



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

TESTING ALTERNATIVE METHODS FOR FORECASTING MAIZE PRICES IN SOUTH AFRICA

By

M.H. Labuschagne

A thesis submitted in partial fulfilment of the requirements for the degree of

M.Com (Agricultural Economics)

University of Pretoria

2010

DECLARATION

I declare that the dissertation, which I hereby submit for the degree M.Com Agricultural Economics at the University of Pretoria, is my own work and has not been submitted for a degree at any other tertiary institution.

SIGNITUTARE:

DATE: October 2010

ACKNOWLEDGEMENTS

I would hereby like to thank the following people for their advice, support and encouragement in helping me to complete this study:

- My study leaders Dr. Ferdinand Meyer and Mr. Paul van Staden. The passion that you have for your respective subject fields is truly inspirational.
- My parents, for teaching me the qualities of hard work and perseverance, through example. I dedicate this study to you.
- Werner, for his unfailing love and faith in me.
- Robyn and Karlien, for all the encouragement. I am truly blessed with friends like you.
- To my Creator, for the blessings and grace that He has bestowed on me. To Him be all the glory and praise!

ABSTRACT

Testing alternative methods for forecasting maize prices in South Africa.

by

Marlene Labuschagne

Degree: M.Com
Department: Agricultural Economics
Supervisor: Dr. F.H Meyer¹
Co-supervisor: Mr. P.J van Staden²
Key Concepts: Higher frequency forecasting, maize prices, time series modeling, short term decision making.

Forecasting is an indispensable tool in the decision making process. People, on a daily basis, make forecasts inherently and decide on the basis of that. Formal forecasting methods are commonly applied to substantiate business decision making. Agriculture is no exception to this. There are numerous examples of agricultural projections on production, consumption, prices and trade. A South-African example of these types of projections is the outlook generated by the Bureau for Food and Agricultural Policy (BFAP). The formal forecasting scene in South Africa is, however, predominated by models that generate annual forecasts, with structural models. This is incidentally also the method used by the BFAP research program, mentioned above.

Logically, more frequent forecasts lead to better decision making. Since formal high(er) frequency forecasts are absent from the South African milieu, this study aims to determine whether the generation of these type of forecasts are feasible and meaningful

¹ University of Pretoria, Department of Agricultural Economics, Extension and Rural Development.

² University of Pretoria, Department of Statistics

within the South African agricultural context. Thus, a forecasting evaluation is applied to average monthly yellow maize prices in South Africa. The variable modelled was developed by taking the average monthly price of daily yellow maize closing prices as reported by the Agricultural Marketing Division (AMD) of the South African Futures Exchange of the Johannesburg Stock Exchange (JSE). This series is considered due to the strategic role that yellow maize plays in the South African Agricultural industry.

Several methods were considered as possibilities to be used in generating monthly projections. The study was, however, steered into time series econometrics arena because of data availability and the initial objectives of the study. Relevant variables are mostly available in a quarterly or yearly frequency. This problem is even more severe in macro-economic variables, such as gross domestic product (GDP) and disposable income, which are also used in agricultural modelling. Practical considerations guided the study to models which are parsimonious and easily updated. Time series models adhere to these qualities, resulting in two time series methods being applied. Firstly, univariate models, which generate forecasts based on the stochastic properties of the underlying data generating process, are regarded. Secondly a multivariate model is estimated to determine whether additional fundamental information improves on univariate models. Both univariate and multivariate models produce unrealistic forecasts, since both models forecast a mean level with little or no variation from month to month. If a 12 month forecasting period is considered, the monthly forecast, with no or little variability around the mean is essentially the same as an annual forecast. This adds little value to short(er) term decision making.

The study attempts to understand why the time series method yielded unrealistic results, since it has been applied successfully in the following studies Gjolberg (1997) and Liew, Shittan and Hussain (2000), as far as univariate models are concerned, and Skaggs and Snyders (1992) and Colino (2008) in the case of multivariate models. Possible causes for the mentioned unrealistic results projected with the univariate model are: Non-linear underlying data generating process, volatility clustering in the data generating process and non-normal distributions of average monthly yellow maize prices. Upon further

inspection it is apparent that these are common characteristics of financial time series, which caused the study to venture into the financial econometric field. These characteristics were not initially considered, since it was not the objective to make forecasts to facilitate investment decisions in financial markets, but rather to serve as a tool for grass root decision makers in the agricultural industry. Multivariate models fared poorly in forecasting as well, since there were already problems in the cointegration establishment stage of model development. A general expectation in the market is that world maize prices and the Rand/Dollar exchange rates should be exogenous variables with respect to South African yellow maize prices and therefore, cointegration could only be established with parity prices. This confirms that world maize prices and Rand/Dollar exchange rates are only drivers of local maize prices when South Africa is trading under an import or export parity regime and prices are formed close to or at parity prices. This, however, has little benefit for forecasting.

A theory that goes hand in hand with the concept of commodity forecasting is that of the efficient market hypothesis (EMH). Although it was not initially the objective to test this hypothesis, discussing the effect of an associated futures market impact on price forecasting necessitated a referral to this theory. A question that beckoned was whether time series methods were simply the wrong method to use when attempting to generate monthly yellow maize price projections, or is forecasting of this price series a futile exercise? The latter was found to be true in that the market was determined to be weakly efficient. There are however still forecasting methods that could be applied even if the EMH holds. These however falls under advanced topics of forecasting and were not included in the scope of the study.

Formal high frequency forecasts in agriculture are to a large extent absent from South African agricultural literature and even more so in agricultural economics literature. This study attempts to address this by testing various methods. However none of the methods that were applied were successful in forecasting monthly yellow maize prices. The possible reasons for the performance lack of the models are explained in detail and a number of interesting conclusions regarding the usefulness of forecasting models for

agricultural future markets are made. Apart from many empirical issues, one of the reasons stems from the fact that forecasting is a subject matter that stands with legs in different, and frequently far removed, study fields. This is even more true for forecasting prices of a commodity with an associated derivative market.

This study however provides a synthesis of various subject matters. Themes that are usually conducted in isolation, such as market efficiency, financial econometrics, agricultural practicalities and price formation, were combined into one study. The need for more diligent record keeping is illustrated by the empirical findings and the importance of exposing agricultural economics scholars to a broader scope of econometric methods and concepts is identified. In the South African context this study is, important, since it serves as a starting point for future research in agricultural forecasting in general, specifically for scholars wanting to explore the “higher” frequency milieu. Forecasting research opportunities that might be explored in the future are *inter alia*, neural networks, bootstrapping techniques and the application of time series methods (as applied in this study) for products or commodities with a different price formation process.



TABLE OF CONTENTS

CHAPTER 1	1
INTRODUCTION	1
1.1 INTORODUCTION	1
1.2 PROBLEM STATEMENT	2
1.2.1 General Problem Statement	2
1.2.2 Specific problem statement	3
1.3 OBJECTIVES	5
1.3.1 General objective.....	5
1.3.2 Specific objective.....	5
1.4 METHODOLOGY	6
1.5 DELIMITATIONS.....	7
1.6 OUTLINE.....	7
CHAPTER 2	9
LITERATURE STUDY	9
2.1 INTRODUCTION.....	9
2.2 LITERATURE OVERVIEW.....	9
2.2.1. Quantitative Methods – Technical techniques	10
2.2.2 Qualitative Forecasting Methods	20
2.2.3 Composite forecasting.....	22
2.3 CONCLUSION.....	27
CHAPTER 3	29
UNIVARIATE FORECASTING	29
3.1 INTRODUCTION.....	29
3.2 THE ARIMA MODEL	30
3.2.1 Introduction.....	30
3.2.2 Data Description.....	30
3.2.3 Univariate Properties of the Data.....	33
3.2.4 Identification	39
3.2.4 Estimated Models.....	41



3.2.5 Diagnostic Tests.....	42
3.2.6 Forecasts.....	42
3.3 A NOTE ON EXPONENTIAL SMOOTHING.....	44
3.4 CONCLUSION.....	45
CHAPTER 4	47
MULTIVARIATE MODEL	47
4.1 INTRODUCRION.....	47
4.2 VARIABLE JUSTIFICATION	47
4.3 UNIVARIATE PROPERTIES OF THE DATA	48
4.3.1 Univariate properties of average monthly world prices	49
4.3.2 Univariate properties of average monthly R/\$ exchange rate.....	54
4.3.3 Univariate properties of monthly deliveries of maize	59
4.3.4 Univariate properties of monthly maize utilization.....	63
4.3.5 Concluding remarks on the univariate properties	66
4.4 TEST FOR COINTEGRATION.....	67
4.5 ESTIMATED MODEL	70
4.6 DIAGNOSTIC TESTS.....	72
4.7 FORECASTING.....	75
4.8 CONCLUSION.....	76
CHAPTER 5	78
TIME SERIES MODELS REVISITED	78
5.1 INTRODUCTION	78
5.2 POSSIBLE CAUSES OF POOR FORECASTS.....	78
5.2.1 Introduction.....	78
5.2.2 Univariate properties revisited	79
5.2.3 Cointegration and multivariate model revisited.....	82
5.3 CONCLUSION.....	87
CHAPTER 6	88
FUTURES MARKETS – FORECASTING FRIEND OR FOE?.....	88
6.1 INTRODUCTION.....	88
6.2 TESTING FOR AN EFFICIENT MARKET	89



6.2.1 Data.....	89
6.2.2 Futures prices as unbiased forecast of subsequent cash prices	90
6.2.3 Cointegration – Testing for a long-term relationship.....	93
6.2.4 Determine if future prices lead to subsequent cash prices	99
6.2.5 Is all available information incorporated into the futures price?	100
6.3 CONCLUSION.....	100
CHAPTER 7	103
CONCLUDING REMARKS	103
7.1 OVERVIEW	103
7.3 FINDINGS OF THE STUDY	106
7.4 FUTURE RESEARCH	108
Appendix 1	119
Appendix 2(a)	124
Appendix 2(b)	124

LIST OF TABLES

- Table 2.1: Summery of forecasting methods
- Table 3.1: Augmented Dickey-Fuller Test for Differenced Monthly Maize Prices
- Table 3.2: Phillips-Perron test for Differenced Monthly Maize Prices
- Table 3.3: KPSS test results for Average Monthly Maize Prices
- Table 3.4: Augmented Dickey-Fuller Test for average Monthly Maize Prices
- Table 3.5: Phillips-Perron Tests for Average Monthly Maize Prices
- Table 3.6: KPSS test results for Average Monthly Maize Prices
- Table 3.7: KPSS test results for Average Monthly Maize Prices
- Table 3.8: Summery of Tentative Order Selection Tests of Differenced Maize Prices
- Table 3.9: Diagnostic test results of MA(1) model residuals
- Table 3.10: Actual vs Forecasted observations (12/2007- 10/2008)
- Table 4.1: Augmented Dickey Fuller Test for Differenced Monthly World Maize Prices
- Table 4.2: Phillips-Perron Test for Differenced Monthly World Prices
- Table 4.3: KPSS test results for Average World Prices of Maize
- Table 4.4: Augmented Dickey-Fuller Test for Average Monthly World Maize Prices in levels
- Table 4.5: Phillips-Perron Test for Average Monthly World Maize Prices
- Table 4.6: KPSS test results for Average Monthly World Prices of Maize
- Table 4.7: KPSS test results for Average Monthly World Prices of Maize
- Table 4.8: Augmented Dickey-Fuller Unit Root Test for Differenced Average Monthly R/\$ Exchange Rate
- Table 4.9: Phillips-Perron Unit Root Test for Differenced Average Monthly R/\$ Exchange Rate
- Table 4.10: KPSS test results for differenced Average Monthly R/\$ Exchange Rate
- Table 4.11: Augmented Dickey-Fuller Test for Average Monthly R/\$ Exchange Rate
- Table 4.12: Phillips-Perron Unit Root Tests for Average Monthly R/\$ Exchange Rate
- Table 4.13: KPSS test results for Average Monthly R/\$ Exchange Rate
- Table 4.14: KPSS test results for average Monthly R/\$ Exchange Rate
- Table 4.15: Augmented Dickey-Fuller Test for Differenced Monthly Deliveries of Maize

- Table 4.16: Phillips-Perron Unit Root Tests for Monthly Deliveries of Maize in First Differences
- Table 4.17: KPSS test results for first difference of Monthly Maize Deliveries
- Table 4.18: Augmented Dickey-Fuller Test for Monthly Deliveries of Maize
- Table 4.19: Phillips-Perron Unit Root Tests for Monthly Deliveries of Maize
- Table 4.20: KPSS test results for Monthly Maize Deliveries
- Table 4.21: Augmented Dickey-Fuller Test for Differenced Average Monthly Utilization of Maize
- Table 4.22: Phillips-Perron Unit Root Tests for Differenced Average Monthly Utilization
- Table 4.23: KPSS test results for first difference of Monthly Maize Utilization
- Table 4.24: Augmented Dickey-Fuller Test for Average Monthly Utilization
- Table 4.25: Phillips-Perron Unit Root Tests for Average Monthly Utilization
- Table 4.26: KPSS test results of Monthly Maize Utilization
- Table 4.27: Johansen Cointegrating Rank Test Using Trace (Case 2)
- Table 4.28: Johansen Cointegrating Rank Test Using Trace (Case 3 under restriction)
- Table 4.29: Test of Restriction (Case 2 or Case 3)
- Table 4.30: Parameter Estimates of Vector Error Correction
- Table 4.31: Cross correlations of estimated VECM residuals.
- Table 4.32: Portmantuau test for significance in cross correlations of residuals.
- Table 4.33: Goodness of fit test statistic of individual equations in VECM
- Table 4.34: Autocorrelation, Normality and Homoscedasticity test of individual equations in VECM
- Table 4.35: Actual vs forecasted observations of VECM (2007/12 – 2008/11)
- Table 5.1: Trading regimes in the South African yellow maize market
- Table 5.2: Long-Run Cointegration Equation (Dependant Variable: SA Average Maize Price)
- Table 5.3: Engle-Granger Cointegration Test (Dependant variable: SA_Price)
- Table 5.4: Error Correction Model (Dependant Variable: D[SA Average Maize Price])
- Table 5.5: Engle-Yoo Third Step
- Table 5.6: Calculated Elasticities for World Price and Exchange Rate

- Table 6.1: Simple linear equation of Futures Prices 10 months before contract maturity, regressed on stock prices
- Table 6.2: Simple linear equation of Futures Prices 6 months before contract maturity, regressed on stock prices
- Table 6.3: Simple linear equation of Futures Prices 3 months before contract maturity, regressed on stock prices
- Table 6.4: Augmented Dickey-Fuller Test for Monthly Yellow Maize Futures Prices
- Table 6.5: Phillips-Perron Unit Root Test for Monthly Yellow Maize Futures Prices
- Table 6.6: KPSS test results for Monthly Yellow Maize Futures Prices
- Table 6.7: KPSS test results for Monthly Yellow Maize Futures Prices
- Table 6.8: Augmented Dickey-Fuller Test for first differences of Monthly Yellow Maize Futures Prices
- Table 6.9: Phillips-Perron Unit Root Test for First Differences of Monthly Yellow Maize Futures Prices
- Table 6.10: KPSS test results for first differences of Monthly Yellow Maize Futures Prices
- Table 6.11: Cointegration regression estimated for yellow maize contracts
- Table 6.12: Engle-Granger Cointegration Test
- Table 6.13: Error Correction estimation results for yellow maize contracts

LIST OF FIGURES

- Figure 3.1: Average Monthly Prices of Maize (11/1997-11/2007)
- Figure 3.2: First Difference of Average Monthly Maize Prices (11/1997 – 11/2007)
- Figure 4.1: Average Monthly World Prices of maize in levels (11/1997-11/2007)
- Figure 4.2: First Differences of Average Monthly World Maize Prices (11/1997-11/2007)
- Figure 4.3: Average Monthly R/\$ Exchange Rate in levels (11/1997 – 11/2007)
- Figure 4.4: First Difference of Average Monthly R/\$ Exchange Rate
- Figure 4.5: Monthly Deliveries in levels (05/2000 – 12/2007)
- Figure 4.6: Monthly Deliveries in first differences (05/2000 – 12/2007)
- Figure 4.7: Monthly Utilization in levels (05/2000 – 12/2007)
- Figure 4.8: Monthly Utilization in first differences (05/2000 – 12/2007)
- Figure 6.1: Monthly Futures Prices in Levels (11/1997-11/2007)
- Figure 6.2: First difference of Monthly Futures Prices (11/1997 – 11/2007)

LIST OF ABBREVIATIONS

2SLS:	Two stage least squares
3SLS:	Three stage least squares
ADF test:	Augmented Dickey Fuller Test
ANN:	Artificial Neural Networks
ARCH:	Autoregressive Conditional Heteroscedaticity
ARMA:	Autoregressive Moving Average
ARIMA:	Autoregressive Integrated Moving Average
BFAP:	Bureau for Food and Agricultural Policy
BIC:	Bayesian Information Criterion
BVAR:	Bayesian Vector Autoregressive
CBOT:	Chicago Board of Trade
DW-statistic:	Durbin Watson Statistic
EMH:	Efficient Market Hypothesis
ES:	Exponential Smoothing
ESACF:	Extended Sample Autocorrelation Function
FAO:	Food and Agricultural Association
FAPRI:	Food and Agricultural Policy Research Institute
GARCH:	Generalised Autoregressive Conditional Heteroscedaticity
JB-statistic:	Jarque-Bera Statistic
KPSS test:	Kwiatkowski, Phillips, Schmidt and Shin
MINIC:	Minimum Information Criterion
OECD:	Organization for Economic Co-operation and Development
OLS:	Ordinary Least Squares
PP test:	Phillips – Perron test
RMSE:	Root Mean Square Error
SACF:	Sample autocorrelation function
SAGIS:	South African Grain Information Service
SAFEX:	South African Futures Exchange
SCAN:	Smallest Canonical Correlation method

SPACF: Sample partial autocorrelation function
VAR: Vector Autoregressive
VECM: Vector Error Correction Model

CHAPTER 1

INTRODUCTION

It is better to foresee even without certainty, than not to foresee at all.

~Henri Poincare in The Foundation of Science p129

1.1 INTRODUCTION

As is apparent from the title, this study concerns itself with forecasting of maize prices. The aim of this introduction is to contextualize the environment and needs that led to the venture into this topic. This section will start with a discussion on forecasting in general and move to the specific case as addressed in this study.

The need for forecasting stems from the time lag between decision making and the time at which the outcome of those decisions reach the market (Freebairn, 1975). In addition to this Ferris (2005) notes that decision making at many levels, depends on reasonably accurate assessments of future developments in the commodity scene. Consequently, it can be concluded that people have the need to make informed decisions. Forecasting is one of the weapons in the information arsenal that can be used to substantiate decisions. Agricultural decision making is no exception to the above facts and market players might be even more inclined to base decisions on forecasts due to the extreme uncertainty that this industry is exposed to. Examples of agricultural forecasts are not uncommon. In the United States research programs such as the Food and Agricultural Policy Research Institute (FAPRI) have a set of models that forecast annual figures on prices, production, consumption and trade for various countries on a wide range of commodities and products. Locally, the Bureau for Food and Agricultural Policy (BFAP) has comparable models that projects the components as mentioned above. Other well known examples that follow the same methodology are the Food and Agricultural Organization (FAO) of the United Nations and the Organization for Economic Co-operation and Development (OECD).

The need for price forecasting, in particular, flows out of the fact that in a free market system prices are key economic variables. It reflects the current state of the market and also provides signals to the various role-players in the market which in turn determines their behaviour. It is a well known economic fact that prices have two distinct roles to play in a market. The first is known as the rationing role of price, which ensures that goods are channelled to consumers who value it the most. The second is the allocative role of price which directs resources away from markets with little or no profit opportunities to markets with larger profit opportunities. It is therefore no surprise that price forecasts are highly valued by all market participants. Price forecasts are also not rare, and are projected by all the research programs mentioned in the first paragraph. The one characteristic that is shared by the models mentioned above is that it provides average *annual* projections. These types of models are typically known as structural econometric models and are predominantly suited to facilitate strategic decision making. For the purpose of this study, strategic decision making is understood to be decision making pertaining to situations that are more than one year into the future.

Under general circumstances, it is reasonable to expect that more frequent information flows will result in more effective decision making than less frequent information flows. This highlights the need for the use of higher frequency forecasts in agricultural decision making, especially decision making pertaining to short(er) time periods.

1.2 PROBLEM STATEMENT

1.2.1 General Problem Statement

The structural models alluded to in the introduction, specifically the BFAP model, represent the current state of agricultural price forecasting in South Africa. These models work on the basis that a so-called baseline projection is generated. Various scenarios are then developed and “plugged” into the model to evaluate the different outcomes relative to the baseline scenario. These methods are applied with great success in order to facilitate strategic decision making. The importance of this method is

acknowledged, but it has a number of limitations. The most significant of these is that it is a very comprehensive method with limited application to higher frequency (quarterly, monthly, weekly) data. The main reason is that these models have tedious data requirements and adequate monthly data is rarely available in these frequencies. An example of the application of these models is a study done by BFAP in order to determine the effect of the soccer world cup on meat consumption and prices. The results of this study indicated *marginal* annual increases in price and consumption of red meat. This might be largely attributed to the fact that the impacts were generated with a structural model based on annual data. If a monthly model had been available the effects might have been isolated to a specific month and the associated effects might have been larger. Blake and Clevenger (1984) also identified the importance of monthly forecasts relative to yearly forecasts in a study on alfalfa hay prices. The authors of the article emphasized monthly prices to be important in assisting market participants with marketing and procurement decisions.

The BFAP system of models does not currently include models that can be used to generate higher frequency forecasts (monthly, weekly, daily) which are often requested by stakeholders in a particular industry. For example, role players in the livestock industry have an interest in the cyclical price behaviour of grain and meat prices *within* a year. Annual models typically project average yearly price levels for the years under consideration. Higher frequency would therefore be beneficial to project the pattern of prices within a year or production period. Contrary to strategic decision making, higher frequency models can facilitate short(er)-term decision making, and are therefore worth analysing in greater detail. Examples of short-term decision making that might benefit from such forecasts are production, procurement, marketing and investment decisions. In short, the general problem to be addressed is the absence of monthly price forecasts for South African agricultural products and commodities.

1.2.2 Specific problem statement

In the context of the general problem as presented above, the specific issue that needs to be addressed is to determine if it is possible to forecast average monthly yellow maize

prices. The series that is to be evaluated in this study is average monthly South African yellow maize spot prices as reported by the Agricultural Marketing Division of the Johannesburg Stock Exchange. Although this series is reported daily, prices was converted into average monthly terms. This series is used because of the strategic nature that maize plays in South Africa. A useful monthly forecast generated for maize prices could be applied in grain, livestock and to some extent the horticultural industry. Due to the fact that yellow maize is traded on South African Futures Exchange (SAFEX) it may appear that the essence of this question can be addressed by testing the efficient market hypothesis (EMH). The EMH is an investment theory that asserts that all relevant information is instantaneously and completely reflected by the market price (Fama, 1970). The implication of this is that no relationship exists between historical, current and future prices. Hence, forecasting with historical data is a futile exercise.

The EMH is conventionally applied to daily closing prices on the market. The effect of the aggregation of daily prices, into monthly time series, on the EMH is however unknown. At this point it would suffice to say that the author takes cognisance of this theory but since this main objective is to determine test the possibility of forecasting *monthly* prices of South African yellow maize, time series modelling techniques are applied and evaluated first. Based on the outcome of the modelling techniques the applicability of the EMH on average monthly data will be debated. Pertaining to daily data Wiseman (1999), Ehlers (2004) and Moholwa (2005) tested the EMH on maize in South Africa and found that the market did display weak form efficiency. It can therefore be concluded that it is not possible to forecast price movements from one day to the next. As stated above it is however unclear if this concept will hold with aggregated average monthly data. This study will therefore proceed with an evaluation of various quantitative forecasting methods. Univariate time series methods have been applied successfully by, *inter alia*, Pope, Kramer and Green (1979) on farmland prices, Gjolberg (1997) on hog prices and Liew, Shittan and Hussain (2000) on pepper prices in Malaysia. Success with Multivariate measures were obtained by Colino (2008) and Skaggs and Snyders (1992).

There seem to be no conclusive results as to which method generates the best or most accurate forecasting results in terms of agricultural commodity and product prices. From literature sources reviewed it is apparent that the majority of studies use univariate and multivariate time series models and structured regression based models. Leuthold (1970) found that a structural econometric model outperformed univariate time series models for daily hog prices. In addition to this Guttormsen (1999) could not identify an obvious winner out of a range of five methods which included univariate time series models such as Box-Jenkins modelling and exponential smoothing. Bessler and Brandt (1983) concluded that different models performed well under different forecasting evaluation techniques and recommended the use of composite forecasting to yield the best results. It is expected, however, that results would differ between different data generating processes and that the results of previous studies cannot be extrapolated. The possibility to forecast yellow maize prices should therefore be determined empirically.

1.3 OBJECTIVES

1.3.1 General objective

The objective of this study is to attempt to find a method, which is relatively simple and easily updateable, to generate forecasts for monthly South African monthly yellow maize prices.

1.3.2 Specific objective

In the context of the above the study aims to:

- Identify methods commonly used in agricultural price forecasting by means of a literature review.
- Evaluate the availability of relevant data in a monthly frequency.
- Apply a number of identified techniques, to average monthly yellow maize prices based on findings of the literature study and available data.
- Generate monthly forecasts with the estimated models.

- Evaluate the forecasts based its usefulness and applicability in the short term decision making process.
- Determine the effect of an associated futures market on high frequency price forecasting and the decision making process.
- Determine if the EMH holds when applied to aggregated monthly data.
- Debate the extendibility of the forecasting techniques applied, to other commodities and products.

1.4 METHODOLOGY

The focus of this study will be forecasting with econometric time series techniques. The reason for this is two-fold. Firstly, limited variables are available in a monthly frequency. Structural models are therefore disqualified since they not only need primary variables to include, but also instrumental variables to use in the two stage least squares procedure. Secondly, time series forecasting is objective and has documented methods that can be followed. A practical consideration of time series methods is that models with higher frequency are maintained and updated more frequently. Simplistic and parsimonious models, such as those used in the time series arena, are therefore more user friendly than structural models with tedious data requirements. The attempt to develop simple models for the forecasting of South African maize prices may appear as incredibly ignorant, as it is a complicated market and should be modeled and represented as such. It should be noted, however, that the models developed in this study are not intended to displace complicated large scale structural econometric models. They should rather be used as a supplement to models with more tedious data requirements and which are more difficult to update and develop on a high frequency basis.

The first focus of the study will be on univariate models. These are models in which the observed values of the time series under consideration are used to develop the model (Bowerman, 2003). These models serve as a good starting point, since the data requirements are fairly simple. Two types of univariate models will be regarded, namely, an Autoregressive Integrated Moving Average (ARIMA) model and an Exponential Smoothing (ES) model. The study will continue with the development of a multivariate

model in order to determine if additional fundamental data might improve the accuracy of forecasts. Lastly, the effect of an associated futures market is determined with conventional EMH test as applied by Ehlers (2004).

1.5 DELIMITATIONS

Forecasting is a very broad topic and could be regarded as a subject matter in its own right. A brief overview of various forecasting methods will be discussed in the next chapter. Here it would suffice to say that the focus of this study will be on objective, time series methods. Subjective forecasting has various advantages/application but will only be touched on theoretically. The data generating process of choice is based on yellow maize prices since this, is a commodity that has far reaching implications for agricultural producers and consumers of staples. The sample period chosen for the estimation of the models are November 1997 to October 2007. The *ex post* forecasting period is from November 2007 to October 2008. This is a twelve month period and according to Scheepers (2003), a maize production and marketing cycle is roughly ten months. The forecasting horizon would therefore be able to comfortably cover one production and marketing cycle. It was initially envisaged that November 2008 to October 2009 be used as an *ex ante* forecasting period. The nature of the generated forecasts did however not warrant *ex ante* forecasts.

1.6 OUTLINE

The remainder of this study is organised into six chapters. Chapter Two is a literature study which provides a comprehensive background on various forecasting methods. This also serves as a substantiation of why the applied methods were chosen. Chapter Three discusses the development and related forecasts of the ARIMA model and the ES model as univariate methods used to generate forecasts. Chapter Four discusses multivariate modelling and the predictions generated from it. Chapter Five, which was originally earmarked for a comparison of the forecasts generated, is rather devoted to discussing possible reasons for the poor results in Chapter Three and Four. Chapter Six is dedicated to determine if the yellow maize market can be regarded as efficient, based on the nature of the generated forecasts and formal a EMH tests. Chapter Seven serves

as the concluding section. In this chapter the usefulness of high frequency price forecasting is debated. The limitations of the forecasts, the extendibility to other industries and future research possibilities are also discussed.

CHAPTER 2 LITERATURE STUDY

2.1 INTRODUCTION

In view of the difficulties in accurately forecasting future price movements, the assessment of forecast performance has to be a relative one - measured by how certain types of forecasts perform in relation to others (Bowman and Husain, 2006). Therefore, a study comparing the predictive performance of various forecasting techniques is not unique (see inter alia Bessler and Brandt, 1983 ; Leuthold and Harris, 1985). This chapter, dealing with the applicable literature, will be divided into three main sections. The first section will provide a concise overview of studies that have been done on agricultural price forecasting, and the comparison thereof. Secondly, the advantages, disadvantages and applications of the various methods identified by the literature study will be included and summarized. Lastly, the reasons for the various forecasting methods chosen for this study will be elaborated on. This chapter will serve as a basis for choosing between different forecasting techniques.

The objective of this chapter is therefore to review literature on agricultural forecasting in order to identify which methods or techniques is appropriate to apply for the purposes of this particular study. The methods are grouped according to qualitative methods or quantitative methods used to generate forecasts. Quantitative methods can further be categorized into technical and structural approaches. After appropriate methods have been identified a brief section of this chapter will be devoted to why the other discussed methods are not utilized.

2.2 LITERATURE OVERVIEW

In this section 45 literature sources have been consulted. The objective of this review was to determine what methods/techniques are commonly utilised when conducting price forecasting in agriculture. The section will be organized around the various techniques that were identified. As mentioned earlier, the methods can be categorized

into two main groups namely qualitative and quantitative forecasting. Quantitative forecasting can be broken down further into forecasting based on technical modeling approaches and forecasting based on structural modeling methods. A summary table of all the identified methods is given at the end of this section.

2.2.1. Quantitative Methods – Technical techniques

2.2.1.1 Technical Analysis

Ferris (2005) describes technical analysis as predictions of future values, based on historical patterns and human psychology, indicating how prices are formed in a market place. This source further states that technical analysis emerged due to the inability of fundamental analyses to adequately explain short term (intra-day, daily, weekly, monthly) price fluctuations.

None of the comparative articles reviewed in this literature study included technical analysis as one of the forecasting techniques evaluated. This might be attributed to the fact that technical analyses closely resemble univariate time series methods such as the ARIMA models discussed below. The difference between these two approaches is that technical analysis is concerned with identifying certain data formations that indicate buy or sell signals. In contrast to this, time series models forecast actual prices or price changes between periods based on the underlying mathematical model of the data series. The two methods overlap in that both can be regarded as “a-theoretical” in terms of economic theory and fundamental analysis. This forecasting method is included here due to its wide spread popularity in, amongst others, commodity futures trading. A forecasting overview would not be complete without a referral to this particular method.

2.2.1.2 Exponential Smoothing

Exponential smoothing provides a forecasting method that is most effective when the components (trend and seasonal factors) of a time series may change over time. It is a method that weighs the observed time series values unequally (Bowerman and Hussain, 2005). The main feature of exponential smoothing is that recent observations are weighted more heavily than observations in the distant past. Holton and Keating (2007)

distinguish between three types of exponential smoothing: simple exponential smoothing; Holt's trend corrected exponential smoothing, used for time series with a trend component; and Holt-Winters exponential smoothing which is used for smoothing of seasonal data. The latter method can be used either in additive or multiplicative form, depending on the type of seasonal variation encountered in the data. The additive method is appropriate for constant seasonal variation, whereas the multiplicative method suits data with increasing seasonal variation. The Holt-Winters forecasting model is an extension of simple exponential smoothing and explicitly recognizes these components in a time series (Mendenhall and Sincich, 2003). These types of time series analysis and forecasting methods can be regarded as deterministic, in that no reference is made to the sources or nature of the underlying randomness in the series. Essentially, these models involve extrapolation techniques that have been standard tools in economic and business forecasting for years (Pindyck and Rubinfeld, 1998). Pindyck and Rubinfeld (1998), identify the main disadvantage of smoothing techniques to be that the development of such models, and therefore the forecasts generated using these models, tend to be *ad hoc*. Stated differently, there is no way to determine the 'correct' value of the smoothing parameters. The result is that the choice of these parameters becomes somewhat subjective. The obvious advantages are that the methods are relatively simple and inexpensive to apply. Newbold and Granger (1974) advise that for time series of less than 30 observations there is almost no alternative to employing the exponential smoothing method. A forecaster with limited data might therefore find this method particularly attractive.

Skaggs and Snyders (1992) evaluated various forecasting techniques applied to alfalfa hay prices. It was found that Holt-Winters exponential smoothing produced better one month ahead forecasts than univariate stochastic models and a number of multivariate models. The accuracy of ES decreased relative to the other methods as the forecast horizon increased. In a study where composite forecasting methods were evaluated Gil and Albisu (1993) used Holt Winters exponential smoothing as one of the individual methods in a composite forecasting evaluation on monthly Spanish maize prices. It was concluded that ES produced a better one month ahead forecasts than structural

methods, which is in accordance with Skaggs and Snyders (1992). It was found, however, that ARIMA methods improved on ES methods. In contrast to this, Brandt and Bessler (1983) found that Holt-Winters and simple exponential smoothing had a higher absolute error than a random walk model and a naïve model. Thus, the application of ES to quarterly hog prices in order to generate forecasts did not produce favourable results. It is therefore apparent that there are no conclusive results as far as the accuracy of forecasts generated from exponential smoothing are concerned.

2.2.1.3 ARIMA (Univariate Stochastic)

A rationale for time series analysis is that structural parameters or fundamental factors are often difficult to measure and apply in short-run forecasting (daily, weekly, monthly etc.). While supply and demand are recognized as fundamental forces working in a particular market, the process of monitoring these factors is often slow (Ferris, 2005). This is also acknowledged by Ke-Young (1978), who emphasised that simultaneous equations are indispensable tools in obtaining an understanding of a particular market. It might however not be the best forecasting method since it is difficult to estimate for short-run prediction of monthly prices, because monthly data for all relevant variables is usually hard to obtain. The main disadvantage of the ARIMA model is that it cannot effectively utilize available information on exogenous variables. Tomek and Myers (1993) explain that ARIMA models rely on information contained on the past values of the series being forecasted. This makes it difficult to include other information that might assist in the forecasting process. Ke-Young (1978) claims that the foremost merit of the ARIMA equation to be its simplicity. Only data of the variable in question is needed to develop a forecasting model and generate forecasts. Bowerman, Khoeler and O' Connel (2005) notes that univariate time series forecasting is ideal when conditions are expected to remain unchanged. The implication of this is that this method is not well suited for policy analysis.

Despite limitations mentioned above, the simplicity and data requirements of this method makes it well suited for high frequency forecasting of prices. Newbold and Granger (1974) stipulate that the Box-Jenkins method, which is the method used to construct an

ARIMA model, can usually be employed with a fair degree of confidence for a time series, based on, between 40 and 50 observations. An understanding of univariate time series analysis paves the way utilising multivariate time series analysis. This is discussed in the following section.

Examples of ARIMA models applied to agricultural price forecasting are, *inter alia*, Spreen, Mayer and Simpson (1979), Oliveira, O'Conner and Smith (1979) and Harris and Leuthold (1985). Oliveira *et al* (1979) evaluated ARIMA models against futures forecasts for weekly beef prices. The authors of this article advocate that ARIMA models may not always produce adequate forecasts and advises that forecasters should not rely solely on ARIMA models. Since ARIMA models are devoid of economic theory it is difficult to determine why forecasts are inaccurate when forecasting results are not favourable.

In contrast to this, Harris and Leuthold (1985) compared econometric and time series techniques applied to quarterly cattle prices and found that an ARIMA model had a lower Root Mean Square Error (RMSE) over a four quarter forecasting horizon than structural models. The ARIMA model performed the poorest when turning point accuracy was regarded. It is therefore apparent that there are no conclusive results when the ARIMA method is utilized as a forecasting method. Extensions of these methods are Autoregressive Conditional Heteroscedasticity (ARCH) and Generalized Autoregressive Conditional Heteroscedasticity (GARCH). According to Gujurati (2003) these models are typically used when the volatility of a series needs to be considered. These methods will not be elaborated on for the purposes of this study. It is however, well worth mentioning since price volatility are often discussed and modelled under the broader subject of price forecasting, - see *inter alia* Jordaan (2007) on South African grain prices.

2.2.1.4 VAR and VECM (Multivariate)

Gujurati (2003) explains that the Vector Autoregressive (VAR) methodology superficially resembles simultaneous-equation modelling, in that it represents several endogenous

variables together. Each endogenous variable is explained by its lagged values and the lags of the other endogenous variables in the model. Several problems are associated with this method and are acknowledged by Gujarati (2003) such as that VAR models are a-theoretical in terms of economic theory and fundamental analysis. In addition to this, and due to the emphasis on forecasting, VAR models are less suited for policy analysis. The biggest practical challenge of the VAR method is to choose the appropriate lag length. Degrees of freedom problems can occur when the sample size is small and the chosen lag length is long.

Brand and Bessler (1984) and Skaggs and Snyders (1992) found that a VAR model, indicated no improvement in forecasting performance when compared to the ARIMA model. According to Kaylen (1988) the VAR approach is one approach that enables a forecaster to incorporate additional information that may be available. Kaylen (1998) applied several VAR forecasting models to quarterly hog prices and found that unrestricted VAR models did not forecast as well as VAR models with certain explanatory variables excluded or VAR where the parameters are determined with Bayesian estimation. Zapata and Garcia (1990) explains that there is a myriad of multivariate models, such as Vector Autoregression models (VAR), Bayesian Vector Autoregression (BVAR) models and Vector Error Correction models, available to forecast non-stationary data. In contrast to the references mentioned above, Zapata and Garcia (1990) found that VAR's and BVAR's provide more accurate forecasts than an ARIMA model. Chen and Bessler (1990) conducted a forecasting study on cotton and found that structural models perform better in the midst of a policy change whereas time series (VAR) models have a greater accuracy in forecasting normal price movement.

As mentioned in Chapter One, time series econometrics in univariate (ARIMA) and multivariate form (VAR and VECM) have been applied successfully in studies in order to generate forecasts. Examples of studies that fruitfully used univariate methods to forecast *are inter alia*, Pope *et al* (1979) on farmland prices, Gjolberg (1997) on hog prices in Scandinavian countries and Liew *et al* (2000) on pepper prices in Malaysia. Studies on multivariate models that proved to be successful are Colino (2008) on hog

prices in Iowa and Skaggs and Snyders (1992) on monthly alfalfa hay prices. In a review of forecasting in agriculture Allen (1994) notes that the VAR approach is the best single method to generate short term forecasts. The appeal of these methods lies in the data requirements. For univariate models one time series, of the variable to be forecasted, is needed. VAR or VECM models, which superficially resemble structural models, also have less stringent data requirements than their structural counterparts, in that no data is needed to determine instrumental variables as used in two or three stage least squares estimation.

2.2.1.5 Futures as Forecasts

Many studies have been done on the precision of futures prices as a basis for forecasting subsequent cash prices (see Kastens and Schroeder (1998), Kenyon, Jones and McGuirk (1993) and Zulauf, Irwin and Ropp (1999)). Leuthold and Hartman (1979) note that evidence from various studies suggests that grain futures are generally good estimates of subsequent spot prices. Leuthold and Hartman (1979) further found that the futures prices for hogs do not give a reliable estimate of subsequent cash prices. Kastens, Jones and Schroeder (1998) used the concept of futures market efficiency, along with simple methods that allows for a basis adjustment. This implies that the futures market reflects all available information in the futures prices, and that such a futures price is a good estimate of the cash price of the underlying commodity at delivery time. Just and Rausser (1981) compared futures forecast for various commodities to that of commercial forecasts generated with large-scale econometric sector models. The results indicated that futures market prices are preferable in the forecasting of grain prices, but that econometric methods were more accurate when livestock forecasts were considered. This analysis was conducted with adjusted quarterly data.

It is apparent from the sources reviewed that time series forecasting and structural econometric forecasting were analyzed extensively in the late seventies to the early nineties. Since then the emphasis has moved away from these topics to futures markets as forecasts. The inherent issue that is evaluated when futures prices as forecasts are considered is that of the efficient market hypothesis. Fama (1970) explains the efficient

market hypothesis as follows: the price reflects all available information so that no speculative trading gains that are based on public information can be made. Since contracts on the South-African futures exchange were established in the mid nineties limited research has been done to determine if these markets were efficient. It was found that these markets did display weak form efficiency (Moholwa, 2005 and Wiseman, Darroch and Ortman(1999)). It is, however unsure if these findings can be extended to average monthly data and if South African yellow maize prices would provide good forecasts of subsequent cash prices.

2.2.1.6 Naïve methods

These types of models are usually used as benchmarks, since they represent the cheapest and simplest forecasting procedure. There are a number of variations on this model. The first, “no change”, which means that conditions prevailing in the current period, will hold in the next period. Another variation, “same change,” assumes that the change from the current period to the next will be the same as the change from the previous period to the current (Gellatly, 1979). This method is usually not used exclusively for forecasting purposes. As mentioned above, it is rightly used as a benchmark to determine whether resources, invested in more sophisticated forecasting techniques, are justified.

Kastens *et al* (1998) utilizes two naïve methods, which are compared to forecasts generated from future prices for several livestock and grain commodities. The first approach uses the previous year’s price, for a particular week, to forecast the price for the same week in the current year. The second approach assumed that the cash price returns to its five year average. The study concluded that naïve models are normally less precise than forecasts generated from futures. However, the naïve forecasts in Kastens *et al* (1998) did not deviate too far from the actual cash price.

2.2.2 Quantitative methods – Structural techniques

2.2.2.1 Structural Single Equation Regression Analysis

According to Gujarati (2003) the emphasis of single equation regression models, is estimating or predicting the average value of the dependant variable conditional on fixed values of the explanatory variables. The cause-and-effect relationship, if any, is therefore from the explanatory variables to the dependent variable (Gujarati, 2003). Single equation regression analysis can be regarded as a simple method to estimate statistical relationships. The main advantage of this method is that a very basic statistical background on regression analyses is needed to conduct analysis of this kind. The drawback of this method is that, as mentioned above, it only allows for a one way causal relationship. This type of statistical analysis is usually conducted with the Ordinary Least Square (OLS) method. In terms of forecasting, forecasts on explanatory variables are needed to generate a forecast of the dependent variable. Another drawback of this method is that results obtained with OLS might be spurious when technical econometric concepts such as cointegration are not regarded.

Although articles containing single equation regression models as forecasting models are limited, the concept of single equation regression provides an attractive forecasting alternative for a forecaster with limited time or forecasting experience. As is the case with exponential smoothing, the method is widely applied in business forecasting. The method is intuitive in that a basic understanding of the market to be modeled provides one with the knowledge of which explanatory variables should be incorporated into the equation. In an award winning undergraduate paper Wilson (1993) discussed six simple regression equations used to forecast quarterly cattle prices. It was found that single equation regression models that take seasonality and the most important market factors into account produce better forecasts than forecasts that are simply derived from futures prices.

2.2.2.2 Structural Econometric Models

In many instances, a single equation model and time series analysis do not capture the dynamics of a particular market or industry. This occurs when the dependent variable is determined by explanatory variables, and some of the explanatory variables are, in turn, determined by the dependent variable (Gujurati, 2003). In other words, there is a unidirectional causal relationship between the various variables. In contrast to single equation regression analysis, OLS is not suited to estimate the parameters of the various equations. The equations should be estimated simultaneously, thus capturing the information on all the equations when estimating the parameters. These types of models are estimated by means of two stage least squares (2SLS) or three stage least squares (3SLS). Park and Tomek (1988) argue that structural econometric models that are specified correctly are nevertheless at a disadvantage to time series models in that these models require the forecasting of exogenous variables. Thus, these *ex ante* exogenous forecasts could have large associated errors, causing the actual series being forecasted to be erroneous.

The most prominent examples of structural equation models are those utilised by organisations such as the United States Department of Agriculture. These organizations use structural econometric models to produce an annual outlook on prices, production, consumption and trade for various products and commodities. Other institutions that utilize this method to are, as mentioned in Chapter One, the Food and Agricultural Organization (FAO) of the United Nations, the Organization for Economic Co-operation and Development (OECD) and locally the Bureau for Food and Agricultural Policy (BFAP). These organizations all produce an annual outlook on various commodities that serves as a decision making and policy analysis tool for various role players in the agricultural industry. The tendency of the above organisations to use structural models for policy analysis are confirmed by Chen and Bessler (1990), in that this article advocates the use of structural models where decision making regarding changing policy is concerned. Just and Rausser (1981) found that structural models performed well for livestock forecasts but grains and oilseeds were forecasted better with futures prices. Emmett and Holder (1981) compared the US rice outlook, generated from

structural econometric models, to ARIMA models for average seasonal prices. It was found that the rice outlook had a lower forecasting error, but that this was not significantly lower than forecasts generated with an ARIMA model.

This is a very useful approach due to the fact that it not only enables you to generate forecasts once the model is completed; it also captures the structure of the market. The data requirements for this approach are however stringent and it is doubted if there will be data to estimate such models with associated instrumental variables, used in two and three stage least squares, in a higher frequency context.

2.2.1.3 Neural Networks

Problems associated with time series and statistical methods, as discussed in the previous sections, are that they have certain limiting assumptions like normality, linearity of the parameters etc. A possible solution to the above problems has been identified in artificial neural networks (ANN's). ANN is a soft computing method which is based on the way the human brain functions. According to Kumar (2006), an ANN is able to 'learn'. ANN's are initially set up with weights that are random numbers. As it is trained the network learns to adjust the weights. Kumar (2006), states that the advantages of ANN's are that they are useful when the underlying model is unknown. The disadvantage of this approach is that it is hard to understand how the ANN come to a specific conclusion since it only provides a final outcome, with no explanation of how it came to the specific results. This can be regarded as a relatively new method with most literature pertaining to agriculture and specifically agricultural prices published in the last decade. Examples are Ahmad, Dozier and Roland (2001) and Zou, Xia and Yang (2007). The former study is a forecasting evaluation of egg prices with an ANN. It was found that neural networks provided an improvement in forecasting on conventional statistical techniques. Zou *et al* (2007) forecasted Chinese wheat prices with an ARIMA model and an ANN model. It was found that the ANN model outperformed the ARIMA model and that the best forecasting results were obtained by combining the two methods.

2.2.2 Qualitative Forecasting Methods

2.2.2.1 Judgmental Expert Opinion

Makridakis (1994) acknowledges judgmental forecasts as an indispensable alternative tool for predicting systematic changes from established patterns and/or existing relationships. According to Armstrong and Calopy (1998) judgmental forecasts are invaluable for three reasons. Firstly, industry experts often have knowledge of recent events which effects have not yet been reflected in the time series under consideration. Secondly, industry experts might be able to identify past events that are unlikely to have an influence again. This is vital, since occurrences in the past may distort data, which can cause forecasts to be less accurate. Thirdly, experts might be able to predict future events that have not occurred in the past, and are therefore not represented in historical data. At the same time, it must be noted that this method is prone to biases that will have a negative influence on forecasting accuracy.

A variety of judgmental forecasts are identified by Hanke (2005). These methods are, amongst others, expert opinion, the Delphi method and scenario analysis. For the purposes of this literature overview, scenario analysis will be treated as a separate approach in a later section and this section will focus on expert opinion. Expert opinion is an inexpensive and speedy method by which forecasts can be generated. The limitations of these methods, in the form of personal biases, should however be thoroughly understood. Most advocates of these methods encourage the use of judgmental methods in combination with some or other quantitative method. An additional advantage of expert opinion forecasts is that a wider range of information can be incorporated in the forecasting process and information that could normally not be included in quantitative analysis can also be incorporated. Confirming this, Armstrong and Collopy (1998) note that, in order for judgmental forecasts to be useful, it must incorporate information that is not captured by statistical forecasts. Time series and statistical methods capture historical changes to a large extent. Therefore, judgmental forecasts are most useful when very recent, pending or expected occurrences are to be incorporated (Armstrong and Collopy, 1998).

Bowman and Hussain (2006) compared judgmental forecasts, forecasts generated by futures prices and multivariate econometric models. They concluded that judgmental forecasts outperformed model based forecasts over the short term. Futures as forecasts were the best alternative to longer range forecasting. The forecasting methods mentioned above were applied to various quarterly commodity prices. Bessler and Brandt (1992) found that expert opinion forecasting outperformed future prices as forecasts for quarterly cattle and hog prices.

2.2.2.2 Scenario Analysis

Ferris (2005) regards scenario analysis as most appropriate for long term forecasting of five years and more, due to the fact that structural relationships change over such a long period of time. The Bureau for Food and Agricultural Policy (BFAP) uses scenario planning in conjunction with a set of econometric models to evaluate different possible future scenarios. The process basically entails that the exogenous variables included in the econometric models are analyzed and scenarios are based on how these variables and possible changes may develop. The advantage of this method is that “out of the box” thinking can be incorporated into the model. The disadvantage is that it will be of little use to a forecaster who has a limited understanding of a specific sector and how this sector is affected by changes in the broader economic, political and social environments.

Meyer, Strauss and Funke (2008) used scenario analysis in conjunction with a set of structural econometric methods to analyse the economic feasibility of bio-fuels in South-Africa. The article focuses on a forecasting period of eight years. This reference therefore serves to confirm that scenario analysis are typically suited for long range forecasts and that this type of forecasts should facilitate strategic decision making in the broader economic, social and political environment. It is, further, an example of how scenario analysis can be combined with structural econometric results to cover a broad range of possible futures.

2.2.3 Composite forecasting

The reason for discussing composite forecasting in its own section is that it can be qualitative or quantitative or both. The logic behind composite forecasting is that different individual forecasting methods can contain information which is independent of information contained in other forecasting methods (Brandt and Bessler, 1981). The study found that composite methods perform better than individual methods and that if forecasting performance history can be incorporated into the composite forecasting process, with a weighting scheme, it can improve forecasting accuracy.

Park and Tomek (1988) applied various forecasting methods such as; *inter alia*, an econometric method, ARIMA method, naïve method and an exponential smoothing method to commodity prices. After the individual methods were applied a composite forecast of these methods were generated. Park and Tomek (1998) found that composite forecasting have the greatest advantage when specification error is thought to be potentially important. It is further noted that errors of individual forecasts are often highly correlated, which may reduce the improvements usually associated with composite forecasts. Composite forecasts should be regarded as a risk management tool that may reduce variance and forecast error over a period of time (Park and Tomek, 1988). Kastens, Schroeder and Plain (1998) found that composite forecasts are more accurate than forecasts obtained solely from government agencies or universities (extension forecasts). They conclude that a combination of forecasts obtained from different sources such as government, private sector and extension forecasts, has the potential to improve forecasting accuracy.

Gil and Albisu (1993) found that, composite forecasting is more accurate than quantitative methods such as ARIMA and structural econometric methods for forecasting Spanish maize prices. In accordance with this, Bessler and Brandt(1983) note that combining forecasts has smaller forecasting errors than any individual approach applied to quarterly hog prices.

Table 2.1: Summary of forecasting methods

Technique	Technical Analysis	Exponential Smoothing	ARIMA	VAR and VECM	Single Regression	Structural Econometric	Futures as forecasts	Naïve Methods	ANN	Judgmental	Scenario Analysis	Composite
Description	The analysis of a price time series in order to identify known patterns that indicate buy or sell signals.	Recent observations are weighted more heavily than observations in the distant past. Based on this, data is extrapolated to generate forecasts.	Emphasis on analysing the underlying stochastic properties of a specific time series using Autoregressive and Moving Average terms.	Several endogenous variables are considered together. Each endogenous variable is explained by its lagged values and the lagged values of the other endogenous variables.	Functionally relates independent variables to the dependant variable by estimating an equation with a least-squares technique.	A system of interdependent regression equations. The parameters of the equations are usually estimated simultaneously.	Prices achieved by futures contracts are used as forecasts of subsequent cash prices.	No change or same change approach for value in successive period.	A soft computing method based on the functions of the human brain. Network learns from previous errors or corrects outcomes and makes adjustments accordingly.	Based on opinions and expectations of experts.	The development of possible future scenarios with an elaborate analysis of how the various scenarios might affect the future outcome of a variable or set of variables.	A combination of forecasting techniques. A variety of methods can be used to combine forecasts.
Short Term Forecasting Accuracy	Fair to very good (Patterns only).	Fair to very good.	Very good to excellent.	Very good to excellent.	Fair to very good.	Fair to very good.	Fair to very good.	Poor to fair (One period ahead).	Fair to excellent.	Poor to fair.	Very good to excellent	Fair to excellent.
Medium Term Forecasting Accuracy	Fair (Patterns only).	Poor to good.	Poor to good.	Poor to good.	Poor to good.	Poor to good.	Fair to very good.	Not applicable.	Fair.	Poor to fair.	Fair to very good.	Poor to very good.
Long Term Forecasting Accuracy	Not Applicable.	Poor.	Very poor.	Very poor.	Poor.	Fair.	Poor.	Not applicable.	Fair.	Poor	Fair to very good.	Fair.
Identification of Turning Points	Good.	Poor.	Fair.	Fair.	Good.	Good.	Good.	Not applicable.	Excellent.	Poor to fair.	Poor to fair.	Good.

Technique	Technical Analysis	Exponential Smoothing	ARIMA	VAR and VECM	Single Regression	Structural Econometric	Futures as Forecasts	Naïve methods	ANN	Judgmental	Scenario Analysis	Composite
Typical Application	Identification of buying and selling signals in commodity price and other financial data.	Appropriate for short term forecasting. Are commonly used in business forecasting.	Short term price forecasting. Commonly used in combination with regression analysis (exogenous variables).	Short term price forecasting. Commonly used in combination with regression analysis.	Used in a wide range of price forecasting applications such as policy analysis and supply response analysis.	Used in a wide range of price forecasting applications and policy analysis. Mostly concerned with annual outlook forecasts and cross commodity analysis.	Typically applied when market efficiency and rationale expectations tests are considered.	Benchmark against which other forecasts can be evaluated.	When formal statistical models are unknown. Where statistical assumptions does not hold.	Appropriate for longer range forecasts where structure of the variables are expected to change.	Typically used to evaluate future outcomes that have key uncertainties associated with it.	Where combining forecasting methods can potentially increase accuracy due to a variety of information contained in the various forecasts.
Data Requirements	Graphed data of one year or more.	A minimum of two years of price data required if seasonal factors are present in the data.	At least 40-50 observations. Single series.	At least 40-50 observations. Multiple series.	At least 30 observations.	At least 30 observations.	The prevailing futures price of some subsequent month in the future.	Previous observation.	At least 40-50 observations	Knowledge of industry experts is used. No data requirements.	Development of future scenarios. No historical data required.	As many as the method with the largest data requirement.
Time Requirements	Minimal.	Minimal.	A relatively long time period because model is developed by a process of trial and error.	A relatively long time period because model is developed by a process of trial and error.	Depends on ability to identify relationships.	Time consuming.	Minimal.	Minimal.	A relatively long period.	Depends on the number of experts consulted and how their opinions differ.	Minimal.	Relatively time consuming since individual methods are developed first and a composite are then calculated..

Technique	Technical Analysis	Exponential Smoothing	ARIMA	VAR and VECM	Single Regression	Structural Econometric	Futures as Forecasts	Naïve methods	ANN	Judgmental	Scenario analysis	Composite
Cost of Development (with computer software)	None.	Minimal.	Relatively expensive since econometric computer software is required.	Relatively expensive since econometric computer software is required.	Varies with application.	Expensive.	None.	None.	Expensive.	Minimal.	Minimal	Composite section can usually be forecasted easily when forecast generated with other method are available.
Expertise Required	Experience in price patterns and an understanding of how these patterns affect market sentiment.	Basic knowledge of business statistics.	Thorough knowledge of time series econometrics is required.	Thorough knowledge and experience in time series econometrics is required.	Basic statistical knowledge with a focus on regression analysis.	Extensive knowledge in econometrics with a focus on simultaneous equations and structural systems.	Understanding of commodity markets and related aspects.	None.	Extensive knowledge of forecasting and soft computing methods.	Thorough knowledge of industry and related aspects.	Thorough understanding of key drivers and uncertainties. Good insight into possible future outcomes.	Comprehensive knowledge of forecasting as a discipline.
Main Advantage	Very short term price movements can be predicted without taking fundamental factors into account.	Simple and easy to apply.	Appropriate when data on related time series are limited.	Information on related time series can be included.	Logical and intuitive.	Captures simultaneous two way relationships between variables.	Unbiased forecasts based on market perception. Readily available in general media and publications.	Can be used as a benchmark against which other forecasts can be tested.	Can accommodate models that does not conform to statistical assumptions such as normality and linearity.	Can accommodate factors not captured by quantitative methods.	Can accommodate extraordinary factors not captured by quantitative methods or historical data.	Information of various forecasting methods is incorporated into one process.

Technique	Technical Analysis	Exponential Smoothing	ARIMA	VAR and VECM	Single equation regression	Structural Econometric	Futures as Forecasts	Naïve methods	ANN	Judgmental	Scenario Analysis	Composite
Main Disadvantage	Does not forecast actual prices but rather buy and sell signals used by investors and speculators.	Forecast generating procedure is ad hoc since the determination of smoothing parameters is somewhat arbitrary.	Cannot capture information other than that of the particular time series.	Using this method in an unrestricted form can result in degrees of freedom problems.	Assumes a one way causal relationship between variables which may not be the case.	Difficult and costly to develop. Tedious data requirements.	No apparent disadvantage. Associated futures markets do however not always exist.	Does not incorporate any information, except the price in the previous period, into the forecasting process.	System functions like a black box. The reasons for a particular outcome are not apparent.	Prone to personal bias.	Limited to the knowledge and beliefs of the scenario developer.	Adds an additional step to a process that might already be time consuming.
Examples	No application in scientific literature. Used extensively in practice by market traders and analysts.	Relatively accurate method applied by Skaggs and Snyders (1992) in forecasting alfalfa hay prices.	Sprien et al (1979) on slaughter cow prices.	Kaylen (1988) forecast of quarterly hog prices in USA. Zapata (1990) forecast of monthly cattle prices in USA.	Harris and Leuthold (1985) forecast of US livestock prices.	BFAP annual baseline. Arzac (1979) Quarterly models of livestock and feed grain markets.	Used to in comparison studies and to test EMH. See inter alia Moholwa (2005).	Mostly contained in literature that compared forecasting techniques such as Guttormsen (1999).	Ahmad (2001) on egg prices and Zou (2007) on Chinese grain prices.	Mostly contained in literature that used judgemental methods in combination with quantitative methods as in Bessler and Brandt (1983).	BFAP (2009) on scenarios for the red meat industry pertaining to the soccer world cup.	Bessler and Brandt (1983) on US hog prices.

Adapted from Roche, J.R (1995)

2.3 CONCLUSION

It is acknowledged that the set of forecasting procedures discussed above by no means exhausts the various methods of forecasting methods available. A basic comprehension of all the techniques included in time series modelling and forecasting is a discipline in itself and all the methods could therefore not be discussed in this study.

After the above literature was reviewed, twelve methods were identified as being recurrently applied to agricultural price forecasting. These methods can basically be categorized as quantitative methods and qualitative methods. For the purposes of this study, qualitative methods will not be elaborated on further. This is because there are no published forecasts for the maize industry that were generated by judgmental methods. Experts are reluctant to give forecasts for South African commodities and products. If common literature such as the *Farmers weekly* and *Landbou weekblad* are regarded, it is apparent that experts will at most disclose what direction they anticipate prices to go. The lack of local judgmental forecasts, therefore, forces this study in a quantitative direction. Quantitative methods are favoured in most published literature since models used in these methods are widely applied in practice. The measurement statistics used to evaluate forecasts are established and common procedure. They therefore provide an objective measure of various forecasts generated.

In the above paragraph, forecasting was categorised based on the general groups of qualitative forecasting and quantitative forecasting. It is now important to move from the general to the specific. Ten of the above forecasting techniques can be regarded as quantitative. These ten methods are naïve forecasting, futures as forecasts, technical analysis, exponential smoothing, Box-Jenkins univariate time series forecasting, multivariate time series forecasting, forecasting by means of single equation regression, structural econometric model forecasting, composite forecasting and artificial neural networks. Due to the fact that the forecasts that are generated are monthly and can therefore be regarded as “high frequency,” technical time series models are preferred

since it has less stringent data requirements. The focus of the following two chapters will therefore be on time series methods, to generate forecasts monthly forecasts.

CHAPTER 3

UNIVARIATE FORECASTING

3.1 INTRODUCTION

The preceding chapter provided an overview of the various methods that can be used to generate forecasts. One of the methods identified was the univariate Box-Jenkins procedure as successfully applied by Gjolberg (1997), Liew *et al* (2000) and Colino (2008) as indicated on p.13 of Chapter Two. Consequently the first forecasting procedure tested is univariate models. The procedure, results and forecasts are represented and discussed below.

Univariate models are models that forecasts future values of a time series based solely on the past values of that particular time series (Bowerman 2003 p.11). In this chapter two methods will be considered, namely the Box Jenkins ARIMA method and exponential smoothing. It is important to note that they are in fact two sides of the same coin. Even though the two methods are based on different assumptions and methods, the various types of exponential smoothing can essentially be regarded as special types of ARIMA models. This will be discussed later in greater detail. The reader might wonder why both methods are included if they yield the same outcome. The reason is that exponential smoothing is a simple and intuitive method, used frequently in business statistics and in several agricultural price forecasting articles such as Skaggs and Snyders (1992), Guttormsen (1999), Gil(1993) and Brandt and Bessler (1983). At the end of this chapter, the question of which of the above mentioned methods is more appropriate in modeling agricultural prices, will be discussed.

3.2 THE ARIMA MODEL

3.2.1 Introduction

A generic representation of an ARIMA model is symbolized as follows:

$$\text{ARIMA}(p,d,q)$$

with p : The number of Autoregressive terms

q : The number of Moving Average Terms

d : The order of integration

The Box-Jenkins methodology is an iterative four step process used to estimate ARMA models and generate forecasts with it. The Box-Jenkins methodology should be applied to a stationary time series. Gujarati (2003) defines a weakly stationary process as a stochastic process with a mean and variance that is constant over time and the covariance is time invariant. Although the method under discussion requires stationarity, it is common for economic time series to be non-stationary. If the series under consideration is non-stationary, the order of integration, represented by d above, should first be established. This is done by testing for a unit root or unit roots. The non-stationary series can be made stationary by differencing the series in accordance to the number of unit roots (level of integration) identified. Subsequent to the stationarity transformation, in the form of differencing, the four steps of the Box–Jenkins methodology should be applied sequentially as follows:

1. Identification: Find the appropriate levels of p and q .
2. Estimation of parameters: Estimate parameters for the autoregressive and moving average terms with maximum likelihood estimation.
3. Verification of parameters with diagnostic tests.
4. Forecasting.

3.2.2 Data Description

When constructing an ARIMA model the model developer has the option of transforming the original series into logarithms. It is common practice for financial time series to be transformed into logarithms before the identification stage as was done by, amongst

others, Granger (1992). This transformation can be done in order to stabilize the variance and because of the following:

$$\Delta[\ln(\text{price}_t)] \approx \% \Delta[\text{price}_t]$$

with: $\ln(\text{price}_t)$: average monthly prices in logarithms

$\% \Delta [\text{price}_t]$: percentage change in price

In layman's terms, this statement basically says that if the difference of a series in logarithms is taken it approximates a percentage change in price between two consecutive periods. The popularity of this transformation can be attributed to the fact that developers might be more interested in forecasting the percentage price change, or returns, from one period to the next, than the actual price level. It was decided not to transform the price range due to two prominent reasons. Firstly, the objective of the study is to develop simplistic models that can be used to forecast price. The correct transformation of the variables complicates the model unnecessarily. The transformation of logarithmic forecasts back into levels by simply taking the exponents, leads to biased forecasts (Mayr and Ulbright, 2007). Secondly, it is envisaged that short term forecasts should be usable for all market participants. Working in logs would by implication model returns rather than prices. Forecasted returns would only benefit investment decisions for those involved in maize trading on the financial market.

As mentioned above, it is imperative to apply the Box-Jenkins procedure to a stationary time series. The Augmented Dickey-Fuller (ADF) test, Phillips-Peron (PP) and Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test were conducted to determine the univariate properties of the data in levels and differenced once.

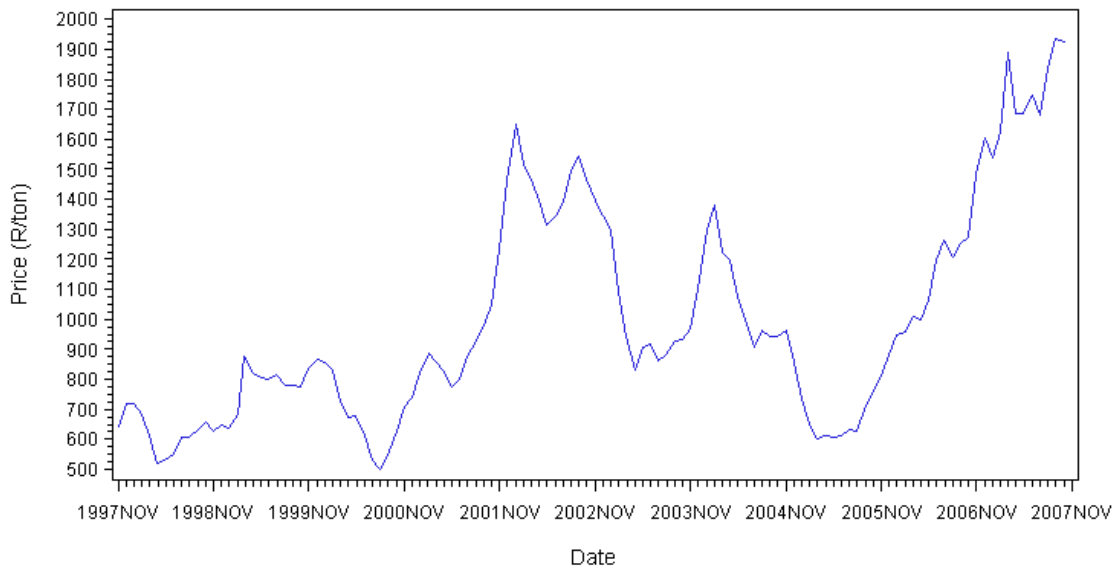


Figure 3.1 Average monthly prices of maize (11/1997 – 11/2007)

The graph above represents average monthly SAFEX yellow maize prices, in levels. Daily prices are reported by the Agricultural Marketing Division of the Johannesburg Stock Exchange. The data was converted into average monthly prices from the daily prices mentioned above. As with the logarithm decision discussed above, prices were left in nominal terms. A conversion to real prices would make the forecasting process clumsy and tedious since real prices and inflation would need to be forecasted in order to calculate a nominal forecast. Since it is nominal prices that are of interest in this study prices were not deflated.

If the univariate properties of monthly yellow maize prices were to be evaluated at face value, one might conclude that the series is non-stationary. This is due to the apparent spikes in 2002 and the steep increase from 2005 to the end of 2007. The spikes in 2002 were a result of the significant depreciation of the exchange rate. The increases in 2005 through 2007 can be attributed to a combination of factors such as world food and energy markets becoming more integrated as a result of bio-fuels. Oil prices for this period also increased significantly and as a result of this strong growth, pulled all commodity prices to unprecedented levels. Furthermore, the local maize market shifted from a surplus in 2005 to a shortage in 2007. In other words, prices moved from export

parity level in 2005 to import parity level in 2007. The shift was caused by a sharp decline in the area planted under maize in 2006 and a subsequent drought in 2007.

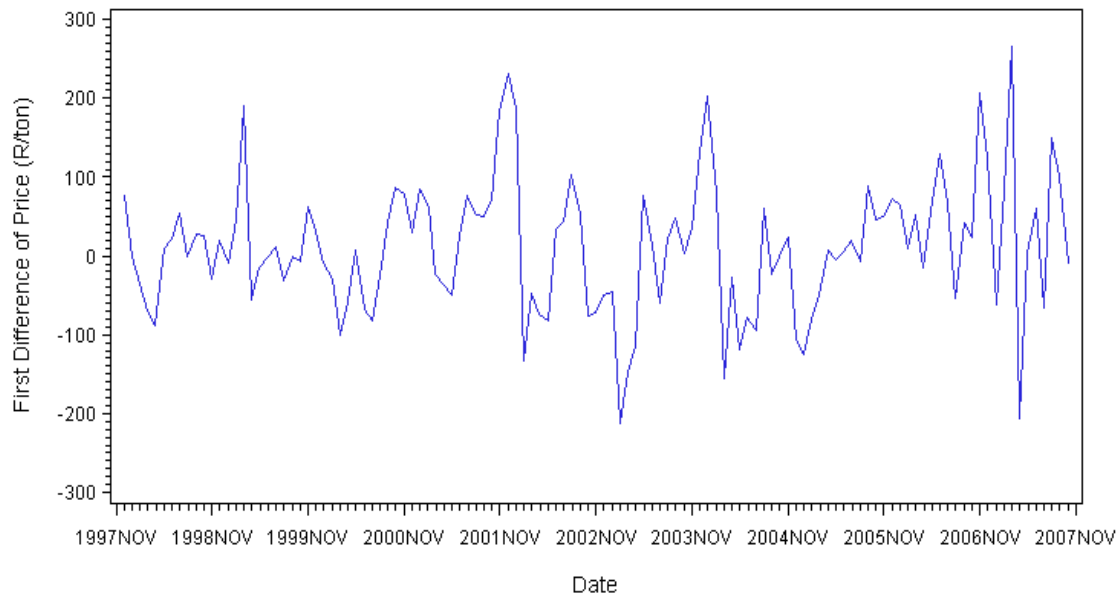


Figure 3.2: First Difference of Average Monthly Maize Prices (11/1997 – 11/2007)

The figure above represents average monthly yellow maize prices in first differences which appear to be stationary. Formal unit root test will be conducted in the next section to determine the exact univariate properties of the series discussed above.

3.2.3 Univariate Properties of the Data

In order to establish whether the data contains unit roots three, test were conducted namely the Augmented Dickey-Fuller (ADF), The Phillips-Perron(PP) and the Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test. These tests were conducted for the differenced series and the series in levels. The ADF and PP tests were conducted since these are the tests that are commonly found in agricultural time series studies (see *inter alia* Guttormsen, 1999 and Jordaan, Grové and Jooste, 2007). The KPSS tests were included for three reasons. Firstly, it serves as a compliment for the conventional ADF test, since ADF tests have power problems when applied to small samples (Gujurati, 2003 p.819). Secondly, the KPSS test, in contrast to the ADF test, has a $H_0=stationary$. According to Kwiatkowski, Phillips, Schmidt and Shin (1992), by testing

for a null hypothesis of both stationarity (KPSS) and non-stationarity (ADF), one can determine if a series that appear stationary, are in fact non-stationary and appear stationary only due to inadequate information. Thirdly, the KPSS test is a relatively new advancement in econometrics. To the best of the author’s knowledge, however, there has been no South African study in Agricultural Economics that has applied it and therefore benefited from this advancement in time series analysis.

The ADF tests applied below are for average monthly yellow maize prices in first differences and in levels. The first test is done to determine if the series has one or two unit roots. The tests are all applied in the zero mean form, since the series in first differences appear to fluctuate around a zero mean level.

The hypothesis for the ADF and PP test, applied to the differenced series, is therefore as follows:

$$H_0 = \text{First difference of series has a unit root, } d=1$$

$$H_A = \text{First difference of series have no unit root, } d=1$$

Table 3.1: Augmented Dickey Fuller Test for Differenced Monthly Maize Prices

Augmented Dickey-Fuller Unit Root Tests			
Type	Test Statistics and p-Values		
Zero Mean	Lags	Regression coefficient based test	Studentized test
	0	77.602 (<0.0001)	-7.60 (<0.0001)
	1	-84.476 (<0.0001)	-6.44 (<0.0001)
	2	-62.175 (<0.0001)	-4.85 (<0.0001)
	3	-96.719 (<0.0001)	-5.01 (<0.0001)

Table 3.2: Phillips Perron test for Differenced Monthly Maize Prices

Phillips-Perron Unit Root Tests			
Type	Test Statistics and p-Values		
Zero Mean	Lags	Regression Coefficient based test	Studentized test
	0	-77.60 (<0.0001)	-7.60 (<0.0001)
	1	-78.58 (<0.0001)	-7.62 (<0.0001)
	2	-76.32 (<0.0001)	-7.57 (<0.0001)
	3	-78.53 (<0.0001)	-7.62 (<0.0001)

From the significant p-values, indicated in brackets, of both the test tables it can be concluded that the H_0 , which states that the order of integration (d) is equal to two, can be rejected. The hypothesis for the KPSS test, applied to the differenced series, can be stated as follows:

$$H_0 = \text{Series has a unit root, } d=1$$

$$H_A = \text{Series has a unit root, } d=2$$

Table 3.3: KPSS test results for average monthly maize prices

KPSS test statistics for the null of <u>zero mean</u> stationarity of differenced average monthly prices of maize				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.2204	0.1963	0.18	0.1734

*** 10 % critical value = 1.1950

** 5% critical value = 1.6560

* 1% critical value = 2.7590

From the KPSS test results, it can be seen that none of the test statistics are significant. H_0 , which states that (d) is equal to one, can therefore not be rejected.

It has now been established that the series does not have two unit roots. The next step is to determine whether the series has one unit root. The ADF and PP tests are applied to the series in levels. Both the single mean and trend ADF test were conducted. From inspection of the graph of maize prices in levels it is unclear whether there is a trend or not. The hypotheses, applied to the series in levels, are therefore as follows:

$$H_0 = \text{Series has a unit root, thus } d=1$$

$$H_A = \text{No unit Root, thus } d=0$$

Table 3.4: Augmented Dickey Fuller Test for average Monthly Maize Prices

Augmented Dickey-Fuller Unit Root Tests			
Type	Test Statistics and p-Values		
Trend	Lags	Regression Coefficient Based Test	Studentized Test
	0	-3.18 (0.93)	-0.98 (0.9419)
	1	-9.05 (0.49)	-1.88 (0.6614)
	2	-8.56 (0.53)	-1.75 (0.7210)
	3	-12.29 (0.28)	-2.04 (0.573)
Single Mean	0	-0.4443 (0.9298)	-0.17 (0.9383)
	1	-4.0171 (0.5316)	-1.01 (0.7461)
	2	-3.7031 (0.5676)	-0.94 (0.7734)
	3	-5.8734 (0.3498)	-1.23 (0.6592)

Table 3.5: Phillips Perron Tests for average Monthly Maize Prices

Phillips-Perron Unit Root Tests			
Type	Test Statistics and p-Values		
Trend	Lags	Regression Coefficient Based Test	Studentized Test
	0	-3.185 (0.9269)	-0.98 (0.9419)
	1	-4.909 (0.8224)	-1.32 (0.8789)
	2	-5.719 (0.7599)	-1.45 (0.8403)
	3	-6.382 (0.7054)	-1.55 (0.8047)
Single Mean	0	-0.4443 (0.9298)	-0.17 (0.9383)
	1	-1.5995 (0.8227)	-0.53 (0.8812)
	2	-2.1366 (0.7596)	-0.67 (0.8504)
	3	-2.5735 (0.7056)	-0.77 (0.8238)

From the above tables, it is apparent from the insignificant test statistics, denoted by high p-values, that H_0 cannot be rejected.

The hypothesis for the KPSS test is stated below.

H_0 : No unit root, $d=0$

H_A : Series has a unit root, $d=1$

Table3.6: KPSS test results for average monthly maize prices

KPSS test statistics for the null of <u>single mean</u> stationarity of average monthly prices of maize				
Lag truncation Parameter	1	2	3	4
Test Statistics	2.1244**	1.4570*	1.1240	0.9251

*** 10 % critical value = 1.1950

** 5% critical value = 1.6560

* 1% critical value = 2.7590

Table 3.7: KPSS test results for average monthly maize prices

KPSS test statistics for the null of <u>trend</u> stationarity of average monthly prices of maize				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.3667*	0.2520*	0.1960**	0.1612**

*** 10 % critical value = 0.1190

** 5% critical value = 0.1460

* 1% critical value = 0.2160

It is apparent from the above table that the test statistics are all significant. H_0 is therefore rejected.

It can therefore be concluded that the data is non-stationary with one unit root. Thus, the original series needs to be differenced once in order to obtain stationarity. The discussed issues can be denoted mathematically as follows:

$$Y_t \sim I(1) \text{ thus } d = 1$$

$$\Delta Y_t \sim I(0)$$

with Y_t : Maize prices in levels.

and ΔY_t : Maize prices differenced once.

From the evaluation of the univariate properties of the data, as depicted above, it is apparent that the data need to be differenced once to obtain stationarity. Stationarity is a data requirement, when the Box-Jenkins methodology is applied. This step established the order of integration to be one. The next step is to identify the number of autoregressive and moving average terms and is depicted below.

3.2.4 Identification

There are essentially four methods that can assist in identifying an ARMA or ARIMA model. The first identification method is the Sample Autocorrelation Function (SACF) in combination with the Sample Partial Autocorrelation Function (SPACF) and the resulting correlograms. The sample autocorrelation function is simply the ratio of the sample covariance at a specified lag to the sample variance and measures the correlation between observations at different lags. The partial autocorrelation is the correlation between the observation at time 't' and an observations at a specified lag, after removing the effects of the intermediate observations. The correlograms are plots of the Autocorrelation Function(ACF) and the Partial Autocorrelation Function(PACF) against time (Gujurati, 2003). The autocorrelation function can be used to determine whether the MA process is appropriate for an observed series, and if so, to select the order of the process, q . In a similar fashion, the partial autocorrelation function can be used to determine whether the AR process is appropriate for an observed series and, if so, the appropriate value of p (Van Staden, 2008). Since the maize price time series is non-stationary with a unit root of one, the Box-Jenkins methodology have to be applied to the first difference of the series. Maize prices in first differences are indeed stationary as indicated in the previous data discussion. The ACF's and PACF's that are statistically significant from zero are lags which can be considered when deciding on the orders of the ARIMA model. These ACF's are denoted by bars in the correlogram that extends over the vertical dotted line.

The following paragraph refers to the ACF and PACF correlograms, of maize prices differenced once, in Appendix 1. In order to evaluate the series in terms of the autoregressive properties the PACF needs to be regarded. This correlogram indicates that from lag 2 onwards the partial autocorrelations are not significantly different from zero. This is an indication that an AR(1) process explain the data series relatively well. In order to evaluate the moving average properties of the series, the autocorrelation correlogram should be evaluated. This correlogram also indicate that from lag 2 onwards the autocorrelations are not significantly different from zero. This might suggest that an

MA(1) process might be a good representation of the discussed time series. Unfortunately, it is not permissible to use the correlograms for processes that might contain both an MA and an AR process. It is therefore necessary to resort to other test to provide guidance as to which model to choose.

The SCAN, ESACF and MINIC methods can be used to tentatively identify the orders of a stationary or non-stationary ARMA process. More formally these methods are named as the **S**mallest **C**anonical Correlation method, the **E**xtended **S**ample **A**utocorrelation **F**unction method and the **M**inimum **I**nformation **C**riterion method. The SCAN table that is constructed can be used to find a maximal rectangular pattern in which the test statistics are insignificant. The boundary of this rectangle identifies the proposed ARMA orders. The ESACF table can be used in a similar fashion by finding a triangular pattern in the table of which all the elements are insignificant. The boundaries of the triangular pattern again indicate the order of the ARMA process. The MINIC method can simply be used by finding the minimum value in the MINIC table. The coordinates of the minimum value in the table indicates the order for the ARMA process. The table below is a summary of the results of all the tests mentioned above. For a representation of all the individual tests please refer to Appendix 1.

Table 3.8: Summary of Tentative Order Selection Tests of Differenced Maize Prices

ARMA(p+d,q) Tentative Order Selection Tests					
SCAN			ESACF		
$P+d$	Q	BIC	$P+d$	q	BIC
1	0	8.774	0	1	8.763
0	1	8.763	1	1	8.802
			4	3	8.962

The above table indicates that the MA(1) process provides a good representation of the underlying data generating process, as indicated by the Bayesian information criterion

(BIC). Keep in mind that the order of integration (d) is zero for the above table since it is based on the differenced series.

3.2.4 Estimated Models

A MA(1) model in generic form can be represented as follows:

$$\Delta y_t = \mu + \varepsilon_t - \phi \varepsilon_{t-1}$$

with

$$\Delta y = y_t - y_{t-1}$$

where ε denotes the disturbance, ε_{t-1} the lagged disturbance and μ the possible constant term that should be included in the model. Δ Indicate that y is used in first differences.

The estimated model is depicted by the function below:

$$\Delta y_t = 10.87879 + \varepsilon_t - 0.37435 \varepsilon_{t-1}$$

(1.08) (-4.36)

with y_t : Differenced Average Monthly Prices of Yellow Maize.

ε_t and ε_{t-1} : The disturbance term and the disturbance term lagged, respectively.

It is apparent that the constant term included in the above model is not statistically significant (t-statistic < 1.96). It was therefore decided to estimate a model without a constant.

The resulting zero mean model can therefore be depicted as follows:

$$\Delta y_t = \varepsilon_t - 0.38071 \varepsilon_{t-1}$$

(-4.46)

3.2.5 Diagnostic Tests

Diagnostic tests for a model generated with the Box-Jenkins methodology is usually done by analysing the residuals obtained from the estimated model. The model tested was the MA(1) model without the constant term.

Table 3.9: Diagnostic test results of MA(1) model residuals

Test	H_0	Test Statistic	p-Value	Conclusion
Jarque-Bera	Residuals normally distributed	15.473	0.00044	Residuals not normally distributed
Ljung-Box Q	No autocorrelation up to order 6	3.18	0.6717	Do not reject H_0 , residuals are not significantly correlated.
ARCH LM (lag 6)	Homoscedasticity	16.0882	0.0133	Reject H_0 . Model has heteroscedastic errors

It is apparent from the above table that the residuals generated from the model are in fact, uncorrelated. From this it can be concluded that the above model is adequate. The residual sample auto correlation function (RSACF) and residual partial sample auto correlation function (RPSACF) in Appendix 1 can be used to confirm this fact. What is however cumbersome is the fact that the other diagnostic tests for normality and heteroscedasticity reveal grave problems. These issues will be elaborated on further in Chapter Five.

3.2.6 Forecasts

The expression *“the proof of the pudding is in the eating”* can be directly applied to the section that follows. It can be argued that if the forecast generated with the above models are adequate, the problems of non-normality and heteroscedasticity, identified in the previous section can be excused. Then again, it might be argued that these very problems will undoubtedly cause the generated forecasts to be poor. The table below regards the forecast generated with the zero mean model. This model was used to

generate the forecast for the obvious reason that the constant term in the first model was not statistically significant. The implication of using the zero mean model is that it forecasts a persisting mean level since there is no constant term to add to the moving average term. There are various statistical tests that can be used to evaluate the forecasts below. Whether it is necessary to include these tests is however debatable since it is apparent that generating a mean value as a forecast is unrealistic. The forecasts of a single mean value is of little use to decision making in the short run, since it approximates the mean value forecasted with annual models. The lack of variability between the monthly forecasts makes the usefulness of it uncertain. Producers of yellow maize will typically be interested in a 10 month ahead forecast, since the production and marketing period for maize is 10 months. If the forecast is essentially just an extrapolation of the month in which the producer plants, the forecast is of little value since the producer might as well base its production decisions on prices in the current month. The same applies to possible procurers of maize. Budgetary provisions made on the current price will have the same outcome as budgetary provisions made on forecasts of 10 months ahead, since it is approximately the same price.

The univariate model development was completed in December 2008. *Ex post* forecast were generated in order to evaluate the forecasting capabilities of the model. The initial intension of the study was to evaluate *ex ante* forecasts as well, but as a result of the poor forecasting performance no *ex ante* forecasts were generated and the forecasting horizon was not extended beyond October 2008.

Table 3.10: Actual vs Forecasted observations (12/2007- 10/2008)

	Actual Observation	Forecast	Std. Error
December, 2007	1877.00	1916.11	80.1619
January, 2008	1925.00	1916.11	136.6606
February, 2008	1830.33	1916.11	175.8587
March, 2008	1920.95	1916.11	207.7895
April, 2008	1872.27	1916.11	235.4287
May, 2008	1811.23	1916.11	260.1478
June, 2008	2072.95	1916.11	282.7137
July, 2008	2010.26	1916.11	303.6070
August, 2008	1847.90	1916.11	323.1523
September, 2008	1884.59	1916.11	341.5810
October, 2008	1813.52	1916.11	375.7366

3.3 A NOTE ON EXPONENTIAL SMOOTHING

Exponential smoothing is a forecasting technique commonly utilized in business statistics and forecasting. According to Gardner (1985) its popularity can be attributed to practical considerations such as simplicity, limited computational and data requirements and intuitive meaning of the parameters of the model. It should however be acknowledged that the various types of exponential smoothing models all have equivalent ARIMA procedures as described in McKenzie (1984). Simple exponential smoothing corresponds to an ARIMA (0,1,1) process, as discussed in the sections above.

There are certain advantages of using the ARIMA approach as stipulated in SAS/ETS 9.1:

- The optimal smoothing weight is automatically computed as the estimate of the moving average parameter of the ARIMA model.
- The problem of setting the starting smoothed value is eliminated. This is done in a statistical optimal way when the maximum likelihood method is used.
- The statistical estimates of the forecast confidence limits have a sounder theoretical basis.

In the literature study the methods were treated as unrelated and individual. Studies occasionally include both methods as individual approaches to generate forecasts. These methods will yield different results if smoothing constants for exponential smoothing are determined arbitrarily. It is however advocated that where econometric expertise and software are available, some form of the ARIMA model should be used. Another important aspect to consider is that exponential smoothing only allows for smoothing constants between 0 and 1. Sometimes this constant might be larger than one and it will only be possible to capture this effect in an ARIMA model.

3.4 CONCLUSION

The aim of this chapter was to apply the Box Jenkins methodology to construct a univariate forecasting model for average monthly yellow maize prices. This method was identified in Chapter Two as one of the common methods used in forecasting in articles on forecasting evaluation such as Spreen *et al* (1979), Oliviera *et al* (1979) and Harris and Leuthold (1985). These models and their associated forecasts are popular since they are simplistic with little data requirements. It can however be concluded that these models cannot be used fruitfully as far as forecasting average South African yellow maize prices are concerned. The estimated model was plagued by non-normality and heteroscedasticity and the forecasts were unrealistic.

It might be worthwhile to note at this point that 2008 was a year that saw unprecedented increases in maize prices, locally and abroad. There was as a result no way that a model without a constant term could account for such increases. If 2008 was a “normal” season without the massive increases that occurred, a stable average forecast might not have been that unrealistic. Regardless of this one could however argue that the forecasts are of little value since is to a large degree the same as average yearly price generated with structural models based on fundamental data.

CHAPTER 4

MULTIVARIATE MODEL

4.1 INTRODUCTION

The univariate models discussed in the previous chapter hold a simplistic appeal. A question that still beckons is whether additional fundamental information can improve on forecasts generated by univariate models. Intuitively, it is suspected that this would indeed be the case. There are however, several statistical considerations that need to be accounted for when developing a multivariate model. These issues will be elaborated on later in the chapter. At this point, it would suffice to note that it is not merely extending the model by including information on external factors. It might be statistically incorrect to include these external factors in the model. Keeping the above mentioned issue in mind, the remainder of the chapter is grouped into four sections. The first is devoted to the theoretical consideration for including the chosen additional time series. The second is an overview of the statistical tests conducted on the chosen additional time series. The third section is a representation of the developed model and the last section is the projected forecast with some concluding remarks.

4.2 VARIABLE JUSTIFICATION

According to Geyser and Cutts (2007) fundamental factors are the primary drivers of price. It is further explained that on the South African Futures Exchange, the fundamental factors influencing maize prices can be grouped into local and international factors. The local factors are domestic supply, demand and stock levels. The main international factors are international supply and demand which is reflected in international prices. A good proxy for international prices is the Chicago Board of Trade price of corn. The rand-dollar exchange rate also plays a key role since it directly affects

parity prices. At the hand of the above, the additional proposed fundamental factors to be included in the multivariate model are discussed.

Monthly deliveries, obtained from the South African Grain Information Service (SAGIS), are tested as a possible variable to include in a multivariate model since it represents the local supply of yellow maize. In order to incorporate a local demand time series, monthly utilization, as available of SAGIS, is also included. It should however be noted that there are certain fundamental factors that are not directly manifested in the local demand and supply series, as explained above. Yellow maize is an internationally traded commodity and, as mentioned, international occurrences play a significant role in the domestic price formation of yellow maize. In order to incorporate this into a possible model world maize prices and the Rand/Dollar exchange rate were identified as the main international fundamental factors that impact local yellow maize prices.

4.3 UNIVARIATE PROPERTIES OF THE DATA

As in Chapter Three, three tests were conducted to test the univariate properties of South African maize prices, namely an Augmented Dickey Fuller (ADF) test, Phillips-Perron (PP) test and the Kwiatkowski, Phillips, Schmidt and Shin (KPSS) test. For the sake of completeness, all the variables that are proposed to be included in a multivariate model in section 4.2 are tested in levels and in first differences with the tests as specified and discussed in Chapter Three. The hypothesis of each test will not be stated explicitly as in Chapter Three. Only the results and implication of the various tests will be discussed. Here, it might again however be worthwhile to note that the ADF and PP test, test a null hypothesis of non-stationarity whereas a KPSS test have a reverse null hypothesis of stationarity. For the sake of uniformity, the tests conducted in this chapter have the same format as in Chapter Three. The tests are first conducted on the differenced series to determine if the series are integrated of order one or order two. If it is found that a series is not integrated of order two the test will be applied to the variables in levels, in order to confirm that the series does have one unit root. Graphs of the tested time series are also shown, in order to give the reader a basic idea of what the nature is of the data that are tested.

4.3.1 Univariate properties of average monthly world prices

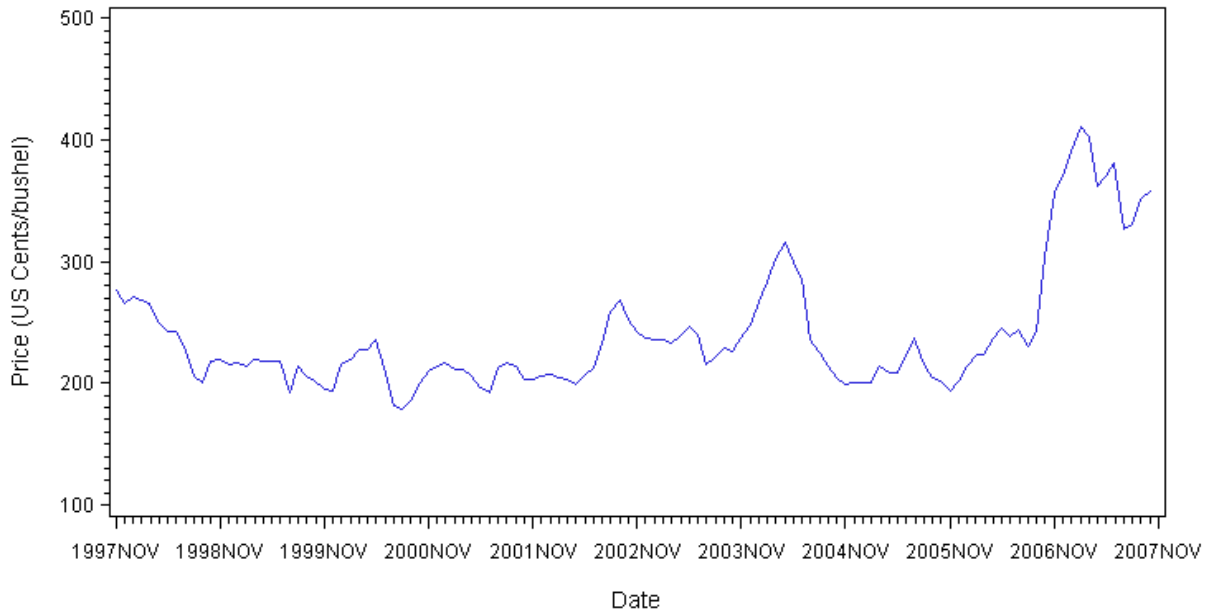


Figure 4.1: Average monthly world prices of maize in levels (11/1997-11/2007)

From the graph above it seems that world prices in levels are non-stationary. This can be deduced from, amongst others the large spikes in April 2004 and the increasing trends experienced since the end of 2006. The spike in 2004 can largely be attributed to U.S. stocks drawn down to relatively low levels at the end of the 2003/04 crop year. This occurred as a result of exports and ethanol use picking up in that year. Demand only strengthened over the course of the 2004 marketing year, explaining the upward trend in prices until April 2004. Prices, however, declined in later months as it became more and more clear that the U.S. would harvest a very big crop in 2004/05—as it turned out, the 2004 yield not only set a record, but it broke the 2003 record by more than 12 percent.

The increasing trend that started towards the end of 2006 is due to the, now familiar commodity price run, attributed to a myriad of factors. The main factors contributing to this severe upward trend are increased bio-fuel production, which caused a demand pull for corn, and increased energy and food market interrelatedness. The oil price increases caused by a large demand for oil in developing countries manifested in the maize price

as well. The trend was also amplified by several droughts in large grain producing regions

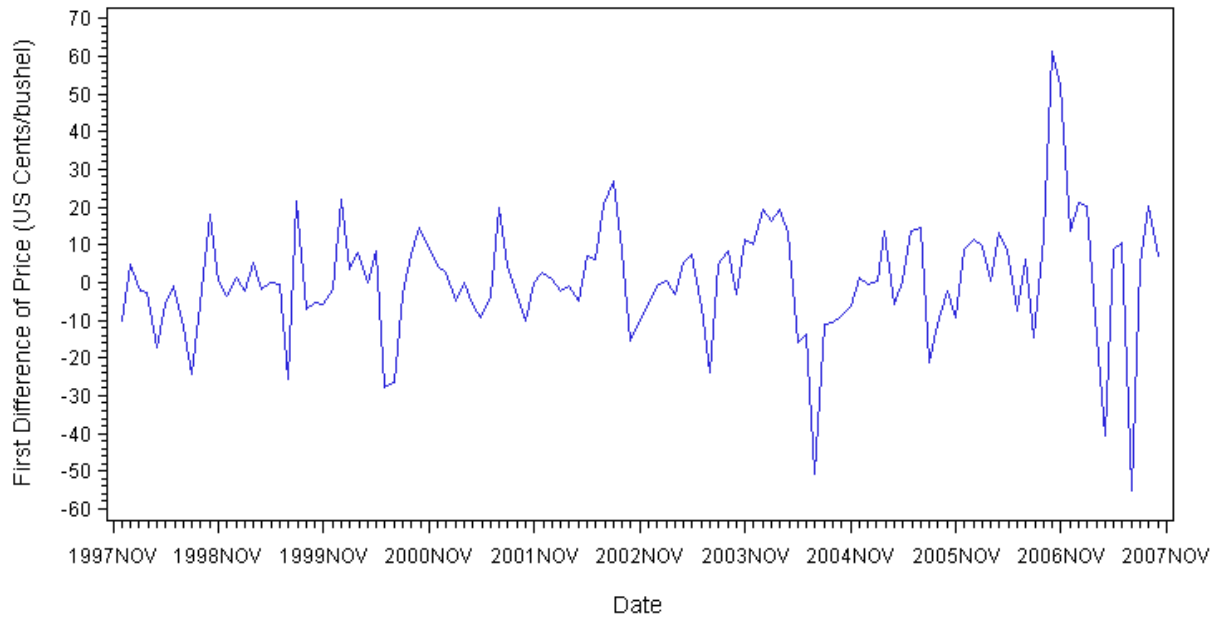


Figure 4.2: First Differences of Average Monthly World Maize Prices (11/1997-11/2007)

The series in differences, depicted in the graph above, appears to be stationary. This is to be confirmed with formal unit root testing as depicted in the tables below.

Table 4.1: Augmented Dickey Fuller Test for Differenced Monthly World Maize Prices

Augmented Dickey-Fuller Unit Root Tests			
		Test Statistics and p-Values	
Type	Lags	Regression Coefficient Based Test	Studentized
Zero Mean	0	-79.8051 (<0.0001)	-7.74 (<0.0001)
	1	-99.8297 (<0.0001)	-6.97 (<0.0001)
	2	-69.8264 (<0.0001)	-5.08 (<0.0001)
	3	-106.265 (<0.0001)	-5.10 (<0.0001)

Table 4.2: Phillips-Perron Test for Differenced Monthly World Prices

Phillips-Perron Unit Root Test			
		Test Statistics and p-Values	
Type	Lags	Regression Coefficient Based Test	Studentized
Zero	0	-79.8051 (<0.0001)	-7.74 (<0.0001)
	1	-81.9407 (<0.0001)	-7.79 (<0.0001)
	2	-77.6174 (<0.0001)	-7.69 (<0.0001)
	3	-78.1284 (<0.0001)	-7.70 (<0.0001)

Table 4.3: KPSS test results for differenced average world prices of maize

KPSS test statistics for the null of <u>zero mean</u> stationarity of differenced monthly average world prices of maize in first differences				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.1035	0.0957	0.09	0.0874

*** 10 % critical value = 1.195

** 5% critical value = 1.656

* 1% critical value = 2.759

Table 4.4: Augmented Dickey Fuller Test for Average Monthly World Maize Prices in levels

Augmented Dickey Fuller Unit Root Test			
		Test statistics and p-Values	
Type	Lags	Regression Coefficient Based Test	Studentized Test
Single Mean	0	-3.2209 (0.6254)	-0.94 (0.773)
	1	-9.2471 (0.1533)	-1.80 (0.3779)
	2	-7.6428 (0.2282)	-1.57 (0.494)
	3	-12.064 (0.0748)	-1.93 (0.3156)
Trend	0	-7.4121 (0.6195)	-1.94 (0.6256)
	1	-15.044 (0.1653)	-2.66 (0.254)
	2	-13.906 (0.2068)	-2.55 (0.302)
	3	-20.740 (0.0489)	-2.92 (0.1606)

Table 4.5: Phillips-Perron Test for Average Monthly World Maize Prices

Phillips-Perron Unit Root Tests			
		Test statistics and p-Values	
Type	Lags	Regression Coefficient Based Test	Studentized Test
Single Mean	0	-3.2209 (0.6254)	-0.94 (0.773)
	1	-5.1563 (0.4135)	-1.30 (0.6263)
	2	-5.8585 (0.3513)	-1.42 (0.5707)
	3	-6.4576 (0.3047)	-1.51 (0.5239)
Trend	0	-7.4121 (0.6195)	-1.94 (0.6256)
	1	-9.7394 (0.4381)	-2.22 (0.4720)
	2	-10.528 (0.3839)	-2.31 (0.4246)
	3	-11.1758 (0.3434)	-2.38 (0.3881)

Table 4.6: KPSS test results for average world prices of maize

KPSS test statistics for the null of <u>single mean</u> stationarity of monthly average world prices of maize				
Lag truncation Parameter	1	2	3	4
Test Statistics	1.7514*	1.2101*	0.9404*	0.7802*

*** 10 % critical value = 0.3470

** 5% critical value = 0.4630

* 1% critical value = 0.7390

Table 4.7: KPSS test results for average world prices of maize

KPSS test statistics for the null of <u>trend</u> stationarity of monthly average world prices of maize				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.4955*	0.3453*	0.2708*	0.2271*

*** 10 % critical value = 0.119

** 5% critical value = 0.1460

* 1% critical value = 0.2160

It is evident from the univariate test results above that, as expected from the graphs, average world prices needs to be differenced once in order to obtain stationarity. It can therefore be concluded that the series is integrated of order one.

4.3.2 Univariate properties of average monthly R/\$ exchange rate

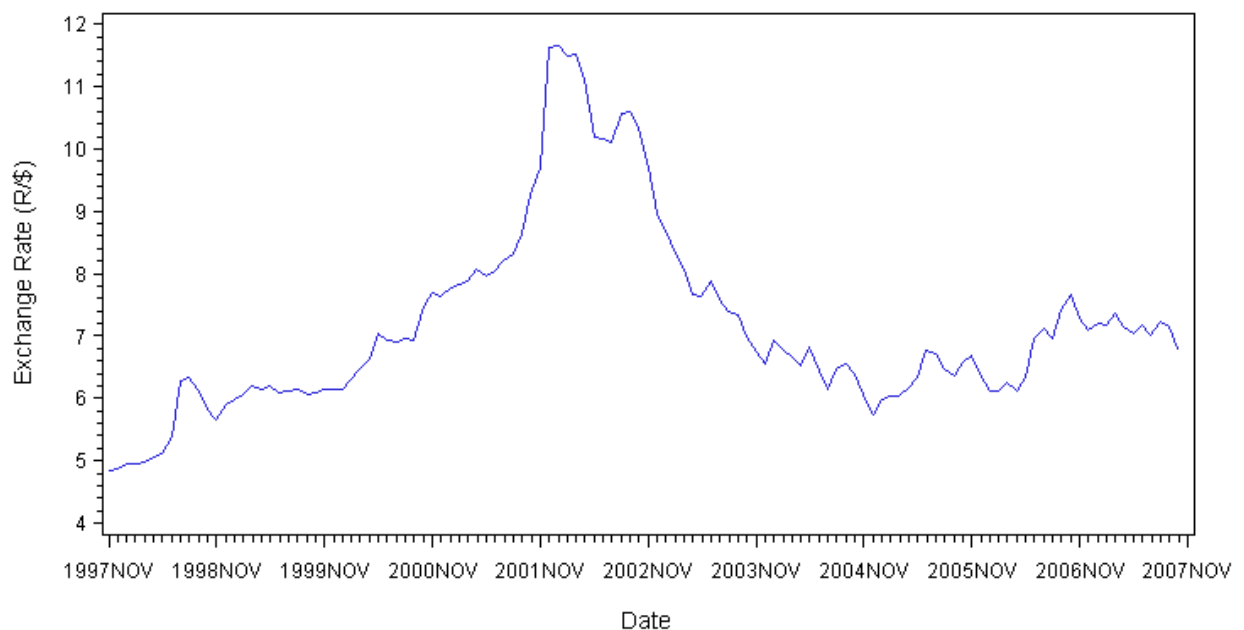


Figure 4.3: Average monthly R/\$ exchange rate in levels (11/1997 – 11/2007)

The Rand/Dollar exchange rate depicted in the graph above appears to be non-stationary. This is due to the significant depreciation experienced from the beginning of the time series to the end of 2001. This trend can largely be attributed to following South African developments:

- Worsening of current account balance in 2001
- Developments in financial markets that caused the financial account in the balance of payments to change from a surplus at the beginning of 2001 to a deficit at the end of 2001
- Continued slowdown in global economic activity

The rand recovered from the levels identified and discussed above, but as a result of this “spike” the series is expected to be non-stationary.

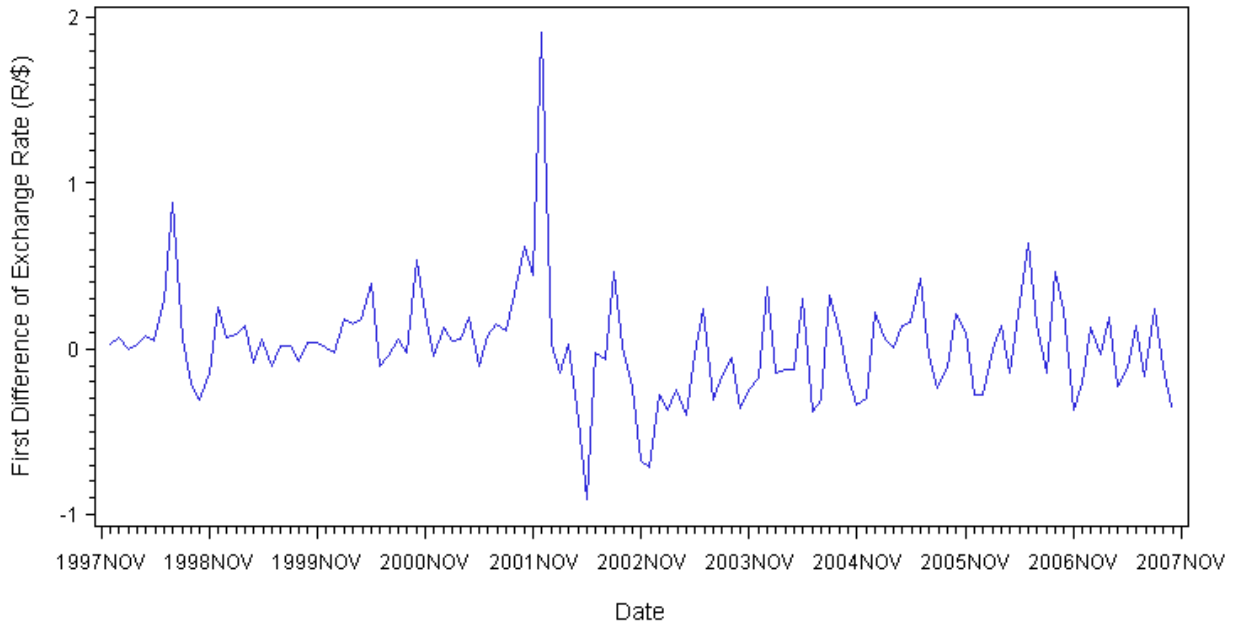


Figure 4.4: First Difference of average monthly R/\$ Exchange Rate

From the graph above, it seems that the series is in fact difference stationary. The exchange rate shocks in the 2001/2002 series are however still apparent, by the large spikes, in the differenced series and consequently the order of integration should be determined formally with a set of unit root tests.

Table 4.8: Augmented Dickey-Fuller Unit Root Test for Differenced Average Monthly R/\$ Exchange Rate

Augmented Dickey-Fuller Unit Root Test			
Type	Test Statistics and p-Values		
Zero Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-80.5626 (<0.0001)	-7.74 (<0.0001)
	1	-84.2613 (<0.0001)	-6.39 (<0.0001)
	2	-68.1173 (<0.0001)	-5.04 (<0.0001)
	3	-84.1889 (<0.0001)	-4.82 (<0.0001)

Table 4.9: Phillips-Perron Unit Root Test for Differenced Average Monthly R/\$ Exchange Rate

Phillips-Perron Unit Root Tests			
Type	Test Statistics and p-Values		
Zero Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-80.5626 (<0.0001)	-7.74 (<0.0001)
	1	-81.0017 (<0.0001)	-7.75 (<0.0001)
	2	-79.4471 (<0.0001)	-7.72 (<0.0001)
	3	-80.8216 (<0.0001)	-7.75 (<0.0001)

Table 4.10: KPSS test results for differenced average monthly R/\$ exchange rate

KPSS test statistics for the null of zero mean stationarity of monthly average world prices of maize				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.4861	0.4349	0.4019	0.3855

*** 10 % critical value = 1.195

** 5% critical value = 1.656

* 1% critical value = 2.759

Table 4.11: Augmented Dickey Fuller Test for Average Monthly R/\$ Exchange Rate

Augmented Dickey Fuller Unit Root Test			
		Test Statistics and p-Values	
Type	Lags	Regression Coefficient Based Test	Studentized Test
Single Mean	0	-3.81 (0.5553)	-1.66 (0.4484)
	1	-6.385 (0.31)	-1.96 (0.302)
	2	-6.303 (0.316)	-1.92 (0.3211)
	3	-7.5965 (0.2308)	-2.05 (0.2642)
Trend	0	-3.45 (0.9141)	-1.49 (0.8267)
	1	-5.9813 (0.7385)	-1.85 (0.6764)
	2	-5.8726 (0.7473)	-1.81 (0.693)
	3	-7.074 (0.6473)	-1.96 (0.6193)

Table 4.12: Phillips-Perron Unit Root Tests for Average Monthly R/\$ Exchange Rate

Phillips-Perron Unit Root Test			
		Test statistics and p-values	
Type	Lags	Regression Coefficient Based Test	Studentized Test
Single Mean	0	-3.81 (0.5553)	-1.66 (0.4484)
	1	-4.6298 (0.4655)	-1.76 (0.3978)
	2	-5.0387 (0.4247)	-1.81 (0.3728)
	3	-5.3586 (0.3947)	-1.85 (0.3536)
Trend	0	-3.45 (0.9141)	-1.49 (0.8267)
	1	-4.2536 (0.8676)	-1.61 (0.7825)
	2	-4.6381 (0.8418)	-1.67 (0.7595)
	3	-4.9298 (0.8209)	-1.71 (0.7415)

Table 4.13: KPSS test results for average monthly R/\$ exchange rate

KPSS test statistics for the null of <u>single mean</u> stationarity of monthly average R/\$ exchange rate				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.9569*	0.6482**	0.4942**	0.4022***

*** 10 % critical value = 0.3470

** 5% critical value = 0.4630

* 1% critical value = 0.7390

Table 4.14: KPSS test results for average monthly R/\$ exchange rate

KPSS test statistics for the null of <u>trend</u> stationarity of monthly average R/\$ exchange rate				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.9339*	0.6322*	0.4817*	0.3917*

*** 10 % critical value = 0.119

** 5% critical value = 0.1460

* 1% critical value = 0.2160

It is apparent from the results depicted above that the average monthly exchange rate series are integrated of order one. Stated differently the R/\$ exchange rate has a unit root.

4.3.3 Univariate properties of monthly deliveries of maize

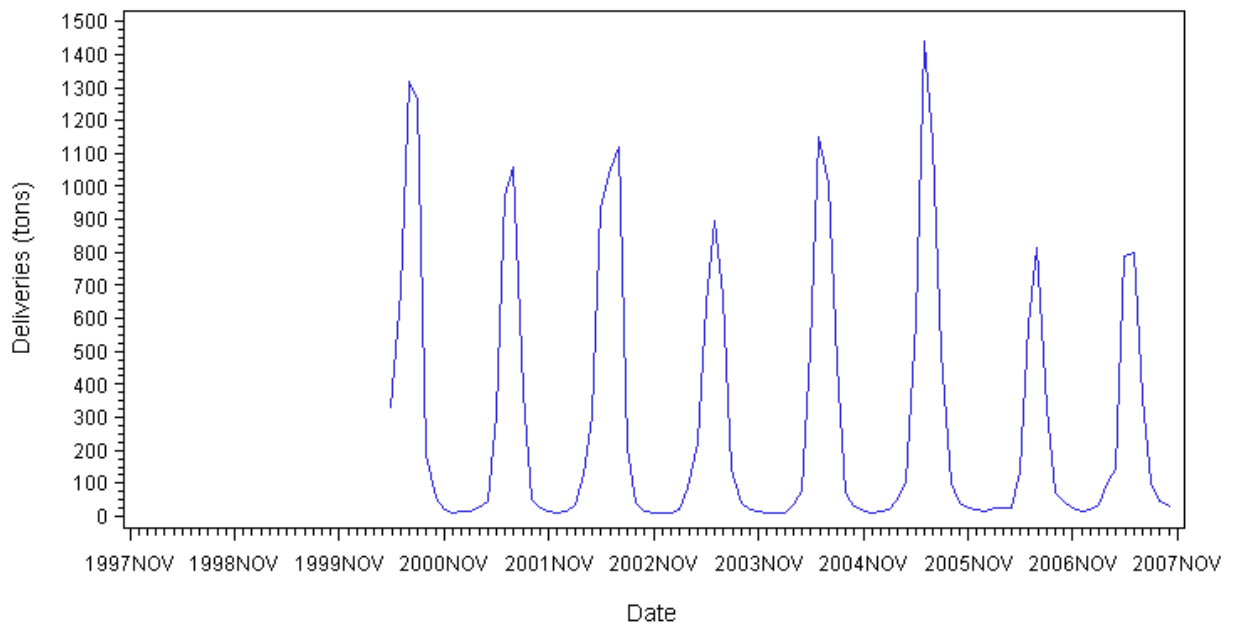


Figure 4.5: Monthly Deliveries in levels (05/2000 – 12/2007)

From the above graph, it is apparent that monthly yellow maize deliveries can be associated with an extreme seasonal pattern. Deliveries reached levels of between 800,000 tons and 1,450,000 tons, depending on the size of the harvest, in June and July and decreased to almost no deliveries during the pre-harvest months. When using seasonal data it may be necessary to allow for seasonal differencing.

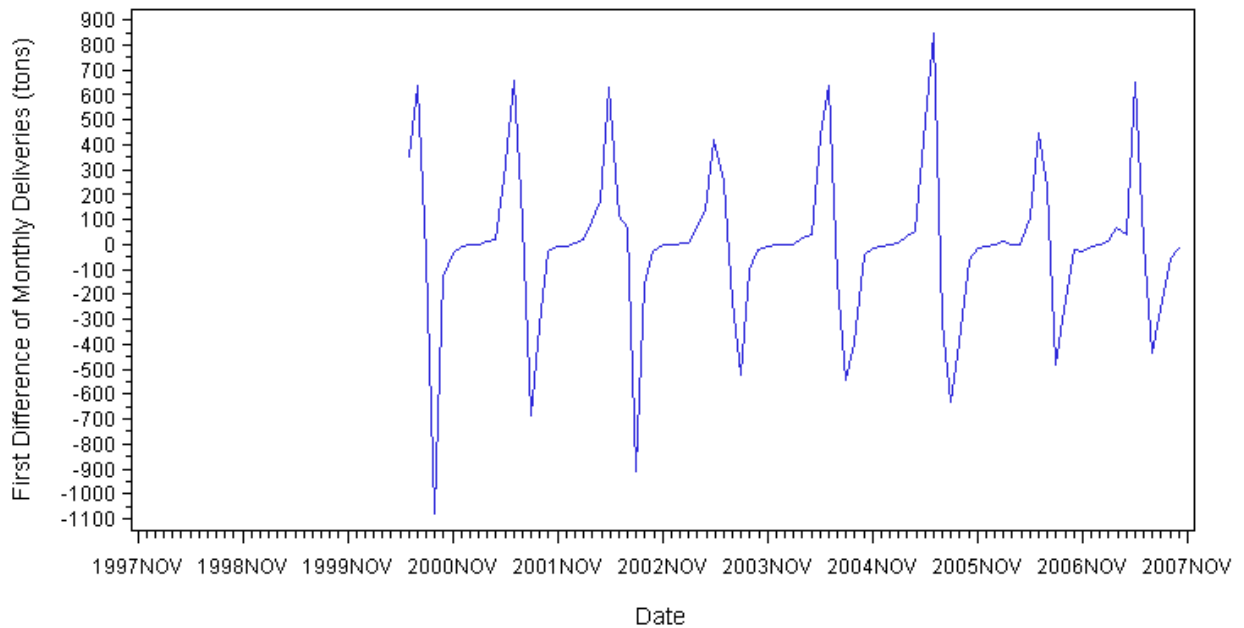


Figure 4.6: Monthly Deliveries in first differences (05/2000 – 12/2007)

Although heavily plagued by seasonality, the series was differenced and tested for unit roots for the sake of completion. It is apparent from the graph that the seasonal effect was not removed. The results of the unit root tests are depicted below.

Table 4.15: Augmented Dickey Fuller Test for Differenced Monthly Deliveries of Maize

Augmented Dickey Fuller Unit Root Test			
Type	Test Statistics and p-Values		
Zero Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-58.57 (<0.0001)	-6.64 (<0.0001)
	1	-146.76 (0.0001)	-8.81 (<0.0001)
	2	-225.06 (0.0001)	-6.94 (<0.0001)
	3	-297.89 (0.0001)	-5.89 (<0.0001)

Table 4.16: Phillips-Perron Unit Root Tests For Monthly Deliveries of Maize in First Differences

Phillips-Perron Unit Root Test			
Type	Test Statistics and p-Values		
Zero Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-58.5706 (<0.0001)	-6.64 (<0.0001)
	1	-63.3011 (0.0001)	-6.78 (<0.0001)
	2	-56.4452 (0.0001)	-6.58 (<0.0001)
	3	-48.4268 (0.0001)	-6.39 (<0.0001)

Table 4.17: KPSS test results for first difference of monthly maize deliveries

KPSS test statistics for the null of <u>single mean</u> stationarity of monthly deliveries of maize in first differences				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.0134	0.0142	0.017	0.0205

*** 10 % critical value = 1.1950

** 5% critical value = 1.6560

* 1% critical value = 2.7590

Table 4.18: Augmented Dickey Fuller Test for Monthly Deliveries of Maize

Augmented Dickey Fuller Unit Root Test			
Type	Test Statistics and p-Values		
Single Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-28.255 (0.0009)	-4.03 (0.002)
	1	-109.24 (0.0001)	-7.34 (<0.0001)
	2	-90.181 (0.0008)	-5.78 (<0.0001)
	3	-248.06 (0.0001)	-5.96 (<0.0001)

Table 4.19: Phillips-Perron Unit Root Tests For Monthly Deliveries of Maize

Phillips-Perron Unit Root Test			
Type	Test Statistics and p-Values		
Single Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-28.255 (0.0009)	-4.03 (0.002)
	1	-37.832 (0.0009)	-4.58 (0.0004)
	2	-37.908 (0.0009)	-4.58 (0.0004)
	3	-34.023 (0.0009)	-4.37 (0.0007)

Table 4.20: KPSS test results for monthly maize deliveries

KPSS test statistics for the null of <u>single mean</u> stationarity of monthly average R/\$ exchange rate				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.0876	0.0735	0.0716	0.0756

*** 10 % critical value = 0.3470

** 5% critical value = 0.4630

* 1% critical value = 0.7390

What is peculiar about monthly maize deliveries is the series in levels and the series in differences yield stationary conclusions. The cause of this is that the series is highly seasonal which causes it to move in seasonal patterns, around a mean level which, in turn, creates the illusion of stationarity. Seasonal series can however, not be regarded as stationary and as a result of this, the order of integration is unclear. It was therefore opted to exclude this series from the proposed multivariate model. In addition to this, the yellow maize delivery series, as available from the South African Grain Information Service (SAGIS) is not available for the full time series under consideration. This makes seasonal differencing impractical, since an extended lag structure and long time series is required to account for a seasonal unit root.

4.3.4 Univariate properties of monthly maize utilization

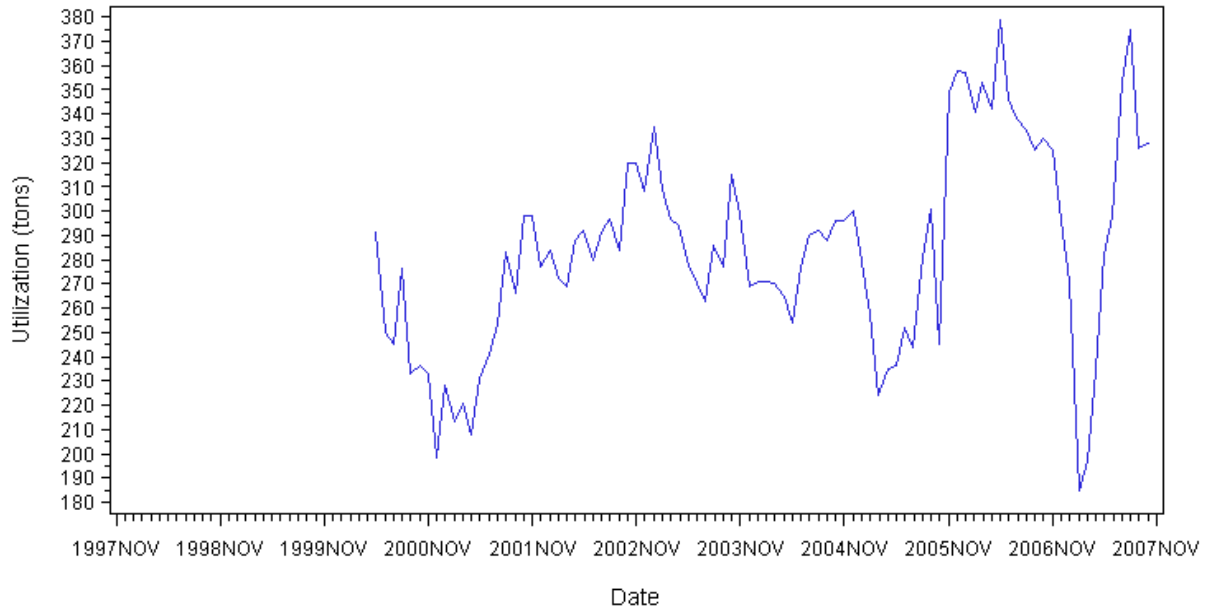


Figure 4.7: Monthly Utilization in levels (05/2000 – 12/2007)

From the graph, it seems the yellow maize utilization in levels is non-stationary. It appears that the quantities as well as the variation in quantities increased over the length of the time series. The large variation towards the end of the time series can possibly be attributed to the fact that maize prices in 2005 were low which encouraged utilization. In contrast to this utilization in the beginning of 2006 was very low. This might be due to the fact that sufficient utilization took place during the end of 2005 in combination with higher prices that are frequently seen in February months. These months tend to have higher prices since stocks are low and early harvest maize is not available yet.

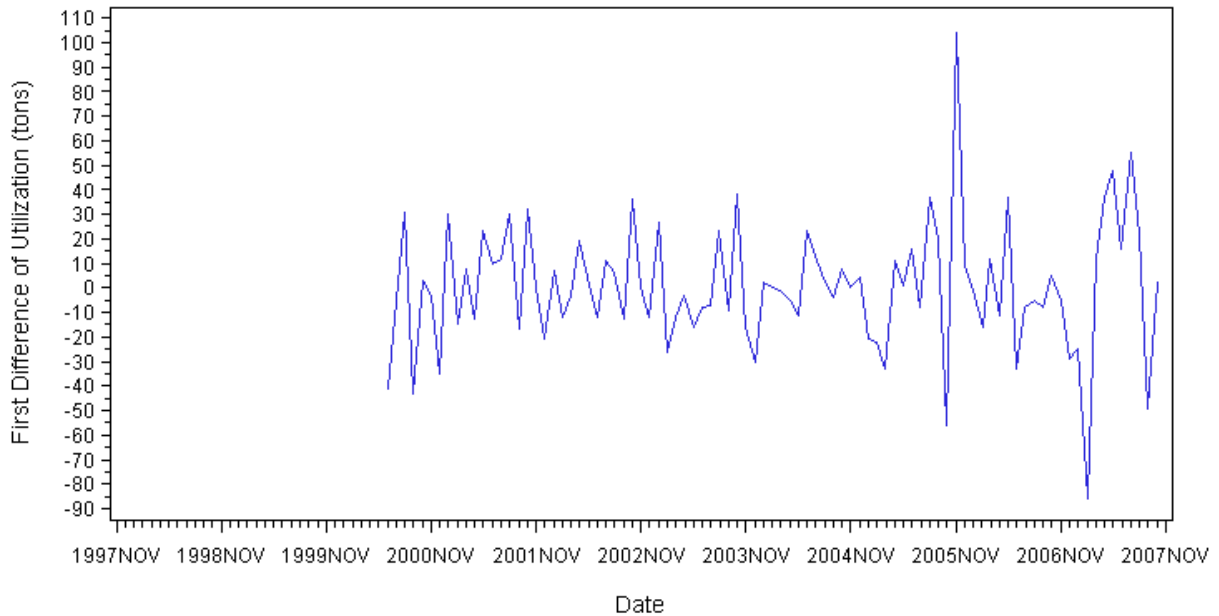


Figure 4.8: Monthly utilization in first differences (05/2000 – 12/2007)

The series above appears stationary but formal unit root tests are depicted below to establish if this is, in fact, the case.

Table 4.21: Augmented Dickey Fuller Test for Differenced Average Monthly Utilization of Maize

Augmented Dickey Fuller Unit Root Test			
Type	Test Statistics and p-Values		
Zero Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-97.5859 (<0.0001)	-10.56 (<0.0001)
	1	-94.3420 (<0.0001)	-6.79 (<0.0001)
	2	-67.9078 (<0.0001)	-4.82 (<0.0001)
	3	-150.447 0.0001	-5.06 (<0.0001)

Table 4.22: Phillips-Perron Unit Root Tests for Differenced Average Monthly Utilization

Phillips-Perron Unit Root Test			
Type	Test Statistics and p-Values		
Zero Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-97.5859 (<0.0001)	-10.56 (<0.0001)
	1	-97.3666 (<0.0001)	-10.56 (<0.0001)
	2	-98.5935 (<0.0001)	-10.54 (<0.0001)
	3	-100.083 (<0.0001)	-10.52 (<0.0001)

Table 4.23: KPSS test results for first difference of monthly maize utilization

KPSS test statistics for the null of <u>single mean</u> stationarity of monthly utilization of maize in first differences				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.0339	0.0346	0.0334	0.0352

*** 10 % critical value = 1.1950

** 5% critical value = 1.6560

* 1% critical value = 2.7590

Table 4.24: Augmented Dickey Fuller Test for Average Monthly Utilization

Augmented Dickey Fuller Unit Root Test			
Type	Test Statistics and p-Values		
Single Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-16.747 (0.0204)	-2.95 (0.0438)
	1	-15.865 (0.026)	-2.73 (0.0735)
	2	-21.011 (0.0061)	-2.95 (0.0443)
	3	-21.127 (0.0008)	-3.49 (0.0104)

Table 4.25: Phillips-Perron Unit Root Tests for Average Monthly Utilization

Phillips-Perron Unit Root Test			
Type	Test Statistics and p-Values		
Single Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-16.747 (0.0204)	-2.95 (0.0438)
	1	-16.587 (0.0214)	-2.94 (0.0452)
	2	-17.589 (0.0162)	-3.02 (0.0368)
	3	-19.100 (0.0106)	-3.14 (0.0270)
	4	-19.311 (0.01)	-3.16 (0.0258)

Table 4.26: KPSS test results of monthly maize utilization

KPSS test statistics for the null of <u>single mean</u> stationarity of monthly utilization of maize				
Lag truncation Parameter	1	2	3	4
Test Statistics	14.3086*	9.5976*	7.2401*	5.8259*

*** 10 % critical value = 1.1950

** 5% critical value = 1.6560

* 1% critical value = 2.7590

The differenced series is stationary and it can therefore be concluded that the series is integrated of order one. It was decided not to include utilization in the multivariate model since the time series available is much shorter than the time series under consideration.

4.3.5 Concluding remarks on the univariate properties

Based on the above results it was decided to include South African maize prices, world maize prices and R/\$ exchange rate as variables for a multivariate model. The reason for this is twofold. Firstly, data on monthly utilization and deliveries are only available from May 2000. This reduces the length of the time series significantly and could result in degrees of freedom problems when used in a multivariate model. The second reason

is that deliveries have a definite seasonal pattern that should be incorporated in the model as such. This will however, complicate the model development and forecasting generation stages. The length of the time series might also limit the model developer to take seasonality into account.

It is apparent from the above tables that the univariate properties of the data show that the series that are proposed for model are both integrated of order one. That is, both series are non-stationary and this aspect should be taken into consideration when developing the model. Non stationary data could still be used successfully in a model if the series under consideration are found to be cointegrated. A test and discussion on cointegration are included in the next section.

4.4 TEST FOR COINTEGRATION

Cointegration is an econometric concept which explains a phenomenon where non-stationary time series are combined in order to form a stationary series that mimics equilibrium between the two series. There are various tests that have been developed of which the Johansen test is common in agricultural economic articles. The popularity of the Johansen test can largely be attributed to the ability of this test to identify more than one cointegrating relationship, should it exist. There are two separate tests for the Johansen approach, namely the Trace test and the Eigen value test. Due to space limitations only the Trace test are shown. It should be noted that there are five different cases of the Johansen test that can be applied to test for cointegration. The case applied is determined by the properties of the data series under consideration. For the purposes of this study case 2 and case 3 were applied. Cases 2 and 3 are used when the cointegrating relationships have intercepts. The difference between the two cases is that case 2 have no deterministic trend in the original normalized series whereas case 3 have linear trends in the original normalised series. SAS 9.2 provides a test where it can be determined if case 2 or case 3 should be considered.

Table 4.27: Johansen Cointegrating Rank Test Using Trace (Case 2)

H0: Rank =r	H1: Rank>r	Trace	10% Critical Value
0	0	28.6096***	26.70
1	1	6.6213	13.31
2	2	0.2765	2.71

The Johansen test is conducted by means of matrix algebra. The objective of the test is to determine the rank of a particular matrix (based on the variables proposed to be included in the multivariate model). The determined rank is an indication of how many cointegrating relationships there are between the included variables. The above table are set up so that the first column displays the null hypothesis of $rank=i$. The second column displays the alternative hypothesis that $r>i$. From the above table it is therefore apparent that the null hypothesis in the first row can be rejected on a 10% level of significance since the trace test statistic is larger than the critical value in the last column. It can therefore be interpreted that the rank of the tested matrix is one and as a result there is one cointegrating relationship. This is however not the case for hypothesis tested for higher ranks as displayed in the table. It can thus be concluded that there is one cointegrating relationship between the variables that are included.

Table 4.28: Johansen Cointegrating Rank Test Using Trace (Case 3 under restriction)

H0: Rank =r	H1: Rank>r	Trace	10% Critical Value
0	0	29.4391	31.88
1	1	7.4505	17.79
2	2	1.0891	7.5

If we apply the same principles and critical values as discussed above it is apparent that there is no cointegrating relationship when the Johansen case 3 test are applied. The

next step is to apply a test to determine if case two or case three should be applied. This is done by testing a hypothesis of restriction as indicated below.

H_0 =Case 3 (Drift in ECM constant, drift in process linear)

H_A =Case 2 (Drift in ECM and process constant)

Table 4.29: Test of Restriction (Case 2 or Case 3)

Hypothesis Test of Restriction					
Rank	Eigen Value	Restricted Eigen Value	DF	Chi-Square	Pr>Chi Square
1	0.0523	0.0525	2	0.83	0.66

From the above table it is apparent that the null hypothesis cannot be rejected. It can therefore be concluded that Case 3 should be applied when using the Johansen test for cointegration. This leaves us with a dilemma since cointegration was obtained at a 10 percent level with case 2. Despite the above, the author went ahead with the model estimation of a VECM based on the cointegration obtained with case 2. It is expected that the model estimated from this and the associated forecasts will not be satisfying. Keeping in mind that this study is a forecasting evaluation, it was however decided to go ahead and generate a model and forecasts. This could possibly aid in the understanding and rectification of the problem at hand.

4.5 ESTIMATED MODEL

A generic representation of the VECM model are represented below

$$\Delta y_t = \alpha \beta' y_{t-1} + \sum_{i=1}^{p-1} \phi_{t-1}' \Delta y_{t-i} + \varepsilon_t$$

This generic depiction can be rewritten in the VECM(2) representation of the estimated model:

$$\begin{pmatrix} \Delta sa_price_t \\ \Delta world_price_t \\ \Delta exc_rate_t \end{pmatrix} = \begin{pmatrix} -0.03528 & 0.21501 & 4.50382 & -48.76924 \\ 0.03263 & -0.19888 & -4.16609 & 45.1122 \\ -0.0009 & 0.00056 & 0.01167 & -0.12633 \end{pmatrix} \begin{pmatrix} sa_price_{t-1} \\ world_price_{t-1} \\ exc_rate_{t-1} \end{pmatrix} \\ + \begin{pmatrix} 0.27847 & 0.2809 & 33.73069 \\ -0.02517 & 0.4154 & 2.3163 \\ 0.00044 & -0.00162 & 0.26703 \end{pmatrix} \begin{pmatrix} \Delta sa_price_{t-1} \\ \Delta world_price_{t-1} \\ \Delta exc_rate_{t-1} \end{pmatrix}$$

Table 4.30: Parameter Estimates of Vector Error Correction

Model Parameter Estimates						
Equation	Parameter	Estimate	Std. Error	t-value	Pr>t	Variable
D_sa_price	Const1	-48.76929	63.48602			1, EC
	AR1_1_1	-0.03528	0.04592			Sa_price(t-1)
	AR1_1_2	0.21501	0.27989			World_price(t-1)
	AR1_1_3	4.50382	5.86290			Exc_rate(t-1)
	AR2_1_1	0.27847	0.10548	2.64	0.0095	D_sa_price(t-1)
	AR2_1_2	0.28090	0.51497	0.55	0.5865	D_world_price(t-1)
	AR2_1_3	33.73049	26.17594	1.29	0.2002	D_exc_rate(t-1)
D_world_price	Const2	45.11220	11.00946			1,EC
	AR1_2_1	0.03263	0.00796			Sa_price(t-1)
	AR1_2_2	-0.19888	0.04854			World_price(t-1)
	AR1_2_3	-4.16609	1.01672			Exc_rate(t-1)
	AR2_2_1	-0.02517	0.01829	-1.38	0.1715	D_sa_price(t-1)
	AR2_2_2	0.41540	0.08930	4.65	0.0001	D_world_price(t-1)
	AR2_2_2	2.31630	4.53931	0.51	0.6109	D_exc_rate(t-1)
D_exc_rate	Const3	-0.12633	0.24241			1,EC
	AR1_3_1	-0.00009	0.00018			Sa_price(t-1)
	AR1_3_2	0.00056	0.00107			World_price(t-1)
	AR1_3_3	0.01167	0.02239			Exc_rate(t-1)
	AR2_3_1	0.00044	0.00040	1.09	0.4126	D_sa_price(t-1)
	AR2_3_2	-0.00162	0.00197	-0.82	0.0087	D_world_price(t-1)
	AR2_3_3	0.26703	0.09995	2.67	0.0087	D_exc_rate(t-1)

The table above shows the parameter estimates of the VECM(2) model. What is worthwhile to note is that the t-statistics of the AR(2) terms, show in each case, that there is only one variable that is significant. With closer inspection it is apparent that this is a lagged term of the dependent variable in question. What can be concluded from this, is that there is not really a simultaneous relationship between the variables, but that each variable is only affected by lagged values of itself. This is most probably an outflow of the fact that it was a struggle to obtain cointegration between the variables, which in turn can be explained by prices moving between different trade regimes. In different regimes different factors influence prices. This concept will be elaborated on further in the following chapter. The parameters in the table above, without t-statistics, were estimated with data in levels. It is known from the univariate test conducted earlier that the variables in levels are not stationary and as a result of this no t-statistics could be calculated.

4.6 DIAGNOSTIC TESTS

There exist a number of tools for determining if a VECM provides a satisfactory representation of the data generating process in question. The two options to be applied here are a graphical representation of the cross correlations of the residuals and the Pormanteau test.

Table 4.31: Cross correlations of estimated VECM residuals.

Schematic Representation of Cross Correlations of Residuals													
Lag/Variable	0	1	2	3	4	5	6	7	8	9	10	11	12
sa_price	+++	+.
world_price	++.
exc_rate	+.	+++

The table above is a graphical representation of the residual vectors' cross correlations. A good model will have no cross correlations of the residuals. In the diagram cross

correlations are indicated by plusses. It can be seen the only significant issue is at lag eight. This could've been rectified by adjusting the lag structure in the model. It should however be noted that this would not be parsimonious and with lag structures reaching 8 lags, one might encounter degrees of freedom problems.

Table 4.32: Portmantuau test for significance in cross correlations of residuals.

Portmanteau Test of Cross Correlation of Residuals		
Up to Lag	Chi-Square	Pr>ChiSq
3	15.17	0.0863
4	24.95	0.1263
5	34.87	0.1422
6	42.50	0.2112
7	51.90	0.2229
8	76.87	0.0221
9	81.39	0.0594
10	90.48	0.0695
11	92.86	0.1732
12	96.63	0.2974

The Portmanteau test is a formal statistical test that can be conducted to determine if there is any autocorrelation different from zero for a residuals group. From the table above it is apparent that the autocorrelations does not differ from zero, except when applying the test up to lag eight. This confirms the findings of Table 4.31.

Table 4.33: Goodness of fit test statistic of individual equations in VECM

Univariate Model ANOVA Diagnostics				
Variable	R-Square	Std Deviation	F-Value	Pr>F
Sa_price	0.1344	79.18152	2.87	0.0122
World_price	0.2180	13.73130	5.16	0.0001
Exc_Rate	0.1086	0.30233	2.25	0.0433

The R-square statistic is an indication of how well the variations in the various variables are explained by the estimated equations. It is apparent that all the R-squares are poor and it can therefore be concluded that the model has a poor fit. For the sake of completion the other ANOVA diagnostics were also included in the table. All these statistics corroborate the fact that the estimated model is not a good fit to the data.

Table 4.34: Autocorrelation, Normality and Homoscedasticity test of individual equations in VECM

Univariate Model White Noise Diagnostics					
Variable	Durbin Watson	Normality		ARCH	
		Chi-Square	Pr>ChiSq	F Value	Pr>F
Sa_price	1.96011	10.4	0.0055	6.99	0.0093
World_price	1.94487	34.57	<0.0001	2.00	0.1598
Exc_rate	1.95719	343.09	<0.0001	0.65	0.4230

The table above is a representation of additional residual tests. The Durbin-Watson (DW) tests are a confirmation of the autocorrelation structure as discussed by Table 4.31 and Table 4.32. The rule of thumb that can be used to evaluate the DW statistic is 2. Since all the Durbin Watson statistics approximates two, autocorrelation does not seem to be a problem in the model. The normality tests and heteroscedasticity test does however indicate that the residuals of the model are non-normal and have unequal

variance over time. This is undesirable. At this point, it would suffice to take note of the problems. Further discussions on these issues will follow in Chapter Five.

4.7 FORECASTING

Table 4.35: Actual vs forecasted observations of VECM (2007/12 – 2008/11)

	Actual Observation	Forecast	Std. Error
December, 2007	1877.00	1902.67	79.18
January, 2008	1925.00	1888.58	132.74
February, 2008	1830.33	1880.83	176.27
March, 2008	1920.95	1876.86	212.63
April, 2008	1872.27	1874.97	243.95
May, 2008	1811.23	1874.17	271.68
June, 2008	2072.95	1873.92	296.76
July, 2008	2010.26	1873.92	319.83
August, 2008	1847.90	1874.00	341.30
September, 2008	1884.59	1874.10	361.49
October, 2008	1813.52	1874.18	380.60
November, 2008	1741.15	1874.24	398.79

As with the forecasts generated by the univariate model, the above forecasts are disappointing. This is to some extent expected since there were already problems in the estimation stages of the multivariate model. Again, as with the univariate model, the forecasts tend towards a mean level, which has the same decision making application as an average annual model. The main reason for the similar results obtained by the two models is that exchange rate and world prices are for all practical considerations, exogenous to the estimated system. Each of the variables included in the model is only affected by its own lag structure. In accordance with the forecasts represented in Chapter Three, *ex post* forecast were generated in order to evaluate the forecasting

capabilities of the model. The initial intension of the study was to evaluate *ex ante* forecasts as well, but as a result of the poor forecasting performance no *ex ante* forecasts were generated and the forecasting horizon was not extended.

4.8 CONCLUSION

The objective of this chapter was to determine if additional fundamental factors included in a forecasting model, improves on a forecasts generated from a univariate model. The short answer to this is, in the case of South African yellow maize prices, it does not. The long answer to this would include the reasons for this and are shortly mentioned in this conclusion. In the chapters that follow, some of the problems are however further explored.

The first and most apparent obstacle encountered, is the fact that limited data on related series are available to include in the forecasting process. Utilities and deliveries were not included in the model because of the reduced time series and the seasonal properties associated with it. Seasonal properties do not indicate a problem, *per ce*, but with the shorter time series it is not desirable to use a lag structure that would account for seasonality. The only data, in monthly frequency, that were left to include were R/\$ exchange rates and world prices of maize, which are key determinants of price, only when prices are trading at parity levels. As a result, it was a struggle to obtain cointegration between South African yellow maize prices, R/\$ exchange rates and world prices of maize. Cointegration is needed between the variables in order to develop a model since all the variables under consideration are non-stationary. Cointegration was however obtained with a case 2 application of the Johansen test at a 10 percent level of significance. A restriction test indicated that case 3 of the Johansen test should rather be applied, but no cointegration was obtained with this case.

Despite the problems mentioned above, cointegration obtained with the case 2 application of the Johansen test was used to estimate a VECM. This might not be statistically sound, but as mentioned earlier, this study is first and foremost concerned

with forecasts. A forecasting evaluation and forecasts are therefore needed to evaluate the method in question and also to attempt to identify the reasons for it.

As expected, the estimated model was plagued by problems such as insignificant variables, poor fit, non-normal residuals and heteroscedasticity. Poor fit and insignificant variables can possibly be addressed if there were other variables that could've been included in the model. Exchange rates and world prices could then rather be included to fulfill an exogenous role the model. Unfortunately, relevant data, on endogenous variables, in a monthly frequency is unavailable.

Lastly, the forecasts yield more or less the same results as the univariate model estimated in Chapter Three. The reason for this is that each variable is only affected by its own lag structure and they do not function as a system. The forecasts change marginally from one period to the next and essentially tend toward a mean level. This is not useful in short(er)-term decision making since it will have more or less the same application as an average annual model.

CHAPTER 5

TIME SERIES MODELS REVISITED

5.1 INTRODUCTION

The forecasting results in Chapter Three and Four are somewhat disappointing. This, however, can be attributed to the poor quality of the estimated models. Both models were plagued by non-normal residuals, heteroscedasticity etc. According to initial planning, Chapter Five was earmarked to be a comparison between the generated univariate and multivariate forecasts, and futures prices as forecasts. The logic behind this is that futures prices are widely published and readily available. If futures prices turned out to yield better forecasting results than those generated with estimated models, the resources spent on model development would not be justified. In addition to this the EMH as applied to average monthly data could be confirmed. As a result of the mentioned disappointing forecasts, the line of thought of the study had to be adapted.

The rest of this chapter is devoted to identifying possible explanations for the poor performance of the models in Chapter Three and Four. A section is also dedicated to identify circumstances in which time series modelling might yield satisfying results. Lastly, futures as forecasts will be evaluated. The reason for this is two fold. Firstly, Chapter One stated that an investigation into the possibility of forecasting prices is in essence a test of the efficient market hypothesis. The second reason is that if futures prices are the best short term forecast available, it is worthwhile to gain an in depth understanding how it can be incorporated into the decision making process.

5.2 POSSIBLE CAUSES OF POOR FORECASTS

5.2.1 Introduction

It seems that the strategic nature of maize in the agricultural industry is the very reason why the models considered previously, yield unsatisfying forecasts. The fact that this is a key commodity with an underlying futures exchange makes it difficult to model. The

existence of an associated futures market forces the study into the financial econometric field, which was not considered in the previous chapters. In the light of the above statement, the preceding models and some of the diagnostic tests conducted are revisited.

5.2.2 Univariate properties revisited

5.2.2.1 Goodness of fit of the ARIMA model

The measure used to test the goodness of fit of an ARIMA model is to evaluate the autocorrelation structure of the error term. This was done in Chapter Three by conducting the Ljung-Box test which is common to apply in determining if there are autocorrelation in a series. Svetlozar (2007 p. 264) noted the following about the above mentioned test:

“ the Lung-Box test check only for uncorrelatedness of the error terms and not for independence or true white noise”

This means that if data is generated by a non-linear process, we may find an ARIMA model that explains the autocorrelation structure of the data. The implication is that the error terms would be uncorrelated and tests conducted would show the ARIMA model to be satisfactory. However, more complex sequential dependencies may still exist in the model. This means that a linear ARIMA model is unsuitable to model the series in question.

A test that could be conducted to determine if there are some non-linear dependencies is the MacLeod and Li (1983) test. This tests the autocorrelation structure of the model with squared residuals. It however falls out of the scope of this study to determine the presence of non-linear dependencies. Future studies in time series forecasting might however find it useful to consider and apply the above mentioned test.

5.2.2.2 Volatility

In Chapter Three a test was done to determine if the univariate model that was developed does have time invariant variance. The hypothesis was however rejected (refer to ARCH LM test p.41). Variance is an indication of the volatility of the market. For

the purposes of this study, we understand volatility to be the magnitude and frequency with which prices change. It is somewhat logical to expect that a price series with an associated derivative market will have unequal variance. There are, however, ways in which this variance can be addressed. This is by constructing an ARCH or GARCH model. The reason why this route was not pursued was that ARCH and GARCH methods model volatility and not prices in levels. As mentioned this study was not initiated to facilitate investment decisions of financial market participants but served to be an extension of the current BFAP models so that intra period (annual) price movements can be projected. The requirements of these models and projections was that it had to be easy to update and easily understood so that regular market players in the cash market (as opposed to the futures and options market) could use this to facilitate short term decision making such as procurement and marketing decisions. Granger and Poon (2003 p. 478) concluded in an article on volatility forecasting that *“...because of the complex issues involved...volatility will continue to remain a specialist subject....”*

The essence of the quotation above is that volatility forecasts is a specialized field with appeal mostly to economic agents in the financial sphere. This contradicts with the objectives of the study as set out in Chapter One and was as a result not perused further.

The implication of unequal variance for time series forecasting for price level is not that detrimental for point forecasts. It is however difficult to determine the prediction interval in which the price might fall. Heteroscedasticity can make the forecast error variance imprecise since the predicted forecast variance is based on the average variance instead of the variability at the end of the series.

In addition to this, Geysers and Cutts (2007) stated that yellow maize price volatility increases during the planting season and initial growth stages. It decreases throughout the growth period and reaches a minimum when a final crop estimate is established. One might therefore argue that based on these facts alone, the usefulness of generating

a forecast based on an econometric model becomes pointless. A decision maker (producer or processor) will value forecasts the most at the initial stages of the planting/growing season to facilitate production, marketing and procurement decisions (regardless if he/she uses a futures market or not). This is when the range of possible prices is at its widest and the forecasts' variance may be incorrect due to heteroscedasticity present in the model.

5.2.2.3 Fat tail phenomenon – Non-normal distributions

Fat tails are usually a phenomenon discussed when risk and returns are regarded. It is however the opinion of the author, that it is just as applicable to forecasting of a price series in levels. Fat tails have the implication that probabilities assigned to certain occurrences are undervalued causing market players to undervalue the probability for the occurrence of outliers.

For price forecasting, fat tails mean that a very low probability is assigned to very low or very high prices. If we think of the nature of agriculture this is not the case. Think of the high prices that we experienced in 2008 or the extremely low prices in 2000. Agriculture, even more so than other industries, is exposed to natural and economic conditions that cause extreme price levels. The implication of fat tails is therefore that the probability of prices that deviate far from the mean is higher than with a normal or Gaussian distribution. From this it might be concluded that forecasting based on econometric models is a futile exercise since most econometric models are derived with the mean and normality assumption playing a central role. Methods that are based on distributions that are adjusted for fat tails should therefore be kept in mind when agricultural prices, with a tendency to extreme levels, are being modelled.

5.2.2.4 Summary and Conclusion

The section above attempted to identify possible causes for the poor models and the dismal forecasts resulting from it. After consultation of various literature sources (Wei and Leuthold, 1998 and Lux and Marchesi, 1998) it became apparent that the

phenomena listed above are common characteristics of most financial time series and might very well be the cause, or at least part of the cause, for the disappointing results in Chapter Three and Four. Although common, these characteristics are difficult to explain and even harder to model with econometric time series techniques. It is argued that decision makers who want to make use of forecasts should venture into techniques such as artificial neural networks or qualitative forecasting approaches. These approaches might yield better forecasting results since it is less affected by the univariate properties of the underlying data generating process, which are in turn influenced by factors such as those discussed above.

5.2.3 Cointegration and multivariate model revisited

5.2.3.1 Introduction

The previous section on the reasons of poor forecasting performance of the univariate model is very technical. The reasons for this are that there is no logical economical theory incorporated into this model and as a result the conclusions derived from this are a-theoretical. Since a VAR or VECM superficially replicates a structural econometric model it might be easier to derive some fundamental conclusions from the poor forecasting performance of the multivariate model.

Referring to Chapter Four, it is apparent that the cointegration tests conducted yield results that are not completely sufficient to construct a multivariate model. The test used to determine if there is cointegration between the variables is the Johansen test for cointegration. As mentioned in Chapter Four, this is a popular test since it tests for more than one cointegrating vector using matrix algebra. The problem with this test, is in order to obtain cointegration and construct a useful Vector Error Correction model, the variables included in the model are all regarded as endogenous. That is, all variables included in the model should be affected by their own lagged values and lagged values of all the other variables. From the model depicted in Chapter Four it is apparent that this is not the case. This is to be expected, for two reasons. The first is that South Africa

is a small maize producer, by international standards, and it is therefore not to be expected that South African maize prices will have an effect on world maize prices. Similarly, maize is thought to have no effect on the rand/dollar exchange rate. Geysers and Cutts (2007 p. 293) however note the following:

“...fundamental factors determining the price of maize...are: supply and demand at the international level, as reflected in the CBOT price....as well as the Rand-Dollar exchange rate...”

It can therefore be concluded that despite the fact that a two way relationship is not present, world prices and exchange rates are important in determining South African maize prices. It is therefore opted to test for cointegration with a single equation method as opposed to the system of equations used in the Johansen approach, as applied in Chapter Four. There are various single equation methods that can be used to test for cointegration. The Engle-Granger approach is the method applied below. Usually, the most important criticism associated with this method is that all the explanatory variables are considered to be exogenous. In this case, it is not a problem, since it addresses the problem that there is not a simultaneous relationship between the variables included in the model in Chapter Four. The Engle-Granger cointegration test for the full period is included in the Appendix 2(a). The reason for this is that cointegration was still not obtained with this test for the period as a whole. With further inspection and reading, the reasons for this became apparent. Meyer, Westhoff, Binfield and Kirtsen (2006) point out that the formation of prices in a specific market changes as the market switches between different trade regimes. The various regimes according to Meyer *et al* (2006) are explained in the table below:

Table 5.1 Trading regimes in the South African yellow maize market

Regime:	Definition:
Import Parity	The difference between import parity price and the domestic price exceeds transfer costs and the possibility of arbitrage integrates the local and world markets at prevailing international prices. This would trigger imports of the commodity into the South African market. One would expect the market price in South Africa to move with the price on international markets, plus the cost of shipping commodities to South Africa.
Autarky	If domestic prices are below that which triggers imports, but not low enough to be competitive on international markets, domestic prices will be determined by supply and demand conditions in that market.
Export Parity	The difference between export parity price and the domestic price exceeds transfer costs and the possibility of arbitrage integrates the local and world markets at prevailing international prices. The country can export grain to the world market.

The definitions above, along with visual inspection of annual graphs included in the Appendix 2(b), were loosely applied to determine the periods in which prices traded at or near parity levels. The term “loosely” is used since the original time series is relatively short (120 observations) and a strict definition used to identify prices at parity levels might limit the restricted time series so that cointegration testing might not be feasible. Out of the 120 observations of the original time series, 62 were identified as at, or close to, parity levels. The reduced number of observations is also tested for cointegration to determine if there is a long run relationship between variables when prices are trading at parity levels. The method used to test for cointegration is the Engle Granger test. The results of the test are depicted below:

Table 5.2: Long-Run Cointegration Equation (Dependant Variable: SA Average Maize Price)

Variable	Coefficient
World Price	6.453089
Exchange Rate	166.9102
Constant	-1766.344

Table 5.3: Engle-Granger Cointegration Test

Series	Model	Lags	ADF	Critical Value (5%)
Residual Cointegration Equation	Constant, No Trend	3	-4.413040**	-3.88

Table 5.2 represents the long-run relationship between the variables in question. The Engle-Granger cointegration test confirms that the variables are cointegrated close to or at parity levels. Stated differently, the non-stationary series can be combined to obtain a series with a stationary residual, which signifies a long run equilibrium between the variables.

Table 5.4: Error Correction Model (Dependant Variable: D[SA Average Maize Price])

Variable	Coefficient	Std Error	t-statistic	p-value
Resid(-1)	-0.256687	0.093643	-2.741134	0.0081
D(World Price)	3.122220	0.597042	5.229479	0.0000
D(Exchange Rate)	169.8168	24.93404	6.810639	0.0000
R-Square: 0.60				
Adjusted R Square: 0.59				
DW Statistic: 1.61				

The error correction model in the table above, gives an indication of how the short run behaviour of the relationship estimated above fluctuates around a long-run equilibrium. The model can be evaluated statistically, since it is based on differenced data. It is apparent that all the variables are significant, with an adequate R-square. What is somewhat alarming is the Durbin-Watson, which is slightly less than the rule of thumb value of 2. This might be an indication that there is still some relevant information contained in the residuals and the model can be improved by including additional variables or extended lag structures, provided that the length of the time series is long enough.

Table 5.5: Engle-Yoo Third Step

Variable	Adjusted Coefficient	Adjusted t-statistic
World Price	$6.453+0.514=6.967$	$6.453/0.723=8.9198$
Exchange Rate	$166.9102-10.2695=156.64$	$166.9102/23.782=7.018$

An Engle-Yoo step is performed to adjust the long-run coefficients and t-statistics for initial bias. This step is performed so that the variables included in the long-run equation can be evaluated statistically and economically. From relatively high (>1.96) adjusted t-statistics depicted in the table above, it is apparent that the variables are both statistically significant. The economic evaluation considers the sign and magnitude of the coefficients. The signs on both variables are positive, which is as expected. Magnitudes are evaluated by regarding elasticities. The following elasticities have been calculated and provide adequate results.

Table 5.6: Calculated Elasticities for World Price and Exchange Rate

Variable	Elasticity
World Price	1.64
Exchange Rate	1.12

It seems from the above results that cointegration was obtained and an adequate error correction model was developed. The problem associated with this type of model is two fold. Firstly, the model is developed based only on parity prices. The implication of this is that it is only able to generate forecast when markets are expected to function at import or export parity. The challenge that flows out of this is that the decision maker/forecaster needs to determine, in advance whether the market will be at, or close to parity levels. Secondly, forecasts for all the variables included in the model are not generated by the model. Forecasts of the exogenous variable, which in the above model is exchange rate and world prices, are therefore not generated by the model. These problems are not impossible to overcome, but the forecaster will have to venture into qualitative forecasting techniques to address it. A qualitative technique that can be used to

generate forecasts to address the above mentioned problems is intuitive scenario analysis as applied by Strauss (2010).

5.3 CONCLUSION

The objective of this chapter was to establish the possible causes for the poor models and forecasts in Chapter Three and Four. Three possible problems were identified with the univariate model. The first is non-linear dependencies in the yellow maize price data generating process which makes linear time series methods the wrong method to apply. Secondly, volatility clustering causes the residuals of the models to be heteroscedastic. Thirdly, models are based on the assumption that the data generating process is normally distributed. This is rarely the case for financial time series. Models that are based on distributions that take this into account should therefore be considered. The majority of these problems result from the fact that maize prices have an associated futures market. The impact of this was not previously considered, since the objective of the study steered the study away from investment decision making pertaining to financial markets.

The multivariate model had problems with cointegration. This was addressed by constructing and using a series based only on parity prices. It therefore became apparent that the R/\$ exchange rate and world maize prices are only key drivers of local maize prices when prices are formed at, or close to, parity levels. The implication of this on forecasting is that it should be known in advance if the maize market is going to trade at parity levels. A further implication is that the exogenous variables, which in this case are exchange rate and world prices, should be determined from sources outside the model. This can possibly be done by using scenario analysis as advocated by Strauss (2010).

CHAPTER 6

FUTURES MARKETS – FORECASTING FRIEND OR FOE?

6.1 INTRODUCTION

It seems that the futures market as a forecasting tool has won the forecasting contest by default. None of the other methods applied in the study, were able to generate realistic forecasts, with the exception of the single equation multivariate model, in Chapter Five. Although the model in Chapter Five proved exchange rate and world prices to be key determinants in the price formation of maize, it was not without its problems. The model cannot generate forecasts for the exogenous variables and can only be used for periods when equilibrium pricing takes place at either import or export parity. To a large extent, the poor nature of the preceding forecasts can be attributed to the fact that there is an associated futures market which complicates the modelling process severely.

The deduction that can be made from the above is that the EMH, as mentioned on p.4 cannot be rejected when applied to average monthly data. Conventional methods of testing market efficiency were not used to test this, but by implication, the fact that no model forecasted successfully might be considered as an indication that the market is efficient. In order to confirm this, a cointegration test of average monthly spot prices and average monthly futures prices was done. If the two series are found to be cointegrated, it can be concluded that the yellow maize market is weakly efficient. Stated in layman's terms, if the EMH holds, the series in question which is in this case average monthly yellow maize prices, is deemed not possible to forecast, and futures prices can be considered to be the best forecast available. According to Ehlers (2004), the key feature of well-functioning futures markets are their ability to predict prices at a specified future date, both efficiently and in an unbiased fashion.

The remainder of the chapter is devoted to a market efficiency test as applied by Ehlers (2004) and Carter and Mohapatra (2008). This is done to establish whether the market is

truly efficient, or if time series approaches are simply the incorrect method to apply for forecasting purposes. The implications of market efficiency on forecasting and decision making will also be discussed.

6.2 TESTING FOR AN EFFICIENT MARKET

According to Carter and Mohapatra (2008) futures prices as forecasts can be evaluated using the following criteria:

- Determine if futures prices is an unbiased forecast of cash prices.
- Determine if there exists a long term relationship between futures markets and cash markets.
- Determine if future prices lead to cash prices in the long run. Stated differently, futures prices can be used as an explanatory variable for cash prices.
- All information is aggregated into futures prices and there is no leftover information.

The above conditions are evaluated in the following sections, after a brief overview of the data has been given.

6.2.1 Data

Ehlers (2004) tested for market efficiency with daily spot prices and futures prices. The data had to be adapted slightly in order to apply the efficiency test to average monthly data. Ehlers (2004) notes that in order to test for market efficiency effectively, the gap between futures and spot prices should be held constant. This was done as far as possible but it should be noted that converting data to average monthly prices made it difficult. An average futures price had to be generated. Unlike spot prices, futures prices are not a continuous series. Scheepers (2003) advised that a time gap of 300 days should be used between spot and futures prices, for two reasons. Firstly, the production period for yellow maize is more or less 300 days. Secondly, if a time gap of more than 300 days is used, one will get breaks in the data series, since futures contracts might not

be available yet. The formulas to generate the futures price series, as used by Scheepers (2003) are as follows:

$$\text{Futures Price} = \text{Price of Nearest Futures Contract Expiring at Calculated Future Date} \quad (6.1.1)$$

where

$$\text{Calculated Future Date} = \text{Spot Date} + 300 \text{ Days} \quad (6.1.2)$$

What is important to note is that the above formula pertains to days. This has to be converted into average monthly prices in order to determine if the EMH applies to monthly data. 300 days translates into roughly 10 months. Equation 6.2 can therefore be rewritten as follows:

$$\text{Calculated Future Date} = \text{Spot Date} + 10 \text{ months} \quad (6.1.3)$$

6.2.2 Futures prices as unbiased forecast of subsequent cash prices

In order to test if future prices are in fact unbiased forecasts of subsequent cash prices the following simple linear regression model can be estimated:

$$SP_t = \alpha + \beta FP_{t-j} \quad (6.2.1)$$

where SP = Spot prices at time t when futures contract matures

FP = Future prices for a contract with maturity at time t, j months before maturity

Different time horizons for the above regressions were applied and are depicted in the tables below.

Table 6.1: Simple linear equation of Futures Prices 10 months before contract maturity, regressed on spot prices

Dependant Variable: Spot Price				
Method: Least Squares				
Sample: 1 to 50				
Included Observations: 50				
Variable	Coefficient	Std Error	t-statistic	Probability
FP(10)	0.694681	0.174387	3.983556	0.0002
C	369.2525	170.6989	2.163180	0.0355
Adjusted R-Square: 0.232801				

Table 6.2: Simple linear equation of Futures Prices 6 months before contract maturity, regressed on spot prices

Dependent Variable: Spot Price				
Method: Least Squares				
Sample: 1 to 50				
Included Observations: 50				
Variable	Coefficient	Std Error	t-statistic	Probability
FP(6)	1.0381	0.038092	27.25229	0.0000
Adjusted R-Square: 0.515281				

Table 6.3: Simple linear equation of Futures Prices 3 months before contract maturity, regressed on spot prices

Dependent Variable: Spot Price				
Method: Least Squares				
Sample: 1 to 50				
Included Observations: 50				
Variable	Coefficient	Std Error	t-statistic	Probability
FP(3)	1.021094	0.02254	45.30236	0.0000
Adjusted R-Square: 0.817377				

According to Kofi (1973) the coefficient of determination (R-square) in the above estimation can be regarded as a measure of market performance. A high R-Square is indicative of reliable predictions of futures prices as forecasts. The tables above show that futures prices 10 months before contract maturity has a very low coefficient of determination and it can therefore be concluded that the predictive power of these futures prices is poor. As the future price used for possible projection purposes comes closer to the maturity date, however, the R-square improves. With a six month forecasting horizon 51% of the variation in spot prices is explained by preceding futures prices and 81% of the variation in spot prices is explained by preceding futures prices over a three month forecasting horizon.

This section aims to establish whether the futures market provides an unbiased forecast of forthcoming spot prices. According to Carter and Mohapatra (2008) futures prices can be considered unbiased forecasts of subsequent cash prices if the following holds:

$$\alpha = 0 \text{ and } \beta = 1$$

If the tables above are regarded, it is apparent that the condition does hold for the three and six month futures price to maturity. The constant term for these two regressions was not found to be statistically different from zero and therefore excluded from the model. In addition to this the beta coefficients are approximately one. It can therefore be

concluded that futures prices of three and six months to maturity, are unbiased forecasts of subsequent cash prices and have adequate predictive power.

6.2.3 Cointegration – Testing for a long-term relationship

The first step, before cointegration tests can be conducted, is to determine the order of integration of the series to be tested for cointegration. It was established in Chapter Three, by using various unit root tests, that monthly spot price series is integrated of order one. The results of this test are represented in the table below.

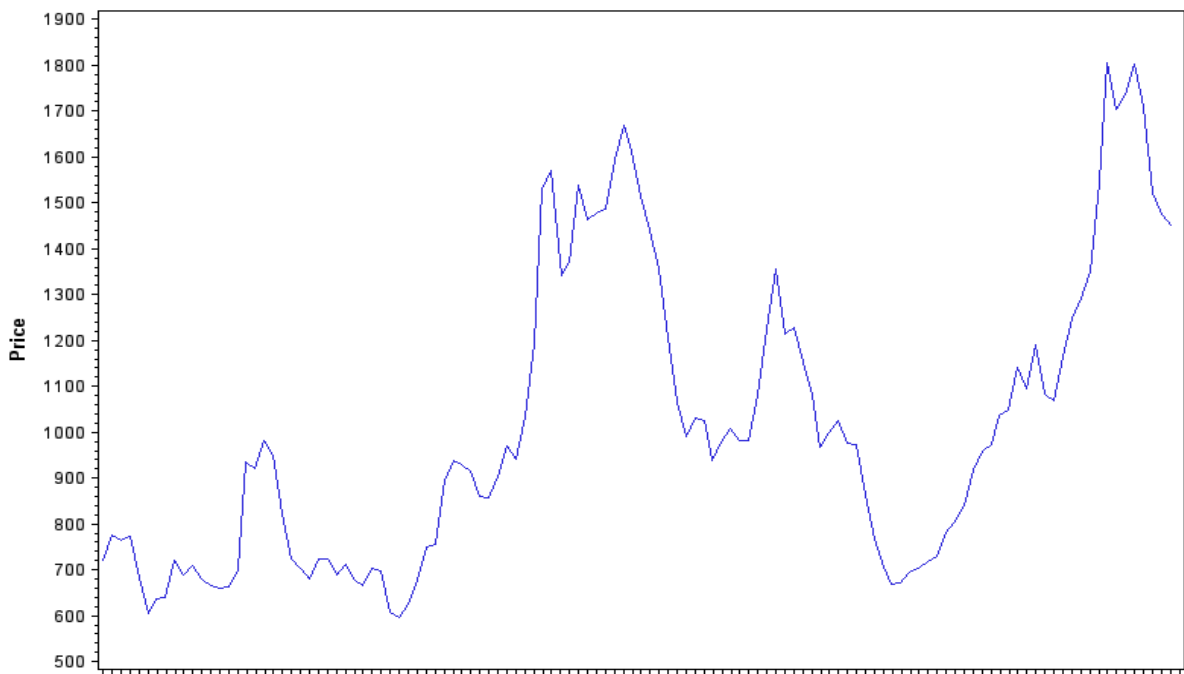


Figure 6.1: Monthly Futures Prices in Levels

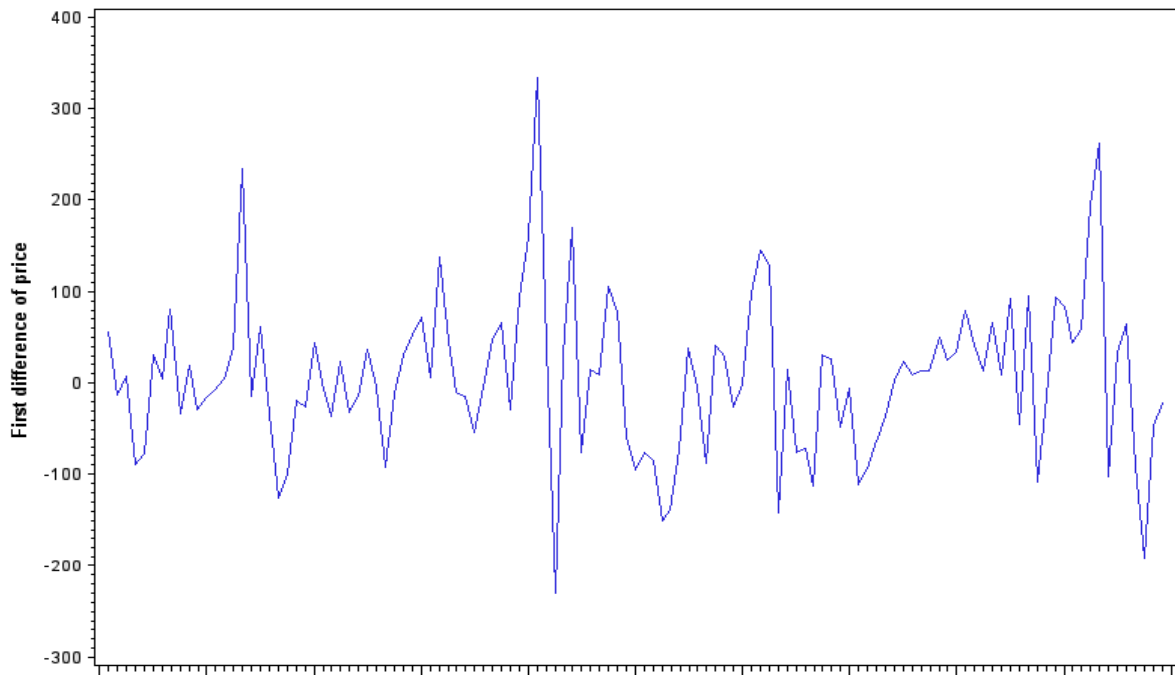


Figure 6.2: First difference of monthly futures prices

Table 6.4: Augmented Dickey Fuller Test for monthly yellow maize futures prices

Augmented Dickey-Fuller Unit Root Tests			
Type	Test Statistics and p-Values		
Single Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-3.7147 (0.5664)	-1.27 (0.6412)
	1	-7.4327 (0.2404)	-1.81 (0.3725)
	2	-6.6571 (0.2903)	-1.68 (0.4400)
	3	-9.2043 (0.1548)	-1.92 (0.3200)
Trend	0	-5.8711 (0.7476)	-1.69 (0.7485)
	1	-11.5109 (0.3235)	-2.35 (0.4037)
	2	-10.4078 (0.3915)	-2.17 (0.5035)
	3	-14.2900 (0.1916)	-2.43 (0.3605)

Table 6.5: Phillips-Perron Unit Root Test for Monthly Yellow Maize Futures Prices

Phillips-Perron Unit Root Tests			
Type	Test Statistics and p-Values		
Single Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-3.7414 (0.5664)	-1.27 (0.6412)
	1	-4.9761 (0.4308)	-1.50 (0.5318)
	2	-5.4526 (0.3862)	-1.57 (0.4926)
	3	-5.8282 (0.3538)	-1.63 (0.4629)
Trend	0	-5.7811 (0.7476)	-1.69 (0.7485)
	1	-7.6556 (0.5994)	-1.94 (0.6283)
	2	-8.3454 (0.5436)	-2.03 (0.5813)
	3	-8.8836 (0.5015)	-2.09 (0.5452)

Table 6.6: KPSS test results for monthly yellow maize futures prices

KPSS test statistics for the null of <u>single mean</u> stationarity of monthly futures prices of yellow maize				
Lag truncation Parameter	1	2	3	4
Test Statistics	1.9212*	1.311*	1.0072*	0.8267*

*** 10 % critical value = 0.347
 ** 5% critical value = 0.463
 * 1% critical value = 0.739

Table 6.7: KPSS test results for monthly yellow maize futures prices

KPSS test statistics for the null of <u>trend</u> stationarity of monthly futures prices of yellow maize				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.3910*	0.2680*	0.2096*	0.1706*

*** 10 % critical value = 0.347
 ** 5% critical value = 0.463
 * 1% critical value = 0.739

Table 6.8: Augmented Dickey Fuller Test for first differences of monthly yellow maize futures prices

Augmented Dickey-Fuller Unit Root Tests			
Type	Test Statistics and p-Values		
Zero Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-83.5659 (<0.0001)	-8.02 (<0.0001)
	1	-98.0329 (<0.0001)	-6.94 (<0.0001)
	2	-76.3350 (<0.0001)	-5.23 (<0.0001)
	3	-58.7933 (<0.0001)	-4.25 (<0.0001)

Table 6.9: Phillips-Perron Unit Root Test for First Differences of Monthly Yellow Maize Futures Prices

Phillips-Perron Unit Root Tests			
Type	Test Statistics and p-Values		
Zero Mean	Lags	Regression Coefficient Based Test	Studentized Test
	0	-83.5659 (<0.0001)	-8.02 (<0.0001)
	1	-85.0018 (<0.0001)	-8.05 (<0.0001)
	2	-81.8208 (<0.0001)	-7.98 (<0.0001)
	3	-81.1703 (<0.0001)	-7.97 (<0.0001)

Table 6.10: KPSS test results for first differences of monthly yellow maize futures prices

KPSS test statistics for the null of <u>zero mean</u> stationarity of monthly futures prices of yellow maize in first differences				
Lag truncation Parameter	1	2	3	4
Test Statistics	0.1679	0.1556	0.1474	0.1391

*** 10 % critical value = 1.195
 ** 5% critical value = 1.656
 * 1% critical value = 2.759

From the above, it can be concluded that average monthly spot prices and the constructed average monthly yellow maize futures prices are non-stationary and integrated of order one. Carter and Mohapatra (2008) argue that with non-stationary prices, cointegration theory provides a comprehensive approach by taking both long- and short run variation of spot and futures prices into account. The next step would therefore be to apply cointegration theory as explained by Engle and Granger (1987). The Engle-Granger test results are presented in the next table.

Table 6.11: Cointegration regression estimated for yellow maize contracts

Dependant Variable: LNSPOT				
Method: Least Squares				
Sample: 1997/11 – 2007/10				
Included observations: 120				
Variable	Coefficient	Std Error	t-stat	Prob
LNFUT	0.995713	0.001329	749.1603	0.000000
Adjusted R-Square: 0.916291				
Durbin-Watson Statistic: 0.272176				

Table 6.12: Engle-Granger Cointegration Test

Series	Model	Lags	ADF	Critical Values
Res_COINT	No Constant, No Trend	1	-3.087***	10% = -1.62 5% = -1.94 1% = -2.56

It can be seen from the above table, that there is cointegration between monthly spot prices and monthly futures prices for yellow maize. It can therefore be concluded that there exists a long run relationship between these variables. The results below depict the short run variation around the long-run relationship.

Table 6.13: Error Correction estimation results for yellow maize contracts

Dependent Variable: DLNSPOT				
Method: Least Squares				
Sample (adjusted): 1997/12 – 2007/10				
Included observations: 119 (after adjusting end points)				
Variable	Coefficient	Std Error	t-statistic	Probability
Res(-1)	-0.088379	0.0048289	-1.830204	0.0698
DLNFUT	0.831170	0.059437	13.98399	0.0000
Adjusted R-Square: 0.617554				
Durbin Watson Statistic: 1.403				

6.2.4 Determine if future prices lead to subsequent cash prices

The third requirement, according to Carter and Mohapatra (2008), for a market to be weakly efficient is for futures prices to be exogenous. Stated differently, future prices should *cause* subsequent cash prices. The conventional test for exogeneity is done by estimating a VECM, for futures and spot prices, and determining if the error correction term associated with the futures variable is zero. This is an indication that the series is unaffected by spot prices included in the system and that futures prices cause subsequent cash prices. Since the Johansen test and VECM estimation are cumbersome procedures, the Engle Granger single equation method was applied to test for cointegration and to estimate an error correction model, as depicted in the previous section. As stated in Chapter Five, this method assumes that variables included in the

long-run relationship are exogenous. For the purposes of the market efficiency discussion, we therefore assume that the futures price series is an exogenous variable and as a result, cause spot prices in the long run.

6.2.5 Is all available information incorporated into the futures price?

The last condition that future prices should adhere to, in order to regard the market as efficient, is that all available information should be aggregated into the futures price and that there is no outstanding information. According to Carter and Mohapatra (2008) this implies that there is no serial correlation in the residuals of the ECM as estimated in section 6.2.3. The degree of serial correlation is indicated by the Durbin Watson statistic. As mentioned earlier, a rough guideline is that a DW statistic of 2 signifies no serial correlation, values tending toward zero is indicative of positive serial correlation and values close to four can be regarded as showing negative serial correlation. The Durbin Watson statistic, as depicted in the estimation results of the above error correction estimation is 1.4, which is somewhat less than the rule neutral value of 2. It can therefore be concluded that there is some positive serial correlation present in the model. The implication of this is that futures prices do not incorporate all pertinent information and prediction power can be improved by incorporating additional information as contained in the residuals (Carter and Mohapatra, 2008).

6.3 CONCLUSION

Although not a conventional test of the EMH, the weak performance of the time series models in the previous sections alluded to the possibility that this hypothesis holds for monthly yellow maize prices. The presence of an associated futures market was to some extent, part of the problem of the poor models and forecasts in Chapter Three and Four, since it made it difficult to capture the underlying complexities, in the data generating process, with basic time series models. The poor forecasts generated by the time series models are however saved by the fact that there is a futures market. The possibility of market efficiency of monthly yellow maize prices was indicated by the poor

models in Chapter Three and Four. This was appropriately tested in the earlier section of this chapter with a method as applied by Carter and Mohoptra (2008). It was found that the market can be regarded as weakly efficient, based on the criteria as provided by the authors above. There are two issues, however, that should not be overlooked. Firstly, futures markets were assumed exogenous and no formal test was conducted to prove this. Future studies in markets efficiency might find it worthwhile to establish if this is in fact the case. Secondly, serial correlation was still present in the residual terms of the estimated ECM. This is an indication that all relevant information are not incorporated into the futures prices. This is also the reason for concluding that the market is weakly efficient.

The implications of these findings on forecasting and decision making are as follows:

- The EMH holds, and as a result futures, prices can be regarded as the best forecast of subsequent cash prices.
- Three month futures prices explain roughly 80 % of the variation in subsequent spot prices and a six month forecasting horizon describes approximately 50 % of the variation in subsequent spot prices. A three-month forecasting horizon can therefore be used with more confidence than a six month forecasting horizon. The question is: is such a short forecasting horizon useful? With a production and marketing period stretching over ten months, the ideal would have been a ten month projection period. However, it was found that the futures prices preceding cash prices by ten months are not unbiased predictors of the latter and their predictive powers are weak (roughly 20 percent).
- Future prices are quoted daily in circulatory literature. As a result, these “forecasts” are freely available to any market participant or decision maker. If decisions are based on these prices, this would in itself alter the path of prices and lead to forecasts that are unreliable. This, however, opens the door of behavioural finance, which falls beyond the scope of this study. Here, it would suffice to say that futures can be regarded as the most reliable forecast possible.
- Lastly, and closely related to the above point, is an argument made by Timmermann and Granger (2004). They argue that if it is possible to forecast

prices, market players would use it to generate unlimited profits. Logically, it can be assumed that as profitable opportunities become available and known, they are exploited and disappear after a while as more market players become aware of it and uses it. This conforms to the fact that financial markets are expected to be efficient. Stated differently, and as explained in the previous point, the fact that profitable scenarios can be forecasted would change the path of prices and make the initial forecasts obsolete. Timmermann and Granger (2004) therefore argue that efficient markets could be forecasted if a method is applied that could change quickly and that could detect instances of temporary forecastability that might arise and quickly disappear. A method that Timermann and Granger (2004), amongst others, names as a possible contender to satisfy the above are neural networks. Neural networks are emphasized here since the concept was already introduced in Chapter Two. The other methods proposed by Timermann and Granger (2004) ranges from non-linear models to predictive density functions. All these methods provide opportunities to forecast without violating the EMH.

CHAPTER 7

CONCLUDING REMARKS

An economist is an expert who will know tomorrow, why the things he predicted yesterday didn't happen today.

~ Peter J Laurence (Canadian Academic and Writer)

7.1 OVERVIEW

The objective of this study was to attempt to find a method, which is relatively simple and easily updateable, to generate monthly forecasts for South African yellow maize prices. Maize was chosen due to the strategic role it plays in South African agriculture. Maize is a key driver of numerous agricultural prices in the role that it plays as an input and a competitive product. Yellow maize was chosen specifically, because of its interrelatedness with world markets. As pointed out in Chapter One, monthly forecasts have the potential to aid agricultural decision makers with short(er)-term decision making. Currently, there are annual models that facilitate strategic decision making such as the BFAP structural econometric method. There are, however, no models or literature pertaining to agricultural forecasts in a higher frequency.

A literature study on the various types of methods as applied to price forecasting in agriculture is represented in Chapter Two. This study identified numerous possible methods which can be used to generate price forecasts. Due to the fact that limited international and macro economic data is available in a monthly frequency, the study steered away from structural models based on fundamental economic theories. The focus is consequently on technical time series models that are based on the properties of underlying data generating processes of the variable under consideration.

Based on the findings of the literature study and the discussed data issues, two time series approaches are applied. The first is a univariate model estimated according to the Box-Jenkins methodology. The forecasts generated with this model are, however, a

persisting mean level over the whole forecasting period. This is unrealistic and after some consideration, it more or less gives the same forecasting result as an annual model which forecasts the average annual price over a period of twelve months. The second method applied is a multivariate model, as depicted in Chapter Four. The model that is estimated is a Vector Error Correction, model based on a cointegrating relationship found with the Johansen test for cointegration. Finding a cointegrating relationship between South African maize prices, world maize prices and the exchange rate proved to be a tall order. Cointegration was found at a 10 % level and a VECM was constructed with dismal forecasting results. The reason for this proved to be the fact that there is not a simultaneous relationship between the variables. World prices and exchange rates should rather be treated as exogenous variables. The implication of this is that a vector autoregressive or vector error correction approach is not suitable, since no other fundamental data of factors affecting maize prices are available in adequate monthly format. The forecasts of the VECM show little improvement on the univariate model since, it also only reverts to a mean level.

Chapter Five was originally earmarked for a forecast comparison. As a result of the unrealistic nature of the forecasts the chapter is rather devoted to a discussion of possible reasons for the poor performance of the models and forecasts. The main cause of the poor forecasts identified is that yellow maize prices, world prices and the rand-dollar exchange rate have associated futures markets. This impacts adversely on the univariate properties of the variables in question and make it extremely difficult to model. According to Aydemir in Knight and Satchell (1998), financial time series, such as those mentioned above, exhibit certain patterns such as fat tails and volatility clustering. Svetlozar and Fabiozzi (2007) further mention that financial time series are frequently characterized by a non-linear data generating process, which renders conventional linear time series models ineffective. It might be worthwhile to note that these problems were not identified at the start of the study, since it was never the objective to aid investment decision and identify profit seeking opportunities. As a result no financial econometrics literature were reviewed. Only after the dismal forecasting results financial econometric literature was regarded to look for possible reasons. The problem

associated with the multivariate model was the difficulty to obtain cointegration between the variables (South African prices, world prices and rand-dollar exchange rates). As a result of the limited data available in a monthly frequency, no other fundamental variables could have been included in the model. The reason for the lack of cointegration between the variables mentioned above can be attributed to the fact that there is regime switching in the price formation mechanism of yellow maize. The various regimes, as mentioned in Chapter Five, are import parity, export parity and autarky. Exchange rates and world prices are key drivers of local maize prices *only* when the prices are at parity levels. Of the 120 observations that were originally used, 62 observations were identified to be close to or at parity levels. These observations were tested for cointegration by making use of a single equation cointegration test as opposed to a vector approach as used in Chapter Four. This addressed the problem that exchange rate and world prices were found to be exogenous. Cointegration is obtained between the variables if the 62 observations, as discussed above, are used. The associated error correction model was also estimated with relative success. There are however, still two major forecasting problems associated with this modelling approach. The first is that the decision maker would have to know in advance whether prices are going to trade at import and export parity. The above mentioned model is not usable at autarky. The second is that since exchange rate and world price are regarded as exogenous, forecasts for these variables have to be obtained for them from external sources or by deduction. A solution to this problem is the scenario analysis approach as applied by Strauss (2010). This can, however, be regarded as qualitative forecasting and falls outside of the scope of this study.

Chapter Six resulted from the fact that the futures market served as the best method for obtaining forecasts. The above conclusion alluded to the possibility that the EMH does hold. A chapter was therefore devoted to determine, with conventional EMH tests, whether this is in fact the case. The test applied shows that the market can be considered as weakly efficient. The question that stems from this is: If futures markets are the best forecasts for subsequent cash prices, exactly how good are futures prices as forecasts? It was found that for shorter forecasting horizons the South African yellow

maize futures price could provide an adequate forecast of subsequent cash prices. For longer time horizons, however, this was not the case. This makes the usefulness of such forecasts doubtful, since longer forecasting horizons are what is needed to aid production and procurement decisions. In addition to this, futures prices are seen as public information. Decisions based on this might in itself change the path of prices and make that subsequent cash prices differ greatly from those forecasted and used in decision making.

7.3 FINDINGS OF THE STUDY

After reading the preceding sections, the reader, as the author at a certain point in time, might wonder why the models in Chapter Three and Four were applied in the first place, if it is apparent from the topics discussed above that it is not a sensible exercise. The reason for this stems from the fact that forecasting is a subject matter that stands with legs in different, and frequently far removed, study fields. This is even more true when it comes to forecasting prices of a commodity with an associated derivative market. Even though the literature study, at the start of the study, was done as complete as possible within the available time, the scope was focused on agricultural economics since this is, first and foremost, a study in agricultural economics. In addition to this, the objectives of the study guided it towards a decision making tool for all market participants and not, as mentioned previously, a profit seeking tool for investors. The literature gave no indication of the severity of the problems that one might encounter when developing and forecasting with the time series models discussed. Not one comparative study reviewed included diagnostic testing on estimated models. It should also be noted that numerous comparison forecasting studies, included in the literature study, had an associated futures market. The effect of this on price forecasting was never discussed. At best, futures prices were included as a forecasting method. It can therefore be concluded that there is a large void in Agricultural Economics with respect to time series models that did not work for forecasting purposes and the reasons for it. To some extent, this is to be expected since authors might be reluctant to publish models and forecasts that do not produce positive results.

Despite the results in Chapter Three and Four, the study still has contributions to South African Agricultural Economics, specifically in the forecasting milieu. These contributions are set out below

- It is expected that the usability of the price forecasts depends, to a large extent on the length of the production period of the commodity or product in question. For yellow maize, with a production cycle of roughly ten months, accurate monthly forecasts ten months ahead are needed. Time series models were not able to generate such forecasts. However, if one would regard a product with a shorter production cycle, such as certain vegetables, time series models might be sufficient. This needs to be tested empirically.
- The study provided ample exposure to various econometric methods and concepts. This is something in which South African agricultural economics scholars often do not have a lot of experience in. It is as a result of the hard work that went into the understanding and estimation of the models in Chapter Three and Four that it was decided to persist with it rather than discard it for another method or study with more appealing results.
- The study highlights the need for more diligent record keeping. Monthly data on maize deliveries, utilization and stocks are only available from May 2000. This significantly reduces the time series that are available for the development of the multivariate model in Chapter Four, especially when seasonal factors have to be accounted for. The need for monthly South African field crop data was also emphasized by Ehlers and Frick (2002) in an article on South African Agricultural data needs. In addition to agricultural data needs, macro- economic data such as per capita GDP (which is a key variable in a demand function) in a monthly frequency might enable the model builder, interested in monthly forecasts, to move away from a-theoretical time series models to structural models such as BFAP's annual models.
- The study provided the opportunity for the synthesis of various subject matters. Themes that are usually conducted in isolation, such as market efficiency, financial econometrics, agricultural practicalities and price formation were combined into one study.

- There is no recipe for agricultural price forecasting that one can follow in order to obtain adequate results. Although it was not attempted to forecast prices of other products, it is expected that the forecasting procedure needs to be custom fitted for different products in different regions.

7.4 FUTURE RESEARCH

As mentioned in several of the preceding chapters, neural networks might provide the solutions to some of the problems encountered in the study. To the knowledge of the author, this approach has not been applied to agricultural price forecasting in South Africa. Due to the possibilities that this method hold it, might be worthwhile to explore it in future forecasting studies.

Timmermann and Granger (2004) mention various methods that can be applied to generate forecasts, without violating the EMH. These methods include bootstrapping techniques and predictive density methods. These are, however, highly advanced topics in forecasting and should consequently be investigated only after sufficient knowledge in this field has been obtained.

It is not to claim that the methods applied in this study would not work for other products or commodities. This study has shown that time series methods are not appropriate for South African commodities with a futures market, for the reasons as mentioned in Chapter Five and Six. These methods might, however, work for products or commodities where prices are formed in another way, for instance at fresh produce markets and abattoirs. Policy and trade measures are also expected to impact in the price formation process, and as result, these factors might influence the ability to successfully forecast prices.

This study might be seen as negligible in terms of agricultural forecasting in general. Forecasting studies that obtained more successful results might be attributed to scholars with a stronger econometric/mathematical foundation that enabled them to account for

factors such as non-linearity and non-normality. Sadly, South-Africa's econometric capacity, pertaining to time series econometrics, is limited when compared to countries such as the United States. This is particularly true in the Agricultural Economics subject field. Another factor that might be worthwhile to consider is that other world regions might have less price variability, due to policy measures set in place. As a result, forecasting with time series methods, applied to such prices, might be more successful.

It is still the opinion of the author that quarterly, monthly and weekly price forecasting have a key role to play in short(er)-term decision making. The feasibility thereof should however be determined empirically. It should be noted that, despite of the issues alluded to above, this study serves as a starting point for future higher frequency agricultural price forecasting studies in the South African context. Something that has, to date, not been addressed at all.

References

Ahmad, H.A., G.V. Dozier and D. A. Roland .2001. *Egg Price Forecasting Using Neural Networks*. Poultry Science Association Inc 2001.pp. 162-171.

Allen, P.G.1994. *Economic Forecasting in Agriculture*. International Journal of Forecasting. Vol. 10(1). pp.81-135.

Armstrong, J.S. and F. Collopy. 1998. *Integration of statistical forecasts and judgment for time series forecasting: Principles from Empirical Research*. Editors: G. Wright and P.Goodwin . Forecasting with Judgement. John Wiley and Sons.

Arzac, E.R. and M. Wilkinson .1979. *A Quarterly Econometric Model of the United States Livestock and Feed Grain Markets and Some of its Policy Implications*. American Journal of Agricultural Economics. Vol. 61(2). pp. 297-308.

Bessler, D.A. and J. A. Brandt. 1992. *An analysis of livestock prices*. Journal of Economic Behaviour and Organization. Vol 18. pp. 249-263.

BFAP .2009. Scenario Analasis: Soccer World Cup 2010.

www.bfap.co.za

Bowerman, B.L, A Khoeler and R O'Connel. 2005. *Forecasting, Times Series and Regression – 4th Edition*. Duxbury Applied Series. Belmont, CA.

Bowman, C and A.M. Husain. 2006. *Agricultural Commodity Markets and Trade – New approaches to analyzing Market Structure and Instability*. Edited by Alexander Sarris and David Hallam. Published by the Food and Agricultural Organization of the united nations.

Brandt, J.A. and D.A. Bessler. 1981. *Composite Forecasting: An application to U.S Hog Prices*. American Journal of Agricultural Economics. Vol 63(1). pp. 135-140

Brandt, J.A. and D.A. Bessler. 1983. *Price Forecasting and Evaluation: An Application in Agriculture*. Journal of Forecasting. Vol.2. pp. 237-248

Brandt, J.A. and D.A. Bessler. 1984. *Forecasting with Vector Autoregressions versus a Univariate ARIMA process: An Empirical Example to U.S. Hog Prices*. North Central Journal of Agricultural Economics. Vol.6(2). pp. 29-36

Carter, C.A. and S. Mohapatra. 2008. *How Reliable are hog futures as forecasts?* American Journal of Agricultural Economics. Vol. 90(2). pp. 367-378.

Chen, D.T. and D.A. Bessler. 1990. *Forecasting Monthly Cotton Price: A structural and Time Series Approach*. International Journal of Forecasting. Vol 6. pp. 103-113.

Colino, E.V. 2008. *How much can outlook forecasts be improved? An application to the U.S. hog market*. Proceedings of the NCCC-134. Conference on applied commodity price analysis, forecasting and risk management. St. Louis, MO.
<http://www.farmdoc.uiuc.edu/ncc134>

Ehlers, A.C and A. Frick. (2002). *The need for Agricultural Data in South Africa: with Specific Reference to the Western Cape Province*. Paper prepared for presentation at the 13th International Farm Management Congress. Wageningen.

Ehlers, A.C. 2004. *Testing the efficiency of the South African Derivatives market for yellow maize*. Submitted in partial fulfilment of a master's in business administration degree. University of Pretoria, Pretoria.

Emmett, W.E. and S.H. Holder. 1985. *An Evaluation of the Rice Outlook and Situation Price Forecasts*. Southern Journal of Agricultural Economics. December, 1985. pp.155-161.

Fama, E.F. 1970. *Efficient capital markets: A review of theory and empirical work*. Journal of Finance. Vol 25. pp. 383-417.

Ferris, J. 2005. *Agricultural Prices and Commodity Market Analysis*. MacGrawHill.

Freebairn, J.W. 1975. *Forecasting for Australian Agriculture*. Australian Journal of Agricultural Economics. Vol. 19. pp. 154-174

Gardner, E.S. 1985. *Exponential Smoothing: The State of the Art*. Journal of Forecasting. Vol. 4(1). pp. 1-28.

Gellatly, C. 1979. *Forecasting New South Wales Beef Production: An Evaluation of Alternative Techniques*. Review of Marketing and Agricultural Economics. Vol. 47(2). pp. 81- 94.

Geyser, M. and M. Cutts. 2007. *SAFEX maize price volatility*. Agrekon. Vol46(3). pp. 291-305

Gil, J.M. and L.M. Albisu 1993. *Composite Forecasting methods: An application to Spanish maize prices*. Journal of Agricultural Economics. Vol 44. pp. 264-271

Granger, C.W.J.1992. *Forecasting Stock Market Prices: Lessons for Forecasters*. International Journal of Forecasting. Vol. 8,pp. 3-13.

Granger, C.W.J and S. Poon. 2003. *Volatility in Financial Markets: A Review*. Journal of Economic Literature. Vol 41 (2). pp. 478-539

Gujurati, D.N. 2003. *Basic Econometrics – 4th Edition*. McGrawHill. Boston.

Gjolberg, O. 1997. *Forecasting Quarterly Hog Prices*. *Agribusiness*. Vol. 13(6). pp. 673-679.

Guttormsen, A.G. 1999. *Forecasting Weekly Salmon Prices: Risk Management in Fish Farming*. *Aquaculture Economics and Management*. Vol. 3(2). pp.159-166

Hanke, J.E. and D.W. Wichern. 2005. *Business Forecasting – 8th Edition*. Pearson Prentice Hall. New Jersey.

Harris, S.K. and R.M. Leuthold. 1985. *A Comparison of Alternative Forecasting Techniques for Livestock Prices: A Case Study*. *North Central Journal of Agricultural Economics*. Vol 7(1). pp. 40-50.

Holton, J. and B. Keating. *Business Forecasting – 5th Edition*. MacGraw-Hill International Edition. Boston.

Jordaan H., B. Grove, A. Jooste and Z.G. Alemu. 2007. *Measuring the Price Volatility of Certain Field Crops in South Africa using the ARCH/GARCH Approach*. *Agrekon*, Vol.46(3). pp. 306-322.

Just, R.E. and G.C. Rausser.1981. *Commodity Price Forecasting with Large-Scale Econometric Models and the Futures Market*. *American Journal of Agricultural Economics*. Vol.63(2). pp. 197-208.

Kastens, T.L., R. Jones and T.C. Schroeder .1998. *Futures-Based Price Forecasts for Agricultural Producers and Businesses*. *Journal of Agricultural and Resource Economics*. Vol 23(1). pp. 294-307.

- Kastens, T.L, T.C Schroeder and R. Plain. 1998. *Evaluation of Extension and USDA Price and Production Forecasts*. Journal of Agricultural and Resource Economics. Vol 23(1). pp. 244-261.
- Kaylen, M.S. 1988. *Vector Autoregression Forecasting Models: Recent Developments Applied to the U.S. Hog Market*. American Journal of Agricultural Economics. Vol.70(3). pp. 701-712.
- Kenyon, D. , E Jones and A McGuirk. 1993. *Forecasting Performance of Corn and Soybean Harvest Futures Contracts*. American Journal of Agricultural Economics. Vol 75(2). pp. 399-407.
- Ke-Young, C. 1978. *Short-run Forecasting of Commodity Prices: An application of Autoregressive Moving Average Models*. International Monetary Fund Staff Papers. Vol. 25(1). pp. 90-111.
- Knight, J. and S. Satchell. 1998. *Forecasting volatility in the financial markets*. Butterworth-Heinemann Publishers. Oxford.
- Kofi, T.A. (1973). *A framework for comparing the efficiency of futures markets*. American Journal of Agricultural Economics. Vol 55(4). pp. 584-594.
- Kumar, K. (2006) . *The role of artificial neural networks in agricultural research*.
<http://www.iasri.res.in/icsi2006/theme2/kuldeep.pdf>
- Kwiatkowski, D., P.C.B. Phillips, P. Schmidt and Y. Shin. 1992. *Testing the null hypothesis of stationarity against the alternative of a unit root*. Journal of Econometrics. Vol. 54. pp. 159-179.

Leuthold, R.M. and P.A. Hartman.1979. *A Semi-Strong Evaluation of the Efficiency of the Hog Futures Market*. American Journal of Agricultural Economics. Vol. 61(3). pp. 482-489

Liew, V.K.S, M. Shittan and H. Hussain. 2000. Time series modelling and forecasting of Sarawak Black Pepper price. Online available at <http://mpra.ub.uni.muenchen.de/791>. Posted November 2007.

Lux, T. and M. Marchesi. 1998. *Volatility Clustering in Financial Markets: A Micro-Simulation of Interactive Agents*.

<http://www.ge.infm.it/~ecph/papers/marchesipapers/microsimul.pdf>

Makridakis, S and G. Spyros. 1998. *Forecasting: Methods and Applications – 3rd Edition*. John Wiley and Sons Inc. New York.

Mayer, J. and D. Ulbricht. 2007. *Log vs. Level in VAR forecasting: 16 million empirical answers – expect the unexpected*. IFO working papers series no.42. Munich Germany. <http://citeseerx.ist.psu.edu>.

McKenzie, E. 1984. *General Exponential Smoothing and the Equivalent ARMA process*. *Journal of Forecasting*. Vol 3. pp. 449-458

McLeod, A.I and W.K Li. (1983). *Diagnostic Checking ARMA time series models using squared residual autocorrelations*. *Journal of time series analysis*. Vol 4. pp. 269-273.

Mendenhall, W and T. Sincich. 2003. *A second course in statistics: Regression analysis – 6th Edition*. New York.

Meyer, F.H., P. Westhoff, J. Binfield and J.F. Kirsten. 2006. *Model closure and price formations under switching market regimes*. *Agrekon*. Vol. 45(4). pp. 370-380.

- Meyer, F.H, P.G Strauss and T.B Funke 2008. *Modelling the impacts of macro-economic variables on the South African bio-fuels industry*. *Agrekon*.Vol47(3). pp. 327-345.
- Moholwa, M.B.2005. *Testing for weak-form efficiency in South African Futures Markets for white and yellow maize*. Unpublished Paper. Department of agricultural Economics, Michigan State University, USA.
- Newbold, P. and C.W.J Granger. 1974. *Experience with Forecasting Univariate Time Series and the Combination of Forecasts*. *Journal of the Royal Statistical Society: Series A*. Vol. 137(2). Pp. 131-165.
- Oliveira, R.A. et al 1979. *Short-Run Forecasting Model of Beef Prices*. *Western Journal of Agricultural Economics*. Vol.4. pp. 45-56
- Park, D.W. and W.G. Tomek.1988. *An appraisal of composite forecasting methods*. *Applied Economic Perspectives and Policy*. Vol.10(1). pp.1-11.
- Pindyck, R.S. and D.L. Rubinfeld. 1998. *Econometric Models and Forecasting – 4th Edition*. MacGraw-Hill. Boston.
- Pope, D.P, R.A Kramer and R.D Green (1979). *An Evaluation of econometric methods of U.S. farmland prices*. *Western Journal of Agricultural Economics*. Vol. 4(1). Pp107-120.
- Roche, J.R. 1995. *Forecasting Commodity Markets*. Probus Publishing. London.
- Scheepers, D. 2003. Unpublished MSc thesis. University of Pretoria, Pretoria.
- Skaggs, R.K. and D.L. Snyders. 1992. *The comparison of Selected Methods for Forecasting Monthly Alfalfa Hay Prices*. *Agribusiness*. Vol. 8(4). pp. 309-321.

Spreen, T.H., R Mayer and J.R. Simpson. 1979. *Forecasting Monthly Slaughter Cow Prices with a Subset Autoregressive Model*. Southern Journal of Agricultural Economics. July 1979.

Svetlozar, T.R. and F.J Fabiozzi. 2007. *Financial Econometrics – From Basics to advances Modelling Techniques*. John Wiley and Sons. New Jersey.

Strauss, P.G. 2010. *Scenario thinking and stochastic modelling for strategic and policy decision in agriculture*. PhD Dissertation. University of Pretoria.

Timmermann, A. and C.W.J Granger. 2004. Efficient market hypothesis and forecasting. Journal of Forecasting. Vol 20. pp. 15-27.

Tomek, W and G. Myers. *Empirical analysis of Agricultural Commodity Prices: A Viewpoint*. Review of Agricultural Economics, Vol. 15 no.1 (Jan 1993) pp.181-102.

Van Staden, P.J (2008). *Mathematical Statistics 321: Course Notes 2008*. University of Pretoria, Pretoria.

Wei, A. and R. M. Leuthold. 1998. *Long agricultural futures prices: ARCH, long memory, or chaos process*. OFOR paper number 98-03.

Wilson, C.A. 1993. *Award Winning Undergraduate Paper: Feeder Cattle Forecasting Models – An Econometric Study of Development and Performance*. American Journal of Agricultural Economics. Vol 75(5). pp. 1326-1332.

Wiseman, J.A, M.A.G. Darroch and G.F. Ortman.1999. *Testing the efficiency of the South African futures market for white maize*. Agrekon. Vol38(3). pp.321-335.

Zapata, H.O. and P. Garcia. 1990. *Price Forecasting with Time-Series Methods and Nonstationary Data: An Application to Monthly US cattle prices*. Western Journal of Agricultural Economics. Vol. 15(4). pp. 123-132.

Zulauf, C.R., S.H Irwin and J.E Ropp. 1993. *A Reappraisal of the Forecasting Performance of Corn and Soybean New Crop Futures*. The Journal of Futures Markets. Vol 19(5). pp. 603-618.

Zou, H.F., G.P. Xia and F.T Yang. 2007. *An investigation and comparison of artificial neural networks and time series models for Chinese food grain price forecasting*. Neural Computing. Vol 70. pp. 2913-2923.

Appendix 1

Table A1.1: Correlogram of Autocorrelation Function of First Differences Maize Price

Autocorrelations				
Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1	Std Error
0	7201.404	1.00000	*****	0
1	2192.685	0.30448	.*****	0.090536
2	535.175	0.07432	. .*****	0.098572
3	639.016	0.08873	. .*****	0.099031
4	-603.615	-.08382	. .*****	0.099680
5	-284.294	-.03948	. .*****	0.100256
6	-143.121	-.01987	. .*****	0.100384
7	-734.857	-.10204	. .*****	0.100416
8	1038.140	0.14416	. .*****	0.101262
9	686.412	0.09532	. .*****	0.102931
10	-69.571472	-.00966	. .*****	0.103652
11	-27.091248	-.00376	. .*****	0.103659
12	-142.214	-.01975	. .*****	0.103660
13	-637.488	-.08852	. .*****	0.103691
14	20.009934	0.00278	. .*****	0.104309
15	-754.382	-.10475	. .*****	0.104309
16	-1065.727	-.14799	. .*****	0.105168
17	-454.076	-.06305	. .*****	0.106861
18	-737.831	-.10246	. .*****	0.107166
19	-589.493	-.08186	. .*****	0.107966
20	-225.701	-.03134	. .*****	0.108473
21	-531.767	-.07384	. .*****	0.108547
22	230.678	0.03203	. .*****	0.108958
23	871.912	0.12108	. .*****	0.109036
24	752.549	0.10450	. .*****	0.110132

Table A1.2: Correlogram of Partial Autocorrelation Function for Difference Maize Price

Partial Autocorrelations		
Lag	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
1	0.30448	. *****
2	-0.02027	. .
3	0.07919	. ** .
4	-0.14680	. *** .
5	0.03135	. * .
6	-0.02346	. .
7	-0.07801	. ** .
8	0.21620	. ****
9	-0.02129	. .
10	-0.02463	. .
11	-0.05435	. * .
12	0.02405	. .
13	-0.07994	. ** .
14	0.05271	. * .
15	-0.10143	. ** .
16	-0.10851	. ** .
17	-0.02798	. * .
18	-0.06734	. * .
19	-0.01025	. .
20	-0.04533	. * .
21	-0.01804	. .
22	0.03204	. * .
23	0.10993	. ** .
24	0.09230	. ** .

Table A1.3: SCAN Test Probability Values for Differenced Maize Prices

SCAN Chi-Square[1] Probability Values					
Lags	MA 0	MA 1	MA 2	MA 3	MA 4
AR 0	0.0005	0.4463	0.3652	0.3949	0.6875
AR 1	0.7962	0.3889	0.4065	0.5916	0.9789
AR 2	0.3652	0.8945	0.3388	0.6989	0.5508
AR 3	0.0817	0.3994	0.5369	0.4768	0.5958
AR 4	0.6356	0.8282	0.4268	0.5242	0.4761

Table A1.4: ESACF Probability Values for Differenced Maize Prices

ESACF Probability Values					
Lags	MA 0	MA 1	MA 2	MA 3	MA 4
AR 0	0.0008	0.4509	0.3702	0.4004	0.6938
AR 1	0.3938	0.0824	0.5930	0.1978	0.9912
AR 2	0.0024	0.6934	0.2242	0.1829	0.8045
AR 3	<.0001	0.1795	0.0011	0.2377	0.4702
AR 4	0.0052	0.1739	0.0005	0.1090	0.5294

Table A1.5: MINIC Criterion for Differenced Monthly Maize Prices

Minimum Information Criterion					
Lags	MA 0	MA 1	MA 2	MA 3	MA 4
AR 0	8.853217	8.785261	8.820682	8.841325	8.876639
AR 1	8.798016	8.824388	8.858463	8.874335	8.912008
AR 2	8.836456	8.852463	8.890801	8.911457	8.949653
AR 3	8.866333	8.882324	8.917331	8.950833	8.988718
AR 4	8.880179	8.904071	8.940098	8.977656	9.005538

Table A1.6: Autocorrelation Function Correlogram of Residuals of MA(1) process

Autocorrelation Plot of Residuals				
Lag	Covariance	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1	Std Error
0	6619.372	1.00000	*****	0
1	-17.756145	-.00268	. .	0.090536
2	241.132	0.03643	. * .	0.090536
3	754.565	0.11399	. ** .	0.090656
4	-751.594	-.11354	. ** .	0.091824
5	-95.830758	-.01448	. .	0.092968
6	263.752	0.03985	. * .	0.092986
7	-1126.550	-.17019	. *** .	0.093126
8	1196.872	0.18081	. **** .	0.095641
9	343.409	0.05188	. * .	0.098403
10	-239.793	-.03623	. * .	0.098627
11	79.775370	0.01205	. .	0.098736
12	86.275625	0.01303	. .	0.098748
13	-753.021	-.11376	. ** .	0.098762
14	428.162	0.06468	. * .	0.099831
15	-536.874	-.08111	. ** .	0.100174
16	-813.724	-.12293	. ** .	0.100711
17	5.945899	0.00090	. .	0.101933
18	-576.896	-.08715	. ** .	0.101933
19	-385.305	-.05821	. * .	0.102542
20	100.859	0.01524	. .	0.102813
21	-550.968	-.08324	. ** .	0.102831
22	164.982	0.02492	. .	0.103382
23	698.298	0.10549	. ** .	0.103431
24	211.217	0.03191	. * .	0.104309

Table A1.7: Partial Aurocorrelation Function Correlogram of Residuals of MA(1) process.

Partial Autocorrelations		
Lag	Correlation	-1 9 8 7 6 5 4 3 2 1 0 1 2 3 4 5 6 7 8 9 1
1	-0.00268	. .
2	0.03642	. *
3	0.11434	. **
4	-0.11547	. **
5	-0.02351	. .
6	0.03697	. *
7	-0.14719	. ***
8	0.17986	. ****
9	0.04941	. *
10	-0.01443	. .
11	-0.06273	. *
12	0.03587	. *
13	-0.08765	. **
14	0.03766	. *
15	-0.03396	. *
16	-0.12813	. ***
17	-0.04149	. *
18	-0.06944	. *
19	-0.01701	. .
20	-0.03921	. *
21	-0.02638	. *
22	-0.00241	. .
23	0.08978	. **
24	0.08469	. **

Appendix 2(a)

Table A2.1: Long-run Cointegration Equation for full time series (Dependant variable: SA Average Maize Price).

Variable	Coefficient
World Price	5.083
Exchange Rate	142.23
Constant	-1241.205

Table A2.2: Engle Granger Cointegration test for full time series.

Series	Model	Lags	ADF	Critical Value (5%)
Residual Cointegration Equation	Constant, No Trend	1	-3.77	-3.91

Appendix 2(b)

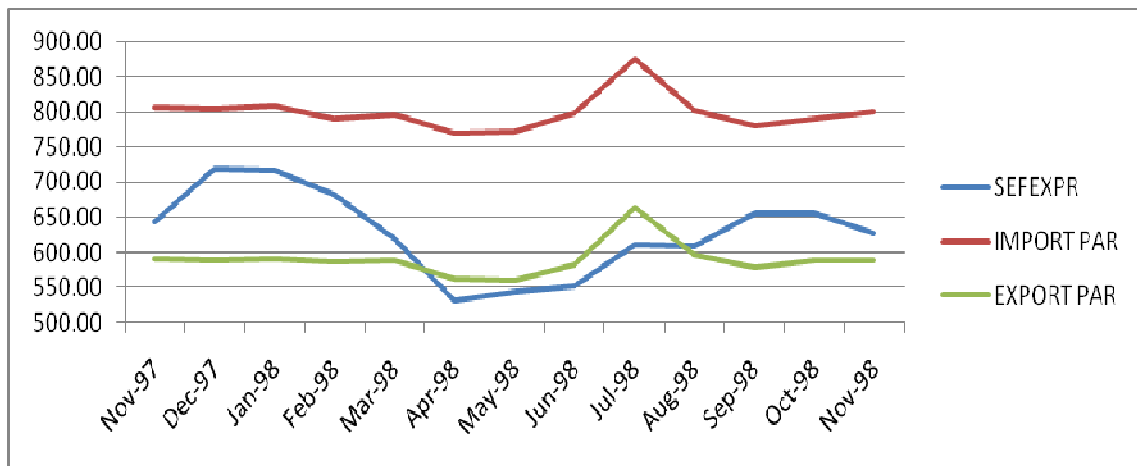


Figure A2.1: Average monthly SAFEX price with respect to import and export parity levels (11/1997-11/1998)

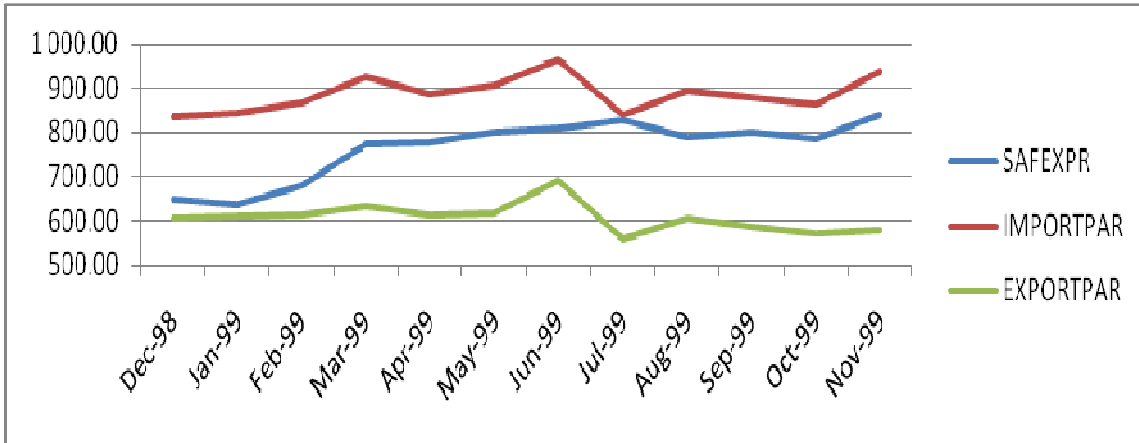


Figure A2.2: Average monthly SAFEX price with respect to import and export parity levels (12/1998-11/1999)

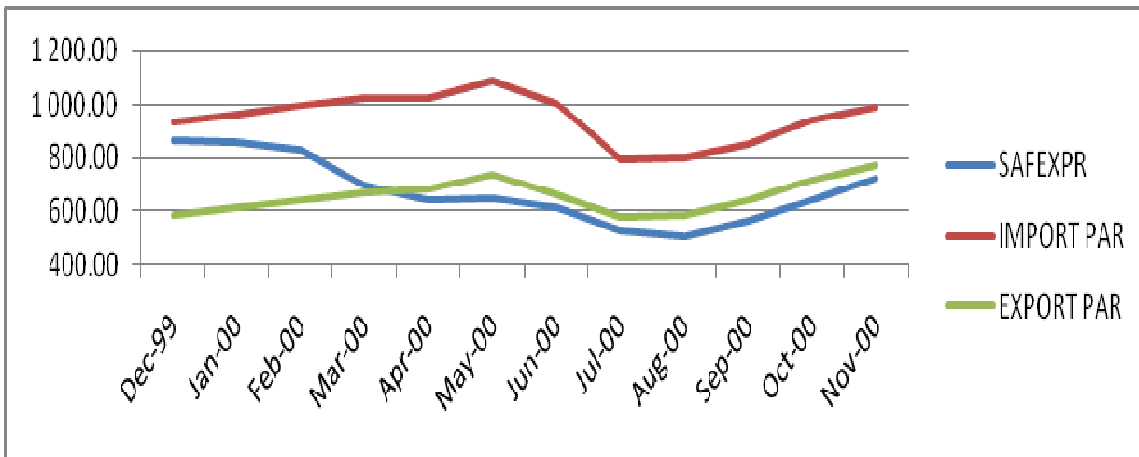


Figure A2.3: Average monthly SAFEX price with respect to import and export parity levels (12/1999-11/2000)

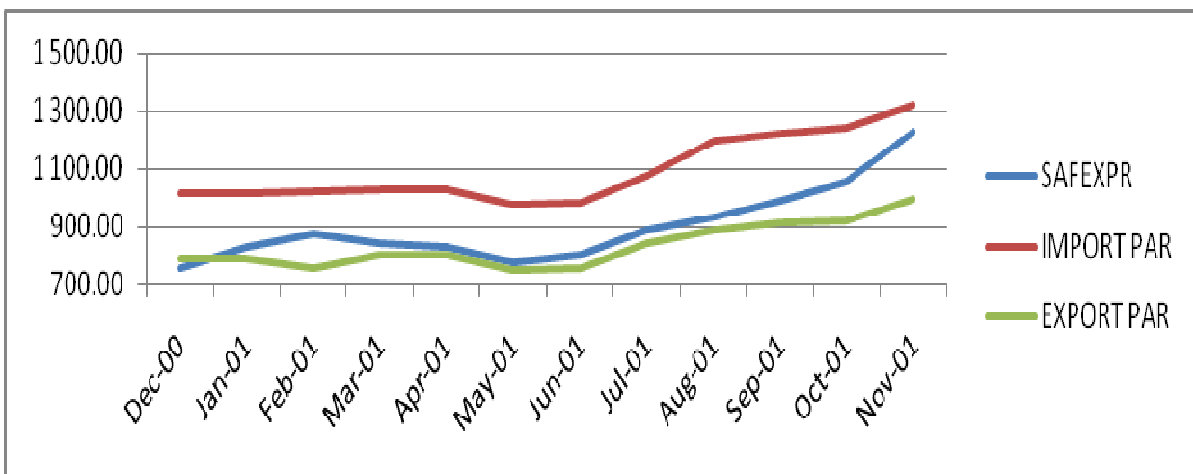


Figure A2.4: Average monthly SAFEX price with respect to import and export parity levels (12/2000-11/2001)

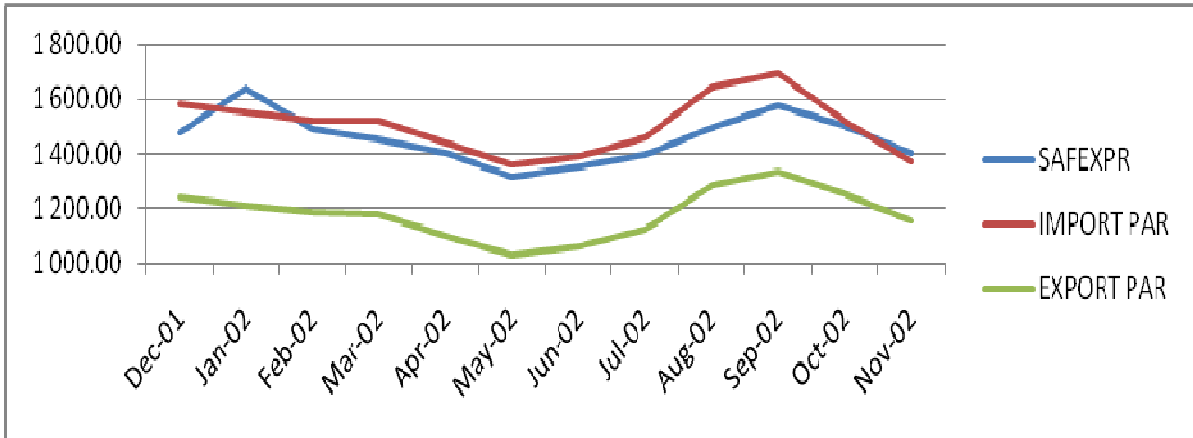


Figure A2.5: Average monthly SAFEX price with respect to import and export parity levels (12/2001-11/2002)

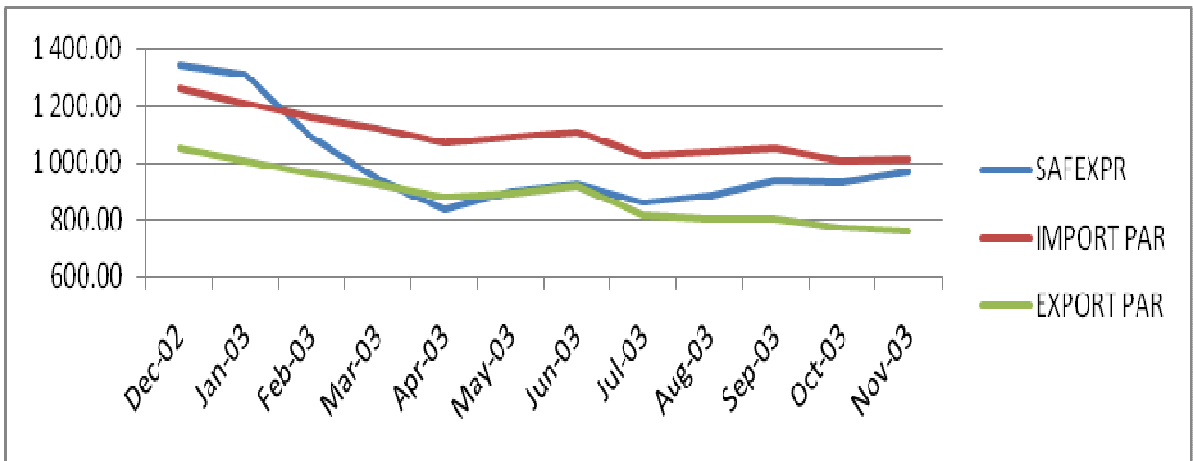


Figure A2.6: Average monthly SAFEX price with respect to import and export parity levels (12/2002-11/2003)

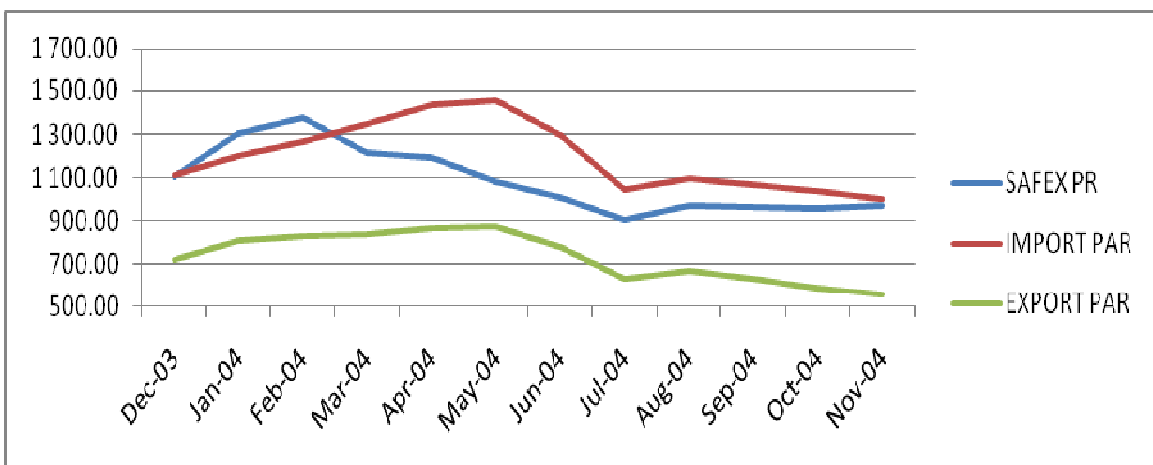


Figure A2.7: Average monthly SAFEX price with respect to import and export parity levels (12/2003-11/2004)

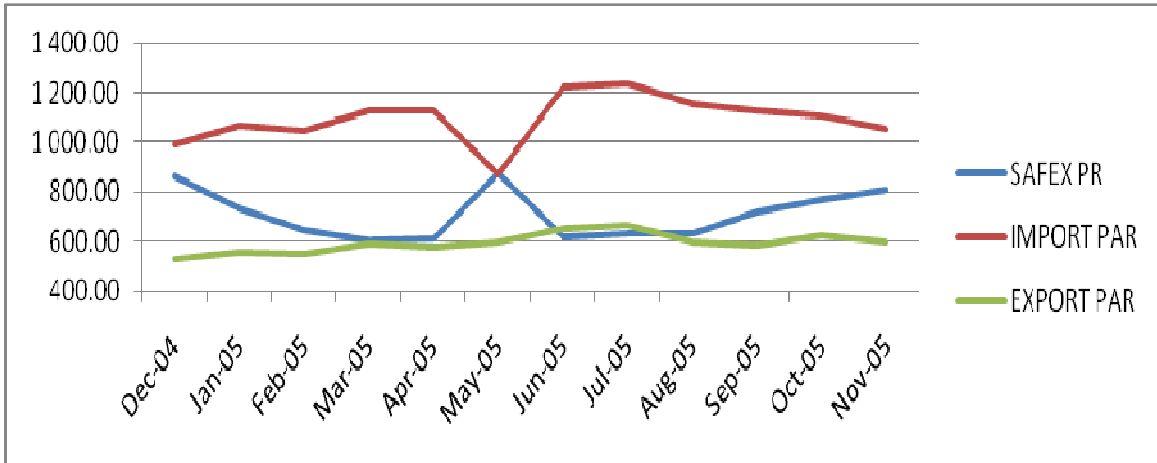


Figure A2.8: Average monthly SAFEX price with respect to import and export parity levels (12/2004-11/2005)

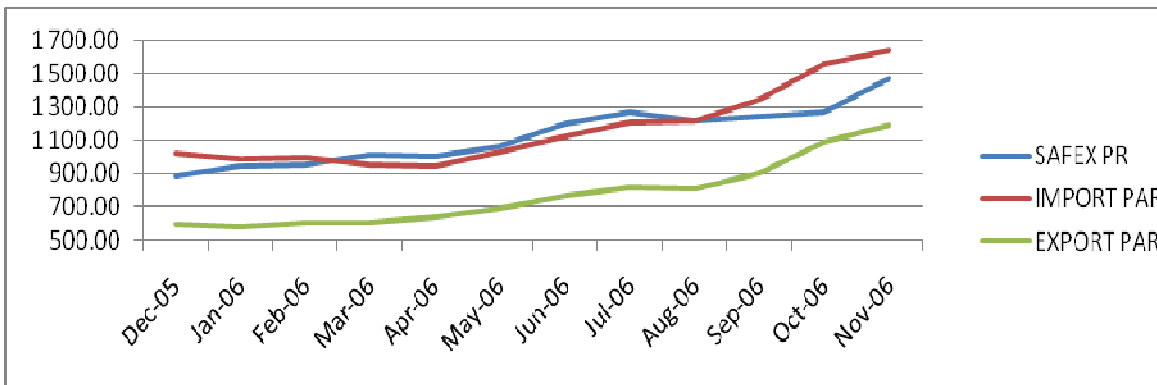


Figure A2.9: Average monthly SAFEX price with respect to import and export parity levels (12/2005-11/2006)

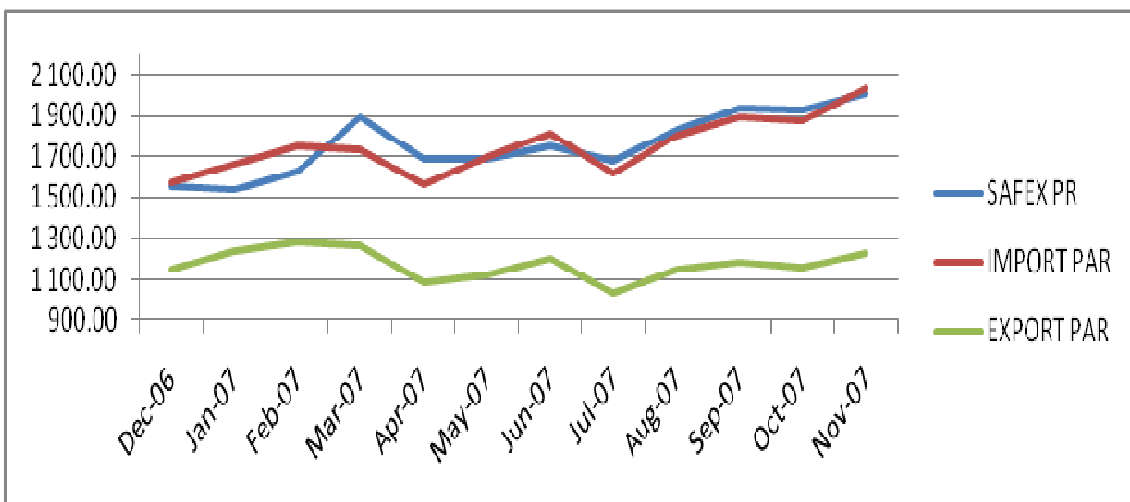


Figure A2.10: Average monthly SAFEX price with respect to import and export parity levels (12/2006-11/2007)