KWAME NKRUMAH UNIVERSITY OF SCIENCE AND

TECHNOLOGY, KUMASI

MODELLING OF GRAIN YIELD IN MAIZE

By

DWAMENA HARRIET ACHIAA

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS,

KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY IN

PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE DEGREE

OF M.PHIL APPLIED MATHEMATICS

May 12, 2016

Declaration

I hereby declare that this submission is my own work towards the award of the M. Phil degree and that, to the best of my knowledge, it contains no material previously published by another person nor material which had been accepted for the award of any other degree of the university, except where due acknowledgement had been made in the text.

Dwamena, Harriet Achiaa(PG1098913) ....

Student Signature Date

Certi ed by:

<u>Mr. K. F Darkwah</u> ....

Supervisor Signature Date

Certi ed by:

Prof. S.K Amponsah ....

Head of Department Signature Date

Dedication

I dedicate this project work to my loving husband Dennis .S. O ei and my lovely kids Akua O eibea O ei, Kwaku Dwamena O ei and for Nhyira Nyamekye O ei their maximum care and support. Also to my parents Mr. S.Y.Dwamena and Mrs. Felicia Dwamena for the huge investment they have made in my life. Lastly to my siblings Sylvia Dwamena , Ernest Dwamena , Joyce Dwamena and Stephen Dwamena.

Abstract

Grain yield is very important in maize production for breeders at the Crops Research Institute (CRI) of Ghana. However the yields of most varieties that are high yielding released by breeders does not perform so well after some of
their release.

This study was carried out to nd what causes the reduction in yield of these maize varieties of CRI over the years.

An autoregressive moving average model (ARMA) was tted using a 20 year data (1995-2014) from CRI Fumesua. A multiple linear regression model was also tted to study factors a ecting grain yield in maize. Flowering data recorded on a trial eld at Fumesua research station in 2014 was used for the regression model.

The study revealed that ARMA (2, 2) was found to be most suitable model for the di erenced series of maize yield. The multiple regression model showed that the factors plants height, days to owering and eld weight were statistically signi cant at 0.05 level. These factors (plants height, the days to ower and eld weight) are signi cant factors a ecting maize grain yield in Ghana.

Acknowledgements

I thank the almighty God for giving me the strength and wisdom to study this course. I am most grateful to my supervisor Mr. K. F. Darkwa who guided me, read and corrected this work. I would like to thank Dr. Dwommoh Sarpong of the Department of Mathematics, UMaT who helped with supervision. Mr. Martin Tengan CSIR-CRI who also helped me acquired my data I say many thanks for your concern. My heartfelt thanks go to Mr. Ewool of CSIR-CRI who also assisted me with my data. I am most grateful to Dr. Adelaide Agyemang of

CSIR-CRI who assisted me with my data analysis.

Contents

List of Abbreviation

VIF ....Variance In ated Factors

MSE ....Mean Square of Error

MSR .... Mean Square of Regression SSE

.... Sum of Squares due to Error

SSR ....Sum of Squares due to Regression SST

.... Sum of Squares Total

DLM .... Dynamic Linear Model ARIMA

..........................Autoregressive Integrated Moving Average

ARMA ....Autoregressive Moving Average

GARCH ....Generalizes Autoregressive Conditional Heteroskedasticity

TARCH .... Threshold general Autoregressive Conditional Heteroskedasticity EGARCH

.Exponential General Autoregressive Coditional Heteroskedasticity

FIGARCH .... Fractionally Integrated General Autoregressive Coditional Heteroskedasticity

CGARCH Component General Autoregressive Conditional Heteroskedasticity ACF

....Autocorrelation Function

PACF .... Partial Autocorrelation Function AR

....Autoregressive

MA ....Moving Average WAP
....Weeks After Planting

MSV .... Maize Streak Virus

OPVs .... Open Pollinated Varieties KEP

.... Kalahari Early Pearl

DMP .... Days to Mid-Pollen emergency DMS

.... Days to Mid-Silk emergency

CIMMYT ....Centro International de Mojoramieno de Maize ASI

.... Anthesis Silking Interval

CGMS .... Crop Growth Monitoring System

LAI ....Leaf Area Index

GDD ....Growing Degree-Days CSMs

.... Crop Simulation Models

AWC ....Available Water Content

List of Tables

List of Figures

Chapter 1

Introduction

1.1 INTRODUCTION

This chapter presents the background of the study, problem statement, its objectives, the methodology, and signi cance of the study and the organization of the thesis.

1.2 BACKGROUND OF STUDY

Agriculture will be faced with major challenges in the next few decades. The number of undernourished people was estimated at 868 million for the period 2010 - 2012 (FAOSTAT, 2013), indicating that food demand has not yet been satis ed in some parts of the world. The situation may worsen in the near future due to current demographic trends, with the world population likely to reach 9.3 billion by 2050 (World Population Prospects, 2011; Calderini and Slafer, 1998).

Several studies have recently shown that, after a period of strong yield increase, yield levels are currently stagnating in several countries. The average yield of cereal crops increased by more than 98% worldwide, and by more than 187% in France from 1960 to 1990. There were several reasons for this positive trend: genetic improvement of crop cultivars, increase in the use of chemical inputs (fertilizers, insecticides, herbicides and fungicides), mechanization, and irrigation. These improvements led to an approximately linear increase in crop yields in many countries. However, since the 1990s, yield increases for several major cereal crops (wheat, maize, rice, barley or oat) have slowed down. In some countries, yield levels have remained constant or have even declined for some crops (Brisson et al, 2010; Calderini and Slafer, 1998). This is the case for (maize ,wheat) for which several authors have recently shown much slower rates of yield increase than the period

prior to 1990s, with yield stagnation in several countries, including France (Brisson et al, 2010; Lin and Huyber,2012;Ray et al,2012) and Switzerland (Finger R ,2010). These results have raised signi cant concerns in the scienti c community about the ability of agriculture to feed the world in the
future.

Statistical analyses play a key role in current research studies on food security (Ray et al, 2012), where yield time series analysis is used to estimate past yield trends and to predict future yield trends. Various types of statistical models have been used for the analysis of yield time series. Linear regression has been used in many studies (Brisson et al, 2010; Calderini and Slafer, 1998; Hafner, 2003; Kumar, 2000). Other regression models, such as quadratic regression, bilinear, tri-linear, and linear-plus-plateau models, have been used in a smaller number of papers. Several authors have shown that quadratic and linear-plusplateau models tend to perform better in cases of yield stagnation Brisson et al, 2010; Finger R ,2010; Hafner ,2003; Lin and Huybers , 2012; Ray et al,2012).

Statistical methods other than regression models have been used to predict future yield trends (Kumar, 2000) compared the performances of linear and quadratic regression models with those of exponential smoothing (also known as the Holt-Winters method) and moving averages. Exponential smoothing has been shown to perform well in a large range of applications (Kumar, 2013; Brockwell and Davis, 2002), but (Kumar, 2000) showed that the lowest mean square error (MSE) for yield predictions was obtained with the quadratic regression model. However, this result was obtained with a small dataset: yield predictions were assessed for three years at a speci c location in Canada. The Autoregressive Moving Average Model is a developed time series method that can be used to estimate past trends and to predict future trends. It has been
applied in diverse domains, such as econometrics, signal processing, genetics and population dynamics (Petris et al, 2009; Petris ,2010; Prado and West,2010) Their values are, therefore, not xed they vary from year to year and could thus account for

changes in yield trends (e.g., stagnation, increase or decline in the rate of yield increase).

1.3 PROBLEM STATEMENT

Most high yielding maize varieties does not perform after some years of their release. Several studies have recently shown that, after a period of strong yield increase, yield levels are currently stagnating in several countries. The average yield of cereal crops increased by more than 98% worldwide, and by more than 187% in France from 1960 to 1990(FAOSTAT,2013). However, since the 1990s, yield increases for several major cereal crops (wheat, maize, rice, barley or oat) have slowed down. In some countries, yield levels have remained constant or have even declined for some crops (Brisson et al, 2010; Calderini and Slafer, 1998). This is the case for maize, for which several authors have recently shown much slower rates of yield increase than the period prior to 1990s, with yield stagnation in several countries (Brisson et al, 2010; Lin and Huybers , 2012;Ray et al,2012).These results have raised signi cant concerns in the scienti c community about the ability of agriculture to feed the world in the future. Scientists are always coming out with maize cultivars whose yields decline after some years and it has been a challenge as to what causes yield reduction in all these varieties that were giving out high yields sometime ago. I therefore want to take that challenge and undertake a study to access the problem and come out with appropriate
recommendations.

1.4 OBJECTIVES

1. To t an autoregressive moving average (ARMA) model and forecast maize grain yield based on the model
2. Formulate a multiple linear regression model to nd out factors a ecting maize yield.

Research Questions

1. Does the series change signi cantly over time? (Checking for stationarity and variability).

2. Is the empirical distribution of the series and Q-Q plot obtained normal?

3. Does the Auto Correlation Function and the Partial Auto Correlation Function (ACF/PACF) for the returns show dependence in the di erenced series?

H_o : There is a unit root in the series

H_A : There is no unit root in the series

1.5 SIGNIFICANCE OF THE STUDY

The thesis results serves as a basis for future studies into modeling of grain yield in maize using ARMA models and multiple linear regression. The thesis contributes to already literature on factors a ecting grain yield in maize. It will assist Researchers to come out with high yielding varieties that are stable and allow developing countries to predict short falls in grain yields, with bene ts of food security

1.6 METHODOLOGY

The major purpose of this study is modelling grain yield in maize and studying of factors a ecting it. The multiple linear regression approach is considered because it can be used to predict a quantity of interest depending on known values of other quantities .In this case, grain yield will depend on several independent variables such plant height, ear height, days to owering and eld weight.

An autoregressive moving average (ARMA) is a mathematical model which consists of two parts, an autoregressive (AR) part and a moving average (MA) part. The model is usually then referred to as the ARMA (p, q) where p is the order of the

autoregressive part and q is the order of the moving average part. ARMA models can also be used to predict behaviour of a time series from past values alone. Such a prediction can be used as a baseline to evaluate the
possible importance of other variables to the system.

R software will be used for the analyses. Reading materials were sourced from KNUST library and CSIR-CRI library as well as the internet.

1.7 ORGANISATION OF THE THESIS

The thesis work is made up of ve chapters. Chapter one introduces the problem, outlines the objectives of the project, describes the methods used in the work and facilities employed. Chapter two reviews literature on similar topics done by other researchers. Chapter three is basically the methodology. Chapter four is made of the data acquisition and e ectiveness of the technology. Chapter ve looks at the conclusions and recommendations.

Chapter 2

LITERATURE REVIEW

Jackson (2004) worked on spatial distribution of surface soil moisture under a corn eld. Autocorrelation within surface soil moisture (SSM) data may be used to produce high-resolution spatial maps of SSM from point samples. The objective of this study was to characterize the temporal and spatial properties of SSM (0-5 cm) in a Beltsville, MD corn eld using a capacitance probe. The range of spatial autocorrelation was approximately 10 m and the highest sill values were found at water contents (θ) between 20-27%. Nugget values represented a signi cant portion of the total variance (up to 50% for θ > 20% and 73% for θ < 12%). The patterns of SSM under wet conditions exhibited large, continuous polygons while drier conditions resulted in smaller, discreet regions. Early season (< 60 days) Auto-Regressive Moving-Average (ARMA) forecasts of SSM plotted against observed data resulted in R^2 values from

0.15-0.26, while late season (> 80 days) forecasts improved to 0.46-0.65. Forecasts were improved by autoregressive

coe cients and additional SSM datasets.

Lobell et al (2010) came out with the fact that predicting the potential e ects of climate change on crop yields requires a model of how crops respond to weather. Although the general strengths and weaknesses of statistical models are widely understood, there has been little systematic evaluation of their performance relative to other methods. Here we use a perfect model approach to examine the ability of statistical models to predict yield responses to changes in mean temperature and precipitation, as simulated by a process-based crop model. The CERES-Maize model was rst used to simulate historical maize

yield variability at nearly 200 sites in Sub-Saharan Africa, as well as the impacts of hypothetical future scenarios of 2^oC warming and 20% precipitation reduction.

Statistical models of three types (time series, panel, and cross-sectional models) were then trained on the simulated historical variability and used to predict the responses to the future climate changes. The agreement between the processbased and statistical models' predictions was then assessed as a measure of how well statistical models can capture crop responses to warming or precipitation changes. The performance of statistical models di ered by climate variable and spatial scale, with time-series statistical models ably reproducing site-speci c yield response to precipitation change, but performing less well for temperature responses. In contrast, statistical models that relied on information from multiple sites, namely panel and cross-sectional models were better at predicting responses to temperature change than precipitation change. The models based on multiple sites were also much less sensitive to the length of historical period used for training. For all three statistical approaches, the performance improved when individual sites were rst aggregated to country-level averages. Results suggest that statistical models, as compared to CERES-Maize, represent a useful if imperfect tool for projecting future yield responses, with their usefulness higher at broader spatial scales. It is also at these broader scales that climate projections are most available and reliable, and therefore statistical models are likely to continue to play an important role in anticipating future impacts of climate change.

Chen et al (2004) conducted a statistical investigation on yield variability as in uenced by climate. One of the issues with respect to climate change involves its in uence on the distribution of future crop yields. Many studies have been done regarding the e ect on the mean of such distributions but few have addressed the e ect on variance. Furthermore, those that have been done generally report the variance from crop simulators, not from observations. This paper examines the potential e ects of climate change on crop yield variance in the context of current observed yields and then extrapolates to the e ects under projected climate change. In particular, maximum likelihood panel data estimates of the impacts of climate on year-to-year yield variability are constructed for the major U.S. agricultural crops. The panel data technique used embodies a variance estimate developed along the lines of the stochastic production function approach suggested by Just and Pope. The estimation results indicate that changes in climate modify crop yield levels and variances in a crop-speci c fashion. For sorghum, rainfall and temperature increases are found to increase yield level and variability. On the other hand, precipitation and temperature are individually found to have opposite e ects on corn yield levels and variability.

Naylor et al(1972) made more extensive and detailed comparison of alternative methods and examined Box-Jenkins approach in contrast to Wharton econometric model for the year 1963 through 1967. They observed that the accuracy of ARMA models of Box-Jenkins methodology was considerably better than the accuracy of Wharton econometric model.

Nelson (1972) compared econometric (regression) and time-series (ARMA) methods for a longer time horizon. He concluded that the simple ARMA models are relatively more robust with respect to post sample predictions than the complex econometric models. If the mean square error is an appropriate measure of loss' an unweighted assessment clearly indicated that a decision maker will be better o relying simply on ARMA predictions in the post sample period i.e. in the forecasting phase.

Adam (1973) reported that the factors including the number of observations in the series, seasonality of the data, the number of periods in the time horizon are to be forecast to the extent in which randomness in the series and others had a substantial impact on the accuracy and performance of individual forecasting models.

Kimball and Gutterrez (1973) adopted Mincer-Zarnwitz technique of the goal of forecasting the minimization of the mean square error (MSE) i.e. the squared di erence between the actual and forecast values, which is a measure of dispersion around the line of perfect forecast. They indicated that a least square straight line must be tted to a scatter diagram of actual realization and estimates. They gave the idea that on the overall forecast accuracy, i.e. the square root of the MSE has been computed and expressed as a percent of the actual mean value. They also said that R^2 is not a reliable guide since it merely represents errors explained by a linear adjustment of the forecast series.

Leuthold et.al (1970) in their study of forecasting daily hog price and quantities' used Theil's inequality coe cient for comparing the predicative accuracy of the di erent forecasting approaches. For price forecast to hog market they compared econometric model, random walk model, and mean model and for supply forecasts they compared econometric model, random walk model, mean model and time-series models. They concluded that the data required for time series modelling was the concerned data on the variable to be forecast, whereas for econometric models data are needed on both the regress or and regress and. Therefore the forecasts using econometric model are slightly better than those using a stochastic non-casual frame-

work. Further, the cost of making slightly greater error in using the latter will be less than the additional cost involved in
setting up an econometric model and collecting the data.

Chambers et.al. (1971) in their study of 'how to choose the right forecasting technique' discussed time-series analysis also. They discussed the di erent forecasting techniques viz, qualitative method, time-series analysis, and projection (moving average, exponential smoothing, Box-Jenkins and trend projections), and casual methods (regression model, econometric model, Inputoutput model, leading indicator and life-cycle analysis). For each method they provided description, accuracy, identi cation or turning points, typical application and requirement of data. They tried to explain the potential of forecasting to the manager focusing special attention on sales forecasting for products of Corning glass works as these have matured through the product life cycle. They indicated that the need to-day is not for better forecasting methods but the better application of the techniques at hand. Similar ndings were also observed by Gross and Rain.

Both Reid and New bold (1971) concluded that the Box-Jenkins approach of ARIMA models gave more accurate results than exponential smoothing or step-wise regression methods.

Soliman (1971) worked out several major relationships that explained the behavior of the United States turkey industry in 1946-66. He constructed the model that consisted of four structural equations and he also used four di erent estimation techniques to derive the values of the structural parameters. He further observed that no one method proved superior to another with respect to their
forecasting ability against the observed data in the post 1966 period.

Kpongor (2007) evaluated the application of the APSIM-Sorghum model version 4.0 to predict grain and biomass yield response of sorghum to inorganic N and P fertilizer in a semi-arid region of Ghana under two management systems. The model performed well in predicting grain and biomass yield with an average R^2 of 0.81 and 0.86, respectively.

Delve et al. (2009) simulated P responses in annual crops on contrasting soil types using the APSIM-model for maize and beans in Kenya. The goodness t (r^2) between simulated and observed grain yield of maize was 0.81 and 0.74, whereas for biomass, this was 0.88 on Oxisol and 0.83 on an Andisol. An average r^2 of 0.79 and 0.69 was reported for grain and biomass of beans. The authors concluded that the model performed creditably in predicting the growth of maize and bean crops for the di erent P sources (fertilizer or chicken manure) and treatments (rates and frequency of application).

Chen et al. (2010) used a model to analyze the response of crop productivity to irrigation in the North China Plain, where excessive use of water for irrigation has caused a rapid decline in the groundwater table. Using data from three sites (Luancheng ,Yucheng and Fengqiu), they parameterized and evaluated the APSIM-Wheat model. The results showed that the model was able to simulate growth and yield of wheat and maize in a double cropping system.
Root mean squared error (RMSE) of yield and biomass simulations was 0.83 and 1.40 t/ ha for wheat, and 1.07 and 1.70 t /ha for maize, respectively. Soil water and evapotranspiration (ET) were also reasonably predicted. The simulated rain fed yields ranged from 0 to 6.1 t /ha for wheat and for maize 0 to 9.7 t /ha in a double cropping system. It was reported that for each 60mm additional irrigation water, crop yield increased by 1.2 t /ha; to achieve a yield potential of 7.1 t /ha of wheat and 8.3 t /ha of maize, 540 mm irrigation water would be required. The authors concluded that the model predicted grain yield, soil water and ET quite well.

Abuzar et al (2009) conducted a eld experiment to determine the e ect of plant population densities on maize at the Agricultural Research Institute, Dera Ismail Khan,in mid July 2009.The e ect of six plant population densities i.e. T1(40000 plants /ha), T2(60000 plants/ ha), T3(80000 plants ha^{-1}), T4(100000 plants /ha), T5(120,000 plants /ha) and T6(140,000 plants /ha) was investigated using maize variety Azam. Results showed that plant population of 40000 plants/ha produced maximum

number of grains per row (32.33) and grains per ear (447.3).However, 60000 plants ha^{-1} produced the maximum number of ears per plant (1.33), number of grain rows per ear (15.44), biomass yield (16890 kg/ha) and grain yield(2604 kg /ha).Therefore, planting density of 60000 plants/ha(keeping plant to plant distance of 22.70 cm) is recommended for
obtaining higher yield of maize.

Bruce et al (2001) worked on molecular and physiological approaches to maize improvement for drought tolerance. It came out average maize yields have increased steadily over the years in the USA and yet the variations in harvestable yield have also markedly increased. Much of the increase in yield variability can be attributed to varying environmental stress conditions; improved nitrogen inputs and better weed control; and continuing sensitivity of di erent maize lines to the variation in input supply, especially rainfall. Drought stress alone can account for a signi cant percentage of average yield losses. Yet despite variable environments, new commercially available maize hybrids continue to be produced each year with ever-increasing harvestable yield. Since many factors contribute to high plant performance under water de cits, e orts are being made to elucidate the nature of water-stress tolerance in an attempt to improve maize hybrids further. Such factors include better partitioning of biomass to the developing ear resulting in faster spikelet growth and improved reproductive success. An emphasis on faster spikelet growth rate may result in a reduction in the number of spikelets formed on the ear that facilitates overall seed set by reducing water and carbon constraints per spikelet. To understand the molecular mechanisms for drought tolerance in improved maize lines better, a variety of genomic tools are being used. Newer molecular markers and comprehensive gene expression pro ling methods provide opportunities to direct the continued breeding of genotypes that provide stable
grain yield under widely varied environmental conditions.

Kwesiga et al (2003) did a research on the e ect of short rotation Sesbania sesban planted fallows on maize yield. Two provenances of Sesbania sesban var.

$nubica$ (Kakamega and Chipata) were planted in fallows for 1, 2 and 3 years at $0.5m \times 0.5m$, $0.7m \times 0.7$ and $1.0m \times 1.0m$ spacing. Maize crop (MM604) was grown after fallow period at 0, 37, 74 and 112 kg N ha^{-1} to evaluate the e ects of nitrogen (N) and fallow on grain yield. There were no signi cant di erences between the two provenances of $S. sesban$. Wood biomass after 1, 2 and 3 years fallow at close spacing was 8.3, 17.6 and 21.4 t ha^{-1} for the Kakamega provenance and 10.8, 14.5 and 21.2 t ha^{-1} for the Chipata provenance. Litter fall in both provenances ranged from 0.6 t ha^{-1} in June to 0.01 t ha^{-1} in November. Stand mortality increased with plant density and fallow years: 27% in the rst year and about 90% by the end of the third year. Weed biomass ranged from 6.8 t ha^{-1} to 6.0 t ha^{-1} at close and wide spacing respectively. Maize grain yield without N was 2.27, 5.59 and 6.02 t ha^{-1} after 1, 2 and 3 years fallow respectively compared with the control plots with 1.6, 1.2, 1.8 tha^{-1} after 1, 2 and 3 years of continuous cropping. Even with addition of 112 kg N ha^{-1}, yield in the control plots declined from 6.09 to 4.88 and 4.28 t ha^{-1} after 1, 2 and 3 years of continuous cropping. In the planted fallows at 112 kg N ha^{-1}, maize yield increased from 6.75 to 7.16 and 7.57 t ha^{-1} following 1, 2 and 3 years fallow. It is concluded that short fallow rotations of 1-3 years using S. $sesban$ have a potential in increasing maize yield even without fertilizers. Thus, increasing the fallow period decreases the e ectiveness of inorganic fertilizers but increases grain yield for low fertilizer input.

Lobell and Field (2007) worked on Global scale climate-crop yield relationships and the impacts of recent warming. Despite the complexity of global food supply, here we show that simple measures of growing season temperatures and precipitation-spatial averages based on the locations of each crop-explain $\sim 30\%$ or more of year-to-year variations in global average yields for the world's six most widely grown crops. For wheat, maize and barley, there is a clearly negative response of global yields to increased temperatures. Based on these sensitivities and observed climate trends, we estimate that warming since 1981 has resulted in annual combined losses of these three crops representing roughly 40 Mt or $5 billion per year, as of 2002. While these impacts are small relative to the technological yield gains over the same period, the results demonstrate already occurring negative impacts of climate trends on crop yields at the global scale.

Kpotor (2012) evaluated newly released maize varieties in Ghana for yield and stability under three nitrogen application rates in two agro-ecological zones Farmers' adoption of hybrid varieties would reduce the large discrepancy between current low yields and achievable yields reported by maize researchers in yield evaluation trials as hybrids wield superior genetic potential over improved open pollinated varieties (OPVs) and local varieties due their heterozygosity resulting in their exhibition of high heterosis in yield and general performance. The current low yield necessitated the need to undertake this study to assess the relative yielding abilities and stability of 3 hybrid varieties, 5 OPVs, 1 local variety and 4 inbred lines under three levels of nitrogen fertilization at Kwadaso, a forest ecology, and Ejura, a transitional ecology, both in the Ashanti region of Ghana, in the major and minor seasons of 2011, respectively. Analysis of variance (ANOVA) revealed signi cant interactions for genotype by location ($G \times L$), genotype by nitrogen (G × N) and genotype by nitrogen by location ($G \times N \times L$) for grain yield. GGE biplot analysis for mean yield and stability also showed that hybrids had better yielding abilities than OPVs under both low and high nitrogen fertilization and at di erent environment. Economic bene t analysis also revealed that best option for highest net bene t is the cultivation of hybrid varieties under 90 kg Nha^{-1}. In order to bridge the gap between the current low yields and achievable yields in Ghana, farmers need hybrid seeds together with adequate levels of fertilizers.

Reidsma et al (2009) examined Regional crop modelling in Europe: The impact of climatic conditions and farm characteristics on maize yields Impacts of climate variability and climate change on regional crop yields are commonly assessed using process-based crop models. These models, however, simulate potential and water limited yields, which do not always relate to observed yields. The latter are largely in uenced by crop management, which varies by farm and region. Data on speci c management strategies may be obtained at the eld level, but at the regional level information about the diversity in management strategies is rarely available and di cult to be considered adequately in process-based crop models. Alternatively, understanding the factors in uencing management may provide helpful information to improve simulations at the regional level. In this study, we aim to identify factors at the regional level that explain di erences between observed and simulated yields. Observed yield data were provided by the Farm Accountancy Data Network (FADN) and Eurostat. The Crop Growth Monitoring System (CGMS), based on the WOFOST model, was used to simulate potential and water limited maize yields in the EU15 (i.e., the old member states of the European Union). Di erences between observed and simulated maize yields were analyzed using regression models. We assumed that the highest yields observed in a region were close to the yield potential as determined by climate and considered the average regional yields as also in uenced by management. Model performance was analyzed with respect to spatial and temporal yield variability. Results indicate that for potential yield, the model performed unsatisfactory in southern regions, where high temperatures increased observed yields which were in contrast to model simulations. When considering management e ects, we nd that especially irrigation and the maize area explain much of the di erences between observed and simulated yields across regions. Simulations of temporal yield variability also diverted from observed data of which about 80% could be explained by the climatic factors (35%) and farm characteristics (50%) considered in the analysis. However, e ects of speci c factors di ered depending on the regions. Accordingly, we propose di erent groups of regions with factors related to management which should be considered to improve regional yield simulations with CGMS.

Lizaso et al (2007) developed a sweet corn simulation model. Current production of sweet corn (*Zea mays* L.) in the United States is 4.0 million Mg with a value of $807 million. The fresh market component amounts to three-fourths of this value with California, Florida, and Georgia harvesting half of the U.S. fresh market production. Existing maize simulation models have limited potential to assist sweet corn production as a result of the distinctive nature of the marketed end product (i.e., fresh market ears versus dry mature kernels). The purpose of this study was to develop a sweet corn simulation model. The Cropping System Model-Crop-Environment Resource Synthesis CSM-CERESMaize simulation model, version 4.0, was modi ed to improve the simulation of ear growth, to predict ear fresh market yield, and to predict fresh market ear quality according to U.S. standards. A eld experiment conducted in Florida in 2003 was used for model development. Five nitrogen fertilization levels (0, 67, 133, 200, and 267 kg ha^{-1} N) were applied to a $sh2$-based commercial hybrid with a Bt gene sown at 8.2 plants/m^2. Three additional experiments conducted in 2002, 2004, and 2005 provided independent data to evaluate the new model. In 2002, the treatments and hybrid were the same as mentioned, but the population density was 5.5 plants $/m^2$. A yellow $sh2$-based hybrid with a Bt gene was planted at 6.1 plants $/m^2$ in 2004. In 2005, a bicolor $sh2$-based hybrid with a Bt gene was planted at 8.1 plants $/m^2$. The 2004 and 2005 experiments had 100% and 150% of the Florida N recommendations applied to the crop. Results indicated that the new model was able to simulate adequately crop and ear growth of sweet corn. The ear dry weight simulation was improved as indicated by 30% reduction of root mean square of the error (RMSE) when the new model was compared with the original CSM-CERES-Maize. Total ear fresh weight yield and marketable yield were also simulated reasonably well with RMSE values of 3367 and 3502 kg ha^{-1}, respectively. The simulation of ear quality was consistently over predicted at intermediate levels of N fertilization, indicating the need to further examine the impact of limited N on ear quality.

Shin et al (2009) did an evaluation of crop yield simulations with various seasonal climate data performed to improve the current practice of crop yield projections. The El Niæo Southern Oscillation (ENSO)-based historical data are commonly used to predict the upcoming season crop yields over the southeast United States. In this study, eight di erent seasonal climate data are generated using the combinations of two global models, a regional model, a statistical downscaling technique, and two convective schemes. These data are linked to maize and peanut dynamic models to assess their impacts on crop yield simulations compared to the ENSO-based approach. Improvement of crop yield simulations with the climate model data is varying, depending on the model con guration and the crop type. While the global climate model data provide no improvement, the dynamically and the statistically downscaled data show increased skill in the crop yield simulations. A statistically downscaled operational seasonal climate model shows a statistically signi cant (5%) inter annual predictability in the peanut yield simulation. Since the yield amount simulated by the dynamical crop model is highly sensitive to wet/dry spell sequences (water stresses) during the growing season, a proper parametrization of precipitation physics is essential in climate models to improve the crop yield projection.

Carberry et al (2012) came out with the fact that simulation models have proven bene cial to commercial farmers in Australia when applied within a participatory action research approach. This paper reports on an attempt to combine a participatory research approach and computer-based simulation modelling to engage smallholder farmers in Africa on issues of soil fertility management. A three-day interaction with farmers in one village in Zimbabwe provided evidence that the farmers found the simulation outputs to be credible and meaningful in a manner that allowed 'virtual' experiential learning to take place. The paper concludes that simulation applied within an action research framework may have a role in direct interventions with smallholder farmers in
such regions

The number of weeds in each of the plots was counted before herbicide application. Reduced rates of atrazine (562.5, 742.5, 1507.5, 1687.5 and 2250 g ha-1 (active ingredient)) were applied at the two to three-leaf stage of weed growth. Three weeks after herbicide application the number of surviving weeds was counted and recorded. Weed dry weight and maize grain yield per plot were determined. Percent control of the resident weed populations was calculated. At UZ farm, the 75% rate had as high percent weed control, as did the full label rate for all weed species. Amaranthus hybridus and Nicandra physalodes were the most susceptible species and were completely controlled even by the lowest rate (25%) at the UZ farm, while the same happened with Galinsoga parvi ora, Bidens pilosa and N. physalodes with the lowest rate (33%) at the UZ Department of Crop Science. Grass weed species (Setaria spp and E. indica) were more tolerant to reduced atrazine dosages than broadleaf weeds. There was no signi cant di erence ($P > 0.05$) in percent weed control at the UZ Department of Crop Science for the 100%, 67% and 33% rates for Commelina benghalensis and the grass species. Reduced rates of atrazine can be used to achieve equitable weed control as the full label rates without loss of maize yield. Results agree with
work done by other authors.

Maxmillan (1995) evaluated Maize Streak Virus Resistance in sixteen maize genotypes. Sixteen genotypes of maize were grown and arti cially infested with the maize streak virus (MSV), using reared leaf hoppers (Cicadulina mbila) that had acquired the virus from streak infested plants, to determine the level of MSV resistance in each genotype and to determine the relationship between plant height, yield and the MSV disease score. Most cultivars except Pool 16, R201, Pool 16 SR and Maka SR had MSV disease scores of greater than 3 showing that
they were relatively susceptible to the disease. There was a negative relationship between MSV disease score and plant height. Uninfected plants were higher
yielding than their infected counterparts. However, correlation analysis showed that there was no signi cant relationship between MSV scores, yield and owering dates.

The MSV disease results in plant height and yield reduction in susceptible genotypes. The screening technique is useful for identi cation of MSV resistance vis-a-vis ratings of entries.

Chifamba(1995) worked on response to simulated drought conditions of tall and short maize (Zea mays) near isogenic lines .The responses for yield, anthesis-silking, harvest index and wilt rating of four tall and four short maize (Zea mays) near-isogenic lines were studied at 100, 50, 25 and 10 per cent soil AWC (available water content) in the greenhouse. In short lines, grain yield showed signi cant positive correlations with harvest index ($r = 0.9760$) and height ($r = 0.3556$). In tall lines, grain yield was found to have signi cant positive correlations with harvest index ($r = 0.9616$) and height ($r = 0.5210$). A signi cant negative correlation was found between the overall grain yield and anthesis-silking interval (ASI) for both tall and short lines, with correlation coe cients of -0.8565 and -0.9054 respectively. There was no signi cant correlation between overall grain yield and wilt rating for both tall and short lines. The tall genotypes signi cantly yielded better than the dwarf lines at 100% and 50% soil AWC. At 25% and 10% soil AWC, the short plants performed better than their tall counterparts but this was not statistically signi cant. At 100% soil AWC, there was no signi cant di erences in the overall mean ASI between the tall and short genotypes. At 50%, 25% and 10% soil AWC, the tall lines had signi cantly higher overall mean ASI than those of short counterparts. There were signi cant di erences between the overall mean wilt ratings and overall mean height of tall and short genotypes at each moisture level. The e ects of decreasing moisture levels were more pronounced in tall plants than in short plants.

Moyo (1995) worked on response to selection for drought tolerance of two maize populations. Two populations of maize were studied for their response to selection for drought tolerance by: 1) determining the relationship between yield and anthesis-silking interval (ASI) of lines from cycle 2 (C2) and cycle 3 (C3) of recurrent selection of ZM601 (DR) and lines from Tuxpeno Sequia (a CIMMYT population which has undergone six cycles of recurrent selections); 2) comparing ASI, grain yield, and other physiological traits and to con rm the utility of various secondary traits; and 3) con rming the use of secondary traits as selection criteria when breeding for drought tolerance. The secondary traits were: leaf rolling, leaf erectness and tassel size. Under stressed conditions Tuxpeno Sequia produced a signi cantly higher yield than ZM601. Yield was negatively correlated with ASI ($r = -0.08$), but not signi cantly so ($P > 0.05$), for Tuxpeno Sequia and ($r = -0.22$) for ZM601, meaning yield increased as ASI decreased. Mean ASI was signi cantly shorter for Tuxpeno Sequia (0.63 d) under moisture stress than for ZM601 (2.23 d) where as under adequate moisture conditions ASI did not di er for the two populations. Linear correlation analysis indicated signi cant correlations with yield for both populations ($P < 0.05$) for the secondary traits ($r = -0.11$ to -0.18). The results seem to indicate that Tuxpeno Sequia performs better than ZM601 under moisture stressed conditions. The fact that the correlation coe cient of yield with ASI ($r = -0.08$) was not signi cant (P>0.05) for Tuxpeno Sequia is an indication that the population has been more for this character than ZM601 ($r = -0.22$), ($P < 0.05$).

Mugweni (1996) conducted a eld experiment in summer to determine the e ect of maize residue at 0 t/ha, 5 t/ha, 10 t/ha and 15 t/ha and the herbicide dimethenamid and alachlor applied as pre-emergence treatments on the control of ve weed species, namely Nicandra physalodes, Richardia scabra, Galinsoga parvi ora, Chenopodium album and Amaranthus hybridus in a soyabean crop. Rainfall was abundant at the time of herbicide application. Two di erent assessments, weed counts and percent control, were carried out at three weeks after planting (WAP) and 6 WAP. At three weeks after herbicide treatment, emergence of all weeds had not been signi cantly suppressed by maize residues at all levels of mulch. Both herbicides signi cantly ($P < 0.01$) reduced emergence of all weeds except C. album and A. hybridus but their e ect was not in uenced by maize residues at all levels although a gradual decline in weed emergence as mulch increased was noticeable. For all the weeds added together there was a signi cant ($P < 0.01$) mulch herbicide interaction.

Dimethenamid was more e ective than alachlor when no mulch was applied but had similar levels of control to alachlor on weed emergence at all levels of mulch. Visually assessed percent control of weeds 6 WAP showed that herbicide e cacy was reduced by the presence of maize residues. There was a consistently signi cant decline in percent control as mulch level was increased for all weeds except C. album. C. album had the lowest percent control and its control was not in uenced by the presence of maize residues.

Bwerazuva (1996) conducted a eld study to evaluate resistance to maize stalk borer and minimum sample size requirement for leaf damage rating among selected maize genotypes. Leaf damage ratings showed the susceptible entry DMRESR-w to have the highest mean rating (on a scale of 1 to 9) of 7.5, while the hybrid entries which were crosses between a susceptible and a resistant source showed no signi cant di erences between the di erent mean leaf damage ratings. The hybrid entries gave progeny with an average leaf rating of 6.3 (on a scale of 1 to 9) showing intermediate resistance to maize stalk borer. On a yield reduction basis, CML123/DMRESR-w showed the greatest yield reduction of 46% compared to the thionex-protected entry. CML139/DMRESR-w showed the least yield reduction of 10% in comparison to the thionex-protected entry. Thus, resistance should not be evaluated on a leaf damage basis only but should also include measurements of yield reduction as some material may have poor resistance to leaf damage but still show little yield reduction. Calculation of minimum sample size showed that DMRESR-w/CML123 required the largest sample size (49 plants), while CML67/DMRESR-w had the lowest sample requirement of 31 plants. Di erences in sample sizes and yield between reciprocal crosses suggest the possible presence of maternal e ects.

Mwashaireni(1996) conducted an experiment to assess the e ects of intrarow intercropping of maize and pumpkin was conducted at the University of Zimbabwe Farm during the 1995/96 season. Maize hybrid R201 was intra-row intercropped with a pumpkin landrace, Nzunzu. Intra-row intercropping with pumpkins did not signi cantly reduce maize grain yield and maize stover biomass when compared with the sole-maize stands. However, maize and pumpkin intrarow intercropping led to poor pumpkin spread and subsequently resulted in absence of pumpkin fruit yield. The sole-pumpkins spread better than pumpkins in the intercrop situation. As a result, the pumpkins in the sole situation yielded signi cant amount of fruits. Weed biomasses, unlike weed numbers, were signi cantly higher in the pumpkin sole-crops compared to intercrops. Therefore, pumpkin intra-row intercropped in maize was e ective in suppressing weed growth but did not have any e ect on weed germination. Besides weed growth suppression, maize and pumpkin intra-row intercropping will allow farmers to bene t from the pumpkin leaf (useful as a vegetable) without losing out on maize

yield.

Gumunyu (1996) conducted a study to evaluate maize experimental varieties from CIMMYT at high and low N levels, using a selection index which included grain yield, ear leaf chlorophyll content, ear leaf area under low N only, anthesis-silking interval (ASI) and plant height. Signi cant interactions between maize varieties and N levels were observed for ear leaf chlorophyll content and grain yield. Signi cant di erences among N levels were observed for grain yield, ear leaf chlorophyll content, ear leaf area, plant height and ASI. The experimental varieties showed signi cant di erences in grain yield, ear leaf chlorophyll concentration, plant height and ASI under high or low N levels. Low N levels resulted in reduced grain yield, reduced ear leaf chlorophyll content, reduced plant height and increased ASI. Correlation analysis showed grain yield to be highly increased with chlorophyll content, ear leaf area and plant height
under low N.

Munyaka(1996) worked on four CIMMYT maize (Zea mays L.) genotypes which had each undergone four cycles of recurrent selection for other agronomic characteristics like yield were evaluated using the penetrometer method for changes

in stalk strength. The maize varieties used were ZM601, ZM605, ZM607 and ZM609. Penetrometer measurements were done at the internode below the ear and the rst well-developed internode above ground. Root lodging counts in the eld were negatively correlated with root force (r = -0.35; P < 0.05). Plant height was highly signi cantly correlated with ear height (r = 0.81;P < 0.001). There was little change in rind puncture resistance over the four cycles of recurrent selection. No de nite trend in changes in stalk strength was observed in the cycles of selection. Recurrent selection for other agronomic traits such as nitrogen use e ciency, high yield and disease resistance did not a ect the stalk quality
signi cantly.

Chihande (1996) studied the e ect of seed size, simulated moisture stress and planting depth on germination and establishment of dwarf maize (Zea mays) hybrids. Three experiments were carried out on the germination of ve maize hybrids. The hybrids were three dwarf hybrids $EM42$? $6L57$, $10LK2$? $12Dr$ and $32Y$? $12Dr$, and two conventional hybrids R201 and SC701, which were all obtained from the African Centre for Fertilizer Development. The objectives of these experiments were to determine the e ect of seed size and planting depth on germination and emergence of the maize hybrids, and to determine how seed size in uences germination under simulated moisture stress conditions. The experiments were carried out in a greenhouse, incubator and growth room at a maximum temperature of 25° C. Seed size did not in uence the percent emergence of the hybrids and there was no signi cant di erence in the percent emergence among the ve hybrids. However, all the dwarf hybrids had a higher rate of water uptake than the conventional hybrids, and within the dwarf hybrids the small seed size had a higher rate of water uptake than the large seed size lots. Under moisture stress conditions the dwarf hybrids had a higher percent germination when compared with the conventional hybrids. The advantages of the dwarf hybrids were attributed to the characteristic small seed size which tends to have a high surface area to volume ratio. It is also probable that di

erences between the hybrids were due to genetic factors or variation in the composition
of the maize seeds.

Musara(1996) studied the e ect of di erent basal fertilizers and seed generations on maize growth and yield .Two fertilizer types, ammonium nitrate (AN) (70 kg/ha) and compound D (300 kg/ha) were used as basal fertilizer dressings on the F1 and the F2 generation of maize variety SC601 and on the F1 generation of maize variety R201 to determine their e ect on growth and grain yield. This work was conducted on a sandy soil at Domboshava Training Centre near Harare. The growth pattern in terms of leaf area index and dry
matter production were signi cantly a ected by type of basal fertilizer dressing. The F2 generation of SC601 and those treatments with ammonium nitrate had a lower leaf area index, and dry matter production than the F1's during the early stages of growth. However, these treatment di erences disappeared after four weeks .Genotype (F1 and F2 of SC601) had no e ect on grain yield. However,
grain yield was also signi cantly ($P<0.05$) a ected by the type of basal fertilizer used, with F1 treatments which had compound D giving higher yields than the treatments which had AN. There was no interaction e ect of basal fertilizer and maize generation. Inorganic fertilizers are expensive and recommendations to resource poor farmers as a blanket application are not often pro table. For farmers that can a ord fertilizer, there is urgent need to increase the pro tability of this input by using better-targeted recommendations. Basically, if a farmer has more money then he or she can use compound D as well as AN, otherwise it
is better to put AN as a basal and topdressing than not apply any fertilizer.

Mafa (1996) did an intercropping experiment in the 1995/96 rainy season to study the e ects of two maize (Zea mays) densities and two dry bean (Phaseolus vulgaris) row arrangements on maize and bean yield, yield components, and phenological development. A factorial arrangement was used with two maize densities (i.e. $90cm \times 30cm$ and $90cm \times 45cm)$× two row arrangements of two bean varieties to give a total of four sole-crop and eight intercrop treatments. The trial was planted at two sites, Domboshava (Natural Region IIa) and Chinyika (Natural Region IIb), with Chinyika receiving one dry bean variety. Growth and yield data were collected and analyzed statistically to see if treatment means di ered signi cantly ($P = 0.05$). Grain yield of both components were

signi cantly decreased with intercropping, but the overall total yield was higher in intercrop than in sole-crops. Although grain yield of beans was severely depressed in intercrops, maize yield compensated for the loss giving a system's advantage of up to 30%. Maize grain yield also decreased with increase in number of rows of beans. The bean variety Natal Sugar depressed maize yield more than Carioca. Planting pattern and density had no e ect on phenological development.

Manzira (1997) analyzed di erent maize hybrids for resistance to grey leaf spot disease .Grey Leaf Sport Disease, a new disease in Zimbabwe, can cause up to 50% yield loss if fungicides are not used to control the disease. The objectives of this project were to quantify the levels of resistance to grey leaf spot on di erent maize hybrids, to nd the e ectiveness of spraying on the severity of grey leaf spot and to nd the best scoring date in screening for resistance. Eleven commercially available maize hybrids were planted at the Agricultural Research Trust Farm in the 1996/97 season in two trials with and without fungicide spray to control grey leaf spot. Scores of the severity of the disease were taken on three di erent days, 90, 120 and 150 after planting in the unsprayed trials. In the sprayed trials Benomyl was applied at 750 g/ha. Grain yield was measured in both trials. The varieties SC709 and SC625 were tolerant to grey leaf spot giving yields of 9.93 and 8.53 t/ha respectively. These varieties also had the lowest scores in the unsprayed trials. Spraying signi cantly increased the yields of all the varieties. There were signi cant negative correlations between yield and disease scores at days 120 and 150 with r-values of 0.601 and 0.634 respectively.

Fusire (1997) designed a project to determine the optimum planting depth desirable for each seed size in a single environment in Zimbabwe. An experiment was carried out at the International Centre for Maize and Wheat Improvement (CIMMYT). Five 5 cm and 10 cm depths were used for eight

di erent seed grades classi ed according to a CIMMYT classi cation and grading system. Total percentage of germination, plant height at the fth and tenth leaf stage, and yield were used to evaluate the e ect of maize seed size, plant performance and yield as in uenced by planting depth. The Minitab statistical package was used for variance analysis. The 5 cm depth produced the tallest pants and the highest percentage of germination especially for mixed, thick and at seeds. The mixed (largest) seed had the tallest height at both fth and tenth leaf stage as in uenced by the initial seed size. The large seed also had the highest yield due to higher vigour or plant performance as compared to the smaller seeds. Even though the latter had the largest count, it had the least seed weight resulting in the lowest tonnage. It was also concluded that a planting depth of 5 cm was

ideal for early plant establishment/germination and better yields in environments with adequate moisture.

Khun and Smith (1997) evaluated maize for resistance under greenhouse conditions was used in the eld. The progress of symptom development was assessed at 6, 11, 16 and 28 days post-inoculation using the severity rating method. Both the severity rating and index methods indicated three varieties to be tolerant while the other six varieties were shown to be susceptible. In the susceptible hybrids, 10-32% of the plants had developed symptoms by six days after the rst inoculation. Analysis of plant heights before; during and after tasselling showed that growth of two of the tolerant varieties was not signi cantly a ected. Grain yield of all the tolerant cultivars, including one of the susceptible

ones, was not signi cantly reduced.

Chitokomere(1997) carried out a study at the University of Zimbabwe to nd out the e ect of genotype and seed size on germination and establishment of maize seed. The study consisted of three experiments. The rst experiment was designed to determine the e ect of genotype (variety) and seed size on the amount of water taken up by maize seeds in the rst 48 hours after immersion in water. The second experiment was designed to determine the e ect of di erent levels of simulated moisture stress (0, -5, -10 and -12 bars), seed size and genotype on maize seed germination. The third experiment was designed to determine the e ect of genotype, seed size and planting depth on emergence of maize seed in pots. Water uptake rate, as indicated by the change in mass per unit seed mass, signi cantly ($P < 0.01$) increased with a decrease in seed size. Water uptake signi cantly ($P < 0.01$) varied among maize varieties, but the di erences were not consistent at 24 and 48 hours after immersion in water. Germination under simulated moisture stress signi cantly ($P < 0.01$) decreased with increase in level of moisture stress. Large seeds had signi cantly lower germination than medium and small seeds under simulated moisture stress. Genotype had signi cant e ect ($P < 0.01$) on percent germination under simulated moisture stress. However, at eight days after seed immersion in osmotic solutions, there was a signi cant genotype and simulated moisture stress level interaction indicating that germination percentage of varieties di ered according to simulated moisture stress level. Emergence decreased with planting depth and there were varietal di erences in their ability to emerge from depth, but seed size had no e ect. At 20 days after planting, there was a signi cant interaction ($P < 0.05$) between variety and planting depth, showing that the ability of the di erent varieties to emerge di ered with the depth at which they are planted. There was a clear relationship between water uptake in the initial 48 hours and germination of varieties under simulated moisture stress.

Zheke (1997) conducted a study on the e ect of season quality on the time to owering and on pollen shed synchronization with silk emergence of di erent maize hybrids (Zea mays).Eight maize hybrids, SC701, ZS206, SC601, 93MW5, SC501, R201, ZS225 and SC401 were grown for four consecutive season 1993/94, 1994/95, 1995/96 and 1996/97 at Rattray Arnold Research Station (RARS). The hybrids, SC701, ZS206, 93MW5, SC501 and R201 were stable in their days to mid-pollen shed (DMP) and days to mid-silk emergence (DMS). But ZS206, SC601 and ZS225 varied signi

cantly ($P < 0.05$) from season to season. Six of the eight hybrids SC701, ZS206, SC601, ZS225, SC501 and 93MW5 were more stable in their synchronization of pollen shed with silk emergence than SC401
and R201. Thus SC401 and R201 was the least stable of the eight hybrids.

Madzokere (1997) conducted a study on grain yield, light interception and canopy development in intercrop systems of maize and sorghum with cowpea in Natural Region Five . An intercropping experiment is described for the 1996/97 growing season, at Save Valley Experiment Station. The experiment was a row intercrop in which maize at 11 000 and 22 000 plants/ha and sorghum at 33 000 and 66 000 plants/ha were intercropped with a single or double rows of cowpeas spaced at $0.75m \times 0.1m$ and $0.385m \times 0.1m$ between the cereal
component respectively. Combined grain yields of maize + cowpea and sorghum + cowpea were more than the yield of sole components and therefore the total productivity of the system was well above that of pure stands. Maize at 22 000 plants/ha intercropped with one row of cowpea and sorghum at 66 000 plants/ha intercropped with one row of cowpea proved to be the best intercropping systems in terms of grain yield and land equivalent ratio (LER). Leaf area index of maize and sorghum (LAI) was lower in intercropped systems than in pure stands. In terms of light interception, intercropping systems were more e cient in light utilization compared to sole crops. The most e ciency was realized in intercrop systems of cereals at high populations with two rows of cowpea during the rst weeks due to early canopy cover but during the later weeks there is was no
signi cant di erence in light interception between the intercrop systems.

Chipomho (1997) conducted an intercropping experiment in the 1996/97 rainy season to study the e ect of maize (Zea mays) canopy architecture and nitrogen rate on yield in a maize and bean intercrop trial. Two maize hybrids, which di er morphologically, SC 501 with a planophile canopy architecture and PHB 3442 with an erectophile canopy architecture, were used. The maize hybrids were intercropped on the row with Umkuzi variety of bean. The experiment had a total of eight intercrop treatments and ve sole crop treatments. A two by four factorial arrangement in a randomized complete block design with three replicates was used. The trial was conducted at Domboshava Training Centre in Natural Region IIa. Growth measurements were taken at six weeks after emergence during the bean owering phase. Yield data was collected and analyzed statistically to see if treatment means di ered signi cantly ($P = 0.05$). Maize yield, in the intercrop averaged with no signi cant di erence between the two hybrids. Fertilizer level had an e ect on maize with both hybrids responding well to increase in nitrogen level. Maize yield increased by 38 kg per single kilogram unit of fertilizer. Bean yield responded to applied nitrogen between zero and 30 kg N per ha, however there was no signi cant di erence between the di erent nitrogen rates of 30, 60 and 90 kg N per ha. Maize variety had an e ect on bean yield with PHB 3442 intercropped bean giving a higher yield than SC 501. Yield reduction of the bean crop with SC 501 was between 25 and 56% and with PHB 3442 between 13 and 30%.

Chiyanike(1999) conducted an intercropping experiment was done in the 1996/97 rainy season to evaluate eight bean (Phaseolus vulgaris) varieties in maize (Zea mays L.) bean intercrop. Of the eight varieties, Natal Sugar was used as a control as it is the most widely grown in the smallholder sector. Results from national trials have shown that there are other varieties that yield higher than Natal Sugar in sole cropping. This trial was to evaluate some of these varieties in intercrops as this is a fast growing practice in the smallholder sector. The trial was conducted at two sites, Domboshava (Natural Region IIa) and Chinyika (Natural Region IIb). A split plot design was used, the maize and beans were planted in the same row (maize at $90cm \times 30cm$ spacing and beans at 90 cm

x 10 cm spacing). Yield data was collected and analyzed statistically to see if treatment means di ered signi cantly ($P \leq 0.05$) both the maize and bean yields in the intercrop were reduced. There were no signi cant di erences in the maize yield from the di erent treatments but there were signi cant di erences in the bean varieties. The dry varieties tested yielded higher than Natal Sugar in the intercrop with A286, MC5001 and 36/6/10 giving the highest yields.

Chakauya(1999) conducted a eld experiment on the Response of Maize Inbred Lines to Seedling Drought Stress .Ten maize inbred lines obtained from the Centro International de Majoramiento de Maiz Trigo (CIMMYT) were evaluated in a greenhouse experiment for genotype responses to seedling drought stress. Inbred line CML202 was used in the rst experiment to determine the most suitable parameter for measuring desired drought stress intensity. Parameters that were used of predicting drought stress included actual evapotranspiration (ET), growing degree-days (GDD), days after planting (DAP) and potential evapotranspiration (PET). Potential evapotranspiration seemed to be the most suitable of genotype response to drought stress. Fifty per cent plant survival was at 38 mm PET. Genotypes were provided with an initial irrigation to ensure germination and re-watered after signs of drought stress. Soil water potential and ET were determined by weighing each pot. Potential evapotranspiration was measured using evaporation pans. Among the ten inbred lines evaluated, inbred line Z180017 had the highest plant survival (58%) and the least leaf senescence (58%). Inbred line Z180028 had the least plant survival (8%) and highest leaf senescence (92%). The results indicated that genotype di erences to the response to seedling drought stress existed in maize inbred lines.

Kimbini(1999) conducted a study to evaluate the e ect of grey leaf spot on some locally available commercial maize hybrids on yield and to evaluate the e ect of grey leaf spot on days to owering of the commercial maize hybrids. Twenty-four commercial maize hybrids were planted at Cargill Research Station during the 1997/98 agricultural season, in two trials which included fungicide spraying and no fungicide spray to control the grey leaf spot. Three hybrids of maize, which are resistant to grey leaf spot, were used as checks. The resistant checks showed no yield reductions between sprayed and unsprayed treatments whilst susceptible entries showed signi cant reduction of yield in unsprayed treatments when compared to the respective sprayed treatments. Furthermore grey leaf spot also delayed days to owering for those hybrids which succumbed

to the disease.

Chapter 3

METHODOLOGY

3.1 Introduction

This chapter considers thoroughly the basic plots, de nitions and concepts of

time series analysis, assumptions, conditions, principles processes involved in the application of autoregressive moving average (ARMA), multiple linear regression and estimation of regression coe cients.

3.2 Basic Concepts and De nitions of Time Series

3.2.1 Basic de nitions

Time series is de ned as a collection of observations or measurements on quantitative variables made sequentially, usually daily, weekly, monthly, quarterly, annually, and so on. Examples include semi-annual grain yield of maize for a period of fteen years, daily stock prices of a rm for a period of one year, monthly electricity consumption for a household for a period of ve years and so on. Time series analysis comprises methods that break down a

series into components and explainable portions that allow trends to be identi ed, estimates and forecasts to be made. Basically time series analysis attempts to understand the underlying context of the data points through the use of a model to forecast future values based on known past values. Such time series models

include GARCH, TARCH, EGARCH, FIGARCH, CGARCH, ARMA, ETC but the main focus of this study is based on ARMA models.

3.2.2 Time Series Graph

Time series plot is simply a graph which displays observations on the y-axis against equally spaced time intervals on the x-axis. The time series plot speci cally consists of: Time scale (index, calendar, clock, or stamp column) on the x-axis; and lines displaying each time series as shown in the Figure 3.1 below for a given hypothetical data. The plots are usually used to: detect seasonality in your data; and compare trends across groups.

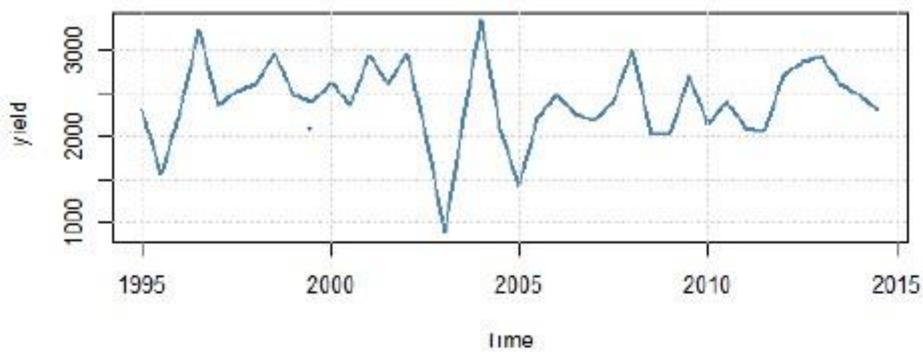


Figure 3.1: Time Series Plot of grain yield from 1995 to 2014

3.3 Components of Time Series

A virtual step in choosing appropriate modelling and forecasting procedure is to consider the type of data patterns exhibited from the time series graphs of the time plots. The sources of variation in terms of patterns in time series data are mostly classi ed into four main components. These components include seasonal variation; trend variation; cyclic changes; and the remaining ’irregular’ uctuations.

3.3.1 The Trend (T)

The trend is simply the underlying long term behaviour or pattern of the data or series. The Australian Bureau of Statistics (ABS,2008) de ned trend as the ’long term’ movement in a time series without calender related and irregular e ects, and is re ection of the underlying level. It is the result of in uences such as population growth, price in ation and general economic changes.

3.3.2 Seasonal variation (S)

A seasonal e ect is a systematic and calendar related e ect. Some examples include the sharp escalation in most Retail series which occurs around December in response to the Christmas period, or an increase in water consumption in summer due to warmer weather. Other seasonal e ects include trading day e ects (the number of working or trading days in a given month di ers from year to year which will impact upon the level of activity in that month) and moving holidays (the timing of holidays such Easter varies, so the e ects of the holiday will be experienced in di erent periods each year). Seasonal adjustment is the process of estimating and then removing from a time series in uences that are systematic and calendar related. Observed data needs to be seasonally adjusted as seasonal e ects can conceal both the true underlying movement in the series, as well as certain non-seasonal characteristics which may be of interest to analysts. Seasonality in a time series can be identi ed by regularly spaced peaks and troughs which have a consistent direction and approximately the same magnitude every

year, relative to the trend as depicted in the gure 3.2 below. Other techniques

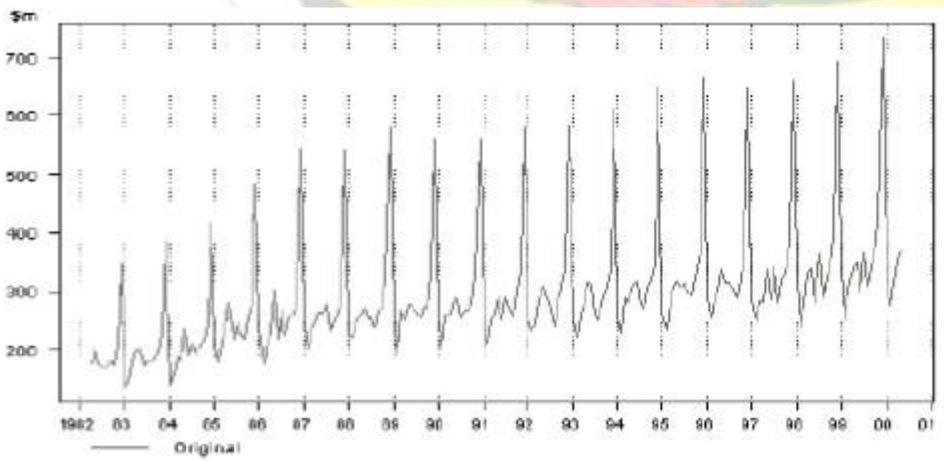


Figure 3.2: Graphical display of seasonal e ect of a hypothetical data

that can be used in time series analysis to detect seasonality include:

1. A seasonal subseries plot is a specialized technique for showing seasonality.

2. Multiple box plots can be used as an alternative to the seasonal subseries plot to detect seasonality.

3. The autocorrelation plot can help identify seasonality.

3.3.3 Cyclical variations(C)

Cyclical variations are the short term uctuations (rises and falls) that exist in the data that are not of a xed period. They are usually due to unexpected or unpredictable events such as those associated with the business cycle sharp rise in in ation or stock price, etc. The main di erence between the seasonal and cyclical variation is the fact that the former is of a constant length and recurs at regular intervals, while the latter varies in length. More so, the length of a cycle is averagely longer than that of seasonality with the magnitude of a cycle usually being more variable than that of seasonal variation.

3.3.4 Irregular variations (I)

The irregular component (sometimes also known as the residual) is what remains after the seasonal and trend components of a time series have been estimated and removed. It results from short term uctuations in the series which are neither systematic nor predictable. In a highly irregular series, these uctuations can dominate movements, which will mask the trend and seasonality. The Figure 3.2 below is a graph which is of a highly irregular hypothetical time series.

3.4 A common assumption in Time Series

Techniques

A common assumption in many time series techniques is that data are stationary. A stationary process has the property that the mean, variance and autocorrelation

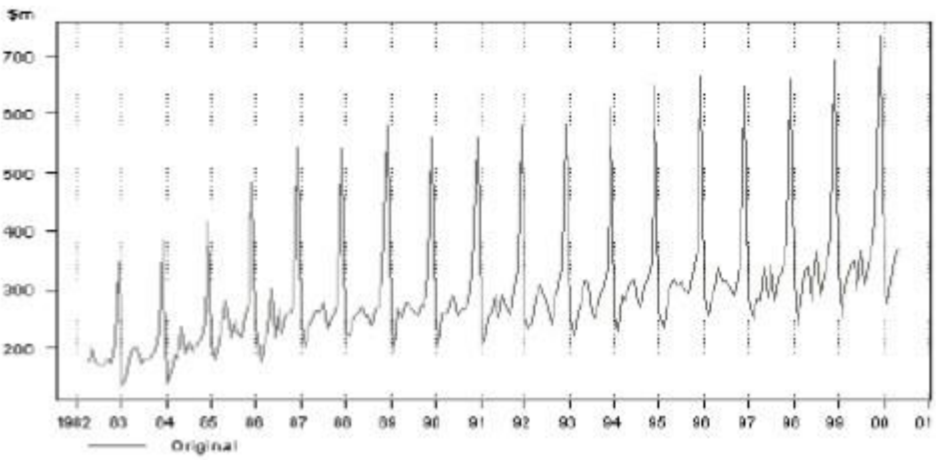


Figure 3.3: Typical irregular e ect graph of a hypothetical time series data.

structure do not change over time. Stationary can be de ned in precise mathematical terms as

(i) the mean $\mu(t) = E(y_t)$

(ii) the variance $\sigma^2(t) = var(y_t) = \gamma(0)$

(iii) The autocovariances $\gamma(t_1,t_2) = Cov(y(t_1),y(t_2))$ hence a time series is said to be strictly stationary if the joint distribution of any set of n observations $y(t_1,t_2) = Cov(y(t_1),y(t_2))$ is the same as the joint distribution of $y(t_1),y(t_2),...,y(t_n)$ for all n and k. If the time series is not stationary, we can often transform it to stationary with one of the following techniques.

1. We can di erence the data. That is, given the series Z_t, we create the new series

$$Y_i = Z_i - Z_{i-1}$$

The di erence data will contain one less point than the original data. Although you can di erence the data more than once, one di erence is usually su cient.

2. If the data contain a trend, we can t some type of curve to the data and then model the residuals from that t. Since the purpose of the t is to simply remove long term trend, a simple t, such as a straight line, is typically used.

3. For non-constant variance, taking the logarithm or square root of the series may stabilize the variance. For negative data, you can add a suitable constant to make the entire data positive before applying the transformation. This constant can then be subtracted from the model to obtain predicted (i.e.,the tted) values and forecasts for future points.

White noise process $x_t \sim i.i.d\ N(0,\sigma^2)$

A sequence x_t is a white noise process if each value in the sequence has zero-mean $E(x_t) = E(x_{t-1}) = 0$ constant conditional variance $E(x_t^2) = E(x_{t-1}^2) = ..\sigma^2 =$ $Var(x_{t-1})$ is uncorrelated with all other realizations $E(x_t x_{t-s}) = E(x_{t-j} x_{t-j-s}) =$ $.. = 0 = Cov(x_{t-j} x_{t-j-s})$

Property 1 and 2: absence of serial correlation or predictability Property 3:

Conditional homoscedasticity (constant conditional variance)

3.4.1 Covariance Stationarity (weakly stationarity)

A sequence x_t is covariance stationary if the mean, variance and autocovariance do not grow over time, i.e. it has nite mean $E(x_t) = E(x_{t-1}) = .. = \mu$ nite variance $E[(x_t - \mu)^2] = E[(x_{t-1} - \mu)^2] = .. = \sigma_x^2 = Var_x$ nite autocovariance $E[(x_t - \mu)(x_{t-s} - \mu)] = E[(x_{t-j} - \mu)(x_{t-j-s} - \mu)] = .. =$

$\sigma_s = Cov(x_{t-j} x_{t-j-s})$

Example: autocovariance between $x_t, x_{t-s} = \sigma_s = \dfrac{\gamma_s}{\gamma_0}$

But white noise process does not explain macro variables characterized by persistence so we need AR and MA features.

AR(1): $x_t = \rho x_{t-1} + e_t$ $\quad e_t \sim i.i.d(0,\sigma^2)$, $\quad$ (random walk: ρ = 1)

MA(1): $x_t = e_t + \theta e_{t-1}$

More generally:

AR(p): $x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \ldots + \rho_p x_{t-p} + e_t$

MA(q): $x_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q}$

ARMA(p,q): $x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \ldots + \rho_p x_{t-p} + e_t + \theta_1 e_{t-1} + \ldots + \theta_q e_{t-q}$

Using the lag operator $x_t = \rho_1 x_{t-1} + \rho_2 x_{t-2} + \ldots + \rho_p x_{t-p} +$

$e_t + \theta_1 e_{t-1} + \ldots + \theta_q e_{t-q}$

AR(1): $(1 - \rho L)x_t = e_t$

MA(1): $x_t = (1 + \theta L)e_t$

AR(p): $(1 - \rho_1 L - \rho_2 L^2 - \ldots - \rho_p L^p)x_t = \rho(L)x_t = e_t$

MA(q): $x_t = (1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q)e_t = \theta(L)e_t$

ARMA(p,q): $a(L)x_t = b(L)e_t$

3.5 Univariate Time Series Models

There are time series analysis models with only one series of is a typical example of a univariate time series. Univariate time series models usually view their series as a function of its own past, random shocks and time. Some basic univariate time series models and their processes are discussed below as follows.

3.5.1 Common Approaches to Univariate Time Series

There are a number of approaches to modelling time series. A few of the most common approaches are outlined below.

a) Decomposition

One approach is to decompose the time series into a trend, seasonal, and residual component. In other words decomposition refers to separating a time series into trend, cyclical, and irregular e ects. Decomposition may be linked to de-trending and de-seasonalizing data so as to leave only irregular e ects, which are the main focus ot time series analysis. Triple exponential smoothing is an example of this approach. Another example, called seasonal loss, is based on locally weighted least squares and
is discussed by Cleveland(1993).

b). The spectral plot

Another approach, commonly used in scienti c and engineering applications, is to analyze the series in the frequency domain. An example of this approach is modelling a sinusoidal type data set as shown in the beam de ection case study. The spectral plot is the primary tool for the frequency analysis of time series. Detailed discussions of frequency-based methods are included in Bloom eld(1976), Jenkins and Watts(1968), and Chat eld(1996).

c). Autoregressive(AR) models

Another common approach for modelling univariate time series is the autoregressive(AR) model. An autoregressive model is simply a linear regression of the current value of the series against one or more prior values of the series AR_p. The value of p is called the order of the AR model. AR models can be analysed with one of the various methods, including
standard linear least squares techniques. They can have a straightforward interpretation.

AR Process

Stationary Conditions for an AR(1) process

$(1-\rho L)x_t = e_t$ with $\rho(L) = 1-\rho L$ and substituting for L : $\rho(z) = 1-\rho z$

The process is stable if $\rho(z)$ 6= 0 for all numbers satisfying $|z| \leq 1$. Then we can write

$$x_t = (1-\rho L)^{-1} e_t = \sum_{i=0}^{\infty} \rho^i e_{t-i}$$

If x is stable, it is covariance stationary:

$E(x_t) = \mu$ or $0 - finite$

$Var_x = E[x_t^2] = E(\sum_{i=0}^{\infty} \rho^i e_{t-i})^2 = \sigma_e^2 \sum_{i=0}^{\infty} \rho^{2i} = \dfrac{\sigma_e^2}{1-\rho^2} = \gamma_0$ - nite covariances

$\gamma_1 = E(x_t x_{t-1}) = E[(\rho x_{t-1} + e_t) x_{t-1}] = \rho \sigma_{x2}$ $\gamma_2 = E(x_t x_{t-2}) = E[((\rho x_{t-2} +$

$e_{t-1}) + e_t) x_{t-2}] = \rho 2 E(x2_{t-2}) = \rho 2 \sigma_{x2}$ $\gamma_s = E(x_t x_{t-s}) = \rho^s \sigma_x^2 = \rho^s Var(x)$

Autocorrelations between x_t, x_{t-s} :

$$r_s = \frac{\gamma_s}{\gamma_0} = \rho^s$$

Plot of r_s over time = Autocorrelation function(ACF) or correlogram.

For stationary series, ACF should converge to 0:

$$\lim_{s\to\infty} r_s = 0 \text{ if } |p| < 1$$

$\rho > 0 \rightarrow$ direct convergence $\rho < 0 \rightarrow$ dampened oscillatory path around 0.

Partial Autocorrelation (PAC)

In AR(p) processes all x's are correlated even if they don't appear in the regression equation. Example AR(1)

$\gamma_s = \rho^s Var(x_t)$

$r_1 = \dfrac{\gamma_1}{\gamma_0} = \rho; \quad r_2 = \dfrac{\gamma_2}{\gamma_0} = \rho^2 = r_1\rho = r_1^2; \quad r_3 = \dfrac{\gamma_3}{\gamma_0} = \rho^3 = r_2 r_1$

We want to see the direct autocorrelation between x_{t-j} and x_t by controlling for all x's between the two. We construct the demeaned series and form regressions to get the PAC from the ACs.

1^{st} PAC:

$x_t^* = \phi_{11}x_{t-1}^* + e_t$ $\quad\quad$ $\Longrightarrow \varphi_{11} = r_1$

2^{nd} PAC:

$x_t^* = \phi_{21}x_{t-1}^* + \phi_{22}x_{t-2}^* + e_t \quad\quad \Longrightarrow \;\phi_{22} = \dfrac{r_2 - r_1^2}{1 - r_1^2}$

In general, for $s \geq 3$, s^{th} PAC:

$$\phi_{ss} = \frac{r_s - \quad {}_{j=1}\phi_{s-1,j}r_{s-j}}{1 - \sum_{j=1}^{s-1}\phi_{s-1,j}r_j} \quad \mathrm{P}_s$$

and $\varphi_{sj} = \varphi_{s-1,j} - \varphi_{ss}\varphi_{s-1,s-j}$

Example: for $s = 3$,

$$\phi_{33} = \frac{r_3 - (\phi_{21}r_2 + \phi_{22}r_1)}{1 - (\phi_{21}r_1 + \phi_{22}r_2)}$$

Identi cation for an AR(p) process

PACF for $s > p$: $\varphi_{ss} = 0$

Hence AR(1): $\phi_{22} = \dfrac{r_2 - \phi_{11}r}{1 - \phi_{11}r_1} = \dfrac{r_2 - r_1^2}{1 - r_1^2} = \dfrac{r_1^2 - r_1^2}{1 - r_1^2} = 0$

$$\phi_{33} = \frac{r_3 - (\phi_{21}r_2 + \phi 22r_1)}{1 - (\phi_{21}r_1 + \phi 22r_2)}$$

To evaluate it, use the relation $\varphi_{sj} = \varphi_{s-1,j} - \varphi_{ss}\varphi_{s-1,s-j}$:

$\phi_{21} = \phi_{11} - \phi_{22}\phi_{11} = \phi_{11} = r_1$

$\phi_{33} = \dfrac{r_3 - r_1r_2}{1 - r_1^2} = \dfrac{r_1^3 - r_1^3}{1 - r_1^2} = 0$, substitute it to get:

Stability condition for an AR(p) process

$(1 - \rho_1L - \rho_2L^2 - \ldots - \rho_pL^p)x_t = \rho(L)x_t = e_t$

The process is stable if $\rho(z)$ 6= 0 for all z satisfying $|z| \leq 1$, or if the roots of the characteristic polynomial lie outside the unit circle. Then, we can write:

$$x_t = \rho(L)^{-1}e_t = a(L)e_t = \sum_{j=0}^{\infty} a_i e_{t-i}$$

Then we have the usual moment conditions:

$E(x_t) = \mu$ or $0 - finite$

$Var_x = E[x_t^2] = E(\sum_{i=0}^{\infty} a_i e_{t-i})^2,\;\; a_0 = 1,\; e_{t-i}e_{t-j} = 0 \longrightarrow Var_x = \sigma_e^2 \sum_{i=0}^{\infty} a_i^2 = \dfrac{\sigma_e^2}{1-a^2} = \gamma_0$ - nite variance, hence time independent. Covariances

$$\gamma_s = E(x_t x_{t-s}) = E[(e_t + a_1 e_{t-1} + a_2 e_{t-2} + \ldots)(e_{t-s} + a_1 e_{t-s-1} + \ldots)] = \rho^2 E(x_{t-2}^2) = \rho^2 \sigma_x^2 = (a_s + a_1 a_{1+s} + a_2 a_{2+s} + \ldots)\sigma_e^2 = \sigma_e^2 \sum_{i=0}^{\infty} a_i a_{i+s}$$

$$r_1 = \frac{\gamma_1}{\gamma_0} = \frac{a_1 a_2 + a_2 a_3 + a_3 a_4 + \ldots}{(a_1 + a_2 + \ldots)}$$

$$r_s = \frac{\gamma_s}{\gamma_0} = \frac{\sum a_i a_{i+s}}{\sum a_i^2}$$

nite and time independent.

If the process is non-stationary, then there is a unit root, i.e. the polynomial has a root for $z = 1$

$\Longrightarrow \rho(1) = 1$. We can thus factor out the operator and transform the process into a rst-di erence stationary series:

$$\rho(L) = (1 - \rho_1^* L - \rho_2^* L^2 - \ldots - \rho_{p-1}^* L^{p-1})(1 - L)$$

$\Longrightarrow \;(1 - \rho_1^* L - \rho_2^* L^2 - \ldots - \rho_{p-1}^* L^{p-1})\Delta x_t = e_t$ - an AR(p-1) model.

If $\rho^*(L) = (1 - \rho_1^* L - \rho_2^* L^2 - \ldots - \rho_{p-1}^* L^{p-1})$ has all its roots outside the unit circle, Δx_t is stationary: $x_t \sim I(1)$

If $\rho^*(L)$ still has a unit root, we must di erence it further until we obtain a stationary process: $x_t \sim I(d)$

An integrated process= a unit root process.

Unconditional men is still nite but

Variance is time dependent

d) Moving Average(MA) models Another common approach for modelling univariate time series models is the moving average(MA) model:

MA process $x_t = e_t + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q}$, $e = 0$ mean white noise error term. $= (1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q)e_t = \theta(L)e_t$

If $\theta(z)$ 6= 0 for $|z| \leq 1$, the process is invertible, and has an AR(∞) representation:

$\theta(L)^{-1} x_t = \alpha(L) x_t = \sum_{j=0}^{\infty} \alpha_j x_{t-j} = e_t \qquad \Longrightarrow \; x_t - \sum_{j=1}^{\infty} \alpha_j x_{t-j} = e_t$

$$\implies x_t = \sum_{j=1}^{\infty} \alpha_j x_{t-j} + e_t$$

Stability condition for MA(1) process

$x_t = e_t + \theta_1 e_{t-1}$

Invertibility requires $|\theta_1| < 1$

Then the AR representation would be:

$$x_t = (1 + \theta_1 L)e_t \implies (1 + \theta_1 L)^{-1} x_t = e_t \implies x_t = \sum_{i=1}^{\infty} \theta_{1i} x_{t-1} + e_t$$

$E(x_t) = 0$ nite

$$Var(x_t) = \gamma(0) = Var(e_t + \theta_1 e_{t-1}) = (1 + \theta_1^2)\sigma_e^2$$

$$\gamma(1) = \theta_1 E e_{t-1}^2 = \theta_1 \sigma_e^2 \implies r_1 = \frac{\gamma(1)}{\gamma(0)} = \frac{\theta_1}{1 + \theta_1^2}$$ nite

$\gamma(2) = \gamma(3) = 0 \Longrightarrow r_2 = r_3 = 0$, hence autocorrelations' cut o point = lag 1

More generally: AC for MA(q)=0 for lag q.

PAC

$$\phi_{11} = r_1 = \frac{\theta_1}{1 + \theta_1^2}$$

$$\phi_{22} = \frac{r_2 - r_1^2}{1 - r_1^2} = \frac{-\theta_1^2}{(1 + \theta_1^2)(1 + \theta_1^2 - \theta_1)} = -\phi_{11}\left(\frac{\phi_{11}}{1 - \phi_{11}^2}\right)$$

$$\phi_{33} = \frac{-\phi_{22} r_1}{1 - \phi_{21} r_1} = \frac{-\phi_{22}\phi_{11}}{1 - \phi_{21}\phi_{11}} = \frac{-\phi_{22}\phi_{11}}{1 - (\phi_{11}(1 - \phi_{22}))\phi_{11}} = -\phi_{11}\frac{\phi_{22}}{1 - \phi_{22}}$$

For AR

AC depends on the AC coe cient(rho), thus tapers o PAC depends

on r_s or ρ^s, cuts o 0 at s(AR(1): cut o at $L = 1$)

For MA:

AC depends on var of error terms: abrupt cut o

PAC depends on the MA coe cient θ, thus tapers o .

ARMA Process

ARMA(p,q): $\rho(L)x_t = \theta(L)e_t$

$(1 - \rho_1 L - \rho_2 L^2 - \ldots - \rho_p L^p)x_t = (1 + \theta_1 L + \theta_2 L^2 + \ldots + \theta_q L^q)e_t$

If $q = 0 \rightarrow$ pure AR(p) process

If $p = 0 \rightarrow$ pure MA(q) process

If all characteristics roots of $x_t = \sum_{i=1}^{p} \rho_i x_{t-i} + \sum_{i=0}^{q} \theta_i e_{t-i}$ are within the unit circle, then this is an ARMA(p,q) process. If one or more roots lie outside the unit circle, then this is an integrated ARMA(p,d,q) process.

Stability condition for ARMSA(1,1) process

$$x_t = c_1 x_{t-1} + e_t + a_1 e_{t-1}$$

$$(1 - c_1)x_t = (1 + a_1 L)e_t \implies x_t = \left(\frac{1 + a_1 L}{1 - c_1 L}\right) e_t$$

If $|c_1| < 1$, then we can write

$x_t = (1 + a_1 L)(1 + c_1 L + (c_1 L)^2 + \ldots)e_t$ (3.1) $= (1 + c_1 L + (c_1 L)^2 + \ldots + a_1 L + a_1 c_1 L^2 + \ldots)e_t$

(3.2)

$= (1 + (a_1 + c_1)L + c_1(c_1 + a_1)L^2 + c_1^2(c_1 + a_1)L^3 + \ldots)e_t \implies$ an $MA(\infty)$ representation.

(3.3)

$E(x_t) = 0$ nite

$Var(x_t) = \rho(0) = \left(1 + \frac{(a_1 + c_1)^2}{1 - c_1^2}\right)\sigma_e^2$ nite

Covariances nite $\gamma(1) = c_1 Var(x_t) + a_1 \sigma_e^2$ $\gamma(2) = c_1(c_1 Var(x_t) + a_1 \sigma_e^2) = c_1 \gamma(1) \implies \gamma(j) = c_1 \gamma(j-1), j \geq 2$

Autocovariance functions:

$$r_1 = \frac{\gamma(1)}{\gamma(0)}$$

$$r_2 = c_1 r_1$$

$$r_3 = c_1 r_2 = c_1^2 r_1$$

Any stationary time series can be represented with an ARMA model:

AR(p) $\implies$ MA(∞)

MA(p) $\implies$ AR(∞)

3.6 Box-Jenkins ARMA Process

The Box-Jenkins methodology, named after the statisticians George Box and Gwilym Jenkins, applies ARMA Models to nd the best t of a time series to past values of this time series, in order to make forecasts. The model is generally referred to as an ARMA(p,q) model where p and q are non-negative integers that refer to the order of the autoregressive and moving average parts of the model respectively.

Modelling approach

The Box-Jenkins model uses an iterative three-stage modelling approach which is:

1. Model identi cation and model selection: making sure that the variables are stationary, (seasonally di erencing it if necessary), and using plots of the autocorrelation and partial autocorrelation functions of the dependent time series to decide which autoregressive or moving average part should be used in the model.
2. Parameter estimation using computation algorithms to arrive at coe cients which best t the selected ARMA model. The most common methods use maximum likelihood estimation or non-linear least-squares estimation.
3. Model checking by testing whether the estimated model conforms to the speci cations of a stationary univariate process. In particular, the residuals should be independent of each other and constant in mean and variance over time. (Plotting the mean and variance of residuals over time and autocorrelation, partial autocorrelation of the residuals are helpful in identifying misspeci cation.) If the estimation is inadequate, we have to return to step one and attempt to build a better model.

3.6.1 Box-Jenkins model identi cation

Detecting stationarity

The rst step developing a Box-Jenkins model is to determine if the time series is stationary. Stationarity can be accessed from a run sequence plot. The run sequence plot should show constant location and scale. It can also be detected from an autocorrelation plot. Speci cally, non-stationarity is often indicated by an autocorrelation plot with very slow decay. Finally, unit root tests provide a more formal approach to determining the degree of di erencing.

Di erencing to achieve stationarity

Box and Jenkins recommend the di erencing approach to achieve stationarity. However, tting a curve and subtracting the tted values from the original data can also be used in the context of Box-Jenkins models.

Identify p and q

Once stationarity and seasonality have been addressed, the next step is to identify the order(i.e., the p and q) of the autoregressive and moving average terms. These are determined by examining the values of the autocorrelations and the partial autocorrelation s with their corresponding plots as explained below.

Autocorrelation and partial autocorrelation plots

The primary tools for doing this are the autocorrelation plot and the partial autocorrelation plot. The sample autocorrelation plot and the sample partial autocorrelation plot are compared to the theoretical behaviour of these plots when the order is known.

Order of autoregressive process(p)

Speci cally, for an AR(1) process, the sample autocorrelation function should have an exponentially decreasing appearance. However, higher-ordered AR processes are often a mixture of exponentially decreasing and damped

sinusoidal components. For higher-ordered autoregressive processes, the sample autocorrelation needs to be supplemented with a partial autocorrelation plot. The partial autocorrelation of an AR(p) process becomes zero at lag $p + 1$ and greater, so we examine the sample partial autocorrelation function to see if there is evidence of a departure from zero. This is usually determined by placing a 95% con dence interval on the sample partial autocorrelation plot(most software programs that generate sample autocorrelation plots will also plot this

con dence interval). If the software does not generate the con dence band, it is approximately $\pm 2/\sqrt{N}$, with N denoting the sample size.

Order of moving-average process(q)

The autocorrelation function of an MA(q) process becomes zero at lag $q + 1$ and greater, so we examine the sample autocorrelation function to see where it essentially becomes zero. We do this by replacing the 95% con dence interval for the autocorrelation function on the sample autocorrelation plot.

Most software that can generate the autocorrelation plot can also generate this con dence interval. The sample partial autocorrelation function is generally not helpful for identifying the order of the moving average process.

3.6.2 Model Estimation

After identifying the order of the model, the parameters of the model are estimated using the maximum likelihood estimation to determine the AR and MA parameters, as well as all other parameters reported in the study.

he Akaike Information Criteria(AIC) is one of the statistics used to verify the adequacy of the chosen models. In general, the AIC is de ned as: AIC= $2k - 2 * ln(L)$ Where: k is the number of model parameter and ln(L) is the log-likelihood function for the statistical model. Comparatively, models with the smallest AIC have residuals which resembles a white noise process.

Each parameter estimate reports standard error for that particular parameter. Using the parameter estimate and its standard error, a test for statistical signi cance(t-value) are then conducted. For statistically signi cant parameters, the absolute values of the t-ratios are expected to be greater than 1.96 or 2 in order for the parameters to be maintained in the model whereas parameters which are not signi cant are removed from the model.

Furthermore, the estimated AR and MA parameters must also conform to certain boundary condition that is they must lie between 1 and 1. If the AR and MA parameters do not lie within those bounds of stationarity then the parameters of the model are re-estimated or if possible a di erent candidate model is alternatively considered for estimation. All these checks when strictly adhered to would lead to obtaining reliable results from the model.

3.6.3 Diagnostic Checking

A key question inARMA modeling is does the model e ectively describe the persistence? If so, the model residuals should be random or uncorrelated in time and the autocorrelation function (ACF) of residuals should be zero at all lags except lag zero. For sample series, the ACF will not be exactly zero, but should uctuate close to zero. The ACF of the residuals can de examined in two ways. First, the ACF can be scanned to see if any individual coe cients fall outside some speci ed con dence interval around zero. Approximate con dence intervals can be computed. The correlogram of the true residuals (which are unknown) is such that is normally distributed. Besides the randomness of the residuals, we are concerned with the statistical signi cance of the model coe cients. The estimated coe cients should be signi cantly di erent than zero. If not, the model should probably be simpli ed by reducing the model order.

3.7 Forecasting ARMA Processes

The purpose of forecasting is to predict future values of a Time Series based on the data collected to the present. In this section we will discuss a linear function of

$X = (X_n, X_{n-1}, \ldots, X_1)^T$ predicting a future value of X_{n+m} for $m = 1,2,\ldots$

The function

$$f_n(X) = \beta_0 + \beta_1 X_n + \ldots + \beta_n X_1 = \beta_0 + \sum_{i=1}^{n} \beta_i X_{n+1-i}$$

is called best linear predictor(BLP) of X_{n+m} if it minimizes the prediction error.

$$s(\beta) = E[X_{n+m} - f_n(X)]^2,$$

where β is the vector of the coe cients β_i and X is the vector of variables X_{n+1-i}. Since $S(\beta)$ is a quadratic function of β and is bounded below by zero there is at least one value of β that minimizes $S(\beta)$. It satis es the equations

$$\frac{\partial S(\beta)}{\partial \beta_i} = 0, i = 0, 1, \ldots, n$$

Evaluation of the derivatives gives the prediction equations;

$$\frac{\partial S(\beta)}{\partial \beta_0} = E[X_{n+m} - \beta_0 - \textstyle\sum_{i=1}^{n} \beta_i X_{n+1-i}] = 0 \qquad (3.7.1)$$

$$\frac{\partial S(\beta)}{\partial \beta_j} = E[(X_{n+m} - \beta_0 - \textstyle\sum_{i=1}^{n} \beta_i X_{n+1-i})X_{n+1-j}] = 0 \qquad (3.7.2)$$

Assuming that $E(X_t) = \mu$ the rst equation can be written as

$$\mu - \beta_0 - \sum_{i=1}^{n} \beta_i \mu = 0$$

Which gives

$$\beta_0 = \mu\left(1 - \sum_{i=1}^{n}\beta_i\right) = 0$$

The set of equations (3.7.3) gives

$$\begin{aligned}
0 &= E(X_{n+m}X_{n+1-j}) - \beta_0\mu - \sum_{i=1}^{n}\beta_i E(X_{n+m}X_{n+1-j}) \\
&= E(X_{n+m}X_{n+1-j}) - \mu^2(1 - \sum_{i=1}^{n}\beta_i) - \sum_{i=1}^{n}\beta_i E(X_{n+m}X_{n+1-j}) \\
&= \gamma(m-(1-j)) - \sum_{i=1}^{n}\beta_i\gamma(i-j), \qquad j = 1,\ldots,n
\end{aligned}$$

That is we obtain the same set of equations when $E(X_t) = 0$. Hence, we assume further that the Time Series is zero-mean stationary process. Then $\beta_0 = 0$ too

3.8 Multiple Linear Regression

In a regression analysis we study the relationship, called the regression function, between one variable Y, called the dependent variable, and several others X_i, called the independent variables. Regression function also involves a set of unknown parameters. If a regression function is linear in the parameters, we term it a linear regression model. Otherwise, the model is called non-linear. Linear regression models with more than one independent variable are referred to as multiple linear models, as opposed to simple linear models with one independent variable.

3.9 Multiple Regression Model with k independent variables

General Case: We have k variables that we control, or know in advance of outcome, that are used to predict Y, the response (dependent variable). The k independent variables are labelled $X_1, X_2, \ldots, X_k$. The levela of these variables for the i^{th} case are

labelled $X_{1i},\ldots,X_{ki}$. Note that simple linear regression is a special case where $k = 1$, thus the methods used are just basic extensions of what
we have previously done.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \ldots + \beta_k X_{k-1} + \varepsilon_i$$

where the change in mean for Y when variable X_j increases by 1 unit, while holding the k-1 remaining independent variables constant (partial regression coe cient). This is also referred to as the slope of Y with variable X_j holding
the other predictors constant.

Least Squares Fitted(Prediction) Equation

$$\hat{Y}_i = b_0 + b_1 x_{1i} + .. + b_k x_{ki}$$

Coe cient of Multiple Determination

Proportion of variation in Y explained by the regression on the k independent variables.

$$R^2 = r^2_{y.1,\ldots,k} = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

Adjusted-R^2

It is used to compare models with di erent sets of independent variables in terms of predictive

$$Adj - R^2 = r^2_{adj} = 1 - \left(\frac{n-1}{n-k-1}\right)\left(\frac{SSE}{SST}\right) = 1 - \left[(1 - r^2_{Y.1,\ldots,k})\left(\frac{n-1}{n-k-1}\right)\right]$$

capabilities and also penalizes models with unnecessary or redundant predictors.

3.10 Residual Analysis for Multiple Regression

It is very similar to the case for Simple Regression. The only di erence is to plot residuals versus each independent variable. Below is a review of plots and interpolations:

Residuals: $e_t = Y_t - \hat{Y}_t = Y_t - (b_0 + b_1X_{1t} + \ldots + b_kX_{kt})$ Plots:

a. Plot e_i versus $\hat{Y}$ Can be used to check for linear relation, constant variance

   If relation is non-linear, U-shaped pattern appears

   If error variance is non constant, funnel shaped pattern appears

   If assumptions are met, random cloud of point appears

b. Plot of e_i versus X_{ji} for each j Can be used to check for linear relation with respect to X_j

   If relation is nonlinear, U-shaped pattern appears

   If assumptions are met, random cloud of points appear

c. Plot e_i versus i Can be used to check for independence when collected over time

   If errors are dependent, smooth pattern will appear

   If errors are independent, random cloud of points appears

d. Histogram of e_i

   If distribution is normal, histogram of residuals will be mound-shaped, around 0

3.11 F-Test for the Overall Model

F-test for a model is used in testing whether any of the independent variables are linearly associated with Y.

Analysis of Variance

Total Sum of Squares and degrees of freedom(df): SST=$\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ $df_T = n - 1$

Regression Sum of Squares: SSR=$\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2$ $df_R = k$

Error Sum of Squares: SSE=$\sum_{i=1}^{n}(Y_i - \bar{Y}_i)^2$ $df_E = n - k - 1$

Source	df	SS	MS	F
Regression	k	SSR	$= SSR/k$	$F_{obs} = \frac{MSR}{MSE}$
Error	n-k-1	SSE	$= \frac{SSE}{(n-k-1)}$	
Total	n-1	SST		

F-test for Overall Model

$H_0 : \beta_1 = \dots = \beta_k = 0$ (Y is not linearly associated with any of the independent variables)

H_A : Not all $\beta_j = 0$ (At least one of the independent variables is associated with Y)

TS: $F_{obs} = \frac{MSR}{MSE}$

RR: $F_{obs} \geq F_{\alpha,k,n-k-1}$

P-value: Area in the F-distribution to the right of F_{obs}

3.12 Inferences Concerning Individual Regression Coefficients

Used to test or estimate the slope of Y with respect to X_j, after controlling for all other predictor variables. t-test for β_j

$H_0 : \beta_j = \beta_{j0}$

$H_A : \beta_j \neq \beta_{j0}$ TS:

$$t_{obs} = \frac{b_j - \beta_j^0}{S_{bj}}$$

RR: $|t_{obs}| \geq t_{\frac{\alpha}{2}, n-k-1}$

P-value: Twice the area in the t-distribution to the right of $|t_{obs}|$

$(1 - \alpha)100$ % Confidence Interval for β_j

$$b_j \pm t_{\frac{\alpha}{2}, n-k-1} S_{bj}$$

If entire interval is positive, conclude $\beta_j > 0$ (Positive association)

If interval contains 0, conclude(do not reject) $\beta_j = 0$ (No association)

If entire interval is negative, conclude $\beta_j < 0$ (Negative association)

3.13 Testing Portions of the Model

Consider 2 blocks of independent variables:

Block A containing k-q independent variables

Block B containing q independent variables

We wish to test whether the set of independent variables in block B can be dropped from a model containing the set of independent variables. That is, the variables in block B add no predictive value to the variables in block A.

Notation:

SSR(A&B) is the regression sum of squares for model containing both blocks

SSR(A) is the regression sum of squares for model containing A MSE(A&B)

is the error mean square for model containing both blocks k is the

number of predictors in the model containing both blocks q is the number

of predictors in block B

H_0: The β^{0s} in block B are all 0, given that the variables in block A have been included in model

H_A: Not all β^{0s} in block B are 0, given that the variables in block A have been included in model

TS: $$F_{obs} = \frac{\left[\frac{SSR(A\&B)-SSR(A)}{q}\right]}{MSE(A\&B)} = \frac{\left[\frac{SSR(B|A)}{q}\right]}{MSE(A\&B)} = \frac{MSR(B|A)}{MSE(A\&B)}$$

RR: $F_{obs} \geq F_{\alpha,q,n-k-1}$

P-value: Area in F distribution to the right of F_{obs}

Coe cient of Partial Determination

Measures the fraction of variation in Y by X_j, that is not explained by other independent variables.

$$r^2_{r2.1} = \frac{SSR(X_1, X_2) - SSR(X_1)}{SST - SSR(X_1)} = \frac{SSR(X_2|X_1)}{SST - SSR(X_1)}$$

$$r^2_{r3.12} = \frac{SSR(X_1, X_2, X_3) - SSR(X_1, X_2)}{SST - SSR(X_1, X_2)} = \frac{SSR(X_3|X_1, X_2)}{SST - SSR(X_1, X_2)}$$

$$r^2_{rk.12\ldots k-1} = \frac{SSR(X_1,\ldots,X_k) - SSR(X_1,\ldots,X_{k-1})}{SST - SSR(X_1,\ldots,X_{k-1})} = \frac{SSR(X_k|X_1,\ldots,X_{k-1})}{SST - SSR(X_1,\ldots,X_{k-1})}$$

Note that since labels are arbitrary, these can be constructed for any of the independent variables.

3.14 Multicollinearity

Problem: In observational studies and models with polynomial terms, the independent variables may be highly correlated among themselves. Two classes of problems arise. Intuitively, the variables explain the same part of the random variation in Y. This causes the partial regression coe cients to not appear signi cant for individual terms(since they are explaining the same variation). Mathematically, the standard errors of estimates are in ated, causing wider con dence intervals for individual parameters, and smaller t-statistics for testing whether individual partial regression coe cients are 0. Variance In ation Factor(VIF)

$$VIF_j = \frac{1}{1 - r_j^2}$$

where r_j^2 is the coe cient of determination when variable X_j is regressed on the $j-1$ remaining independent variables. There is a VIF for each independent variable. A variable is considered to be problematic if its VIF is 10.0 or larger.

When multicollinearity is present, theory and common sense should be used to choose the best variables to keep in the model.Complex methods such as principal components regression and ridge regression have been developed, we will not pursue them here, as they don't help in terms of the explanation of which independent variables are associated and cause changes in Y.

Chapter 4

ANALYSIS AND DISCUSSION OF RESULTS

4.1 Data Collection

Maize grain yield data of the two farming seasons in Ghana was obtained from the Council for Scienti c and Industrial Research (CSIR) of Ghana. The data is the mean semi-annual grain yield of maize from 1995 to 2014 which is taken as yield data. A second set of data (Flowering data) was obtained for the year 2014 on the variables days to ower, plant height, ear height and eld weight as against yield at Fumesua Research Station. The statistical computing package R is used for modelling the two sets of data.

4.2 Display of Data

For the purposes of the ow of the analysis, the time series data for maize grain yield at Fumesua in the Ashanti Region of Ghana is displayed in Appendix I.

4.3 Time Series Plot of Maize Grain Yield

Figure 4.1 is a time series plot of maize grain yield of the semi-annual maize grain yield from 1995 to 2014.

From Figure 4.1 above the mean of grain yield changes over time which indicates the series is non stationary as a result there is the need to test for stationarity.

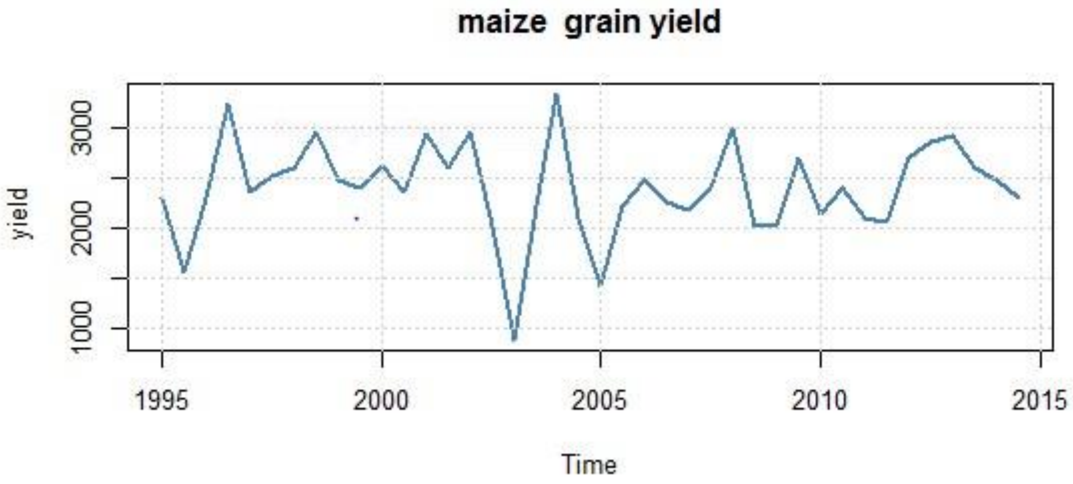


Figure 4.1: Time Series Plot of maize grain yield from 1995 to 2014

4.4 Stationarity Test

In testing for stationarity unit root tests are carried out. Table 4.1 shows the results of a unit root testing procedure of Augmented Dickey-Fuller (ADF).

In the table is a summary of the test statistics. Values for the test type, test statistic, critical value and p-value are shown in the table below. From Table Table

4.1: Unit root test for maize grain yield series

Summary of Test Statistic			
Test type	Test statistic	Critical value	p-value
ADF	-2.6809	-2.86431	0.3062

4.1 the Augmented Dickey-Fuller (ADF) root test statistic (-2.6809) is higher than the critical value (-2.86431) with a p-value of 0.3062, at a 5% signi cance level. Hence the null hypothesis which states that there is a unit root in the series is not rejected. In conclusion the ADF test indicates that the maize grain yield series is non-stationary and has to be di erenced to become stationary.

4.5 Di erencing of Maize Grain Yield Series

After the maize grain yield series was found to be non-stationary through the ADF test, the series is transformed by di erencing. Below is a graph of the di erenced grain yield series. Figure 4.2 is a time series plot of the di erenced grain yield series with yield on the y-axis and time on the x-axis. From Figure

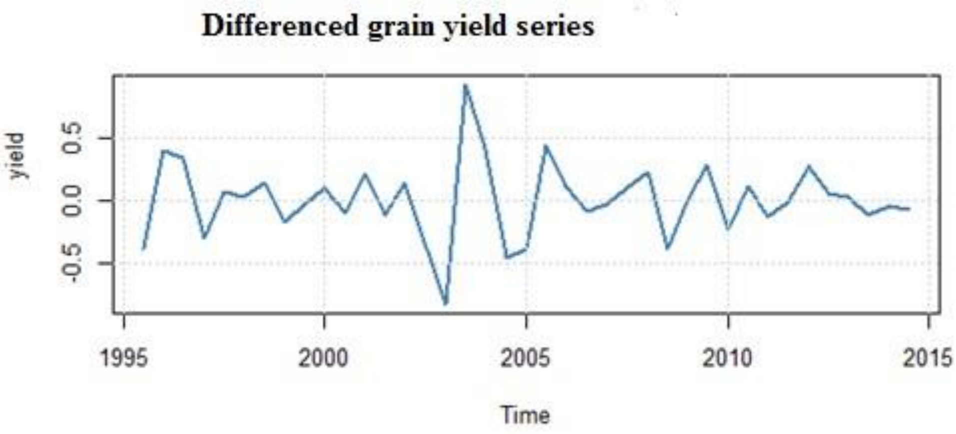


Figure 4.2: Time series plot of the di erenced grain yield series

4.2, at the series is now stationary with a spike in 2004

The ADF test is performed again on the di erenced maize grain yield series. The results obtained are shown in Table 4.2 below.

In the table is a summary of the test statistics. Values for the test type, test statistic, critical value and p-value are shown in the table below. From Table 4.2

Table 4.2: Unit root of the di erenced maize grain yield series

Summary of Test Statistic			
Test type	Test statistic	Critical value	p-value
ADF	-5.212	2.86431	0.01

the ADF test statistic for the di erenced grain yield series is (-5.212) and the critical value is (-2.86431) with p-value of 0.01 at a 5% signi cance level. Hence the null hypothesis which states that there is a unit root in the series is rejected.

In conclusion the di erenced grain yield series is stationary.

Below is a Histogram of the di erenced grain yield series. On the histogram are frequencies and mean of the di erenced series on the y?axis and x-axis respectively.

From Figure 4.3, the histogram is symmetric with heavy tails indicating the

Distribution of di erence yield

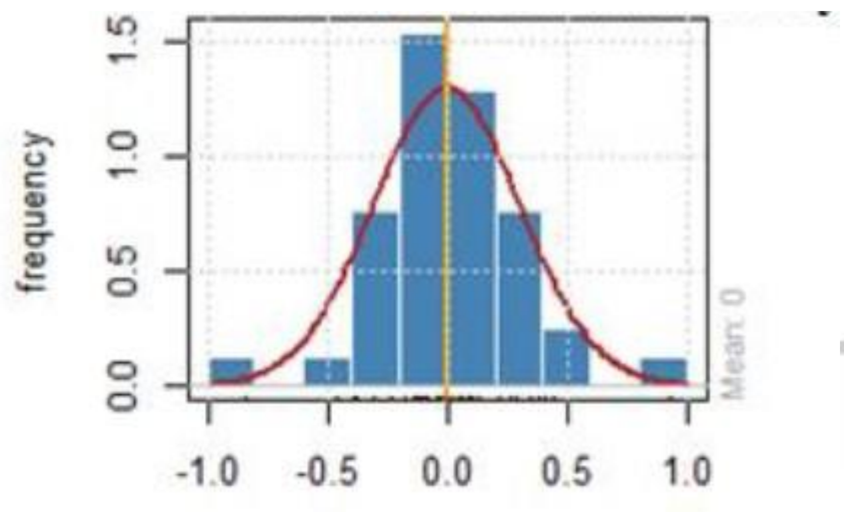


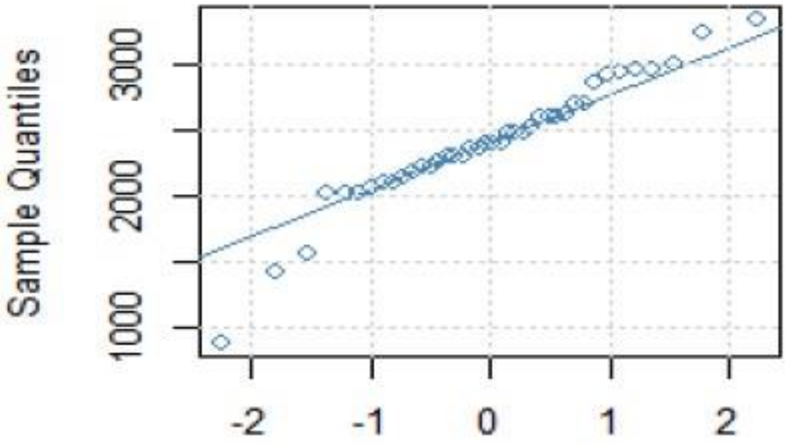
Figure 4.3: Histogram of the di erenced grain yield series

distribution is normal.

Below shows a normal Q-Q plot of the di erenced grain yield series. The graph below is a plot of sample quantiles against theoretical quantiles From Figure 4.4, the normal Q-Q plot the deviations from the straight line are minimal. This indicates normal distribution.

4.6 Determining Order of Dependency of Di erenced yield series

The autocorrelation function(ACF) for the di erenced grain yield series are illustrated in Figure 4.5 below. On the y-axis are autocorrelation functions and on the x-axis are the lags. From Figure 4.5 the Autocorrelation graph showed dependency at lag zero(0) and 2. Below is the Partial Autocorrelation function of Normal Q-Q Plot



Theoretical Quantiles

Figure 4.4: Normal Q-Q plot of the di erenced grain yield series

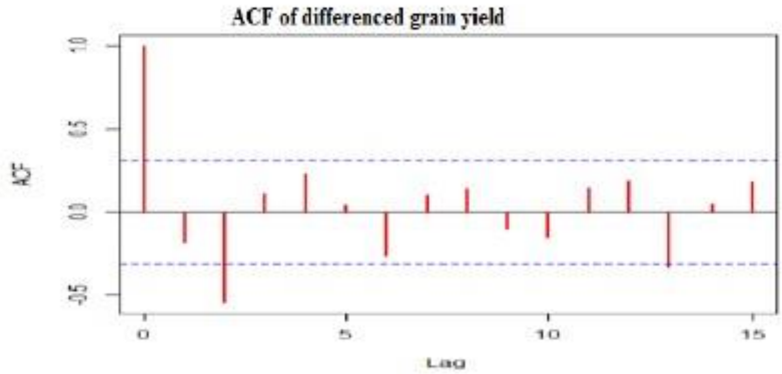


Figure 4.5: Autocorrelation function of di erenced grain yield the di erenced grain

yield. From Figure 4.6 the Partial Autocorrelation function

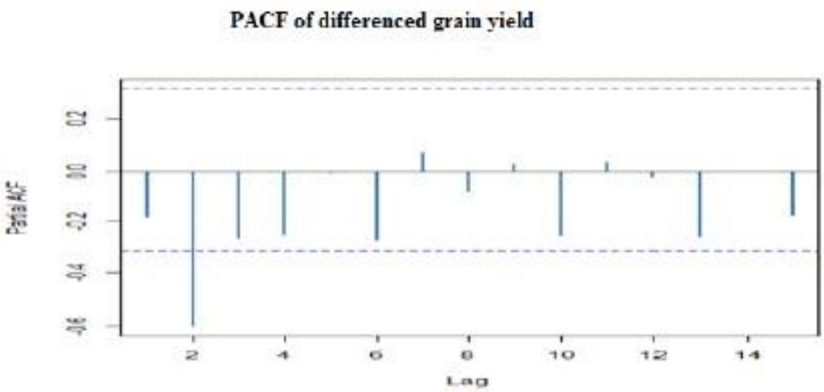


Figure 4.6: Partial Autocorrelation function of di erenced grain yield

Table 4.3: Model selection by AIC with non-zero mean

ARMA(p,q)	AIC
ARMA(1,0)	20.84
ARMA(0,1)	13.55
ARMA(1,1)	15.44
ARMA(2,0)	3.82
ARMA(2,2)	-10.42

showed dependency at lag 2 in the di erenced grain yield series. The ACF and PACF plots above include that ARMA(2,2) is the model for the di erenced grain yield series. Hence the order of the model is (2,2).

4.7 Model Selection and Estimation of Parameters

In this section the Akaike Information Criterion(AIC) is used in selecting the appropriate model. Parameters of the selected ARMA model are also estimated in this section. The ACF and PACF plots above indicates ARMA(2,2) as the model for di erenced grain yield series but other ARMA models [(1,0),(0,1),(1,1) and (0,2)] are included in Akaike Information Criterion(AIC) for comparison. In choosing the modl, the one with minimum AIC is selected. Results of the AIC are shown below in Table 4.3. From Table 4.3, ARMA(2,2) which has the minimum AIC(-10.42) is selected.

Table 4.4 shows results of parameter estimates of ARMA(2,2) model for the di erenced grain yield series. On the table are variable, coe cient, standard error, T-statistics and probability values. From Table 4.4 above all four parameters are signi cant at $\alpha = 0.05$ since the p-values (0.0206, 0.0412, 2.00e-16 and 1.11e-15) are less than α. Therefore ARMA(2,2) is considered as the model suitable for the di erenced grain yield series.

Table 4.4: ARMA(2,2) Model's Parameter Estimates

Variable	Coe cient	Standard Error	T-Statistics	Probability
AR(1)	0.182091	0.078672	2.315	0.0206
AR(2)	-0.301556	0.147734	-2.041	0.0412

MA(1)	-0.820056	0.023489	-24.912	2.00e-16
MA(2)	-0.510013	0.063544	-8.026	1.11e-15
Intercept	0.008489	0.006147	1.381	0.1673

$\sigma^2 = 0.03469$, conditional sum of squares=1.27 AIC=-10.42 $\alpha = 0.05$

4.8 ARMA Model for the Di erenced grain yield series

The ARMA(p,q) model states that the current value of some series r_t depends on its own previous values plus a combination of current and previous values of a white noise error term ε_t. From Table 4.4 above the model for the di erenced maize grain yield series ARMA(2,2) is given by $r_t = 0.1822091r_{t-1} - 0.30155r_{t-2} - 0.820056\varepsilon_{t-1} - 0.510013\varepsilon_{t-2} + \varepsilon_t$

4.9 Model Diagnostics of ARMA(2,2)

Figure 4.7 is a plot of the standardized residuals. On the graph is a plot of values against time From Figure 4.7 the time series plot shows the series is stationary. Since the mean does not change over time.

Below is an ACF plot of the standardized residuals. On the y-axis are autocorrelation functions and on the x-axis are the lags. Figure 4.8 shows no apparent departure from the model assumption showing dependency at lag 2. Figure 4.8 shows no apparent departure from the model assumption showing dependency at lag 2.

Below is a histogram of the standardized residuals. On the graph is a plot of probability against mean. From Figure 4.9 above histogram is symmetric with heavy tails indicating normality. Results of the descriptive statistics shown in

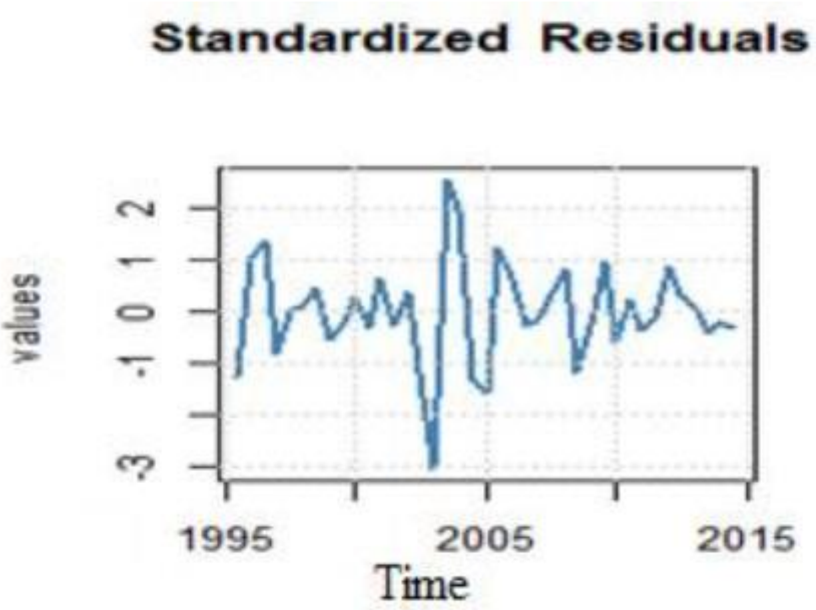


Figure 4.7: Time Plot of standardized residuals

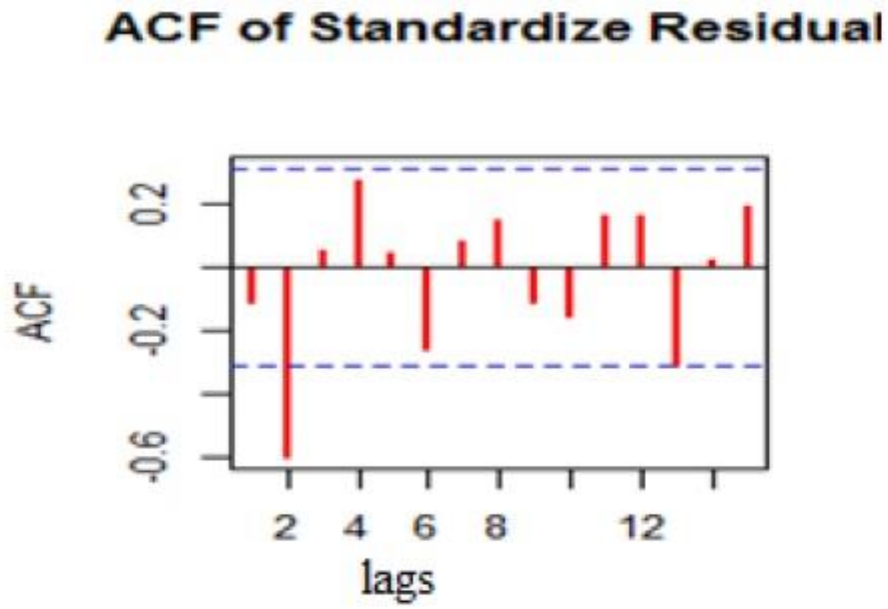


Figure 4.8: ACF of standardized residuals

table 4.5 below con rm that the distribution is normal. From Table 4.5 there is a standard deviation (1.01306) with a general mean (-0.00568). The empirical distribution of residuals shows symmetric with kurtosis (0.25935) and skewness (0.01956). This indicates normality of standardized residuals, is supported by

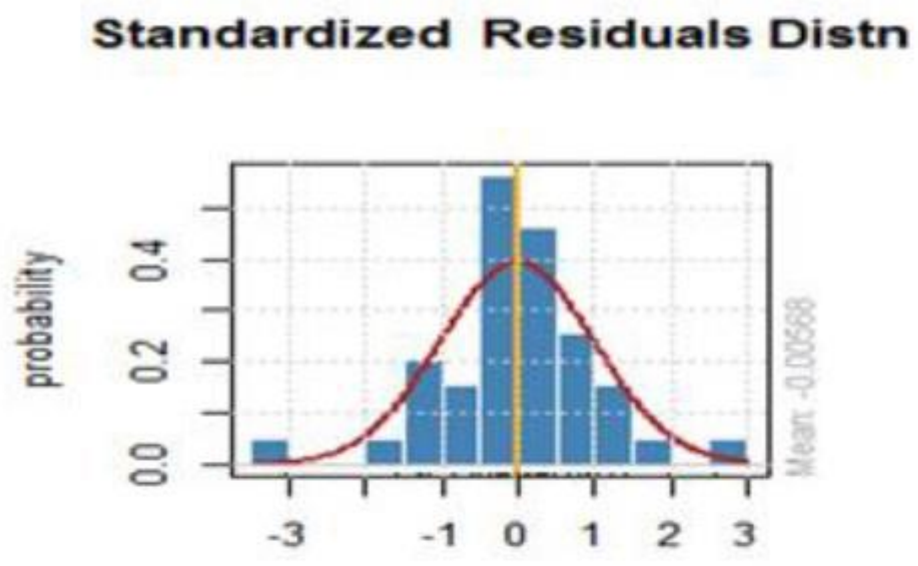


Figure 4.9: Histogram of the standardized residuals

Table 4.5: Descriptive statistics

Statistics	Value	Statistics	Value	
Mean	-0.00568	Standard deviation	1.01306	
Median	-0.15136	Skewness	0.01956	p-value=(0.3459)
Minimum	-3.05210	Kurtosis	0.25935	
Maximum	2.57038	Shapiro-Wilk	0.96855	

Shapiro-Wilk test value of 0.96855 and p-value of 0.3459. These results indicate that the residuals are uncorrelated. Hence the model ts well for the di erenced maize yield.

As a result the ARMA(2,2) model is adequate in describing the conditional mean of the di erenced grain yield series. Therefore the model is suitable and appropriate for the data.

Below is a normal Q-Q plot of the standardized residuals. The graph is a plot of sample quantiles against theoretical quantiles. The normal Q-Q plot

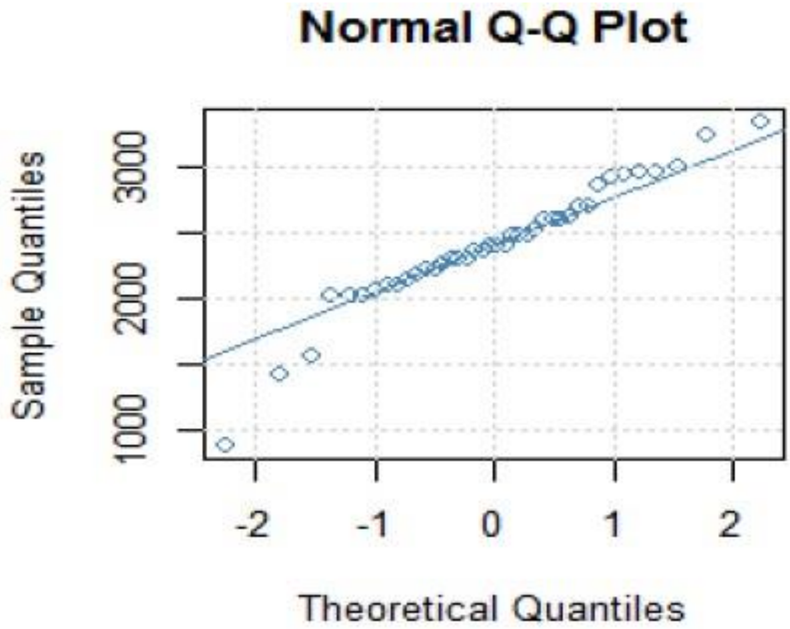


Figure 4.10: Normal Q-Q plot of the standardized residuals

indicates normality wiyh few outliers. In conclusion the above plot quantiles shows normality.

4.10 Forecasting of Maize Grain yield

Figure 4.5 shows forecast for the di erenced maize grain yield series. The yield tends to approach the mean of the series with wider forecast limit. The two curves seen are the upper and lower limit of the 95% con dence interval constructed for the forecast. It can be seen that grain yield returns forecasted lie within the con dence interval indicating that the tted ARMA(2,2) model is appropriate for the maize grain yield data and can be used .

The ARMA(2,2) model for the di erenced grain yield series was also used to make an intra-prediction of di erenced grain yield for 15years by constructing a model each with one-step-ahead prediction of the next observations. The results of the forecast are shown below in Table 4.6. It is seen from Table 4.6 that the grain yields predicted are very close to the actual yield and all their values

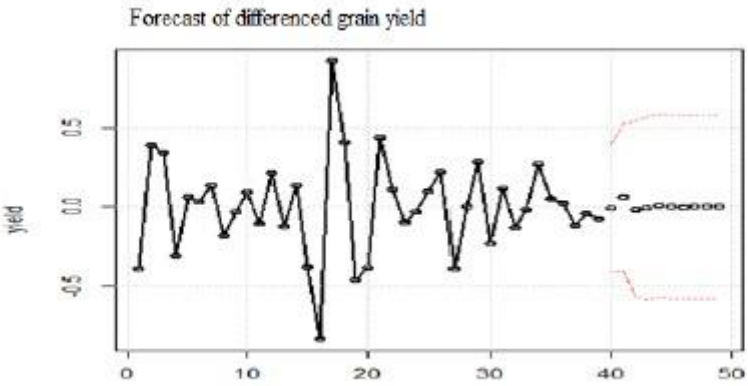


Figure 4.11: Forecast of di erenced grain yield

Table 4.6: Forecasting of 15 years Intra-Predicted yield against Actual yield

Year	Actual yield	Predicted yield
1995	2242.847	2361.87
1996	2198.089	2430.24
1997	2652.609	2731.29
1998	2367.134	2471.03
1999	2459.434	2762.159
2000	2502.057	2762.159
2001	2824.171	3469.45
2002	2839.225	2953.38
2003	2231.288	2281.22
2004	1750.877	1869.23
2005	2870.974	2963.28
2006	2322.534	2460.54
2007	2825.346	2975.36
2008	2215.951	2395.34
2009	2396.228	2556.23

lies within the 95% con dence interval created. Hence the tted ARMA(2,2) model is appropriate for the data.

4.11 Multiple Regression Analysis

In this section a multiple linear model is tted to determine how the factors(plant height, ear height, days to ower and eld weight) a ect the yield of maize.

4.11.1 Display of Regression Data

Below is maize owering data used for the regression modelling. On the table are values for the variables entry, days to ower, plant height, eld weight and grain yield respectively. Table 4.8 shows the Pearson's correlation between all the four

ENTRY	Days to ower	plant height	ear height	Field Weight	Grain yield
1	51.7	202.25	111	5.025	4566.4
2	51.25	187.75	90.25	5.225	4649.7
3	52	203.75	110.5	4.4	3877.2
4	51.25	190	88	3.525	3184.3
5	52	225.25	122.25	5.55	4789.5
6	53.75	199.25	99.25	3.4	2957.8
7	51.75	192.25	110.25	4.2	3748.7
8	51.5	187	93.75	4.575	4171
9	52.25	200.5	129.75	5.225	4502.9
10	51.75	216.5	115.5	5.525	4801.2
11	52	219.75	110.25	5.6	4886.6
12	51.25	207.75	111.75	4.525	4014.2
13	51.75	195	104.25	4.1	3567.4
14	50.809	198.53	96.88	4.6007	4111.7
15	51.75	202.75	98.8	5.35	4698.4
16	51.652	195.45	93.21	3.1656	2718.3
17	52.25	200	102	4.725	3225.6
18	50.5 207.75	107.75	4.225	3806.6	

factors and yield. From Table 4.8 there is a strong positive linear relationship

Table 4.7: Pearson's Correlation between Variables(Factors)

Factors	Yield	Days to ower	Plant Height	Ear Height	Field Weight
Yield	1.000	0.821	0.641	-0.360	0.731
Days to ower	0.821	1.000	0.478	0.545	0.319
Plant Height	0.641	0.478	1.000	0.713	0.657

Ear Height	-0.360	0.545	0.713	1.000	0.424
Field Weight	0.731	0.310	0.656	0.424	1.000

between grain yield and plant height, days to ower and weight of the eld; however there is a weak negative linear relationship between yield and ear height with a correlation coe cient of -0.360. The response variable and the independent variables are highly correlated.

4.11.2 Regression Analysis for Maize Grain yield

The multiple regression model is

Grain yield $=\alpha + \beta_1$Plant height$+\beta_2$Ear height$+\beta_3$Days to ower$+\beta_4$Field weight$+\varepsilon_t$ where α is the intercept, β_1 is the coe cient of Plant height, β_2 is the coe cient of Ear height, β_3 is the coe cient of Days to ower, β_4 is the coe cient of Field weight and ε_t is the error term. Grain yield is the independent variable Table 4.9 below shows results of the regression analysis for the maize grain yield coe cients. Column 1 gives coe cients, column 2 gives estimates, column 3 gives standard errors, column 4 gives values and column 5 gives p-values. From Table 4.8 the

Table 4.8: Regression Analysis for Maize Grain yields

Coe cients	Estimates	Standard Error	T-Value	P-Value
Intercepts	8739.11	4679.54	1.878	0.0463
Plant height	44.31	13.31	3.329	0.0016
Ear height	-18.07	15.50	-1.166	0.2492
Days to ower	185.44	100.64	1.843	0.0007
Field weight	704.39	158.12	4.455	4.87e-05

Multiple R-squared, $R^2 = 0.713$, F(4,49)=13.23$\alpha = 0.05$ P-value=0.0002501 variables(plant height, days to ower and eld weight) are statistically signi cant at 0.05 level of signi cance, while ear height is statistically insigni cant. This means that plant height, days to ower and eld weights are signi cant factors a ecting maize grain yield. The regression gives a coe cient of determination(R-square) value as 0.713, which indicates 71.3 percent of the total variation in the response variable accounted for by the predictor variables.

From the regression equation, grain yield increases by 44.31 kilograms per hectare for a unit change of centimetre in plant height, 185.44 kilogram per hectare for a unit change in days to ower and 704.39 kilogram per hectare for a unit change of kilogram in eld weight when all other variables are held constant. However, a unit of change of a centimetre in ear height causes grain yield to decrease by 18.07 kilogram per hectare.

From Table 4.8, the multiple regression equation for a maize grain yield is given by

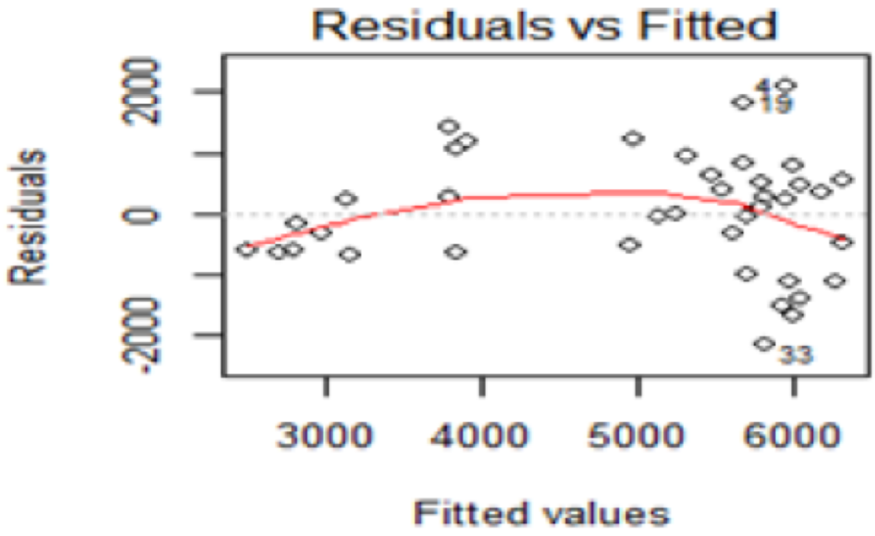
Grain yield=8739.11+44.31 Plant height+185.44 Days to ower+704.39 Field weight.

4.11.3 Diagnostics of the grain yield regression model

The gures below are four individual plots indicating the diagnostics of the grain yield regression model. Figure 4.12 below is a plot of residuals against tted errors It can be seen from Figure 4.6a that the residuals are randomly distributed around the horizontal line(in short dashes on the graph) including a good t for a linear model.

Figure 4.6b shows a plot of the squared root of the standardized residuals against the tted values. From the plot the residuals are randomly distributed and for good model values of the residuals should be more or less randomly distributed. Hence the model is good and t for data.

Figure 4.14 below is a normal Q-Q plot Figure 4.14 is the normal Q-Q plot which shows no departure from the model assumption. Figure 4.15 shows a



plot of residuals against leverage. On the plots are contour lines for the Cook's distance.As such the variable (ear height) which is not statistically signi cant can be dropped from the model and the model can still perform as expected. In Figure 4.12: Plot of residuals errors against tted values

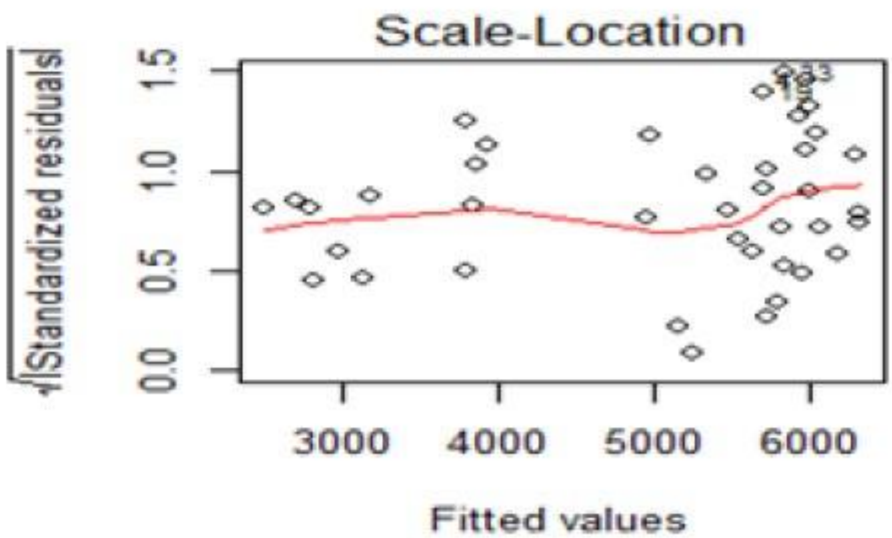


Figure 4.13: Scale Location plot of the standardized residuals

conclusion the model tted is suitable and appropriate for data. Hence it can be used to predict future grain yields.

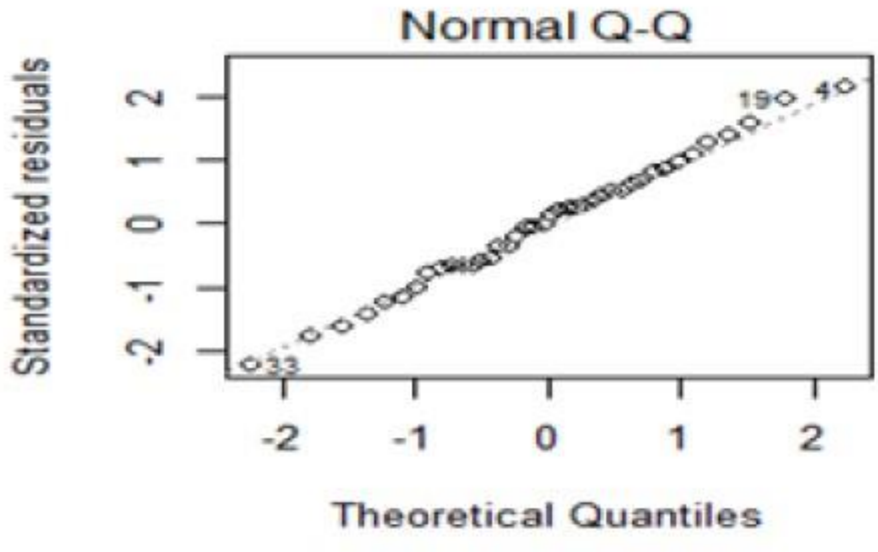


Figure 4.14: Normal Q-Q plot of the standardized residuals

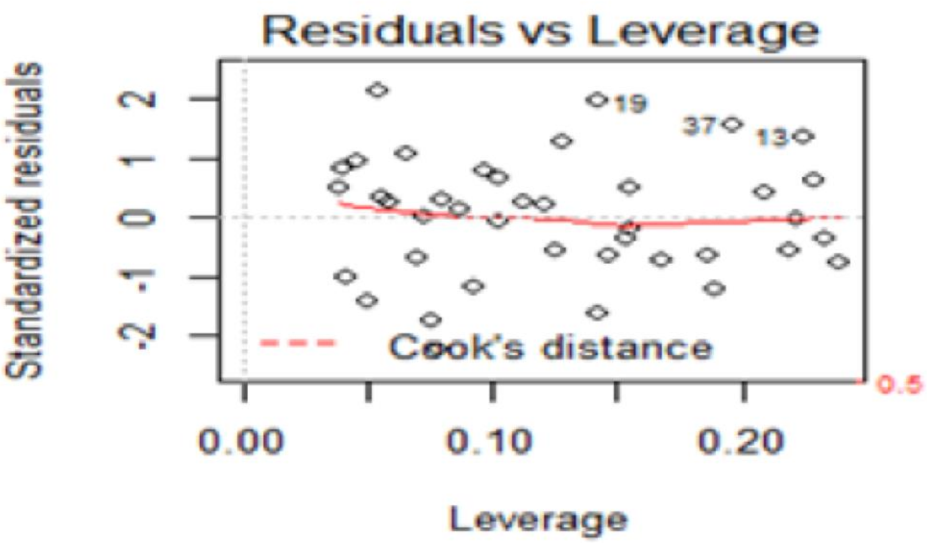


Figure 4.15: Plot of Residuals and leverage

4.12 Discussion of Results

4.12.1 Time Series Model

The time series was found to be non-stationary which resulted from the presence of a unit root in it.

The series became stationary after eliminating the unit root by di erencing the grain yield series.

The Autocorrelation function showed dependency at a lag 0 and 2 whiles the Partial Autocorrelation function showed dependency at lag 2 in the yield series. ARMA(2,2) was found to be the most suitable model for the conditional mean of the di erenced maize grain yield series.The ACF of the standardized residual show no apparent departure from the model assumption.The empirical distribution shows normality which is supported by Shapiro-Wilk normality test value of 0.96855 and a p-value of 0.3459. Hence the model ts well for the maize yield returns. The ARMA(2,2) model for the di erenced maize grain yield series
is given as

$$r_1 = 0.1822091r_{t-1} - 0.301556r_{t-2} - 0.820056\varepsilon_{t-1} - 0.510013_{t-2} + \varepsilon_t$$

The ARMA (2, 2) model tted was used to make intra predictions for grain yield for 15 years and the results showed that the grain yields foretasted are very close to the actual grain yield lies within the 95% con dence interval constructed indicating that the ARMA (2,2) model is appropriate for the maize grain yield data.
The model can be used to make prediction into the future. The results in this study also con rms the ndings of the research article of Anup et al (2006) titled crop yield estimation model for Iowa using remote sensing and surface parameters.

4.12.2 Multiple Regression Model

The multiple regression model:

$Grain\ yield = 8739.11 + 44.31 Plant\ height - 18.07 Ear\ height + 185.44 Days\ to\ flower + 704.39 F$ tted was suitable with about 71.3% of the variation in the yield of grains being explained by the variables plant height, days to ower and the eld weight.

The results also shows that, plants height, the days to ower and eld weight are all statistically signi cant, with p-values 0.0016, 0.0007 and 0.0000487 respectively. Ear height is not statistically signi cant with p-values (0.2492),

hence it can be dropped from the model. Hence the model now becomes:

$$Grainyield = 8739.11 + 44.31 Plantheight + 185.44 Daystoflower + 704.39 Field\ weight$$

Chapter 5

CONCLUSION AND RECOMMENDATION

5.1 Conclusions

Time Series Model

The research examined grain yield data obtained from the Crops Research Institute of Ghana for a period of 1995-2014; the objective of the study was to t an ARMA model for the data and make forecast of future grain yields with it. The observed grain yield series was not stationary at rst but it became stationary after di erencing. The distribution of the di erenced maize yield series showed normality with symmetric histogram which is supported by a Shapiro-Wilk test value of 0.96855 and a p-value of 0.3459.The model that explains the stochastic mechanism of the observed series is ARMA(2, 2). Grain yields forecasted were very close to the actual grain yield and lied within the 95% con dence interval
constructed.

Multiple Regression Model

A multiple regression model was also tted using maize owering data for 2014 in nding out factors a ecting grain yield in maize. The study revealed that the variables plants height, days to ower and the eld weight are signi cant to the model at 0.05 level. This indicates that they are factors that a ect the grain yield as they account for a high percentage (71.3) of the variation in the grain yield of maize.

5.2 Recommendations

It is recommended that the ARMA (2, 2) model should be used by researchers to forecast since the predictions made with it had values very close to the actual grain yield and lies within the 95% con dence interval. Variables that are found to be statistically signi cant in in uencing grain yield in maize such as plants height, the days to ower and eld weight should be given more attention by maize breeders in Ghana when they are coming out with high yielding varieties.

The researcher also recommends that further studies should be conducted on other agricultural factors that might a ect the maize grain yield in the near future.

REFERENCES

Abisai M. (1997)Maize/Bean Intercropping Planting Pattern and Density Studies - University of Zimbabwe

Abuzar, M. R., (2009) "E ect of plant population densities on yield of maize." The Journal of Animal & Plant Sciences 21.4 (2011): 692-695.

Adam R. Wychulis, L., Ann P., Malcolm B. Dockerty, Lila R. Elveback, Oliver H. Beahrs, and Leonard T. Kurland. (1973)"Colorectal cancer in Rochester, Minnesota 1940?1969." Cancer 31, no. 3 (1973): 606-615.

AGRESTE. Available at http://agreste.agriculture.gouv.fr/. Accessed 2013 September 26.

Agrimonde-Terra: Foresight on land use and food (in) security. Available: http://www.agrimonde.org/. Accessed 2013 September 26.

Amon M. E ect of Maize (Zea mays L.) and Pumpkin (Curcubita moschata (Duch. ex Lam.) Duch. ex Poir) Intra-row Intercropping on Yield and Weed Pressure - University of Zimbabwe

Brockwell PJ, Davis RA (2002), The Holt-Winters Algorithm. Introduction to Time Series and Forecasting. Springer. p. 434.

Berita M. (1997)The E ect of Di erent Basal Fertilizers and Seed Generations on Maize Growth and Yield - University of Zimbabwe

Borchers, S., M., Lehrnbecher, T., Dammann, E., Pabst, B., W lk, B., &Koehl, U. (2013)Sequential anti-cytomegalovirus response monitoring may allow prediction of cytomegalovirus reactivation after allogeneic stem cell transplantation. p.113

Brisson N, Gate P, Gouache D, Charmet G, Oury F-X, et al. (2010) ,Why are wheat yields stagnating in Europe? A comprehensive data analysis for France. Field Crops Research 119: 201-212.

Bruce, W.B., Edmeades, G.O and Thomas C. Barker.(2001) "Molecular and physiological approaches to maize improvement for drought tolerance." Journal of experimental botany 53.366 (2002): 13-25.

Calderini DF, Slafer GA (1998), Changes in yield and yield stability in wheat during the 20th century. Field Crops Research 57: 335-347.

Cain M., (1998)Analysis of Di erent Maize Hybrids for Resistance to Grey Leaf Spot Disease - University of Zimbabwe 117

Carberry, B.A, P. S., Bindraban, P. S., Asseng, S., Meinke, H., & Dixon, J. (2012). Eco-e cient agriculture: concepts, challenges, and opportunities.
Crop Science, 50(Supplement_1), S-109.

Cassman KG (1999), Ecological intensi cation of cereal production systems:
Yield potential, soil quality, and precision agriculture. Proceedings of the National Academy of Sciences 96: 5952?5959. doi: 10.1073/pnas.96.11.5952
108

Chambers, T. C.., & Wolanski, B. S (1971). The multiplication of lettuce necrotic yellows virus. Virology, 44(3), 582-591.

Chen, X. P., Cui, Z. L., Vitousek, P. M., Cassman, K. G., Matson, P. A., Bai, J. S.,& Zhang, F. S. (2004). Integrated soil-crop system management for food security. Proceedings of the National Academy of Sciences, 108(16), 6399-6404.

Chifamba, R. (1995)Response to Simulated Drought Conditions of Tall and
Short Maize (Zea mays) Near-isogenic Lines - University of Zimbabwe Chipomho J.(1999) E ect of Maize Canopy Architecture and Nitrogen Fertilizer
Rates on Bean Yield in a Maize+Bean Intercrop - University of Zimbabwe

Dadirai Chihande (1996)The E ect of Seed Size, Simulated Moisture Stress and Planting Depth on Germination and Establishment of Dwarf Maize (Zea mays) Hybrids -, University of Zimbabwe

DiYang, Yu, F., Li, H., Meng, Y (2011). Extraction optimization of Angelica sinensis polysaccharides and its antioxidant activity in vivo. Carbohydrate polymers, 94(1), 114-119.115

111

Eremencia H. (1995), E cacy of Reduced Dosage Rates of Post-emergence Applied Atrazine to Control Broad Leaf Weeds in Maize.

Ereck Chakauya (1999)The Response of Maize Inbred Lines to Seedling Drought Stress - University of Zimbabwe

Delve, E., Zingore, S., GonzÆlez-Estrada, R. J., Herrero, M., Dimes, J. P., & Giller, K. E. (2009). An integrated evaluation of strategies for enhancing productivity and pro tability of resource-constrained smallholder farms in Zimbabwe. Agricultural systems, 101(1), 57-68.

FAO: hunger web page. Available: http://www.fao.org/hunger/en/. Accessed 2013 August 21.

FAOSTAT. Available: http://faostat.fao.org/. Accessed 2013 September 26.

Finger, R (2010) ,Evidence of slowing yield growth?the example of Swiss cereal yields. Food Policy 35: 175-182.

Hammer, G.L., Greg McLean, Scott Chapman, Bangyou Zheng, Al Doherty, Erik van Oosterom, and David Jordan.(2014) "Crop design for speci c adaptation in variable dryland production environments." Crop and Pasture Science 65, no. 7 (2014): 614-626.

Hopwell G. Zheke (1996)The E ect of Season Quality on the Time to Flowering and on Pollen Shed Synchronization with Silk Emergence of Di erent Maize Hybrids (Zea mays) - University of Zimbabwe

Hafner S (2003), Trends in maize, rice, and wheat yields for 188 nations over the past 40 years: a prevalence of linear growth. Agriculture, Ecosystems & Environment 97: 275-283. doi: 10.1016/s0167-8809(03)00019-7

Jackson K. F. (2004) Spatial Distribution of Surface Soil Moisture under a Corn eld.

Jerita T. Chiyan Dry Bean Variety Evaluation for Maize/Bean Intercropping University of Zimbabwe

Kalra N, Chakraborty D, Sharma A, Rai HK, Jolly M, et al. (2008), E ect of increasing temperature on yield of some winter crops in northwest India. Current science-Bangalore 94: 82.

Kimbini, G. (1999)The E ect of Grey Leaf Spot on some Commercial Maize Hybrids in Zimbabwe - University of Zimbabwe

Kpongor D.S (2007)"Spatially explicit modeling of sorghum (Sorghum bicolor (L.) Moench) production on complex terrain of a semi-arid region in Ghana using APSIM." PhD diss., Universit ts-und Landesbibliothek Bonn, 2007.

Kumar-Boken V (2000), Forecasting spring wheat yield using time series analysis. Agronomy Journal 92: 1047-1053. doi: 10.2134/agronj2000.9261047x

Kumar V, Haque CE (n.d.) Forecasting wheat yield in the Canadian Prairies using climatic and satellite data. Available: http://pcag.uwinnipeg.ca/Prairie-Perspectives/PP-Vol01/Kumar-Haque.pdf. Accessed 2013 February 21.109

Kpotor, P.(2012) Evaluation of newly released maize varieties in Ghana for yield and stability under three nitrogen application rates in two agro-ecological zones. Diss. SCHOOL Of Graduate Studies, Kwame Nkrumah University Of Science And Technology, 2012.

Kimball, B. A ,Wall, G. W., &. Amthor, J. S (1973). COTCO2: a cotton growth simulation model for global change. Agricultural and forest meteorology, 70(1), 289-342.

Reid, C. W. J., & Newbold, P. (1971). Forecasting transformed variables, Nottingham University Forecasting Project. Note 6.

Kwesiga, P.L., Sileshi, G., MafongoyaF., &Nkunika, P. (2003). Termite damage to maize grown in agroforestry systems, traditional fallows and monoculture on nitrogen?limited soils in eastern Zambia. Agricultural and Forest Entomology, 7(1), 61-69.

Lizaso, J. I., A. E. Fonseca, and M. E. Westgate. (2007)"Simulating sourcelimited and sink-limited kernel set with CERES-Maize." Crop science 47.5 (2007): 2078-2088.

Ladha JK, Dawe D, Pathak H, Padre AT, Yadav RL, et al. (2003),How extensive are yield declines in long-term rice?wheat experiments in Asia? Field Crops Research 81: 159?180. doi: 10.1016/s0378-4290(02)00219-8

Lin M, Huybers P (2012), Reckoning wheat yield trends. Environmental Research Letters 7: 024016. doi: 10.1088/1748-9326/7/2/024016

Licker R, Kucharik CJ, DorØ T, Lindeman MJ, Makowski D (2013), Climatic impacts of winter wheat yields in Picardy, France and Rostov, Russia: 19732010, Agricultural and Forest Meteorology 176, : 25-37.110

Lobell, D. B., & Burke, M. B. (2010). On the use of statistical models to predict crop yield responses to climate change. Agricultural and Forest Meteorology, 150(11), 1443-1452.

Nelson, Charles R. "The rst-order moving average process: Identi cation, estimation and prediction." Journal of Econometrics 2.2 (1974): 121-141.

Linwood A. Ho man, Scott H. Irwin, Xiaoli L. Etienne, Evelyn V. Colino, Jose Toasa(2014) Forecast Performance of WASDE Price Projections for U.S. Corn

Lobell, David B., Kenneth G. Cassman, and Christopher B. Field. (2007)"Crop yield gaps: their importance, magnitudes, and causes." Annual Review of Environment and Resources 34.1 (2009): 179.112

Leuthold, R. M., Garcia, P., Irwin, S. H., & Yang, L. (1970). The value of public information in commodity futures markets. Journal of Economic Behavior& Organization, 32(4), 559-570.

Madzokere, G. (1998)Grain Yield, Light Interception and Canopy Development in Intercrop Systems of Maize and Sorghum with Cowpea in Natural Region Five - University of Zimbabwe

Makowski, M.L (2013) Comparison of Statistical Models for Analyzing Wheat Yield Time Series. PLoS ONE 8(10): e78615. doi:10.1371/journal.pone.0078615

Marciline F. (1999)E ect of Seed Size and Planting Depth on Germination/Emergence, Plant Performance (Vigour) and Yield of Maize (Zea mays L.) - University of Zimbabwe

Michel, L., &Makowski, D. (2013). Comparison of Statistical Models for Analyzing Wheat Yield Time Series.

Misheck C. (1998)The E ect of Seed Size, Simulated Moisture Stress and Planting Depth on Germination and Establishment of Dwarf Maize (Zea mays) Hybrids - University of Zimbabwe

Mokhtarpour, "Non-destructive estimation of maize leaf area, fresh weight, and dry weight using leaf length and leaf width." Communications in Biometry and Crop Science 5.1 (2010): 19-26.

Moyo, S. (1995). Three decades of agrarian reform in Zimbabwe. Journal of Peasant Studies, 38(3), 493-531.

Mugabe, M. (1995) Evaluation of Maize Streak Virus Resistance in 16 Maize Genotypes -

Naylor, T. H., Seaks, G.T., and Wichern, D.W (1972). "Box-Jenkins methods: An alternative to econometric models." International Statistical Review/Revue Internationale de Statistique (1972): 123-137.

Nelson M. (1995) Evaluation of Four CIMMYT Maize (Zea mays L.) Varieties after Cycles of Recurrent Selection for Changes in Stalk Strength Using the Penetrometer - University of Zimbabwe

Oury F-X, Godin C, Mailliard A, Chassin A, Gardet O, et al. (2012), A study of genetic progress due to selection reveals a negative e ect of climate change on bread wheat yield in France. European Journal of Agronomy 40: 28?38. doi: 10.1016/j.eja.2012.02.007

Peltonen-Sainio P, Jauhiainen L, Laurila IP (2009), Cereal yield trends in northern European conditions: Changes in yield potential and its realisation. Field Crops Research 110: 85?90. doi: 10.1016/j.fcr.2008.07.007

Petris G, Campagnoli P, Petrone S (2009), Dynamic Linear Models with R. Springer. 258 p.

Petris G (2010), An R package for dynamic linear models. Journal of Statistical Software 36: 1-16.

Prado R, West M (2010), Dynamic linear models. Time Series: Modeling, Computation, and Inference. Chapman and Hall/CRC. p. 368.

Prost L, Makowski D, Jeu roy M-H (2008), Comparison of stepwise selection and Bayesian model averaging for yield gap analysis. Ecological Modelling 219: 66-76. doi: 10.1016/j.ecolmodel.2008.07.026

Ray DK, Ramankutty N, Mueller ND, West PC, Foley JA (2012), Recent patterns of crop yield growth and stagnation. Nature Communications 3: 1293. doi: 10.1038/ncomms2296

Reidsma, P, .(2009) "Regional crop modelling in Europe: the impact of climatic conditions and farm characteristics on maize yields." Agricultural Systems 100.1 (2009): 51-60.

Rondanini DP, Gomez NV, Agosti MB, Miralles DJ (2012), Global trends of rapeseed grain yield stability and rapeseed-to-wheat yield ratio in the last four decades. European Journal of Agronomy 37: 56-65. doi: 10.1016/j.eja.2011.10.005

Saseendran, S.A., L. R. Ahuja,D. C. Nielsen,T. J. Trout,L. Ma(2008)Use of crop simulation models to evaluate limited irrigation management options for corn in a semiarid environment AGU Journal

Shin, D. W., et al.(2009) "The role of an advanced land model in seasonal dynamical downscaling for crop model application." Journal of applied meteorology and climatology 45.5 (2009): 686-701.

Soliman, H. M. (2003). Myco ora and Mycotoxins of Cereal Grains in Delta, Egypt. Mycobiology, 31(4), 183-190.

Tanaka G. (1997)Evaluation of Maize (Zea mays L.) Experimental Varieties for Nitrogen Use E ciency - University of Zimbabwe

TojoSoler, C. M. Sentelhas, P. C, Hoogenboom, G (2007)Application of the CSM-CERES-Maize model for planting date evaluation and yield forecasting for maize grown o -season in a subtropical environment,European Journal of Agronomy

Tsungai B. (1995)Resistance Evaluation and Minimum Sample Size Determination for Maize Stalk Borer (Busseola fusca) in Di erent Maize (Zea mays) Genotypes - University of Zimbabwe 116

United Nations (UN) (2011),World Population Prospects: The 2010 Revision

Vercammen. "Marketing and distribution: theory and statistical measurement." Handbook of agricultural economics 1 (1970): 1137-1181.114

Xu, H., X., & Wu, Q. (2006). High quality biodiesel production from a microalga Chlorella protothecoides by heterotrophic growth in fermenters. Journal of biotechnology, 126(4), 499-507.

Ye L, Xiong W, Li Z, Yang P, Wu W, et al. (2013), Climate change impact on China food security in 2050. Agronomy for Sustainable Development 33: 363-374.

Zhang, L. Z., W., Zhang, S. P., Li, B., &Spiertz, J. H. J. (2007). Growth, yield and quality of wheat and cotton in relay strip intercropping systems. Field Crops Research, 103(3), 178-188.

Appendix A

Grain yield	year	Grain yield	Year
2296.47	1st half of 1995	1430.59	1st half of 2005
1556.08	2nd half of 1995	2221.18	2nd half of 2005
2296.47	1st half of 1996	2484.71	1st half of 2006
3237.65	2nd half of 1996	2258.82	2nd half of 2006
2371.76	1st half of 1997	2182.54	1st half of 2007
2522.35	2nd half of 1997	2409.41	2nd half of 2007
2597.65	1st half of 1998	3011.76	1st half of 2008
2974.12	2nd half of 1998	2032.94	2nd half of 2008
2484.72	1st half of 1999	2032.94	1st half of 2009
2409.41	2nd half of 1999	2710.59	2nd half of 2009
2635.29	1st half of 2000	2145.88	1st half of 2010
2371.76	2nd half of 2000	2409.41	2nd half of 2010
2936.47	1st half of 2001	2108.24	1st half of 2011
2597.65	2nd half of 2001	2070.59	2nd half of 2011
2032.94	2nd half of 2002	2861.18	2nd half of 2012
878.43	1st half of 2003	2926.47	1st half of 2013
2221.18	2nd half of 2003	2597.65	2nd half of 2013
335.59	1st half of 2004	2484.71	1st half of 2014
2108.24	2nd half of 2004	2296.47	2nd half of 2014