HQ(n)		SC(n)		FPE(n)				
3	3	3	3	3	3					
\$criteria										
	1	2	3	4	5	6	7	8	9	10
AIC(n)	-21.250	-21.333	-21.350	-21.336	-21.312	-21.290	-21.327	-21.387	-21.348	-21.311
HQ(n)	-21.174	-21.191	-21.274	-21.121	-20.970	-20.881	-20.851	-20.845	-20.739	-20.636
SC(n)	-21.046	-20.951	-20.311	-21.775	-20.394	-20.194	-20.052	-19.935	-19.717	-19.502
FPE(n)	0	0	0	0	0	0	0	0	0	0

The lag selection is obtained from AIC where its criteria had the minimal value.

#### 4.2.6 The model system output for zone two

The Roots of the characteristic polynomial were obtained and confirmed to be inside the unit circle. They included: 0.8092 0.4988 0.4987 0.4601 0.4222 0.4221 0.4163 0.4161 0.4104 0.3603 0.3613 0.3479 0.3477 0.3440 0.3439 0.3213 0.3211 0.3152 0.3151 0.1289 0.1287

It is important to observe that all the roots are inside the unit circle, which means that the system is stable. The zone two model coefficient were obtained and inserted to form the model.

$$x_0 = 0.037 + 0.571x_{0.l1} - 2.01x_{1.l1} + 0.727x_{3.l1} + 0.922x_{4.l1} + 0.326x_{6.l1} + 0.856x_{0.l2} + 1.47x_{1.l2} + 0.945x_{3.l2} + 0.954x_{4.l2} - 0.292x_{6.l2} + 0.294x_{0.l3} - 0.418x_{1.l3} + 0.449x_{3.l3} + 0.286x_{4.l3} + 0.00414x_{6.l3}$$

### Zone two model coefficient testing

**Table 4.4: Zone two model coefficient**

Variables	Estimate	Std. Error	t value	Pr(> t )
$x_{0.l1}$	0.57135	0.03947	14.477	8.13e-10 ***
$x_{1.l1}$	-2.01085	0.12175	-16.516	1.42e-10 ***
$x_{3.l1}$	0.72686	0.10209	7.120	5.17e-06 ***
$x_{4.l1}$	0.92207	0.08577	10.751	3.78e-08 ***
$x_{6.l1}$	0.32639	0.10769	3.031	0.008986 **
$x_{0.l2}$	0.85648	0.11749	7.290	3.97e-06 ***
$x_{1.l2}$	1.47058	0.18822	-7.813	1.80e-06 ***
$x_{3.l2}$	0.94533	0.16376	5.773	4.83e-05 ***
$x_{4.l2}$	0.95473	0.15654	6.099	2.75e-05 ***
$x_{6.l2}$	-0.29238	0.15900	1.839	0.087237.
$x_{0.l3}$	0.29352	0.07515	3.905	0.001585 **
$x_{1.l3}$	-0.41775	0.08478	-4.927	0.000223 ***
$x_{3.l3}$	0.44999	0.08137	5.530	7.41e-05 ***
$x_{4.l3}$	0.28647	0.07908	3.623	0.002771 **
$x_{6.l3}$	0.00415	0.06538	0.063	0.0095024 **
<i>Const</i>	0.03735	0.12818	0.291	0.0017508 ***

The model coefficients were statistically significant except  $x_{6.l2}$  which had a p-value= 0.087237, but it was still used to form the model. Two variables  $x_2$  and  $x_5$  were dropped from the model since they had no granger causality to the endogenous variable  $x$  and their parameters were not significant. The model also had a great fit, as the adjusted  $R^2$  was 0.9961, meaning that 99.61% of the change in the response variable was explained by this regression model. It was also worth mentioning that the trend coefficients were highly significant including the constant. This result is presented in the appendix III output of the analysis.

#### 4.2.7 Model diagnostic test

The stability test used Portmanteau Test (asymptotic) which found that the serial correlation had the p-value = 0.1328, therefore no serial correlation of the residual

inside the model and thus the test was approved. The test for heteroscedasticity was done using ARCH test that gave a p-value = 0.1622 therefore the test was good and the model does not suffer from heteroscedasticity. The normality test failed, but it did not rule out the development of the model and its application. The stability test was represented in the graphs which displayed no structural breaks, since no point were outside the confidence interval lines. Thus, the system was stable and suitable for the forecasting. All these tests were well presented in appendix III.

#### 4.2.8 Zone three

In zone three the lag order was 3 as shown below

**Table 4.5: Zone three Lag order and criteria**

\$selection										
	AIC(n)	HQ(n)	SC(n)	FPE(n)						
	3	5	4	3						
\$criteria										
	1	2	3	4	5	6	7	8	9	10
AIC(n)	-21.418	-21.401	-21.459	-21.410	-21.377	-21.352	-21.340	-21.347	-21.301	-21.277
HQ(n)	-21.035	-21.258	-21.150	-21.143	-21.334	-20.943	-20.864	-20.805	-20.693	-20.602
SC(n)	-20.206	-21.019	-20.799	-21.679	-20.460	-20.256	-20.065	-19.894	-19.670	-19.468
FPE(n)	0	0	0	0	0	0	0	0	0	0

The lag order was determined from criteria with the minimal value, which happened to be AIC with order three.

#### 4.2.9 Zone three model

The Roots of the characteristic polynomial were obtained and confirmed to be inside the unit circle. They included: 0.8296 0.4461 0.412 0.422 0.4123 0.4122 0.3735 0.3521 0.3542 0.3460 0.3459 0.341 0.342 0.3301 0.3311 0.3248 0.3247 0.2998

0.2996 0.1693 0.1691. All the roots are inside the unit circle, which means that the system is stable.

The observed roots values are all inside the unit circle, which means that the model is generally stable.

$$x_0 = 0.23 - 0.67x_{0.11} - 1.51x_{1.11} + 2.42x_{3.11} + 2.39x_{4.11} - 0.54x_{6.11} + 0.93x_{0.12} + 1.48x_{1.12} + 2.83x_{3.12} + 2.37x_{4.12} + 2.37x_{6.12} - 0.61x_{0.13} + 0.49x_{1.13} + 1.20x_{3.13} + 0.85x_{4.13} + 0.41x_{6.13}$$

### Zone three model coefficient testing

**Table 4.6: Zone three model coefficient**

Variables	Estimate	Std. Error	t value	Pr(> t )
$x_{0.11}$	-0.67334	0.09010	-7.473	1.24e-05 ***
$x_{1.11}$	-1.51131	0.11488	-13.156	4.50e-08 ***
$x_{3.11}$	2.42137	0.30024	8.065	6.05e-06 ***
$x_{4.11}$	2.38636	0.31760	7.514	1.18e-05 ***
$x_{6.11}$	-0.54220	0.09695	-5.593	0.0162 *
$x_{0.12}$	-0.93043	0.11799	-7.886	7.49e-06 ***
$x_{1.12}$	-1.48498	0.14198	-10.459	4.71e-07 ***
$x_{3.12}$	2.83377	0.38570	7.347	1.45e-05 ***
$x_{4.12}$	2.36645	0.39394	6.007	8.84e-05 ***
$x_{6.12}$	-0.58036	0.12982	-4.470	0.000946 ***
$x_{0.13}$	-0.61184	0.06635	-9.221	1.65e-06 ***
$x_{1.13}$	-0.49540	0.08979	-5.518	0.000181 ***
$x_{3.13}$	1.20002	0.19285	6.223	6.51e-05 ***
$x_{4.13}$	0.84650	0.19622	4.314	0.001227 **
$x_{6.13}$	-0.40831	0.08150	-5.010	0.001396 **
<i>Const</i>	0.23067	0.35266	-0.654	0.045264 *

The model coefficients were statistically significant and they were used to develop the model for zone three. The analysis found that  $x_2$  and  $x_5$  p-values were not significant and thus not included in the model. The VAR model was able to confirm that the response variable was affected by both its own lagged values and the lagged values of the predictor variables. These coefficients were all highly significant. The model also had a great fit, as the adjusted  $R^2$  was 0.9994, meaning that 99.94% of a change in the response variable was explained by this regression model. It was also worth mentioning





The above table show the value of the smallest selection criteria which was taken to be AIC of lag order 2.

#### 4.2.12 The model system output for zone four

The Roots of the characteristic polynomial were obtained using R-software and they were confirmed to be inside the unit circle. They included: 0.9207 0.4219 0.4066 0.3963 0.3913 0.3913 0.3743 0.3568 0.3553 0.257 0.2474 0.18 0.09407 0.09417. All the roots were inside the unit circle, which made the model to be stable.

$$x_0 = 0.11 - 5.16x_{0.11} + 1.80 x_{1.11} + 1.46x_{2.11} + 1.71x_{3.11} + 1.74x_{4.11} + 2.33x_{5.11} + 2.29x_{6.11} + 3.09x_{0.12} + 1.19 x_{1.12} + 2.10 x_{2.12} + 1.77x_{3.12} + 5.61x_{4.12} + 0.36x_{5.12} - 2.43x_{6.12}$$

#### Zone four model coefficient testing

**Table 4.8: Zone four model coefficient**

Variables	Estimate	Std. Error	t value	Pr(> t )
$x_{0.11}$	-5.1629	1.4287	-3.614	0.004738 **
$x_{1.11}$	1.8011	0.5829	-3.090	0.011449 *
$x_{2.11}$	1.4644	0.2707	5.410	0.000297 ***
$x_{3.11}$	1.7064	0.9135	1.868	0.091323.
$x_{4.11}$	1.7417	0.5853	-2.976	0.013913 *
$x_{5.11}$	2.3322	0.8748	2.666	0.023659 *
$x_{6.11}$	2.2922	1.0235	2.240	0.049037 *
$x_{0.12}$	3.0851	0.9321	-3.310	0.007883 **
$x_{1.12}$	1.1852	0.3807	-3.113	0.011006 *
$x_{2.12}$	2.0965	0.6506	3.222	0.009137 **
$x_{3.12}$	1.7720	0.3811	-4.649	0.000909 ***
$x_{4.12}$	5.6150	1.7671	-3.178	0.009860 **
$x_{5.12}$	0.3640	0.3348	1.087	0.0302471 *
$x_{6.12}$	-2.4316	0.6745	-3.605	0.004809 **
<i>Const</i>	0.1099	0.6123	0.179	0.000861 ***

The model coefficients were statistically significant except  $x_{3.11}$  which had a p-value= 0.091323 however, they were used to develop the model for zone four. The adjusted  $R^2$  is 0.994, meaning that 99.4% of a change in the response variable were explained by this regression model. It was also worth mentioning that the trend coefficients were highly significant.

#### 4.2.13 Model four diagnostic test

The stability test used asymptotic test which found that the serial correlation had the p-value =  $1.624e-09$ , therefore no serial correlation of the residual in the model and thus the test was approved. The test for heteroscedasticity was done using ARCH test that gave a p-value =  $2.2e-16$  therefore the test was good and the model does not suffer from heteroscedasticity. The normality test used the JB-test, kurtosis and skewness test had their p-value was great than 0.05 which passed the normality test. The stability test graphs had no structural breaks, since no point were outside the confidence interval lines. Thus, the system was stable and suitable for the forecasting. All these tests were well presented in appendix III.

#### 4.2.14 Zone five model

In zone five the lag order was 1 as shown below

**Table 4.9: Lag order and criteria for zone five**

\$selection										
	AIC(n)	HQ(n)	SC(n)	FPE(n)						
	1	1	1	1						
\$criteria										
	1	2	3	4	5	6	7	8	9	10
AIC(n)	13.193	24.078	17.843	19.553	16.506	14.214	22.265	20.592	20.880	22.227
HQ(n)	13.403	24.154	17.987	19.830	16.850	14.625	22.743	21.137	20.269	21.549
SC(n)	3.756	24.282	18.227	10.295	7.427	5.314	13.545	12.051	10.757	10.411
FPE(n)	6.145	28.631	87.900	15.000	53.675	20.04	93.52	69.22	71.808	10.108

The above table show the value of the smallest selection criteria which was taken to be AIC of lag order 1.

#### 4.2.15 The model system output for zone five

The Roots of the characteristic polynomial were obtained using R-software and they were confirmed to be inside the unit circle. They included: 0.7837 0.7449 0.7449 0.7419 0.7419 0.7225 0.6662. All the roots were inside the unit circle, which made the model to be stable.

$$x_0 = 0.46 + 0.75x_{0.l1} - 0.157x_{1.l1} + 0.59x_{2.l1} + 0.80x_{3.l1} + 1.22x_{4.l1} - 0.38x_{5.l1} + 0.74x_{6.l1}$$

#### Zone five model coefficient testing

**Table 4.10: Zone five model coefficient**

Variables	Estimate	Std. Error	t value	Pr(> t )
$x_{0.l1}$	0.75462	0.19954	3.782	0.00359 **
$x_{1.l1}$	-0.15744	0.38071	-0.414	0.03687 *
$x_{2.l1}$	0.58847	0.05364	-10.972	6.75e-07 ***
$x_{3.l1}$	0.80424	0.18164	-4.428	0.00128 **
$x_{4.l1}$	1.21706	0.37087	-3.282	0.00827 **
$x_{5.l1}$	-0.38317	0.13787	-2.779	0.01948 *
$x_{6.l1}$	0.73979	0.30570	2.420	0.03606 *
<i>Const</i>	0.45538	1.05863	0.430	0.67620 ***

The coefficients were highly significant including the constant. The model had also a great fit, as the adjusted  $R^2$  was 0.9823, meaning that 98.23% of a change in the response variable were explained by the regression model.

#### 4.2.16 Model diagnostic test

The stability test used asymptotic test which found that the serial correlation had the p-value = 0.03993, therefore no serial correlation of the residual in the model and thus the test was approved. The test for heteroscedasticity was done using ARCH test that gave a p-value = 4.23e3 therefore the test was good and the model suffers from heteroscedasticity. The normality test used the JB-test, kurtosis and skewness test had their p-value was great than 0.05 which passed the normality test. The stability test

graphs had no structural breaks, since no point were outside the confidence interval lines. Therefore, the system was stable and suitable for the forecasting. All these tests were well presented in appendix III.

#### 4.2.17 Global vector

In Global vector the lag order was 3 as shown below

**Table 4.11: Lag order and criteria for Global vector**

\$selection

AIC(n)	HQ(n)	SC(n)	FPE(n)
3	3	3	3

\$criteria

	1	2	3	4	5	6	7	8	9	10
AIC(n)	-20.691	-20.654	-20.829	-20.704	-20.710	-20.687	-20.732	-20.696	-20.752	-20.712
HQ(n)	-20.615	-20.512	-20.719	-20.428	-20.368	-20.278	-20.257	-20.154	-20.144	-20.037
SC(n)	-20.487	-20.272	-20.568	-19.965	-19.792	-19.591	-19.458	-19.243	-19.121	-18.903
FPE(n)	8.631	7.900	6.145	15.000	13.675	20.04	13.52	19.22	7.808	10.108

The lag order of global vector was 3 as obtained from the above tables.4.113...18

#### 4.2.18 The model system output for global vector

The Roots of the characteristic polynomial were obtained using R-software and they were confirmed to be inside the unit circle. They included: 0.7834 0.4474 0.423 0.423 0.4173 0.4173 0.4105 0.4105 0.3943 0.36 0.36 0.336 0.336 0.3359 0.3359 0.3256 0.3256 0.3248 0.3248 0.2634 0.2634. All the roots were inside the unit circle, which made the model to be stable.

$$\begin{aligned}
 x_0 = & 0.05 + 1.31x_{0.11} + 0.70x_{1.11} + 2.68x_{3.11} - 4.15x_{4.11} + 1.89x_{6.11} + \\
 & 1.08x_{0.12} + 0.56x_{1.12} + 3.12x_{3.12} + 4.83x_{4.12} + 3.66x_{6.12} + 0.67x_{0.13} + \\
 & 0.38x_{1.13} + 1.18x_{3.13} + 1.01x_{4.13} + 1.93x_{6.13}
 \end{aligned}$$

## The Global model coefficient testing

**Table 4.12: Global model coefficient**

Variables	Estimate	Std. Error	t value	Pr(> t )
$x_{0.11}$	1.31387	0.13867	-9.475	1.26e-06 ***
$x_{1.11}$	0.69994	0.10293	6.800	2.95e-05 ***
$x_{3.11}$	2.68024	0.24603	-10.894	3.12e-07 ***
$x_{4.11}$	-4.14922	0.48832	-8.497	3.67e-06 ***
$x_{6.11}$	1.88787	0.23913	7.895	7.41e-06 ***
$x_{0.12}$	1.07847	0.20245	-5.327	0.000242 ***
$x_{1.12}$	0.55664	0.18559	2.999	0.012095 *
$x_{3.12}$	3.11604	0.28575	-10.905	3.09e-07 ***
$x_{4.12}$	4.83226	0.72821	-6.636	3.68e-05 ***
$x_{6.12}$	3.66172	0.42100	8.698	2.92e-06 ***
$x_{0.13}$	0.66616	0.12831	-5.192	0.000298 ***
$x_{1.13}$	0.38099	0.09545	3.991	0.002117 **
$x_{3.13}$	1.17905	0.13707	-8.602	3.26e-06 ***
$x_{4.13}$	1.01060	0.31844	-3.174	0.008863 **
$x_{6.13}$	1.92924	0.23391	8.248	4.88e-06 ***
<i>Const</i>	0.05123	1.14477	-0.045	0.000156 ***

The fitted model explained 96.9% of the model variables the value that was obtained from adjusted R- squared which showed that the model fitted the data better. R- squared measures the strength of the relationship between your model and the dependent variable on a convenient of 0% to 100% scale.

### 4.2.19 Model diagnostic test for global vector.

The stability test used asymptotic test which found that the serial correlation had the p-value = 1.256e-05, therefore no serial correlation of the residual in the model and thus the test was approved. The test for heteroscedasticity was done using ARCH test that gave a p-value = 3.21e-6 therefore the test was good and the model suffers from heteroscedasticity. The normality test used the JB-test, kurtosis and skewness test which had their p-value was less than 0.05 which failed the normality test. The stability test graphs had no structural breaks, since no point were outside the confidence interval lines. Thus, the system was stable and suitable for the forecasting. All these tests were well presented in appendix III.

#### 4.2.20 Models Performance Accuracy Measure

The accuracy was measured using sensitivity analysis. Each model was tested using MAE, MSE and RMSE.

#### 4.2.21 Performance Measures

The different accuracy measure gave the following values in different zones as show in the table below;

The table below contain accuracy measure values for all the Models

**Table 4.13: Performance Analysis for the developed models.**

ZONES	MAE	MSE	RMSE	AIC	BIC
<b>ZONE 1</b>	0.897018	0.9131199	0.9555731	33.37781	42.70324
<b>ZONE 2</b>	0.8354289	0.7486969	0.8652727	31.32569	43.48938
<b>ZONE 3</b>	0.9090333	0.778894	0.8825497	37.57943	49.74311
<b>ZONE 4</b>	0.7339015	0.698275	0.835628	45.21378	56.29842
<b>ZONE 5</b>	0.9006911	0.9596382	0.979611	33.37781	42.70324
<b>GLOBAL VEC</b>	0.889702	0.923199	0.960832	31.73782	43.90374

In addition, the findings also indicated that the Performance of the models ranged between 73.39% to 90.90% for MAE, 69.82% to 95.96% for MSE and 83.56% to 97.08% for RMSE. The performance criterion used were AIC and BIC where AIC gave the minimal values. These shows that the models Performances were good.

#### 4.2.22 Sensitivity Analysis

The sensitivity analysis was conducted by use of confusion Matrix. The matrix was represented by Zone Three extracted as shown below

	<i>status1</i>	
<i>pred1</i>	<i>correct</i>	<i>incorrect</i>
<i>correct</i>	530	18
<i>incorrect</i>	11	121

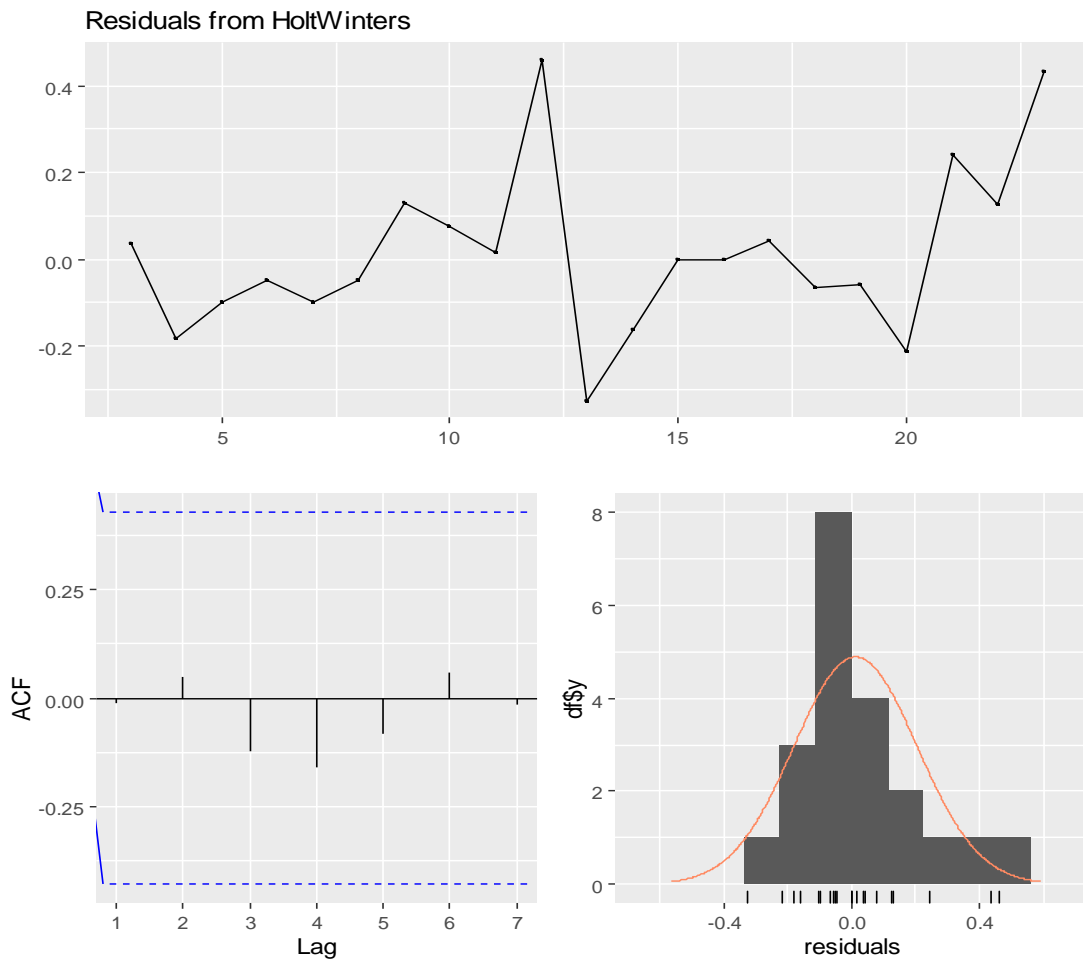
From the confusion matrix;

$$\text{Accuracy} = \frac{651}{680} = 0.9574, \text{ Sensitivity} = \frac{530}{541} = 0.9797, \text{ Specificity} = \frac{121}{139} = 0.8705$$

The model gives an accuracy of 96.55% of the prediction and a 96% sensitivity level.

### 4.3 Forecasting the rainfall pattern

#### 4.3.1 Zone one: Residual analysis

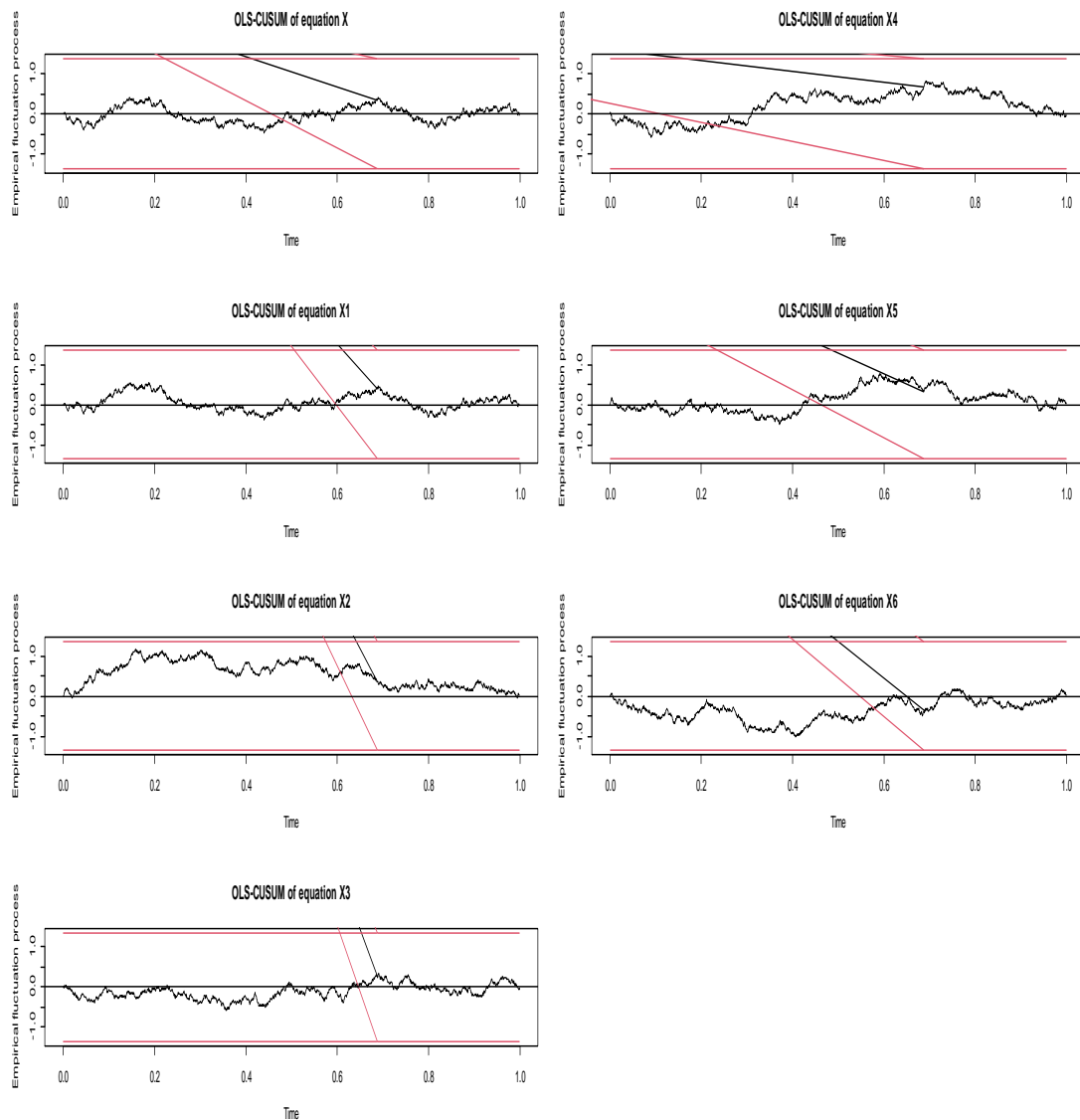


**Figure 4.3: Ljung-Box test**

The Ljung-box test shows three items; the graph of the residuals, which displays the deviations from the actual values, it also displayed the ACF graph, which helps to check for uncorrelation in the residuals. It is the standard residual diagnostic to check if they behave as white noise and therefore the model can be used for forecasting. In this case the developed model can be used for the intended purposes of forecasting. The last part is the histogram, which is used to check for the gaussian behavior. The bell shape is

well displayed in the histogram, and since a good forecast method should have normally distributed residuals, then the model would give a good forecast.

### Stability test for zone one



**Figure 4.4: Stability graph**

The study employed to investigate model stability by developing OLS-CUSUM graphs for each variable. The model is found to be stable because the graph oscillates within the control limit lines.



### 4.3.2 Zone one forecasting analysis

**Table 4.14: Zone one forecasting analysis**

Points	fcst	lower	upper	CI
1	0.4435105	0.3638516944	0.8508728	0.4873558
2	0.5907031	0.0723471584	1.2537533	1.1814061
3	0.4862335	0.3390821988	1.1115492	0.7724670
4	0.7652912	0.0008073628	1.5297750	1.5289676
5	0.5235901	0.2554446203	1.3026249	1.0471803
6	0.4256933	0.3559468631	1.2073335	0.8513866
7	0.5165178	0.4687087960	1.1017444	0.6330356
8	0.6283320	0.4594947612	1.1161588	0.6566640
9	0.6991094	0.4991364994	1.0973553	0.5982188
10	0.5 431256	0.4608876550	1.1471390	0.6862513

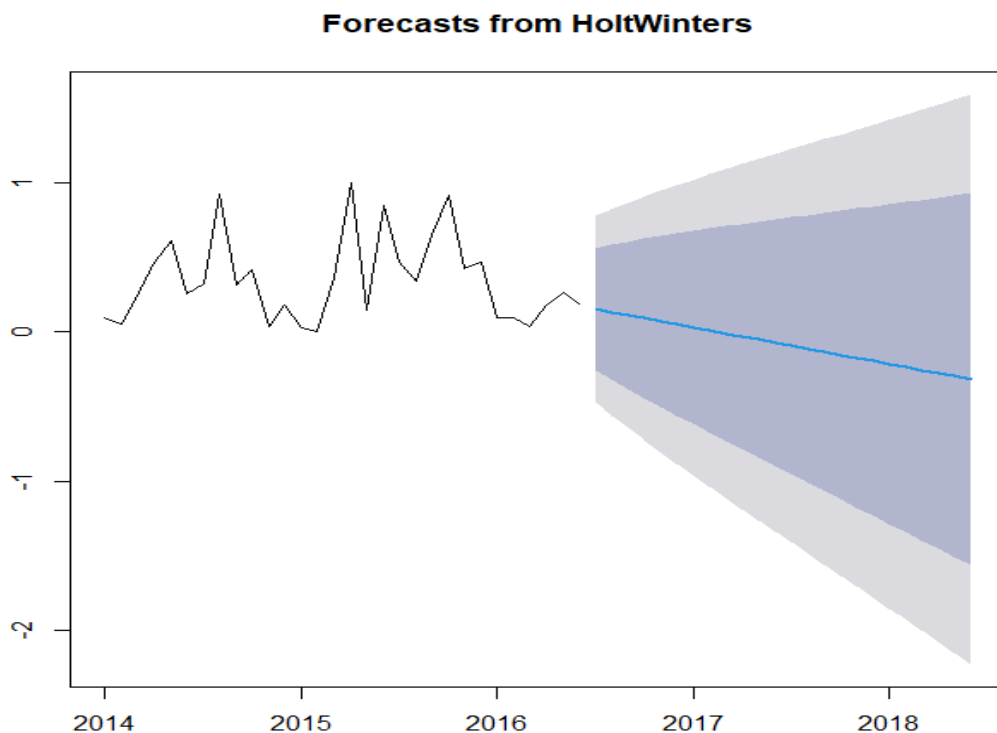
The table shows the forecasted values and the intervals between the lower and upper limit values where the forecasting values will lie between. It also shows the confident interval from the mean forecasted value.

### 4.3.3 Zone one model comparison results of actual and predictive

**Table 4.15: The Prediction Accuracy Test**

Points	Actual value	Predicted value	Error
1	0.09292385	0.09307039	-0.00014653
2	0.60660266	0.55920463	0.04739802
3	0.25305800	0.18979422	0.06326378
4	0.36735499	0.31946700	0.04788799
5	0.64869131	0.52741294	0.12127836
6	0.08976720	0.06899376	0.02077343
7	0.18768907	0.50749100	-0.3198019
8	0.1282763	0.1236670	0.0046093
9	0.1582632	0.1406648	0.0175984
10	0.7836330	0.7575771	0.0260559

This table above was used to compare the forecasted value from BVAR model and the actual values. The p-value of t-test is  $p = 0.8613$  which is greater than the significance level 0.05. In conclusion, there is no difference between the two variance of actual and predicted values. The developed model had a high forecasting ability.



**Figure 4.5: HoltWinters forecast for zone one**

The figure shows the 80% and 90%, ten months forward forecasting, where the thick blue represents the 80% while the light blue is the 90%.

#### 4.3.4 Zone Two

##### Zone Two Forecasting Analysis

**Table 4.16: Zone Two Forecasting Analysis**

Points	fcst	lower	upper	CI
1	0.16261840	0.0496275	0.1710383	0.1214108
2	0.16079294	0.1454337	0.3327519	0.1873182
3	0.23503271	0.2104697	0.6805351	0.4700654
4	0.27284283	0.1855569	0.7312426	0.5456857
5	0.42222782	0.3651445	0.6096001	0.2444556
6	0.53352417	0.4732197	0.5402680	0.0670483
7	0.47436820	0.4399424	0.5886788	0.1487364
8	0.55271189	0.3665874	0.6720112	0.3054238
9	0.46461172	0.3558062	0.6850296	0.3292234
10	0.51846059	0.4034188	0.6403400	0.2369212

The table shows the forecasted values and the intervals between the smallest lower limit = 0.4034188 and maximum upper limit = 0.7312426 values where the forecasting values will lie between. It also shows the confident interval from the mean forecasted value.

### Zone two model comparison results of actual and predictive

**Table 4.17: Zone Two Prediction Error Analysis**

Points	Actual value	Predicted value	Error
1	0.2099334357	0.397611215	-0.187677780
2	0.0087685612	0.187554743	-0.178786182
3	0.0009600614	0.136364679	-0.135404617
4	0.0641001024	0.008314018	0.055786085
5	0.0698604711	0.004570093	0.065290378
6	0.0239375320	0.032621564	-0.008684032
7	0.0501472094	0.435995180	-0.385847971
8	0.3279249872	0.304371484	0.023553503
9	0.0451228879	0.071784835	-0.026661948
10	0.3874017	0.3775781	0.0098236

This table above was used to compare the forecasted value from BVAR model and the actual values. The p-value of t-test is  $p = 0.5609$  which is greater than the significance level 0.05. In conclusion, there is no difference between the two variance of actual and predicted values.

#### 4.3.5 Zone Three Forecasting

##### Zone three forecasting analysis

**Table 4.18: Forecasting Intervals Analysis**

Points	fcst	lower	upper	CI
1	0.81191764	0.7894719	0.8343634	0.0448915
2	0.46466077	0.1709029	1.1002245	0.9293216
3	0.55993465	0.4613027	0.8811720	0.4198693
4	0.32951045	0.1329601	0.5919810	0.5919810
5	0.63947370	0.5047473	0.8257999	0.3210526
6	0.08500537	0.0193076	0.1893183	0.1700107
7	0.15270713	0.0975858	0.2700716	0.1724858
8	0.56771934	0.4283934	0.8070453	0.3786519
9	0.56980961	0.4212342	0.9183850	0.4971508
10	0.32853181	0.1020572	0.6550064	0.5529492

The table shows the forecasted values and the intervals between the smallest lower limit = 0.0193076 and maximum upper limit = 1.1002245 values where the forecasting values will lie between. It also shows the confident interval from the mean forecasted value

### Zone three BVAR model comparison results of actual and predictive

**Table 4.19: Prediction Error Analysis**

Points	Actual value	Predicted value	Error
1	0.29505728	0.07323858	0.22181870
2	0.36242070	0.64851021	-0.28608951
3	0.19424034	0.07051018	0.12373016
4	0.10825613	0.01842611	0.08983002
5	0.13472685	0.16657751	-0.03185066
6	0.06909771	0.02545807	0.04363964
7	0.22121309	0.57787942	-0.35666633
8	0.44165031	0.57137946	-0.12972915
9	0.14476747	0.42009414	-0.27532667
10	0.1266253	0.1445741	-0.0179488

This table above was used to compare the forecasted value from BVAR model and the actual values. The p-value of t-test is  $p = 0.2698$  which is greater than the significance level 0.05. In conclusion, there is no difference between the two variance of actual and predicted values.

#### 4.3.6 Zone Four

##### Zone four forecasting analysis

**Table 4.20: Forecasting Analysis**

Points	fcst	lower	upper	CI
1	0.6486216	0.3591513	0.8563945	0.4972432
2	0.6304823	0.3801009	0.7610655	0.3809646
3	0.5133975	0.4104680	0.7372629	0.3267949
4	0.5894680	0.4746094	0.8535454	0.3789360
5	0.6031851	0.3865770	0.9829473	0.5963703
6	0.4109852	0.2913824	0.6133528	0.3219704
7	0.4273526	0.2789988	0.5337039	0.2547051
8	0.3834961	0.3230151	0.7900073	0.4669922
9	0.3445475	0.3033779	0.5524729	0.2490950
10	0.4294349	0.3802816	0.6391513	0.2588697

The table shows the forecasted values and the intervals between the smallest lower limit = 0.2789988 and maximum upper limit = 0.9829473 values where the forecasting values will lie between. It also shows the confidence interval from the mean forecasted value.

### Zone four BVAR model comparison results of actual and predictive

**Table 4.21: Zone four Prediction Error Analysis**

Points	Actual value	Predicted value	Error
1	0.71353620	0.6721457	0.04139053
2	0.01932494	0.2453208	-0.22599591
3	0.01403463	0.1868350	-0.17280033
4	0.35768625	0.6882216	-0.33053531
5	0.23229276	0.3104205	-0.07812777
6	0.68419902	0.7472730	-0.06307402
7	0.37644281	0.6778950	-0.30145214
8	0.17095138	0.2950825	-0.12413116
9	0.29122945	0.4463421	-0.15511263
10	0.20264953	0.2509196	-0.14827010

The table above was used to compare the actual and predicted values and calculate the accuracy level. The p-value of t-test is  $p = 0.8119$  which is greater than the significance level 0.05. In conclusion, there is no difference between the two variance of actual and predicted values.

#### 4.3.7 Zone Five

##### Zone five forecasting analysis

**Table 4.22: Forecasting Analysis for Zone five**

Points	fcst	lower	upper	CI
1	0.4486216	0.3591513	0.6563945	0.2972432
2	0.5904823	0.3801009	0.7610655	0.3809646
3	0.6133975	0.5104680	0.7372629	0.2267949
4	0.6894680	0.4746094	0.8535454	0.3789360
5	0.5031851	0.3865770	0.9929473	0.6063703
6	0.4109852	0.2913824	0.7133528	0.4219704
7	0.4273526	0.2789988	0.6337039	0.3547051
8	0.3834961	0.3230151	0.5900073	0.2669922
9	0.4445475	0.3633779	0.6524729	0.2890950
10	0.5294349	0.3802816	0.7391513	0.3588697

The table shows the forecasted values and the intervals between the smallest lower limit = 0.2789988 and maximum upper limit = 0.9929473 values where the forecasting values will lie between. It also shows the confidence interval from the mean forecasted value.

### Zone five model comparison results of actual and predictive

**Table 4.23: Prediction Error Analysis for Zone five**

Points	Actual value	Predicted value	Error
1	0.09834357	0.3677607	-0.26941716
2	1.00000000	0.7352895	0.26471051
3	0.59967566	0.6345414	-0.03486578
4	0.47565543	0.6351184	-0.15946294
5	0.13904012	0.4547479	-0.31570774
6	0.15614502	0.3401993	-0.18405425
7	0.08130997	0.06777652	0.01353345
8	0.17715939	0.14582165	0.03133774
9	0.36519333	0.32300049	0.04219284
10	0.62482868	0.62798413	0.00315545

This table above was used to compare the forecasted value from zone five model predicted values and the actual values. The p-value of t-test is  $p = 0.3893$  which is greater than the significance level 0.05. In conclusion, there is no difference between the two variance of actual and predicted values.

### 4.3.8 Global Region forecasting analysis

**Table 4.24: Forecasting Analysis for Global Vector**

Points	fcst	lower	upper	CI
1	0.486216	0.296353	0.563945	0.267592
2	0.304823	0.221118	0.610663	0.389545
3	0.233975	0.104680	0.372637	0.267957
4	0.494680	0.474609	0.535462	0.060853
5	0.631851	0.465770	0.929481	0.464872
6	0.198852	0.123824	0.333536	0.209712
7	0.513526	0.389988	0.637047	0.247059
8	0.734961	0.630151	0.900171	0.270020
9	0.445475	0.233779	0.524737	0.290958
10	0.294349	0.102816	0.391522	0.288706

The table shows the forecasted values and the intervals between the smallest lower limit = 0.104680 and maximum upper limit = 0.900171 values where the forecasting values will lie between. It also shows the confident interval from the mean forecasted value.

### 4.3.9 Global region model comparison results of actual and predictive

**Table 4.25: Prediction Error for Global Vector**

Points	Actual value	Predicted value	Error
1	0.53396298	0.53604143	-0.002078454
2	0.14318096	0.31297298	-0.169792019
3	0.06667926	0.22819949	-0.161520233
4	0.15531923	0.19788896	-0.042569729
5	0.30681526	0.52991659	-0.223101324
6	0.11578769	0.07328579	0.042501899
7	0.34267473	0.52983639	-0.187161659
8	0.43071402	0.60557426	-0.174860248
9	0.14357764	0.31468782	-0.171110183
10	0.32300049	0.36519333	-0.04219284

The p-value of t-test is  $p = 0.6988$  which is greater than the significance level 0.05. In conclusion, there is no difference between the two variance of actual and predicted values for the ten point values of the months of May 2016.

## CHAPTER FIVE

### SUMMARY, CONCLUSION AND RECOMMENDATION

#### 5.0 Introduction

This chapter summarizes the findings of the study. It provided the conclusions that were drawn from the results discussed in chapter four. Finally, it provided the statistical contributions of the study and the recommendations for further research.

#### 5.1 Summary of the Findings

This study was propelled by the requirement of the meteorological need for the determinants of the climate changes in assisting the meteorologists to predict the future weather changes.

##### 5.1.1 Diagnostic analysis of the Variables

The graphs of the initial data were plotted and they exhibited a time series behavior which contained non stationarity. The preprocessing of data was required and at least first differencing was done. Two tests were conducted to confirm data stationarity and causality.

##### 5.1.2 Model Development

The lag order for each region were determined which was three for most of the zones, two for zone four and one for zone five. The models for each zone were developed after their coefficients were statistically obtained. The model coefficients were tested to ascertain their usefulness where most of them were found to be statistically significant. The constant in all the models were found to be statistically significant.

##### 5.1.3 The Model Performance Measures

The performance analysis for the developed models were tested using MAE, MSE and RMS. Their critical values were determined using AIC and BIC.



#### 5.1.4 Forecasting Model

The ten-point interval estimation and actual against predicted values were forecasted using the developed models. This clearly was exhibited by use of graphs and HoltWinters diagrammatic representation for forecasting. The achieved accuracy of the models was found to be good for forecasting.

### 5.2 Conclusion of the Study

The following conclusions were arrived at on the basis of the study objectives:

#### 5.2.1 Diagnostic analysis

After first differencing, the study variables were found to be stationary. The graphs of the differenced data depicted stationarity, where the mean and variance were constant. The stationarity test was evaluated using two tests, ADF and PP test. Their ADF and PP values had a significant p- value  $< 0.05$ . The variables in all the regions were found to have a causal effect to the exogenous variable rainfall ( $x_0$ ), except in zone one and two where  $x_2$  and  $x_5$  had a p-value  $> 0.05$ . The conclusion was arrived at if both ADF and PP models rejected the null hypothesis, thus the data was treated to be stationary. All zones were found to be stationary from the ADF and confirmed by PP test which gave a strong statistical significance of the p – values. The granger causality test, which was to test if there was any serial correlation and if the lags of the predictor variables influenced that of response variable were conducted. It was concluded that the temperature, relative humidity, atmospheric pressure, wind speed, granger caused rainfall in all the zones but radiation and wind gust failed in zone one and two. This was clearly given by the statistically significant p-values in the zones. Finally, the ADF and PP test for unit roots was conducted for time series data used for the study. The result of unit root test at the first or second difference level was shown in different tables

appended in the appendices I. The null hypothesis of non-stationarity was performed at 5% significance levels. The result of the test illustrated that the data series was nonstationary before the first differencing. However, the result of the test on the first difference strongly supports that all data series were stationary at 5%, significance level. The test was a fundamentally a statistical significance test. That means, there was a hypothesis testing involved with a null and alternative hypothesis and as a result a test statistic was computed and p-values obtained tested for its significant. From the statistic test and the p-values, an inference was made onto whether the given time series was stationary or not. In conclusion, the data was stationary and this gave an allowance for conducting time series analysis and developing a forecasting model.

### **5.2.2 Model Development for specific zones**

The determined lag order ranged between one and three for the zones and global vector. The roots of the characteristic's polynomial were equal to the number of variables and they were confirmed to be inside the unit circle. The coefficients of the models were tested for their significant and found to be less than the p-value (0.05). The models had also a great fit as reflected in adjusted R squared, which was more than 82%. The models were diagnosed by checking model suitability through; serial correlation test, heteroscedasticity test, normality test and stability test. The model had no serial correlation and their p-value was greater than 0.05. The models did not suffer from heteroscedasticity since their p-value  $< 0.05$ . In most of the models, the normality test failed, but these was attributed to the type of data used. The graphs of stability test had no structural breaks, since on point were outside the confidence interval lines. These depicted the strength of the model coefficients and their predictability ability. Therefore, the developed models were suitable for forecasting ability.

### **5.3 Performance and Sensitivity of the Model**

The model performances ranged between 73.56% to 90.64% for MAE, 69.82% to 95.95% for MSE and 83.56% to 97.96% for RMSE with AIC of 33.37% which were termed as good. The sensitivity analysis of the models was tested using confusion matrix. The accuracy for zone one was 96.55% and its sensitivity was 96%. This shows a high level of accuracy and sensitivity of the developed models.

### **5.4 Forecasting**

The researcher established and estimated a forecasting model from the daily variables. This was done by applying Vector Autoregressive (VAR) method and the Bayesian method of multivariate time series analysis. It was found that the rainfall variable and other related variables namely temperature, humidity, atmospheric pressure, and wind speed were interrelated. A VAR was then developed and a BVAR was generated and customized for forecasting purposes. Using the training data, it was possible to test the model and come up with the future predicted values. The variances of the actual and the predicted values were compared using t-test, and it was found that, there was no significant difference between their variances. Hence, rainfall patterns were forecasted for a maximum of ten future days. It was also found that different zones adopted different variables depending on their location; in zone one radiation and wind gust had no influences on the dependent variable. In zone two wind gust had no effect and thus it was dropped. In zone five radiation had very minimal effects and thus it was not included in the predictive model.

### **5.5 Contribution of the Study**

The outcomes and the technique implemented in this study may contribute as a source of model for forecasting prospective of weather for other countries within the tropical

region. The techniques utilized and the outcomes displayed as a part of this study give experience into the impact of these factors on the meteorological forecasting. This study discovered that weather changes are influenced by several factors that need to be considered and applied to give a good forecast performance. The results shows that the use of the VAR coupled with the recognition of BVAR produced better forecasts over long forecasting horizons. The most important contribution of this study was the forecasting model developed could be used in artificial intelligent areas of meteorological department which would greatly improve their predictability performance.

## **5.6 Recommendations**

This part consists of two sections recommendation which is derived from the study findings and recommendation for further research that the research feels that may be investigated further.

### **5.6.1 Recommendations of the study**

Since the predictive model had a more accurate prediction, it is recommended that it be adopted in the area of Artificial Intelligence in weather forecasting. When comparing the weather variables for different regions, the rainfall pattern was been influenced by a number other weather variable, therefore the study recommends that for any other weather variable to be forecasted, influencing factors should be put into consideration. Different regions gave specific weather prediction models, where some of the weather variables had no contribution to the final models. Therefore, it is important to handle model per region before developing the global forecasts model. The sample sizes of secondary data were limited to 4 years of daily observations and the researcher

recommends on an increased amount of historic data which would better the accuracy of the model developed.

### **5.6.2 Recommendations for further study**

For further research, it is recommendable to use more weather variables like topography, cloud cover, sun shine duration among others to improve the predictability. Other researchers may use other variable like temperature since the effects of global warming are noted through the irregular changes of precipitation or temperature. The study used secondary data for 2014 to 2018, therefore the researcher recommends that current data for 2019 to 2022 may be used to make current future predictions. Climate models only predict a range of possible future scenarios, the extent of how far the future would be should be studied. The study recommends for further inclusion of more weather variables in the Bayesian vector autoregression area. The researcher recommends for application of other technique like Random Forest and Bootstrapping technique to check whether the accuracy may be further improved from these models.

## REFERENCES

- Andrews, D. W. (2016). *Asymptotic results for generalized Wald tests*. *Statistical Theory*, 348-358.
- Asteriou Dimitrios och and Stephen G. Hall. (2017) *Applied Econometrics: A Modern Approach Using EViews and Microfit*. Rev. ed. Basingstoke: Palgrave Macmillan.
- Bañbura M, Giannone D, Reichlin L (2010). “Large Bayesian Vector Auto Regressions.” *Journal of Applied Statistics*, 25(1), 71–92.
- Banner, K. M., K. M. Irvine, and T. J. Rodhouse. 2020. *The use of Bayesian priors in ecology: the good, the bad and the not great*. *Methods in Ecology and Evolution* 11:882–889.
- Bauer P., A. Thopre and G. Brunet, (2015): *The quiet revolution of numerical weather prediction*. *Nature*, 525:47–55.
- Belmonte, M. A. G., Koop, G., and Korobolis, D. (2014). *Hierarchical shrinkage in time varying parameter models*. *Journal of Forecasting*, 33:80–94.
- Bernadinelli, L., Pascutto, C., Best, N. G., & Gilks, W. R. (2016). *Bayesian Multi-Disease modelling approach to inform health planning policies*. *Int. J. Environ. Res. Public Health* 2016, 13(4)
- Bhattacharya, Anirban, Debdeep Pati, Natesh S. Pillai, and David B. Dunson (2015) “Dirichlet-Laplace Priors for Optimal Shrinkage,” *Journal of the American Statistical Association*, Vol. 110, No. 512, pp. 1479–1490. PMID: 27019543.
- Blunden, J., D.S. Arndt, and G. Hartfield (eds.), (2018): *State of the Climate in 2017*. *Bull. Amer. Meteor. Soc.*, 99:8, Si–S310, doi:10.1175/2018BAMS State of the Climate. 1.
- Braun, P. A. and S. Mittnik. “Misspecifications in Vector Autoregressions and Their Effects on Impulse Responses and Variance Decompositions.” *Journal of Econometrics* 59 (October 2013)
- Brockwell, P.J. and Davis, R.A. (2010) *Introduction to Time Series and Forecasting*. 2nd Edition, Springer Texts in Statistics, New York.
- Caraiani, P. (2010). *Forecasting Romanian GDP using BVAR model*. *Romanian Journal of Economic Forecasting*, 13(4), 76-87
- Carriero, A., Clark, T. E., & Marcellino, M. (2016). *Common drifting volatility in large Bayesian VARs*. *Journal of Business & Economic Statistics*, 34(3), 375-390.
- Campbell, S.D. & Diebold, F.X. (2018). *On the Comparison of Interval Forecasts* *Journal of Time Series Analysis*. 39:953-965. DOI:10.1111/Jtsa.12426
- Chan, J. and Eisenstat, E. (2019). *Efficient estimation of Bayesian VARMA with time-varying coefficients*. Springer Texts in Statistics, New York.

- Clark TE (2011). “*Real-Time Density Forecasts from Bayesian Vector Autoregressions with Stochastic Volatility.*” *Journal of Business & Economic Statistics*, 29(3), 327–341. doi: 10.1198/jbes.2010.09248.
- Cogley, T., & Sargent, T. J. (2018). *Drifts and volatilities: monetary policies and outcomes in the post WWII US.* *Review of Economic dynamics*, 8(2), 262-302.
- Coumou, D. and S. Rahmstorf, (2012): *A decade of weather extremes.* *Nature Climate Change*, 2:491–49
- David Donoho. 2020. “50 Years of Data Science.” *Journal of Computational and Graphical Statistics* 26 (4): 745–66. <https://doi.org/10.1080/10618600.2017.1384734>.
- D’ees, S, di Mauro F, Pesaran HM, Smith LV. 2017. *Exploring the international linkages of the euro area: a global VAR analysis.* *Journal of Applied Statistics* 22.
- Del Negro M, Schorfheide F (2007). “*Priors from General Equilibrium Models for VARs.*” *International Economic Review*, 45(2), 643–673. doi:10.1111/j.1468-2354.2004.00139.x.
- Dezfuli AK, Zaitchik BF, Gnanadesikan A (2015) *Regional atmospheric circulation and rainfall variability in south equatorial Africa.* *J Clim* 28(2):809–818
- Doan, T., Litterman, R., & Sims, C. (2014). *Forecasting and conditional projection using realistic prior distributions.* *Econometric and Statistics reviews*, 6(1), 221-340.
- Fabbian, D., De Dear, R., & Lellyett, S. (2007). *Application of artificial neural network forecasts to predict fog at Canberra International Airport.* *Weather and forecasting*, 22(2), 372-381.
- Forni, M., Hallin, M., Lippi, M., & Reichlin, L. (2018). *Do financial variables help forecasting inflation and real activity in the euro.* *Journal of Monetary Economics and Statistics*, 50(6), 1243-1255.
- Gabry, J., D. Simpson, A. Vehtari, M. Betancourt, and A. Gelman. (2019). *Visualization in Bayesian work-flow.* *Journal of the Royal Statistical Society: series A (Statistics in Society)* 182:389–402
- Gelman, A., A. Vehtari, D. Simpson, C. C. Margossian, B. Carpenter, Y. Yao, L. Kennedy, J. Gabry, P.-C. Bürkner, and M. Modrák. (2020). *Bayesian work-flow.* arXiv preprint arXiv:2011.01808
- Giannone D, Lenza M, Primiceri GE (2015). “*Prior Selection for Vector Autoregressions.*” *Review of Economics and Statistics*, 97(2), 436–451
- Giannone Ricco (2018). “*Bayesian Inference Vector Autoregressions.*” *Oxford Encyclopedia Economics and Statistics*, oxford university press
- Griffin, Jim and Phil Brown (2017) “*Hierarchical Shrinkage Priors for Regression Models,*” *Bayesian Analysis*, Vol. 12, No. 1, pp. 135–159, 03.

- Hafer, R. W. and R. G. Sheehan. “*The Sensitivity of VAR Forecasts to Alternative Lag Structures.*” *International Journal of Forecasting* (2020), 399-408.
- Hannan, E. J., & Deistler, M. (2012). *The statistical theory of linear systems.* Society for Industrial and Applied Mathematics.
- Hjort, N. L., & Claeskens, G. (2013). *Frequentist model average estimators.* *Journal of the American Statistical Association*, 98(464), 879-899.
- Hoeting, J. A., Madigan, D., Raftery, A. E., & Volinsky, C. T. (2019). *Bayesian model averaging: a tutorial.* *Statistical science*, 382-401.
- Huber F, Feldkircher M (2019). “*Adaptive Shrinkage in Bayesian Vector Autoregressive Models.*” *Journal of Business & Economic Statistics*, 37(1), 27–39. doi:10.1080/07350015.2016.1256217.
- Intergovernmental Panel on Climate Change (IPCC), 2019: *Special Report on the Ocean and Cryosphere in a Changing Climate* [<https://www.ipcc.ch/srocc>]
- Jarociński, Marek and Bartosz Maćkowiak (2017) “*Granger Causal Priority and Choice of Variables in Vector Autoregressions,*” *The Review of Economics and Statistics*, Vol. 99, No. 2, pp. 319–329, May.
- Jaynes, E. T. (2013). *Probability theory: The logic of science.* Cambridge university press.
- Ji, M., A. Kumar and A. Leetmaa (2018): *A multi season climate forecast system at the National Meteorological Center.* *Bull. Amer. Meteor. Soc.*, 75:569–577
- Kapetanios, G., Carriero, A & Marcellino, M. (2019). *Forecasting exchange rates with a large Bayesian VAR.* *International Journal of Forecasting*, 25(2), 400-417.
- Karlsson, Sune (2015). *Forecasting with Bayesian Vector Autoregression.* *Handbook of Economic Forecasting.* 2 B: 791-897
- Kastner G, Frühwirth-Schnatter S (2014). “*Auxiliary-Sufficiency Interweaving Strategy (ASIS) for Boosting MCMC Estimation of Stochastic Volatility Models.*” *Computational Statistics & Data Analysis*, 76, 408 – 423. doi:10.1016/j.csda.2013.01.002
- Kilian L, Lütkepohl H (2017). *Structural Vector Autoregressive Analysis.* Cambridge University Press. doi:10.1017/9781108164818.
- Koop, G., & Korobilis, D. (2013). *Large time-varying parameter VARs.* *Journal of Econometrics and Statistics*, 177(2), 185-198.
- Koop GM, Korobilis D (2010). “*Bayesian Multivariate Time Series Methods for Empirical Macroeconomics.*” *Foundations and Trends in Econometrics*, 3(4), 267–358. doi:10.1561/08000000013.
- Korobilis, Dimitris (2013) “*VAR Forecasting Using Bayesian Variable Selection,*” *Journal of Applied Econometrics*, Vol. 28, No. 2, pp. 204–230.



- Lemoine, N. P. 2019. *Moving beyond non informative priors: why and how to choose weakly informative priors in Bayesian analyses*. *Oikos* 128:912–928
- Liew, Venus Khim, Sen, L., (2014) "Which Lag Length Selection Criteria Should We Employ. *Statistics Bulletin*, Vol. 3, No. 33 pp. 1–9
- Linacre, E and Geerts, B. (2017). *Climates and Weather Explained Routledge London*, pp. 321 – 345
- Litterman RB (2019). "A Bayesian Procedure for Forecasting with Vector Autoregressions." MIT Working Paper, (275).
- Lutgens, F. K. and TarBuck, E. J. (2015). *The Atmosphere: An Introduction to Meteorology, Fourth edition*. Prentice Hall, New Jersey, pp. 299 – 331.
- Lütkepohl, H. (2017). *New introduction to multiple time series analysis*. Springer Science & Business Media.
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer Science & Business Media, Berlin.
- Malakoff, D. (2019). *Bayes offers a 'new' way to make sense of numbers*. *Science*, 286(5444), 14-64.
- Mary K, Dave M, Maurine A and Joanne R. (2018). *Extreme Rainfall and Flooding over Central Kenya Including Nairobi City during the Long-Rains Season*.
- Moffatt, Mike. "What Is the Augmented Dickey-Fuller Test?" Thought Co, Aug. 29, 2020, [thoughtco.com/the-augmented-dickey-fuller-test-1145985](https://thoughtco.com/the-augmented-dickey-fuller-test-1145985).
- Mudelsee, M. (2019). *Climate time series analysis: classical statistical and bootstrap methods*: Atmospheric and Oceanographic Sciences Library, v. 42, 474 p.
- Nicenboim, B., Schad, D. J., & Vasisht, S. (2021). *An introduction to bayesian data analysis for cognitive science*.
- Otiende P, & Brian M. (2009). *The economic impacts of climate change in Kenya: Riparian flood impacts and cost of adaptation*
- Pesaran, M. H., & Smith, R. (2018). *Macro econometric modelling with a global perspective*. The Manchester School, 74, 24-49.
- Pesaran, M. H., Schuermann, T., & Weiner, S. M. (2014). *Modeling regional interdependencies using a global error-correcting macro econometric model*. *Journal of Business & Economic Statistics*, 22(2), 129-162.
- Primiceri, G. E. (2015). *Time varying structural vector autoregressions and monetary policy*. *The Review of Economics and Statistics Studies*, 72(3), 821-852.
- Raftery, A. E., Madigan, D., & Hoeting, J. A. (2017). *Bayesian model averaging for linear regression models*. *Journal of the American Statistical Association*, 92(437), 179-191.

- Ramos, F. F. R. (2003). *Forecasts of market shares from VAR and BVAR models: a comparison of their accuracy*. *International Journal of Forecasting*, 19(1), 95-110.
- Reichert, P., & Omlin, M. (2017). *On the usefulness of overparameterized ecological model*. *Ecological Modelling*, 95(2-3), 289-299.
- Robert, C. P. (2020), *The Bayesian Choice, Second Edition*, New York: Springer-Verlag.
- Rue H, Martino S, and Lindgren F. (2015). *R-INLA: Approximate Bayesian Inference Using Integrated Nested Laplace Approximations*.
- Sharon E. N. (2018). *The ITCZ and the Seasonal Cycle over Equatorial Africa. Department of Earth, Ocean, and Atmospheric Science*, The Florida State University, Tallahassee, Florida, Page(s): 337–348
- Silvia M. A and Giovanni R (2018) “*Forecasting with a Large, Mixed Frequency, Bayesian VAR*,” Working Papers 2018-30, Federal Reserve Bank of St. Louis
- Sims, C. A. (2006) *Bayesian Methods in Applied Econometrics, or, Why Econometrics Should Always and Everywhere Be Bayesian*. Technical Report, Princeton University.
- Sims, Christopher A. (2010) “*Bayesian Methods for Dynamic Multivariate Models*,” *International Economic Review*, Vol. 39, No. 4, pp. 49–68, November.
- Sims CA, Zha T (1998). “*Bayesian Methods for Dynamic Multivariate Models*.” *International Economic Review*, pp. 948–988. doi:10.2307/2527347.
- Statisticat, L. L. C. (2016). *Laplaces Demon: Complete Environment for Bayesian Inference*. R package version 16.0. 1.
- Stock, J. H., & Watson, M. W. (2017). *Generalized shrinkage methods for forecasting using many predictors*. *Journal of Business & Economic Statistics*, 54(2), 88 - 96.
- Stock JH, Watson MW (2020) *Chapter 10 forecasting with many predictors*. In: Elliott CWJG, Timmermann A (eds) *Handbook of economic forecasting*, vol 1. Elsevier, Amsterdam, pp 515–554
- Stock, J. H., & Watson, M. W. (2015). *Generalized shrinkage methods for forecasting using many predictors*. *Journal of Business & Economic Statistics*, 30(4), 481-493.
- Verbeek Marno. (2018) *A guide to modern econometrics. 3rd ed*. Chichester, England; Hoboken, NJ: John Wiley & Sons.
- Villani M (2009). “*Steady-State Priors for Vector Autoregressions*.” *Journal of Applied Econometrics*, 24(4), 630–650. doi:10.1002/jae.1065
- Wasserman, L. (2014). *All of Statistics: A Concise Course in Statistical Inference*. Springer Science & Business Media.

- Wolf, C., M. Novak, and A. I. Gitelman. 2017. *Bayesian characterization of uncertainty in species interaction strengths*. *Oecologia* 184:327–339
- Wooldridge, J. M., 2015. *Introductory Econometrics: A modern approach*. s.l.:South-Western Cengage Learning.
- Yang, D., Sharma, V., Ye, Z., Lim, L. H., Zhao, L., & Aryaputera, A. W. (2015). *Forecasting of global horizontal irradiance by exponential smoothing, using decompositions*. *Energy*, 81
- Youssef Hmamouche , Lotfi Lakhal , Alain Casali. (2020). *A scalable framework for large time series prediction*, Aix Marseille University, LIS, CNRS UMR 7020, 413 Avenue Gaston Berger, Aix en Provence, France.
- Yuriy Gorodnichenko & Byoungchan Lee (2020) *Forecast Error Variance Decompositions with Local Projections*, *Journal of Business & Economic Statistics*,

## APPENDICES

## Appendix I: Analysis Tables

## Stationarity Test

## Zone One

Table 1: Zone One Stationarity Test

Variables	ADF Test Statistics	Phillips- Perron	Truncation lag parameter	P-Value ADF	P- Value P.P	Remarks
$X_0$	-28.164	-273.58	3	0.01	0.0127	Stationary
$X_1$	-27.153	-272.51	3	0.0121	0.0103	Stationary
$X_2$	-27.879	-79.12	3	0.0213	0.0182	Stationary
$X_3$	-28.808	-283.58	3	0.0031	0.0323	Stationary
$X_4$	-29.758	-297.84	3	0.0274	0.0145	Stationary
$X_5$	-28.762	-286.58	3	0.0015	0.0113	Stationary
$X_6$	-28.415	-290.42	3	0.0093	0.0393	Stationary

From table 1, it shows that after differencing once all the variables are stationary, and they have a unit root. Thus, we make a conclusion that zone one data is stationary.

## Stationarity test for zone two

Table 2: Zone two show that the variables are Stationary

Variables	ADF Test Statistics	Phillips- Perron	Truncation lag parameter	P-Value ADF	PValue P.P	Remarks
$X_0$	-11.019	-36.716	3	0.0156	0.0124	Stationary
$X_1$	-12.984	-36.762,	3	0.0126	0.073	Stationary
$X_2$	<b>-10.066</b>	-37.444	3	0.0524	0.0473	Stationary
$X_3$	-10.249	-35.759	3	0.0026	0.0149	Stationary
$X_4$	-11.108	-11.108	3	0.0138	0.0231	Stationary
$X_5$	-37.055	-10.715	3	0.0361	0.0302	Stationary
$X_6$	-10.715	-35.215	3	0.0054	0.0215	Stationary

### Stationary test for zone three

**Table 3:** Zone three Stationarity Test

Variables	ADF Test Statistics	Phillips-Perron	Truncation lag para	P-Value ADF	P-Value P. P	Remarks
$X_0$	-11.173	-36.636	3	0.0155	0.0352	Stationary
$X_1$	-12.238	-36.714	3	0.0268	0.01	Stationary
$X_2$	-10.472	-37.385	3	0.0246	0.0165	Stationary
$X_3$	-10.907	-36.66	3	0.0276	0.0103	Stationary
$X_4$	-11.555	-36.763	3	0.0371	0.0124	Stationary
$X_5$	-11.633	-36.884	3	0.0188	0.01	Stationary
$X_6$	-11.207	-35.193	3	0.0101	0.0297	Stationary

### Stationarity Test for Zone Four

**Table 4:** Zone four Stationarity Test

Variables	ADF Test Statistics	Phillips-Perron	Truncation lag param.	P-Value ADF	P-Value P. P	Remarks
$X_0$	-28.245	-273.61	3	0.01	0.0242	Stationary
$X_1$	-28.327	-274.36	3	0.0260	0.0173	Stationary
$X_2$	-27.892	-277.82	3	0.0124	0.0732	Stationary
$X_3$	-28.507	-281.73	3	0.0121	0.0493	Stationary
$X_4$	-29.518	-301.02	3	0.0386	0.0231	Stationary
$X_5$	-29.222	-287.64	3	0.0361	0.0302	Stationary
$X_6$	-28.453	-291.18	3	0.0142	0.0215	Stationary

### Stationarity Test for Zone Five

**Table 5:** Zone five Stationarity Test

Variables	ADF Test Statistics	Phillips- Perron	Truncation lag param	P-Value ADF	P-Value P.P	Remarks
$X_0$	-11.508	-37.064	3	0.01	0.0173	Stationary
$X_1$	-12.403	-36.616	3	0.0387	0.0158	Stationary
$X_2$	-10.33	-37.838	3	0.0237	0.01	Stationary
$X_3$	-10.901	-36.79	3	0.0275	0.0253	Stationary
$X_4$	-11.438	-36.762	3	0.0421	0.0149	Stationary
$X_5$	-10.702	-37.208	3	0.0173	0.0334	Stationary
$X_6$	-10.586	-35.248	3	0.0192	0.0146	Stationary

### Stationarity Test for Global Vector

**Table 6:** Global Region Stationarity Test.

Variables	ADF Test Statistics	Phillips- Perron	Truncation lag param	P-Value ADF	P-Valu P.P	Remarks
$X_0$	-11.606	-37.085	3	0.01	0.01	Stationary
$X_1$	-11.54	-37.063	3	0.0379	0.01	Stationary
$X_2$	-10.315	-37.456	3	0.0242	0.0186	Stationary
$X_3$	-10.888	-36.3	3	0.0274	0.0101	Stationary
$X_4$	-11.35	-36.638	3	0.0298	0.0199	Stationary
$X_5$	-11.57	-36.931	3	0.0151	0.0437	Stationary
$X_6$	-11.088	-35.423	3	0.0228	0.0382	Stationary

## Granger causality test

### Zone One

The hypothesis was that rainfall is not granger caused by the independent variables.

Hypothesis Testing for Granger causality Zone one

**Table 7: Granger Causality test zone one**

$X^*X_i$	F	P	Null hypothesis rejected
$X_1$	-6 2.1133	0.0491 *	√
$X_2$	-3 3.2283	0.02173 *	√
$X_3$	-6 1.9587	0.006851 **	√
$X_4$	-3 3.2381	0.02144 *	√
$X_5$	-6 2.7284	0.01223 *	√
$X_6$	-3 3.6161	0.01279 *	√

### Zone Two

The hypothesis was that rainfall is not granger caused by temperature, humidity, wind, wind gust, atmospheric pressure and radiation.

Hypothesis Testing for Granger causality Zone two

**Table 8: Granger test zone two**

$X^*X_i$	F	P	Null hypothesis rejected
$X_1$	-6 2.8097	0.01012 *	√
$X_2$	-3 0.2966	0.8279	Not
$X_3$	-3 4.5503	0.003519 **	√
$X_4$	-3 2.7022	0.04426 *	√
$X_5$	-6 2.1268	0.04767 *	√
$X_6$	-3 2.5982	0.03723 *	√

Therefore, the granger causality test for zone two shows that all the variables were having a strong significant influence on the causes of dependent variable  $x_0$ . Their level of significant was less than 0.05.

### Zone Three

The hypothesis was that rainfall is not granger caused by temperature, humidity, wind, wind gust, Atmospheric pressure and radiation.

### Hypothesis Testing for Granger causality Zone three

**Table 9: Granger test zone three**

$X^*X_i$	F	P	Null hypothesis rejected.
$X_1$	-3 2.899	0.03396 *	√
$X_2$	-6 1.5613	0.1548	√
$X_3$	-3 4.361	0.004577 **	√
$X_4$	-3 2.681	0.04553 *	√
$X_5$	-6 2.1863	0.04184 *	√
$X_6$	1 -3 2.5628	0.0007054 ***	√

Therefore, the granger causality test for Zone three shows that all the variables had a strong significant influence on the Causes of dependent variable X.

#### Zone Four.

The hypothesis was that rainfall is not granger caused by temperature, humidity, wind, wind gust, Atmospheric pressure and radiation.

### Hypothesis Testing for Granger causality Zone four

**Table 10: Granger test zone four**

$X^*X_i$	F	P	Null hypothesis rejected
$X_1$	-8 2.1204	0.03119 *	√
$X_2$	-8 1.9036	1.032e-05 ***	√
$X_3$	-6 3.0123	0.006269 **	√
$X_4$	-6 2.1685	0.04352 *	√
$X_5$	-6 2.3997	0.026 *	√
$X_6$	-6 1.8441	0.002428 **	√

So, the granger causality test for Zone four shows that all the variables had a strong significant influence on the Causes of dependent variable X.

#### Zone Five

The hypothesis was that rainfall is not granger caused by temperature, humidity, wind, wind gust, Atmospheric pressure and radiation.

### Hypothesis Testing for Granger causality Zone five

$X^*X_i$	F	P	Null hypothesis rejected
$X_1$	-3 3.0811	0.02655 *	√
$X_2$	-3 5.8745	0.0005519 ***	√
$X_3$	-6 2.16	0.04434 *	√
$X_4$	-6 4.4176	0.000196 ***	√
$X_5$	-6 1.8439	0.003104 **	√
$X_6$	-6 1.9563	0.001708 **	√



**Table 11: Granger test zone five****Global Region**

The hypothesis was that rainfall is not granger caused by temperature, humidity, wind, wind gust, Atmospheric pressure and radiation.

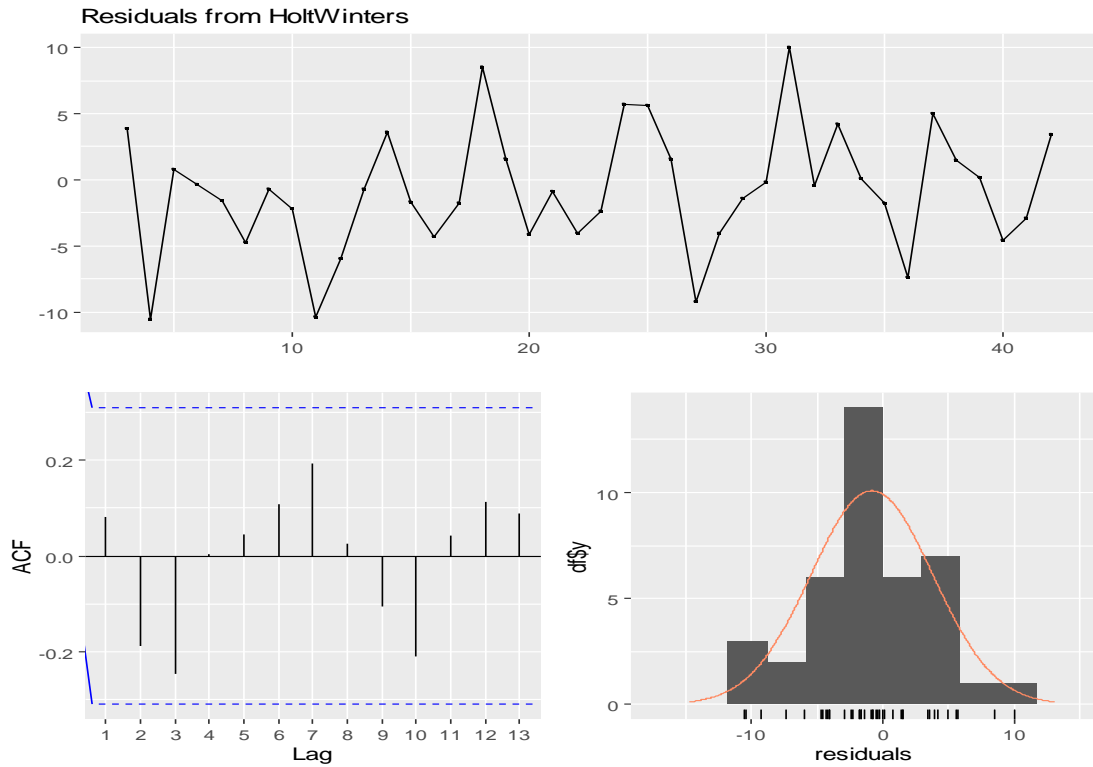
**Hypothesis Testing for Granger causality global****Table 12: Granger Causality test**

<b><math>X^*X_i</math></b>	<b>F</b>	<b>P</b>	<b>Null hypothesis rejected</b>
$X_1$	-3 2.2026	0.01859 *	√
$X_2$	-3 0.4471	0.7194	Not
$X_3$	-3 3.5176	0.01464 *	√
$X_4$	-3 2.5578	0.006 ***	√
$X_5$	-6 2.1318	0.04715*	√
$X_6$	-3 2.6221	0.0274 *	√

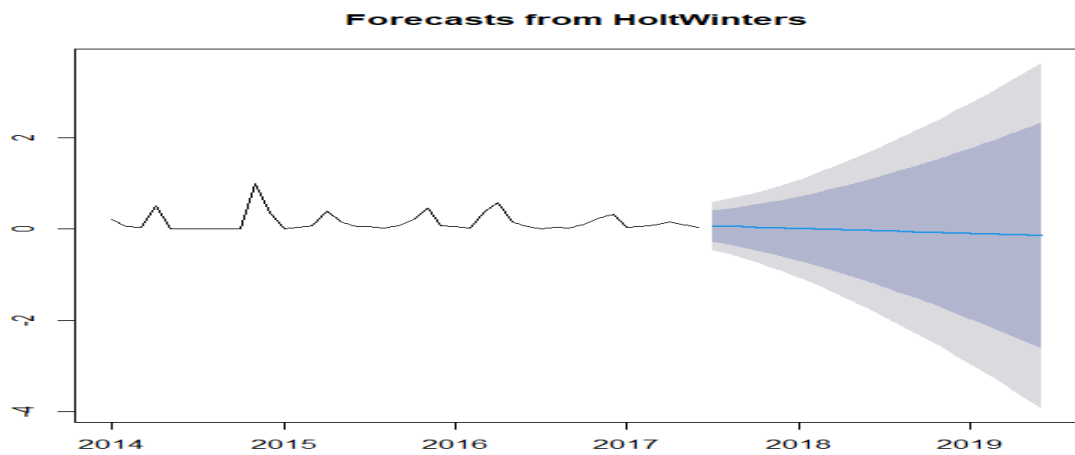
**Appendix II: Figures**

**Holtwinters Forecasting Analysis**

**Zone Two: Forecasting**

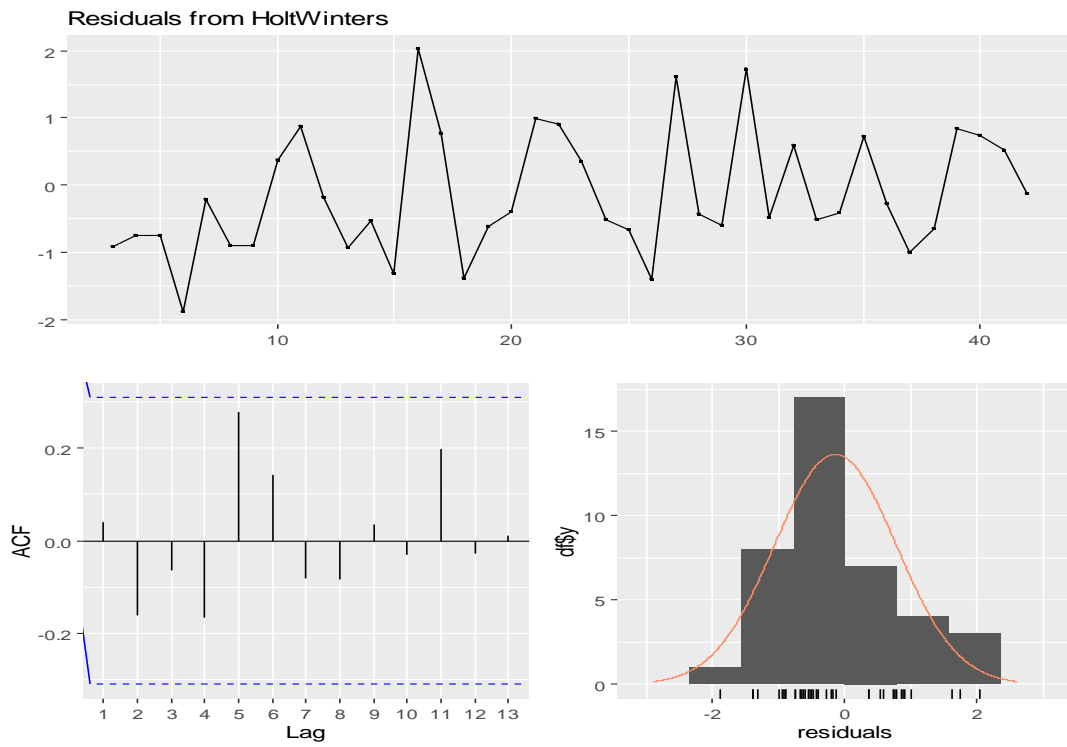


**Figure 4: forecast for zone two**

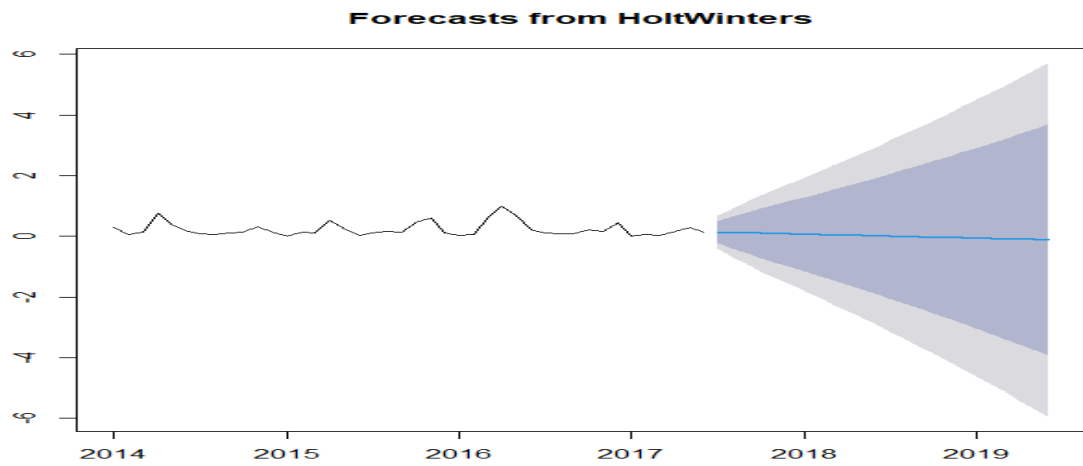


**Figure 5: HoltWinters forecast for zone two**

**Zone Three: Forecasting analysis**



**Figure 6: forecast for zone three**



**Figure 7: HoltWinters forecast for zone three**

### Zone Four: Forecasting Analysis

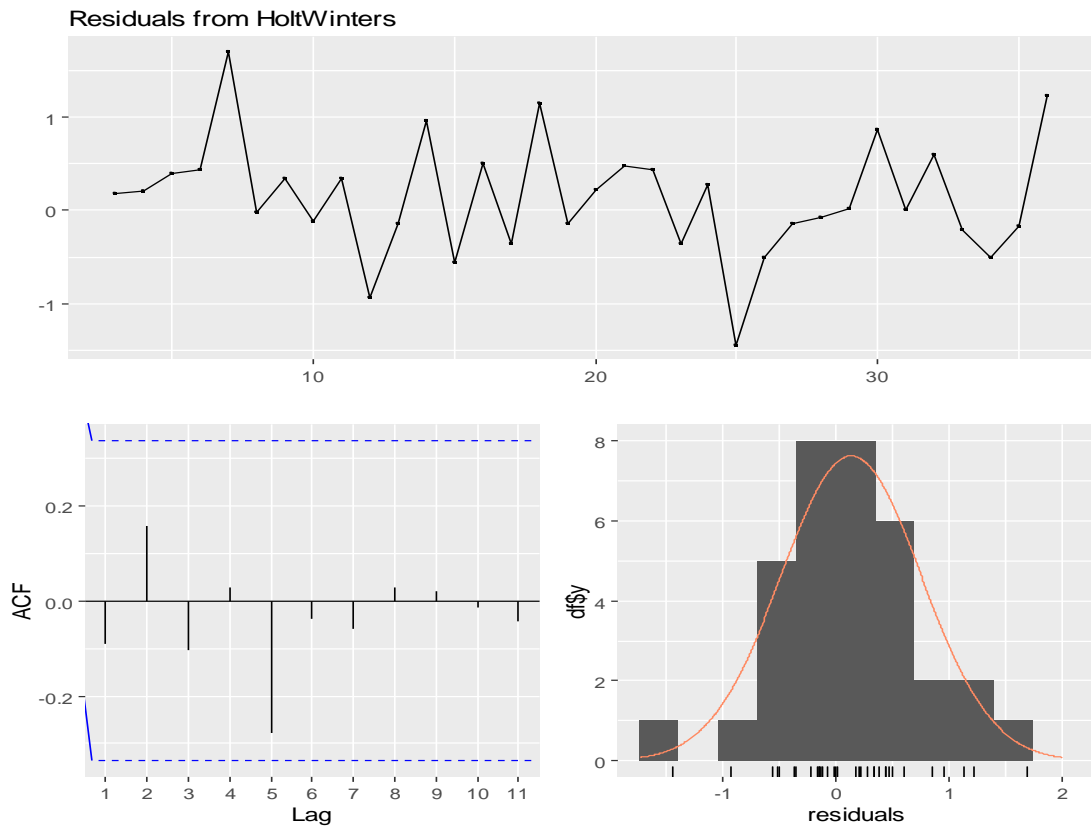


Figure 8: forecast for zone four

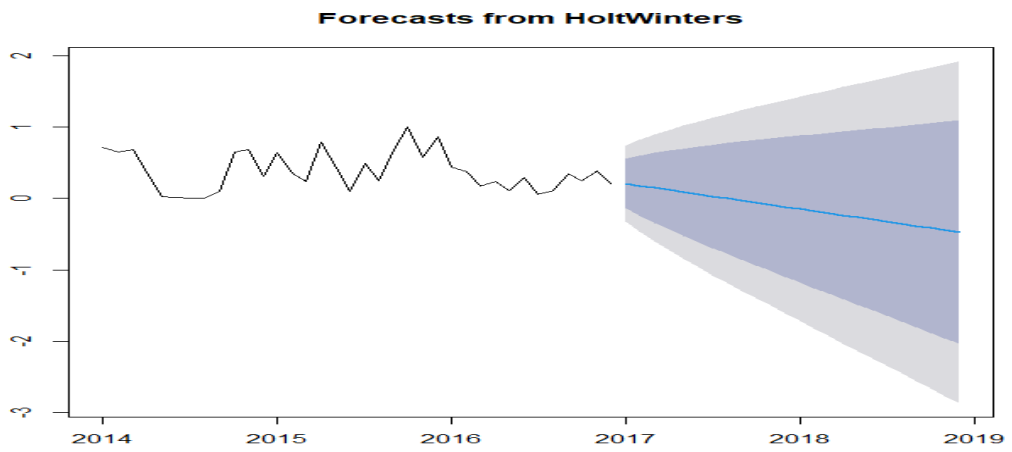
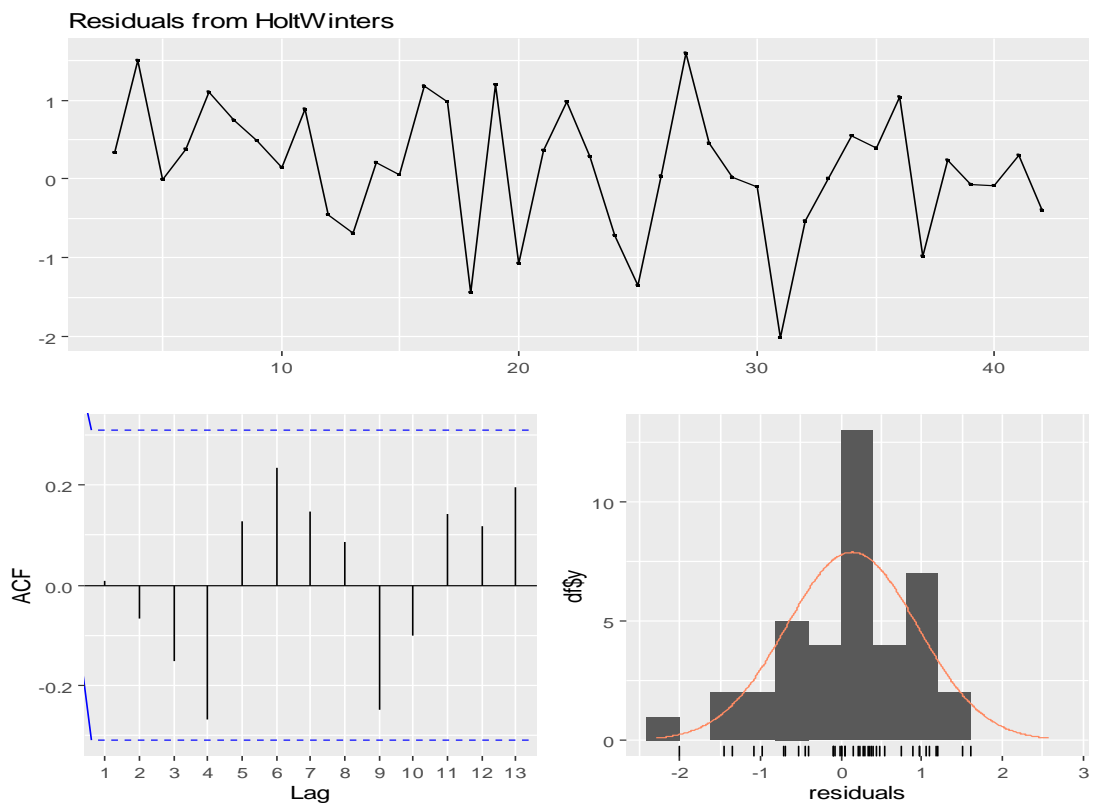
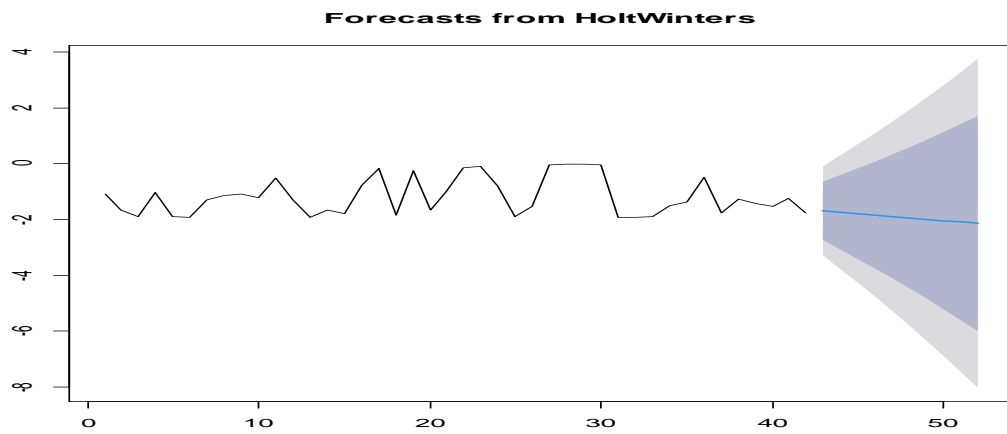


Figure 9: HoltWinters forecast for zone four

**Zone Five: Forecasting analysis**

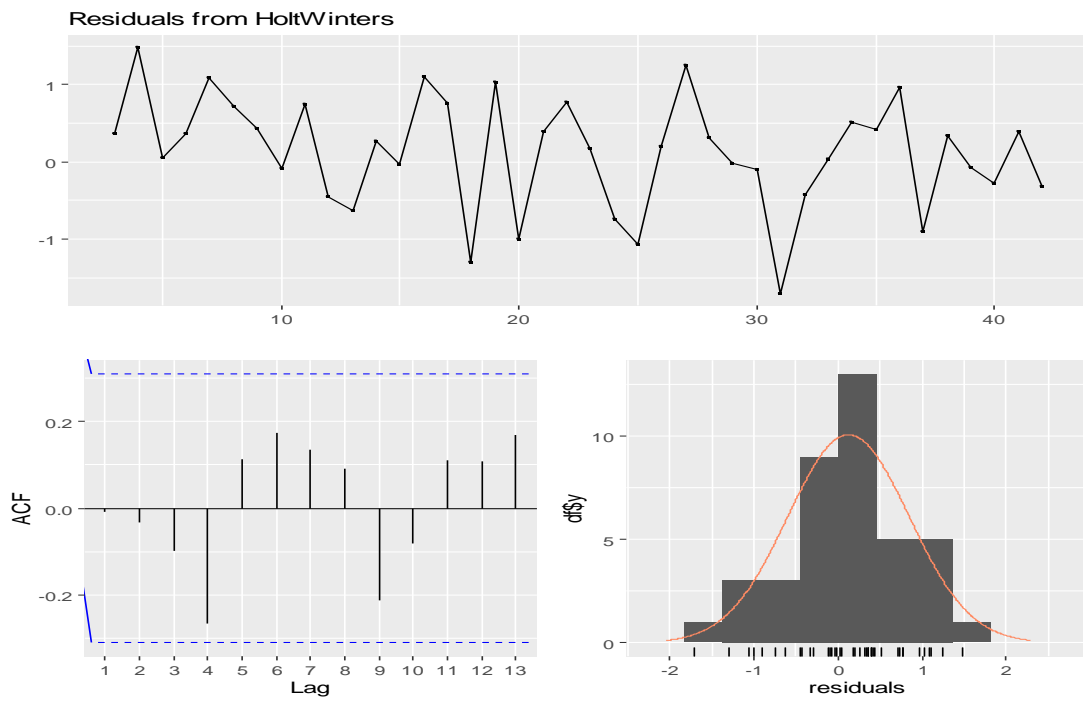


**Figure 10: forecast for zone five**

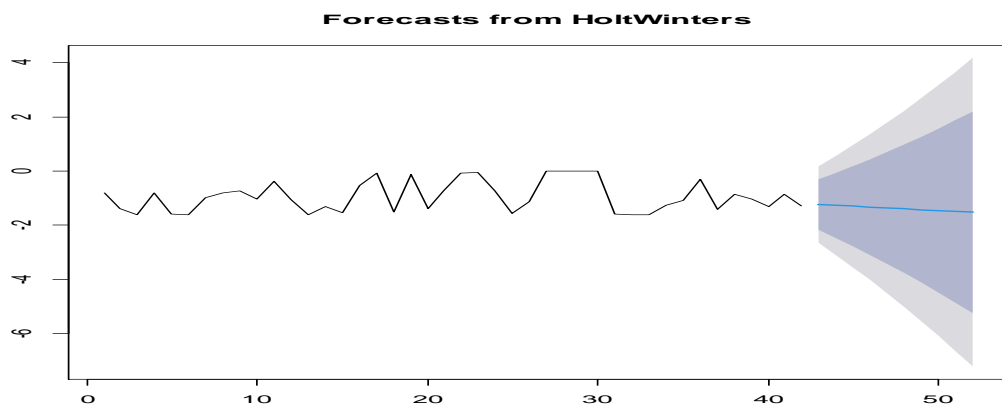


**Figure 11: HoltWinters forecast for zone two**

**Global Region: Forecasting Analysis**



**Figure 12: forecast for Global Region**



**Figure 13: HoltWinters forecast for zone two**

## Appendix III: Output

### Zone One Model Output

Zone one model had the following variables,  $x_{0.l1}$ ,  $x_{1.l1}$ ,  $x_{3.l1}$ ,  $x_{4.l1}$ ,  $x_{6.l1}$ ,  $x_{0.l2}$ ,  $x_{1.l2}$ ,  $x_{3.l2}$ ,  $x_{4.l2}$ ,  $x_{6.l2}$ ,  $x_{0.l3}$ ,  $x_{1.l3}$ ,  $x_{3.l3}$ ,  $x_{4.l3}$  and  $x_{6.l3}$  with a constant. As indicated in table 2 the findings showed variable  $x_{0.l1}$  had t value of -9.303 and the **Pr(>|t|)** of 3.07e-06 \*\*\* variable  $x_{1.l1}$  had t value of -4.527 and the **Pr(>|t|)** of 0.001097 \*\*Variable  $x_{3.l1}$  had t value of -4.217 and the **Pr(>|t|)** of 0.001781 \*\*. Variable  $x_{4.l1}$  had t value of -12.79 and the **Pr(>|t|)** of 1.59e-07 \*\*\* while Variable  $x_{6.l1}$  had t value of -5.368 and the **Pr(>|t|)** of 4.00e-04 \*\*\*. Variable  $x_{0.l2}$  had t value of -2.479 and the **Pr(>|t|)** of 0.032598 \* Variable  $x_{1.l2}$  had t value of -3.777 and the p-value of 0.003621 \*\* Variable  $x_{3.l2}$  had t value of -4.848 and the **Pr(>|t|)** of 0.000673 \*\*\*. while  $x_{4.l2}$  had t-value 0.462 and the p-value of 0.007417 \*\*\* while,  $x_{6.l2}$  had t value of 1.839 and the **Pr(>|t|)** of 0.012372 \*. As indicated in table .... the findings showed variable  $x_{0.l3}$  had t value of 4.905 and the **Pr(>|t|)** of 0.001755 \*\* variable  $x_{1.l3}$  had t value of -4.792 and the **Pr(>|t|)** of 0.000215 \*\*\* Variable  $x_{3.l3}$  had t value of 3.535 and the **Pr(>|t|)** of 4.31e-06 \*\*\*. Variable  $x_{4.l3}$  had t value of 2.263 and the **Pr(>|t|)** of 0.002771 \*\* while Variable  $x_{6.l3}$  had t value of 0.036 and the **Pr(>|t|)** of 0.009524 \*\* and the constant had at value of 0.264 and the **Pr(>|t|)** of 0.007147 \*\*\*.

### Zone Two Model Output

Zone two model had the following variables  $x_{0.l1}$ ,  $x_{1.l1}$ ,  $x_{3.l1}$ ,  $x_{4.l1}$ ,  $x_{6.l1}$ ,  $x_{0.l2}$ ,  $x_{1.l2}$ ,  $x_{3.l2}$ ,  $x_{4.l2}$ ,  $x_{6.l2}$ ,  $x_{0.l3}$ ,  $x_{1.l3}$ ,  $x_{3.l3}$ ,  $x_{4.l3}$  and  $x_{6.l3}$  with a constant. As indicated in table 4 the findings showed variable  $x_{0.l1}$  had t value of 14.477 and the **Pr(>|t|)** of 8.13e-10 \*\*\*, variable  $x_{1.l1}$  had t value of -16.516 and the **Pr(>|t|)** of 1.42e-10 \*\*\*Variable  $x_{3.l1}$  had t value of 7.120 and the **Pr(>|t|)** of 5.17e-06 \*\*\*. Variable  $x_{4.l1}$  had t value of 10.751 and the **Pr(>|t|)** of 3.78e-08 \*\*\*while Variable  $x_{6.l1}$  had t value of 3.031and the **Pr(>|t|)** of 0.008986 \*\*. Variable  $x_{0.l2}$  had t value of 7.290 and the **Pr(>|t|)** of 3.97e-06 \*\*\*, Variable  $x_{1.l2}$  had t value of -7.813 and the p-value of 4.83e-05 \*\*\* Variable  $x_{3.l2}$  had t value of 5.773 and the **Pr(>|t|)** of 4.83e-05 \*\*\* while  $x_{4.l2}$  had t-value 6.099 and the p-value of 2.75e-05 \*\*\* while,  $x_{6.l2}$  had t value of 1.839 and the

**Pr(>|t|)** of 0.087237. As indicated in table 4 the findings showed variable  $x_{0.13}$  had t value of 3.905 and the **Pr(>|t|)** of 0.001585 \*\* variable  $x_{1.13}$  had t value of -4.927 and the **Pr(>|t|)** of 0.000223 \*\*\* Variable  $x_{3.13}$  had t value of 5.530 and the **Pr(>|t|)** of  $7.41e-05$  \*\*\*. Variable  $x_{4.13}$  had t value of 3.623 and the **Pr(>|t|)** of 0.002771 \*\* while Variable  $x_{6.13}$  had t value of 0.063 and the **Pr(>|t|)** of 0.950286 and the constant had a t value of 0.291 and the **Pr(>|t|)** of 0.775018.

### Zone Three Model Output

Zone three model had the following variables,  $x_{0.11}$ ,  $x_{1.11}$ ,  $x_{3.11}$ ,  $x_{4.11}$ ,  $x_{6.11}$ ,  $x_{0.12}$ ,  $x_{1.12}$ ,  $x_{3.12}$ ,  $x_{4.12}$ ,  $x_{6.12}$ ,  $x_{0.13}$ ,  $x_{1.13}$ ,  $x_{3.13}$ ,  $x_{4.13}$  and  $x_{6.13}$  with a constant. As indicated in table 6 the findings showed variable  $x_{0.11}$  had t value of -7.473 and the **Pr(>|t|)** of  $1.24e-05$  \*\*\*, variable  $x_{1.11}$  had t value of -13.156 and the **Pr(>|t|)** of  $4.50e-08$  \*\*\* Variable  $x_{3.11}$  had t value of 8.065 and the **Pr(>|t|)** of  $6.05e-06$  \*\*\* Variable  $x_{4.11}$  had t value of 7.514 and the **Pr(>|t|)** of  $1.18e-05$  \*\*\* while Variable  $x_{6.11}$  had t value of -5.593 and the **Pr(>|t|)** of 0.000162 \*\*\* Variable  $x_{0.12}$  had t value of -7.886 and the **Pr(>|t|)** of  $7.49e-06$  \*\*\*, Variable  $x_{1.12}$  had t value of -10.459 and the p-value of  $4.71e-07$  \*\*\*, Variable  $x_{3.12}$  had t value of 7.347 and the **Pr(>|t|)** of  $1.45e-05$  \*\*\*, while  $x_{4.12}$  had t-value 6.007 and the p-value of  $8.84e-05$  \*\*\*, while  $x_{6.12}$  had t value of -4.470 and the **Pr(>|t|)** of 0.000946 \*\*\*. As indicated in table 6 the findings showed variable  $x_{0.13}$  had t value of -9.221 and the **Pr(>|t|)** of  $1.65e-06$  \*\*\*, variable  $x_{1.13}$  had t value of -5.518 and the **Pr(>|t|)** of 0.000181 \*\*\*, Variable  $x_{3.13}$  had t value of 6.223 and the **Pr(>|t|)** of  $6.51e-05$  \*\*\*, Variable  $x_{4.13}$  had t value of 4.314 and the **Pr(>|t|)** of 0.001227 \*\*, while Variable  $x_{6.13}$  had t value of -5.010 and the **Pr(>|t|)** of 0.000396 \*\*\* and the constant had a t value of -0.654 and the **Pr(>|t|)** of 0.526497.

### Zone Four Model Output

Zone four model had the following variables,  $x_{0.11}$ ,  $x_{1.11}$ ,  $x_{2.11}$ ,  $x_{3.11}$ ,  $x_{4.11}$ ,  $x_{5.11}$ ,  $x_{6.11}$ ,  $x_{0.12}$ ,  $x_{1.12}$ ,  $x_{2.12}$ ,  $x_{3.12}$ ,  $x_{4.12}$ ,  $x_{5.12}$  and  $x_{6.12}$  with a constant. As indicated in table 8 the findings showed variable  $x_{0.11}$  had t value of -3.614 and the **Pr(>|t|)** of 0.004738 \*\* variable  $x_{1.11}$  had t value of -3.090 and the **Pr(>|t|)** of 0.011449 \*, Variable  $x_{2.11}$  had t value of 5.410 and the **Pr(>|t|)** of 0.000297 \*\*\*, Variable



$x_{3.l1}$  had t value of 1.868 and the  $\text{Pr}(>|t|)$  of 0.091323. Variable  $x_{4.l1}$  had t value of -2.976 and the  $\text{Pr}(>|t|)$  of 0.013913 \* Variable  $x_{5.l1}$  had t value of 2.666 and the  $\text{Pr}(>|t|)$  of 0.023659 \* while Variable  $x_{6.l1}$  had t value of 2.240 and the  $\text{Pr}(>|t|)$  of 0.049037 \* Variable  $x_{0.l2}$  had t value of --3.310 and the  $\text{Pr}(>|t|)$  of 0.007883 \*\* Variable  $x_{1.l2}$  had t value of -3.113 and the  $\text{Pr}(>|t|)$  of 0.011006 \*, Variable  $x_{2.l2}$  had t value of 3.222 and the p-value of 0.009137 \*\* Variable  $x_{3.l2}$  had t value of -4.649 and the  $\text{Pr}(>|t|)$  of 0.000909 \*\*\*, while  $x_{4.l2}$  had t-value -3.178 and the p-value of 0.009860 \*\*, variable  $x_{5.l2}$  had t value of 1.087 and the  $\text{Pr}(>|t|)$  of 0.302471, while Variable  $x_{6.l2}$  had t value of -3.605 and the  $\text{Pr}(>|t|)$  of 0.004809 \*\* and the constant had a t value of 0.179 and the  $\text{Pr}(>|t|)$  of 0.861135.

### Zone Five Model Output

Zone five models had the following variables;  $x_{0.l1}$ ,  $x_{1.l1}$ ,  $x_{2.l1}$ ,  $x_{3.l1}$ ,  $x_{4.l1}$ ,  $x_{5.l1}$  and  $x_{6.l1}$  with a constant. As indicated in table 10 the findings showed variable  $x_{0.l1}$  had t value of 3.782 and the  $\text{Pr}(>|t|)$  of 0.00359 \*\*, variable  $x_{1.l1}$  had t value of -0.414 and the  $\text{Pr}(>|t|)$  of 0.68793 Variable  $x_{2.l1}$  had t value of -10.972 and the  $\text{Pr}(>|t|)$  of 6.75e-07 \*\*\* Variable  $x_{3.l1}$  had t value of -4.428 and the  $\text{Pr}(>|t|)$  of 0.00128 \*\*. Variable  $x_{4.l1}$  had t value of -3.282 and the  $\text{Pr}(>|t|)$  of 0.00827 \*\* Variable  $x_{5.l1}$  had t value of -2.779 and the  $\text{Pr}(>|t|)$  of 0.01948 \* while Variable  $x_{6.l1}$  had t value of 2.420 and the  $\text{Pr}(>|t|)$  of 0.03606 \* and the constant had at value of 0.430 and the  $\text{Pr}(>|t|)$  of 0.67620.

### Global Vector Model Output

Global region models had the following variables;  $x_{0.l1}$ ,  $x_{1.l1}$ ,  $x_{3.l1}$ ,  $x_{4.l1}$ ,  $x_{6.l1}$ ,  $x_{0.l2}$ ,  $x_{1.l2}$ ,  $x_{3.l2}$ ,  $x_{4.l2}$ ,  $x_{6.l2}$ ,  $x_{0.l3}$ ,  $x_{1.l3}$ ,  $x_{3.l3}$ ,  $x_{4.l3}$  and  $x_{6.l3}$  with a constant. As indicated in table 12 the findings showed variable  $x_{0.l1}$  had t value of -9.475 and the  $\text{Pr}(>|t|)$  of 1.26e-06 \*\*\* variable  $x_{1.l1}$  had t value of 6.800 and the  $\text{Pr}(>|t|)$  of 2.95e-05 \*\*\*, Variable  $x_{3.l1}$  had t value of -10.894 and the  $\text{Pr}(>|t|)$  of 3.12e-07 \*\*\*, Variable  $x_{4.l1}$  had t value of -8.497 and the  $\text{Pr}(>|t|)$  of 3.67e-06 \*\*\*, Variable  $x_{6.l1}$  had t value of 7.895 and the  $\text{Pr}(>|t|)$  of 7.41e-06 \*\*\*, Variable  $x_{0.l2}$  had t value of --5.327 and the  $\text{Pr}(>|t|)$  of 0.000242 \*\*\*, Variable  $x_{1.l2}$  had t value of 2.999 and the  $\text{Pr}(>|t|)$  of 0.012095 \*, Variable  $x_{3.l2}$  had t value of -10.905 and the  $\text{Pr}(>|t|)$  of 3.09e-07 \*\*\*, while  $x_{4.l2}$  had t-value

-6.636 and the p-value of  $3.68e-05$  \*\*\*, while Variable  $x_{6.l2}$  had t value of 8.698 and the  $\text{Pr}( > |t| )$  of  $2.92e-06$  \*\*\*, variable  $x_{0.l3}$  had t value of -5.192 and the  $\text{Pr}( > |t| )$  of 0.000298 \*\*\*, variable  $x_{1.l3}$  had t value of 3.991 and the  $\text{Pr}( > |t| )$  of 0.002117 \*\*, variable  $x_{3.l3}$  had t value of -8.602 and the  $\text{Pr}( > |t| )$  of  $3.26e-06$  \*\*\*, variable  $x_{4.l3}$  had t value of -3.174 and the  $\text{Pr}( > |t| )$  of 0.008863 \*\*, variable  $x_{6.l3}$  had t value of 8.248 and the  $\text{Pr}( > |t| )$  of  $4.88e-06$  \*\*\*, and the constant had a t value of -0.045 and the  $\text{Pr}( > |t| )$  of 0.000156 \*\*\*.

### The Roots of the characteristic polynomial for zone one

VAR Estimation Results:

=====

Endogenous variables:  $x_0, x_1, x_2, x_3, x_4, x_5, x_6$

Deterministic variables: const

Sample size: 1457

Log Likelihood: 1079.708

Roots of the characteristic polynomial:

0.4781 0.4782 0.4649 0.4289 0.4288 0.4287 0.3997 0.3936 0.3935 0.3659  
 0.3658 0.3526 0.3528 0.3085 0.3185 0.2918 0.2915 0.2899 0.2555 0.2434  
 0.2432

Call:

VAR(y = Data1, p = 3, type = "const", exogen = NULL)

### Zone one Residual standard error

Residual standard error: 0.9763 on 14 degrees of freedom

Multiple R-Squared: 0.9895, Adjusted R-squared: 0.9761

F-statistic: 1919 on 15 and 14 DF, p-value:  $< 1.27e-16$

### Model diagnostic test

#### Serial correlation test

Portmanteau Test (asymptotic)

data: Residuals of VAR object model

Chi-squared = 59.144, df = 14, p-value = 0.152

## Heteroscedasticity test

ARCH (multivariate)

data: Residuals of VAR object model

Chi-squared = 9784.2, df = 14, p-value < 1.2e-16

## Normality test

\$JB JB-Test (multivariate)

data: Residuals of VAR object model

Chi-squared = 1074307, df = 14, p-value < 2.2e-16

\$Skewness

Skewness only (multivariate)

data: Residuals of VAR object model

Chi-squared = 303.24, df = 7, p-value < 2.2e-16

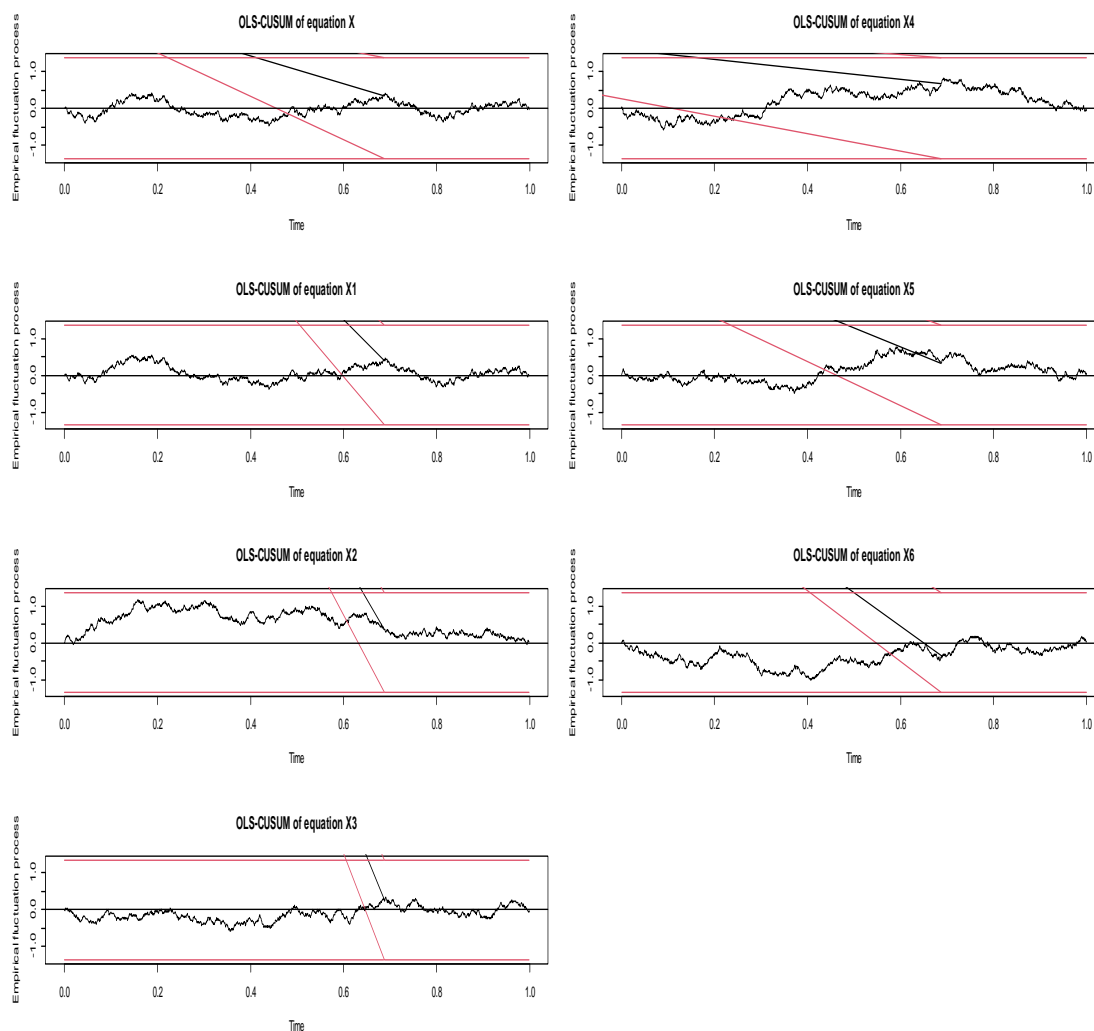
\$Kurtosis

Kurtosis only (multivariate)

data: Residuals of VAR object model

Chi-squared = 1074004, df = 7, p-value < 2.2e-16

## Stability test for Zone one



## Zone two output

### The Roots of the characteristic polynomial for zone two

```

VAR Estimation Results:
=====

Endogenous variables: x0, x1, x2, x3, x4, x5, x6

Deterministic variables: const

Sample size: 1457

Log Likelihood: 762.749
Roots of the characteristic polynomial:
0.8092 0.4988 0.4987 0.4601 0.4222 0.4221 0.4163 0.4161 0.4104 0.3603
0.3613 0.3479 0.3477 0.3440 0.3439 0.3213 0.3211 0.3152 0.3151 0.1289
0.1287
Call:
VAR(y = Data1, p = 3, type = "const", exogen = NULL)

```

### Zone two Residual standard error

```

Residual standard error: 0.6973 on 14 degrees of freedom

Multiple R-Squared: 0.9995, Adjusted R-squared: 0.9961

F-statistic: 1919 on 15 and 14 DF, p-value: < 1.3e-16

```

## Model diagnostic test

### Serial correlation test

```

Portmanteau Test (asymptotic)

data: Residuals of VAR object model

Chi-squared = 107.05, df = 14, p-value = 0.1328 .

```

### Heteroscedasticity test

```

ARCH (multivariate)

data: Residuals of VAR object model

Chi-squared = 8556.8, df = 14, p-value < 0.1622 .

```

### Normality test

```

$JB
JB-Test (multivariate)
data: Residuals of VAR object model

Chi-squared = 848864, df = 14, p-value < 1.2e-10

$Skewness

Skewness only (multivariate)

data: Residuals of VAR object model

Chi-squared = 3399.1, df = 7, p-value < 4.3e-9

```

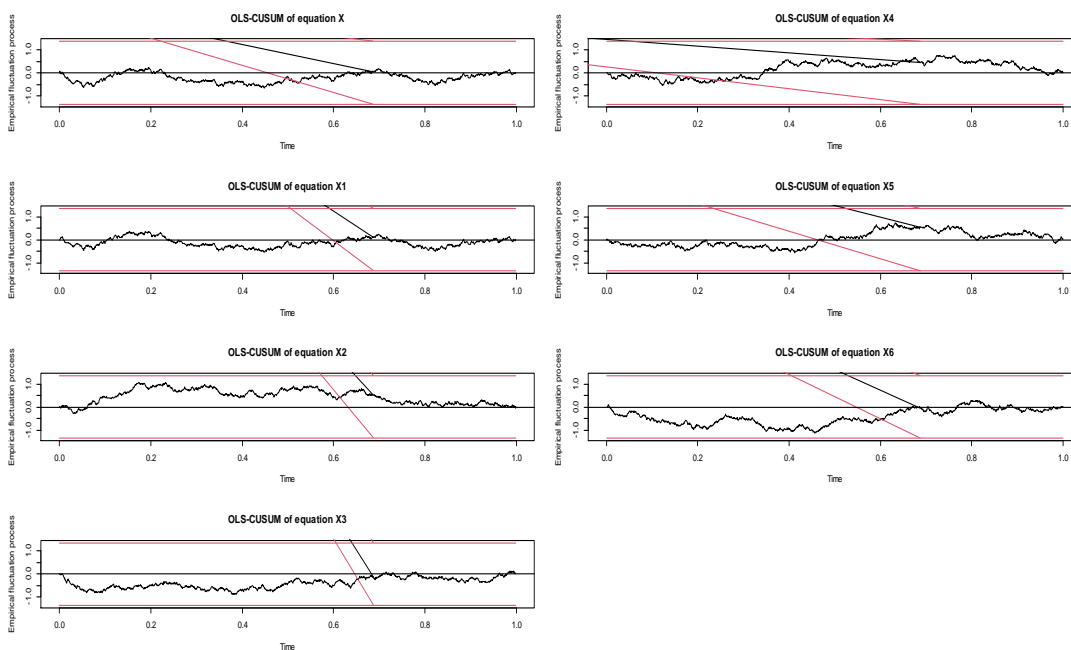
\$Kurtosis

Kurtosis only (multivariate)

data: Residuals of VAR object model

Chi-squared = 845464, df = 7, p-value < 2.2e-11

### Stability test graph for Zone two



### Zone three output

#### Zone three Roots of the characteristic polynomial

VAR Estimation Results:

=====

Endogenous variables:  $x_0, x_1, x_2, x_3, x_4, x_5, x_6$

Deterministic variables: const

Sample size: 1457

Log Likelihood: 1090.255

Roots of the characteristic polynomial:

0.8296 0.4461 0.412 0.422 0.4123 0.4122 0.3735 0.352 0.352 0.3459  
 0.3459 0.34 0.34 0.3301 0.3311 0.3248 0.3247 0.2996 0.2996 0.1691  
 0.1691

Call:

VAR(y = Data1, p = 3, type = "const", exogen = NULL)

#### Zone three Residual standard error

Residual standard error: 1.812 on 11 degrees of freedom

Multiple R-Squared: 0.9997, Adjusted R-squared: 0.9994

F-statistic: 2745 on 15 and 11 DF, p-value: < 3.2e-16

## Model three diagnostic test

### Serial correlation test

```
Portmanteau Test (asymptotic)
data: Residuals of VAR object model
Chi-squared = 94.003, df = 14, p-value = 0.0001167
```

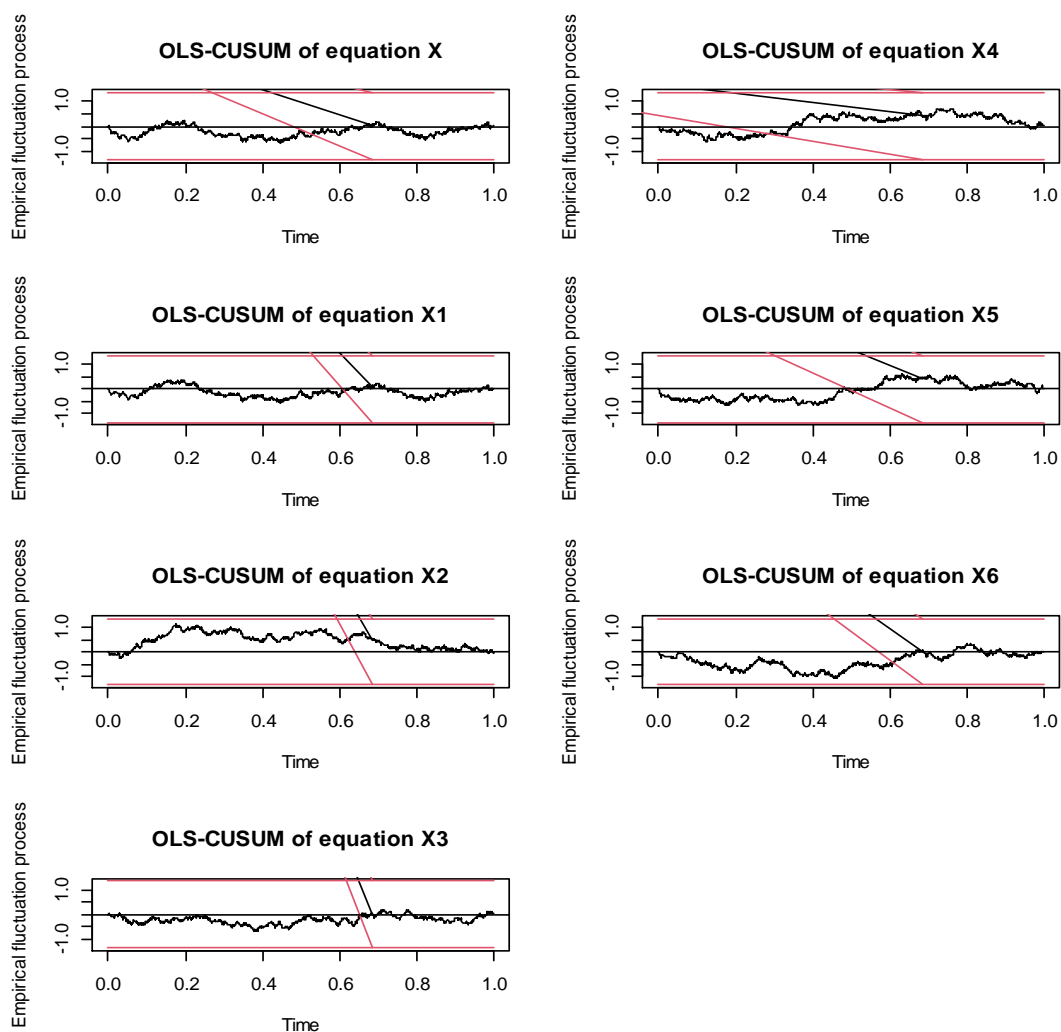
### Heteroscedasticity test

```
ARCH (multivariate)
data: Residuals of VAR object model
Chi-squared = 6860.3, df = 14, p-value < 1.4e-16
```

### Normality test

```
$JB
JB-Test (multivariate)
data: Residuals of VAR object model
Chi-squared = 769276, df = 14, p-value < 2.1e-16
$Skewness
Skewness only (multivariate)
data: Residuals of VAR object model
Chi-squared = 661.84, df = 7, p-value < 1.3e-16
$Kurtosis
Kurtosis only (multivariate)
data: Residuals of VAR object model
Chi-squared = 768614, df = 7, p-value < 2.2e-16
```

### Stability test graph for Zone three



### Zone four model

#### The Roots of the characteristic polynomial output for zone four

VAR Estimation Results:

=====

Endogenous variables:  $x_0, x_1, x_2, x_3, x_4, x_5, x_6$

Deterministic variables: const

Sample size: 1457

Log Likelihood: 609.47

Roots of the characteristic polynomial:

0.9207 0.4219 0.4066 0.3963 0.3913 0.3913 0.3743 0.3568 0.3553 0.257  
0.2474 0.18 0.09407 0.09417

Call:

VAR(y = Data1, p = 2, type = "const", exogen = NULL)

### **Zone four Residual standard error**

Residual standard error: 3.008 on 10 degrees of freedom

Multiple R-Squared: 0.9975,      Adjusted R-squared: 0.994

F-statistic: 282.9 on 14 and 10 DF, p-value: 4.622e-11

### **Model diagnostic test for zone four**

#### **Serial correlation test**

Portmanteau Test (asymptotic)

data: Residuals of VAR object model

Chi-squared = 131.79, df = 14, p-value = 1.624e-09

#### **Heteroscedasticity test**

ARCH (multivariate)

data: Residuals of VAR object model

Chi-squared = 8684.1, df = 14, p-value < 2.2e-16

#### **Normality test**

JB-Test (multivariate)

data: Residuals of VAR object model

Chi-squared = 1099592, df = 14, p-value < 2.2e-16

\$Skewness

Skewness only (multivariate)

Data: Residuals of VAR object model

Chi-squared = 839.71, df = 7, p-value < 2.2e-16

\$Kurtosis

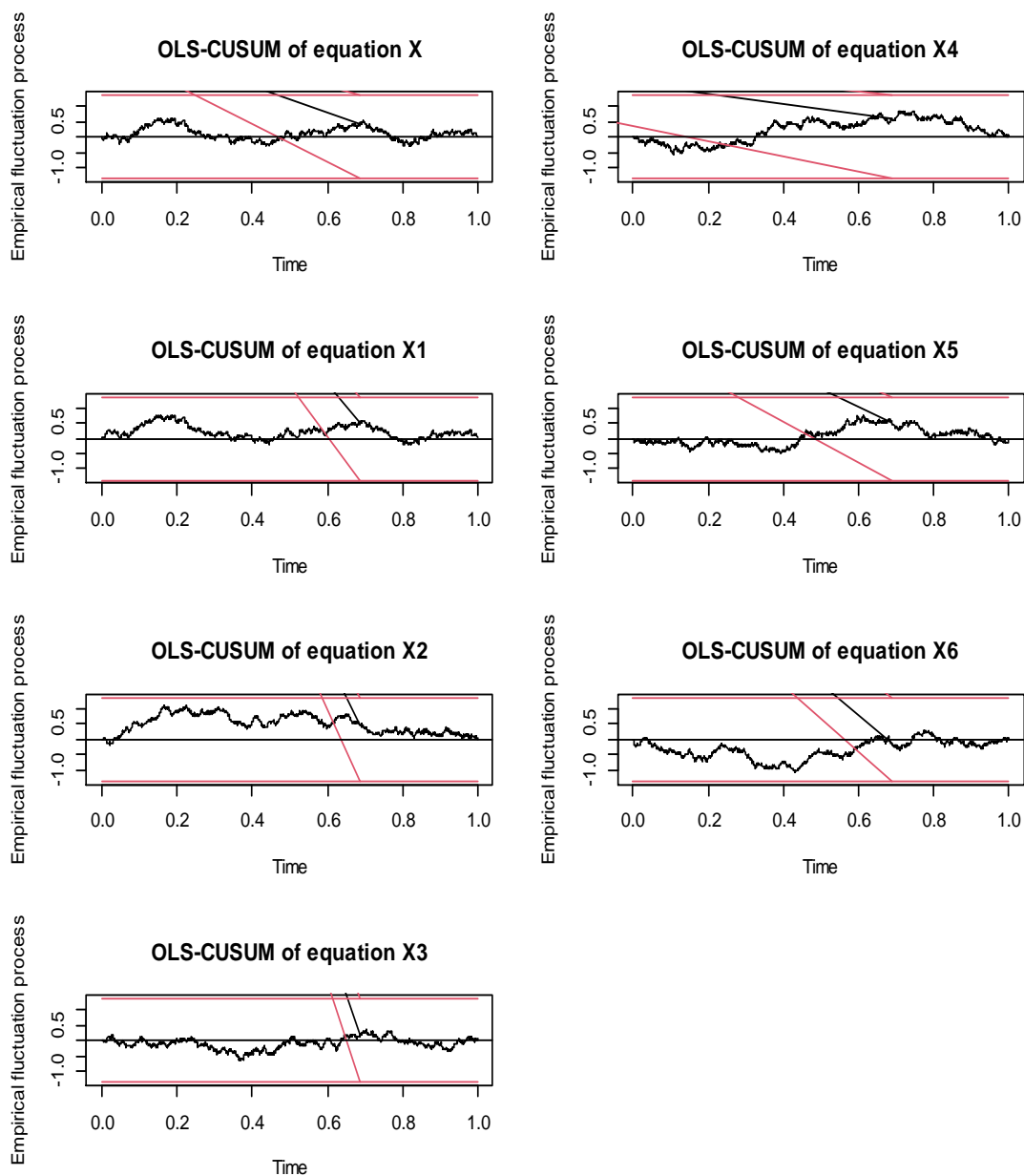
Kurtosis only (multivariate)

data: Residuals of VAR object model

Chi-squared = 1098752, df = 7, p-value < 2.2e-16



### Stability test graph for zone four



### Zone five model

#### The Roots of the characteristic polynomial for zone five

```

VAR Estimation Results:
=====
Endogenous variables: x0, x1, x2, x3, x4, x5, x6
Deterministic variables:
const
Sample size: 1456
Log Likelihood: -9521.494
    
```

```

Roots of the characteristic polynomial:
0.7837 0.7449 0.7449 0.7419 0.7419 0.7225 0.6662

Call:
VAR(y = dat2, p = 1, type = "const", exogen = NULL)

```

### Zone five Residual standard error

```

Residual standard error: 4.463 on 10 degrees of freedom
Multiple R-Squared: 0.9896, Adjusted R-squared: 0.9823
F-statistic: 136.1 on 7 and 10 DF, p-value: 4.161e-09

```

### Model diagnostic test for zone five

#### Serial correlation test

```

Portmanteau Test (asymptotic)
data: Residuals of VAR object model
Chi-squared = 50.885, df = 7, p-value = 0.3993

```

#### Heteroscedasticity test

```

ARCH (multivariate)
data: Residuals of VAR object model
Chi-squared = 5813.5, df = 7, p-value < 0.2216

```

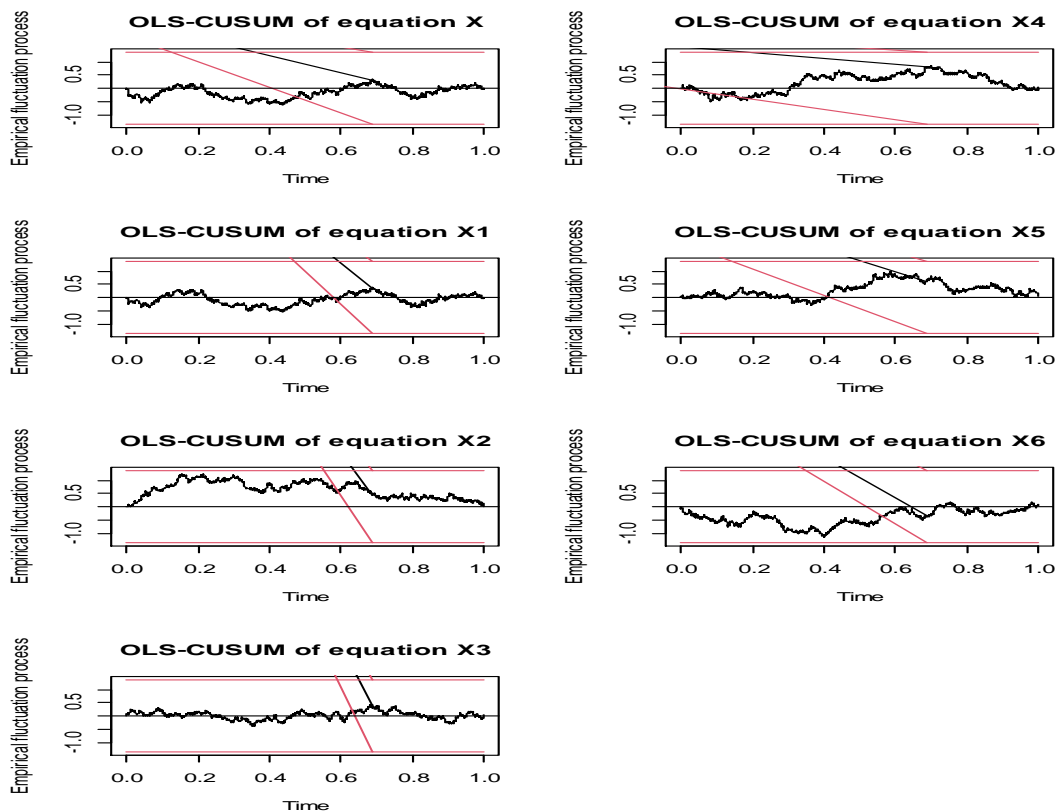
#### Normality test

```

$JB
JB-Test (multivariate)
data: Residuals of VAR object model
Chi-squared = 1286491, df = 7, p-value < 4.23e-1
$Skewness
Skewness only (multivariate)
data: Residuals of VAR object model
Chi-squared = 5743.7, df = 7, p-value < 1.5e-7
$Kurtosis
Kurtosis only (multivariate)
data: Residuals of VAR object model
Chi-squared = 1280747, df = 7, p-value < 3.26e-10

```

### Stability test graph for zone five



### Global vector output

#### Roots of the characteristic polynomial

VAR Estimation Results:

=====

Endogenous variables:  $x_0, x_1, x_2, x_3, x_4, x_5, x_6$

Deterministic variables: const

Sample size: 1457

Log Likelihood: 633.572

Roots of the characteristic polynomial:

0.7834 0.4474 0.423 0.423 0.4173 0.4173 0.4105 0.4105 0.3943 0.36 0.36  
 0.336 0.336 0.3359 0.3359 0.3256 0.3256 0.3248 0.3248 0.2634 0.2634

Call:

VAR(y = Data1, p = 3, type = "const", exogen = NULL)

## Global model

### Residual standard error

Residual standard error: 5.84 on 11 degrees of freedom  
 Multiple R-Squared: 0.9496, Adjusted R-squared: 0.969  
 F-statistic: 1769 on 15 and 11 DF, p-value: < 2.2e-16

### Model diagnostic test for global vector.

#### Serial correlation test

Portmanteau Test (asymptotic)  
 data: Residuals of VAR object model  
 Chi-squared = 102.3, df = 14, p-value = 1.256e-05

#### Heteroscedasticity test

ARCH (multivariate)  
 data: Residuals of VAR object model  
 Chi-squared = 7799.6, df = 14, p-value < 3.21e-6

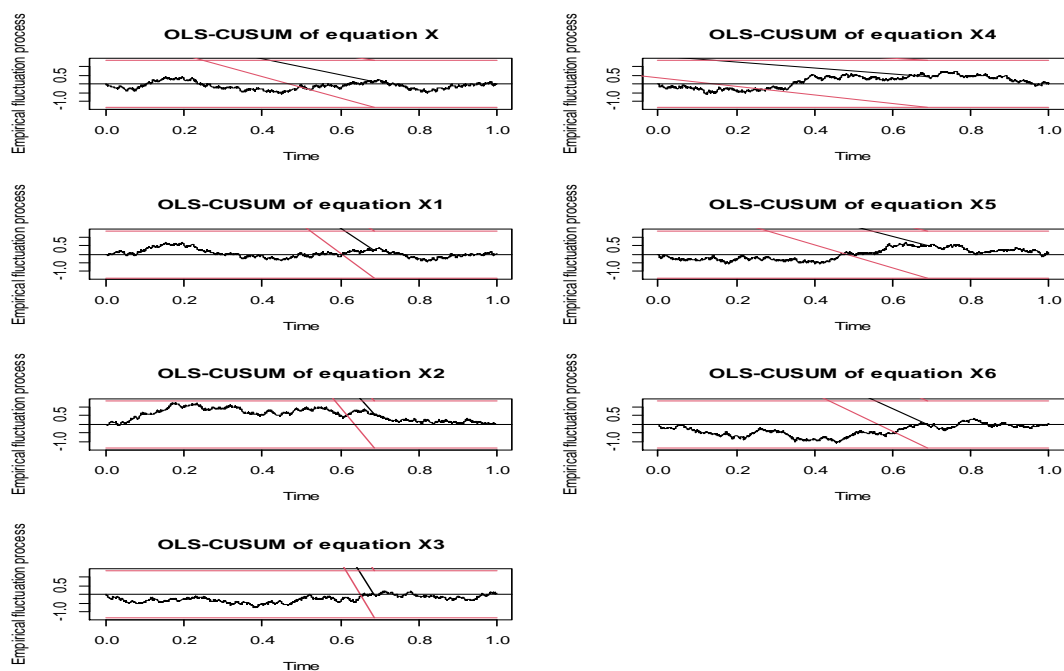
#### Normality test

JB-Test (multivariate)  
 data: Residuals of VAR object model  
 Chi-squared = 493886, df = 14, p-value < 6.2e-10  
 \$Skewness

Skewness only (multivariate)  
 data: Residuals of VAR object model  
 Chi-squared = 2639.4, df = 7, p-value < 4.13e-7  
 \$Kurtosis

Kurtosis only (multivariate)  
 data: Residuals of VAR object model  
 Chi-squared = 491247, df = 7, p-value < 2.2e-16

## Stability test



## Performance Measures

### Accuracy measure in different zones,

ZONE 1 showed MAE was 0.897018, MSE of 0.9131199, RMSE of 0.9555731, AIC of 33.37781 and BIC of 42.70324.

ZONE 2 showed MAE was 0.8354289, MSE of 0.7486969, RMSE of 0.8652727, AIC of 31.32569 and BIC of 43.48938.

ZONE 3 showed MAE was 0.9090333, MSE of 0.778894, RMSE of 0.8825497, AIC of 37.57943, and BIC of 49.74311.

ZONE 4 showed MAE of 0.7339015 MSE of 0.698275, RMSE of 0.835628, AIC of 45.21378, and BIC of 56.29842.

ZONE 5 showed MAE of 0.9006911 MSE of 0.9596382, RMSE of 0.9796113, AIC of 33.37781 and BIC of 42.70324.

GLOBAL VEC showed MAE of 0.889702 MSE of 0.923199, RMSE of 0.960832, AIC of 31.73782, BIC of 43.90374.

### F- Test Comparing variances of Actual and Predicted

#### **ZONE ONE**

```
a<-
c(0.53396298,0.14318096,0.06667926,0.15531923,0.30681526,0.11578769,0.342674
73,0.43071402,
0.14357764,0.32300049)
```

```
a
```

```
p<-
```

```
c(0.53604143,0.31297298,0.22819949,0.19788896,0.52991659,0.07328579,0.529836
39,
0.60557426,0.31468782,0.36519333)
```

```
p
```

#### **#Results**

```
var.test(a,p)
```

*F test to compare two variances*

data: a and p  
 F = 0.76676, num df = 9, denom df = 9, p-value = 0.6988  
 alternative hypothesis: true ratio of variances is not equal to 1  
 95 percent confidence interval:  
 0.1904516 3.0869587

sample estimates:  
 ratio of variances 0.7667569

#### **ZONE TWO**

var.test(a,p)  
 F test to compare two variances

data: a and p  
 F = 0.67045, num df = 9, denom df = 9, p-value = 0.5609  
 alternative hypothesis: true ratio of variances is not equal to 1  
 95 percent confidence interval:  
 0.1665307 2.6992350

sample estimates:  
 ratio of variances 0.6704518

#### **ZONE THREE**

var.test(a,p)  
 F test to compare two variances

data: a and p  
 F = 0.46531, num df = 9, denom df = 9, p-value = 0.2698  
 alternative hypothesis: true ratio of variances is not equal to 1  
 95 percent confidence interval:  
 0.1155753 1.8733177

sample estimates:  
 ratio of variances 0.4653056

#### **ZONE FOUR**

var.test(a,p)  
 F test to compare two variances

data: a and p  
 F = 1.1773, num df = 9, denom df = 9, p-value = 0.8119  
 alternative hypothesis: true ratio of variances is not equal to 1  
 95 percent confidence interval:  
 0.2924151 4.7396473

sample estimates:  
 ratio of variances 1.177261

#### **ZONE FIVE**

var.test(a,p)  
 F test to compare two variances

data: a and p  
 F = 1.8116, num df = 9, denom df = 9, p-value = 0.3893  
 alternative hypothesis: true ratio of variances is not equal to 1  
 95 percent confidence interval:  
 0.4499753 7.2934822

sample estimates:  
 ratio of variances 1.811598

#### **GLOBAL VECTOR**

var.test(a,p)  
 F test to compare two variances

data: a and p  
 F = 0.76676, num df = 9, denom df = 9, p-value = 0.6988  
 alternative hypothesis: true ratio of variances is not equal to 1  
 95 percent confidence interval:  
 0.1904516 3.0869587

sample estimates:  
 ratio of variances 0.7667569

Training data

	$X_0$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
2	0.453955751	0.166666667	1.0000000	0.32352941	0.9217221	0.02096680	0.1679688
3	0.605699062	0.24358974	0.6460107	0.79411765	0.6497065	1.00000000	0.1289062
4	0.403181590	0.23076923	0.6789415	0.61764706	0.6281800	0.52145215	0.0000000
7	0.153171937	0.20512821	0.5250445	0.82352941	0.5146771	0.51252184	0.6484375

```

8 0.082358392 0.11538462 0.7073847 0.70588235 0.6183953 0.91322073 0.1953125
9 0.007027298 0.06410256 0.6075365 0.82352941 0.5146771 0.92351000 0.2382812
10 0.000000000 0.08974359 0.8090807 0.29411765 0.6790607 0.00000000 0.2226562
11 0.018263253 0.00000000 0.0000000 0.14705882 0.7710372 0.02834401 0.3828125
12 0.070543264 0.00000000 0.9946545 0.02941176 0.9941292 0.03222675 0.5117188
13 0.114405962 0.89743590 0.4738996 0.70588235 0.1878669 0.43409047 0.1835937
16 0.263214796 0.94871795 0.3947733 0.85294118 0.0000000 0.51562803 0.1132812
17 0.284953087 1.00000000 0.3512836 0.85294118 0.1624266 0.02795574 0.1523438
18 0.405343836 0.96153846 0.4595130 0.76470588 0.2544031 0.19685498 0.1679688
19 0.219043206 0.83333333 0.6954398 0.52941176 0.4383562 0.20093186 0.2109375
20 0.290783428 0.83333333 0.5997492 0.61764706 0.5107632 0.37817899 0.2109375
22 0.342522877 0.83333333 0.5288062 0.64705882 0.5381605 0.21199767 0.1679688
23 0.033669254 0.85897436 0.7277767 0.23529412 0.7142857 0.07202485 0.2187500
24 0.033823700 0.83333333 0.8427374 0.05882353 0.8747554 0.08328480 0.3046875
25 0.037105680 0.84615385 0.8334323 0.00000000 1.0000000 0.09124442 0.3164062
28 0.167110699 0.75641026 0.1323170 1.00000000 0.5048924 0.19180742 0.1015625
> testing data

```

```

      X0      X1      X2      X3      X4      X5      X6
1 0.09834357 0.2051282 0.9555864 0.2941176 0.9452055 0.09939818 0.1210937
5 1.00000000 0.3076923 0.3451462 0.9117647 0.4227006 0.92583964 0.2929687
6 0.59967566 0.1923077 0.4892761 0.8529412 0.4422701 0.96000777 1.0000000
14 0.60504267 0.9358974 0.3607206 0.8529412 0.3287671 0.43777907 0.0781250
15 0.47565543 0.83333333 0.5086122 0.8823529 0.2446184 0.51718113 0.1132812
21 0.13904012 0.8076923 0.6607932 0.5294118 0.6203523 0.05785284 0.2460938
26 0.15614502 0.83333333 0.7094305 0.2647059 0.9941292 0.55193166 0.2851562
27 0.18935094 0.8589744 0.5736158 0.5588235 0.7710372 0.21976315 0.2343750

```

```

=====
=====
                                     Dependent variable:
-----
                                     X
(6)      (7)      (1)      (2)      (3)      (4)      (5)
-----
X0.11      0.252***  -0.538***  -0.041  0.180**  -0.083
0.190**    0.080
(0.079)   (0.080)      (0.077)  (0.078)  (0.080)  (0.081)  (0.080)

X1.11      -0.087  0.712***  0.205***  0.010  0.245***
0.054     0.085
(0.076)   (0.078)      (0.075)  (0.075)  (0.077)  (0.078)  (0.077)

X2.11      0.155***  0.153***  0.164***  0.150***  0.142***
0.193***  0.135***
(0.026)   (0.026)      (0.025)  (0.026)  (0.026)  (0.026)  (0.026)

X3.11      0.172***  0.177***  0.189***  0.219***  0.146***
0.175***  0.120***
(0.024)   (0.025)      (0.024)  (0.024)  (0.025)  (0.025)  (0.025)

X4.11      0.119***  0.115***  0.148***  0.187***  0.196***
0.131***  0.122***
(0.026)   (0.026)      (0.025)  (0.025)  (0.026)  (0.026)  (0.026)

X5.11      0.138***  0.137***  0.130***  0.116***  0.134***
0.151***  0.173***
(0.026)   (0.026)      (0.025)  (0.025)  (0.026)  (0.026)  (0.026)

X6.11      0.184***  0.182***  0.144***  0.118***  0.142***
0.168***  0.244***
(0.025)   (0.026)      (0.025)  (0.025)  (0.025)  (0.026)  (0.026)

```

Observations	1,459	1,459	1,459	1,459	1,459
1,459	1,459				
R2	0.715	0.727	0.709	0.718	0.699
0.716	0.714				
Adjusted R2	0.714	0.726	0.707	0.717	0.698
0.715	0.713				
Residual Std. Error (df = 1452)	0.301	0.303	0.310	0.315	0.311
0.307	0.311				
F Statistic (df = 7; 1452)	520.220***	552.490***	504.350***	527.980***	482.740***
524.130***	518.390***				

Note:

\*\*p<0.05; \*\*\*p<0.01

\*p<0.1;



## Appendix IV: R-Codes

```

#Test for the Causality using Granger Causality test.
library(vars)
library(mFilter)
library(tseries)
library(TSstudio)
library(forecast)
library(tidyverse)
library(lmtest)
library(BVAR)
#NET WORK MODELS
#zone 1
#data=read.csv(file.choose(),header=TRUE)# Gitonga PhD Thesis
data1=read.csv("C:\\Users\\User\\Desktop\\phd\\z1N.csv",header=TRUE)
data1
#attach(data1)
#str(data1)
#names(data1)
#Normalization of data
data1$x0<-(data1$x0-min(data1$x0))/(max(data1$x0)-min(data1$x0))
data1$x1<-(data1$x1-min(data1$x1))/(max(data1$x1)-min(data1$x1))
data1$x2<-(data1$x2-min(data1$x2))/(max(data1$x2)-min(data1$x2))
data1$x6<-(data1$x6-min(data1$x6))/(max(data1$x6)-min(data1$x6))
data1$x3 <- (data1$x3 -min(data1$x3))/(max(data1$x3)-min(data1$x3))
data1$x5<-(data1$x5-min(data1$x5))/(max(data1$x5)-min(data1$x5))
data1$x4<-(data1$x4-min(data1$x4))/(max(data1$x4)-min(data1$x4))
#data1$timestamp<-(data1$timestamp-min(data1$timestamp))/(max(data1$timestamp)-
min(data1$timestamp))
x0 <-data1$x0
x1 <-data1$x1
x2 <-data1$x2
x3 <-data1$x3
x4 <-data1$x4
x5 <-data1$x5
x6 <-data1$x6
data1<-data.frame(x0, x1, x2, x3, x4, x5, x6)
data1
x0 <- diff(x0, differences = 1)
x1 <- diff(x1, differences = 1)
x2 <- diff(x2, differences = 1)
x3 <- diff(x3, differences = 1)
x4 <- diff(x4, differences = 1)
x5 <- diff(x5, differences = 1)
x6 <- diff(x6, differences = 1)
# Access a subset of the fred_qd dataset
dat1 <- data1[, c("x0", "x1", "x2","x3","x4","x5","x6")]
dat1
# Transform it to be stationary
dat<- fred_transform(data1, codes = c(2,2, 2, 2, 2, 2, 1), lag = 2)
dat
# Estimate a BVAR using one lag, default settings and very few draws
BV <-bvar(data1,lags=4,n_draw = 1000,n_burn = 500,n_thin = 5,
fcst = NULL,irf = NULL,verbose = FALSE)
BV
# Calculate and store forecasts and impulse responses
predict(BV) <- predict(BV, horizon = 8)
predict(BV)
irf(BV) <- irf(BV, horizon = 8, fevd = FALSE)
irf(BV)
bv_minnesota(lambda = bv_lambda(),alpha = bv_alpha(),psi = bv_psi(),var =
10000000,b = 1)
bv_mn(lambda = bv_lambda(),alpha = bv_alpha(),psi = bv_psi(),var = 10000000,b
= 1)
bv_lambda(mode = 0.2, sd = 0.4, min = 0.0001, max = 5)

```

```

bv_alpha(mode = 2, sd = 0.25, min = 1, max = 3)
bv_psi(scale = 0.004, shape = 0.004, mode = "auto", min = "auto", max = "auto")
bv_minnesota
# Adjust alpha and the Minnesota prior variance.
bv_mn(alpha = bv_alpha(mode = 0.5, sd = 1, min = 1e-12, max = 10), var = 1e6)
# Optionally use a vector as shorthand
bv_mn(alpha = c(0.5, 1, 1e-12, 10), var = 1e6)
# Only adjust lambda's standard deviation
bv_mn(lambda = bv_lambda(sd = 2))
# Provide prior modes for psi (for a VAR with three variables)
bv_mn(psi = bv_psi(mode = c(0.7, 0.3, 0.9)))
## Not run:
# Check convergence of the hyperparameters with a trace and density plot
plot(BV)
# Plot forecasts and impulse responses
plot(predict(BV))
plot(irf(BV))
# Check coefficient values and variance-covariance matrix
summary(BV)
#Differencing to attain stationarity
data1<-data.frame(x0, x1, x2, x3, x4, x5, x6)
data1
x0 <- diff(x0, differences = 6)
x1 <- diff(x1, differences = 1)
x2 <- diff(x2, differences = 6)
x3 <- diff(x3, differences = 6)
x4 <- diff(x4, differences = 6)
x5 <- diff(x5, differences = 6)
x6 <- diff(x6, differences = 6)

x0 <- (x0 -min(x0)) / (max(x0)-min(x0))
x1 <- (x1 -min(x1)) / (max(x1)-min(x1))
x2 <- (x2 -min(x2)) / (max(x2)-min(x2))
x3 <- (x3-min(x3)) / (max(x3)-min(x3))
x4 <- (x4 -min(x4)) / (max(x4)-min(x4))
x5 <- (x5-min(x5)) / (max(x5)-min(x5))
x6 <- (x6 -min(x6)) / (max(x6)-min(x6))

Da1 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
Da1
fit11<-glm(x0~x1+x2+x3+x4+x5+x6, family = gaussian,data = Da1)
fit11
summary(fit11)
a<-coef(fit11)[1]
a
b<-coef(fit11)[2]
b
c<-coef(fit11)[3]
c
d<-coef(fit11)[4]
d
e<-coef(fit11)[5]
e
f<-coef(fit11)[6]
f
g<-coef(fit11)[7]
g
Xhat<-a+b*x1+c*x2++d*x3++e*x4++f*x5++g*x6+
Xhat
deviation<-x0-Xhat
cbind(x0,Xhat,deviation)
SSE<-0
for(i in 1:length(x0)){
SSE= SSE+(x0[i] - Xhat[i])^2
}
print(SSE)

```

```

#Obtaining th Mean Squared error(MSE)
MSE<- SSE/length(x0)
MSE
#ROOT MEAN SQUARED ERROR (RMSE)
RMSE<-sqrt(MSE)
print(RMSE)
summary(fit11)
D1 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
D1
fit1<-glm(x0~x1+x2+x3+x4+x5+x6-1,family = gaussian, data = Data1)
fit1
summary(fit1)
#Obtaining granger causality
grangertest(x0 ~ x1, order = 6)
grangertest(x0 ~ x2, order = 6)
grangertest(x0 ~ x3, order = 6)
grangertest(x0 ~ x4, order = 6)
grangertest(x0 ~ x5, order = 6)
grangertest(x0 ~ x6, order = 6)
Acf1<-acf(Data1)
Acf1
plot(acf1)
v1 <- cbind(x0, x1, x3,, x4)
colnames(v1) <- cbind("x0", "x1", "x3", "x4")
lagselect <- VARselect(v1, lag.max = 15, type = "const")
lagselect$selection
Modell <- VAR(v1, p = 2, type = "const", season = NULL, exog = NULL)
#Modell
summary(Modell)
Serial1 <- serial.test(Modell, lags.pt = 5, type = "PT.asymptotic")
Serial1
Arch1 <- arch.test(Modell, lags.multi = 15, multivariate.only = TRUE)
Arch1
Norm1 <- normality.test(Modell, multivariate.only = TRUE)
Norm1
Stability1 <- stability(Modell, type = "OLS-CUSUM")
plot(Stability1)
#NN ANALYSIS
set.seed(222)
ind<-sample(2,nrow(Data1),replace = TRUE, prob = c(0.7,0.3))
training<-Data1[ind==1,]#first sample
testing<-Data1[ind==2,]#second sample
training
testing
library(neuralnet)
set.seed(333)
library(neuralnet)
names(Data1)
#training=training[-c(1,9)]
training=data.frame(training)
#attach(training)
nn<-neuralnet(x0~x1+x3+x4+x6,data = Data1, hidden = 1, linear.output = FALSE)
plot(nn)
#PREDICTION ANALYSIS
out=compute(nn, training[, -c(1)])#excludes the dependent variable
out#net.result =probabilities
head(out$net.result)
head(data[1,])#first row values#
#confusion matrix
p1=out$net.result
pred1=ifelse(p1>0.5,1,0)
X=training$x
length(X)
length(pred1)
tab1=table(pred1, x0)
tab1
1-sum(diag(tab1))/sum(tab1)

```

```

par(mfrow=c(2,3))
hist(Data1$x0, breaks=4)
hist(Data1$x1, breaks=4)
#hist(Data1$x2, breaks=4)
hist(Data1$x3, breaks=4)
hist(Data1$x4, breaks=4)
#hist(Data1$x5, breaks=4)
hist(Data1$x6, breaks=4)
#CONFUSION MATRIX
library(caret)
status=as.numeric(training$x)
status1=ifelse(status>0.05,0,1)
status1
cm=table(pred1, status1)
colnames(cm)=c("correct", "incorrect")
rownames(cm)=c("correct", "incorrect")
cm
pred=as.factor(pred1)
x0=as.factor(x0)
library(caret)
library(MLmetrics)
MSE(as.numeric(pred), training$x0)
RMSE(as.numeric(pred), training$x0)
MAE(as.numeric(pred), training$x0)
MAPE(as.numeric(pred), training$x0)
aicmodel<-glm(x0~x1+x3+x4+x6, family = binomial, data = data1)
AIC(aicmodel)
BIC(aicmodel)

#code for zone 2
data2=read.csv("C:\\Users\\User\\Desktop\\PHD\\z2N.csv", header = TRUE)
data2
attach(data2)
str(data2)
names(data2)
data2$x0<-(data2$x0-min(data2$x0))/(max(data2$x0)-min(data2$x0))
data2$x1<-(data2$x1-min(data2$x1))/(max(data2$x1)-min(data2$x1))
data2$x2<-(data2$x2-min(data2$x2))/(max(data2$x2)-min(data2$x2))
data2$x3<-(data2$x3-min(data2$x3))/(max(data2$x3)-min(data2$x3))
data2$x4<-(data2$x4-min(data2$x4))/(max(data2$x4)-min(data2$x4))
data2$x5<-(data2$x5-min(data2$x5))/(max(data2$x5)-min(data2$x5))
data2$x6<-(data2$x6-min(data2$x6))/(max(data2$x6)-min(data2$x6))
#data$timestamp<-(data$timestamp-min(data$timestamp))/(max(data$timestamp)-
min(data$timestamp))
x0 <-data2$x0
x1 <-data2$x1
x2 <-data2$x2
x3 <-data2$x3
x4 <-data2$x4
x5 <-data2$x5
x6 <-data2$x6
data1<-data.frame(x0, x1, x2, x3, x4, x5, x6)
data1
# Access a subset of the fred_qd dataset
dat1 <- data1[, c("x0", "x1", "x2", "x3", "x4", "x5", "x6")]
dat1

Data2 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
Data2
fit2<-glm(x0~x1+x2+x3+x4+x5+x6-1, family = gaussian, data = Data2)
fit2
summary(fit2)
#Differencing to attain stationarity
x0 <- diff(x0, differences = 9)
x1 <- diff(x1, differences = 9)

```

```

x2 <- diff(x2, differences = 9)
x3 <- diff(x3, differences = 9)
x4 <- diff(x4, differences = 9)
x5 <- diff(x5, differences = 9)
x6 <- diff(x6, differences = 9)
#Data2 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
#Data2
#Obtaining granger causality
grangertest(x0 ~ x1, order = 6)
grangertest(x0 ~ x2, order = 6)
grangertest(x0 ~ x3, order = 6)
grangertest(x0 ~ x4, order = 6)
grangertest(x0 ~ x5, order = 6)
grangertest(x0 ~ x6, order = 6)

v2 <- cbind(x0, x1, x2, x3, x4, x5, x6)

colnames(v2) <- cbind("x0", "x1", "x2", "x3", "x4", "x5", "x6")
lagselect <- VARselect(v2, lag.max = 15, type = "const")
lagselect$selection
Model2 <- VAR(v2, p = 3, type = "const", season = NULL, exog = NULL)
summary(Model2)
Serial2 <- serial.test(Model2, lags.pt = 5, type = "PT.asymptotic")
Serial2
Arch2 <- arch.test(Model2, lags.multi = 15, multivariate.only = TRUE)
Arch2
Norm2 <- normality.test(Model2, multivariate.only = TRUE)
Norm2
Stability2 <- stability(Model2, type = "OLS-CUSUM")
plot(Stability2)

```

### #Bayesian Analysis

```

library(bvartools)
data1=read.csv("C:\\Users\\User\\Desktop\\PhD
FINAL\\PHD\\z12N.csv",header=TRUE)
data1

x0<-log(data1$x0)
x1<-log(data1$x1)
x2<-log(data1$x2)
x3<-data1$x3
x4<-log(data1$x4)
x5<-log(data1$x5)
x6<-data1$x6

Data1<-data.frame(x0, x1, x2, x3, x4, x5, x6)
Data1
Dat1 = ts(Data1,start=c(2014,1,1),frequency=365.25)
plot.ts(Dat1)

data <- gen_var(Dat1, p = 2, deterministic = "const")
data
y <- t(data$data$Y)
Y
x <- t(data$data$Z)
x
A_freq <- tcrossprod(y, x) %*% solve(tcrossprod(x)) # Calculate estimates
round(A_freq, 3) # Round estimates and print

u_freq <- y - A_freq %*% x
u_sigma_freq <- tcrossprod(u_freq) / (ncol(y) - nrow(x))
round(u_sigma_freq, 2)

# Reset random number generator for reproducibility
set.seed(1234567)

```

```

iter <- 30000 # Number of iterations of the Gibbs sampler
burnin <- 15000 # Number of burn-in draws
store <- iter - burnin
tt <- ncol(y) # Number of observations
k <- nrow(y) # Number of endogenous variables
m <- k * nrow(x) # Number of estimated coefficients

# Set priors
a_mu_prior <- matrix(0, m) # Vector of prior parameter means
a_v_i_prior <- diag(1, m) # Inverse of the prior covariance matrix

u_sigma_df_prior <- 6 # Prior degrees of freedom
u_sigma_scale_prior <- diag(1, k) # Prior covariance matrix
u_sigma_df_post <- tt + u_sigma_df_prior # Posterior degrees of freedom

# Initial values
u_sigma_i <- solve(u_sigma_freq)

# Data containers for posterior draws
draws_a <- matrix(NA, m, store)
draws_sigma <- matrix(NA, k * k, store)

# Start Gibbs sampler
for (draw in 1:iter) {
  # Draw conditional mean parameters
  a <- post_normal(y, x, u_sigma_i, a_mu_prior, a_v_i_prior)

  # Draw variance-covariance matrix
  u <- y - matrix(a, k) %*% x # Obtain residuals
  u_sigma_scale_post <- solve(u_sigma_scale_prior + tcrossprod(u))
  u_sigma_i <- matrix(rWishart(1, u_sigma_df_post, u_sigma_scale_post)[,,
1], k)
  u_sigma <- solve(u_sigma_i) # Invert Sigma_i to obtain Sigma

  # Store draws
  if (draw > burnin) {
    draws_a[, draw - burnin] <- a
    draws_sigma[, draw - burnin] <- u_sigma
  }
}
A <- rowMeans(draws_a) # Obtain means for every row
A <- matrix(A, k) # Transform mean vector into a matrix
A <- round(A, 3) # Round values
dimnames(A) <- list(dimnames(y)[[1]], dimnames(x)[[1]]) # Rename matrix
dimensions

A # Print

Sigma <- rowMeans(draws_sigma) # Obtain means for every row
Sigma <- matrix(Sigma, k) # Transform mean vector into a matrix
Sigma <- round(Sigma, 2) # Round values
dimnames(Sigma) <- list(dimnames(y)[[1]], dimnames(y)[[1]]) # Rename matrix
dimensions

Sigma # Print

bvar_est <- bvar(y = data$data$Y, x = data$data$Z, A = draws_a[1:18,],
C = draws_a[19:21, ], Sigma = draws_sigma)
summary(bvar_est)

bvar_est <- thin_posterior(bvar_est, thin = 15)

bvar_pred <- predict(bvar_est, n.ahead = 10, new_d = rep(1, 10))

plot(bvar_pred)

FEIR <- irf(bvar_est, impulse = "X", response = "cons", n.ahead = 8)

```

```

plot(FEIR, main = "Forecast Error Impulse Response", xlab = "Period", ylab =
"Response")

OIR <- irf(bvar_est, impulse = "X", response = "cons", n.ahead = 8, type =
"oir")

plot(OIR, main = "Orthogonalised Impulse Response", xlab = "Period", ylab =
"Response")

GIR <- irf(bvar_est, impulse = "X", response = "cons", n.ahead = 8, type =
"gir")

plot(GIR, main = "Generalised Impulse Response", xlab = "Period", ylab =
"Response")

bvar_fevd_oir <- fevd(bvar_est, response = "cons")

plot(bvar_fevd_oir, main = "OIR-based FEVD of consumption")

bvar_fevd_gir <- fevd(bvar_est, response = "cons", type = "gir")

plot(bvar_fevd_gir, main = "GIR-based FEVD of consumption")

#CONFUSION MATRIX
library(caret)
status=as.numeric(training$x)
status1=ifelse(status>0.05,0,1)
status1
cm=table(pred1,status1)
colnames(cm)=c("correct", "incorrect")
rownames(cm)=c("correct", "incorrect")
cm
pred=as.factor(pred1)
X=as.factor(x)
library(caret)
library(MLmetrics)
MSE(as.numeric(pred),training$x)
RMSE(as.numeric(pred),training$x)
MAE(as.numeric(pred),training$x)
MAPE(as.numeric(pred),training$x)
aicmodel<-glm(x0~x1+x3+x4+x6,,data = Data2,family = binomial(link = "identity"))
AIC(aicmodel)
BIC(aicmodel)

#code for zone 3
#zone 3
#data3=read.csv(file.choose(),header=TRUE)
# Gitonga PhD Thesis
data3=read.csv("C:\\Users\\User\\Desktop\\PHD\\z3N.csv",header=TRUE)
attach(data3)
str(data3)
names(data3)
data3$x0<-(data3$x0-min(data3$x0))/(max(data3$x0)-min(data3$x0))
data3$x1<-(data3$x1-min(data3$x1))/(max(data3$x1)-min(data3$x1))
data3$x2<-(data3$x2-min(data3$x2))/(max(data3$x2)-min(data3$x2))
data3$x3<-(data3$x3-min(data3$x3))/(max(data3$x3)-min(data3$x3))
data3$x4<-(data3$x4-min(data3$x4))/(max(data3$x4)-min(data3$x4))
data3$x5<-(data3$x5-min(data3$x5))/(max(data3$x5)-min(data3$x5))
data3$x6<-(data3$x6-min(data3$x6))/(max(data3$x6)-min(data3$x6))

x0 <-data3$x0
x1 <-data3$x1
x2 <-data3$x2
x3 <-data3$x3
x4 <-data3$x4

```

```

x5 <-data3$x5
x6 <-data3$x6
data1<-data.frame(x0, x1, x2, x3, x4, x5, x6)
data1
#Data3 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
#Data3
#Differencing to attain stationarity
x0 <- diff(x0, differences = 12)
x1 <- diff(x1, differences = 12)
x2 <- diff(x2, differences = 12)
x3 <- diff(x3, differences = 12)
x4 <- diff(x4, differences = 12)
x5 <- diff(x5, differences = 12)
x6 <- diff(x6, differences = 12)

Data3 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
Data3
fit3<-glm(x0~x1+x2+x3+x4+x5+x6-1,family = gaussian,data = Data3)
fit3
summary(fit3)
#Obtaining granger causality
grangertest(x0 ~ x1, order = 9)
grangertest(x0 ~ x2, order = 9)
grangertest(x0 ~ x3, order = 9)
grangertest(x0 ~ x4, order = 9)
grangertest(x0 ~ x5, order = 9)
grangertest(x0 ~ x6, order = 9)

v3 <- cbind(x0, x1, x3, x4, x6)
colnames(v3) <- cbind("x0", "x1", "x2", "x3", "x4", "x5", "x6" )
lagselect <- VARselect(v3, lag.max = 15, type = "const")
lagselect$selection
Model3 <- VAR(v3, p = 3, type = "const", season = NULL, exog = NULL)
Model3
summary(Model3)
Serial3 <- serial.test(Model3, lags.pt = 5, type = "PT.asymptotic")
Serial3
Arch3 <- arch.test(Model3, lags.multi = 15, multivariate.only = TRUE)
Arch3
Norm3 <- normality.test(Model3, multivariate.only = TRUE)
Norm3
Stability3 <- stability(Model3, type = "OLS-CUSUM")
plot(Stability3)
#code for zone 4
#zone 4
#data4=read.csv(file.choose(),header=TRUE)# Gitonga PhD Thesis
data4=read.csv("C:\\Users\\User\\Desktop\\PHD\\z4.csv",header=TRUE)
attach(data4)
str(data4)
names(data4)
data4$x0<-(data4$x0-min(data4$x0))/(max(data4$x0)-min(data4$x0))
data4$x1<-(data4$x1-min(data4$x1))/(max(data4$x1)-min(data4$x1))
data4$x2<-(data4$x2-min(data4$x2))/(max(data4$x2)-min(data4$x2))
data4$x6<-(data4$x6-min(data4$x6))/(max(data4$x6)-min(data4$x6))
data4$x3 <-data4$x3 -min(data4$x3)/(max(data4$x3)-min(data4$x3))
data4$x5<-(data4$x5-min(data4$x5))/(max(data4$x5)-min(data4$x5))
data4$x4<-(data4$x4-min(data4$x4))/(max(data4$x4)-min(data4$x4))
#data$timestamp<-(data$timestamp-min(data$timestamp))/(max(data$timestamp)-
min(data$timestamp))
x0 <-data4$x0
x1 <-data4$x1
x2 <-data4$x2
x3 <-data4$x3
x4 <-data4$x4

```



```

x5 <-data4$x5
x6 <-data4$x6
data1<-data.frame(x0, x1, x2, x3, x4, x5, x6)
data1
#Data4 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
#Data4
#fit4<-glm(X~X1+X2+X3+X4+X5+X6,family=gaussian,data=Data4)
#fit4
#summary(fit4)
#Differencing to attain stationarity
Data4 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
Data4
fit4<-glm(x0~x1+x2+x3+x4+x5+x6,family = gaussian, data = Data4)
fit4
summary(fit3)
#Obtaining granger causality
grangertest(x0 ~ x1, order = 7)
grangertest(x0 ~ x2, order = 7)
grangertest(x0 ~ x3, order = 7)
grangertest(x0 ~ x4, order = 7)
grangertest(x0 ~ x5, order = 7)
grangertest(x0 ~ x6, order = 7)

v4 <- cbind(x0, x1, x2, x3, x4, x5, x6)
colnames(v4) <- cbind("x0", "x1", "x2", "x3", "x4", "x5", "x6")
lagselect <- VARselect(v4, lag.max = 15, type = "const")
lagselect$selection
Model4 <- VAR(v4, p = 3, type = "const", season = NULL, exog = NULL)
Model4
summary(Model4)
Serial4 <- serial.test(Model4, lags.pt = 5, type = "PT.asymptotic")
Serial4
Arch4 <- arch.test(Model4, lags.multi = 15, multivariate.only = TRUE)
Arch4
Norm4 <- normality.test(Model4, multivariate.only = TRUE)
Norm4
Stability4 <- stability(Model4, type = "OLS-CUSUM")
plot(Stability4)
#code for zone 5
#zone 5
#data5=read.csv(file.choose(),header=TRUE)# Gitonga PhD Thesis
data5=read.csv("C:\\Users\\User\\Desktop\\PHD\\z5.csv",header=TRUE)
attach(data5)
str(data5)
names(data5)
data5
Data5$x0<-(data5$x0-min(data5$x0))/(max(data5$x0)-min(data5$x0))
data5$x1<-(data5$x1-min(data5$x1))/(max(data5$x1)-min(data5$x1))
data5$x2<-(data5$x2-min(data5$x2))/(max(data5$x2)-min(data5$x2))
data5$x6<-(data5$x6-min(data5$x6))/(max(data5$x6)-min(data5$x6))
data5$x3 <-(data5$x3 -min(data5$x3))/(max(data5$x3)-min(data5$x3))
data5$x5<-(data5$x5-min(data5$x5))/(max(data5$x5)-min(data5$x5))
data5$x4<-(data5$x4-min(data5$x4))/(max(data5$x4)-min(data5$x4))
#data$timestamp<-(data$timestamp-min(data$timestamp))/(max(data$timestamp)-
min(data$timestamp))
x0 <-data5$x0
x1 <-data5$x1
x2 <-data5$x2
x3 <-data5$x3
x4 <-data5$x4
x5 <-data5$x5
x6 <-data5$x6
data1<-data.frame(x0, x1, x2, x3, x4, x5, x6)
data1

```

```

#Differencing to attain stationarity
x0 <- diff(x0, differences = 9)
x1 <- diff(x1, differences = 9)
x2 <- diff(x2, differences = 9)
x3 <- diff(x3, differences = 9)
x4 <- diff(x4, differences = 9)
x5 <- diff(x5, differences = 9)
x6 <- diff(x6, differences = 9)

Data5 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
Data5
fit5<-glm(x0~x1+x2+x3+x4+x5+x6-1,family=gaussian,data=Data5)
fit5
summary(fit5)
#Obtaining granger causality
grangertest(x0 ~ x1, order = 5)
grangertest(x0 ~ x2, order = 5)
grangertest(x0 ~ x3, order = 5)
grangertest(x0 ~ x4, order = 5)
grangertest(x0 ~ x5, order = 5)
grangertest(x0 ~ x6, order = 5)

v5 <- cbind(x0, x1, x2, x3, x4, x5, x6)

colnames(v5) <- cbind("x0", "x1", "x2", "x3", "x4", "x5", "x6")
lagselect <- VARselect(v5, lag.max = 15, type = "const")
lagselect$selection
Model5 <- VAR(v5, p = 3, type = "const", season = NULL, exog = NULL)
Model5
summary(Model5)
Serial5 <- serial.test(Model5, lags.pt = 5, type = "PT.asymptotic")
Serial5
Arch5 <- arch.test(Model5, lags.multi = 15, multivariate.only = TRUE)
Arch5
Norm5 <- normality.test(Model5, multivariate.only = TRUE)
Norm5
Stability5 <- stability(Model5, type = "OLS-CUSUM")
plot(Stability5)
#CODE FOR GLOBAL VECTOR
#dataG=read.csv(file.choose(),header=TRUE)# Gitonga PhD Thesis
dataG=read.csv("C:\\Users\\User\\Desktop\\PHD\\GZ1.csv",header=TRUE)
attach(dataG)
str(dataG)
names(dataG)
dataG$x<-(dataG$x-min(dataG$x))/(max(dataG$x)-min(dataG$x))
dataG$x
dataG$x0<-( dataG$x0-min(dataG$x0))/(max(dataG$x0)-min(dataG$x0))
dataG$x1<-( dataG$x1-min(dataG$x1))/(max(dataG$x1)-min(dataG$x1))
dataG$x2<-( dataG$x2-min(dataG$x2))/(max(dataG$x2)-min(dataG$x2))
dataG$x6<-( dataG$x6-min(dataG$x6))/(max(dataG$x6)-min(dataG$x6))
dataG$x3<-( dataG$x3-min(dataG$x3))/(max(dataG$x3)-min(dataG$x3))
dataG$x5<-( dataG$x5-min(dataG$x5))/(max(dataG$x5)-min(dataG$x5))
dataG$x4<-( dataG$x4-min(dataG$x4))/(max(dataG$x4)-min(dataG$x4))
#data$timestamp<-(data$timestamp-min(data$timestamp))/(max(data$timestamp)-
min(data$timestamp))

x0 <-dataG$x0
x0 <- dataG$x0
x1 <- dataG$x1
x2 <- dataG$x2
x3 <- dataG$x3
x4 <- dataG$x4
x5 <- dataG$x5
x6 <- dataG$x6

```

```

DataG <- data.frame(x0, x1, x2, x3, x4, x5, x6)
DataG
fitG<-glm(x0~x1+x2+x3+x4+x5+x6,family = gaussian,data = DataG)
fitG
summary(fitG)
#Differencing to attain stationarity
x0 <- diff(x0, differences = 12)
x1 <- diff(x1, differences = 12)
x2 <- diff(x2, differences = 12)
x3 <- diff(x3, differences = 12)
x4 <- diff(x4, differences = 12)
x5 <- diff(x5, differences = 12)
x6 <- diff(x6, differences = 12)

DataG1 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
DataG1
fitG1<-glm(x0~x1+x2+x3+x4+x5+x6-1,family=gaussian,data=DataG1)
fitG1
summary(fitG1)
#Obtaining granger causality
grangertest(x0 ~ x1, order = 8)
grangertest(x0 ~ x2, order = 8)
grangertest(x0 ~ x3, order = 8)
grangertest(x0 ~ x4, order = 8)
grangertest(x0 ~ x5, order = 8)
grangertest(x0 ~ x6, order = 8)

set.seed(222)
ind<-sample(2,nrow(DataG1),replace=TRUE,prob=c(0.7,0.3))
training<-DataG1[ind==1,]#first sample
testing<-DataG1[ind==2,]#second sample
training
testing
library(neuralnet)
set.seed(333)
library(neuralnet)
names(DataG1)
training=training[-c(1,9)]
training=data.frame(training)
attach(training)
nn=neuralnet(x0~x1+x2+x3+x4+x5+x6,data= DataG1, hidden=1,act.fct = "logistic",
             linear.output = FALSE)
# plot neural network
plot(nn)
vG <- cbind(x0, x1, x3, x4, x6)
colnames(vG) <- cbind("x0", "x1", "x2","x3","x4","x5","x6")
lagselect <- VARselect(vG, lag.max = 15, type = "const")
lagselect$selection
ModelG <- VAR(vG, p = 3, type = "const", season = NULL, exog = NULL)
summary(ModelG)
SerialG <- serial.test(ModelG, lags.pt = 5, type = "PT.asymptotic")
SerialG
ArchG <- arch.test(ModelG, lags.multi = 15, multivariate.only = TRUE)
ArchG
NormG <- normality.test(ModelG, multivariate.only = TRUE)
NormG
StabilityG <- stability(ModelG, type = "OLS-CUSUM")
plot(StabilityG)
predict(ModelG,n.ahead=3,newxreg=NULL,se.fit=TRUE)
plot(forecast(ModelG))
forecasts <- HoltWinters(x, beta=FALSE, gamma=FALSE)
HoltWinters(x, beta=FALSE, gamma=FALSE, l.start=23.56)
fit <- HoltWinters(x,gamma=FALSE)
plot(forecast(fit))
#Test for the Causality using Granger Causality test.
library(vars)

```

```

library(mFilter)
library(tseries)
library(TSstudio)
library(forecast)
library(tidyverse)
library(lmtest)
library(BVAR)
#NET WORK MODELS
#zone 1
#data=read.csv(file.choose(),header=TRUE)# Gitonga PhD Thesis
data1=read.csv("C:\\Users\\User\\Desktop\\phd\\z1N.csv",header=TRUE)
data1
attach(data1)
str(data1)
names(data1)
#Normalization of data
data1$x0<-(data1$x0-min(data1$x0))/(max(data1$x0)-min(data1$x0))
data1$x1<-(data1$x1-min(data1$x1))/(max(data1$x1)-min(data1$x1))
data1$x2<-(data1$x2-min(data1$x2))/(max(data1$x2)-min(data1$x2))
data1$x6<-(data1$x6-min(data1$x6))/(max(data1$x6)-min(data1$x6))
data1$x3 <-(data1$x3 -min(data1$x3))/(max(data1$x3)-min(data1$x3))
data1$x5<-(data1$x5-min(data1$x5))/(max(data1$x5)-min(data1$x5))
data1$x4<-(data1$x4-min(data1$x4))/(max(data1$x4)-min(data1$x4))
#data1$timestamp<-(data1$timestamp-min(data1$timestamp))/(max(data1$timestamp)-
min(data1$timestamp))
x0 <-data1$x0
x1 <-data1$x1
x2 <-data1$x2
x3 <-data1$x3
x4 <-data1$x4
x5 <-data1$x5
x6 <-data1$x6
Data1 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
Data1
#Differencing to attain stationarity for X1
x0 <- diff(x0, differences = 9)
x1 <- diff(x1, differences = 9)
x2 <- diff(x2, differences = 9)
x3 <- diff(x3, differences = 9)
x4 <- diff(x4, differences = 9)
x5 <- diff(x5, differences = 9)
x6 <- diff(x6, differences = 9)

#Obtaining granger causality
grangertest(x0 ~ x1, order = 6)
grangertest(x0 ~ x2, order = 6)
grangertest(x0 ~ x3, order = 6)
grangertest(x0 ~ x4, order = 6)
grangertest(x0 ~ x5, order = 6)
grangertest(x0 ~ x6, order = 6)

v1 <- cbind(x, x1, x3, x4,x6)
colnames(v1) <- cbind("x0", "x1", "x2", "x3", "x4", "x5", "x6")
lagselect <- VARselect(v1, lag.max = 15, type = "const")
lagselect$selection
Modell <- VAR(v1, p = 3, type = "const", season = NULL, exog = NULL)
summary(Modell)
Serial1 <- serial.test(Modell, lags.pt = 5, type = "PT.asymptotic")
Serial1
Arch1 <- arch.test(Modell, lags.multi = 15, multivariate.only = TRUE)
Arch1
Norm1 <- normality.test(Modell, multivariate.only = TRUE)
Norm1
Stability1 <- stability(Modell, type = "OLS-CUSUM")
plot(Stability1)

```

```

#code for zone 2
data2=read.csv("C:\\Users\\User\\Desktop\\PHD\\z2N.csv",header=TRUE)
data2
attach(data2)
str(data2)
names(data2)
data2$x0<-(data2$x0-min(data2$x0))/(max(data2$x0)-min(data2$x0))
data2$x1<-(data2$x1-min(data2$x1))/(max(data2$x1)-min(data2$x1))
data2$x2<-(data2$x2-min(data2$x2))/(max(data2$x2)-min(data2$x2))
data2$x6<-(data2$x6-min(data2$x6))/(max(data2$x6)-min(data2$x6))
data2$x3<-(data2$x3-min(data2$x3))/(max(data2$x3)-min(data2$x3))
data2$x5<-(data2$x5-min(data2$x5))/(max(data2$x5)-min(data2$x5))
data2$x4<-(data2$x4-min(data2$x4))/(max(data2$x4)-min(data2$x4))
#data$timestamp<-(data$timestamp-min(data$timestamp))/(max(data$timestamp)-
min(data$timestamp))
x0 <-data2$x0
x1 <-data2$x1
x2 <-data2$x2
x3 <-data2$x3
x4 <-data2$x4
x5 <-data2$x5
x6 <-data2$x6

Data2 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
Data2
#Differencing to attain stationarity
x0 <- diff(x0, differences = 9)
x1 <- diff(x1, differences = 9)
x2 <- diff(x2, differences = 9)
x3 <- diff(x3, differences = 9)
x4 <- diff(x4, differences = 9)
x5 <- diff(x5, differences = 9)
x6 <- diff(x6, differences = 9)

#Obtaining granger causality
grangertest(x0 ~ x1, order = 6)
grangertest(x0 ~ x2, order = 6)
grangertest(x0 ~ x3, order = 6)
grangertest(x0 ~ x4, order = 6)
grangertest(x0 ~ x5, order = 6)
grangertest(x0 ~ x6, order = 6)

v2 <- cbind(x0, x1, x2, x3, x4, x5, x6)
colnames(v2) <- cbind("x0", "x1", "x2", "x3", "x4", "x5", "x6")
lagselect <- VARselect(v2, lag.max = 15, type = "const")
lagselect$selection
Model2 <- VAR(v2, p = 3, type = "const", season = NULL, exog = NULL)
summary(Model2)
Serial2 <- serial.test(Model2, lags.pt = 5, type = "PT.asymptotic")
Serial2
Arch2 <- arch.test(Model2, lags.multi = 15, multivariate.only = TRUE)
Arch2
Norm2 <- normality.test(Model2, multivariate.only = TRUE)
Norm2
Stability2 <- stability(Model2, type = "OLS-CUSUM")
plot(Stability2)
#code for zone 3
#zone 3
#data3=read.csv(file.choose(),header=TRUE)# Gitonga PhD Thesis
data3=read.csv("C:\\Users\\User\\Desktop\\PHD\\z3N.csv",header=TRUE)
attach(data3)
str(data3)
names(data3)
data3$x0<-(data3$x0-min(data3$x0))/(max(data3$x0)-min(data3$x0))
data3$x1<-(data3$x1-min(data3$x1))/(max(data3$x1)-min(data3$x1))

```

```

data3$x2<-(data3$x2-min(data3$x2))/(max(data3$x2)-min(data3$x2))
data3$x6<-(data3$x6-min(data3$x6))/(max(data3$x6)-min(data3$x6))
data3$x3<-(data3$x3-min(data3$x3))/(max(data3$x3)-min(data3$x3))
data3$x5<-(data3$x5-min(data3$x5))/(max(data3$x5)-min(data3$x5))
data3$x4<-(data3$x4-min(data3$x4))/(max(data3$x4)-min(data3$x4))
#data$timestamp<-(data$timestamp-min(data$timestamp))/(max(data$timestamp)-
min(data$timestamp))
x0 <-data3$x0
x1 <-data3$x1
x2 <-data3$x2
x3 <-data3$x3
x4 <-data3$x4
x5 <-data3$x5
x6 <-data3$x6
Data3 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
Data3
#Differencing to attain stationarity
x0 <- diff(x0, differences = 12)
x1 <- diff(x1, differences = 12)
x2 <- diff(x2, differences = 12)
x3 <- diff(x3, differences = 12)
x4 <- diff(x4, differences = 12)
x5 <- diff(x5, differences = 12)
x6 <- diff(x6, differences = 12)

#Obtaining granger causality
grangertest(x0 ~ x1, order = 9)
grangertest(x0 ~ x2, order = 9)
grangertest(x0 ~ x3, order = 9)
grangertest(x0 ~ x4, order = 9)
grangertest(x0 ~ x5, order = 9)
grangertest(x0 ~ x6, order = 9)

v3 <- cbind(x0, x1, x2, x3, x4, x5, x6)
colnames(v3) <- cbind("x0", "x1", "x2", "x3", "x4", "x5", "x6")
lagselect <- VARselect(v3, lag.max = 15, type = "const")
lagselect$selection
Model3 <- VAR(v3, p = 3, type = "const", season = NULL, exog = NULL)
summary(Model3)
Serial3 <- serial.test(Model3, lags.pt = 5, type = "PT.asymptotic")
Serial3
Arch3 <- arch.test(Model3, lags.multi = 15, multivariate.only = TRUE)
Arch3
Norm3 <- normality.test(Model3, multivariate.only = TRUE)
Norm3
Stability3 <- stability(Model3, type = "OLS-CUSUM")
plot(Stability3)
#code for zone 4
#zone 4
#data4=read.csv(file.choose(),header=TRUE)# Gitonga PhD Thesis
data4=read.csv("C:\\Users\\User\\Desktop\\PHD\\z4.csv",header=TRUE)
attach(data4)
str(data4)
names(data4)
data4$x0<-(data4$x0-min(data4$x0))/(max(data4$x0)-min(data4$x0))
data4$x1<-(data4$x1-min(data4$x1))/(max(data4$x1)-min(data4$x1))
data4$x2<-(data4$x2-min(data4$x2))/(max(data4$x2)-min(data4$x2))
data4$x6<-(data4$x6-min(data4$x6))/(max(data4$x6)-min(data4$x6))
data4$x3 <-data4$x3 -min(data4$x3)/(max(data4$x3)-min(data4$x3))
data4$x5<-(data4$x5-min(data4$x5))/(max(data4$x5)-min(data4$x5))
data4$x4<-(data4$x4-min(data4$x4))/(max(data4$x4)-min(data4$x4))
#data$timestamp<-(data$timestamp-min(data$timestamp))/(max(data$timestamp)-
min(data$timestamp))
x0 <-data4$x0

```

```

x1 <-data4$x1
x2 <-data4$x2
x3 <-data4$x3
x4 <-data4$x4
x5 <-data4$x5
x6 <-data4$x6
Data4 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
Data4
#Differencing to attain stationarity
x0 <- diff(x0, differences = 9)
x1 <- diff(x1, differences = 9)
x2 <- diff(x2, differences = 9)
x3 <- diff(x3, differences = 9)
x4 <- diff(x4, differences = 9)
x5 <- diff(x5, differences = 9)
x6 <- diff(x6, differences = 9)

#Obtaining granger causality
grangertest(x0 ~ x1, order = 7)
grangertest(x0 ~ x2, order = 7)
grangertest(x0 ~ x3, order = 7)
grangertest(x0 ~ x4, order = 7)
grangertest(x0 ~ x5, order = 7)
grangertest(x0 ~ x6, order = 7)

v4 <- cbind(x0, x1, x2, x3, x4, x5, x6)
colnames(v4) <- cbind("x0", "x1", "x2", "x3", "x4", "x5", "x6")
lagselect <- VARselect(v4, lag.max = 15, type = "const")
lagselect$selection
Model4 <- VAR(v4, p = 3, type = "const", season = NULL, exog = NULL)
summary(Model4)
Serial4 <- serial.test(Model4, lags.pt = 5, type = "PT.asymptotic")
Serial4
Arch4 <- arch.test(Model4, lags.multi = 15, multivariate.only = TRUE)
Arch4
Norm4 <- normality.test(Model4, multivariate.only = TRUE)
Norm4
Stability4 <- stability(Model4, type = "OLS-CUSUM")
plot(Stability4)
#code for zone 5
#zone 5
#data5=read.csv(file.choose(),header=TRUE)# Gitonga PhD Thesis
data5=read.csv("C:\\Users\\User\\Desktop\\PHD\\z5.csv",header=TRUE)
attach(data5)
str(data5)
names(data5)
#data5
Data5$x0<-(data5$x0-min(data5$x0))/(max(data5$x0)-min(data5$x0))
data5$x1<-(data5$x1-min(data5$x1))/(max(data5$x1)-min(data5$x1))
data5$x2<-(data5$x2-min(data5$x2))/(max(data5$x2)-min(data5$x2))
data5$x6<-(data5$x6-min(data5$x6))/(max(data5$x6)-min(data5$x6))
data5$x3 <-(data5$x3 -min(data5$x3))/(max(data5$x3)-min(data5$x3))
data5$x5<-(data5$x5-min(data5$x5))/(max(data5$x5)-min(data5$x5))
data5$x4<-(data5$x4-min(data5$x4))/(max(data5$x4)-min(data5$x4))
#data$timestamp<-(data$timestamp-min(data$timestamp))/(max(data$timestamp)-
min(data$timestamp))
x0 <-data5$x0
x1 <-data5$x1
x2 <-data5$x2
x3 <-data5$x3
x4 <-data5$x4
x5 <-data5$x5
x6 <-data5$x6
Data5 <- data.frame(x0, x1, x2, x3, x4, x5, x6)

```

```

Data5
#Differencing to attain stationarity
x0 <- diff(x0, differences = 9)
x1 <- diff(x1, differences = 9)
x2 <- diff(x2, differences = 9)
x3 <- diff(x3, differences = 9)
x4 <- diff(x4, differences = 9)
x5 <- diff(x5, differences = 9)
x6 <- diff(x6, differences = 9)

#Obtaining granger causality
grangertest(x0 ~ x1, order = 3)
grangertest(x0 ~ x2, order = 3)
grangertest(x0 ~ x3, order = 3)
grangertest(x0 ~ x4, order = 3)
grangertest(x0 ~ x5, order = 3)
grangertest(x0 ~ x6, order = 3)

v5 <- cbind(x0, x1, x3, x4, x6)
colnames(v5) <- cbind("x0", "x1", "x2", "x3", "x4", "x5", "x6")
lagselect <- VARselect(v5, lag.max = 15, type = "const")
lagselect$selection
Model5 <- VAR(v5, p = 3, type = "const", season = NULL, exog = NULL)
Model5
summary(Model5)
Serial5 <- serial.test(Model5, lags.pt = 5, type = "PT.asymptotic")
Serial5
Arch5 <- arch.test(Model5, lags.multi = 15, multivariate.only = TRUE)
Arch5
Norm5 <- normality.test(Model5, multivariate.only = TRUE)
Norm5
Stability5 <- stability(Model5, type = "OLS-CUSUM")
plot(Stability5)
#CODE FOR GLOBAL VECTOR
#dataG=read.csv(file.choose(),header=TRUE)# Gitonga PhD Thesis
dataG=read.csv("C:\\Users\\User\\Desktop\\PHD\\GZ1.csv",header=TRUE)
attach(dataG)
str(dataG)
names(dataG)
dataG$X<-(dataG$X-min(dataG$X))/(max(dataG$X)-min(dataG$X))
dataG$X
dataG$x0<-( dataG$x0-min(dataG$x0))/(max(dataG$x0)-min(dataG$x0))
dataG$x1<-( dataG$x1-min(dataG$x1))/(max(dataG$x1)-min(dataG$x1))
dataG$x2<-( dataG$x2-min(dataG$x2))/(max(dataG$x2)-min(dataG$x2))
dataG$x6<-( dataG$x6-min(dataG$x6))/(max(dataG$x6)-min(dataG$x6))
dataG$x3<-( dataG$x3-min(dataG$x3))/(max(dataG$x3)-min(dataG$x3))
dataG$x5<-( dataG$x5-min(dataG$x5))/(max(dataG$x5)-min(dataG$x5))
dataG$x4<-( dataG$x4-min(dataG$x4))/(max(dataG$x4)-min(dataG$x4))
#data$timestamp<-(data$timestamp-min(data$timestamp))/(max(data$timestamp)-
min(data$timestamp))

x0 <-dataG$x0
x0 <- dataG$x0
x1 <- dataG$x1
x2 <- dataG$x2
x3 <- dataG$x3
x4 <- dataG$x4
x5 <- dataG$x5
x6 <- dataG$x6

DataG <- data.frame(x0, x1, x2, x3, x4, x5, x6))
DataG
#Differencing to attain stationarity
x0 <- diff(x0, differences = 12)
x1 <- diff(x1, differences = 12)

```



```

x2 <- diff(x2, differences = 12)
x3 <- diff(x3, differences = 12)
x4 <- diff(x4, differences = 12)
x5 <- diff(x5, differences = 12)
x6 <- diff(x6, differences = 12)

DataG1 <- data.frame(x0, x1, x2, x3, x4, x5, x6)
DataG1
#Obtaining granger causality
grangertest(x0 ~ x1, order = 8)
grangertest(x0 ~ x2, order = 8)
grangertest(x0 ~ x3, order = 8)
grangertest(x0 ~ x4, order = 8)
grangertest(x0 ~ x5, order = 8)
grangertest(x0 ~ x6, order = 8)

vG <- cbind(x0, x1, x3, x4, x6)
colnames(vG) <- cbind("x0", "x1", "x2", "x3", "x4", "x5", "x6")

lagselect <- VARselect(vG, lag.max = 15, type = "const")
lagselect$selection
ModelG <- VAR(vG, p = 3, type = "const", season = NULL, exog = NULL)
summary(ModelG)
SerialG <- serial.test(ModelG, lags.pt = 5, type = "PT.asymptotic")
SerialG
ArchG <- arch.test(ModelG, lags.multi = 15, multivariate.only = TRUE)
ArchG
NormG <- normality.test(ModelG, multivariate.only = TRUE)
NormG
StabilityG <- stability(ModelG, type = "OLS-CUSUM")
plot(StabilityG)
Theil's U

Ljung-Box test

Q* = 21.054, df = 5, p-value = 0.0007912
Model df: 0. Total lags used: 5
predict (Modell1,n.ahead=10,newxreg=NULL,se.fit=TRUE)
$X
> predict (Modell1,n.ahead=10,newxreg=NULL,se.fit=TRUE)
> plot(forecast (Modell1))
Smoothing parameters: > HoltWinters(X0, beta=FALSE, gamma=FALSE,
l.start=23.56)
From the findings, alpha was 0.3511712, beta was FALSE and the gamma was
FALSE with a Coefficients of 0.2042217
Holt-Winters exponential smoothing without trend and without seasonal
component.
Showed that HoltWinters (x0 = X0, beta = FALSE, gamma = FALSE, l.start = 23.56)
and the Smoothing parameters showed that alpha was 0.9879182, beta was
FALSE and the gamma was FALSE. The Coefficients was showed as 0.188547 > fit
<- HoltWinters(X0,gamma=FALSE) > plot(forecast(fit)
Theil's U showed that Training set was NA while the Test set was 0.9265776
Ljung-Box test
The Residuals from Naive method shwed that Q* = 5.1557, df = 5, p-value =
0.3972, With Model df of 0 leading to a Total lags used of 5
> predict (Modell1,n.ahead=10,newxreg=NULL,se.fit=TRUE)
$X0
> plot(forecast (Modell1))
> forecasts <- HoltWinters(X0, beta=FALSE, gamma=FALSE)
> forecasts
Holt-Winters exponential smoothing without trend and without seasonal
component.
Call:
HoltWinters(x0 = X0, beta = FALSE, gamma = FALSE)
Smoothing parameters:
alpha: 0.02949134

```

```

beta : FALSE
gamma: FALSE
Coefficients: [,1]
a 0.1623501
> HoltWinters(X0, beta=FALSE, gamma=FALSE, l.start=23.56)
Holt-Winters exponential smoothing without trend and without seasonal
component.
Call:
HoltWinters(x0 = X0, beta = FALSE, gamma = FALSE, l.start = 23.56)
Smoothing parameters:
  alpha: 0.999319
  beta : FALSE
  gamma: FALSE
Coefficients:
  [,1]
a 0.04514694
> fit <- HoltWinters(X0, gamma=FALSE)
> plot(forecast(fit))
data: Residuals from Naive method
Q* = 4.3446, df = 5, p-value = 0.5009
Model df: 0. Total lags used: 5
> predict (Modell, n.ahead=10, newxreg=NULL, se.fit=TRUE)
$X0
> plot(forecast (Modell))
> forecasts <- HoltWinters (X0, beta=FALSE, gamma=FALSE)
> forecasts
Holt-Winters exponential smoothing without trend and without seasonal
component.
Call:
HoltWinters(x0 = X0, beta = FALSE, gamma = FALSE)
Smoothing parameters:
  alpha: 0.0242375
  beta : FALSE
  gamma: FALSE
Coefficients:
  [,1]
a 0.2550029
> HoltWinters(X0, beta=FALSE, gamma=FALSE, l.start=23.56)
Holt-Winters exponential smoothing without trend and without seasonal
component.
Call:
HoltWinters(x0 = X0, beta = FALSE, gamma = FALSE, l.start = 23.56)
Smoothing parameters:
  alpha: 0.9967289
  beta : FALSE
  gamma: FALSE
Coefficients:
  [,1]
a 0.1452681
> fit <- HoltWinters(X0, gamma=FALSE)
> plot(forecast(fit))
> checkresiduals(fc)
      Ljung-Box test
data: Residuals from Naive method
Q* = 16.168, df = 5, p-value = 0.006381
Model df: 0. Total lags used: 5
> predict (Modell, n.ahead=10, newxreg=NULL, se.fit=TRUE)
$X
> plot(forecast (Modell))
> forecasts <- HoltWinters (X0, beta=FALSE, gamma=FALSE)
> forecasts
Holt-Winters exponential smoothing without trend and without seasonal
component.
Call:
HoltWinters(x0 = X0, beta = FALSE, gamma = FALSE)
Smoothing parameters:
  alpha: 0.6309507

```

```
beta : FALSE
gamma: FALSE
Coefficients:
  [,1]
a 0.2524445
> HoltWinters(X0, beta=FALSE, gamma=FALSE, l.start=23.56)
Holt-Winters exponential smoothing without trend and without seasonal
component.
Call:
HoltWinters(x0 = X0, beta = FALSE, gamma = FALSE, l.start = 23.56)
Smoothing parameters:
alpha: 0.9970127
beta : FALSE
gamma: FALSE
Coefficients:
  [,1]
a 0.2032076
> fit <- HoltWinters(X0, gamma=FALSE)
> plot(forecast(fit))
> accuracy(fc, my.test)
> checkresiduals(fc)
      Ljung-Box test
data:  Residuals from Naive method
Q* = 3.605, df = 5, p-value = 0.6076
Model df: 0.  Total lags used: 5
> predict(Model1, n.ahead=10, newxreg=NULL, se.fit=TRUE)
```