

**Discriminatory Performance of Error Metrics in selected Non-
linear Models**

**Dissertation Submitted in fulfilment of the Degree Master of
Commerce in Statistics in the Faculty of Commerce and
Administration, School of Economics and Decision Sciences at
North-West University, Mafikeng Campus**

Submitted by

Nthabiseng Charmaine Moitse

(22619992)

Supervisor: Prof. N.D. Moroke

April 2017

DECLARATION

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

Nthabiseng Charmaine Moitse

(22619992)

.....

Signature

.....

Date

ACKNOWLEDGEMENT

Above all, I would like to praise the almighty God for granting me the wisdom and capability to progress magnificently throughout. I would also like to express my sincere gratitude and appreciation to my supervisor, Prof N.D Moroke who continuously gave valuable comments, guidance and contribution throughout the expansion of this thesis. Not excluding the rest of the economic and decision making research committee for their comments. I am also in debt to my family, specifically my husband (Eugene Khokhong), my son (NeoEntle Eugene-Junior Khokhong), My mother (Kelebogile Moitse), my brother (Moemedi Moitse), my baby sister (Chantal Moitse) and my in-laws for the enthusiasm and warm encouragement. Much appreciation is also awarded to My Masters colleagues, Mr Katleho Makajane, Kebone Molefi and my other friends for their overwhelming support.

DEDICATION

All the hard work is dedicated to all the late people whom I knew that they wanted all that is best for me; my father (Jacky Patrick Moitse), my Aunt (Gladness Tlalang), my uncles (Sandford, Lawrence and Oageng Moletsane), my grandmother (Betty Moletsane) and my grandfather (Jack Moletsane).

ACRONYMS AND ABBREVIATIONS

AIC	Akaike Information Criterion
ANN	Artificial Neural Networks
APE	Absolute Percentage Error
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroscedasticity
ARIMA	Autoregressive Integrated Moving Averages
BDS	Brock, Dechert and Scheinkman
BG	Breusch-Godfrey
CPI	Consumer Price Index
CUSUM	Cumulative Sum
EM	Expectation Maximum
EXPAR	Exponential Autoregressive
GDP	Gross Domestic Products
GMSE	Geometric Mean of Squares of Error
HQC	Hannan Quinn Criterion
<i>Iid</i>	identically independently distributed
JB	Jarque-Bera
LM	Langrange Multiplier
LR	Likelihood Ratio
MAPE	Mean Absolute Percentage Error
MAPE _{reg}	Mean Absolute Percentage Error
MAPE _{sym}	Symmetric Mean Absolute Percentage Error
MdAPE	Median Absolute Percentage Error
MS	Markov-Switching

MS-AR	Markov Switching Autoregressive
MSD	Mean Square Deviation
MSE	Mean Square Error
NN	Neural Network
OLS	Ordinary Least Square
RWd	Random Walk with drift
RESET	Ramsey Regression Equation Specification Error Test
RMSE	Root Mean Square Error
SARB	South African Reserve Bank
SBC	Schwartz Bayesian Criterion
SETAR	Self-Exciting Threshold Auto Regressive
SA	South Africa
SSA	Singular Spectrum Analysis
StatsSA	Statistics South Africa
TAR	Threshold Autoregressive
US	United States
VAR	Vector Autoregressive

LIST OF TABLES	Pages
Table 3.1: Data Partitioning	33
Table 4.1: Descriptive statistics summary	48
Table 4.2: BDS Test	50
Table 4.3: RESET test	51
Table 4.4: Model selection criterion results	52
Table 4.5: MS(2) – AR(2)Model	53
Table 4.6: Regime switches of inflation rates (1993-2016)	53
Table 4.7: States of inflation rate	54
Table 4.8: Network Architecture-Model Summary	55
Table 4.9: Residual Diagnostics	56
Table 4.10: Error metrics results	58

ABSTRACT

The purpose of this study is to determine the discriminatory performance of the error metrics on two non-linear models, specifically the Markov Switching Autoregressive (MS-AR) models and the Artificial Neural Networks (ANN). The inflation rate of South Africa was used as an experimental unit and quarterly data for the first quarter of 1993 to the second quarter of 2016, serving 90 observations were used. Brock, Dechert and Scheinkman (BDS) test, Cumulative Sum (CUSUM) and the Ramsey Regression Equation Specification Error Test (RESET) were employed to confirm the presence of non-linearity and instability of the data. In case of MS-AR models, the Akaike Information Criterion (AIC) was used as a best model selection criterion, while for the ANN, different learning rates and momentum values were employed for selecting the best model. The following error metrics were employed for evaluating the forecasting performance of the two competing models; Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE_{reg}), Theil's U Test, Symmetric Mean Absolute Percentage Error (MAPE_{sym}), Geometric Mean of Squares of Error (GMSE) and the Median Absolute Percentage Error (MdAPE). The verdict of the study was that the ANN out-performs the MS-AR models because of lesser errors produced when forecasting. With these results, the central bank can strive to achieve price stability.

Table of Contents

DECLARATION.....	i
ACKNOWLEDGEMENT.....	ii
DEDICATION.....	iii
ACRONYMS AND ABBREVIATIONS.....	iv
LIST OF TABLES.....	vi
ABSTRACT.....	vii
CHAPTER ONE	1
STUDY ORIENTATION	1
1.1 Introduction.....	1
1.2 Problem Statement	3
1.3 Objectives of the study	4
1.4 Research Methodology	5
1.4.1 Data	5
1.4.2 Methods.....	5
1.5 Significance of the Study	6
1.6 Scope limitations and delimitations of the study	7
1.7 Organisation of the study	7
1.8 Conclusion	8
CHAPTER TWO	9
LITERATURE REVIEW	9
2.1 Introduction.....	9
2.2 Inflation rate in South Africa.....	9
2.2.1 Inflation Targeting.....	11
2.3 Significance of modelling inflation rate	12
2.4. Time Series Models	12
2.4.1 Linear models.....	13
2.4.2 Non-Linear Models	14

2.4.2.1 Markov Switching Autoregressive Models.....	14
2.4.2.2 Artificial Neural Network.....	18
2.5 Conclusion	22
CHAPTER THREE	23
METHODOLOGY	23
3.1. Introduction.....	23
3.2 Data	23
3.3 Preliminary data analysis.....	24
3.3.1. Test for Non-linearity	24
3.3.1.1 Brock, Dechert and Scheinkman test.....	25
3.3.1.2 Ramsey Regression Equation Specification Error Test (RESET) test.....	26
3.3.1.3. Cumulative Sum (CUSUM).....	27
3.4 Primary data analysis	28
3.4.1. Markov Switching-Autoregressive Models.....	29
3.4.2 Artificial Neural Networks.....	31
Step 1: Variable selection.....	32
Step 2: Data collection.....	33
Step 3: Selection of data.....	33
Step 4: Partition the data used.....	33
Step 5: Designing of a Neural Network.....	33
Step 6: Training ANN.....	35
3.5 Model selection criterion	37
3.6 Residual diagnostics.....	37
3.6.1 Normality test.....	38
3.6.2 Homoscedasticity.....	39
3.6.3 Test for Autocorrelation.....	41
3.7 Error Metrics	42
3.7.1 Mean Absolute Error.....	42
3.7.2 Mean Square Error.....	43
3.7.3 Root Mean Square Error	44

3.7.4 Mean Absolute Percentage Error (MAPE _{reg})	44
3.7.5 Theil's U-Statistic (U-Statistics).....	45
3.7.6 Symmetric Mean Absolute Percentage Error	45
3.7.7 Geometric Mean of Squares Error	46
3.7.8 The Median Absolute Percentage Error	46
CHAPTER FOUR.....	47
RESULTS	47
4.1 Introduction.....	47
4.2 Preliminary Data Analysis	49
4.2.1 Visual inspection	49
4.3.2 Test for non-linearity and stability.....	50
4.4 Markov Switching-Autoregressive Model	52
4.5 Artificial neural networks results.....	54
4.6 Residual diagnostics results.....	56
4.7 Forecasts from the $MS(2) - AR(2)$ and ANN	57
4.8 Forecasting accuracy	58
CHAPTER FIVE	61
CONCLUSIONS AND RECOMMENDATIONS.....	61
5.1 Summary.....	61
5.2 Research objectives and conclusions	61
5.3 Recommendations	64
5.4 Conclusion	65
REFERENCES.....	66
APPENDICES	74

CHAPTER ONE

STUDY ORIENTATION

1.1 Introduction

Financial time series has a tendency to change dramatically. Such occurrences are defined as regime shifts and therefore, cannot be modelled with the usage of a linear equation model. Usually these dramatic changes are caused by internal factors such as political instability and external factors such as financial crises that negatively affect inter-dependent countries. Moreover, forecasting financial time series such as macro-economic fundamentals has and is still gaining attractiveness amongst researches on the basis of both technical and theoretical reasons. Both public and private agents have developed interest in the movement of such, in order to make sound and informed decisions. Nevertheless, forecasting of a series only provides estimates and retrieves future occurrences; it does not reduce uncertainties and complications that might occur. This process is useful in indicating expectations that will be incorporated into the process of decision-making.

Forecasting has been a very challenging issue for financial experts and economic researchers, due to difficulties in finding the best forecasting model. However, many studies have developed time series models based on the linearity assumption, for example; through univariate Autoregressive Integrated Moving Averages (ARIMA) and multivariate Vector Autoregressive (VAR) models for forecasting macro-economic variables, without considering issues such as structural break points. One of the well-known possibilities of a structural break point that has disturbed many economic variables is the United States financial crisis that occurred between 2007 and 2009. The results of this crisis were heat-felt by many countries including South Africa (SA) in 2009.

For such data with structural breaks, non-linear models are more relevant as they have the ability to capture unforeseen and uncontrollable changes occurring in the economy (Chauvet *et al.*, 2002).

To date, most of the researchers have popularised non-linear models particularly in empirical economics. This has become a trend and has also enticed interest in forecasting economic variables with non-linear models. Reference can be made to Tsay (2002) and Clements *et al.* (2004) for recent developments on this topic. Though there is a promising growth in literature which consists of forecast comparisons involving a rather lesser number of time series and non-linear models, the trend is not good enough.

Very few studies considered comparisons of models which are entirely focused on simulated series. Due to innumerable non-linear forecasting models present, this study restricts its consideration to only two non-linear models. That is, Markov Switching (MS) models specifically Markov Switching Autoregressive (MS-AR) models and Artificial Neural Network (ANN).

Generally MS is a useful tool for analysing dependent random events, whose probability depends on previous occurrences. This model is a ground-breaking work of Hamilton (1989, 1990). Neural Network (NN), more preferably referred to as ANN is developed from the biology concept (Haykin, 1994), it is defined as an information processing model which is stimulated by the way that the nervous system works. It is made up of many interconnected processing elements called neurons working harmoniously to solve the problem at hand. Having this in mind, the two models are more relevant to the current study since they have the ability to capture non-linearity and handle data which contain structural breaks with ease.

An important point to note is that this study only limits its focus to the discriminatory performance of forecasting error metrics in non-linear models with reference to the two proposed models. The study does not cover volatility forecasting.

1.2 Problem Statement

For many decades, linear models have been confidently used for forecasting without considering the element of non-linearity that might be present within the time series data. According to Tiao and Tsay (1994), many macro-economic variable models are developed on the basis of linearity assumptions, while there is increasing evidence that those macro-economic variables have elements of non-linearity. Hence, the results and conclusions drawn from such models could be misleading.

According to literature, most researchers used other non-linear models such as the Exponential Autoregressive (EXPAR) and Threshold Autoregressive (TAR) models. Evidence has been provided that no theories attached to these models were incorporated as they necessitate the burden of assumptions concerned with accurate types of non-linearity. This poses a disadvantage since there are numerous types of non-linearity within a particular data which may be insufficiently captured within the pre-specified non-linear model. Furthermore, this leads to exclusion of important characteristics present in the data concerned. Such shortfall can be modelled well with the proposed models in this study.

The comparative discriminatory performance of the error metrics associated with these models will be applied to the South African inflation data. This series was chosen due to the sudden structural breaks and the non-linearity it possesses. Due to these characteristics, it has become almost impossible to produce accurate and reliable forecasts of inflation rates of South Africa.

This necessitates policy makers to embark on relevant and effective policies since their decision is based on wrong forecasts of this sector. Based on these highlights, the main question posed is, *“how effective are the ANN and MS-AR models in forecasting the non-linear time series data with structural breaks?”* and *“which of the two models is discriminated against by the forecasting error metrics?”*

1.3 Objectives of the study

The primary objective of the current study is to explore the discriminatory performance of the error metrics in the ANN and MS-AR models in the context of the South African inflation rates. Secondary objectives are as follows:

- 1.3.1 To determine the characteristics of inflation rate of South Africa.
- 1.3.2 To develop non-linear models for forecasting inflation rate using ANN and MS-AR frameworks.
- 1.3.3 To determine the discriminatory performance of error metrics in the ANN and MS-AR models.
- 1.3.4 To formulate recommendations for practice and further studies.

1.4 Research Methodology

This section mainly covers the discussion about the data to be used, the procedures to be followed, methods and statistical packages to be employed.

1.4.1 Data

The proposed data used was the inflation rate for total country including all the items of SA, that is; Food and non-alcoholic beverages, alcoholic beverages and tobacco, clothing and footwear, housing and utilities, household contents, health, transport, communication, recreation and culture, education, restaurants and hotels and lastly, miscellaneous goods and services, including insurance and financial services. Such data was obtained from Quantec on quarterly bases.

1.4.2 Methods

The study is comparative in nature and adapted both ANN and MS-AR methodology. The existing literature by accredited publishers and books was used and acknowledged as consultation sources to gain better understanding about each framework. The study strictly followed quantitative approach due to the nature of the models adopted and the data used. The preliminary data analysis included assessment of existence of non-linearity in the data through the Brock, Dechert and Schienkman (BDS) Test(1987), Ramsey Regression Equation Specification Error test (RESET) (1969) and Cumulative Sum(CUSUM)developed by Brown *et al.* (1975).The study further performed a discriminatory performance using the eight error metrics namely: the Mean Absolute Error(MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error MAPE_{reg}, Theil's U test, Symmetric Mean Absolute Percentage Error (MAPE_{sym}), Geometric Mean of Squares of Error (GMSE),

and The Median Absolute Percentage Error (MdAPE) in order to evaluate the forecasting ability and reliability.

1.5 Significance of the Study

In accordance with Green (1996), an inflation targeting framework necessitates a policy implemented by Central Bank's policy actions in order to be conversant with the estimated future inflation rate in alignment to a well-known inflation targeting. In February 2000, South Africa declared that inflation rate must range between 3 and 6 percent (Bahramian *et al.*, 2014). With this in mind, the South African Reserve Bank (SARB) requires a suitable model that can assist the monetary policy to obtain a favourable range of inflation rate (Van den Heever, 2001). The intension of this study is to explore the predictive performance of ANN and MS-AR models by forecasting inflation rate.

The study is unique since no similar study has been conducted in South Africa. The findings of this study present functional information to scholars and analysts by providing statistical background on the chosen non-linear models. This might also assist them to have a better choice between the competing frameworks. It may also provide functional information to monetary policy decision makers whose core objective is to achieve and maintain price stability. This study is expected to contribute optimistically to the development of concrete monetary policies and also assist in embarking on relevant strategies for addressing inflationary related problems. The compilation of this dissertation might also increase the existing empirical literature on the subject by applying both the innovative ANN and MS-AR models, inform and influence the buying power of households, and generally assist to make accurate estimates on the general increase of prices of goods and services.

1.6 Scope limitations and delimitations of the study

- This study focuses on the discriminatory performance of the eight error metrics to the two non-linear models. As a result, other non-linear models are not considered.
- Though literature reports about many error metrics used in selecting models with least forecast error, the current study is limited to those available on the software.
- There is dearth of literature on non-linear models compared to linear models; this poses a limitation to this study with regard to the sources available. The current study may therefore, be forced to consult sources older than favourable interval of 10 years.
- There are no possible delimitations to this study.

1.7 Organisation of the study

The dissertation is structured as follows:

Chapter 2-Gives empirical literature on the study, it is significant in giving a broader perspective of the topic of interest achieved by accredited publishers who have conducted a similar study.

Chapter 3-Methodology: accommodates all the methods and techniques employed, the research design and brief information regarding the data used.

Chapter 4-Includes analysis and interpretation of results obtained in order to achieve the objectives.

Chapter 5-Accommodates the summary and recommendations.

1.8 Conclusion

There is increasing interest in modelling data based on the assumptions of non-linearity because of certain unexpected events happening around the world. The study examines two types of non-linear models that is, Markov-Switching Autoregressive Models and Artificial Neural Networks using inflation of South Africa. The subsequent chapter gives a brief literature review to get a broader perspective about the two models.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

This section briefly discusses literature around the proposed non-linear models. Reference is made to studies that have adopted and/or applied the proposed models. Section 2.2 gives a brief discussion on the inflation rate of South Africa and its proposed targeting, followed by the significance of modelling inflation rate in section 2.3. Section 2.4 provides a discussion about theory and literature on time series modelling. The last section is the conclusion.

2.2 Inflation rate in South Africa

Inflation rate is regarded as a measure of inflation and is defined as the percentage change in the Consumer Price Index (CPI). CPI identifies the general increase of prices of goods and services in the economy. The known and expected inflation rate acts as an important tool for serving as an encouragement for saving, investment, consumption and production in the country (Pretorius and Janse Van Rensburg, 1996).

It is compulsory for central banks to effectively manage the inflation rate as this sector is one of the most important macro-economic variables. Inflation rates have also been found to have direct or indirect effect on changes that occur in other macro-economic variables such as investments, interest rates, and exchange rates (Molwantoa, 2013). Therefore, many policy-makers have gained massive interest in developing the best model for forecasting inflation rate, thus leading to many countries have developed their inflation targeting strategies.

South Africa being one of these countries adapted its own strategies by South African Reserve Bank (SARB) in 2000 in order to identify the inflationary state of the country. Uko and Nkoro (2012) mentioned that usefulness of monetary policy has much influence on the success of inflation targeting.

With reference to Pretorius and Van Rensburg (1996), accurate forecasts of inflation rates do not only have positive influence on concrete decision making and policy enhancement. They also have a significant impact on the overall economy of the country at large. Nevertheless, forecasting inflation rate remains to be a challenging task to be performed because of instability in prices of food and structural changes (Aron and Muellbauer, 2000).

Inflation rate is one of the sensitive macro-economic variables, because every negative or positive event that occurs in the economy somehow affects it, and hence its trend mostly indicates fluctuations and growth impulsively. According to Statistics South Africa (StatsSA), in South Africa, the categories that mostly contribute to consumer price index(CPI) are; housing and utility standing at 24.5 percent of total weight, transport at 16.4 percent and food and Non-Alcoholic Beverages recorded at 15,4 percent. Other categories include miscellaneous goods and services such as personal care, finance and insurance standing at 14.7 percent; followed by Alcoholic beverages and tobacco; Household contents, equipment and maintenance; Recreation and culture; and Clothing and Footwear standing at 5.4 percent; 4.8 percent; 4.1 percent and another 4.1 percent respectively. The remaining 10 percent represents Restaurants and Hotels; Education Communication and Health and Education.

Since 1993 to date, there have been economic events that caused changes in the inflation rate of South Africa. The highest peak was experienced around the end of 2002 and 2008. The

lowest inflation rate was experienced at the beginning of 2004 which almost approached a zero. The current study uses the ANN and the MS-AR models to develop a good forecasting model for inflation rate of South Africa with improved accuracy measures. Since both of the techniques are non-linear and therefore can handle volatility of inflation rate.

2.2.1 Inflation Targeting

The formal adaptation of inflation targeting in South Africa by the SARB was in April 2000. In accordance to Jonsson (1999), the main reason for proposing to adopt this strategy was mainly to assist the central bank to achieve price stability. According to Molwantoa (2013), the main advantage of moving from “eclectic” monetary policy approach to approach of directly targeting inflation is that the decision increases the chances of maintaining and achieving low or stable inflation rate. Moreover, a well-structured target also makes it easy for the economy to recover quickly from economic shocks such as recession.

The following are some of the reasons for changing to a formal inflation targeting outlined by Van der Merwe (2004):

- The informal inflation targeting mostly created reservations among the public concerning the monetary policy stance developed by the authorities.
- The inflation targeting adapted in year 2000 is able to improve co-ordination between monetary policy and other economic policies having the condition that the targets are dependable to other objectives.
- Another reason for developing the new inflation targeting framework is to increase accountability of the SARB with respect to monetary policy.
- To facilitate a decrease in inflation and have an influence in the inflationary expectations.

2.3 Significance of modelling inflation rate

Inflation targeting is one of the most important objectives to be achieved by the monetary policy. Adaptation of the new inflation targeting framework has increased the need for estimating reliable inflation rates. Nevertheless, Nakamura (2006) mentioned that forecasting inflation might be a very difficult mission to accomplish because of the disturbance in the economy caused by the global recession that occurred in 2008.

Moreover, Aron and Muellbauer (2000) states that, forecasting inflation rate remains to be a challenge to many statisticians and economists considering that there are other factors such as change in food prices. This is the reason why this study is considering the application of the two time series analysis models ANN and MS-AR to model and forecast inflation.

2.4. Time Series Models

Time Series models are developed from a time series which is defined as any ordered sequence of data of any variable having equally arranged spaced intervals. Application of these models can be used to make forecasts in the economic world, finance, medicine, environmental states, budget, Stock markets, sales and etc.

With reference to Clements and Hendry (2002), time series models are primary methods of forecasting in the world of economics. In addition to that, Moshiri (1997) stated that time series models are able to give better forecasts as compared to econometric systems, and this is one of the main reasons why they are opted for. Time series models easily show the historical trend of data which provides assistance in what to expect in the future through the assumption that the data has the ability to reveal the behaviour for the variable at hand.

Linear and non-linear models are the two widely used time series models that are used for forecasting variables such as financial and economic variables. It is very important to understand the assumptions and characteristics of each model, as each model possesses its unique use, advantage, fault, character and assumptions.

2.4.1 Linear models

Many experiments in several disciplines are designed, analysed, and interpreted in the context of general linear theory. Hays (1973) mentioned that in most cases, the theory defines at least indirectly in the text of any statistics and explicitly in any mathematical based statistics text, assuming that the response variable is composed of the sum of explanatory variables and the interactions between those explanatory variables.

For the past many years, linear models have played a very significant role in the world of forecasting. The most commonly used traditional models are the multivariate vector Autoregressive (VAR) models and the univariate Autoregressive Integrated Moving Averages (ARIMA) models.

Although application of linear models is flexible and easy to compute, visual inspection of many variables such as inflation rate, stock prices, exchange rates, interest rates, etc., reflects sudden changes of data, indicating the presence of non-linearity of which linear models cannot handle (Makridakis *et al.*, 1982). Hence, there is a high probability that models developed from linearity theory produces misleading results and conclusions.

2.4.2 Non-Linear Models

The interest of modelling financial and economic time series data using non-linear models have increased tremendously and have gradually been opted for as alternatives in modelling and forecasting. This is due to the reason that many researchers have discovered dissimilar behaviour of both financial and economic time series data which is caused by changes within the economy, expectations of investors, Global financial crisis and sudden weather changes that occur. However, like any other models, non-linear models have their own limitations. That is, its implementation is unmanageable and their flexibility is limited since their designed to describe certainty of non-linearity. Therefore, to develop a reliable non-linear model, the data set used should be well observed. Nevertheless, one of the well-known non-linear model that can accommodate this blunder is the ANN because one of its property is the capability to handle any pattern of non-linearity (Kuan and White, 1994).

Literature review based on forecasting macro-economic variables, including inflation rate has increased over the years, to name a few Chauvet *et al.* (2002), Clements *et al.* (2003) and Binner *et al.* (2005). To show the forecasting accuracy of non-linear models, most authors have conducted a comparative study whereby different non-linear models were compared to linear models. Most of the studies indicated that non-linear models outperform linear models (Chauvet *et al.*, 2002), Binner *et al.*, 2005), Balcilar *et al.*, 2012),

2.4.2.1 Markov Switching Autoregressive Models

Markov Switching (MS) model is the groundwork of Hamilton (1989). Since its development it has been used in non-linear modelling in the world of statistics. These models are more appropriate for modelling time series data that consist of periods with dissimilar behaviour.

For example, one cannot expect recessionary economy and expansionary economy to behave the same way. Deviating to inflation, Friedman's (1997) hypothesis that high inflation results in a growth of volatility supports the possibility of irregular regimes in inflation. The MS model is opted for its flexibility since it permits data to expose the nature of incidence of significant changes. This model is also well-known for its capability to capture shifts performed by the mean and variance.

A number of scholars and analysts have applied MS model in their different areas of studies. Among others, Ismail and Isa (2006) conducted a study which compared the two regime switching models, that is the Self Exciting Threshold Autoregressive (SETAR) Model and MS-AR model to examine which model can best capture changes within the time series data of exchange rates of ASEAN Countries namely, Malaysia, Singapore and Thailand. The authors intended to assess the fitness of each model through the use of Akaike Information Criterion (AIC), Hannan Quinn Criterion (HQC), Schwartz Bayesian Criterion (SBC) and Likelihood Ratio (LR).

Primarily, the study sought to compare the efficiency of the switching models with the linear autoregressive model. This also helped in assessing whether regime switching models can best describe nonlinearity features of the data of interest using McLeod test, RESET Test and BDS Test. The exchange rates data used was from February 1990 to June 2005 resulting in total observation of 180. The results indicated that MS-AR had the minimum values of the four information criteria as compared to other competing models, drawing a conclusion that, MS-AR model outperform SETAR model and the linear AR model.

Ismail (2007) used univariate MS-AR model to detect regime shifts behaviour of exchange rates in Malaysian ringgit verses four other currencies using 92 observations. Portmanteau test revealed that non-linear models are more adequate as compared to linear models, meaning that MS-AR model is superior to Autoregressive (AR) model. The study went as far as comparing the significance of each model with the Likelihood Ratio (LR). The findings confirmed that MS-AR model performs better than Autoregressive (AR) Model.

Furthermore, the results indicated a regime shift which was as a result of the financial crisis that occurred in 1997 leading to a conclusion that any disturbance in the economy has an effect on the exchange rate. This further confirms that MS-AR just like the ANN perform well in non-linear data which has structural breaks, hence their application in the current study.

Yu (2007) undertook a study on forecasting inflation rate in Philippines through determining and comparing the predictive power of linear and a non-linear model. The data used was the CPI with 2000 as a base year. The linear model used was based on the Ordinary Least Square (OLS) projections of the Phillips relation and the non-linear model used was MS model. The conclusion that was made was that MS model can best make better forecast of the inflation rate based on RMSE, which was the only error metric that was used for evaluating the forecasting models.

This study slightly differs with the current study because two time series models having different assumptions are compared, which is one of the popular shortfalls performed by many authors. Another deviation is that, only one error metrics was used to evaluate the forecasting performance which is not enough.

Mostafaei and Safaei (2012), applied MS-AR by comparing the in-sample forecast of MS-AR and Random Walk with drift (RWd) processes. The United States (US) Dollars to one Euro data was used and was collected from January 2003 to April 2011 on monthly basis. A total of 124 observations were used. The forecasting performance was based on the minimum value of MSE which represented the MS-AR model. The study concluded that the MS-AR outperforms RWd processes in terms of forecasting. Their study deviates from the current study as the conclusion was based on only one error metric. To avoid the issue of biasness, the current study uses eight different types of error metrics to confirm the results.

Yarmohammadi *et al.* (2012) conducted a similar study to Ismail (2007) but the study was comparing eight different time series modelling approaches (linear and non-linear models) including MS-AR. The data used was the monthly Iran's rail per the US dollar which was collected from 1995 to 2009. The objective of the study was to identify the best fitted model amongst different linear and non-linear models using AIC and the BIC. The results indicated that MS-AR model has the least AIC and BIC as compared to the other five time series modelling approaches. Furthermore, the results revealed that the MS-AR have better statistically significant estimated parameters than other models.

Based on these findings, the authors concluded that MS-AR model is a superior tool for modelling time series data having dramatic jumps. The study recommended Singular Spectrum Analysis (SSA) which is also well-known for forecasting time series data with unstable behaviour.

2.4.2.2 Artificial Neural Network

The development of ANN was motivated by biology. According to Haykin (1994) a human brain is made out of 10 billion nerve cells (neurons) and about 6000 times connections which are known as synapses. Therefore, all the information that is received by any human being is processed in the brain. An ANN is defined as the information processing technique which is inspired by the way the biological nervous system works. There are many interconnected processing elements known as neurons that work harmoniously so as to solve a problem at hand (Dase and Pawar, 2010).

There has been an increased interest of forecasting macro-economic variables using the ANN and the increasing number of published papers proves so. This approach has the ability to accommodate disturbances with time series data. The method is also capable of working equivalent with input variables and can easily handle large data sets. Another strength of this method is that, it has the ability to finding patterns (Chang *et al.*, 2007).

One of the most important characteristics of ANN is that, it can efficiently capture non-linearity in data used within the many quantitative fields such as business, science, economic and financial world. Further the ANN can effectively capture underlying relationship in real world problems. The main application of ANN is making projections for future occurrences which result in concrete decision making, planning, purchasing, formulation strategy and policy making.

One of the recent studies which applied the ANN and other econometric models was conducted by Nunoo (2013) in Ghana. The primary objective of the study was to make a comparison between econometric methods and ANN models. The sample size of 240 was used for conducting the whole analysis. Applying Root Mean Square Error (RMSE) as an error metric, the results retrieved that ANN models had lower errors compared to econometric models. Moreover, the study concluded that ANN can best forecast inflation rate. For future studies, Nunoo (2013) recommended application of ANN models for forecasting other macro-economic variables such as exchange rates and Gross Domestic Products (GDP) and increased sample size to better the forecasts.

This study informs the current study since it applied data collected over longer period of time and applied the ANN as one of the models proposed. A weakness of this study is that only one error metrics was considered and this does not necessitate better judgement about the performance of this model on the basis of one forecast error metric. To address this biasness, the current study appoints a number of error metrics to evaluate the accuracy of each model and make a sound judgement about the performance of each model.

One of the studies that also voted for ANN was conducted by Maliki *et al.* (2011). The objective of the study was to model the electronic power generated in Nigeria on year basis using 35 observations. MAE, MSE, and RMSE were used for comparing the errors involved when developing forecasts of each model. For ANN MAE, MSE, and RMSE was standing at 0.049, 0.0012 and 0.62 respectively.

For the regression model, the MAE was standing at 0.056, MSE was standing at 0.003 and lastly the RMSE was recorded as 0.076. With the above results, the study concluded that the

ANN model has the ability to give better forecast as compared to linear regression model. One of the flaws of the study is that the two models compared are meant to address conflicting assumptions at some point.

Tjung *et al.* (2010) conducted a comparative study between ANN and Ordinary Least Square (OLS) methods. The primary objective of their study was to check which method could retrieve better forecasts of stock prices. The authors used seven financial stocks based on daily changes in stock prices from 1998 to 2008. The R^2 was the only error metric used for evaluating the accuracy of each model. The results indicated that, the accuracy of ANN was 96% while OLS recorded a percentage of only 68%. The study therefore concluded that ANN outperforms OLS methods on the basis of R^2 . The study carries the same flaw as many others, e.g. Maliki *et al.* (2011) who made a comparison of two models with different assumptions which is one of the issues that the current study avoids.

The ANN is good at capturing the non-linearity, the task which the OLS cannot handle. Another shortfall is that the conclusion was drawn from R^2 which is greatly used for regression analysis but irrelevant in the world of forecasting (Makridakis and Hibon, 1995). Stock prices are known to be non-linear and non-stationary by their nature. Therefore, the choice and application of models must be done judiciously. Based on the results of Tjung *et al.* (2010) study, it is clear that linear models do not have the ability to handle non-linear and non-stationary data; hence the ANN was found to be 28% more efficient and robust over the linear models.

Vaisia *et al.* (2010) compared NN with multiple regression analysis which is referred to as multivariate statistical technique. Daily stock prices ranging from 1 April 2005 to 30 March 2007 was used as an application. The objectives of the study were to determine the model that is best suited for forecasting the daily stock prices. The forecasting ability was determined by the lowest values of Mean Absolute Percentage Error (MAPE), Mean Square Error (MSE) and Root Mean Squared Error (RMSE). The results indicated that when the data is well trained and has proper input then Neural Network (NN) can best forecast the daily stock prices as compared to multiple regression. In general, regression analysis is the best linear technique for indicating the relationship of variables, while ANN is for modelling and forecasting time series data with non-linearity. Hence, the study compared two models with different assumptions. For the current study, two models having similar assumptions are compared and evaluated.

Ahangar and Yahyazadehfar (2010) did a comparative analysis between the ANN and the linear regression model using 10 macro-economic variables (growth rates of industrial production, inflation rate, interest rate, exchange rate, rate of return on stock public, unemployment rate, oil price, gross domestic product, Money supply 1(M1), and money supply 2 (M2) and three financial variables (book value per share, sales per share, and earnings per share in order to predict stock prices in Terhran Stock exchange. The population that was included was all the companies that acted in the Tehran stock exchange between 1380-1386. The study employed the MSE, MAPE and the coefficient of determination to evaluate the forecasting accuracy.

The results obtained indicated that ANN outperforms linear regression model because it reported lower values of MSE, MAPE and the coefficient of determination. Like other mentioned studies, the study compared two models characterised with different features and capabilities. The current study attempts not to repeat the mistake by firstly assessing characteristics of the data at hand such as instability, and the presence of non-linearity.

Faria and Gonzalez (2009) conducted a comparative study between ANN and Adaptive exponential smoothing method using daily Brazilian stock market collected between 1998 September to April 2007. The primary objective of the study was to evaluate the ANN and the Adaptive exponential smoothing method through the RMSE and the N_{TEND} , (which was defined as the number of times that the predictions traces of the real data used). Both techniques produced similar results in terms of index return forecasts. Interestingly, the two error metrics were in favour of the ANN over the Adaptive exponential smoothing method. The findings are in support of other studies on similar areas.

2.5 Conclusion

The above literature shows that both ANN and MS-AR Models have significantly outperformed other models. It is clear that most authors opted to compare these models with linear models, a mistake which the current study will try to avoid. To avoid biasness, the current study will use eight error metrics to decide on the best forecasting model between the ANN and the MS-AR models. All the eight error metrics are suitable for comparison of non-linear models according to Makridakis and Hibon (1995).

CHAPTER THREE

METHODOLOGY

3.1. Introduction

Markov switching Autoregressive (MS-AR) models and Artificial Neural Network (ANN) are non-linear models that should be handled with proper procedures. The current chapter accommodates all the techniques, methods and procedures, tests and necessary visual inspections that were previously used by accredited publishers on similar models and retrieved reliable results. The chapter gives brief information on the data used, the formal and graphical presentations of testing non-linearity within the data, necessary steps to follow when conducting MS-AR and ANN. Lastly, the chapter provides a procedure for obtaining and applying each error metrics used in the study.

3.2 Data

The proposed data that was used is the inflation rate that is the Consumer Price Index (CPI) for total country including all the items of SA. The main reason for using inflation rate is because, it is one of the macro-economic variables that never show stability or consistency hence regarded as non-linear in nature. The data was obtained from Quantec on quarterly bases covering the period from first quarter of 1993 to the second quarter of 2016, serving 90 observations.

The reason for having a large sample is to have the population representativeness within the sample (Boot and Pick, 2014). The selected period also accommodates assumed structural break points in South Africa.

The use of quarterly and large data also helps to safeguard the normality and heteroscedasticity assumptions which are usually present in time series data (Moroke, 2015).

Two different statistical packages were used for data analysis. OxMetrics Version 6.0 was used to produce results for MS-AR models since the package is able to regime any parameter independently. Zaitun time series version 0.2.1 was used for obtaining the results of the ANN.

3.3 Preliminary data analysis

Preliminary data analysis is important for assessing the key features of the time series data that are used for the data analysis, to make a summary of all the information about the data into an easily understandable format. For the current study, preliminary analysis is achieved firstly by observing the visual inspection of the data to be used and to conduct the descriptive statistics which includes the mean, median, standard deviation, kurtosis and skewness in order to get a better perspective about the data.

3.3.1. Test for Non-linearity

In order to proceed with the primary data analysis, the data should with certainty be non-linear and non-stationary in nature. The following were the tests that were proposed to check the presence of nonlinearity within the time series data. Brock, Dechert and Scheinkman (BDS) test which tests the null hypothesis of independently and identically distributed (*iid*) within the data, Ramsey Regression Equation Specification Error Test (RESET) test which is relevant for testing specification, and a graphical representation of Cumulative Sum (CUSUM) which best demonstrate the variance instability. The three tests were selected as each of them has the unique ability to assess the type of non-linearity involved in the data used. The three tests are further discussed in preceding sections.

3.3.1.1 Brock, Dechert and Scheinkman test

The BDS Test was named after developers Brock, Dechert and Scheinkman (1987). It is best known for testing dependence in non-linear time series data. The test can assist when avoiding false detections of critical transitions which are mostly caused by misspecification of a model (Barnett *et al.*, 1997). After using different ways of transforming data, the BDS test is then applied to assess null hypothesis that the remaining residuals are *iid*. Rejection of the null hypothesis can act as an ad-hoc diagnostic test that detects non-linearity within the data and hidden non-stationarity or any other structure that is not included due to the model fitting (Brock *et al.*, 1991). Non-linear responses are mostly caused by critical transitions; therefore the null hypothesis of the BDS test should not be accepted.

When the test is applied to the residuals of the fitted linear model, the BDS test can be relevant for detecting the remaining dependence and all the structures of non-linearity that are not included within the model. If the null hypothesis is accepted, then the original model should also be accepted. In this case, if the null hypothesis is rejected, the conclusion made from that is that, the fitted linear model is mis-specified and therefore, can be treated as a test for non-linearity. The following procedure for conducting the BDS test was developed by Potter (1999).

The integral concept to be captured in this test is the correlation integral, which is a measurement of the frequency whereby temporary patterns are recurring. The correlation integral at bedding dimension m can be calculated by using the following equation:

$$C_{m,\epsilon} = \frac{2}{T_m(T_m-1)} \sum_{m \leq s < \sum_{t-T} I(X_t^m, x_s^m; \epsilon) \quad (3.1)$$

Considering the time series data $x_t, x_t \text{ for } t = 1, 2, \dots, T, m$, m - history is recorded as $x_t^m = (x_t, x_{t-1}, \dots, x_{t-m+1})$. $T_m = T - M + 1$ and is defined as an indicator function that is equivalent to zero if $|x_{t-1} - x_{s-1}| < \epsilon$ for $i = 0, 1, \dots, m - 1$ and zero otherwise. Spontaneously, the probability of any two m -dimensional points are within the distance of ϵ is estimated by correlation integral, meaning that it estimates the following joint probability:

$$P_r(X_t - x_s | < \epsilon | X_{t-1} - x_{s-1} | < \epsilon, \dots, |X_{t-m+1} - x_{s-m+1}| < \epsilon) \quad (3.2)$$

If x_t follows a white noise process, the above probability will be equal to the following limiting case:

$$C_{1,\epsilon}^m = \Pr(|x_t - x_s| < \epsilon)^m \quad (3.3)$$

Brock *et al.* (1996), defined the BDS test as the following:

$$V_{m,\epsilon} = \sqrt{T} \frac{C_{m,\epsilon} - C_{1,\epsilon}^m}{s_{m,\epsilon}} \quad (3.4)$$

where the $s_{m,\epsilon}$ is equal to the standard deviation of $\sqrt{T}(C_{m,\epsilon} - C_{1,\epsilon}^m)$. According to Brock *et al.* (1996) this standard deviation can be continuously estimated. If $|V_{m,\epsilon}| > 1.96$ the null hypothesis *iid* should be rejected at 5% significance level.

3.3.1.2 Ramsey Regression Equation Specification Error Test (RESET) test

The RESET test is a ground work of Ramsey (1969), which is commonly used as a specification test for linear regression test. The first step to compute the RESET test is to estimate the OLS regression model,

$$\hat{\epsilon}_t = X'_{t-1} \lambda_1 + M'_{t-1} \lambda_2 + e_t \quad (3.5)$$

Where e_t is the error term that is *iid* and has a zero mean with constant variance. $M'_{t-1} = (\hat{X}_t^2, \hat{X}_t^3, \dots, \hat{X}_t^{s+t})$ is mostly greater or equal to one. After estimating the error term, the sum of squared residuals will be calculated as follows:

$$SSR_1 = \sum_{t=p+1}^n \hat{e}_t^2 \quad (3.6)$$

If λ_1 and λ_2 are both equal to zero, then it can be said that $AR(P)$ is adequate and therefore having the null hypothesis that the model is correctly specified as a linear model while the alternative hypothesis declares that the model is not correctly specified since the model is non-linear. That is

$$H_0 = \lambda_1 = \lambda_2 = 0$$

$$H_1 = \lambda_j \neq 0$$

where $j = 1, 2, \dots$

The null hypothesis is rejected if p-value is less than the test statistic of RESET test. Implying that the model specified is non-linear in nature.

3.3.1.3. Cumulative Sum (CUSUM)

Stability is another way of confirming non-linearity within the time series data. The current study employs the CUSUM test developed by Brown and Evans (1975), which is based on the recursive residuals. The first step in computing the CUSUM test is to calculate the recursive residuals as follows:

$$\begin{aligned} e_{t+1,t} &= Y_{t+1} - Y_{t+1,t} \\ &= Y_{t+1} - [\alpha_{0,t} + \alpha_{1,t}(t+1) + \dots + \alpha_{s,t}(t+1)^s + \phi_{p,t}Y_{p,t} + \dots + \phi_{p,t}Y_{p,t}Y_{t-p+1}] \end{aligned} \quad (3.7)$$

where subscripts t on the estimated parameters mean that the parameters are estimated on the bases of the sample of the last observation occurred in period t .

The second step in computing the stability test is to let $\sigma_{1,t}$ present the standard error of the one step ahead forecast of the inflation rate, that is Y . Then calculate the standard recursive residuals using the following equation:

$$W_{t+1,t} = \frac{e_{t+1,t}}{\sigma_{1,t}} \quad (3.8)$$

While maintaining the assumption that $W_{t+1,t} \sim iid.N(0,1)$.

The third step involves the calculation of the CUSUM statistics which is defined as follows:

$$CUSUM_t = \sum_{i=k}^t w_{i+1,i} \quad (3.9)$$

where W is the recursive residual, $t = k, k + 1, \dots, T - 1$ and $k = 2_{p+s+1}$ represents the minimum sample size that can be used to fit the model. If β vector continues to be constant throughout all the periods, $E(W)_t = 0$, but if β varies from period to period, then w_t will deviate from zero mean value line. In this case, the importance of the deviation from the zero line is being judged by the reference to two lines at 5 per cent significance level. The distance between the two lines increases with t and the 5 per cent significance line are as the results of connecting the following points:

$$[k, \pm -0.948(T - k)^{\frac{1}{2}}] \text{ and } [T, \pm 3 * 0.948(T - k)^{\frac{1}{2}}] \quad (3.10)$$

If the w_t is outside the 5 per cent significance line then it is concluded that the time series data is instable, illustrating non-linearity.

3.4 Primary data analysis

The primary data analysis presents a brief overview how the two non-linear models are developed, which includes: MS-AR and ANN and how each of those is assessed for accuracy.

3.4.1. Markov Switching-Autoregressive Models

The current study only focuses on one variable; therefore considers a univariate auto regressive process, that is, an AR subjected to regime shift. This section follows Cruz and Mapa (2013), who also developed MS-AR model using the inflation rate.

The variable that is explored is on quarterly basis, therefore the MS-AR model having two regimes with AR process of the order p MS (2)-AR(p)s expressed as follows:

$$y_t = \begin{cases} c_1 + \sum_{i=1}^p \phi_{1,i} y_{t-1} + \alpha_{1t} & \text{if } s_t = 1 \\ c_2 + \sum_{i=1}^p \phi_{2,i} y_{t-1} + \alpha_{2t} & \text{if } s_t = 2 \end{cases} \quad (3.11)$$

y_t is the variable in use, in this case, it is the inflation rate of South Africa. The process of $\{s_t\}$ makes assumption that the values of $\{1,2\}$ imply the regime at time $\{t\}$. To be more precise $s_t = 0$ represents the state of low inflation while $\{s_t = 1\}$ represents the state of high inflation regime. α_{1t} and α_{2t} represents the sequences of *iid* random variables with a mean equivalent to zero and a constant variance. The assumption that is made is that $\{s_t\}$ is independent of time, meaning that it is stationary, a periodic and irreducible Markov Chain with transition probabilities that can be expressed as follows:

$$p_{ij} = p_{i|j} = p(s_t = j | s_{t-1} = i) \quad i, j = 1, 2. \quad (3.12)$$

The p_{ij} is the probability that the Markov chain will move from state i during time $t-1$ till state j during time t . The smaller the value of p_{ij} , the more chances that the Markov model tends to stay longer in the i^{th} state.

One of the assumptions about the process is that it depends on the values of y_t and s_t only through s_{t-1} . The process is then expected to move from one state to another.

The other case is when the state does not move but only stays the same. This state occurs under the assumption that, the process of s_t is an irreducible and periodic.

Important feature of probabilities is, they should be non-negative and this is supported by Franses and Van Dijk (2000). Therefore resulting in the following transition matrix:

$$p = \begin{pmatrix} p_{11} & p_{21} \\ p_{12} & p_{22} \end{pmatrix} \quad (3.13)$$

Where both $p_{11} + p_{12}$ and $p_{21} + p_{22}$ are equal to 1. Since the study uses MS-AR, there are four transition probabilities which can be expressed as follows:

$$\begin{aligned} P(s_t = 1 | s_{t-1} = 1) &= p_{11} \\ P(s_t = 2 | s_{t-1} = 1) &= p_{12} = 1 - p_{11} \\ P(s_t = 2 | s_{t-1} = 2) &= p_{22} \\ P(s_t = 1 | s_{t-1} = 2) &= p_{21} = 1 - p_{22} \end{aligned} \quad (3.14)$$

From the above equation, the transition of probabilities is also reflected with the expected duration which is defined as the expected period of which the system will stay in a particular regime. That is:

$$E(D) = \frac{1}{1 - p_{ij}}, j = 1, 2 \quad (3.15)$$

Where D is the duration of regime j .

Franses and Van Dijk (2000) also stated that there are three types of probabilities of each regime that occurs when developing the maximum likelihood ratio, which is one of the crucial steps in constructing a Markov Switching model. First being the forecast of the probabilities for each regime or shift occurring at time t given on the condition that all the sample in use is up to time $t - 1$.

The second probability is the filtered probability which includes all the observations and the time and it is estimated using iterative algorithm. The last probability is regarded as smoothed (which is estimated by using the entire sample size) inference of the regime probabilities. Generally, filtered and smoothed probabilities yields to similar conclusion.

Lastly, according to Hamilton (1990), the commonly used procedure on estimating the parameters of the model is to maximize the log-likelihood function given:

$$\hat{P}_{ij} = \sum_{t=2}^n p(s_t = j, S_{t-1} = i | I_n, \hat{\theta}) = \frac{i | I_n, \hat{\theta}}{\sum_{t=2}^n P(S_{t-1} = i | I_n, \hat{\theta})} \quad (3.16)$$

Where $\hat{\theta}$ is the maximum likelihood estimate of θ .

The parameters that are obtained from the maximum likelihood are used for obtaining the filtered and smoothed inference. However, Ismail (2007) mentioned this method as a disadvantage because the number of parameters that need to be estimated increases. In such instances, Expectation Maximum (EM) is used. The technique starts with the initial estimates of the unobserved regime variable S_t and then finally produces a new joint distribution that will ultimately increase the probability of the data in use. In EM, each iteration will increase the value of the likelihood function which increases the certainty that the final estimates of the parameters can regarded as the Maximum Likelihood estimates.

3.4.2 Artificial Neural Networks

Designing ANN includes consideration of many parameters such as the number of hidden layers, number of hidden neurons, number of output neurons, the correct transfer function to be used, number of iteration, and the momentum rate.

All these parameters are essential when designing an Artificial Neural Network in order to produce a very reliable and non-spurious model. Figure 3.1 shows an example of the basic graphical representation of an Artificial Neural Network.

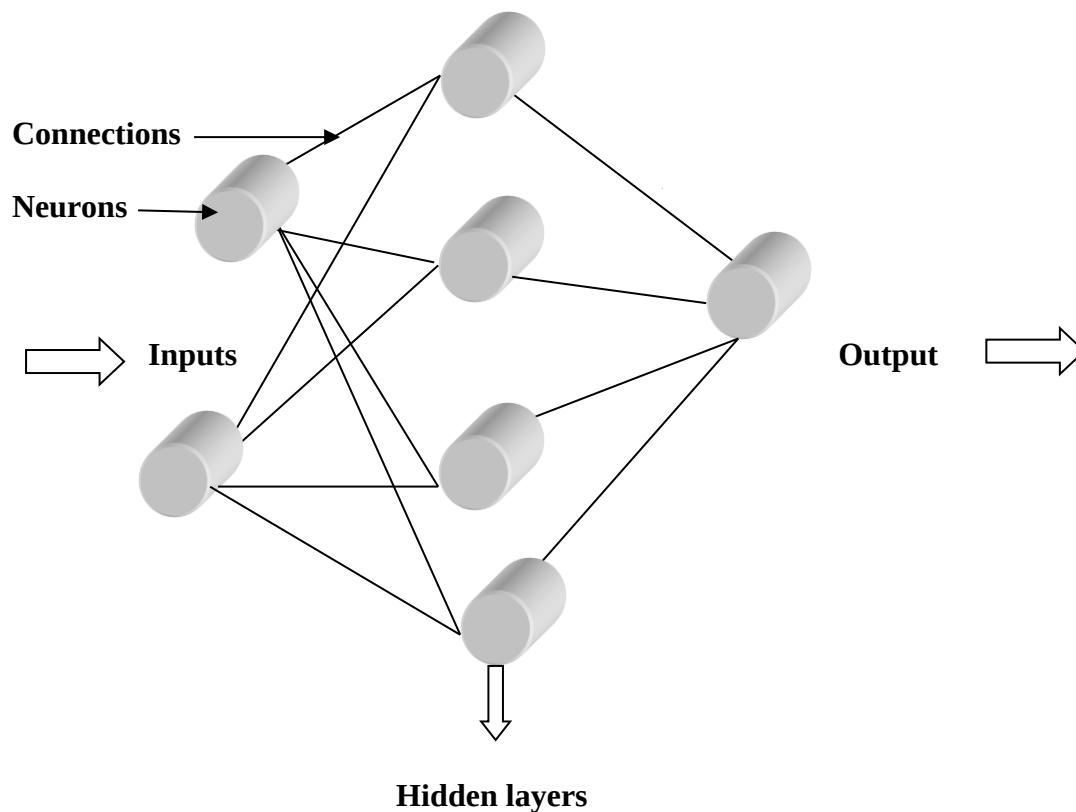


Figure 3.1: Artificial Neural Network

The following are the six steps that are prescribed by Kaastra and Boyd (1996) for preparing data before and when developing an ANN model.

Step 1: Variable selection

As stated the inflation rate of South Africa is used as a variable for conducting the analysis.

Step 2: Data collection

How the data is collected is also outlined on the previous section. As stated, the data was sourced from Quantec on quarterly basis from the first quarter of 1993 to the second quarter of 2016 and therefore the data will be used as raw as it is.

Step 3: Selection of data

Large sample size was selected to accommodate all the assumed structural breaks and instabilities, and inconsistency within the data.

Step 4: Partition the data used

Mostly, the data is divided into three distinct sets namely; training, testing and validation set. As advised by Kaastra and Boyd (1996), the largest set is awarded to the training set mainly because it is used by the whole network for learning the pattern of the data in use, then the testing set ranges between 10 per cent and 30 per cent of the training set, lastly the validation set must only accommodate the most recent observations. The study will follow the percentages used by Leandro and Rosangela (2008).

Training set	Testing set	Validation set
80%	15%	5%

Table 3.1: Data Partitioning

Step 5: Designing of a Neural Network

The following are the sub-steps of the fifth step which is the Neural Network design

Step 5.1: Number of Hidden Layers

In between the input and the output layers, it is always advisable to include one or more hidden layer. The main reason why they are referred to as hidden layers is because they are

hidden/invisible meaning that they do not appear on any external processes that interact with the ANN. The purpose of hidden layers is to enable a network to generalize. Increase of Hidden layers increases computation time and also decreases the problem of over fitting, therefore an increase of hidden layers results in better forecast performed by ANN. The weights included in the Hidden layer will link the number of the hidden layers to the Neurons. The number of observations in use will then determine the probability of over fitting (Baum and Haussler, 1989).

Step 5.2: Number of hidden Neurons

According to Leandro and Rosangela (2008) there is no ideal formula that can be used to compute the number of hidden Neurons. Therefore, most of the researchers opt for experiments. Leandro and Rosangela (2008) further stated that, there is a rule of thumb that has been implemented. That is to use the geometric rule suggested by Masters (1993). With reference to Klimasayskas (1993), there must be five or more times as many training facts which are equivalent to weights which create a boundary on the total number of neurons and inputs. Thus, the current study applies different structures for the whole data and chooses neurons randomly within the hidden layers so as to be able to give a full description of the best structure according to the index.

Step 5.3: Number of output Neurons

There are compelling reasons for always using only one output neuron. For the purpose of compilation of producing a reliable neural network, only one output neuron will be used. Because if multiple neurons are in application, it increases the chances of producing spurious and unreliable results (Masters, 1993).

Step 5.4: The transfer function

There are numerous transfer functions such as, tangent hyperbolicus, arcus tangents, sigmoid and also the linear transfer function. Most of these are unable to handle non-linear data except the sigmoid transfer function which is commonly used. Therefore, sigmoid is used for its relevant feature.

The sigmoid function is computed as follows:

$$f(x) = \frac{1}{1+e^{-ax}}, \quad (3.17)$$

where a permits one set of the slope function. Within this project, the function of $\tanh(x)$ is mostly used.

Step 6: Training ANN

The significance of training a Neural Network assists in learning the patterns of the data. The main objective of this step is to search for sets of weights between the neurons that are useful in determining the global minimum of the error function. The expectation that arises after computing this step is to retrieve favorable generalization.

The gradient descent training algorithm is applied since it is able to make adjustments on the weights to move down the steepest slope. The steps below discuss how the selection of the learning rate, how the learning stopped and the values of momentum used was achieved.

Step 6.1: Number of Iterations

Training is affected by many factors such as the selected learning rate and the value of momentum. Therefore, making it very difficult to have a general and fixed number of maximum runs to make. Kaastra and Boyd (1996) recommended that since there is no fixed number of training iterations, any study can randomly select the number of iterations. A number of the competitive training iteration was used randomly in order to select the best pattern of each index.

Step 6.2: Learning Rate

When training the Neural Network, a learning training that is too high is as a result of error function that is not consistent and does not reflect any improvements. According to Haykin (2001), a very small learning rate requires more time for training the Neural Network, therefore it is always advisable to start training at a higher rate than decrease till it is satisfactory. The learning rates that will be considered will be from 0.1 to 0.9, beginning with 0.9 since it is the highest learning rate to be considered for the purpose of this study.

Step 6.3: Momentum Value

The momentum is mainly necessary because it causes the weight changes not to rely on more than one input pattern. It always ranges from 0 to 1. Similarly to the selection process of the learning rate, the momentum value is also selected randomly.

3.5 Model selection criterion

Wasserman (2000) explicitly defined model selection as the problem of the usage of data to select one model among the list of competing models. There are a number of model selection criteria that can be used to determine the best model that fits the data set of choice.

Model selection criterion uses information from each model to summarize data evidence in support of a model. In summary, model selection criteria accentuate the minimum amount of information that is needed to express both the model and the dataset.

According to Verbeek (2004), the most commonly used model selection criterion is AIC which was developed by Akaike (1969). The AIC is employed in this study because of its popularity. AIC is calculated as follows:

$$AIC = 2k - 2\ln(L) \quad (3.18)$$

where L is the maximum value of the likelihood of the function of the model, then k is the number of all the estimated parameters of the model.

3.6 Residual diagnostics

After developing both the ANN and the MS-AR Model, it is very crucial to determine important model assumptions before performing any inference. If any of the important assumptions is being violated, then succeeding inference techniques may be highly spurious leading to wrong conclusions. Therefore, it is very important to perform relevant model diagnostics. The following diagnostic test will be employed.

3.6.1 Normality test

Normality test is one of the important assumptions made about the selected model. It is essential for checking if the data used conforms to normal distribution standards. This study used a formal test called Jarque-Bera (JB) which was developed by Jarque and Bera (1980).

This test is based on skewness which is supposed to be equivalent to or around zero, reflecting a symmetric distribution where the mean, median and mode are equal. It is also based on kurtosis which should be equal to or around three, indicating perfect tallness of the data.

The first step of constructing JB is to estimate the regression equation

$$y_t = \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \cdots + \beta_k x_{kt} + \mu_t \quad (3.19)$$

Where y_t = dependent variable specifically the inflation rate

β = is the slope

Then make μ_t the subject of the formula;

$$\mu_t = y_t - \beta_1 + \beta_2 x_{2t} + \beta_3 x_{3t} + \cdots + \beta_k x_{kt} \quad (3.20)$$

The second step is to compute the skewness and kurtosis of the data which is expressed as follows respectively:

$$s = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^3}{(N-1)S^3} \quad (3.21)$$

$$k = \frac{\sum_{i=1}^N (Y_i - \bar{Y})^4}{S^4} \quad (3.22)$$

were S is the standard deviation, Y is the actual data $\bar{Y}$ is the mean and N is the grand total of the whole data set in used. After computing the prerequisites of the normality test, the last step is to calculate the JB test which is calculated using the following equation:

$$JB = \frac{N-k}{6} \left[s^2 + \frac{(k-3)^2}{4} \right] \quad (3.23)$$

where

N = number of observations

S = skewness of a variable

k = kurtosis of a variable

When computing the JB test the results are expected to be insignificant because the null hypothesis is that, the data is normally distributed while the alternative is that the data is not normally distributed. The conclusion is based on the benchmark that if p-value is less than 0.05 then the null hypothesis is rejected.

3.6.2 Homoscedasticity

The Autoregressive Conditional Heteroscedasticity (ARCH) effect is mainly concerned with the relationship within heteroscedasticity. It is mostly noticeable when there is instability or a lot of variance in any data when it comes to the following model:

$$Y_t = \beta_0 + \beta_1 X_T + u_t$$

$$u_t \sim N(0, \alpha_0 + \alpha_1 u_{t-1}^2) \quad (3.24)$$

The assumption made about the above equation is that the error term u_t is normally distributed with a zero mean and a conditional variance that depends on the squared error term lagged for only a single period:

$$\theta_t^2 = \text{var}(u_{t-1} u_{t-2} \dots) = E(u_t^2 | u_{t-1} u_{t-2}) \quad (3.25)$$

where θ_t^2 represents the conditional variance of the error term. Then the ARCH effect is modeled by the following equation:

$$\theta_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 \quad (3.26)$$

Equation 3.26 represent the ARCH (1) model and it has only one lag on the squared error term and extending the lags till ARCH (q) is permitted. The ARCH test was developed by Engle (1982), where the following steps were considered for the compilation of the test. Firstly, to develop a regression model using the OLS method, run the secondary regression model using the OLS method then collect and square the residuals. That is:

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \dots + \alpha_p u_{t-p}^2 + v_t \quad (3.27)$$

where u is the residual from the initial regression and p lags are included in this secondary regression. There are two ways to determine the number of lags that is through the information criterion or by the span of data. For the purpose of this study, 4 lags are applicable since the data used is on quarterly basis.

The proceeding step is to calculate R^2 statistic in order to compute T^*R^2 where T is the number of observations. The ARCH test follows a chi-squared distribution with p as degrees of freedom. The null hypothesis of the ARCH test is that there is no Arch effect while the alternative is that the ARCH test is present. That is:

$$H_0: p_1 = p_2 = \dots p_m = 0$$

$$H_1: \exists p_k \neq 0$$

where m is the maximum number of lags that is included in the ARCH effect.

The null hypothesis is then rejected if the p-value is less than the test statistic of the ARCH test effect.

3.6.3 Test for Autocorrelation

According to Gujarati (2010) autocorrelation is defined as the relationship between members of the observations ordered in time and it is always associated with time series data. There are two main formal tests that can be used to detect autocorrelation, that is, Durbin Watson test and the Breusch-Godfrey (BG) test also known as the Lagrange Multiplier (LM) test developed by Saikkonen and Luukonen (1988). The study only focuses on implementing the BG test because it is known for being powerful for high order serial correlation that includes high-order autocorrelation estimators. The first step in computing the BG Test is to run the regression model while including the residuals then run the Auxiliary regression $AR(m)$ as follows:

$$\varepsilon_t = \rho_1 \varepsilon_{t-1} + \rho_2 \varepsilon_{t-2} + \cdots \rho_m \varepsilon_{t-m} + v_t \quad (3.28)$$

The equation simply means that, at time t , regress the residual by including both the intercept and the values that are lagged i^{th} residuals up till time $(t - k)$, where k is determined by error based on the AIC. v_t is the disturbance term that follows the white noise process i^{th} zero mean and a constant variance. Then calculate nR^2 , where n is the sample size and R^2 is the coefficient of determination.

Under the null hypothesis all the coefficients of all the lagged residuals are equal to zero, that is:

$$H_0: \rho_1 = \rho_2 = \cdots = \rho_m = 0$$

$$H_1: \rho_j \neq 0$$

In simple terms the null hypothesis states that there is no autocorrelation and alternative being there is presence of autocorrelation. Having the test statistic:

$$F^* = \frac{\sum_{t=1}^n (\hat{\varepsilon}_t - \bar{\varepsilon}) / (m-1)}{\sum_{t=1}^n (\varepsilon_t - \hat{\varepsilon}_t) / (m-1)} \sim F_{\alpha}(m-1, (n-m)) \quad (3.29)$$

The null hypothesis is rejected if $F^* > F_{\alpha}(m-1, (n-m))$, or alternatively, if the sample size is large

$$nR^2 \sim \chi^2_k \quad (3.30)$$

That is, the product of the coefficient of determination and the sample size follows the chi-square distribution having k as the degree of freedom.

3.7 Error Metrics

Forecasting techniques need to be evaluated to assess the powerfulness and accuracy, and there are different perspectives that can be implemented to achieve that. These include determining the distance between the actual and the forecasted values in order to know the errors involved. Since the gist of this study is the discriminatory performance of error metrics in model selection, eight of these metrics are proposed and discussed in the next section.

3.7.1 Mean Absolute Error

Mean Absolute Error (MAE) is one of the error metrics that is used to measure the closeness of the actual data to the predicted data. The word ‘absolute’ means that all the negative signs of the errors will be ignored. It is defined as follows:

$$MAE = \frac{\sum |X_t - F_t|}{m} = \frac{\sum |e_t|}{m} \quad (3.31)$$

Where $|X_t - F_t|$ is the absolute value of errors.

According to Brown (1962), the biggest advantage of using MAE is that, it can be used as a replacement of MSE when determining the optimal levels of inventory

3.7.2 Mean Square Error

Mean Square Error (MSE) is also referred to as the Mean Square Deviation (MSD) of an estimator which is used for measuring the average of the squares of the errors. In simple terms, it is the difference between the estimator and what has been estimated.

The MSE provides for a quadratic loss function because it squares and averages the errors obtained. When squaring, it gives considerably more weight to large errors as compared to lower errors. Therefore MSE could be more useful when dealing with large errors. The mean square error is defined as follows:

$$MSE = \frac{\sum (X_t - F_t)^2}{m} = \frac{\sum e_t^2}{m} \quad (3.32)$$

where X_t is the actual data at point t

F_t is the forecast from the model at point t

e_t is the forecast error at point t

m is the number of observations used in computing the MSE.

Since one of the objectives of the study is to compare two methods, the summation will move from 1 to m where m is the total number of series summed up to t in order to obtain the average and will vary according to the outcomes of the errors incurred.

3.7.3 Root Mean Square Error

Root Mean Square Error (RMSE) is the alternative way of expressing the MSE. The only difference is that, it is achieved by computing the square root of the expression (3.32) that is,

$$RMSE = \frac{\sqrt{(X_t - F_t)^2}}{m} = \sqrt{\frac{\sum e_t^2}{m}} \quad (3.33)$$

The two main advantages of MSE and RMSE are that they provide a quadratic loss function and they also measure uncertainty in forecasting. In accordance with Chai and Daxler (2014), the formula of RMSE has the square root and this error metric shows large number of deviations which helps to outline more robust results which further prevents cancellation of positive and negative error values. RMSE is also recommended when using a large sample size therefore as is a case in this study. Unlike the MAE, the RMSE is said to weigh more and can easily deal with large errors.

3.7.4 Mean Absolute Percentage Error (MAPE_{reg})

MAPE_{reg} is used as a measure that articulates errors in percentages of the actual data. This is one of the well favoured advantages as its use provides an easy way of making extent judgment and evaluating the importance of errors. It can be expressed as follows:

$$MAPE_{reg} = \frac{\sum \left| \frac{X_t - F_t}{X_t} \right|}{m} (100) \quad (3.34)$$

Which can also be expressed in the following form:

$$MAPE_{reg} = \frac{\sum |e_t|}{m} (100) \quad (3.35)$$

3.7.5 Theil's U-Statistic (U-Statistics)

The Theil's U-Statistic (Theil, 1966) is expressed as follows:

$$U \text{ Statistics} = \sqrt{\frac{\sum_{i=1}^m \frac{(\frac{X_t - F_t}{X_t})^2 / m}{(\frac{X_t - FN_t}{X_t})^2 / m}}{m}} \quad (3.36)$$

Where FN_t is a benchmark forecast, for an example it can represent the latest available value or the latest value after seasonality has been considered. The U-statistic also emphasizes the importance of large errors just as the MSE. It can also be considered when making a comparison amongst forecasting methods.

3.7.6 Symmetric Mean Absolute Percentage Error

The Symmetric Mean Absolute Percentage Error (MAPEsym) is one of the error metrics that can easily be influenced by the presence of outliers. Nevertheless according to Makridakis (1982), the disadvantage can be remedied by dividing the forecasting error by the mean of both actual data and the forecasted value. That is:

$$APE = \left| \frac{X_t - F_t}{(X_t + F_t)/2} \right| (100) \quad (3.37)$$

where APE is the Absolute Percentage Error

3.7.7 Geometric Mean of Squares Error

Geometric Mean of Squares of Error (GMSE) is computed through averaging the product of the squares of the errors. The following is an expression of computing the GMSE

$$GMSE = (\pi_t e_t^2)^{\frac{1}{m}} \quad (3.38)$$

3.7.8 The Median Absolute Percentage Error

The Median Absolute Percentage Error (MdAPE) is similar to regular MAPE or the systematic MAPE, the only difference is that, computing the MAPE involves summing up the APE while when computing the MdAPE the percentage errors are sorted from lowest to highest then find the median from that order. According to Makridakis (1982), the well-known advantage of using MdAPE is that it is not easily influenced by outliers. A value of MdAPE is not equivalent to the APE. For example, if the MdAPE is equal to a percentage of 8, it does not mean that the absolute percentage error is also equal to 8%. Rather, it means that 50 % of the absolute percentage errors are less than 10% and the rest of the 50% are over 10%.

3.8 Conclusion

The chapter accommodates all the methods and the techniques that are implemented in order to find the best non-linear model for forecasting inflation rate of South Africa. The tests and measures for the preliminary and primary data analyses were discussed. The data for this study was described. The analyses of the results of such are detailed in the next chapter.

CHAPTER FOUR

RESULTS

4.1 Introduction

This chapter is set to achieve the study objectives as outlined in Chapter 1. Methods used to achieve the set objectives are discussed in Chapter 3 and the description of the data is also given in the same chapter. OxMetrics Version 6.0 and Zaitun time series version 0.2.1 were used to obtain the results for MS-AR and the ANN models respectively. The chapter makes a clear distinction between the preliminary and primary data analyses.

The data used is inflation rates provided by Quantec. The series covers the period 1993 to 2016. The quarterly data was analysed due to the fact that the researcher needed to expose the current performance and trends of inflation rates in South Africa. The reason for using inflation rate is because it is one of the macroeconomic variables that is directly and indirectly affected by other macro-economic variables (Molwantoa, 2013) resulting in fluctuations. The methods of application for the current study are effective on non-linear series, one of the features displayed by an inflation variable. Moreover, using time series data collected at close intervals is also due to the swift reaction of consumers and producers of goods and services to current information. The data was also available during the period of request. The descriptive measures of the data were calculated and summarised in Table 4.1.

Descriptive statistics are used to describe the basic features and characteristics of the actual data. They give summary about the sample and the measures such as central tendency of distribution which shows the estimates of the centre of a distribution of values.

This includes the mean, median and mode. The other well-known measure is dispersion, which measures how spread the values are, example of such is the maximum value, the minimum value and the standard deviation. Table 4.1 gives summary descriptive statistics of the South African inflation rates for the period 1993 to 2016.

Table 4.1: Descriptive statistics summary

Descriptive measure	Inflation
Mean	6.448
Median	6.152
Maximum	13.430
Minimum	0.438
Std. Dev	2.648
Skewness	0.163
Kurtosis	2.942
Jarque-Bera	0.431331
Probability	0.806

From the above table, the mean reports that, on average, from 1993 to 2016, the South African inflation rate was standing at around 6.4 %. This is above the central banks official range of 3 to 6 %. According to the statsSA, inflation rate will continue to be above range as long as the country is experiencing an increase in the prices of food, non-alcoholic beverages and transport. The median value indicates that it takes 6.2 % quarters for the inflation rate of South Africa to increase or decrease. Other interesting factors include the skewness standing at 0.163 which is close 0 and the kurtosis standing at 2.942 which is approximately equivalent to 3.

Having a kurtosis and skewness value of 3 and 0 respectively shows that the actual data is normal in nature (Joereskog *et al.*, 2003). The positive kurtosis values is also an indication that inflation rate is increasing with time. The probability value 0.806 supports the fact that the inflation series conforms to normal standards. The bell shaped histogram shown as Appendix A also confirms the normality of data. This is not surprising since an excess of 50 observations were used for data analysis to protect violation of this assumption (Mordkoff, 2011).

4.2 Preliminary Data Analysis

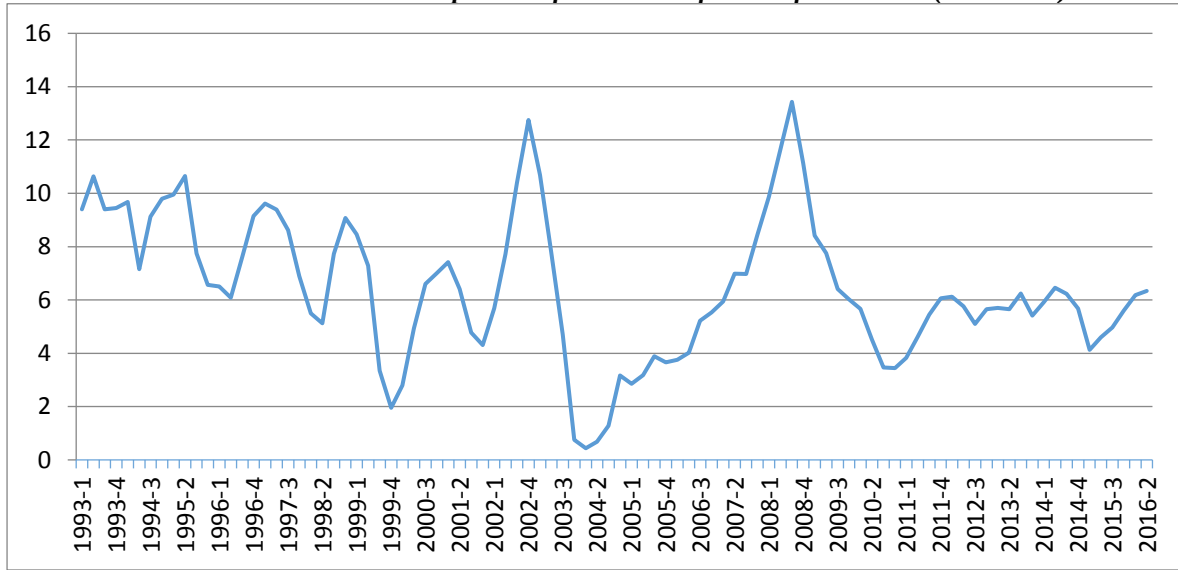
Preliminary data analysis mainly covers the behaviour of the data of interest in order to determine if the data is non-linear in nature, before proceeding to the primary data analysis. The preliminary data analysis includes informal tests and graphs such as the visual inspection of the actual data in order to examine the movement of the South African inflation rates. Lastly, formal tests such as BDS test, CUSUM of squares and the RESET test are carried out to examine if the data follows a nonlinear fashion and that the stability conditions are not violated. These tests are also used to confirm the display by informal tests. The results are discussed in the subsequent sections.

4.2.1 Visual inspection

Exhibit 4.1 displays different patterns and dissimilar behaviour of the data for different quarters across the years. The lowest inflation rates were experienced between 2003Q3 and 2004Q4, followed by 1994Q4. The highest inflation rate in South Africa was experienced between 2002 and 2003 and towards the end of 2008Q4.

With such observations, the critical points of high and low inflation might be in connections to the disturbances and instabilities happening worldwide. For instance, the high inflationary rate in 2008 was triggered by the well-known financial crises that happened in US which affected many inter-reliant countries including South Africa. During this time the spill over effects of this crisis resulted into a recession in South Africa.

Exhibit 4.1: Visual inspection of the South African inflation rates (1993-2016)



The general conclusion that can be made from the visual inspection is that, the South African inflation rate depends on time, *i.e.* does not fluctuate around the same mean therefore, it is non-linear and non-stationary in nature.

4.3.2 Test for non-linearity and stability

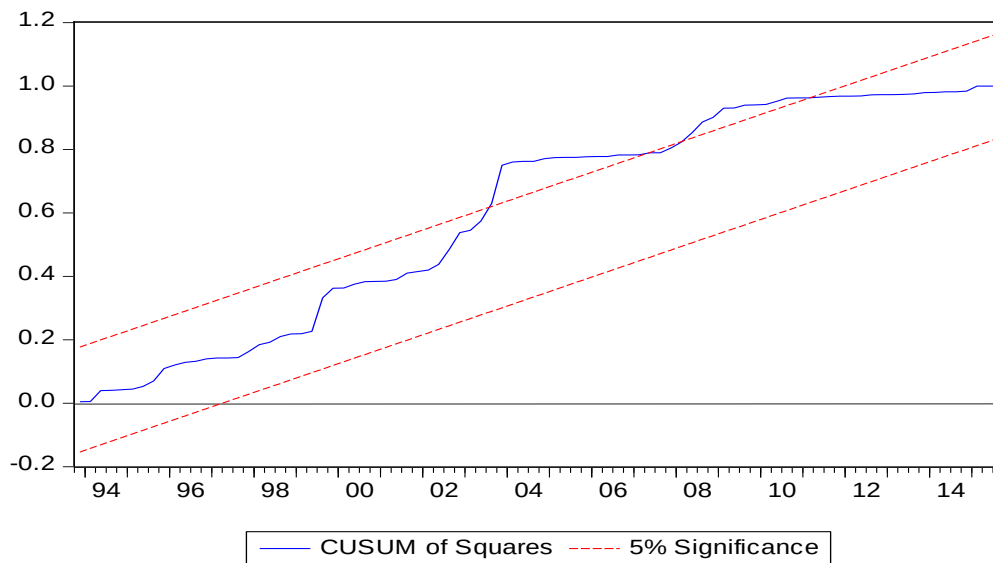
Testing the non-linearity was accomplished through computing the BDS Test, and the RESET test presented in Table 4.2 and Table 4.3 respectively, while the stability test was achieved through CUSUM diagram presented in Exhibit 4.2.

Table 4.2: BDS Test

BDS TEST				
Dimension	BDS Statistic	Std. Error	Z-Statistic	P-value
2	0.112	0.007	16.199	0.000
3	0.176	0.011	15.889	0.000
4	0.212	0.013	16.125	0.000
5	0.226	0.014	16.403	0.000
6	0.227	0.013	17.023	0.000

Table 4.2 shows the results of the BDS test. All the p-values are less than 0.05 for all the specified lags leading to the rejection of identical and independent distribution. Conclusion based on such means that there is a remaining structure within the time series data which has a hidden non-linearity, non-stationary or any other structure missed by model fitting.

Exhibit 4.2 Stability test for inflation rates (1993-2016)



The CUSUM diagram displayed as Exhibit 4.2 confirms that the inflation rate of South Africa had an instable variance during 1993 and 2016 since the time series is not within the 95% critical lines. The findings are similar to those reported by Bonga-Bonga and Kabundi (2015) and Cruz and Mapa (2013).

Table 4.3: RESET test

	T-statistics	F-statistics	Likelihood
Value	4.401	19.371	18.368
P. value	0.000	0.000	0.000

The RESET test displayed in Table 4.3 confirms the results obtained from the CUSUM and the BDS test because the probability value of the RESET test is significant at the value of 0.000.

Hence it can be concluded that the inflation rate of South Africa is unstable and non-linear at levels.

The results of this preliminary analysis were also obtained by numerous authors who were interested in testing the existence of non-linearity within different types of time series data. To name a few, Ismail and Isa (2006) whose objective was to assess whether the regime switching can be used for describing the features on non-linearity using the exchange rate time series data of three different countries reported similar results.

Moshiri and Seifi (2008) who were interested in testing the existence of white noise in the exchange rate data using the BDS test made a conclusion that the data generating process of the original data is non-linear in nature. The study conducted by Xaba *et al.* (2016) reported similar findings using data on the five major banks in South Africa.

4.4 Markov Switching-Autoregressive Model

The Markov switching results were obtain through OxMetrics version 6.1 software. Four-regime Markov switching models ($MS - AR(1)$ to $MS - AR(4)$) were developed using the inflation rates. The main reason for lagging up to 4 was to fit simple models only.

Table 4.4: Model selection criterion results

Models(MS-AR)	Number of states	Number of lags	log likelihood	AIC
1	2	1	-142.89985	3.35332999
2	2	2	-127.247975	3.08422416
3	2	3	-124.330291	3.09841571
4	2	4	-122.627665	3.14086586

The model selection criterion used was the AIC. Based on the results from Table 4.4, $MS(2) - AR(1)$ was excluded as a candidate since the model has the highest AIC of 3.353 as compared to the AIC values of the models. The remaining competing models are $MS(2) - AR(2)$ to $MS(2) - AR(4)$. Amongst the qualifying models competing models $MS(2) - AR(2)$ was selected as the final model because it has the lowest AIC value of 3.084.

Table 4.5: $MS(2) - AR(2)$ Model

Parameter	Coefficients	Stand Error	t value	prob
C1	5.18859	0.5166	10	0
AR1	1.24155	0.1106	11.2	0
AR2	1.38855	0.09637	14.4	0
sigma1	0.603272	0.07489	8.06	0
C2	7.12249	0.5017	14.2	0
AR1	-0.373067	0.09552	-3.91	0
AR2	-0.666631	0.1105	-6.03	0
sigma2	1.1955	0.1757	6.81	0

The parameter estimates of $MS(2) - AR(2)$ reflect that all the coefficients are all significant since they have significant observed p-values equal to 0 meaning that they are all less than the significance levels 0.01, 0.05 and 0.10.

Table 4.6 presents the regime switching results. The intention is to determine the manner in which inflation rates of South Africa switch between high and low levels. Regime 0 represent a state of having a low inflation rate, while regime 1 indicates the probability of having high inflation rate.

Table 4.6: Regime switches of inflation rates (1993-2016)

Regime 0 (Low inflation rate)	Regime 1(high inflation rate)
0.94252	0.057477
0.057119	0.94288

The transition probability metrics is found to be $p(S_t = 0|S_{t-1} = 0) = 0.94252$
 $p(S_t = 1|S_{t-1} = 1) = 0.94288$.

The result exposes that inflation rate is more persistent in regime 1 than in regime 0 with a slight difference of 0.00036. When the process is in regime 0 a probability of switching to that low inflation is $P(S_t = 1|S_{t-1} = 0) = 0.057119$.

Table 4.7: States of inflation rate

Low Inflation	average duration	High Inflation	average Duration
1994(2) - 1994(2)	14.67 quarters.	1993(3) - 1994(1)	15.00 quarters
2003(4) - 2007(2)		1994(3) - 2003(3)	
2008(4) - 2015(3)		2007(3) - 2008(3)	

The results in Table 4.7 are in support of those displayed in Table 4.6. The average duration of inflation rates in Regime 0 (low inflation) and Regime 1 (high inflation) is 14.67 and 15.00 quarters respectively, this is a slight difference of 0.4. This means that the inflation rate of SA is lightly to be in a state of high inflation than in a state of low inflation.

The aim of this study is to determine the discriminatory performance of the error metrics in MS-AR and ANN models. The previous discussions followed a blue print of MS framework to construct and decide on the most applicable model to compete with the ANN model. Subsequent analyses were done through the guidance of the ANN model.

4.5 Artificial neural networks results

The results presented in this section were also obtained for inflation rates collected for the period 1993-2016. Table 4.8 gives a summary ANN model results. Zaitun time series version 0.2.1 was used to generate these results.

Table 4.8: Network Architecture -Model Summary

Variable	Inflation
Included Observation	82 after adjusting endpoint)
Network Architecture	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Functions	Sigmoid Function
Propagation Learning	
Learning Rate	0.8
Momentum	0.9
Criteria	
Error	0.012

The model was selected on the fact that it has the minimum errors approaching zero. As recommended by Masters(1993), only one output layer of Neuron was applied in order to achieve non-spurious and reliable results, the number of hidden layers become higher because the higher the number of hidden layers, the lesser the chances of over fitting.

Selection of 12 input layer neurons was done randomly since this does not affect the fitting of the model. Numerous authors have used 12 input layers in their studies. Sigmoid activation function was applied because it is a function which can handle non-linear data (Panchal and Panchal, 2014). According to the visual inspection of the actual data, CUSUM diagram and BDS test, the inflation series of South Africa possess nonlinear characteristics necessitating the application of the sigmoid function.

The effective learning rate ranges from 0.1 to 0.9. It is advisable to start with the highest learning rate, which is 0.9 because a lower learning rate requires a lot of training which can cause meaningless results. To be accurate, all the possible learning rates were conducted (from 0.9 to 0.1) which are recorded from Appendix B to Appendix I. It was found that the best ANN architecture model was having a learning rate of 0.8 (from Table 4.8) as it resulted in the lowest overall error in comparison to other learning rates.

4.6 Residual diagnostics results

Once the models have been fitted, it is important to assess whether the error terms of both models meet the assumptions before producing the forecasts. If there is an existence of violated consequent inferential procedures, the results are then considered to be meaningless. Therefore, it is important to perform the model diagnosis. This is achieved through estimating the probability values of the tests for serial correlation, heteroscedasticity and normality recorded in Table 4.9 below.

Table 4.9: Residual Diagnostics

	MS-AR		ANN	
Formal Tests	Statistic	p-values	statistic	p-value
Serial correlation(Breusch-Godfrey test)	16.638	0.1637	76.987	0.147
Heteroscedasticity (ARCH test)	0.0060823	0.938	60.066	0.711
Normality(Jarque-Bera test)	2.4151	0.2989	2.264	0.103

All the probability values recorded in Table 4.9 above are above a significant value of 0.05, meaning failure to reject the null hypotheses for the three assumptions made about the model error terms. This implies that there is no autocorrelation, meaning that the covariance and correlations between different disturbances are equal to zero resulting in serial independence. Furthermore, the two model residuals are homoscedastic, *i.e.* their variances are constant as expected and the normality assumption of these residuals is also protected. This finding is expected since more than 50 observations were used to estimate the models that generated the residuals. The use of samples in excess of 50 observations places the analyst in a secure position as far as violation of model residuals is concerned.

4.7 Forecasts from the $MS(2) - AR(2)$ and ANN

Exhibit 4.3 gives a summary of actual and predicted forecasts for inflation rate.

Exhibit 4.3 : In-sample forecasts of $MS(2) - AR(2)$

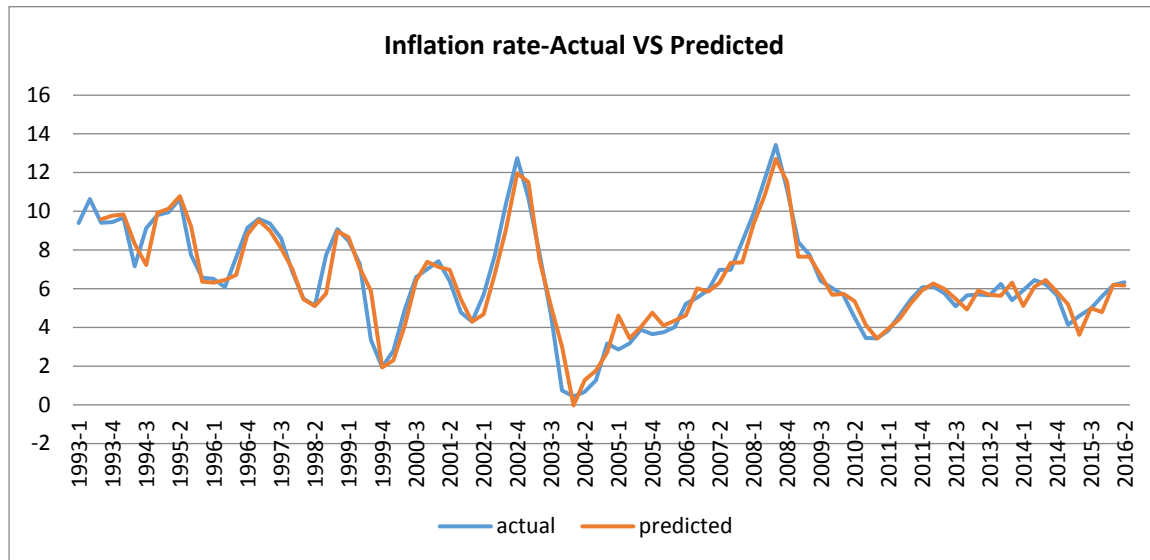
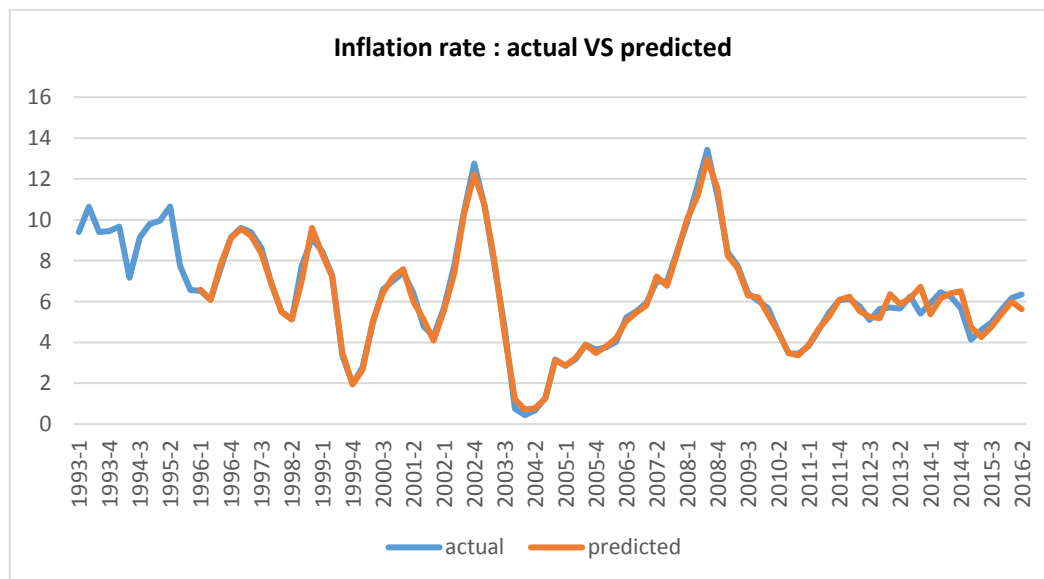


Exhibit 4.4 : In-sample forecasts for ANN



According to the battery of diagnostic test conducted, both the models are able to produce meaningful results but interestingly, exhibition 4.4, which is the in sample of the ANN, clearly indicates that the predicted values traces off the actual values. This implies that, the amount of errors involved in the estimation of the model are less, while exhibition 4.3 shows that through

the predicted values traces the actual data, the fact that they are not closely related shows that there are many errors involved. Appendix K further confirms that by showing that there is accuracy and a strong correlation between the model predictions and the actual values.

Since the $MS(2) - AR(2)$ and the ANN models were found to be fit according to Exhibits 4.3 and 4.4. A battery of diagnostic tests confirmed the two models to be a good fit for the data. 1. The actual and predicted values of the ANN are closely correlated as revealed by Exhibit 4.4 than the regime switching model. This implies that ANN is well fitted for the data.

4.8 Forecasting accuracy

Estimated models need to be evaluated from different perspectives. In this study, the evaluation is done through comparing error metrics. The gist of the study is to determine the discriminatory performance of difference forecast error metrics using the Markov Switching Autoregressive Model and Artificial Neural Networks model. The results are summarised in Table 410. The model which generated minimum forecast errors is favoured.

Table 4.10 : Error metrics results

ERROR METRICS	MS(2)-AR(2)	ANN
MAE	0.525	0.224*
MSE	0.532	0.097*
RMSE	0.076	0.035*
MAPE _{reg}	52.541	22.396*
Theils u test	7.136	0.247*
MAPE _{sym}	200	200*
Geometric Mean of squares	0.532	0.089*
MdAPE	0.525	0.225*

**implies that the associated model generated fewer errors*

The findings displayed in Table 4.10 reveal that the $MS(2) - AR(2)$ model generated large forecast errors when compared with its counterpart. Based on these results, one could conclude that when ANN structure is applied to the inflation rate of South Africa, it released a model with fewer errors. This implies that the ANN increases prediction capability and may be considered for analysis of inflation rates.

Similar results were obtained by Constantinou *et al.* (2005) who tested forecasting accuracy of the Regime switching and ANN models using the RMSE. The results indicated that RMSE of the Regime switching model was equal to 0.121 while the ANN model has RMSE 0.110. This is an indication that ANN performed better than the Regime switching model.

Exhibit 4.5: depicts the residual plots of both $MS(2) - AR(2)$ and ANN models.

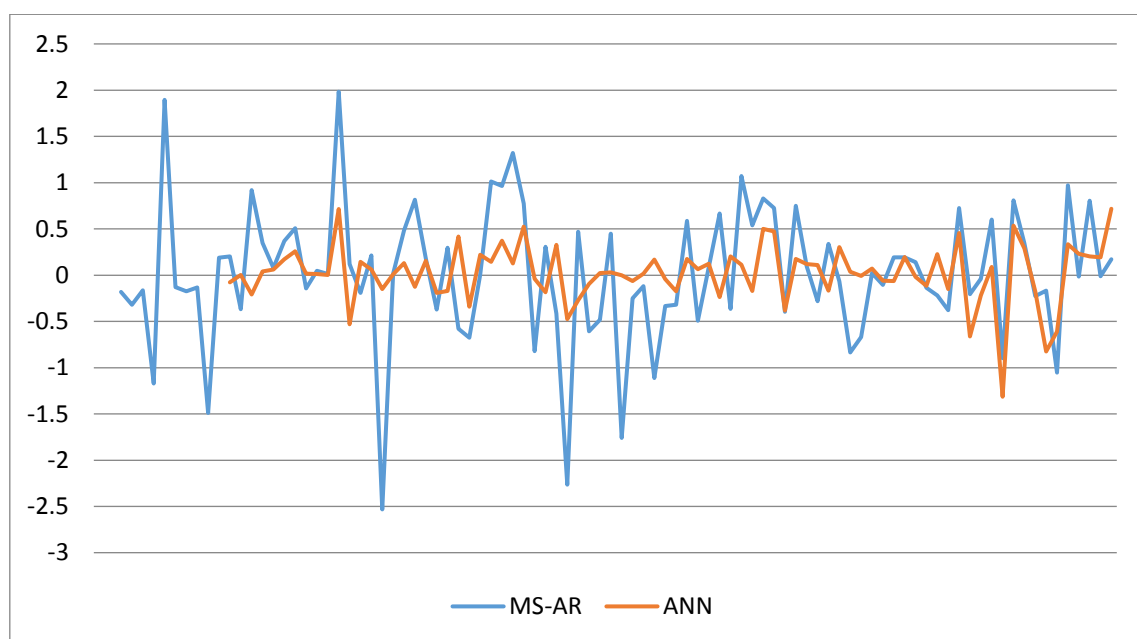


Exhibit 4.5 shows that the residuals of ANN fluctuate around zero while the residuals of $MS(2) - AR(2)$ are within the range of 2 and -2.5. This confirms that ANN outperforms $MS(2) - AR(2)$ model. Moreover, different forecast error metrics were used to calculate the errors associated with each mode and the results obtained are in support of those found in exhibit 4.4.

4.7 Conclusion

This chapter outlined the empirical analysis of the two non-linear models, which is the Markov Switching Autoregressive model and the Artificial Neural Network. The chapter assisted in achieving the objectives as outlined in the first chapter of the dissertation. The primary objective of this study was to explore the discriminatory performance of the error metrics in the Artificial Neural Networks and Markov Switching-Autoregressive models in the context of the South African inflation rates.

The two frameworks were successfully used to estimate the models and the eight error metrics were also computed and used to determine the discriminatory performance of these error metrics in the selected model. The results indicated that inflation rate of South Africa is non-linear in nature, and Artificial Neural Networks out-performs Markov Switching Autoregressive models since lesser error are involved when forecasting.

The next chapter provides a discussion of the findings, the conclusions and recommendations for future studies and policy.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary

There is growing evidence that most of the macro-economic variables such as inflation rate, interest rate, stock prices, exchange rate, etc. have characteristics of non-linearity as there are many factors (both internal and external factors) that can cause sudden fluctuations within the data. The current study considered such facts and performed a discriminatory performance of two non-linear models using the eight error metrics namely: MAE, MSE, RMSE, MAPEreg, Theil's U test, MAPEsym, Geometric Mean of Squares, and MdAPE. The study employed the inflation rate of South Africa from the first quarter of 1993 to the second quarter of 2016. The main reason for selecting inflation rate as a variable is because it is easily influenced by other macro-economic variables such as the exchange rates. The BDS test, RESET test and CUSUM diagram affirmed the choice of variable selected and used.

Forecasting is a key activity in the development of policies. Therefore, it is very crucial when forecasting, that less errors are involved. With this in mind, the findings might be used as guideline for inflation targeting central bank, e.g.; SARB by simply using the forecast as an early warning signal for both high and low inflation rate and it can also help to enhance the monetary policy.

5.2 Research objectives and conclusions

The section accommodates a brief summary of the findings from chapter four in accordance with the objectives outlined in chapter one.

First objective

To determine if inflation rate of South Africa is non-linear in nature.

First conclusion

The data used is the inflation rate of South Africa from the first quarter of 1993 to the second quarter of 2016. The main reason for choosing a large sample was to make certain that most of the events that might have direct or indirect impact on the inflation rate can be captured. Non-linearity is a very important assumption for regime switching and ANN models. It is for this reason that this study sought to confirm the validity of the said assumption.

The results of assessing non-linearity were summarised in Table 4.2 and Table 4.3 representing the BDS test and RESET test respectively. Both the RESET and BDS test revealed that the inflation rate of South Africa is non-linear in nature. These results confirm those of Ismail (2007) study who used exchange rate as a variable to examine non-linearity. Based on this finding, it can be concluded that most of the macro-economic variables or financial time series data is nonlinear in nature. The CUSUM diagram supported this through revealing that the data is also unstable. The values of the series go beyond the 5% significance level as shown in Exhibit 4.2.

Second objective:

To develop models for forecasting inflation rate using ANN and MS-AR frameworks.

Second Conclusion

The development of two models was done after conducting the general primary data analysis. This objective was set to estimate the models using the MS-AR and ANN models.

To develop an MS-AR model, the study followed a framework employed by Cruz and Mapa (2013) who did a univariate analysis of inflation rate. In estimating the ANN model, the study followed a procedure by Leandro and Rosangela (2008). The best model was selected after training the data for several times, through changing the learning rate and the momentum value. The relevant transfer function was the sigmoid function because of one of its characteristics of handling non-linearity within the data. Both the competing models passed all the residual diagnostic assumptions, and were further analysed. $MS(2) - AR(2)$ was selected as the best model based on the value of the AIC.

Third objective:

To assess the discriminatory performance of error metrics in the ANN and MS-AR models.

Third conclusion

One of the important aspects of using error metrics is to use them as an evaluation criterion; they can easily and effectively discriminate among different model results. The lesser the value of error metric the better the model precision power.

The gist of the study was to perform a discriminatory performance of error metrics on the two non-linear models. As indicated, the eight observed error metrics selected the ANN. It was therefore, concluded that the ANN model is a good predictor of inflation rate since the errors produced by this model were lesser than those of the other model.

5.3 Recommendations

This section formulates the recommendations for future studies and policy purpose.

Recommendation 1

The study was only limited to one macro-economic variable, that is inflation rate of South Africa, while there are many other macro-economic variables to consider such as the exchange rate and the interest rate. The study suggests studies that incorporate other macro-economic and financial variables as determinants of inflation rate. In particular, the current study recommends studies that will extend the models used so as to confirm their suitability at multivariate level.

Recommendation 2

The study also considered only two non-linear models, namely: MS-AR model and ANN, whereas there are numerous models such as Smooth Transition Regression (STR), Threshold Autoregressive (TAR) model, etc. It could be interesting to also evaluate the performance of other non-linear models as perhaps they can provide better forecast with lesser errors as compared to that obtained from the ANN model.

Recommendation 3

The results indicated that ANN outperforms the MS-AR model. Therefore, an ANN can be used as a complementary tool for forecasting inflation rate of South Africa. For this reason, the SARB can use the results to achieve price stability that is, keep the inflation rate within the standard and favourable range and generally assist forecasters in using the correct time series technique considering the time of data at hand. The results can also be used to enhance the monetary policy of South Africa.

5.4 Conclusion

The chapter provided a summary of the whole study. The discussions made in this chapter were around the objectives outlined in Chapter 1. The data was successfully assessed for non-linearity, stability and non-linearity prior to the main analyses. The two non-linear methods employed in the study were also successfully used to develop inflation model and the findings confirm the ANN as being the best predictor of this sector for the selected period. Upon being subjected to diagnostic testing the two model residuals did not violate the assumptions confirming their significance in producing the forecasts. Generally, the ANN produced lesser errors compared to the $MS(2) - AR(2)$ model and the eight error metrics further picket the former as the best model.

REFERENCES

- Ahangar, R.G., and Yahyazadehfar, M.** (2010). 'The Comparison of Methods Artificial Neural Network with Linear Regression Using Specific Variables for prediction Stock Price in Stock Exchange'. *International Journal of Computer Science and Information Security*, 7:2.
- Akaike, H.** (1969). 'Fitting autoregressive models for prediction'. *Annals of the Institute of Statistical Mathematics*, 21:243-247.
- Aron, J., and Muellbauer, J.** (2000). 'Inflation and Output Forecasting for South Africa: Monetary Transmission Implications'. Working Paper, Oxford: Centre for the Study of African Economics, Institute of Economics and Statistics.
- Bahramian, P., Balcilar, M., Gupta, R., and Kanada, P.T.** (2014). 'Forecasting South African Inflation Using Non-Linear Models: A Weighted Loss-based Evaluation, Working paper series: 2014-2016.
- Balcilar, M., Gupta, R., and Miller, S.M.** (2012). 'The Out-of-sample Forecasting Performance of Non-linear Models of Regional Housing Prices in the US'. University of Pretoria Department of Economics Working Paper No. 2012-26.
- Barnett, W.A., Gallant, R.A., Hinich, M.J., Jungeilges, J.A., Kaplan D.T., and Jensen, M.J.** (1997). 'A Single-blind Controlled Competition among Tests for Nonlinearity and Chaos', *Journal of Econometrics*, 82:157-192.
- Baum, E.B and Haussler, D.,** (1989). 'What size net gives valid generalization?' .*Neural Computation*, 6:151-160.
- Binner, J.M., Bissoondeal, R.K., Elger, T., Gazely, A.M., and Mullineux, A.W.** (2005). 'A Comparison of Linear Forecasting Models and Neural Networks': *Application to Euro International and Euro Divisia*, *Applied Economics*, 37:665-680.

Bonga-Bonga, L. and Kabundi, A. (2015). Monetary policy instrument and inflation in South Africa: Structural vector error correction model approach.

Boot, T. and Pick, A. (2014). 'Optimal forecasts from Markov switching models'. *DNB Working Paper No. 452*.

Brock, W. A, Dechert, W. D., Scheinkman, J. and LeBaron, B. (1996). 'A test for independence based on the correlation Dimension'. *Econometrics Reviews*, 115:197-235.

Brock, W. A, Dechert, W. D. and Scheinkman, J. A. (1987), 'A Test for Independence Based On the Correlation Dimension', 'Department Of Economics, University Of Wisconsin at Madison, University Of Houston, and University of Chicago.

Brown, R.L and Evans. J.M. (1975). 'Techniques for Testing the Constancy of Regression Relationships over Time'. *Journal of the Royal Statistical Society. Series B (Methodology)*, 37, 2:149-192.

Brown, R.L. (1962). 'Smoothing, Forecasting and Prediction'. *Prentice-Hall, Englewood Cliffs, New Jersey*.

Chai, T., and Draxter, R.R. (2014). 'Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature'—*Geological science. Model Development*. 7:1247–1250.

Chang, P.C., Wang Y.W., and Liu, C.H. (2007). 'The development of a weighted evolving fuzzy neural network for PCB sales forecasting', *Experts system with applications* 32: 86-89.

Chauvet, M., Lima, E.C.R., and Vasquez, B. (2002). 'Forecasting Brazilian Output in the Presence of Breaks'. *A Comparison of Linear and Nonlinear Models, Federal Reserve Bank of Atlanta Working Paper No.2002-28*.

Clements, M P., Frances, P.H., and Swanso, N.R. (2003). 'Forecasting Economic and Financial Time-series with Non-linear Models', Rutgers University Departmental Working Paper No. 200309.

Clements, M. P., Franses, P. H., and Swanson, N. R. (2004). 'Forecasting economic and financial time-series with non-linear models', *International Journal of Forecasting*, 20: 169-183.

Clements, M.P. and Hendry. (2002). 'An overview of Economic Forecasting, Economic Forecasting', *Economic and Social Research Council, Working Paper*.

Constantinou, E., Georgiades, R., Kazandijian, A. and Kouretas, G.P., (2005) 'Regime Switching and Artificial Neural Forecasting of the Cyprus Stock Exchange Daily Returns'. *International Journal of Finance & Economics*.4:371-383

Cruz, C.J. F., and Mapa, D.S. (2013). 'An Early Warning System for Inflation in the Philippines using Markov-Switching and Logistic Regression Models.' *Working paper series*.

Dase, R.K and Pawar D.D. (2010). 'Application of Artificial Neural Network for stock market Predictions: A review of literature'. *International Journal of Machine Intelligence*, ISSN: 2:14-17.

De Faria, E.L. and Gonzalez, J.L. (2009), 'Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods'. *Expert Systems with Applications Article in Press*.

Diebold, F.D. (2012). Comparing predictive Accuracy, Twenty Years Later: A personal Perspective on the Use and Abuse of Diebold-Mariano Tests, *NBER Working Paper No 18391*.

Engle, R.F. (1982). 'Autoregressive Conditional Heteroscedasticity with Estimates of Variance of United Kingdom Inflation', *Econometrica*, 50(4):987-1008.

Franses, H.P. and Van Dijk, D., (2000). *Nonlinear Time Series in Empirical Finance*. Cambridge University Press: Cambridge.

Friedman, M. (1977). 'Nobel lecture: Inflation and Unemployment'. *Journal of Political Economy*, 85:451-472.

Godfrey, L. G. (1988). *Misspecification Tests in Econometrics*. Cambridge: Cambridge University Press.

Green, J. H. (1996). 'Inflation Targeting: Theory and Policy Implication', *IMF staff papers*, 43(4):779-795.

Gujarati, D. (2010). *Basic Econometrics*. New York: McGraw Hill, Inc.

Hamilton, J.D. (1989). 'A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle', *Econometrica*, 57(2):357-384.

Hamilton, J.D. (1990). 'Analysis of Time Series Subject to Changes in Regime'. *Journal of Econometrics*, 45:39-70.

Haykin, S. (1994). 'Neural Network: A comprehensive foundation, Macmillan College Publication, New York.

Haykin, S. (2001). 'Neural Networks for financial Forecasting'. John Wiley and Sons, New York.

Hays, W. (1973). 'Statistics for the social sciences (2nd edition)' New York: Holts.

Ismail, M.T. (2007). 'Detecting regime shifts in Malaysian exchange rates', *Journal of Fundamental Sciences*, 3:221-224.

Ismail, M.T. and Isa, Z. (2006), 'Modelling Exchange Rates Using Regime Switching Models', *SainsMalaysiana*, 35(2):55-62.

Jarque, C.M and Bera, A. (1980), 'Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals'. *Economics Letters* 12:255-259.

Jarque, C.M., and Bera, A.K. (1987). 'A Test for Normality of observations and Regression residuals', *International Statistical Review*, 55:163-172.

Jensen, M.J. (1997). 'A Single-blind Controlled Competition among Tests for Nonlinearity and Chaos '. *Journal of Econometrics*, 82:157-192.

Joereskog, KG., Sorbom, D., DuToit, S. and DuToit, M. (2003) LISREL 8: *New Statistical Features*. Third printing with revisions. Lincolnwood, IL: Scientific Software International.

Jonsson, G. (1999). 'The Relative merits and implications of inflation targeting for South Africa'. IMF Working Paper, International Monetary Fund.

Kaasta, I. and Boyd, M. (1996). 'Designing a Neural Network for forecasting financial and economic time series'. *Neurocomputing*, 10:15-236.

Klimasayskas, C.C. (1993). Applying Neural Networks. 'Neural Networks in Finance and Investing: Using Artificial Intelligence to Improve Real world Performance'. 64-65. Chicago.

Kuan, C.M. and White, H. (1994). 'Artificial Neural Networks: An Econometric perspective (with discussion)', *Econometric Reviews*, 13:1-9, 139-143.

Leandro, S. M. and Rosangela. B. (2008). 'Design a Neural Network For Time Series Financial Forecasting: Accuracy and Robustness Analysis: Brazil'.

Liu, G.Q. (2011). 'Comparison of regression and ARIMA models with neural network models to forecast the daily streamflow of White Clay Creek.'

Makridakis, S. (1982). 'The accuracy of Extrapolative (Time Series Methods): Results of a Forecasting Competition'. *Journal of Forecasting*, 2:15-39.

Makridakis, S., Anderson, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E. and Winkler, R. (1982). The accuracy of extrapolating (time series) methods: Results of a forecasting competition". *Journal of Forecasting*, 1: 111-153.

Maliki, O.S., Agbo, A.O., Maliki, A.O., Ibeh. L.M. and Agwa, C.O. (2011). 'Comparison of Regression Model and Artificial Neural Network Model for the prediction of Electrical Power generated in Nigeria'. *Advances in Applied Science Research*, 2(5):329-339.

Masters, T. (1993). 'Practical Neural Network Recipes in C++'. *Academic press*, New York.

Molwantoa, L. (2013). 'South African Inflation Forecasting Using Genetically Optimised Neural Network: University of Johannesburg'. Retrieved from: <https://ujdigispace.uj.ac.za>
Access date: 6 August 2015.

Mordkoff, J. (2011). The assumption (s) of normality. *Unpublished manuscript*.

Moroke, N.D. (2015). 'Box-Jenkins Transfer Function framework applied to saving-investment nexus in the South African context', *Journal of Governance and Regulation*, 4 (1): 63-77.

Moshiri, S. and Seifi, F. (2008). 'Nonlinearity in Exchange Rates and Forecasting'. *Iranian Economic Review*: 13.21.

Moshiri, S.N. (1997). 'Neural Network Versus Econometric Models in Forecasting inflation, Working paper, Department of Economics', University of Manitoba.

Mostafaei. H. and Safaei. M. (2012). 'Point Forecasting Markov Switching Model for U.S. Dollar/Euro Exchange Rate'. *SainsMalaysiana*, 41:481-488.

Nakamura, E. (2006). 'Inflation forecasting using neural network'. *Economic Letters*, 86: 373-378.

Nunoo, I.K.E. (2013). 'Inflation. Forecasting in Ghana-Artificial Neural Network Model Approach: Ghana'.

Panchal, F.S. and Panchal, M. (2014) 'Review on Methods of Selecting Number of Hidden Nodes in Artificial Neural Network'. *International journal of Computer Science and Mobile Computing*, 3(11): 455-464.

Potter, S.M. (1999). ‘Nonlinear Time Series Modelling: An Introduction’. *Federal Reserve Bank of New York*, 10045.

Pretorius, C. J. and Janse Van Rensburg, T.N. (1996). ‘The Forecast performance of Alternative Models of Inflation. Working Paper’, South African Reserve Bank.

Ramsey, J.B. (1969). ‘Test for Specification error in Classical Linear Least Squares Regression Analysis’. *Journal of Time Series Analysis*, 30:350-371.

Saikkonen, P., and Luukonen, R. (1988). ‘Lagrange Multiplier tests for testing non-linearity in time series models’. *Scandinavian Journal of Statistics*, 15:55-68.

Theil, H. (1966.) *Applied Economic Forecasting*, North-Holland Publishing Co., Amsterdam.

Tiao, G.C. and Tsay, R.S. (1994). ‘Some advances in non-linear and adaptive modelling in time-series’, *Journal of forecasting*, 13: 109-3.

Tjung, L. C., Kwon O., Tseng, K.C., and Bradley-Geist J. (2010). ‘Forecasting Financial Stocks using Data mining’, *Global Economy and Finance Journal*, 3 (2):13-26.

Tsay, R. S. (2002). Nonlinear models and forecasting, in M. P. Clements and D. F. Hendry (eds), ‘A Companion to Economic Forecasting, Blackwell, Oxford’, 453-484.

Uko, A. and Nkoro, E. (2012). ‘Inflation forecasts with ARIMA, Vector Autoregression and Error Correction Models in Nigeria’. *European Journal of Economics, Finance and Administrative Science*, 1:71-87.

Vaisla, K.S., and Bhatt. (2010). ‘Neural Network Technique for Stock Market Forecasting’. *International Journal on Computer Science and Engineering*, 2:2104-2109.

Van den Heever, J. P. (2001). ‘A Note on Inflation Targeting in South Africa’, *The South African Journal of Economics*, 69(1): 168-177.

Van der Merwe, E. (2004). 'Inflation Targeting in South Africa'. *South African Reserve Bank*, 1(Occasional paper No 19).

Verbeek, M. (2004). *A Guide to Modern Econometrics*. 2nd ed. New York: John Wiley and Sons, 234.

Wasserman, L. (2000). 'Asymptotic Inference for Mixture Models Using Data-Dependent Priors. *Journal of Royal Statistical Social Services*, B 25, 318-329.

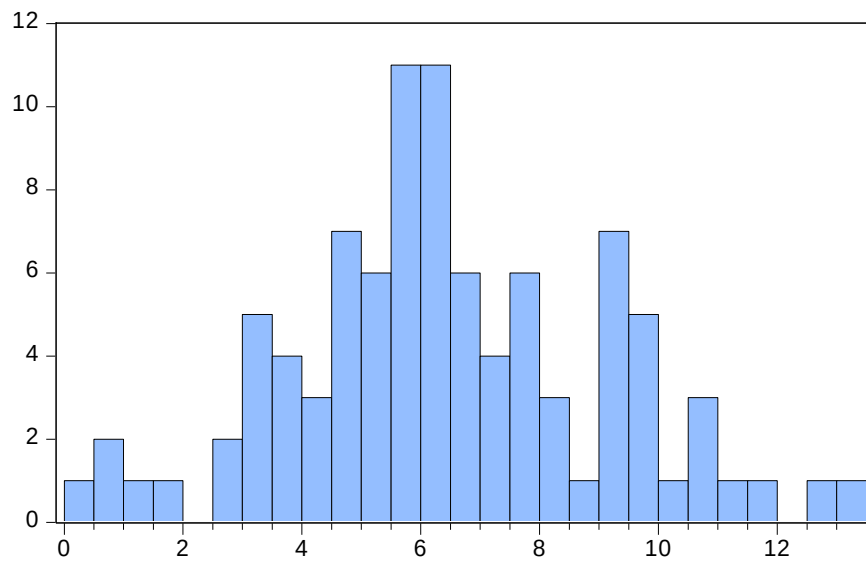
Xaba, D.L., Moroke, N.D., Arkaah Y.J. and Pooe, C. (2016).Modelling South African major banks' closing stock prices: A Markov-Switching Approach, *Journal of Economic and Behavioural Studies*, 8(1): 36-40.

Yarmohammadi, .M., Mostafaei .H., and Safaei, M. (2012), 'Markov Switching Models for Time Series Data with Dramatic Jumps', *Sains Malaysian*, 41(3):371-377.

Yu, J.C. (2007). 'Forecasting Inflation rate in the Philippines: Linear vs Markov- Switching model'. 10th National Convention on Statistics (NCS).

APPENDICES

Appendix A



Series: INFLATION_RATE
Sample 1993Q1 2016Q2
Observations 94

Mean 6.448501
Median 6.151950
Maximum 13.43040
Minimum 0.437600
Std. Dev. 2.648131
Skewness 0.163373
Kurtosis 2.941999

Jarque-Bera 0.431331
Probability 0.806005

Artificial Neural Network

Appendix B

Variable	INFLATION
Included Observation	82 after adjusting endpoint)
Network Architecture	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Functions	Sigmoid Function
Propagation Learning	
Learning Rate	0.9
Momentum	0.9
Criteria	
Error	0.0598

Appendix C

Variable	INFLATION
Included Observation	82 after adjusting endpoint)
Network Architecture	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Functions	Sigmoid Function
Propagation Learning	
Learning Rate	0.7
Momentum	0.9
Criteria	
Error	0.021763

Appendix D

Variable	INFLATION
Included Observation	82 after adjusting endpoint)
Network Architecture	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Functions	Sigmoid Function
Propagation Learning	
Learning Rate	0.6
Momentum	0.9
Criteria	
Error	0.043471

Appendix E

Variable	INFLATION
Included Observation	82 after adjusting endpoint)
Network Architecture	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Functions	Sigmoid Function
Propagation Learning	
Learning Rate	0.5
Momentum	0.9
Criteria	
Error	0.043471

Appendix F

Variable	INFLATION
Included Observation	82 after adjusting endpoint)
Network Architecture	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Functions	Sigmoid Function
Propagation Learning	
Learning Rate	0.4
Momentum	0.9
Criteria	
Error	0.04504

Appendix G

Variable	INFLATION
Included Observation	82 after adjusting endpoint)
Network Architecture	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Functions	Sigmoid Function
Propagation Learning	
Learning Rate	0.3
Momentum	0.9
Criteria	
Error	0.06221

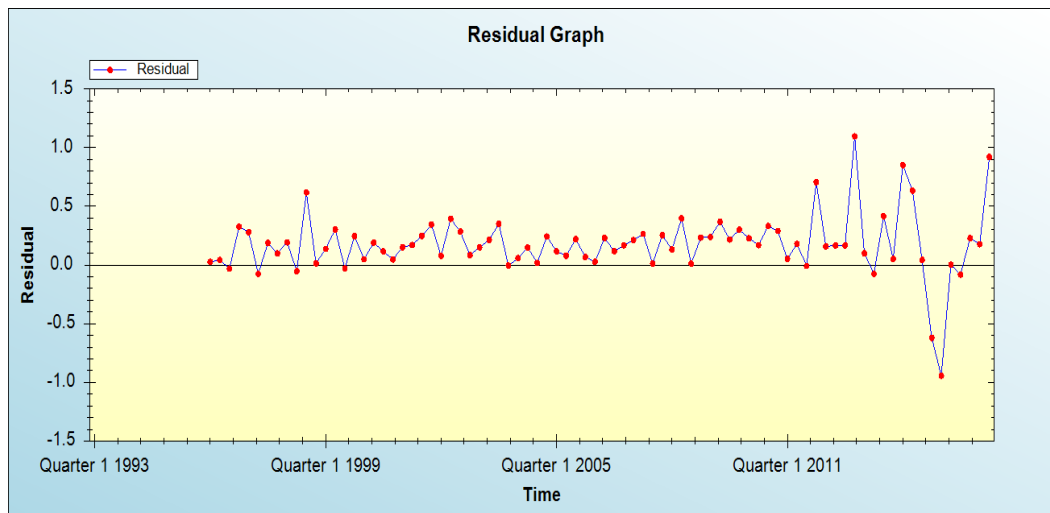
Appendix H

Variable	INFLATION
Included Observation	82 after adjusting endpoint)
Network Architecture	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Functions	Sigmoid Function
Propagation Learning	
Learning Rate	0.2
Momentum	0.9
Criteria	
Error	0.08267

Appendix I

Variable	INFLATION
Included Observation	82 after adjusting endpoint)
Network Architecture	
Input Layer Neurons	12
Hidden Layer Neurons	12
Output Layer Neurons	1
Activation Functions	Sigmoid Function
Propagation Learning	
Learning Rate	0.1
Momentum	0.9
Criteria	
Error	0.08849

Appendix J



Appendix K

