Analysing electricity consumption in South Africa using volatility forecasting models

OS MAKOBEA

orcid.org/0000-0003-1059-0051

Dissertation Submitted in fulfilment of the requirements for the Degree *Master of Commerce in Statistics with Business Statistics* at the North-West University

Supervisor: Prof T.V Montshiwa

Co-supervisor: Mr P.G Seaketso

Graduation: July 2023

Student number: 25389009

DECLARATION

I, the undersigned, Olorato Semantha Makobea, student number 25389009, do hereby declare that this dissertation is my original work as the results of my hard work and investigation except where quotations and references are stated. All the sources used have been acknowledged, and it has never been submitted by anyone at North-West University or any other institution in order to obtain an academic qualification.

………………………………………. **…………………………….**

O. S Makobea **Date**

ACKNOWLEDGEMENTS

Firstly, I would like to send special thanks to God Almighty for getting me through all the difficulties. It was indeed a long and rough journey that taught me resilience and to never quit. Thank you, God, for granting me the strength, wisdom and knowledge I needed the most to keep on pushing in order to make this work a success, I would also like to express my deepest gratitude to North-West University for making it possible by providing financial assistance when I needed it the most in order to further my studies.

My special appreciation and thanks to Professor L.D Metsileng and Professor J.T Tsoku for the help, support and encouragement throughout this journey; you have been tremendous mentors indeed. Words cannot express how grateful I am.

I would also like to express my greatest gratitude for the moral support from my beloved family especially my mother and father (Keaoleboga and Kagiso Makobea). Your prayers for me were what sustained me thus far. To my lovely and supportive friends, colleagues and everyone who contributed to make a completion of this dissertation possible; thank you from the bottom of my heart.

LIST OF ACRONYMS AND ABBREVIATIONS

Autocorrelation Function	ACF
Augmented Dickey-Fuller	ADF
Agriculture, Food and Fisheries Authority - Sugar Directorate	AFFA-SD
Asymmetric Generalised Autoregressive Conditional Heteroskedastic errors	AGARCH
Akaike Information Criteria	AIC
Artificial Neural Network	ANN
Autoregressive	AR
Autoregressive Conditional Heteroscedasticity	ARCH
Autoregressive Integrated Moving Average	ARIMA
Autoregressive Moving Average	ARMA
Breusch-Godfrey	BG
Bayesian Information Criterion	BIC
Box-Lung	BL
Back Propagation	BP
Consumer Price Index	CPI
Dickey-Fuller	DF
Exponential Generalised Autoregressive Conditional Heteroskedastic errors	EGARCH
Generalised Autoregressive Conditional Heteroskedastic errors	GARCH
Data Generating Process	GDP
Gross Domestic Product	GDP
Generalised Error Distribution	GED
Generalised Least Squares	GLS
Group Method of Data Handling Technique	GMDH
Glosten, Jagannathan and Runkle Generalised Autoregressive Conditional Heteroskedastic errors	GRJ-GARCH

Hannan -Quinn information criterion	HQIC
Integrated Generalised Autoregressive Conditional Heteroskedastic errors	IGARCH
International Telecommunication Union	ITU
Jarque-Bera (JB)	JB
Kwiatkowski–Phillips–Schmidt–Shin	KPSS
Lagrange Multiplier	LM
Moving average	MA
Mean Absolute Error	MAE
Mean Absolute Percentage Error	MAPE
Multivariate Minimum-Hellinger-Distance-Based Estimator	MHDE
Mean relative error	MRE
Mean Squared Deviation	MSD
Mean Squared Error	MSE
Partial Autocorrelation Function	PACF
Phillips-Perron	PP
Quadratic Generalised Autoregressive Conditional Heteroskedastic errors	QGARCH
Relative Absolute Error	RAE
Root Mean Square Error	RSME
Seasonal Autoregressive Integrated Moving Average	SARIMA
Back Propagation Seasonal Autoregressive Integrated Moving Average	SARIMA-BP
Schwarz Information Criterion	SIC
Theil Coefficient	TIC
Vector Error Correction Model	VECM
Zimbabwe and Wildlife Management	ZPWMA

ABSTRACT

The study explored the three-phase approach of SARIMA following the Box-Jenkins methodology, GARCH, and hybrid SARIMA-GARCH models. These volatility forecasting models were used to model electricity consumption in South Africa. The SARIMA (1, 1, 2)(0, 1, $1)_{12}$ model was found to be adequate to model South African electricity consumption. However, the series exhibits the presence of the ARCH effect, which led to modelling a GARCH model. The GARCH (1, 1) was modelled and confirmed to be a good fit. Furthermore, the study fitted the residual of the SARIMA model into the GARCH model to make it a hybrid SARIMA (1, 1, 2)(0, 1, $1)_{12}$-GARCH (1,1) model. It was observed that the hybrid model is the best fit model, with the smallest AIC value, and the diagnostic checking also confirmed that the model is adequate for forecasting electricity consumption. The results of the hybrid SARIMA (1, 1, 2)(0, 1, $1)_{12}$-GARCH (1, 1) model showed that the electricity consumption series has the highest volatility persistence value, and unconditional volatility for the series is finite.

Key words: electricity consumption, GARCH ,Hybrid SARIMA-GARCH,SARIMA volatility.

TABLE OF CONTENTS

DECLARATION i

ACKNOWLEDGEMENTS ii

LIST OF ACRONYMS AND ABBREVIATIONS iii

ABSTRACT v

LIST OF FIGURES x

LIST OF TABLES xi

CHAPTER 1 1

BACKGROUND AND MOTIVATION 1

1.1 INTRODUCTION 1

1.2 BACKGROUND OF THE STUDY 1

1.3 PROBLEM STATEMENT 3

1.4 AIM OF THE STUDY 4

1.5 RESEARCH OBJECTIVES 4

1.6 RESEARCH QUESTIONS 5

1.7 LIMITATIONS OF THE STUDY 5

1.8 SIGNIFICANCE OF THE STUDY 5

1.9 SCOPE OF THE STUDY 6

1.10 DEFINITION OF TERMS 6

1.11 ORIENTATION OF THE STUDY 7

1.12 CHAPTER SUMMARY 7

CHAPTER 2 8

LITERATURE REVIEW 8

2.1 INTRODUCTION 8

2.2 THEORETICAL LITERATURE REVIEW 8

2.2.1 Types of stationarity 9

2.2.1.1 Strong stationarity 9

2.2.1.2 Weak stationarity 10

2.2.2 Data transformation 11

2.2.2.1 First-order stationarity 11

2.2.2.2 N-th order stationarity 11

2.2.3 Unit root test 12
2.2.4 Stochastic modelling process 12
2.2.4.1 Autoregressive (AR) model 13
2.2.4.2 Moving average (MA) model 13
2.2.4.3 Autoregressive moving average (ARMA) model 13
2.2.5 Differenced stationary process 14
2.2.6 Model of symmetric volatility 14
2.2.6.1 ARCH models 15
2.2.6.2 GARCH model 15
2.2.6.3 GARCH-M model 16
2.2.7 Models of asymmetric volatility 16
2.2.7.1 EGARCH model 17
2.2.7.2 IGARCH model 18
2.2.7.3 GJR-GARCH model 18
2.2.8 The Hybrid SARIMA-GARCH model 19
2.3 EMPIRICAL LITERATURE REVIEW 19
2.3.1 Application of ARIMA and SARIMA models 19
2.3.2 Application of ARCH and GARCH models 23
2.3.3 Application of hybrid models 24
2.3.4 Review of previous papers in modelling electricity 28
2.4 CHAPTER SUMMARY 29
CHAPTER 3 31
RESEARCH METHODOLODY 31
3.1 INTRODUCTION 31
3.2 DATA DESCRIPTION 32
3.3 DESCRIPTIVE STATISTICS 32
3.3.1 Central Tendency or Location 32
3.3.2 Variability, dispersion or spread 34
3.3.3 Normality test 36
3.4 STATISTICAL METHODS 37
3.4.1 SARIMA model 37
3.4.1.1 Data preparation 39

3.4.1.2 Step 1: Model identification 42
3.4.1.2.1 Seasonality and seasonal ARIMA models...44
3.4.1.3 Step 2: Model estimation 45
3.4.1.4 Step 3: Model selection...46
3.4.1.5 Step 4: Diagnostic checking 47
3.4.1.6 Step 5: Forecasting...48
3.4.2 Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models ..49
3.4.2.1 Model comparison and selection...51
3.4.2.2 Diagnostics checking 52
3.4.2.3 Serial correlation test 52
3.4.2.4 Breusch-Godfrey LM test...52
3.4.2.5 Ljung-Box test...53
3.4.2.6 Heteroscedasticity test 53
3.4.2.7 Lagrange Multiplier (LM) test..54
3.4.2.8 Normality test 54
3.4.2.9 Forecasting accuracy 55
3.4.3 Hybrid of SARIMA-GARCH model 56
3.5 CHAPTER SUMMARY 57
CHAPTER 4 59
DATA ANALYSIS AND INTERPRETATION OF RESULTS..59
4.1 INTRODUCTION..59
4.2 PRELIMINARY DATA ANALYSIS..59
4.3 THE RESULTS OF THE BOX-JENKINS TEST 60
4.3.1 Model identification...60
4.3.2 Unit root test...61
4.3.3 *Autocorrelation plots*...62
4.3.3.1 The ACF and PACF plot after first differencing and seasonal differencing...........64
4.3.4 Estimation and selection..65
4.3.4.1 Parameter estimation of SARIMA (1, 1, 2) ×(0, 1, 1)$_{\mathbf{12}}$ 65
4.3.5 Diagnostic test of the SARIMA (1, 1, 2) (0, 1, 1)12 model 66
4.3.6 Forecasts of the SARIMA (1, 1, 2) (0, 1, 1)$_{\mathbf{12}}$ *model* 67

4.4 AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTIC (ARCH) MODEL 68
4.5 GENERALISED ARCH (GARCH) MODEL 70
4.5.1 Parameter estimates of the GARCH (1, 1) model 70
4.5.2 *Diagnostic test of the GARCH (1, 1) model* 72
4.5.3 Forecasts of the GARCH (1, 1) model...74
4.6 HYBRID SARIMA (1, 1, 2) (0, 1, 1)$_{\mathbf{12}}$- GARCH (1, 1) MODEL 75
4.6.1 Parameter estimates of the SARIMA (1, 1, 2) (0, 1, 1)$_{\mathbf{12}}$*- GARCH (1, 1) model*.77
4.6.2 Diagnostic tests of the SARIMA (1, 1, 2) (0, 1, 1)$_{\mathbf{12}}-\boldsymbol{GARCH}\ (\mathbf{1},\mathbf{1})$ *model* ...79
4.6.3 Forecast of the SARIMA (1, 1, 2) (0, 1, 1)$_{\mathbf{12}}-\boldsymbol{GARCH}\ (\mathbf{1},\mathbf{1})$ *model* 80
4.7 CHAPTER SUMMARY 81
CHAPTER 5 82
DISCUSSION, CONCLUSION AND RECOMMENDATIONS..82
5.1 INTRODUCTION...82
5.2 KEY FINDINGS OF THE STUDY 82
5.3 DISCUSSION 84
5.4 SUGGESTIONS FOR FUTHER STUDY 85
5.5 SUMMARY OF THE DISSERTATION...86
REFERENCES 87
APPENDIX...97

LIST OF FIGURES

Figure 3.1: The steps of the Box Jenkins methodology by Dritsakis and Klazoglou (2018) ..38
Figure 4.1: Original time series plot of electricity consumption in South Africa....................60
Figure 4.2: Plot of the first differenced data for electricity consumption in South Arica 61
Figure 4.3: ACF and PACF of first differenced data..63
Figure 4.4: ACF and PACF plot after first differencing and seasonal differencing................64
Figure 4.5: The graphical approach for checking for the randomness of residuals.................67
Figure 4.6: Series plot of original and forecasted electricity consumption values for the next five years..68
Figure 4.7: Conditional volatility of GARCH (1, 1)...72
Figure 4.8: The volatility forecast plot obtained from GARCH (1, 1) model 75
Figure 4.9: Autocorrelation check of residuals of the SARIMA (1, 1, 2) (0, 1, 1)12- GARCH (1, 1) model...76
Figure 4.10: Conditional volatility of SARIMA (1, 1, 2) (0, 1, 1)12 − GARCH (1, 1) model 78
Figure 4.11: The volatility forecast plot obtained from Hybrid SARIMA (1, 1, 2) (0, 1, 1)12 − GARCH (1, 1) model 81

LIST OF TABLES

Table 3.1: ACF and PACF behaviour (Box & Jenkins, 1994) 44
Table 3.2: Accuracy levels for MAPE test 49
Table 4.1: Descriptive statistics of electricity consumption in South Africa 59
Table 4.2: Unit root test for Electricity Consumption in South Africa...................................62
Table 4.3: Comparison of the models 65
Table 4.4: SARIMA (1, 1, 2) $\times$(0, 1, 1)12 parameters estimations 66
Table 4.5: ARCH Model parameter estimates...68
Table 4.6: ARCH heteroscedasticity test for residuals 69
Table 4.7: ARCH effect test for residual 69
Table 4.8: Information criteria of the fitted conditional distributions of the GARCH (1, 1) ..70
Table 4.9: Estimated model parameter of the GARCH (1, 1) 71
Table 4.10: Heteroscedasticity test of residuals for GARCH (1, 1) model 72
Table 4.11: ARCH effect of the GARCH (1, 1) model 73
Table 4.12: Goodness-of -fit test of the GARCH (1, 1) model..73
Table 4.13: Forecasted values of the GARCH (1, 1) model 74
Table 4.14: Information criteria of the fitted conditional distributions of the SARIMA (1, 1, 2) (0, 1, 1)12 − GARCH (1, 1) model 76
Table 4.15: Estimated parameters for SARIMA (1, 1, 2) (0, 1, 1)12- GARCH (1, 1) model 77
Table 4.16: Heteroscedasticity test of residuals for SARIMA (1, 1, 2) (0, 1, 1)12 − GARCH (1, 1) model 79
Table 4.17: ARCH effect of the hybrid SARIMA (1, 1, 2) (0, 1, 1)12 − GARCH 1, 1 model 79
Table 4.18: Goodness-of-fit test for the SARIMA (1, 1, 2) (0, 1, 1)12 − GARCH (1, 1)model 80

CHAPTER 1

BACKGROUND AND MOTIVATION

1.1 INTRODUCTION

This study explores some volatility forecasting models that can be used for predicting South African electricity consumption. The Seasonal Autoregressive Integrated Moving Average (SARIMA) model, Generalised Autoregressive Conditional Heteroskedastic errors (GARCH) model and hybrid of SARIMA-GARCH model following Box-Jenkins approach will be developed and used.

Electricity is a flexible source of energy, therefore estimating electricity consumption is crucial for decision-making processes because it is an essential component of infrastructure in both the energy industry and for socio-economic development (Fisher, 2008). Decision making usually involves planning under assumptions and predictions, which involves strategic planning for supply capacity and development (Conejo *et al*., 2010).

It is therefore of great importance to produce very accurate and reliable estimations because the effects of either overestimation or underestimation can be quite expensive. Johannesen *et al*.(2019) noted that both electricity suppliers and consumers need reliable short-term forecasts, especially during periods when demand peaks abnormally. Accurate forecasts of electricity consumption are expected to help in determining a consistent and reliable supply of energy during abnormal times. Short-term forecasts are also helpful in economic planning.

1.2 BACKGROUND OF THE STUDY

The SARIMA model is developed from the Autoregressive Integrated Moving Average (ARIMA) models. American statisticians George Edward Pelham Box and Jenkins in 1976 (Ljung *et al*.,2014) developed these models. When data show evidence of non-stationarity, the most widespread systematic class of models invented by Box and Jenkins for modelling and predicting time series are ARIMA models (Kipiński *et al*.,2011).

Transformations such as differencing can be used to eliminate non-stationarity in the mean; however, a suitable variance stabilising transformation is able to reduce non-stationarity in the variance (Stockhammar & Öller, 2012). The orders of transformation are identified using autocorrelation function (ACF) and partial autocorrelation function (PACF) (Dürre *et al*.,2015). Usually, the normal distribution with constant variance and mean of zero is assumed to be the source of the error terms, which are independent random variables with identical distributions.

The observed data in this study is assumed to follow a multiplicative model (Pandey *et al*., 2019). The autoregressive model of order P, AR (p), the moving average model of order q, MA (q), the autoregressive moving average model, ARMA (p,d), the ARIMA model and the SARIMA model are some of these models. The SARIMA $(p,d,q) \times (P,D,Q)$ model is employed when the data being analysed is seasonal. The ARIMA (p,d,q) model is categorised as a non-seasonal ARIMA model.

When analysing time series data, statisticians often employ the SARIMA model, a multiplicative model that is widely used when there is evidence of seasonality in the time series data. Other statisticians and researchers have explored the SARIMA model, including Helman (2011), Ampaw *et al*. (2013), Eni and Adesola (2013), Etuk and Igbudu (2013), Etuk and Ojekudo (2014), and Etuk *et al*. (2014), and they have done extensive work. This model is recommended due to its high level of accuracy and reliability (Fang & Lahdelma, 2016).

A series with time-varying conditional variance is commonly defined using autoregressive conditional heteroscedasticity (ARCH) family models, which were developed by Robert F. Engle in 1982 and generalised by Bollerslev in 1986 (Pepple & Harrison, 2017). The generalised autoregressive conditional heteroscedasticity (GARCH) class models are employed where volatility is a crucial issue. These models were developed to identify volatility clustering and to forecast future volatilities. Standard deviation measures volatility, which is the amount of variance in data over time (Karmakar, 2005). The current study aims to fit a hybrid SARIMA-GARCH model in which the SARIMA model's error term variance follows a GARCH process. To demonstrate the heteroscedastic nature of residuals, the GARCH model is employed.

The current study is intended to extend the scope of previous studies by examining the efficiency of SARIMA, GARCH, and hybrid of SARIMA-GARCH models using electricity consumption data. These models were chosen as they are proven to produce high degree of accuracy when modelling and forecasting other datasets, and are also very helpful where volatility is a crucial issue. Engle and Patton (2007) noted that an effective volatility model should be able to identify and reflect stylised facts, and the authors further explain that GARCH models have shown the ability to capture these features.

1.3 PROBLEM STATEMENT

In South Africa, since 2007 there is a significant demand for power, as indicated by Marwala and Twala (2014). Although there is a scarcity of electrical supply, there is still a strong demand for electricity that has led to load shedding as one of strategies under implementation by the South African energy supplier to address a severe and growing problem of an electricity shortfall in the country (Zhang *et al*., 2008). The crisis of electricity scarcity is caused by the inadequate predicting models employed by Eskom South Africa, the company in charge of supplying the majority of electricity in the country (Butgereit, 2015).

Inadequate forecasting can lead to difficulty when modelling, which produces forecasts that are inaccurate and unreliable. Inglesi and Pouris (2010) found that it is crucial to estimate electricity consumption in order to prevent ineffective planning that will result in a shortfall of power supply. Since there is inefficient distribution of electricity, the demand and consumption in people's daily lives are growing more quickly. The vast growth of poor distribution has an impact on practically everyone, because many homes and industries, and crucial service providers such as education, health care and other crucial utilize electricity. For instance, on average 28% of health institutions rely on energy imports from 11 sub-Saharan African countries (Adair-Rohani *et al*., 2013). Therefore, the growth of South Africa's economy could be negatively impacted by an unreliable and inadequate energy supply.

Traditional time series models like ARIMA and SARIMA models can accurately capture the stochastic process of predicting electricity consumed monthly. But the major challenge when predicting using only traditional time series models is that residuals of the estimated model frequently have heteroscedasticity, despite the assumption that variance of residuals is constant. According to (Zhou *et al*. 2006), traditional models are unable to retain some of the important aspects of the data, and the residual may be autoregressive for some parameters are not considered. However, the GARCH model can effectively eliminate this autoregressive issue. ARCH and the generalised family of models can be employed to estimate and predict future values when there is volatility, and since it was discovered that the GARCH models outperform most of the competing models, numerous improvements to the methodology have been made to adequately capture the unique properties of the data (Awartani & Corradi, 2005). Most researchers often recommend GARCH (1, 1) to be the best model for estimating conditional volatility (Gökbulut & Pekkaya, 2014).

The ARCH family models are proven to be effective in estimating and predicting the volatility of the time series data, although these models do not address the 'leverage effect' that causes the conditional variance to frequently react asymmetrically to both positive and negative shocks in errors. The hybrid of SARIMA-GARCH model will be developed and fitted to address this issue. The current study intends to firstly fit the SARIMA model based on the singularity and seasonality of the electricity consumption data, followed by the GARCH model to alter the fitting error, and lastly fit the hybrid of the SARIMA-GARCH model to check if the prediction accuracy is improved. As such, there is a need to conduct more studies that can determine the efficiency of some novel forecasting models in an attempt to produce accurate forecasts of electricity consumption to resolve the electricity supply crisis in South Africa.

1.4 AIM OF THE STUDY

The study aims to model the electricity consumption in South Africa using some novel volatility forecasting models. The volatility forecasting models tested in this study are SARIMA model, GARCH model and SARIMA-GARCH model to determine the most efficient model.

1.5 RESEARCH OBJECTIVES

The objectives of the study are to:

i. Determine the seasonal effect on electricity consumption.
ii. Model South African electricity consumption using the SARIMA, GARCH and SARIMA-GARCH models.
iii. Use the identified SARIMA, GARCH and SARIMA-GARCH models to forecast South African electricity consumption
iv. Provide recommendations based on the findings of study.

1.6 RESEARCH QUESTIONS

Based on the above-mentioned research objectives, the following research questions were developed:

i. Is there a seasonal effect on electricity consumption?

ii. Which model best fits South African electricity consumption data; SARIMA, GARCH or SARIMA-GARCH model?

iii. Which model best predicts South African electricity consumption?

iv. How the predicted values of electricity consumption in South Africa performs using the forecast accuracy measure?

v. What are recommendations based on the findings?

1.7 LIMITATIONS OF THE STUDY

The study made use of secondary data instead of conducting interviews or surveys with questionnaires. The study uses 244 observations of monthly data on the electricity consumption for the years 2000 to 2020. The study aimed to predict the consumption of electricity in South Africa, therefore the findings of the study cannot be extrapolated to other countries. The study relies on the models originally developed by older studies (older than 10 years) and will cite some of the old literature.

1.8 SIGNIFICANCE OF THE STUDY

The uneven distribution of electricity has negatively impacted many people and different economic sectors, and this has affected the economic growth of South Africa (Okafor, 2008). The current study is expected to propose the most efficient forecasting model(s) that can yield accurate and trustworthy predictions of electricity consumption in South Africa in order to assist future energy planners to generate and supply the correct amounts of electricity needed. The study is also important since it will serve as a reference point for future researchers interested in determining the seasonal influences on South African electricity consumption, and other time series variables. The findings of the study will assist future researchers with knowledge regarding the most efficient model between GARCH, SARIMA, and SARIMA-GARCH model.

1.9 SCOPE OF THE STUDY

The study's main objective is modelling South African electricity consumption using SARIMA, GARCH, and SARIMA-GARCH model. The Box-Jenkins approach will be used in the SARIMA model. The developed models will then be used for forecasting of electricity consumption. The study uses a monthly secondary data set starting from the 3rd of January 2000 until the 3rd of April 2020. The data set consists of 244 observations. The data was obtained from Quantec Easy Data and is available through https://www.easydata.co.za/dataset/P4141/timeseries/P4141-ELEKTR22/. Timely and accurate updates are released by Quantec Easy Data together with its various primary data providers. Quantec is a consultancy entity providing economic and financial data which enables researchers to find and explore relevant information and convey results rapidly. The data analysis was done using SPSS, R and E-views.

1.10 DEFINITION OF TERMS

- ARIMA – Autoregressive Integrated Moving Average models are a general systematic class of models invented by Box-Jenkins for forecasting a time series, and this modelling is used in situations where data exhibits evidence of non-stationarity (Pepple & Harrison, 2017).
- SARIMA – The Seasonal Autoregressive Integrated Moving Average model, developed by American Statisticians G.E.P Box and Jenkins in 1976, is effective in cases where the time series data show signs of seasonality (i.e. timely occurrence with roughly the same intensity (Pepple & Harrison, 2017).
- ARCH – The Autoregressive Conditional Heteroscedasticity model was developed by Nobel laureate Robert Engle in 1982 (Engle, 2004). A common family of econometric models known as ARCH models are used for describing a series with time-varying conditional variance (Gouriéroux, 2012).
- GARCH – The Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model is used when volatility is a crucial issue; compared to the basic models, the model's forecasting performance is superior because it captures long-memory volatility (Hou & Suardi, 2012).
- Volatility – is the amount by how much data fluctuates over time, as measured by the standard deviation (Taylor, 2011).

1.11 ORIENTATION OF THE STUDY

The dissertation contains five chapters that address the above-mentioned aims and objectives of the study, and these chapters are detailed below.

Chapter One: Background and Motivation

This chapter discusses the background of the study as well as the motivation. The chapter further discusses the problem statement, research objectives and the research questions.

Chapter Two: Literature Review

This chapter reviews previous theories. It discusses the relevant literature on forecasting methods with more focus on the use of SARIMA, GARCH, and SARIMA-GARCH models as well as the Box-Jenkins approach. This chapter also discusses the literature on electricity consumption in South Africa.

Chapter Three: Methodology

This chapter outlines the research design and techniques. It describes in detail the parameterisation of the different forecasting techniques used to analyse the data; it also describes the data, its source, and the statistical software that was used

Chapter Four: Data Analysis

The statistical analysis findings as well as their interpretation are presented in this chapter.

Chapter Five: Conclusions and Recommendations

The conclusions are presented in this chapter and recommendations for further studies.

1.12 CHAPTER SUMMARY

The chapter provide an overview of the study outlining the background to the study, the problem statement, and the aims and objectives. Furthermore, the chapter emphasised the importance and the scope of the study in detail.

The next chapter reviews the relevant literature for the current study and provides the background of each model examined.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

This chapter reviews relevant studies relating to modelling and forecasting future values using forecasting techniques. The chapter discusses the application of different techniques and models used in different studies for benchmarking purposes. There has been a lot of interest in volatility modelling and forecasting recently, largely due to their significance in the financial markets. Volatility estimates are commonly used in asset-pricing models as simple risk measures; also, option-pricing formulas are developed from such volatility models. It is essential to have accurate and reliable volatility estimations for portfolio management and risk hedging. There are many models that are available to take into account certain conventional facts such as autoregressive moving average.

The rest of the chapter is structured as follows: in Section 2.2, theoretical literature is reviewed which includes subsections: 2.2.1 Stationarity test, 2.2.2 Data transformation, 2.2.3 Unit root test, 2.2.4 Stochastic modelling process, 2.2.5 Differenced stationary process, 2.2.6 Model of symmetric volatility, 2.2.7 Models of asymmetric volatility and 2.2.8 Hybrid of SARIMA-GARCH model. 2.3 reviews empirical literature, which includes sub-sections: 2.3.1 Application of ARIMA and SARIMA models, 2.3.2 Application of ARCH and GARCH models, 2.3.3 Application of hybrid models, 2.3.4 Review of previous papers in modelling electricity and lastly 2.4 presents the summary of the chapter.

2.2 THEORETICAL LITERATURE REVIEW

In the current study, time series approaches are employed in modelling data for electricity consumption in South Africa. The approach follows the Box-Jenkins methodology, which includes identifying the pattern using visual inspection (time series plots), unit root testing and model order identification using AR and MA and fitting the suitable ARIMA/SARIMA models. According to Pandey, Tripura and Pandey (2019), the SARIMA model is used to stabilise conditional variance and to determine the mean equation. If ARCH effects are present, the GARCH-type models can be utilised to eliminate the effect of heteroscedasticity in the data (Pandey *et al.*, 2019). Studies by Yaziz *et al*. (2013), Tendai and Chikobvu (2017), Dritsakis and Klazoglou (2018), to mention a few, revealed that the hybrid model SARIMA-GARCH model produces better forecasting results because of its ability to assess volatility, and this is one of the reasons why we are interested in this model in the current study.

Determining stationarity of the time series is the first and most important step when modelling data using the Box-Jenkins approach. Stationarity can be verified using both visual inspection and formal unit root tests. Stationarity test is discussed in the following subsection 2.2.1.

2.2.1 Types of stationarity

Adhikari and Agrawal (2013) are of the view that a time series $y_t(t = 1,2 \dots n)$ is said to be stationary if it generates time series' statistical properties that do not change over time (variance, autocorrelation). Stationarity is very important because a stationary process is easier to analyse, hence in time series analyses, testing for stationary has become a standard assumption for many studies (Adhikari & Agrawal, 2013). A time series whose characteristics are unaffected by the observation time is referred to as stationary. Time series data containing patterns or with seasonality are therefore not stationary since this will have an impact on the value of the time series at various points in time (Zhang & Zhao, 2017).

A stochastic process has any of kind of stationarity as a property. A family of real random variables $X = \{x_i(\omega); i\epsilon T\}$ is a real stochastic process; they are all described within the identical space of probability(Ω, F, P) (Adhikari & Agrawal, 2013). The set T is referred to as the index set of the process; if T is an interval of $\mathbb{R}$ the process is referred to as a continuous stochastic process, although if $T \subset Z,$ then it is referred to as a discrete stochastic process (Lindgren *et al*.,2013.) Strong stationarity and weak stationarity are two commonly known and used types of stationarity. The two types are discussed in the following sub-sections.

2.2.1.1 Strong stationarity

Finite-dimensional distributions must be shift-invariant (in time) for a stochastic process to have strong stationarity (Adhikari & Agrawal, 2013). This implies that in the stochastic process the finite sub-sequence of random variables remains constant when shifted along the time index axis.

Statistically, if the discrete stochastic process $X = \{x_i; i\epsilon Z\}$ is stationary, it is given by:

$$F_X\left(x_{t_{1+\tau}}, \dots, x_{t_{n+\tau}}\right)= F_X\left(x_{t_1}, \dots, x_{t_n}\right), \quad (2.1)$$

for $T \subset Z$ with $n \in N$ and any $\tau \in N$. Similar conditions apply for continuous stochastic processes with $T \subset \mathbb{R}$, $n \in N$ and any $\tau \in \mathbb{R}$ instead (Cox & Miller, 1965). Stationarity is commonly defined as a strong-sense or strict-sense stationarity (Adhikari & Agrawal, 2013).

2.2.1.2 Weak stationarity

For weak stationarity, only the cross moment (the auto-covariance) and shift-invariance (in time) of the first moment is required (Shumway *et al.*, 2000). This means that the average of the process remains constant throughout time points, and the covariance between t values at any two time points, t and $t - k$, depend solely on the change in time, k, and not where the points are located on the time axis (Boshnakov, 2011).

According to Boshnakov (2011), if the process$\{x_i; i\epsilon Z\}$ is weakly stationary, then:

- The first moment of x_i is constant, i.e., $\forall t, E[x_i] = \mu$
- The second moment of x_i is finite for all t; i.e., $\forall t, E[{x_i}^2] < \infty$ (which also implies that variance is finite for all t i. e.$E[(x_i - \mu)^2] < \infty$;)
- The cross moment or the auto-covariance is only dependent on the difference between u and v, hence $\forall u, v, a, cov(X_u, X_v) = cov(X_{u+a}, X_{v+a})$

According to the third condition, each lag $\tau \in \mathbb{N}$ is attached to a constant covariance value and it is given by:

$$cov\left(X_{t_1}, X_{t_2}\right)= K_{XX}(t_1, t_2)=K_{XX}(t_2 - t_1, 0)=K_{XX}(\tau). \quad (2.2)$$

Note that since it is found for any $\in \mathbb{N}$, this certainly indicates that the variance of the process remains constant as well.

$$var(X_t) = cov(X_t, X_t) = K_{XX}(t, t)=K_{XX}(0)=d. \quad (2.3)$$

Weak stationarity is also known as second order stationarity, weak-sense stationarity, wide-sense stationarity and covariance stationarity (Shumway *et al.*, 2000). Depending on context, it may also be defined simply as stationarity (Boshnakov, 2011).

Identifying a series that is non-stationary is crucial. A non-stationary series can be made stationary through transformations that include differencing or logarithms. In terms of a random walk, a series can be stationary in trend (Shumway *et al.*, 2000). In order to have accurate results and predictions, the time series data should be free from the effect of trends and seasonality, and if it does not occur naturally in the dataset, this can be achieved from data transformation that is discussed next.

2.2.2 Data transformation

Data transformation is widely used in time series analysis to make a series stationary. Transformations may include differencing, which can assist to stabilise the mean, reduce or eliminate trend and seasonality in time series data. (Lütkepohl & Xu, 2012). Transformation may also include logarithms, which may enable stabilising the variance of a time series, (Lütkepohl & Xu, 2012). The differenced data may not always appear to be stationary at the first attempt; it may require data to be differenced several times before producing a stationary series. Therefore, the next subsection outlines the order of stationarity.

2.2.2.1 First-order stationarity

The phrase "first-order stationarity" is commonly employed when a series mean does not change over time, but variance is not constant. According to the definition by Boshnakov (2011), this is not similar to *N*-th order stationarity for $N = 1$, as the latter demands uniform distribution of all x_i for a process $\boldsymbol{X} = \{x_i;\ i\epsilon Z\}$. For instance, the mean does not vary with time for a process with $x_i \sim \boldsymbol{\mathcal{N}}(\mu, f(i) = 1)$, but x_i is not uniformly distributed, with the $f(i) = 1$ for even values of i, and $f(i) = 2$ for odd values. This implies that the concept of first-order stationarity is relevant to a particular process, but not applicable to the *N*-th order stationarity for $N = 1$.

2.2.2.2 N-th order stationarity

The *N*-th order stationarity and strong stationarity have identical properties. For all *n* up to order N, the distribution of any *n* samples of the stochastic process is required to shift-invariance (in time) in order for the series to be of the *N*-th order stationarity (Adhikari, 2013). Hence, equation (2.1) required the same condition.

Naturally, even while the opposite is true, stationarity of any higher order is not guaranteed by any specific order N. Both a first order stationary process that is not a second order stationary process and that is not a third order stationary process are examples of interesting threads in mathematical overflow.

It should be noted that, despite the fact the latter is usually referred to as second order stationarity, stationarity of the N-th order for $N = 2$ is not equivalent to weak stationarity (Myers, 1989). Similar to strong stationary, second order stationarity does not entail that X has finite moments for the distribution of any two samples of X. A second order stationary process must satisfy the necessary requirement that the second moment be finite in order for it to be weakly stationary.

2.2.3 Unit root test

Formal stationarity tests allow the researcher to verify the visual examination of whether the series is stationary or not. Alternatively, a unit root test can be used for accurately determining whether differencing is necessary or not. There are two different approaches. On the first approach, there are statistical hypothesis tests for stationarity that determine if the null hypothesis that a series is stationary is true. Those tests include the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test, which was proposed by Kwiatkowski *et al.* (1992), and Leybourne and McCabe (1994). On the second approach, the null hypothesis is contrary to that of the first approach, and it is that the series is not stationary. The tests in the second approach include the Dickey-Fuller (DF) test, Augmented Dickey-Fuller (ADF) test developed by Dickey and Fuller (1979), the Phillips-Perron (PP) test developed by Phillips and Perron (1988), and the DF-GLS test proposed by Otero and Baum (2017) as a modification of the Dickey-Fuller test statistic based on the generalised least squares (GLS) approach. The second approach is the most widely employed for determining if a time series has the presence of unit root.

2.2.4 Stochastic modelling process

The common task when examining time series data is forecasting future values. When producing forecasts, it is necessary to establish a few assumptions about the Data Generating Process (DGP); the method used to extract the data (Adhikari & Agrawal, 2013). These assumptions are frequently expressed in an explicit process model and are also employed when modelling stochastic processes for other purposes like causal inference or anomaly detection. The three most common models are discussed in the following subsections.

2.2.4.1 Autoregressive (AR) model

An AR model assumes that a time series model is developed as a linear function using its historical data and random noise or error term, and it is parameterised as follows (Shongwe *et al*.,2019):

$$x_t = c + \vartheta_1 x_{t-1} + \ldots + \vartheta_p x_{t-p} + \varepsilon_t. \qquad (2.4)$$

Since each value is correlated with the *p* preceding values, this indicates that is a memory-based model. An AR model with lag *p* is denoted with AR *(p*). The coefficients ϕ_i are weights that assess how much of an impact these past values have on the value $x[t]$, a univariate white noise process is denoted by ε_i and *c* is a constant intercept (commonly assumed to be Gaussian).

2.2.4.2 Moving average (MA) model

In time series model, of the most recent $q + 1$ random shocks established by ε_i, it is assumed that the univariate white noise process results in a linear function modelled using a moving average model, denoted with $MA(q)$ and it is parameterised as follows (Mahmud *et al*., 2017):

$$x_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_p \varepsilon_{t-q}, \qquad (2.5)$$

where μ is the mean of the series, the $\varepsilon_t, \varepsilon_{t-1}, \ldots, \varepsilon_{t-q}$ are white noise error terms and the $\theta_1, \ldots, \theta_p$ are parameters of the model. The order of the MA process value is given by the value of q. This can also be expressed equivalently using the backshift operator *B*, and be given as (Mahmud *et al*.,2017):

$$x_t = \mu + \left(1 + \theta_1 B + \cdots + \theta_q B^q\right)\varepsilon_t \, . \qquad (2.6)$$

2.2.4.3 Autoregressive moving average (ARMA) model

The assumption is that the time series model is established as a linear function of the most recent p values and the most recent $q + 1$ random shocks produced by ε_i, a univariate white noise process, when modelled using an ARMA(p, q) is defined by the following equation which is adapted from Bopulas (2011).

$$x_t = \vartheta_1 x_{t-1} + \ldots + \vartheta_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_p \varepsilon_{t-q}. \tag{2.7}$$

There are many different ways to generalize ARMA models, such as to handle exogenous variables, or non-linearity, or the multivariate situations (Vector ARMA) or dealing with a particular type of non-stationary data (ARIMA).

2.2.5 Differenced stationary process

Now that there is a fundamental knowledge of standard stochastic process models, the emphasis is on examining the related idea of differenced stationary processes. This concept is predicated on the notion that the stochastic process might be expressed as an AR process of order p, given as AR(p) (Fung *et al*., 2017):

$$y_t = a_0 + a_{1y_{t-1}} + \ldots + a_{py_{t-p}} + \varepsilon_t, \tag{2.8}$$

where ε_i are white-noise processes often uncorrelated (for all times t). The same process can be rewritten as (Fung *et al*., 2017):

$$\left(1 - a_1 L - \cdots - a_p L^p\right) y_t = a_0 + \varepsilon_t. \tag{2.9}$$

The portion on the left side of the parenthesis is the characteristic equation of the process. Considering the roots of equation (2.8) as (Fung *et al*., 2017):

$$m^p - m^{p-1} a_1 - \cdots - a_p = 0. \tag{2.10}$$

Differenced stationary or integrated process is referred to a stochastic process if $m = 1$ is a root of the equation. As a result, by using a transformation known as differencing, the process can be transformed into a weakly stationary process. The order of integration indicates the certain number of times the differencing operator needs to be applied to attain weak differenced stationarity (Adhikari & Agrawal, 2013). Integrated of order *r* refers to a procedure that needs to be differenced *r* times and it is denoted by $I(r)$. The process is integrated of order r, if the multiplicity r characteristic equation has a root of $\mathrm{m} = 1$. The multiplicity of the root $\mathrm{m} = 1$ is perfectly matched by this (Fung *et al*., 2017).

2.2.6 Model of symmetric volatility

The standard GARCH is known to be symmetric; it is often used as a benchmark to illustrate the asymmetry of other GARCH type models (Chen, Zhang, Tao & Tan, 2019). ARCH, GARCH and GARCH-in Mean (GARCH-M) models are more appropriate measurements for risk in the presence of frequent fluctuations (Khan *et al*., 2019). The symmetric models are discussed in the following subsection.

2.2.6.1 ARCH models

The ARCH was introduced in 1982 by Engle and it was the first model with conditional variance that converted ARCH to GARCH and allowed a forecast of conditional volatility. ARCH models are commonly employed in modelling time series data that show time varying volatility.

The general ARCH model is given by Chen *et al*. (2019):

$$\sigma_t{}^2 = a_0 + a_1\theta_{t-1}{}^2 + \ldots + a_q\theta_{t-q}{}^2. \quad (2.11)$$

The parameters of the model are required to be restricted to $a_0 > 0$ and $a_i \geq 0, i = 1, \ldots, q$ ensuring that the conditional variance does not violate the non-negativity constraint (Higgins & Bera, 1992).

The unconditional variance for ARCH is given by Dahlvid and Granberg (2017):

$$\sigma^2 = {}^{\alpha_0}\!/_{1 - \alpha_1 - \alpha_2 -, \ldots, \alpha_q}\,, \quad (2.12)$$

The conditional GARCH variance for ARCH is given by Dahlvid and Granberg (2017):

$$\sigma^2 = VAR(\mu_t|\Omega_{t-1}) = E\left[\mu_t - E\big((\mu_t)\big)^2 \middle| \Omega_{t-1}\right]. \quad (2.13)$$

2.2.6.2 GARCH model

The ARCH model was developed by Engle in 1982 to solve the issues of heteroscedasticity that are present in the time series. The ARCH model, which describes the volatility estimation approach utilising volatile series, was extended by Bollerslev (1986) into the GARCH (*p*, *q*) model. GARCH acts as one of the significant ways to analyse volatility, although GARCH models are more accurate for short-term volatility (Cifter, 2012). The GARCH (*p*, *q*) models enable the conditional variance to be dependent on its residual values (Dahlvid & Granberg, 2017).

The general GARCH $(1, 1)$ model is defined as (Dahlvid and Granberg, 2017):

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha e_{t-1}^2, \tag{2.14}$$

where σ_t^2 is the constant variance, $\omega > 0, \alpha \geq 0, \beta \geq 0$ and $\alpha + \beta < 1$. To model linear dependency on the series, an ARIMA (p, q) process with a GARCH white noise is given by Y_t and the ARCH effect on residuals is best described as (Cifter, 2012):

$$Y_t = \mu + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j}^2 + \sum_{i=1}^{p} \alpha_i Y_{t-i} + \varepsilon_t . \tag{2.15}$$

2.2.6.3 GARCH-M model

The GARCH-M model was invented by Engle, Lilien and Robins in 1987 to establish correlation between conditional variance and return, where the current return is directly related to the current variance through a linear model (Nugroho *et al*., 2019). The GARCH $(1, 1)$-M is expressed as (Nugroho *et al*., 2019):

$$y_t = \eta\sigma^2{}_t + \sigma_t\varepsilon_t,\ \varepsilon_t \sim N(0,1), \tag{2.16}$$

where

$$\sigma^2{}_t = \omega + \alpha \mathrm{K}^2{}_{t-1} + \beta\sigma^2{}_{t-1},$$

where the constraints for parameters are similar to the standard GARCH$(1, 1)$ and

$$\varepsilon_t = Z_t\sigma_t,\ Z_t \sim \text{ i.i.d } N\ (0, 1), \tag{2.17}$$

where

$$\sigma_t{}^2 = a_0 + a_1\varepsilon_{t-1}^2 + a_2\varepsilon_{t-2}^2 + \ldots + a_p\varepsilon_{t-p}^2 + \beta_1\sigma_{t-1}^2 + \beta_2\sigma_{t-2}^2 + \ldots + \beta_q\sigma_{t-q}^2\ , \tag{2.18}$$

where a_i, β_j and μ are parameters and ε_t are residuals of the model. There is a lower likelihood for the GARCH model to ignore non-negativity constraints, because the model uses few parameters (Dahlvid & Granberg, 2017).

2.2.7 Models of asymmetric volatility

Engle (1990) proposed the asymmetric GARCH model to capture asymmetric volatility response in the GARCH process (Cifter, 2012). The asymmetric volatility can be estimated using a variety of widely used models. The models such as Exponential GARCH (EGARCH), Integrated GARCH (IGARCH) and the Glosten, Jagannathan and Runkle GARCH (GRJ-GARCH). These models were all developed from the conventional linear GARCH model. The models are discussed in the following subsection.

2.2.7.1 EGARCH model

Nelson (1991) invented the EGARCH model, which was the first model to integrate the asymmetric volatility. According to Nelson and Cao (1992), the linear GARCH model has the non-negativity constraint, in contrast to the EGARCH model which has no restrictions on the parameters, which means the model is too restrictive. The previous studies showed that the EGARCH yields more accurate findings when compared to the traditional GARCH model (Alberg *et al*., 2008).

The conditional variance in the EGARCH model, $h_t,$ is an asymmetric function of lagged disturbances ε_{t-i} and according to Su (2010):

$$\ln(h_t) = \omega + \sum_{i=1}^{g} \alpha_1\, g(z_{t-i}) + \sum_{j=1}^{p} y_j \ln\left(h_{t-j}\right), \tag{2.19}$$

where

$$g(z_t) = \theta z_t + \gamma[|z_t| - E|Z_t|], \tag{2.20}$$

where:

$$z_t = \frac{\varepsilon_t}{\sqrt{h_t}} \quad , \tag{2.21}$$

The coefficient of the second term in $g(z_t)$ is set to be 1$(y = 1)$. Note $E|Z_t| = \left(2/\pi\right)^{1/2}$ if $Z_t \sim N(0, 1)$. The characteristics of the EGARCH model are outlined as follows (Alberg, Shalit & Yosef, 2008):

- If the function $g(z_t)$ is linear in z_t with slope coefficient $\theta + 1$ if z_t is positive, while $g(z_t)$ is linear in z_t with slope coefficient $\theta - 1$ if z_t is negative
- If $\theta = 0$, large innovations increase the conditional variance if $|Z_t| - E|Z_t| > 0$ and decrease the conditional variance if $|Z_t| - E|Z_t| < 0$.

- If $\theta < 1$. The innovation in variance, $g(z_t)$, is positive if the innovations z_t are less than $\left({(2/\pi)^{1/2}} / {\theta - 1} \right)$. Therefore, the negative innovations in returns ε_t cause the innovation to the conditional variance to be positive if θ is much less than 1.

2.2.7.2 IGARCH model

Engel and Bollerslev (1986) proposed the modified Integrated GARCH (IGARCH) model. The GARCH model requires the covariance to be stationary such that:

$$\sum_{i=1}^{h} \alpha_i + \sum_{j=1}^{p} \beta_j < 1 \ . \qquad (2.22)$$

The GARCH process is assumed to be weakly stationary since the mean, variance and auto - covariance do not vary over time. However, the stationarity does not require a condition restriction such as the unconditional variance not being dependent on time. Practically, the covariance is not stationary (Rastogi *et al*., 2018):

$$\sum_{i=1}^{h} \alpha_i + \sum_{j=1}^{p} \beta_j = 1. \qquad (2.23)$$

There is no unconditional variance in the IGARCH model. But even if it is weakly stationary the IGARCH model tends to exhibit strong stationarity, as a result of omitted structural breaks. IGARCH is given by: (Rastogi *et al*.,2018)

$$\sigma_t{}^2 = a_0 + (1 - \beta_1)\varepsilon^2{}_{t-1} + \beta_1 \sigma^2{}_{t-1} \ . \qquad (2.24)$$

where $\sigma_t{}^2$ is the conditional volatility and $\varepsilon^2{}_{t-1}$ are squared unexpected returns for previous period. a_0 would always be positive, $(1 - \beta_1)$ and β_1 would be non-negative.

2.2.7.3 GJR-GARCH model

The other model that incorporates asymmetric volatility and is widely employed was proposed by Glosten, Jagannathan and Runkle (1993). The model is commonly known as GJR-GARCH model. This model has the advantage of differentiating the negative and positive shocks (Cifter, 2012) in contrast to the EGARCH model, as variance is simply modelled and without using the natural logarithm. As a result, this makes GJR-GARCH easier to employ (Dahlvid & Granberg, 2017). Several studies have employed different GARCH models and concluded that the GJR-GARCH is the most efficient in predicting volatility.

The GJR-GARCH (1,1) model is given by (Asgharian, Christiansen & Hou 2016):

$$\sigma_t{}^2 = \omega + a_1 n_{t-1}{}^2 + a_2 I_{t-1} n_{t-1}{}^2 + \beta \sigma_{t-1}{}^2, \quad (2.25)$$

where: I_{t-1}=1 if $n_{t-1} < 1$ and I_{t-1}=0, otherwise. The impact of $n_{t-1}{}^2$ on $\sigma_t{}^2$ depends on the sign of the error term. Therefore, positive shocks affect a_1, while the negative shock affects $a_1$and a_2(Cifter, 2012)

When compared to the EGARCH, the leverage effect has a different sign (Dahlvid & Granberg, 2017) such that if α_2 = 0, there is symmetry (i.e. absence of asymmetric volatility), if α_2 < 0 then positive shocks increase volatility more than negative shocks and vice versa vice if α_2 > 0.

2.2.8 The Hybrid SARIMA-GARCH model

The ARMA, ARIMA and SARIMA models as conventional time-series models were developed under the presumption that variance does not change over time (seasons) (Pandey *et al.*, 2019). To develop a hybrid of SARIMA and GARCH models, firstly, the seasonal factors in the sequence are eliminated, and then the traditional SARIMA model is employed to model the time series linear data (Yaziz *et al*., 2013). Secondly, the non-linear patterns are modelled using the GARCH model and to correct the residual error (Yaziz *et al.*, 2013). Lastly, the suitable SARIMA model is combined with the suitable GARCH model to make the Hybrid SARIMA-GACH model to examine the univariate series and to forecast the future values. The models discussed in this theoretical review section are further explained in Chapter 3, which explains how the models will be implemented in the current study.

2.3 EMPIRICAL LITERATURE REVIEW

The study further reviews the empirical literature to analyse how other authors applied different models and techniques in pursuit of accurate and reliable results. The focus is on models that are relevant to the current study, and the subsections review applications of SARIMA, GARCH, hybrid models and previous papers on electricity consumption.

2.3.1 Application of ARIMA and SARIMA models

In this subsection, the previous different studies that employed the models that are considered in the current study are discussed. The studies are then linked to the current study to highlight the contribution of the current study to the literature around the models of interest.

The study by Christodoulos *et al*. (2010) used a conventional aggregate diffusion model (a logistic family model) and the Box-Jenkins approach (ARIMA) model to improve the short-term prediction of forecasting models. The study used the data obtained from the International Telecommunication Union (ITU) for the period of 1998 to 2008 to forecast one year ahead. The mean squared error (MSE), the mean absolute percentage error (MAPE) and mean absolute error (MAE) were employed as a selection criterion to select the suitable and best performing model. The study by Christodoulos *et al.* (2010) fitted the diffusion model and ARIMA model separately and then combined their forecasts. The findings of the study revealed that the new methodology (combining the diffusion model and ARIMA model) provided better predictions than each model separately. This current study intends to look at the ARIMA model, but extending the scope of the study by Christodoulos *et al.* (2010) by comparing this model with the hybrid SARIMA-GARCH, SARIMA, and GARCH models.

Etuk *et al*. (2012) used the Box-Jenkins interactive technique to analyse Nigerian inflation rates from the period 2004 to 2011. The competing SARIMA models were fitted. The goodness of fit was tested, and the results of the study revealed that the residuals are normally distributed with a mean of zero, which implies that the model is adequate. The study by Etuk *et al*. (2012) observed that the SARIMA $(0,1,1)\times(0,1,1)_{12}$ model fits the series best. The diagnostic checks show beyond doubt that the identified model is efficient enough to be used for forecasting the future Nigerian inflation rates. The current study intends to look at the SARIMA model but extending the scope of the study by Etuk *et al.* (2012) by comparing this model with the GARCH model and the hybrid SARIMA-GARCH model.
A forecasting study by Osarumwense (2013) used quarterly data of rainfall in Port Harcourt, south Nigeria which was obtained from the Central Bank of Nigeria Statistical Bulletin (2010). The data ranged from 1971 to 2008. The study employed Box-Jenkins methodology to model and forecast time series future values. The study implemented a SARIMA model and the adequacy of the model was shown to be appropriate. The diagnostic checks proved that SARIMA $(0,0,0) \times (2,1,0)_4$ was the suitable model to predict the rainfall in Port Harcourt. Similarly, the current study intends to implement SARIMA model, but extending the scope of

the study by Osarumwense (2013) by also modelling GARCH and hybrid SARIMA-GARCH model and comparing these competing models.

Osarumwense and Waziri (2013) used monthly data for the Nigeria Naira (N) to the USD ($) exchange rate. The study used the monthly time series data for the past 21 years from January 1990 to December 2010 obtained from the Central Bank of Nigeria Statistical Bulletin. The study employed the Box-Jenkins approach (ARIMA) to predict the future values of the monthly data. The findings of the study showed that ARIMA (0,1,1) was the appropriate model which satisfied the diagnostic checks. Therefore, the ARIMA (0,1,1) was employed to predict the Naira to USD exchange rate for the next four years. The current study intends to employ the SARIMA model but extending the scope by also comparing the model with other models (GARCH model and hybrid SARIMA-GARCH).

The study by Boran (2014) used the Box-Jenkins methodology to predict Turkey's net electricity consumption from 2009 to 2013. To eliminate the exponential impact, the study transformed data into natural logarithms. The ARIMA model was employed, and the adequacy of the model was proven. The results of the study identified ARIMA (1, 1, 0) as the best model for predicting Turkish net electricity consumption. The study used the identified ARIMA (1, 1, 0) to forecast the Turkish net consumption for the next five years. Similarly, the current study uses the same electricity consumption data but used a different county (South Africa) and extended the scope by comparing the SARIMA model with other competing models.

The study by Shabri and Samsudin (2014) combined Box-Jenkins methods and the group method of data handling technique (GMDH) to model and forecast the lynx data in the Mackenzie River district of Northern Canada. The data used was from the period of 1821 to 1934 hence it consisted of 114 observations. The study compared the GMDH, Artificial Neural Network (ANN) and Box-Jenkins models. The best performing model was chosen using mean absolute error (MAE) and root mean square error (RSME) as the selection criteria. The study showed that the GMDH model outperformed the Box-Jenkins and ANN models. Different from the study by Shabri and Samsudin (2014), the current study intends to employ SARIMA, GARCH and hybrid SARIMA-GARCH models and compare which model outperforms the others.

The study by Jere and Siyanga (2016) modelled the inflation rate of Zambia using monthly consumer price index data that ranged from May 2010 to May 2014. The data were obtained from the Monthly Bullet organised by the Central Statistical Office of Zambia. The study used ARIMA models and Holt's double exponential smoothing to forecast Zambia's CPI. For selecting the best performing model, the MSE, MAPE and root mean square error (RMSE) were employed as selection criteria. The result revealed that ARIMA (12, 1, 0) is an efficient model to forecast Zambia's CPI; however, Holts' double exponential is as accurate as the ARIMA model when taking into account the minor deviations in the MAPE and MSE, and less complicated when modelled. The current study uses different data when compared to the study by Jere and Siyanga (2016) and intends to implement SARIMA, GARCH and hybrid SARIMA-GARCH model and compare which model outperforms the others.

Nyoni and Nathaniel (2018) modelled and predicted future values of time series data using Nigerian annual inflation rate data. The data was collected from World Bank and ranged from 1960 to 2016. The ARMA, ARIMA and GARCH models were employed to predict the inflation rate in Nigeria. Based on the minimum Theil's U forecast evaluation statistic, the three competing models – ARMA (1, 0, 2) model, the ARIMA (1, 1, 1) model and the AR (3) – GARCH (1, 1) model – were computed. The diagnostic tests proved that the ARMA (1, 0, 2) model is the most efficient model and it provides reliable forecasts of inflation in Nigeria. Similar to the study by Nyoni and Nathaniel (2018), the current study intends to look at the SARIMA and GARCH models but extending the scope of the study by also implementing a hybrid SARIMA-GARCH model and comparing these competing models.

The study by Mao *et al*. (2018) used the monthly incidence of tuberculosis (TB) to develop a forecasting model and also to analyse the seasonality of infections in China. The study used the monthly data from the National Scientific Data Sharing Platform for Population and Health of China, which ranged from January 2004 to December 2015. The Box-Jenkins technique was employed to model and fit the SARIMA model to predict the incidence of TB for the next six months. The measure of accuracy was tested using RMSE, MAE, MAPE and the Theil Coefficient (TIC). Mao *et al.* (2018) proposed the SARIMA model as a helpful tool for monitoring epidemics when examining the seasonal patterns of TB incidence in China. According to the findings of the study, SARIMA $(1,0,0) \times (0,1,1)_{12}$ was the model that best fit predicted monthly TB incidence. Similar to the study by Mao *et al*. (2018), the current study intends to extend the scope of the study by comparing the SARIMA model with the GARCH and hybrid SARIMA-GARCH models.

The study by Mwanga *et al*. (2017) implemented the SARIMA model to forecast quarterly sugarcane yields in Kenya. The study used secondary data obtained from the Agriculture, Food and Fisheries Authority - Sugar Directorate (AFFA-SD) Yearbooks of Statistics. The quarterly data ranged from 1973 to 2014. The study followed the Box-Jenkins modelling approach. The bias corrected Akaike Information Criteria (AICc) was used as a criterion for selecting the appropriate model, where the model with the smallest AICc was chosen. For model diagnostic checking, the Ljung-Box test for independence of residuals was applied. The findings showed the residuals for the best-fit model as being uncorrelated and following a white noise. The results revealed that SARIMA $(2,1,2) \times (2,0,3)_4$ was the suitable model. To assess adequacy of the forecast, the quarterly predictions for the year 2015, were compared against actual quarterly data. The results showed that the selected model predicted a decline in yields until the year 2020 and then it would gradually begin to rise again assuming other factors remained constant. Similar to the study by Mwanga *et al*. (2017), the current study intends to use the same model selection and fit of SARIMA models, but it extends the scope to implement and compare this model with other competing models (GARCH and hybrid SARIMA-GARCH models.)

2.3.2 Application of ARCH and GARCH models

Luo *et al.* (2010) examined the behaviour and modelled the volatility of crude oil price using daily data from 13 February 2006 to 21 July 2009. The study explored the GARCH, EGARCH and GRJ-GARCH models. The Markov regime-switching model was also employed using maximum-likelihood estimation method. Markov regime-switching models are known to have an advantage of classifying observed stochastic behaviour into normal and turmoil states, which can be very helpful in a market' financial crisis (Luo *et al.,* 2010)). The study found that the Markov regime-switching model yielded best fitted results with maximum log likelihood (LLC) value, while the GARCH model produced best-fit results with LLC value assuming that the data follows t-distribution. The current study also uses the GARCH model although different from the study by Luo *et al*. (2010), since the current study compared the SARIMA, GARCH and hybrid SARIMA-GARCH models.

The study by Ping (2013) compared two forecasting methods, which were the ARIMA and GARCH models. The study used daily Kijang Emas prices data for the period July 2001 and September 2012. For identifying the appropriate model, the goodness-of-fit was measured using the smallest SIC. The results revealed that GARCH (1,1) outperformed the ARIMA model as the most accurate model for forecasting future values. The best model was also confirmed by the lower value of MAPE. Similar to the current study by Ping *et al*. (2013) this study intends to look at the SARIMA and GARCH models but extending the scope of the study by comparing these models with the hybrid SARIMA-GARCH.

The study by Ngailo *et al*. (2014) applied the GARCH family models to model Tanzania's inflation rate. The monthly data from the Tanzania National Bureau of Statistics, consisting of 168 observations, which covered the period from January 1997 to December 2010, were gathered. The results showed that the standard GARCH (1,1) model outperformed other GARCH models used in the study. The MSE, AIC and BIC were used as selection criteria to select the appropriate model. Therefore, the selected model was used to forecast Tanzania inflation rates in a sample period (12 months) and results revealed that the predicted series is reasonably close to the actual series. Similarly, the current study intends to use the same selection criteria and also apply the GARCH model, although extending the scope of the study by also applying SARIMA and hybrid SARIMA-GARCH and compare these competing models.

Francq and Sucarrat (2018) forecasted the weekly stock market volatility using the GARCH models and two other non-linear models – Quadratic GARCH (QGARCH) model and GJR-GARCH model. The study revealed that the QGARCH model was recommended as the best forecasting model and improved on the linear GARCH models but only when extreme values were being excluded. The current study also intends using GARCH models when compared to the study by Francq and Sucarrat (2018) by extending the scope of the study by comparing this model with SARIMA and the hybrid SARIMA-GARCH models.

2.3.3 Application of hybrid models

The study by Sigauke and Chikibvu (2011) used data from 1996 to December 2009, which consist of 5097 observations, to forecast the daily peak electricity demand in South Africa. The study examined the SARIMA model with the hybrid SARIMA-GARCH model and the

Regression-SARIMA-GARCH model. In a comparative analysis, a piecewise linear regression model was employed. The AIC was used as a selection criterion while MAPE and RMSE were used as accuracy measures for selecting the best performing model. The results revealed that the Regression-SARIMA-GARCH model produced a more accurate forecast when compared to other competing models. The current study intends to look at the hybrid SARIMA-GARCH model to accommodate possibilities of serial correlation in volatility but extending the scope of the study by Sigauke and Chikibvu (2011) by comparing this model with the performance and contribution of each model independently (SARIMA and GARCH models).

The study by Tendai and Chikobvu (2017) employed the hybrid SARIMA-GARCH to model the monthly international tourist arrivals series data for Victoria Falls Rainforest. The data used was obtained from the Zimbabwe Parks and Wildlife Management (ZPWMA). The study used monthly data collection of foreign visitors from January 2006 to December 2016. The RSME, MAE and MAPE were used as the forecasting performance measures. It was noted that the SARIMA (1,0,0) $(0,1,1)_{12}$ model was the best fit although the presence of an ARCH effect on the residuals led to fitting the SARIMA-GARCH model. The results showed that the SARIMA (1,0,0) $(0,1,1)_{12}$-GARCH (1,0) model best fitted the monthly international tourist arrival volatility data when compared to other models, provided it follows the normal distribution. The current study intends looking at the hybrid SARIMA-GARCH model using different data and also extending the scope by examining the effectiveness of the SARIMA and GARCH models independently.

Dritsakis and Klazoglou (2018) employed SARIMA, GARCH and hybrid SARIMA-GARCH models to provide the best forecast using unemployment rates in the United States of America (USA). The study used the monthly data ranging from January 1955 to July 2017. The study used the RMSE, Theil's inequality coefficient and MAPE as accuracy criteria for selecting the best model. The SARIMA (1, 1, 2) $(0, 1, 1)_{12}$ -GARCH (1, 1) was identified as the appropriate model for predicting the USA's unemployment rates. The current study intends to use different data to compare performance of the abovementioned volatility models, although extending the scope by comparing the performance and contribution of each model independently (SARIMA and GARCH models).

Yaziz *et al*. (2013) modelled 40 daily gold price (per gram) data series for the time frame of November 2005 to January 2006. The study explored the effectiveness of a hybrid of ARIMA (flexible and powerful) and GARCH (superior and strength in handling volatility) models. The Box-Cox transformation was used to normalise data, to reduce heteroscedasticity and to address non-stationarity of variance. The findings revealed that the hybrid ARIMA (1,1,1)-GARCH (0,2) model passed the diagnostic checking and heteroscedasticity tests; therefore, the model produced the best results and improved forecasting accuracy. The current study employs the same hybrid model although extending the scope by using the seasonal time series data and also using SARIMA and GARCH models to compare all competing models.

The study by Tang *et al*. (2015) examined ARIMA, SARIMA and hybrid ARMA-GARCH models along with belief functions to predict the total number of tourists arriving in China monthly. The study conducted analysis using monthly data from EcoWin that covered the period from January 1991 to June 2013. The study firstly modelled the time series using the observed data. The RMSE and MAPE measures were used as forecasting accuracy measures for selecting the best-fit model. The results revealed that a combination of belief function approach and time series models is a simple and effective method to forecast demand for Chinese international tourism. The study by Tang *et al*. (2015) combines the time series models with the belief functions, although the current study intends combining time series models with ARCH family models, which are known volatility forecasting models.

The study by Molebatsi and Raboloko (2016) implemented an ARIMA model and improved the results by including a GARCH model to address volatility in the series. The study used the Botswana Consumer Price Index (CPI) data, and the data was collected from Bank of Botswana reports. The data set used in the study ranges from January 2005 to December 2014. The findings revealed that ARIMA (1,1,1) is the suitable model to model inflation measured by Botswana’s CPI. The study further implemented the GARCH model with the aim of improving the results and to bridge the gap of heteroscedasticity of error terms, since the ARIMA model disregards the volatility in the series. The findings revealed that even though volatility for Botswana’s CPI is weak, the two best models for forecasting were ARIMA (1, 1, 1) and ARIMA (1, 1, 1) - GARCH (1, 2). The current study extends the scope by using the seasonal time series data and also models SARIMA and GARCH models to compare all competing models.

Mukaram and Yusof (2017) applied the ANN, SARIMA, and a hybrid of ANN and SARIMA models using the daily data of Malaysia’s solar radiation. The data were obtained from three stations in Malaysia (Batu Embun, Hospital Jelebu, and Kluang) and then transformed it into monthly data using the average. The MAE, MAPE and RMSE were used as performance and forecasting accuracy measures to aid in the selection the best performing model. The findings revealed that a hybrid of ANN and SARIMA models is the best model for forecasting Malaysia's daily average solar radiation data compared to the SARIMA model alone and/or ANN model alone. The current study intends on using monthly data and comparing the SARIMA, GARCH, and hybrid SARIMA-GARCH models.

The study by Pandey *et al*. (2019) used the monthly rainfall time series data from two distinct climatic environments (humid site and arid region). The monthly data ranging from 1953 to 2012 was used in the study through implementation of the SARIMA model. Furthermore, the best-developed SARIMA model residuals were fitted to the GARCH model which was then used to develop a hybrid SARIMA-GARCH model. Furthermore, in order to eliminate the impact of heteroscedasticity in data and stabilise conditional variance, transformed (Box-Cox and differenced) data was computed. The highest coefficient of determination (R-square) and smallest RMSE were used as selection criteria to select the best-performing model. The results showed that, based on transformed monthly rainfall time series, the hybrid SARIMA-GARCH model produced the best results for both climatic environments. Similarly, the current study intends looking at the SARIMA, GARCH, and the hybrid SARIMA-GARCH model using different data.

The study by Tseng *et al*. (2002) combined SARIMA and the neural network back propagation (BP) models and developed the hybrid SARIMA-BP model. The study examined two seasonal data sets from the Taiwanese machinery industry, the series of total production and the time series of soft drinks. The hybrid model developed, the SARIMA-BP model, was compared with other competing models, which were two neural network models and SARIMA models. These four models were employed using deseasonalized and differenced data. For selecting the most suitable model, MAE, MSE, and MAPE were used as accuracy measures to test their forecasting performance. The results revealed that the SARIMA-BP model was the appropriate model for forecasting the overall production for the Taiwanese machinery industry and the time series for soft drinks. Different from the study by Tseng *et al*. (2002), the current study intends using SARIMA, GARCH and the hybrid SARIMA-GARCH models, and comparing these competing models.

Chakhchoukh *et al*. (2010) modelled the stochastic aspects of electricity consumption in France using the parameters of SARIMA models and multiplicative double seasonal exponential smoothing. Furthermore, two new reliable methods for detecting and suppression outlier were examined, which are the multivariate ratio-of-medians-based estimator (RME) and the multivariate Minimum-Hellinger-distance-based estimator (MHDE) respectively. The forecasting accuracy and effectiveness of the proposed unique models were compared for the French electricity load time series. The findings revealed that the load time series shows breaks, which are long-lasting abrupt shifts in the stochastic pattern. The RME methodology performed better than competing models for “normal days” and has various interesting properties including good robustness, quick execution, simplicity, and simple online implementation.

2.3.4 Review of previous papers in modelling electricity

The study by Hondroyiannis (2004) employed cointegration techniques using the monthly time data of the period spanning from 1986 to 1999 to model the electricity demanded by residents of Greece. The vector error correction model (VECM) estimation was used to estimate three distinct variables: real income, the real price of electricity, and the weighted average temperature. The short-run deviations were also computed. The findings revealed that changes in real income, real price, weighted average temperature and real income of Greece in the long run will have an impact on residential demand for energy. The current study, by contrast to the study by George (2004), intends to analyse electricity consumption data using SARIMA, GARCH and the hybrid SARIMA-GARCH model.

The study by Dergiades and Tsoulfidis (2008) used the data for the years 1965 to 2006 to examine the residential demand for electricity in the US economy as a function of the per capita income, electricity prices, oil prices for heating purposes, weather conditions, and the stock of occupied housing. The analysis had two novel aspects. Firstly, it employed the stock of occupied homes as a proxy for the stock of electrical appliances, and secondly, it used the recently developed ARDL approach of cointegration to identify a potential equilibrium relationship between the variables. The results of the study supported a stable long-run relationship and suggested that short-run and long-run elasticity size and sign can be compared to other relevant studies. The current study intends to use the same data, but for South Africa and applying different models.

Ziramba (2008) examined the relationship between residential electricity demand in South Africa, real gross domestic product (GDP) per capita, and energy price between 1978 and 2005. The study by Ziramba (2008) employed a linear double-logarithmic form with income and price acting as independent variables in the empirical analysis. The findings revealed that in the long run, the income is the primary factor influencing demand for electricity, while price for electricity proved to be statistically insignificant. The current study intends using the same data with a broader focus by employing different volatility models.

2.4 CHAPTER SUMMARY

The chapter reviewed theoretical and empirical literature. Under the theoretical review, the preliminary tests comprising different types of stationarity, data transformation and unit root test were discussed. Furthermore, the stochastic modelling process, differenced stationarity process, model symmetric and asymmetric volatility were also discussed. The chapter also looked at the hybrid model of SARIMA-GARCH model.

Empirical literature was reviewed to look at different models and results revealed by previous studies. The study looked at the application of ARIMA and SARIMA models and studies by Osarumwense *et al*. (2013), Boran (2014), and Nyoni *et al*. (2018) found that ARIMA models produced the best results. For time series data with seasonality, studies by Etuk *et al*. (2012), Osarumwense (2013), Mao *et al.* (2018), and Mwanga *et al*. (2017) revealed that any industry can employ SARIMA models since they are effective for time series with seasonal trends.

On the contrary, the studies by Shabri *et al*. (2014) and Jere *et al.* (2016) found that other competing models (GMDH and Holt's double exponential smoothing,) can produce better results and are less complicated when modelled, considering the minor deviations in the MAPE and MSE when compared to Box-Jenkins family models. The study further examined the use of ARCH and GARCH models, due to the presence of ARCH effects that are a problem in time series data. The studies by Miswan *et al.* (2013), Ngailo *et al*. (2014), and Francq *et al*. (2018) confirmed that GARCH models are the best when compared to ARIMA and GJR-GARCH models and can address volatility in the series.

Various studies revealed that hybrid models provide better predictions. Studies by Sigauke *et al.* (2011), Tendai *et al*. (2017), Dritsakis *et al. (*2018), and Molebatsi and Raboloko (2016), specifically showed a hybrid of ARIMA/SARIMA and GARCH models produced better results, unlike when each model was employed separately (Christodoulos *et al.,* 2010). According to the study by Molebatsi and Raboloko (2016), the hybrid models can enhance results. As a result, the current study will examine the application of hybrid models, although the study by Mukaram and Yusof (2017) also found that a hybrid with SARIMA and other competing models can produce the best predictions. Lastly, the studies reviewed the previous papers in electricity consumption data, and it was found that the data can be used with different statistical models and produce accurate and reliable results.

The current study will examine the performance of the SARIMA model, which is shown to be flexible and powerful, and the GARCH model, which is superior and strength in handling volatility models. The study sets out to implement the GARCH model with the aim of improving the results and to bridge the gap of heteroscedasticity of error terms, since the SARIMA model does not take into consideration the volatility in the series.

CHAPTER 3

RESEARCH METHODOLODY

3.1 INTRODUCTION

This chapter examines the statistical models that can best be applied for addressing the research objectives of the study in the most effective way. The study adopts and discusses the Box-Jenkins methodology, which follows the four-phase method (Box, Jenkins & Reinsel, 1994). The volatility forecasting models are viewed to be competent for other roles rather than just forecasting future values of electricity consumption. Therefore, utilisation of such techniques reflects clearly the importance of electricity consumption from both the empirical and theoretical perspectives, as well as the effects of the predictions when there are unexpected changes.

The use of statistical techniques has led to several more powerful and effective time series tools for modelling and forecasting purposes. Among those various methods, this study aims to compare only a few models, namely, SARIMA, GARCH, and a hybrid of SARIMA-GARCH. These models have been reported to produce a good methodology and results for conditional variance modelling. The process for applying these volatility forecasting models was influenced by the aim to achieve the following objectives:

v. To determine the seasonal effect on electricity consumption.

vi. To model South African electricity consumption using the SARIMA, GARCH and SARIMA-GARCH models.

vii. To use the identified SARIMA, GARCH and SARIMA-GARCH models to forecast South African electricity consumption.

viii. To compare the forecasted values of the three models using forecasting accuracy measures.

ix. To provide recommendations based on the findings.

The chapter is organised as follows: Section 3.2 details the description of the data; section 3.3 discusses methods used in presenting the descriptive statistics, 3.4 outlines statistical methods and 3.5 gives the chapter summary.

3.2 DATA DESCRIPTION

The study makes use of secondary time series data spanning the period from 3rd January 2000 until 3rd May 2020 and consists of 221 observations. The data was obtained from Quantec Easy Data and it is accessible through the following link: https://www.easydata.co.za/dataset/P4141/timeseries/P4141-ELEKTR22/. Timely updates are released by Quantec together with its various primary data providers. Quantec is a consultancy providing economic and financial data that enables researchers to find and explore relevant information and convey results rapidly. Availability of data at Quantec Easy Data starts from 1985, but according to the South African Electrification Programme, in the second phase (1994-2000) there were new approaches such as completing institutional improvements and drafting new policies within the electricity sector. In the third phase, from 2000 to the present, there is an electrification increment level from around 35% to 71% consumption of electricity, hence the researcher decided to use the data set from 2000 with the aim of addressing research objectives. The data analysis is done using statistical forecasting software Eviews 9 (x64) and R-Studio (R 3.4.4). The following Section 3.3 discusses the methods used to compute the descriptive statistics.

3.3 DESCRIPTIVE STATISTICS

The descriptive statistics procedure outlines statistical information about a single variable. This procedure is used to summarise large data statistically, and usually a descriptive statistics report is large and complicated. Therefore, a summary table is used to simplify analyses. The study will focus only on few important statistics with the aim to accurately describe the nature of the data, which are categorised as location, spread, shape indicators, and normality. The measures are discussed in the following subsections.

3.3.1 Central Tendency or Location

The central tendency is the summary statistic that represent the centre point of the dataset. The importance or first impression of the variable is its general location. This measure indicates where the most values in a distribution fall (Lee & Wong, 2001). The main measures of central tendency when analysing the centre of the variable on the number line are the mean (average), median and mode, although there are other averages such as harmonic mean, trimmed mean and geometric mean, which have specialised uses.

The best measure of central tendency may vary depending on the type of data being used. In a symmetric distribution, the mean locates the centre accurately. The mean, median, and mode are equal if the data is normally distributed. If the mean and median are significantly different, there are probable outliers in the data, or a skewed distribution. If this is the situation, the median is likely to be a better measure of location. Each of the common measures has a unique approach to determine the position of the central point. The mean is the arithmetic average of the data values and is known as a common measure of location. One of its main characteristics is that it minimises prediction error for any one value in the data set. Additionally, the sum of the deviations from the mean is always equal to zero when using the mean as a measure of central tendency. Therefore, the researcher will use the following formula to compute the mean is given by:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}, \tag{3.1}$$

where n the number of data points and x_i is each of the values of data. The mean can be seriously contaminated by a change of any variable because it is sensitive to extreme values. The median is the middle number or th*e* *50*th percentile of the sorted data values; it splits the data set into half. Prior to calculating the median, the data should be arranged in ascending order starting with the smallest value to the highest value, then one finds the data point that splits the values equally. The method for locating the median varies when the data set has an odd number or an even number of values.The median is given by:

$$median = z_{50} \tag{3.2}$$

When the data set contains more extreme values, the effect on the median is smaller. When compared to the mean, the median measures more accurately when the distribution is skewed.

The mode is the value that appears the most frequently in the data, or the most common value. The mode is the only measure of central tendency that uses categorical data although one can use discrete and ordinal data. Measures of central tendency, such as mean, median and mode among several others, are important in research to analyse the nature of the data that will be collected.

3.3.2 Variability, dispersion or spread

In statistics, variability, dispersion or spread are all synonyms that denote the overall degree of dispersion in the dataset. Variability describes the width of distribution, how far away the data points are spread or how closely the data points fall to the centre. The data points are likely to spread out from the centre if there is high dispersion. A low dispersion signifies that they are clustered tightly are around the mean. Standard deviation, interquartile range variance, and range are frequently used measures of variability.

Standard deviation is one of the most popular methods often employed for measuring dispersion. It is greatly influenced by use of original values of the data set, which has a significant impact on making interpretation more convenient and straightforward. The sample standard deviation is a popular measure that is denoted by s. It measures the distance each data point is on average from the mean. It is the square root of the sample variance and is given by:

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \ , \qquad (3.3)$$

where $\bar{x}$ is the mean of all values in the data set and all other parameters are defined under equation 3.1. Standard deviation is minimal when the data points are grouped together. However, standard deviation is high when the values are far away, and standard distance is greater.

There are two numeric measures of shape, which are skewness and kurtosis. Skewness gives the degree and direction of deviation from horizontal symmetry. Kurtosis indicates the tallness and sharpness of the centre peak, in relation to the standard bell shape. The distribution is normal when it has exactly excess kurtosis of 0 (kurtosis exactly 3) and skewness. The data is positively, or skewed to the right, when the majority of the data points are on left and the right tail is longer. The distribution is negatively, or skewed to the left, if the peak is towards the right and the long tail is to the left. The data is perfectly symmetrical if skewness = 0 but is quite unlikely for real-world data. The moment coefficient of skewness is given by:

$$skewness: g_1 = \frac{m_3}{{m_2}^{3/2}} \ , \qquad (3.4)$$

where

$$m_r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^r}{n}, \tag{3.5}$$

where n is the sample size, $\bar{x}$ is the mean , m_3 is the third moment of the data and m_2 is the variance. Bulmer (1979) suggested the following rule of thumb:

- The distribution is strongly skewed if the skewness is more than 1 or less than -1
- The distribution is moderately skewed, if skewness is between -1 and -0.5 or between +1 and +0.5.
- The distribution is approximately symmetric if skewness is between -0.5 and +0.5.

Traditionally, kurtosis has been explained in terms of the measure of degree of peakedness or flatness in the distribution. The kurtosis measures the heaviness of the tail in the data distribution. Lower values indicate a lower, less sharp peak. Higher values indicate a higher, more distinct peak. Moment coefficient of kurtosis is given by:

$$kurtosis = \frac{m_4}{{m_2}^2}, \tag{3.6}$$

where

$$m_r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^r}{n}, \tag{3.7}$$

where n is the sample size, $\bar{x}$ is the mean , m_4 is called the fourth moment of the data and m_2 is the variance. Kurtosis is often measured against the normal distribution:

- "Mesokurtic distribution" refers to a distribution with excess kurtosis close to 0 and kurtosis close to 3.
- "Platykurtic distribution" refers to a distribution with lighter tails and usually with a lower and broader centre region, with kurtosis less than 3 and excess kurtosis less than or closer to 0.
- "Leptokurtic distribution" refers to a distribution with heavier tails and often a central peak that is sharper and higher, with kurtosis greater than 3, and excess greater than 0.

The GARCH model frequently assumes that time series data is conditionally normally distributed. The model can be referred to as normal GARCH. However, the kurtosis implied by the normal GARCH model tends to be far smaller than the sample kurtosis that was actually observed. Thus, the non-normal GARCH model is more appropriate with the large leptokurtosis typically observed.

3.3.3 Normality test

The descriptive statistics are also used to determine the normality of the distribution. When using the GARCH model it is commonly assumed that the data is conditionally normally distributed (Ghahramani & Thavaneswaran, 2008). There are several methods for assessing normality of the distribution. For the purpose of this study, the Jarque-Bera (JB) test is employed. The JB test is a type of Lagrange multiplier test that tests for normality. The test is named after Jarque and Bera (1980, 1987). It is usually used for large data sets, because it is reliable when n is large as compared to other normality tests. The data is said to be large when n>30 therefore the current study dataset is large. The test statistics is a function of the measure of kurtosis and skewness of the data to confirm if the distribution is normally distributed. The test statistic of Jarque-Bera is given by Thadewald and Büning (2007) as:

$$JB = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right), \tag{3.8}$$

where

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{\frac{3}{2}}} \tag{3.9}$$

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{2}} \quad , \tag{3.10}$$

where $\hat{\mu}_3$ and $\hat{\mu}_4$ are the estimates of third and fourth moment of the data, respectively. The $\bar{x}$ is the mean and $\hat{\sigma}^2$ is the variance (estimator of the second central moment). To determine whether the data represent a normal distribution, analyse the p-value in relation to the significance level. If the data is not normally distributed, regardless of how non-normally distributed the data is, depending on the model being used, it may still produce reliable conclusions. Transformation of data can be used as a remedy for non-normality, and log transformations are commonly used.

3.4 STATISTICAL METHODS

Volatility forecasting models used in the study are SARIMA, GARCH and the hybrid of SARIMA-GARCH. The models will be used for sample prediction of electricity consumption. Section 3.4.1 discusses the SARIMA model.

3.4.1 SARIMA model

The SARIMA methodology is a multiplicative model that is frequently utilised by statisticians to analyse time series data. The SARIMA model was developed by Box and Jenkins in 1976. The SARIMA model is recommended because of its high level of accuracy when forecasting both short-term and long-term (Fang & Lahdelma, 2016). When seasonal data needs to be analysed, the SARIMA $(p, d, q) \times (P,D,Q)_s$ is employed. An 'ARIMA' (p, d, q) model is used to describe a non-seasonal ARIMA model. The general multiplicative SARIMA $(p, d, q) \times (P,D,Q)_s$, model used in this study is defined as:

$$\emptyset_p(B)\, \Phi_p\, (B^s)\nabla^d {\nabla_s}^D\, y_t = \text{c} + \theta_q\, (B)\Theta_Q\, (B^s)\varepsilon_t, \tag{3.11}$$

where $\varepsilon_t \sim N\,(0,{\sigma_t}^2)$, y_t represent electricity consumption observed monthly t (t= 1, 2…., n) and ε_t represents the error term at time t with variance ${\sigma_t}^2$ and potentially subject to conditional heteroscedasticity. The s is the seasonal length and B is a backshift operator. The SARIMA model follows the Box-Jenkins methodology comprising a four-step procedure. The four steps of the Box-Jenkins techniques are presented in Figure 3.1 and each step's contribution to accurate forecasting is discussed.

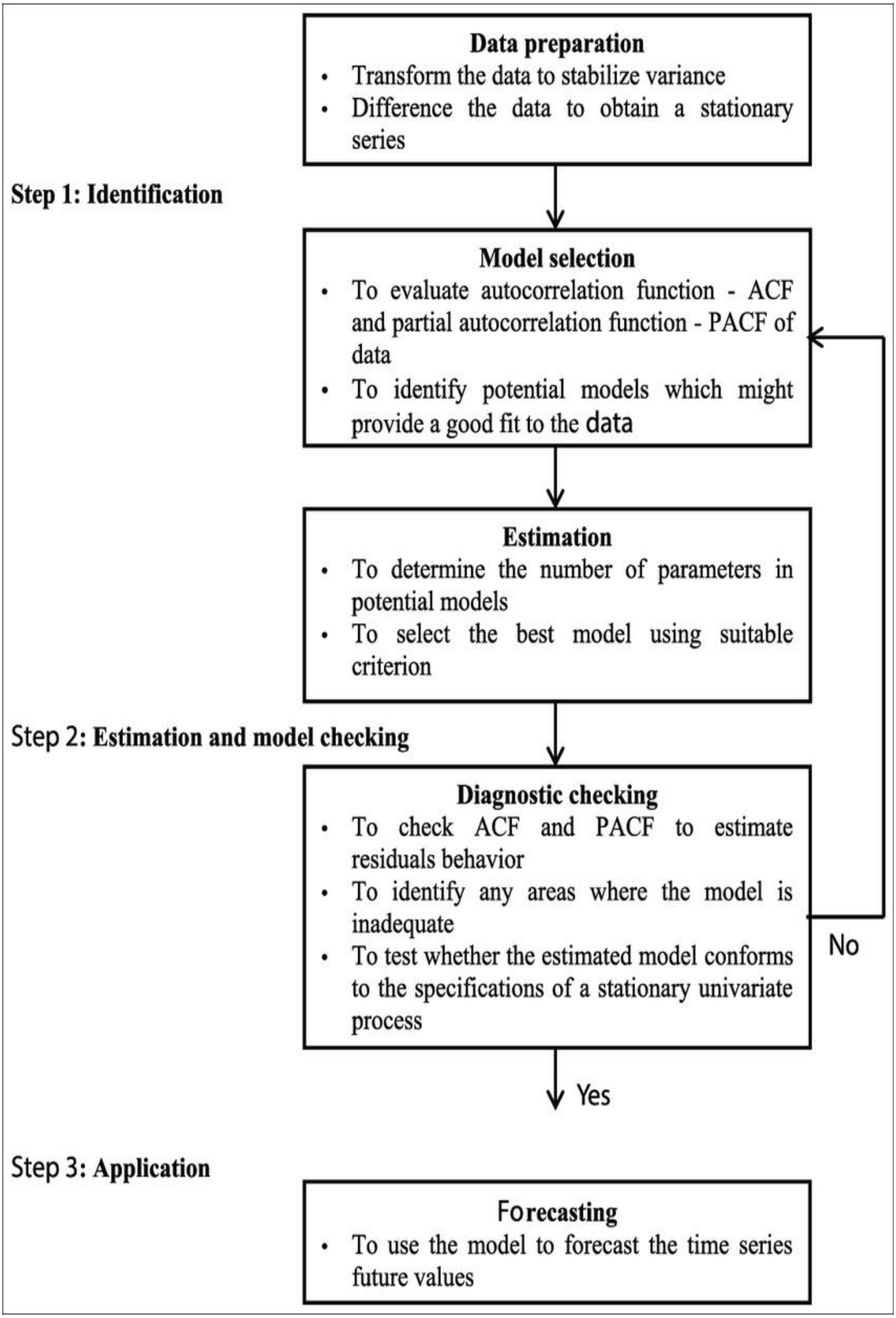


Figure 3.1: The steps of the Box Jenkins methodology by Dritsakis and Klazoglou (2018)

The four-step approach by Box-Jenkins (1976) comprises model identification, model estimation, diagnostic checking, and forecasting. The objectives of the Box-Jenkins (1976) technique are as follows:

i. To check for stationary behaviour and any seasonal elements
ii. To identify the ARMA order (p, q), ARIMA (p, d, q) and SARIMA $(P, d, q\ )(P, D, Q)_s$ and SARIMA-GARCH model
iii. To check if the identified model is adequate.
iv. To use identified "best fit model" to forecast future values for South African electricity consumption.

The steps of the Box-Jenkins methodology are discussed in detail in the following sub-sections.

3.4.1.1 Data preparation

The study will use secondary data collected from Quantec Easy Data. The first step is to examine the nature of the data, and the best way to determine any patterns, such as trends, seasonality, cycles and their significance, is to use a time series plot. Additionally, plotting also helps in identifying if any outliers exist in the data set. Before performing any analysis, the data must first be cleaned in order to enhance the quality of the results.

Using data plots, at this stage the researcher is attempting to identify preliminary specification of the model. Besides looking at the graphical representation of the data, it is crucial to determine whether the data is stationary. The series values given by $y_{1,}y_{2,}\dots,\ y_{n,}$ must be stationary before identifying the pattern whereby the procedure assume zero mean, constant variance and autocorrelation structure.

Stationarity can be accurately defined mathematically or formulaically, but for the purpose of this study, time series plots are utilised for visual evaluation and they describes series without trend or with no periodic irregular rising and falling (flat looking), a constant autocorrelation structure, constant variance that does not change over time, and absence of periodic fluctuations, using an visual assessment. The objective of the model identification stage is to choose values of p, d and then q in the ARIMA (p, d, and q) model, assuming for the time being that there are no seasonal variations. Stationarity can be attained by removing sequence caused by the time-dependent autocorrelation. When the series exhibits non-stationary trend behaviour such trend may need to be handled differently.

A deterministic pattern can be fitted and eliminated; the series can be differenced or be transformed. Box-Jenkins (1994) appears to prefer differencing, although various authors favour the deterministic trend removal (Andrew & Tay, 2017). If a single regular differencing does not result in stationarity, it may be repeated more frequently; although it is uncommon for it to become stationary if is differenced more than twice, so first differencing is usually adequate. For the purpose of this study, time series data may be differenced at most two times. The formula below indicates the first differencing:

P_t, creates the new series.

$$P_t = y_t - y_{t-1} , \quad (3.12)$$

where y_t is the observed value and y_{t-1} is the corresponding fitted value , t=1, 2 3 …. n.

The original data will have one point more than the differenced data. The data can be differenced more than once, and the formula for second differencing is as follows:

$$P_k = (y_k - y_{k-1}) - (y_{k-1} - y_{k-2}) = y_k - 2y_{k-1} + y_{k-2} , \quad (3.13)$$

where k=1, 2 3 …. n.

The simplest illustration of anon-stationary variable is:

$$y_k = y_{k-1} + \varepsilon_k \varepsilon_k : \text{WN}\,(0,\delta^2). \quad (3.14)$$

Complies with the White Noise process. This is an AR (1) process; however, it has only one root ϕ, which is equal to one.

$$y_k = \phi y_{k-1} + \varepsilon_t , \quad (3.15)$$

where ϕ = 1, ,y_{k-1} is the corresponding fitted value for second differencing, and ε_t is the error term

The following are the hypotheses of interest:

H0: $\varphi = 1$ (y_k non-stationarity) ⇒ (unit root in $\varphi(z) = 0$) ⇒ $y_k \sim I(1)$:

H1: $|\varphi| < 1$ (y_k stationarity): ⇒ $y_k \sim I(0)$

The test statistics are given as:

$$t_{\Phi}=1 = \frac{\widehat{\varphi} - 1}{SE(\widehat{\varphi})}, \tag{3.16}$$

where $\widehat{\varphi}$ is the least squares estimate and SE($\widehat{\varphi}$) is the normal standard error estimate (Caner & Kilian, 2001). In addition to visual inspecting the series plots, the study will employ formal tests to check for stationarity. Formal tests for stationary will also be used to determine whether the data is stationary or not. Unit root test are usually used to check for stationarity, and they typically use the null hypothesis that there is existence of unit root, and the alternative hypothesis is either trend stationary or stationarity depending on the test performed (Russo *et al.*,2019). According to Baruník *et al.* (2014), the purpose of the unit test is to detect whether the random process has a unit root or is constant. The unit root testing approach is generally defined as follows:

$$y_t = D_t + Z_t + \mathcal{E}_t \ , \tag{3.17}$$

where D_t denotes the deterministic component (Trend, seasonal component etc.), Z_t is the random process and $\mathcal{E}_t$ is the stationary error process. Several tests such as the Augmented Dickey Fuller (ADF), Phillips-Perron test (PP) and ADF–GLS test are among the tests found in the unit root tests.

The ADF test is frequently conducted on large samples of data. Since the study uses a large sample, the ADF test is relevant. The ADF test is an upgraded model design of the Dickey-Fuller (DF) test utilised when a time series data set is extensive. The test was named after the statisticians Dickey and Fuller who invented the test in 1979. The ADF test determines whether there is unit root in a time series data, and the alternative is different depending on the type of test performed. The hypotheses of the ADF test are stated as follows:

H_0: Unit root exists in first differenced data

H_a: Unit root does not exist in first differenced data

The ADF test procedure and Dickey Fuller (DF) test are more comparable tests computed by the following equation:

$$\Delta y_t = \alpha + \beta t + \gamma y_{t-1} + \ldots + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t , \tag{3.18}$$

where α is a constant, β denotes a coefficient on a time trend and p is the lag order of the autoregressive process. According to Dickey and Fuller (1979), the more negative test statistics value, the stronger the rejection of the hypothesis given by unit root exists in first differenced data at some level of confidence. When y_{t-1} is subtracted from both sides, with the aim to test whether ϕ is equal to 1, AR (1) model can be re-written as:

$$\Delta(y_t) = y_t - y_{t-1} \tag{3.19}$$

$$= (\phi - 1)\, y_{t-1} + \varepsilon_t . \tag{3.20}$$

Now the test of ϕ = 1 is a t-test to test whether parameters of the lagged level are zero is called the D-F test. To accommodate the general ARIMA and ARMA models, the D-F test is modified into the Augmented D-F test if there are higher-order of AR dynamics (or ARMA dynamics that can be approximated by longer AR terms). Suppose an AR (3) is specified as:

$$y_t - \phi y_{t-1} - \phi_2 y_{t-2} - \phi_3 y_{t-3} . \tag{3.21}$$

The backshift operator can be defined as:

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)\, y_t + \varepsilon_t . \tag{3.22}$$

The next sub-section identifies the model.

3.4.1.2 Step 1: Model identification

The autocorrelation plots, partial autocorrelation plots, and other relevant statistical technique are utilised to select the simple ARIMA models by estimating appropriate values for *p, d,* and *q*. When the series is stationary, the arrangement of moving averages (MA) and autoregressive (AR) is examined by using the contiguous approach by including the lags of the order *p*, ADF test considers higher requests autoregressive methods. The autocorrelation and partial autocorrelation plot must therefore determine the lag length *p*.

Autocorrelation refers to characteristic linear reliance of a variable in which there is a relationship between the qualities of the same variables at different periods in time (Gao *et al*.,2014). Autocorrelation between two observations for stationary processes depends on the time lag-*f* between them. The autocorrelation is described as:

$$cov_{\left(y_{t,}y_{t-k}\right)} = \gamma k \ , \tag{3.23}$$

where lag-*f* autocorrelation is given by:

$$p_k = corr_{\left(y_{t,}y_{t-k}\right)} = \frac{\gamma k}{\gamma 0} \ , \tag{3.24}$$

where

denominator $\gamma 0$ is the lag 0 covariance.

Correlative linear dependency on the confounding produces variable independence as the outcome. Partial autocorrelation gives the partial correlation of a time series between y_t and y_{t-k} after removing any linear dependency on y_1,y_2, , $y_{3,}$…,y_{t-k+1}. The partial lag is written as $\Phi_{k,k}$. The ACF does not control any lags; in contrast to the partial autocorrelation it regulates the values of all shorter lags. When analysing data, this function plays a pivotal role that aims at identifying the magnitude of the lag in an autoregressive function. In Box-Jenkins steps for time series modelling, this function is employed, whereby the plots can be used to estimate the appropriate lags *p* in an AR *(P*) model or an expanded ARIMA (*p, d, q*) model.

The ACF for a time series y_t, t = 1, 2…., *N* is the order, p_k,$k_=$ 1, 2, *N*-1 and the PACF order is $\Phi_{k,k.}$*k*= 1, 2….*N* – 1. Theoretically, there is a difference between the AR, MA and ARMA models. There is a significant difference amongst ACF and PACF models which might help when choosing the appropriate model. Box-Jenkins (1994) recommended using Table 3.1 below to explain the behaviour and characteristics of the ACF and PACF models and choose the best model for time series. According to Box-Jenkins (1994), the ACF and PACF can be examined with the well-known theoretical autocorrelation functions as a qualitative model selection tool for time series data.

Table 3.1: ACF and PACF behaviour (Box & Jenkins, 1994)

Conditional Mean model	**ACF**	**PACF**
AR(*p*)	Geometric decay	Cuts off after lag p
MA(*q*)	Cuts off after lag q	Geometric decay
ARMA(*p*,*q*)	Tails off gradually	Tails off gradually
White noise	All zero or close to zero	All zero or close to zero

3.4.1.2.1 Seasonality and seasonal ARIMA models

The discussion has thus far been restricted only to non-seasonal ARIMA models, which are applicable to time series without any seasonal component, despite the fact that many time series show a periodic behaviour. The ARIMA model is expressed by the notation ARIMA(p, d, q) where p is the order of the autoregressive part, d is the order of the differencing and q is the order of the moving average process. In the absence of differentiation $(d = 0)$, the models are referred to as ARMA (p, q) models. Both mathematically and statistically, the pure ARIMA model is defined as:

$$W_t = \frac{\Theta(B)}{\phi(B)} a_t + \mu \,, \tag{3.25}$$

where W_t denotes the difference of the response series, *t* indicates time, *B* denotes the back shift operator and μ is the mean term. The $\Theta(B)$ is the moving average operator, represented as a polynomial in the back shift operator $\Theta(B) = 1 - \Theta_1 B - \ldots - \Theta_q B^q$. The series w_t determines both the original and the estimated model, hence w_t is the response of Y_t or a differenced y_t. Simple non-seasonal differencing can be expressed as $w_t = (1 - B)^D y_t$ and seasonal differencing as:

$$W_t = (1 - B)^d (1 - B^S)^D Y_t \,, \tag{3.26}$$

where *d* is the degree of non-seasonal differencing, *S* is the length of the seasonal cycle and *D* is the degree of seasonal differencing.

The seasonal and non-seasonal ARIMA models are both recognized similarly, with the exception that the autocorrelation patterns of the ACF and PACF lags are the primary consideration when determining seasonality. Differencing operators are often used in the ARIMA models for time series with constant seasonality, when the lags exceed the seasonal cycle's length the AR and MA parameters are used. The constant factor out and appended to the ARIMA *(p, d, q*) notation when all of the lags in an ARIMA models are constant. For the order of a general series for the seasonal ARIMA model and non-seasonal and seasonal factors is defined as:

ARIMA *(p, d, q)* × $(P, D, Q\ )_S$,

where *p, d* and *q* gives the order of the non-seasonal part of the ARIMA (*P, D, Q*) that gives the order of the seasonal part. The *S* is the number of observations in a seasonal cycle (7 for daily series with day-of-week effects, 4 for quarterly and 12 is for monthly). The 12-month seasonal cycle is appropriate for the purpose of this study. Therefore, the seasonal ARIMA model for monthly data is given in the following statistical form:

$$(1-B)(1-B^{12})Y_t = \mu + \left(1 - \Theta_{1,1}B - \Theta_{1,2}\,B^2\right)\left(1-\Theta_{2,1}B^{12}\right)a_t\ . \qquad (3.27)$$

The next step is to estimate the model.

3.4.1.3 Step 2: Model estimation

This phase the 'best-fit model' is chosen using the maximum likelihood techniques, back casting and other techniques as recommended by Box-Jenkins (1976), such that the estimated values are as close to the observed values as much as possible. The Ordinary Least Square (OLS) estimation is a form of statistical regression analysis used to determine the unknown values of the parameters, $\beta_{0\,,}\,\beta_{1\,,}\,\beta_{2\,,}$…, and to find the line of the best fit for a data set that minimizes the sum of the squares errors generated by related equations. The approach uses the minimum residual sum of squares to select the model parameters. Statistically, the minimised least sum of squares criterion for obtaining the parameter estimates is expressed as follows:

$$Q = \sum_{i=1}^{n}\left[y_i - f\left(\overrightarrow{x_i};\ \vec{\beta}\right)\right]^2\ . \qquad (3.28)$$

Estimation of autoregressive model is more likely comparable to a linear regression model exhibiting a delayed effect on the dependent variable. For instance, estimating parameters for an AR(1) for moving average is identical to applying an OLS to y_t; as mentioned above, estimating parameters is quite challenging due to an unobserved error term ε_{t-i} in the model. The next stage is to select the best-fit model.

3.4.1.4 Step 3: Model selection

When selecting the best-fit model, the models that have large R^2 may appear to fit the data extremely well yet produce forecasts that are not accurate and unreliable. The objective is to choose the model that minimises the Information Criterion's value. Amongst an unlimited set of models, the Akaike Information Criterion (AIC), Bayesian information Criterion (BIC) or Schwarz Criterion (SBC) are widely employed as a selection criterion. For the purpose of this study, the BIC that was developed by Schwarz (1978) is considered to be adequate and the best for approximating the data and to minimize the loss of information. By including additional parameters when fitting the models, it could be reasonable to improve the chances, however that can lead to fitting. The BIC addresses this issue by presenting a penalty term, which is greater than in AIC. The BIC assumes that the data is extracted from an exponential family distribution. Consider:

$p(x|k)$,

where *x* is observed data, *k* represent the number of regressors, including the intercept, and *n* denotes the number of data points or numbers of observations. The formula for the BIC is expressed as:

$$-2.\ln(x|k) \approx BIC = -2.\ln L + k\{In(n)\}, \qquad (3.29)$$

where *L* presents the maximized value of the likelihood functions for the estimated model. With the assumption that the model errors are independent and evenly distributed as per the normal distribution, then

$$BIC = n\{In(\hat{\sigma}_\varepsilon^2)\} + K\{Ln(n)\}, \qquad (3.30)$$

where $\hat{\sigma}_\varepsilon^2$ is the error variance and is written as:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_{i-\bar{x}})^2 \quad . \qquad (3.31)$$

The model with smallest BIC is usually regarded to be the appropriate model. According to Granger and Jeon (2004), on average, the BIC generally only fits estimated models when the numerical values of the responding variable are similar across all estimates being compared.

The Box-Lung (BL) test is utilised to determine the fitness of the time series model. After fitting an ARMA (p, q) model to the data, the BL test is employed to the residuals to analyse the autocorrelation of the residuals of a time series. Statistics check whether the autocorrelations are less than zero, and the model is significantly unfit for the data if the autocorrelation is extremely small.

The BL test hypotheses are formulated as follows:

H_0: The model does not exhibit lack of fit
H_a: The model exhibits lack of fit

The BL test statistics is given as:

$$Q = n(n+2)\sum_{k}^{m}\frac{r^2{}_k}{n-k}\ , \qquad (3.32)$$

where *n* represent the number of observations, *m* is the length of coefficient being tested and r_k denotes the number of autocorrelations for lag *k*. The null hypothesis is rejected if the value of Q test is greater than Chi-square, indicating that the model has significant lack of fit. The Chi-square distribution is given as $X^2{}_{1-a,h}$, where α is a significance level and *h* is the degrees of freedom. The next step is diagnostic checking.

3.4.1.5 Step 4: Diagnostic checking

This stage involves evaluating the fitted model identified and estimated to check if it is adequate by taking into consideration the autocorrelations of the residual series (the error or values). The objective of this step is to ensure that the model is suitable and appropriate to forecast future values. According to Vidalis *et al*. (1997), the diagnostic check is used to assess if the residuals of ARMA, ARIMA or SARIMA models satisfy the assumption underlying the innovation series. In the event that the residuals are randomly generated, the model is adequate (Vidalis *et al.*, 1997). Hence the current study examines residual randomness.

In the simple random model in time series, the observation deviates from the mean, which is independent and probably constant. In other words, the random time series is flat, there is no upwards nor downwards trend, the variance does not change over time. A random model is defined as:

$$y_t = \text{A} + e_t \text{ ,} \tag{3.33}$$

where A is a constant and e_t is the error term for residuals with constant variance and zero mean. The autocorrelation plots are widely used when examining the randomness of residuals in time series data at different points in time. The autocorrelation coefficient is established by the vertical axis autocorrelation coefficient:

$$R_h = \frac{C_h}{Co} \text{ ,} \tag{3.34}$$

where *Co* is the variance function computed using the following equation:

$$c_0 = \frac{\sum_{T=1}^{N}(y_t - \bar{y})^2}{N} \text{ ,} \tag{3.35}$$

and C_h is the auto-covariance function computed using the given equation:

$$c_h \frac{1}{N} \sum_{t=1}^{N-h}(y_t - \bar{Y})\,(Y_{t+h\,-\,\bar{Y}}), \tag{3.36}$$

where R_h ranges from -1 to 1. The more the PACF of an AR (*p*) process is lower at lag p + 1, the greater the assumption that the model is accurate. The final model selected, which is referred to as the 'best-fit model,' is then used to predict the future values of the time series data. The next and final phase is forecasting.

3.4.1.6 Step 5: Forecasting

Forecasting future values is the last phase of the Box-Jenkins process for the SARIMA model. According to Boylan (2005), forecasting methods can occasionally be impacted by the historical data, which affects the forecast accuracy. When forecasting the ARIMA model, the most accurate and reliable forecasting models, according to Tsoku *et al*. (2015), have the smallest Mean Absolute Percent Error (MAPE). Warant (2006) compared Box-Jenkins with other forecast methods such as Holt's forecast and combined forecast methods.

There are several straightforward measures that can be employed to test the accuracy of the forecast. The measures such as Mean Squared Error (MSE), Mean Squared Deviation (MSD), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are used to determine the accuracy of the forecasts. According to Hake and Wichern (2005), the MAPE is the best measure that can compare the model performance and the accuracy of the forecasts for two different series, because it is defined in relative and not absolute terms. Therefore, for the purpose of this study, the MAPE is the most suitable simple measure to be used. The MAPE is computed by:

$$\text{MAPE} = \frac{\sum_{t=1}^{n} \frac{|A_{t-F_t}|}{|A_t|}}{n} . \qquad (3.37)$$

According to Lewis (1982), the level of accuracy is divided into four, as is shown in Table 3.2 below.

Table 3.2: Accuracy levels for MAPE test

MAPE percentages	**Accuracy level**
MAPE ≤ 10%	Very accurate
10% < MAPE ≤ 20%	Accurate
20% < MAPE ≤ 50%	Medium accurate
50% ≤MAPE	Less accurate

3.4.2 Generalised Autoregressive Conditional Heteroscedasticity (GARCH) models

The Generalised Autoregressive Conditional Heteroscedasticity (GARCH) model is a leading model used by many researchers in modelling and forecasting volatility. Engle (1982) developed the ARCH model to address the issues of heteroscedasticity present in the time series. The GARCH (p, q) model developed by Bollerslev in 1986 proposes a method for estimating volatility using volatile series. Volatility is defined as the amount of variance in data over time as determined by the standard deviation. The GARCH (p, q) model involves the residual of a time series regression. The model is computed by:

$$Y_t = C + \varepsilon_t , \qquad (3.38)$$

The GARCH model is commonly used to simulate residuals of different models that exhibit an ARCH effect and unequal variance. The GARCH (p, q) model is normally fitted using the three commonly distributional assumptions, namely, normal distribution, Generalised Error Distribution (GED), and Student's t-distribution. According to Johansson and Sowa (2013), the GARCH (1, 1) model can be defined as:

$$\sigma_t^2 = \omega + \beta\sigma_{t-1}^2 + \alpha e_{t-1}^2 , \qquad (3.39)$$

where σ_t^2 is the constant variance, $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$ and $(\alpha + \beta < 1)$. An AA (p, q) process with a GARCH noise expressed by X_t that is used to model ARCH effect on residuals and linear dependence is given by:

$$X_t = \mu + \sum_{j=1}^{q} \beta_j \varepsilon_{t-j}^2 + \sum_{i=1}^{p} \alpha_i X_{t-i} + \varepsilon_t , \qquad (3.40)$$

where

$$\varepsilon_t = Z_t\sigma_t, Z_t \sim \text{ i.i.d N (0, 1)} , \qquad (3.41)$$

and

$$\sigma_t{}^2 = a_0 + a_1\varepsilon_{t-1}^2 + a_2\varepsilon_{t-2}^2 + \ldots + a_p\varepsilon_{t-p}^2 + \beta_1\sigma_{t-1}^2 + \beta_2\sigma_{t-2}^2 + \ldots + \beta_q\sigma_{t-q}^2 , \qquad (3.42)$$

where a_i, β_j and μ are model parameters and ε_t are the model residuals. It is crucial to determine the presence of heteroscedasticity before modelling the data with the GARCH procedure. The presence of heteroscedasticity is tested using the ARCH test. The ARCH test for lag $(p + q)$ is locally equivalent to a test for GARCH effects with lags(p, q). The null hypothesis to be tested is that there is no existence of ARCH effects. The alternative hypothesis is that ARCH effects exist; this is tested using lags up to T.

Furthermore, it is also important to test for the presence of autocorrelation in the residuals. The Ljung-Box test proposed by Ljung and Box (1978) is performed to check the autocorrelation significance of the time series by adjusting the degree of freedom at various lags. The Ljung-Box test statistic is given by:

$$Q(k) = T(T + 2) \sum_{j=1}^{k} \frac{r_j{}^2}{T-j} , \qquad (3.43)$$

where T is the sample size, k denotes the autocorrelations number of residuals and r_j denotes the sample autocorrelation at lag j. The null hypothesis for the test is that no correlation exists. The alternative hypothesis is that correlation exists. The null hypothesis is not accepted if the p-value associated with $Q(k) < \alpha$ is specified, normally it is 5% level of significance. The next step is to apply the GARCH (*p*, *q*) process using the equation stated below:

$$\sigma_t{}^2 = \omega + \alpha\varepsilon_{t-1}{}^2 + \beta\varepsilon_{t-1}{}^2. \tag{3.44}$$

The conditional variance, σ_t in the equation, is a weighted function of its long-run value, information about volatility during previous periods, $\alpha_i\varepsilon_{t-i}{}^2$ (the ARCH term), and the fitted variance from previous periods, $\beta_i\sigma^2{}_{t-i}$ (the GARCH term). The model tells us that tomorrow's variance is a function of today's squared innovations, today's variance, and the weighted average long-term variance (Ding, 2018). The model is subject to non-negativity constraints (as shown above) to ensure that the variance is strictly positive. The stationary condition of $\alpha + \beta < 1$ should hold to ensure weak stationarity of the GARCH process (Lim & Sek, 2013). The error term can be specified as:

$$\varepsilon_t = y_t - E\{y_t|\psi_{t-1}\}, \tag{3.45}$$

where ε_t is a random, unobservable variable with mean and variance conditional on ψ. The GARCH model for ε_t has $E\{\varepsilon_t|\psi_{t-1}\} = 0$ and $\sigma_t = E\{\varepsilon_t{}^2|\psi_{t-1}\}$ and is decomposed as:

$$\varepsilon_t = z_t h_t{}^{1/2} . \tag{3.46}$$

The sequence $\{z_t\}$ is an iid sequence of random variables with mean zero and unit variance.

3.4.2.1 Model comparison and selection

The parameters of the model can be estimated once the model order has been identified. Estimates can be simulated using various software such as R-Studio, which estimates the ARIMA model using MLE (Nakiyingi, 2016). This technique chooses parameter values that maximize the likelihood of obtaining the observed data. For ARIMA models, MLE is quite comparable to the least square estimates that would be obtained by minimising squared errors. Outputs display the log likelihood of the data, which is the logarithm of the probability that the observed data came from the estimated model. R tries to maximize the log likelihood when determining parameter estimates, p, d and q, are known (Hyndman & Athanasopoulos, 2014). The initial step in fitting SARIMA models is to select appropriate values for the two orders of differencing, both seasonal (D) and non-seasonal (d) to eliminate most of the seasonality and make the series stationary. Then an ARMA-type model is fitted to the differenced series with the additional complexity that there may be AR and MA terms at delays that are multiple of the season length s. For the purpose of this study, the model with the smallest BIC will be used as a diagnostic check and to forecast future values of electricity consumption in South Africa.

3.4.2.2 Diagnostics checking

Model diagnostics is a technique that is available for time series analysis that seeks to examine the relevancy and adequacy of the model in different ways. It also helps in assessing if a model satisfies the assumptions of time series. The assumptions include:

- Serial correlation
- Normality
- Heteroscedasticity

This study employs a number of diagnostic tests to test whether the selected model is adequate and efficient by testing the residuals of the models. The different types of diagnostic tests are discussed in the following subsections.

3.4.2.3 Serial correlation test

According to Williams *et al*. (2015), error terms are considered to be serially correlated when they are correlated with error terms from different point of time. Furthermore, Williams *et al.* (2015) indicated that "serial correlation happens when the errors associated with a given time period carry over into future time periods." In the current study, the Breusch-Godfrey (BG) (1980), Lagrange Multiplier (LM) test and the Ljung-Box Q-Test are used to test for serial correlation because they are preferred in most application for overcoming limitations of serial correlation tests

3.4.2.4 Breusch-Godfrey LM test

The BG LM test is used to test the existence of serial correlation in a time series data. The GARCH model described below serves the basis for BG LM test and is given by:

$$\sigma_t{}^2 = \omega + \alpha\varepsilon_{t-1}{}^2 + \beta\varepsilon_{t-1}{}^2 \,. \qquad (3.47)$$

The test statistic is given as:

$$LM_h = T\left[K - tr\left(\widetilde{\sum R}\,\widetilde{\sum e}\right)\right] , \tag{3.48}$$

where $\widetilde{\sum R}$ and $\widetilde{\sum e}$ are the corresponding residual covariance matrices for the confined and unrestricted models respectively. The test statistic LM_h is distributed as chi square χ^2(hK^2). The BG test hypothesis is formulated as:

H_0: $\rho_i = 0\ for\ all\ i$ (There is no serial correlation)
H_1: $\rho_i \neq 0\ for\ all\ i$ (There is serial correlation)

The null hypothesis is rejected given that probability in the BG test statistic is less than 5% level of significance.

3.4.2.5 Ljung-Box test

In time series, the Ljung-Box Q-test is frequently used to examine serial correlation and the volatility clustering impact of asset returns. The test statistic for the Ljung-Box test is expressed as:

$$Q_{LB} = T(T+2)\sum_{j=1}^{k}\frac{{r_j}^2}{T-j} , \tag{3.49}$$

where T is the number of observations, and k is the highest order of autocorrelation for which to test ${r_j}^2$ and j^{th} autocorrelation. The hypothesis for the Q_{LB}test is formulated as:

H_0: $\rho_i = 0\ for\ all\ i$ (There is no serial correlation)
H_1: $\rho_i \neq 0\ for\ all\ i$ (There is serial correlation)

If the probability in the BG test statistic is greater than 1%, 5% or 10% level of significance, the null hypothesis is accepted.

3.4.2.6 Heteroscedasticity test

Error terms are said to be heteroscedastic when they do not have a constant variance. The study will employ White's General test and the LM to test for the existence of heteroscedasticity.

➢ *White's test*

The White's test is a special case of the Breusch-Pagan test, which is a test for linear forms of heteroscedasticity (Williams *et al.*, 2015). The straightforward White's test for heteroscedasticity is uncomplicated to use and does not rely on the normality of the distribution. To test the assumption of heteroscedasticity, consider the following auxiliary equation:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X^2{}_1 + \beta_4 X^2{}_2 + \beta_5 X_1 X_2. \qquad (3.50)$$

3.4.2.7 Lagrange Multiplier (LM) test

The LM test proposed by Engle (1982) is a test for autoregressive conditional heteroscedasticity (ARCH) in the residuals. The test statistic is given by:

$$LM_E = nR^2, \qquad (3.51)$$

where n is the number of observations, and R^2 denotes regression's coefficient of determination. The hypothesis for the LM test is given as:

H_0: Residuals are homoscedastic

H_1: Residuals are heteroscedastic

If the probability of the test statistic is less than 5% level of significance, the null hypothesis is rejected, and the residuals are concluded to be heteroscedastic.

3.4.2.8 Normality test

The original GARCH model with conditional variance changes over time. The OLS residuals and the LM and Q statistics are estimated under the assumption that the disturbances follow a white noise process. The Jacque-Bera (JB) test will be used to test normality of the distribution. The JB test "measures the difference in kurtosis and skewness of a variable compared to those of the normal distribution" (Ssekuma, 2011:19). To test the presence of normality in the data set, the following hypotheses were formulated:

H_0: Residuals are normally distributed

H_1: Residuals are not normally distributed

The test statistic is given as:

$$JB = \frac{N-k}{6}\left[S^2 + \frac{(K-3)^2}{4}\right], \tag{3.53}$$

where N is the number of observations, k is the number of estimated parameters, S is the skewness of a variable and K is the kurtosis of a variable (Ssekuma, 2011). The formula for skewness and kurtosis are given as:

$$s = \frac{\sum_{t=1}^{T}(x_t-\bar{x})^3 / _{T-1}}{s^3} \tag{3.54}$$

and

$$k = \frac{\sum_{t=1}^{T}(x_t-\bar{x})^4 / _{T-1}}{s^4}\ . \tag{3.55}$$

The null hypothesis is rejected if the p-value is less than 5% level of significance.

3.4.2.9 Forecasting accuracy

The approach is used to examine the effectiveness of the developed models. In statistics, the core point for estimating the time series model is to use the estimated, best-fit model to predict the future values for policy evaluation and decision making. The best model for predicting future values, when compared to the computing models, produces minimum forecast errors and the accuracy is higher. Both in-sample and out-sample forecast are also used to check the accuracy of the model. The model producing fewer out-sample forecasting errors is considered the best-fit model. The effectiveness of the models is assessed to compare their adequacy and degree of accuracy using the following indices:

Mean absolute error (MAE)

$$MAE = \frac{1}{n}\left(\left|Q_j - Q_t\right|\right). \tag{3.56}$$

Root-mean-square error (RMSE)

$$RMSE = \sqrt{\frac{\sum_{j=1}^{n}\left(Q_j-Q_t\right)^2}{n}}. \tag{3.57}$$

Relative absolute error (RAE)

$$RAE = \frac{\sum_{j=1}^{n}|Q_j - Q_t|}{\sum_{j=1}^{n}|Q_j - \bar{Q}|}\,. \tag{3.58}$$

Mean relative error (MRE)

$$MRE = \frac{1}{n}\sum_{j=1}^{n}\frac{Q_j - Q_t}{Q_j}. \tag{3.59}$$

Coefficient of the determination

$$R^2 = \left[\frac{\sum_{j=1}^{n}(Q_j - \bar{Q})(Q_t - \bar{Q})}{\sqrt{\sum_{j=1}^{n}(Q_j - \bar{Q})^2 \sum_{j=1}^{n}(Q_t - \bar{Q})^2}}\right]^2, \tag{3.60}$$

where

Q_j is the actual observation.

Q_t is the forecasted value.

$\bar{Q}$ is the mean of the actual value.

n is the number of observation.

The next section 3.6 combines the SARIMA model and the GARCH model to form a SARIMA-GARCH model.

3.4.3 Hybrid of SARIMA-GARCH model

The variance of the error term of the SARIMA models follows a GARCH process in the hybrid of SARIMA-GARCH model. The model that will be used for modelling electricity consumption series can be written as:

$$\emptyset_p(B)\,\Phi_p\,(B^s)(1-B)^d\,(1-B^s)^D y_t = \phi_p(B)\phi_P\varepsilon_t, \qquad \varepsilon_t = z_t\sigma_t \quad , \tag{3.61}$$

where

$z_t \sim$ *i.i.d* with E (z_t)=0,Var (z_t) =1 (3.62)

and

$\sigma_t{}^2 = a_0 + \sum_{i=1}^{q} a_i\, \varepsilon_{t-i}{}^2 + \sum_{j=1}^{p} b_j\, \sigma_{t-j}{}^2$, (3.63)

where y_t represents the time series, p is the order of GARCH process, q is the order of ARCH process. The a_0,a_i and b_j are constants, ε_t is the error term, $\sigma_t{}^2$ is the conditional variance of ε_t, $\varepsilon^2{}_{t-1}$ is the news about the volatility from the i^{th} lag period and $\sigma^2{}_{t-j}$ is the j^{th} lag period forecast error variance,; $Z_t$is a standardized error term.

A model is referred to as a hybrid SARIMA-GARCH model when the variance of the disturbance term in a SARIMA model can be modelled by a GARCH process. The hybrid of SARIMA-GARCH model can be defined by (Sigauke & Chikobvu 2011):

$\phi_p(B)\Phi_P(B^s)\nabla^d\, \nabla_S^D\, Xt = \theta_q(B)\Theta_Q\, (B^s)e_t$, $e_t = z_t\sigma_t$, z_t~ (0, 1), (3.64)

where

$\sigma_t^2 = \mu + \sum_{j=1}^{q} \beta_j\, \varepsilon_{t-j}^2 + \sum_{i=1}^{p} \alpha_i\, \sigma_{t-i}^2$, (3.65)

where X_t is the monthly electricity consumption, p and q are the ARCH and GARCH orders of the model, μ is the constant term, β_j and a_i are the positive model coefficients whose sum should be less than one.

3.5 CHAPTER SUMMARY

The study will employ a three-phase approach – a SARIMA model, GARCH model and hybrid Box-Jenkins (SARIMA model follows Box-Jenkins and GARCH models. The linear data of the time series will be modelled using the Box-Jenkins approach and the residual of this linear model will only contain the non-linear data. The first phase of the study determines whether the data is stationary or not. The visual inspection as well as the formal test of stationary will be used. If the data is stationary, the suitable model is selected using the ACF and the PACF. The smallest BIC will be used to select and estimate the parameters of the best-fit model. For adequacy of the model, the Box-Ljung Q and R-squared tests statistics are used. In the second phase, the standard GARCH model is employed to model the non-linear patterns of the residuals, in order to handle the volatility present in the residuals of the data series. In the third phase, the study combines the SARIMA and GARCH model to form SARIMA-GARCH. The MAPE will be used to determine the best forecasting accuracy. The most appropriate and adequate model with the smallest BIC will then be used for forecasting the future values of electricity consumption.

CHAPTER 4

DATA ANALYSIS AND INTERPRETATION OF RESULTS

4.1 INTRODUCTION

This chapter presents the analyses and interpretation of the results. The results are presented using graphs and tables in line with the research objectives presented in the first chapter. The rest of the chapter is arranged as follows: Section 4.2 presents the preliminary data analysis, In Section 4.3 the results of the Box-Jenkins methodology are drawn, Section 4.4 illustrates the results of the ARCH test, Section 4.5 presents the results of the GARCH test, in Section 4.6 the results of Hybrid SARIMA - GARCH model are presented and finally, Section 4.7 presents the chapter summary.

4.2 PRELIMINARY DATA ANALYSIS

The preliminary data analysis displays the key features of the data before main analysis of the data is conducted using descriptive statistics. The descriptive statistics used in the study are mean, median, minimum (min), maximum (max), standard deviation (Std.Dev.), skewness, kurtosis and the Jarque-Bera (JB) test, together with its probability value. The results are presented in Table 4.1.

Table 4.1: Descriptive statistics of electricity consumption in South Africa

Obs	Mean	Median	Max	Min	Std.Dev.	Skewness	Kurtosis	JB	Prob
244	1436.508	1459.500	1739.000	865.000	147.613	-0.856	3.794	36.188	0.000

Table 4.1 present the summary statistics of electricity consumption in South Africa consisting of 244 observations. According to the results, the electricity consumed in South Africa ranges from 865 to 1739 Gigawatt-hours, with the average of 1436.508 and the standard deviation of 147.613. There is also evidence that the sample kurtosis for electricity consumed is just above 3. Therefore, it is concluded that the electricity consumption is nearly leptokurtic. This implies that the data has heavier tails than a normal distribution. Moreover, the results shows that

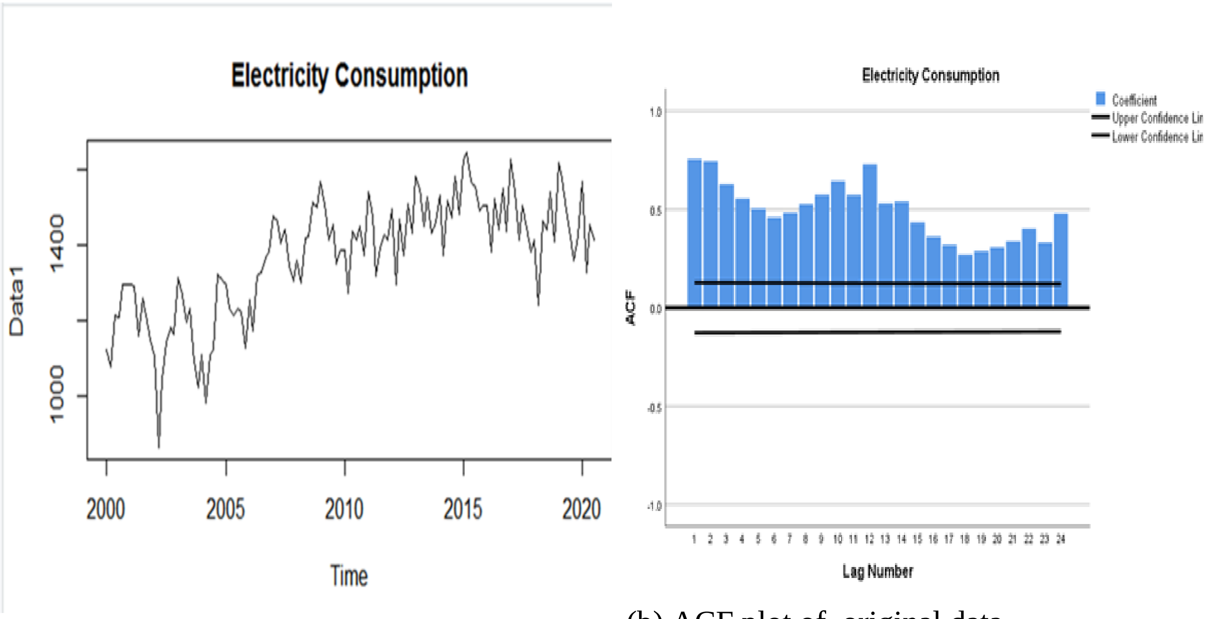
electricity consumed is highly negatively skewed. This implies that the median is greater than the mean. The value of the JB test is 36.188, with a probability of <0.0001. Therefore, one can conclude that the data is not normally distributed. Normality is not a necessary assumption for forecasting therefore forecasting can still be done. The next section presents the Box-Jenkins procedure.

4.3 THE RESULTS OF THE BOX-JENKINS TEST

This section outlines the four steps of the Box-Jenkins procedure followed in analysing the electricity consumption data in South Africa.

4.3.1 Model identification

When developing a model following the Box-Jenkins approach, checking for stationarity and periodic fluctuations is the first and most crucial step by plotting the observations against time. The visual examination for stationarity and other observed pattern can be assessed from the time plot as displayed in Figure 4.1.

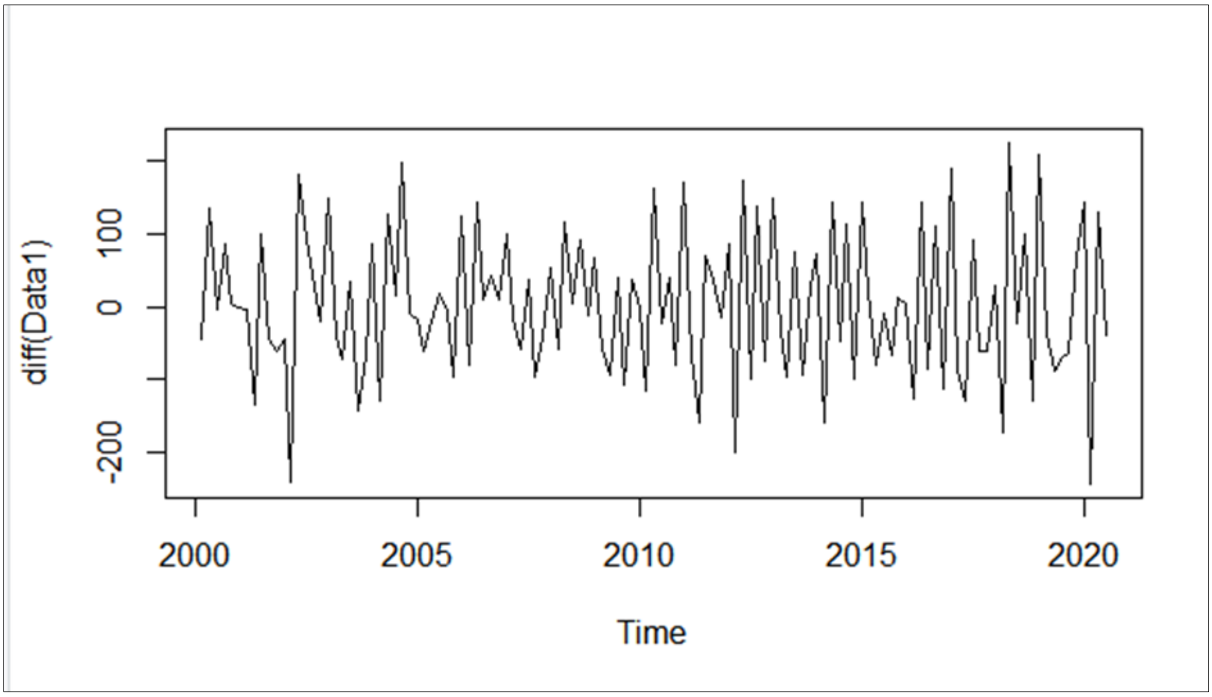


(a) Electricity consumption series (b) ACF plot of original data

Figure 4.1: **Original time series plot of electricity consumption in South Africa**

The results presented in Figure 4.1 (a) revealed that the series seems to show roughly similar trend increasing over time and also reveals a seasonality or cyclic component according to change of seasons. The autocorrelation (ACF) plots are also used to examine stationarity of the series. The ACF plot presented in Figure 4.1 (b) shows that the electricity consumption constantly decays and increases again, which shows irregular fluctuations. By visual inspection, one can conclude that the data is non-stationary. The first differenced plot is presented in Figure 4.2 below.

Figure 4.2: Plot of the first differenced data for electricity consumption in South Arica



The graphical analysis in Figure 4.2 shows the first differenced series fluctuating around the constant mean. The series shows no periodic fluctuations implying that it is stationary with constant variance and autocorrelation structure over time and mean zero. Based on visual inspection, this shows that the series is stationary at first difference (d=1). To confirm the visual inspection, the formal test of stationarity was computed. The results are presented in the following subsection 4.3.2

4.3.2 Unit root test

The unit root test was carried out to validate the visual inspection result as highlighted above. The ADF and PP are used to confirm the visual examination results. The results of the ADF and PP test are summarised in Table 4.2.

Table 4.2: Unit root test for Electricity Consumption in South Africa

	ADF		PP	
	Level	1st difference	Level	1st difference
Test statistic	-1.761	-6.648	-5.742	-40.425
Probability value	0.400	0.000	0.00	0.000

Note: The critical values for ADF and PP test at 1%, 5% and 10% are -3.459,-2.874 and -2.573 respectively.

Table 4.2 present the results of ADF and PP test. The ADF test shows that the series is non-stationary at level since the p-value of 0.400 is greater than 5% level of significance. On the contrary, the PP test revealed that the series is stationary at level. Since the ADF test revealed that the series is non-stationary at level, differencing was then applied. After the series was differenced the results of the ADF and PP test statistics show that the series is stationary. This implies that there is no unit root in the electricity data used; it is stationary after being differenced once. As such, the null hypothesis of the presence of unit root in the series is rejected. The differenced series was used to fit the ARIMA model in order to determine the AR and MA using ACF plots. The ACF plots are discussed in Section 4.3.3

4.3.3 *Autocorrelation plots*

Using the differenced data, the ACF and PACF are used in order to estimates MA and AR respectively. The plot of the ACF and PACF are presented in Figure 4.3.

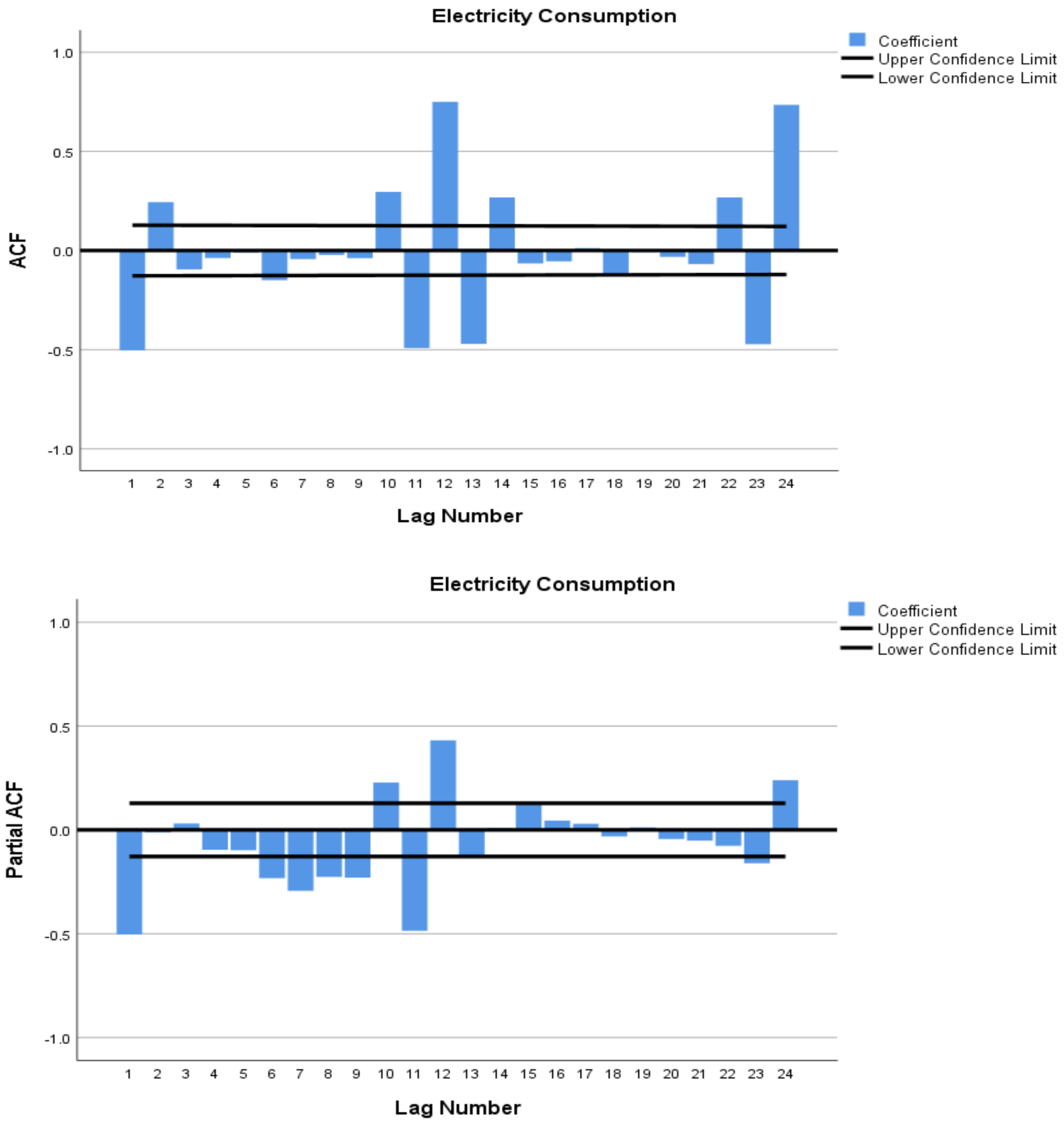


Figure 4.3: ACF and PACF of first differenced data

Figure 4.3 presents the ACF and PACF of the series. Looking at the ACF and PACF plot presented in Figure 4.3, electricity consumption series contains up to 24 lags maximum. The ACF plot spikes up until lag 2 and then oscillates downward, while after lag 1, PACF cuts off and dies down. This suggests that MA (2) and AR (1) models may be suitable for the data. Additionally, in the ACF, the spikes are observed at lag 12 and 24 and the PACF spikes are also observed at lag 11 and 24 as well. From the ACF and PACF plots, it is evident that the electricity consumption series exhibit a significant degree of seasonality, and differencing had no effect on it.

The seasonal differencing is presented in the following Section 4.3.3.1.

4.3.3.1 The ACF and PACF plot after first differencing and seasonal differencing

The ACF and PACF plot of the first differenced and seasonal differenced series is presented in Figure 4.4.

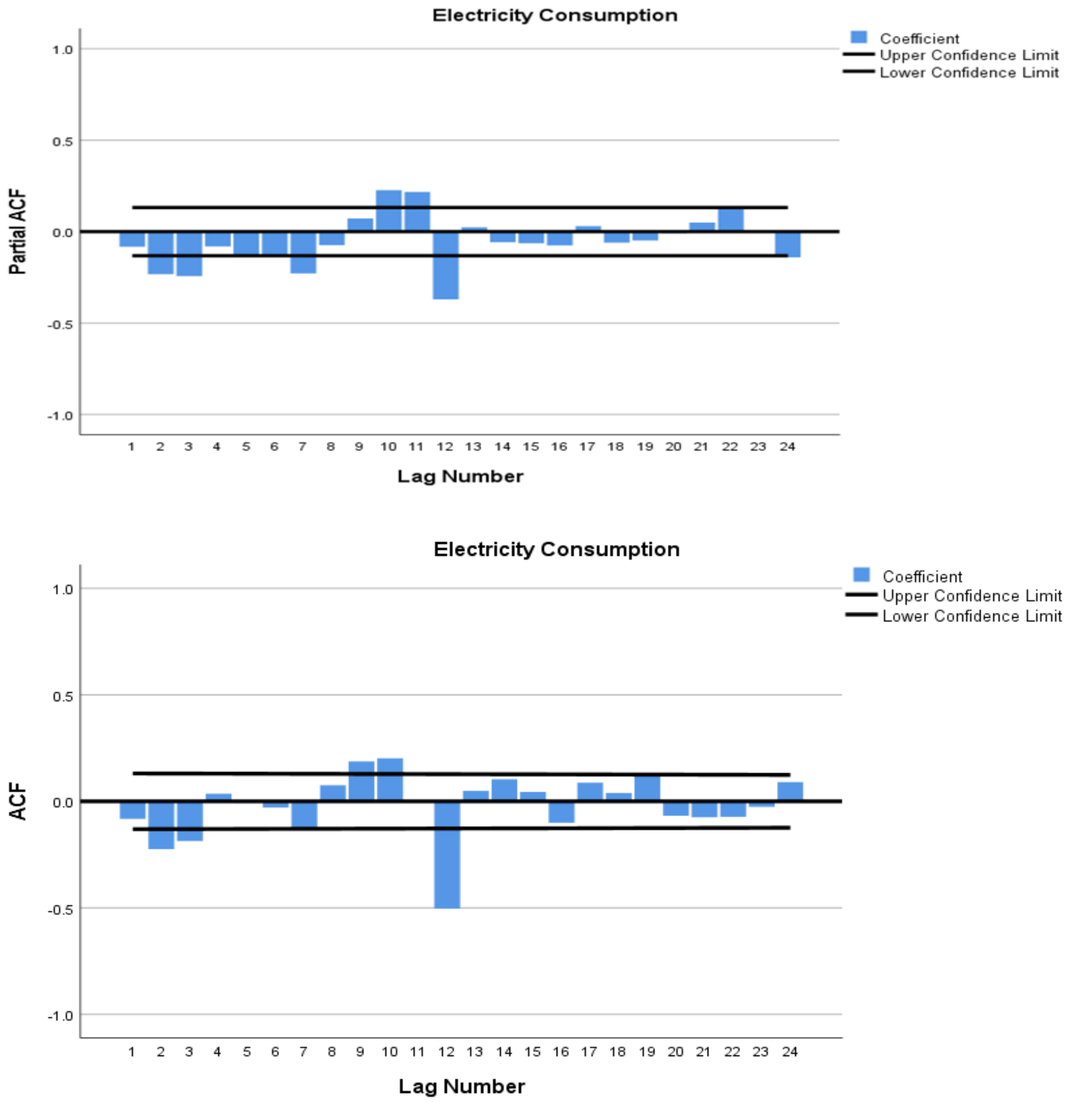


Figure 4.4: ACF and PACF plot after first differencing and seasonal differencing

Figure 4.4 presents the ACF and PACF plots after first differencing and seasonal differencing. Plot analysis reveals that both ACF and PACF quickly decline after lag 3, and again after lag 12, the ACF cuts off and the PACF dies down quickly, which indicates that seasonal differencing is appropriate. According to Bowerman *et al*. (2005), it is sometimes preferable to overlook spikes at high lags in order to simplify the identified model. The spike at lag 23 is ignored. Furthermore, seasonal moving average term ε_{t-12} as stated by Bowerman *et al*. (2005) should be included in the model. Since the ACF and PACF plots identified the non-seasonal data spikes at high process model ARIMA (1, 1, 2), therefore the SARIMA model is estimated as SARIMA (1, 1, 2) × $(0, 1, 1)_{12}$.The next subsection presents the model estimation and selection.

4.3.4 Estimation and selection

The non-seasonal ARIMA model (1, 1, 2) × $(0, 1, 1)_{12}$ and seasonal ARIMA (3, 1, 3) × $(0, 1, 1)_{12}$ are fitted as suggested by the ACF and PACF plots. The objective is to select the model with the minimum BIC. Therefore, the models needed to prove which is the best-fit model; as a result, several models have also been fitted and are compared as part of the selection process step as presented in Table 4.3.

Table 4.3: Comparison of the models

Model number	**Model**	**BIC**
1	**SARIMA(1, 1, 2)×$(1,1,1)_{12}$**	**8.206**
2	**SARIMA(1, 1, 2)×$(0,1,1)_{12}$**	**8.118**
3	**SARIMA(3, 1, 3)×$(1,1,1)_{12}$**	**8.131**
4	**SARIMA(3, 1, 3)×$(0,1,1)_{12}$**	**8.186**

Table 4.3 presents all the four models are appropriate for electricity consumption; however, the model with the smallest BIC is considered to be most appropriate. Model 2, which is SARIMA (1, 1, 2) $(0, 1, 1)_{12}$ has the smallest BIC value of 8.118. Therefore, it is concluded that model 2 is the best fit when compared to the other 3 models. The parameters of the selected model are presented and discussed in sub-section 4.3.4.1.

4.3.4.1 Parameter estimation of SARIMA (1, 1, 2) ×$(0, 1, 1)_{12}$

This section presents the parameter estimation of the SARIMA (1, 1, 2) $(0, 1, 1)_{12}$ model. The estimated parameters are summarised in Table 4.4.

Table 4.4: SARIMA (1, 1, 2) ×(0, 1, 1)$_{12}$ parameters estimations

		Estimates	SE	Statistic	P-value
Consumption- consumption _Model 1	AR 1	0.425	0.141	3.015	0.003
	MA1	-1.729	0.093	-18.587	<0.0001
	MA2	0.747	0.123	6.054	<0.0001
	SMA1	-0.674	0.085	-7.962	<0.0001

In Table 4.4, the parameter estimates are provided together with their corresponding t-statistics and p-values. It is concluded that only lag 1 is significant for the non-seasonal AR (*p*) process. The null hypothesis of the parameter is equal to the hypothesized value, which is rejected at 5% level of significance because the p-value is less than 0.05. Therefore, the non-seasonal AR(1) model is appropriate. Similarly, only lag 2 for non-seasonal MA (*q*) process is significantly different from 0, suggesting that MA (2) is suitable for the electricity consumption data. The seasonal SMA (*P*) progress is highly significant at lag 1, which implies that the SMA $(1)_{12}$ is accurately identified. That is, model parameters are all significant at 5% significance level as indicated by the probabilities.

From Table 4.4, parameter estimates are computed in the following general equation:

$$Z_{t=}\ \varepsilon_t - \varphi_{1Z_{t-1}} - \theta_1\varepsilon_{t-1} - \theta_{12}\varepsilon_{t-12} \quad (4.2)$$

as follow:

$$z\hat{}_t = -0.425_{Z_{t-1}} + 1.729\ \varepsilon_{t-1} - 0.747\ \varepsilon_{t-1} + 0.674\varepsilon_{t-12} + \varepsilon_t \quad (4.3)$$

The next stage is to perform the diagnostic tests to determine whether the chosen model is adequate to forecast South Africa's electricity consumption. The diagnostic tests are discussed in sub-section 4.3.5.

4.3.5 Diagnostic test of the SARIMA (1, 1, 2) (0, 1, 1)12 model

This section presents the diagnostic tests of the SARIMA (1, 1, 2) (0, 1, 1)$_{12}$. The diagnostic tests are presented in Figure 4.5.

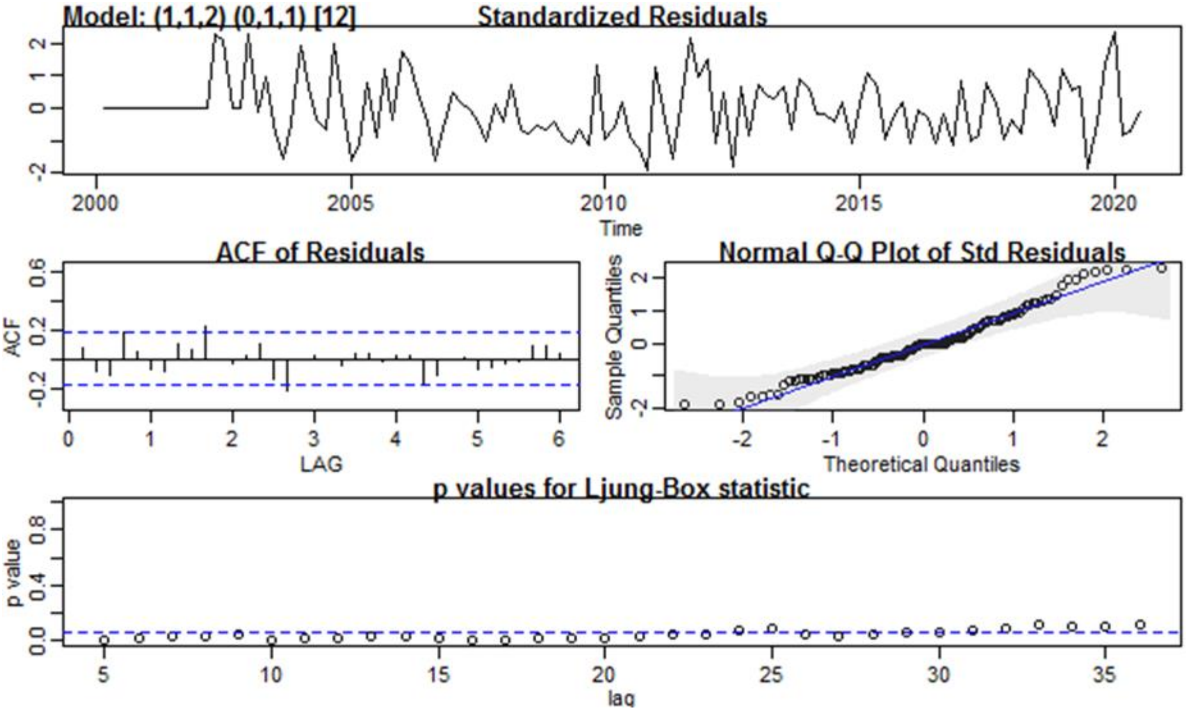


Figure 4.5: The graphical approach for checking for the randomness of residuals

Figure 4.5 presents the graphical approach for checking for the randomness of residuals of the best-fitted model. By eye inspection, the residuals of the ACF seem to be uncorrelated since almost all the spikes are within the boundaries. According to Bowerman *et al.* (2005), in order to streamline the identified model, it is sometimes best to ignore spikes at high lags. The normality plots also show that most of the points seems to fall about the straight line, which confirms that the model SARIMA (1, 1, 2) (0, 1, $1)_{12}$ is appropriate to be used for further analysis. Therefore, it can be concluded that the chosen model is adequate enough and suitable. The model will be used to forecast the future values and the results are presented in subsection 4.3.6.

4.3.6 Forecasts of the SARIMA (1, 1, 2) (0, 1, $1)_{12}$ model

The diagnostic checks have confirmed that the selected SARIMA (1, 1, 2) (0, 1, $1)_{12}$ model satisfies the underlying assumptions well. Therefore, the model will be used to predict values for the next 5 years, from May 2020 to December 2025. Figure 4. 6 is series plot for original values of electricity consumption in South Africa and the forecasted values.

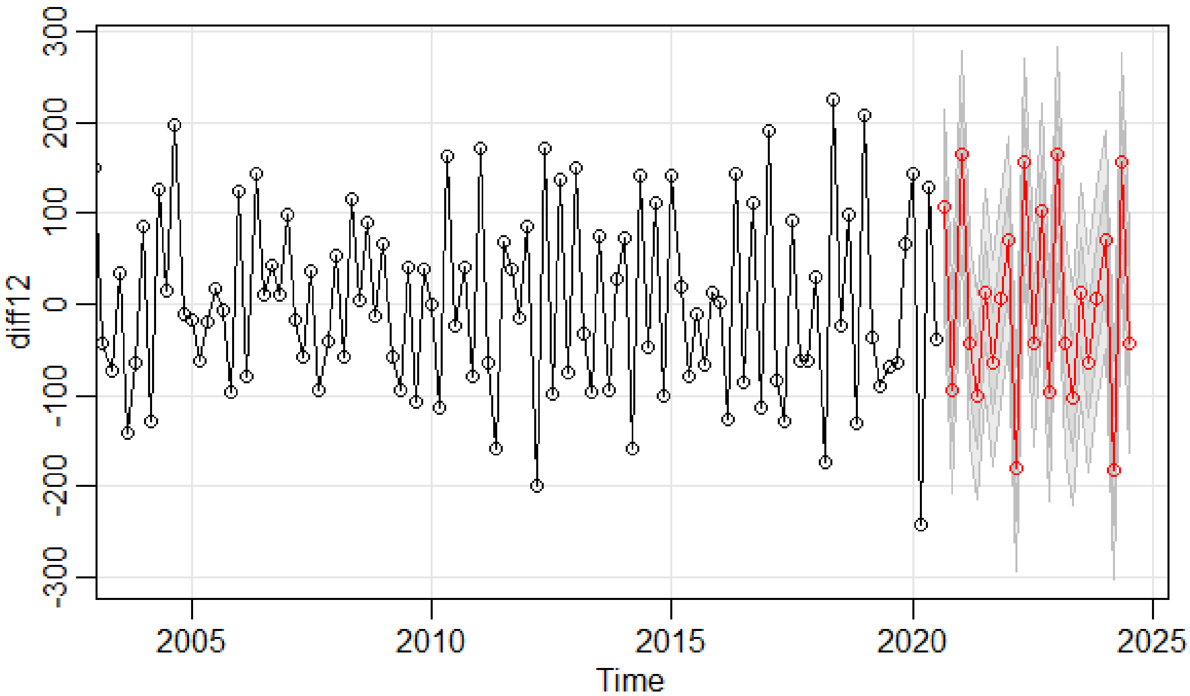


Figure 4.6: Series plot of original and forecasted electricity consumption values for the next five years

Figure 4.6 shows the forecasted plot of electricity consumption. After observing all assumption, the researcher drew the conclusion that seasonal fluctuations have an impact on electricity consumption. The study also used GARCH volatility model to model the electricity consumption. The results of the GARCH model is presented in Section 4.4.

4.4 AUTOREGRESSIVE CONDITIONAL HETEROSKEDASTIC (ARCH) MODEL

The estimated parameter of ARCH model is presented in the following Table 4.5.

Table 4.5: ARCH Model parameter estimates

Parameter	**Estimate**	**Std.Error**	**t-value**	**p-value**
μ	1361.717	339.130	34.800	0.000
AR(1)	0.788	0.055	14.300	0.000

Table 4.5 shows that the estimated parameter of the ARCH (1) effect is statistically significant with the low probability value of 0.000 when compared to 0.01 significance level. Since the

ARCH (1) is statistically significant, it implies that the mean equation can be fitted to the GARCH variance equation. The ARCH heteroscedasticity test for residuals were also computed to confirm the results in Table 4.5 and the results are presented in Table 4.6 below.

Table 4.6: ARCH heteroscedasticity test for residuals

	PQ	**P-value**	**ARCH-LM**	**P-value**
[1,]	1.42	0.841	26.906	6.16e-06
[2,]	4.02	0.856	10.452	8.72e-01
[3,]	29.23	0.004	6.021	1.00e+00
[4,]	29.58	0.020	-0.352	1.00e+00
[5,]	31.17	0.053	-0.361	1.00e+00
[6,]	41.31	0.015	-0.428	1.00e+00

In Table 4.6, the presence of heteroskedasticity was tested using Engle's Lagrange Multiplier (LM) test. The LM test strongly shows that there is existence of heteroskedasticity since the p-values are less than 0.05 level of significance. Furthermore, the ARCH effects were computed and the results are presented in Table 4.7 below.

Table 4.7: ARCH effect test for residual

Test	**Statistics**	**p-value**
Jarque-Bera	0.828	0.661
Box-Ljung	26.239	3.017e-07

Table 4.7 presents the ARCH effect for residuals. The Box-Ljung test shows that the ARCH effect exists in the residual series since the p-value of the test is less than the 5% level of significance. The results confirm the presence of heteroscedasticity; therefore GARCH model can be computed. The results of the GARCH model are presented in the following Section 4.5.

4.5 GENERALISED ARCH (GARCH) MODEL

The previous section revealed that there are ARCH errors in electricity consumption. The first step in computing the GARCH model is to determine the best distribution between the normal distribution, skewed student–t distribution, and student-t distribution. The information criteria were used to determine the best distribution. The results are presented in Table 4.8 below.

Table 4.8: Information criteria of the fitted conditional distributions of the GARCH (1, 1)

Information criterion	**Distribution**		
	Norm	**Sstd**	**Std**
Akaike	12.112	12.137	12.131
Bayes	12.203	12.274	12.245
Shibata	12.110	12.132	12.127
Hannan-Quinn	12.149	12.192	12.177

According to the results presented in Table 4.8, the normal distribution proved to be the best distribution since it has the smallest information criteria values. Therefore, the GARCH (1, 1) model was fitted using normal distribution. The estimated parameters of this model are presented in sub-section 4.5.1.

4.5.1 Parameter estimates of the GARCH (1, 1) model

This section of the study presents the parameter estimated of the GARCH (1, 1) model. The estimated parameters are summarised in Table 4.9 below.

Table 4.9: Estimated model parameter of the GARCH (1, 1)

Parameter	Estimate	Std.Error	t-value	P-value(>|t|)
μ	2.478	9.120	0.272	0.786
ω	23.602	579.979	0.041	0.968
α_1	0.000	0.058	0.000	1.000
β_1	0.999	0.008	131.040	0.000

Table 4. 9 presents the parameter estimates along with their t-ratios and significance level (p-value). The results revealed that β_1 is the only significant parameter since its p-value is less than 5% level of significance.

From Table 4.9, the general equation for the GARCH (1, 1) is formulated using only significant estimates and it is as follows:

$$\sigma_t{}^2 = \omega + \alpha\varepsilon_{t-1}{}^2 + \beta\varepsilon_{t-1}{}^2 \tag{4.5}$$

Using the significant estimates, the following models are deduced from Table 4.9 and the model equation for GARCH (1, 1) is written as follows:

$$\sigma_t{}^2 = 0.999\varepsilon_{t-1}{}^2 \tag{4.6}$$

where $\sigma_t{}^2$ symbolises the volatility in the GARCH (1, 1) model equation for electricity consumption in South Africa. The sum of the estimates $\hat{\alpha}_1$ and $\hat{\beta}_1$ is less than one. This implies that the unconditional volatility for series data is finite. The results further showed that electricity consumption time series data has the highest volatility persistence value of $\hat{\alpha}_1 + \hat{\beta}_1$ =0.999 (0.000 + 0.999). The conditional volatility plot is presented in Figure 4.7 below.

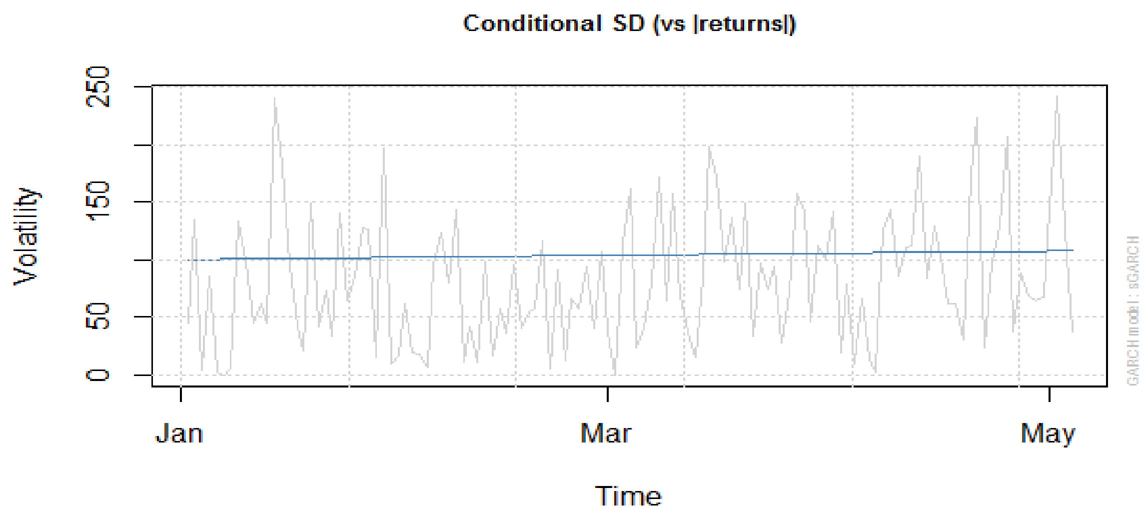


Figure 4.7: Conditional volatility of GARCH (1, 1)

The electricity consumption time series data shows the high volatility as depicted in Figure 4.7. The next subsection 4.5.2 presents the diagnostic tests.

4.5.2 *Diagnostic test of the GARCH (1, 1) model*

This section shows the diagnostic tests of the GARCH (1, 1) model. The diagnostic tests used to test the model adequacy are Ljung-Box(R), Ljung-Box(R-squared), ARCH-LM test and goodness of fit. The results of the Ljung-Box (R) and Ljung-Box (R-squared) test are presented in Table 4.10 below.

Table 4.10: Heteroscedasticity test of residuals for GARCH (1, 1) model

Ljung-Box test (R)			**Ljung-Box test (R-squared)**		
Lag	**statistic**	**p-value**	**Lag**	**Statistic**	**p-value**
[1]	29.430	5.798e-08	[1]	1.643	0.200
[2*(p+q)+(p+q)-1][2]	32.690	2.096e-09	[2*(p+q)+(p+q)-1][5]	4.099	0.242
[2*(p+q)+(p+q)-1][5]	35.400	9.413e-10	[2*(p+q)+(p +q)-1][9]	5.387	0.375

Table 4.10 presents the Ljung-Box test for white noise behaviour in residuals and to test the adequacy of the fitted model. The results shows that in the squared residuals, the model does not suffer from serial correlations since all the p-values are greater than 0.05 significance level, The Box-Ljung test is not rejected and concludes that the fitted GARCH (1, 1) model behaves as white noise, which implies that the model is adequate. Furthermore, the ARCH effects were computed and the results are presented in Table 4.11.

Table 4.11: ARCH effect of the GARCH (1, 1) model

ARCH LM tests	**Statistic**	**Shape**	**Scale**	**P-value**
ARCH lag [3]	2.208	0.500	2.000	0.137
ARCH lag [5]	3.017	1.440	1.667	0.287
ARCH lag [7]	3.389	2.315	1.543	0.443

Table 4.11 present the ARCH effect in the fitted model. The results shows the absence of ARCH errors since all the p-values are greater than 5% significance level. Therefore, the fitted GARCH (1, 1) model is a good fit. The goodness-of-fit test was also computed to confirm the results. The results of goodness-of-fit test are presented in Table 4.12.

Table 4.12: Goodness-of -fit test of the GARCH (1, 1) model

Adjusted Pearson Goodness-of-fit Test			
Obs.	**Group**	**Statistic**	**P-value(g-1)**
1	20	22.20	0.274
2	30	23.10	0.772
3	40	45.13	0.231
4	50	53.83	0.295

Table 4.12 present the results for goodness-of-fit test. The results revealed that all the p-values are greater than 0.05 significance level, which implies that the normal distribution assumption cannot be rejected. Since the fitted residual are normally distributed, the GARCH (1, 1) model appears to fit the normal distribution well and can be used for further analysis of electricity consumption. The next subsection, 4.5.3, presents the forecast of the GARCH (1, 1).

4.5.3 Forecasts of the GARCH (1, 1) model

The forecast of the GARCH (1, 1) of ten periods ahead is presented in the following Table 4.13.

Table 4.13: Forecasted values of the GARCH (1, 1) model

Period	**Mean forecast**	**Mean error**	**Std.Dev.**	**Lower interval**	**Upper interval**
1-May-2020	4.93844	117.2574	117.2574	-196.3796	265.4324
2-June-2020	4.93844	117.3851	117.3851	-196.5987	265.7160
3-July-2020	4.93844	117.5126	117.5126	-196.8177	265.9993
4-Aug-2020	4.93844	117.6399	117.6399	-197.0363	266.2823
5-Sep-2020	4.93844	117.7672	117.7672	-197.2548	266.5649
6-Oct-2020	4.93844	117.8943	117.8943	-197.4730	266.8473
7-Nov-2020	4.93844	118.0212	118.0212	-197.6910	267.1293
8-Dec-2020	4.93844	118.1481	118.1481	--197.9087	267.4111
9-Jan-2021	4.93844	118.2748	118.2748	-198.1262	267.6925
10-Feb-2021	4.93844	118.4013	118.4013	-198.3435	267.9737

Table 4.13 summarises the forecasts obtained from the GARCH (1, 1) model. The forecasted values fall within the 95% confidence interval. Figure 4.10 presents the series plot of the fitted mean and volatility forecast values of electricity consumption in South Africa.

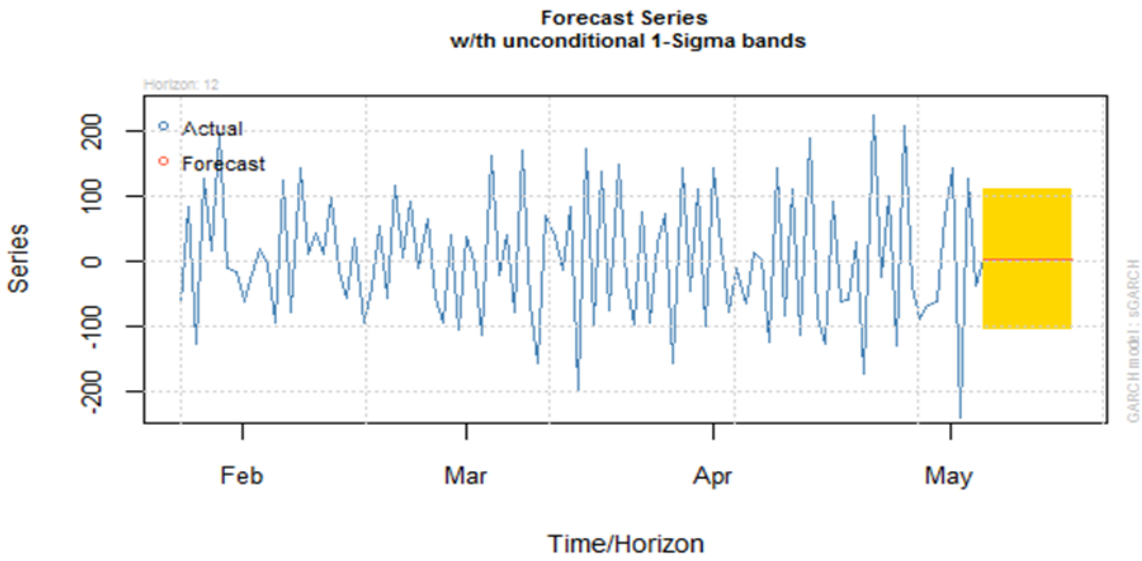


Figure 4.8: The volatility forecast plot obtained from GARCH (1, 1) model

The volatility forecasts presented in Figure 4.8 confirms that the volatility forecasts are within the 95% confidence limit. The volatility forecasts plot for electricity consumption in South Africa seems to be constant for the next 10 periods ahead.

The next section, 4.6, presents the results of the hybrid SARIMA-GARCH model.

4.6 HYBRID SARIMA (1, 1, 2) (0, 1, 1)$_{12}$- GARCH (1, 1) MODEL

It is observed that the SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ model and the GARCH (1, 1) models perform better in modelling the electricity consumption. The study fitted residuals of SARIMA model into GARCH (1, 1) to make a hybrid of a SARIMA-GARCH model. The study intended to determine whether the hybrid model could produce similar or better results than the SARIMA or GARCH models alone. Figure 4.9 present the ACF and PACF plotting of electricity consumption in South Africa for the hybrid SARIMA-GARCH model.

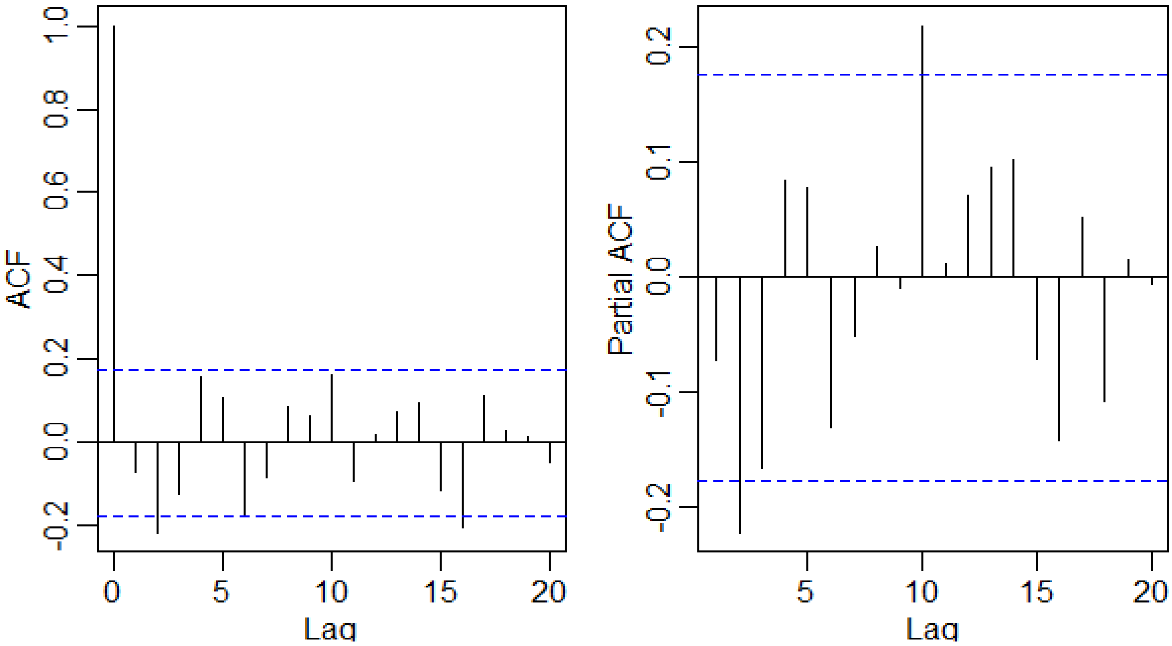


Figure 4.9: Autocorrelation check of residuals of the SARIMA (1, 1, 2) (0, 1, 1)$_{12}$-GARCH (1, 1) model

Figure 4.9 illustrate the ACF plots and PACF for residuals of fitted hybrid SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ $-$ GARCH $(1, 1)$ model. The plots illustrate the presence of ARCH errors in the electricity consumption series because of the taper. The plots also show the presence of serial correlation although the visual examination will be confirmed with the relevant test.

The next step is to determine the best distribution to be used in modelling the SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ $-$ GARCH $(1, 1)$ mode. The results are presented in Table 4.14.

Table 4.14: Information criteria of the fitted conditional distributions of the SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ $-$ GARCH $(1, 1)$ model

Information criterion	**Distribution**		
	Norm	**Sstd**	**Std**
Akaike	10.801	10.789	10.818
Bayes	10.893	10.926	10.932
Shibata	10.799	10.785	10.815

Information criterion	Distribution		
	Norm	**Sstd**	**Std**
Hannan-Quinn	10.839	10.845	10.864

According to the results presented in Table 4.14, normal distribution has the lowest BIC value and HQIC value. The skewed student–t distribution has the lowest AIC value of 10.789 and SIC value of 10.785. Since the AIC and SIC were found to be the lowest values as compared to other values presented in the table, the SARIMA (1, 1, 2) (0, 1, $1)_{12}$- GARCH (1, 1) model was fitted using skewed student–t distribution. The estimated parameters of this model are presented in sub-section 4.6.1.

4.6.1 Parameter estimates of the SARIMA (1, 1, 2) (0, 1, $1)_{12}$- GARCH (1, 1) model

This section of the study presents the parameter estimated of the SARIMA (1, 1, 2) (0, 1, $1)_{12}$ - GARCH (1, 1) model. The estimated parameters are summarised in Table 4.15.

Table 4.15: Estimated parameters for SARIMA (1, 1, 2) (0, 1, $1)_{12}$- GARCH (1, 1) model

Parameter	**Estimate**	**Std.Error**	**t-value**	**P-value(>|t|)**
μ	1.406	4.920	0.286	0.775
ω	10.622	78.479	0.135	0.892
α_1	0.000	0.027	0.000	1.000
β_1	0.997	0.009	108.038	0.000
x_i	1.371	0.191	7.192	0.000
Nu	59.993	237.604	0.252	0.801

Table 4.15 shows the parameter estimates together with their t-ratios and their significance level (p-value). The results revealed that skewed student–t distribution is skewed to the right and significant with the skew value of 1.371 and p-value of 0.000 respectively. The results also revealed that β_1 is also a significant parameter since its p-value is less than 5% level of significance.

From Table 4.15, only significant estimates are used to formulate a general equation for the SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ $-$ GARCH (1, 1) model and it is as follows:

$$\sigma_t{}^2 = \omega + \alpha\varepsilon_{t-1}{}^2 + \beta\varepsilon_{t-1}{}^2 + x_i\, \varepsilon_{t-1}{}^2 + nu\, \varepsilon_{t-1}{}^2 \, . \qquad (4.7)$$

Using the significant estimates, the following models are deduced from Table 4.14 and the model equation for SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ $-$ GARCH (1, 1) model is written as follows:

$$\sigma_t{}^2 = 0.997\varepsilon_{t-1}{}^2 + 1.371\varepsilon_{t-1}{}^2 \, . \qquad (4.8)$$

The sum of the estimates $\hat{\alpha}_1$ and $\hat{\beta}_1$ are less than one. This implies that the unconditional volatility for series data is finite. The results further showed that electricity consumption time series data has the highest volatility persistence value of $\hat{\alpha}_1 + \hat{\beta}_1$ =0.997 (0.000 + 0.997). The conditional volatility is presented in Figure 4.10.

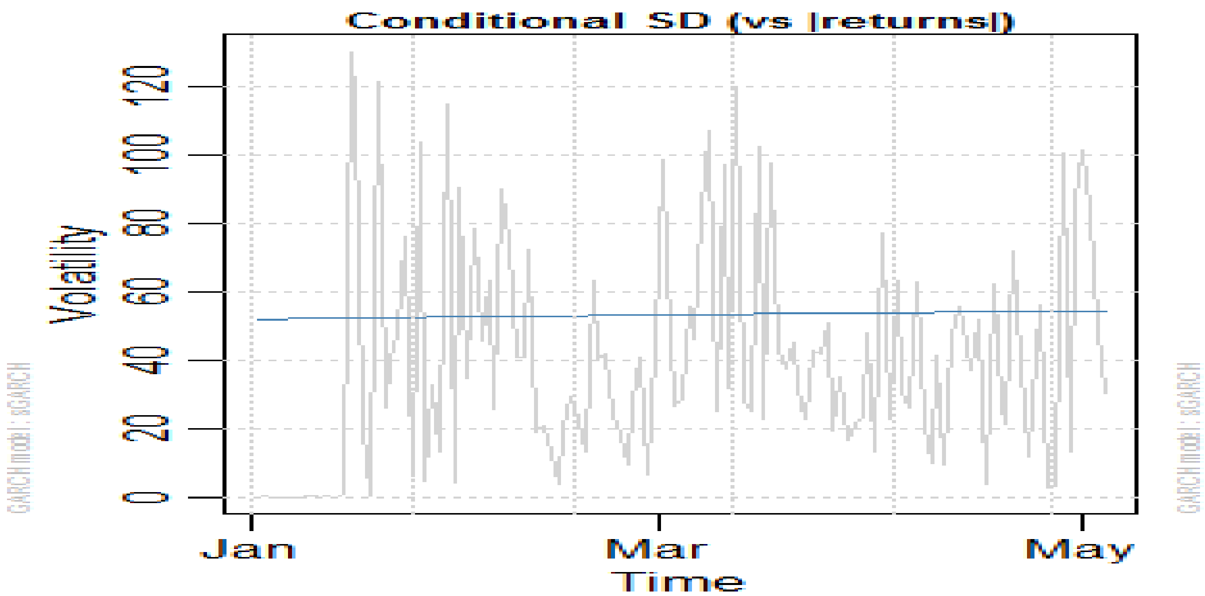


Figure 4.10: Conditional volatility of SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ $-$ GARCH (1, 1) mode

The electricity consumption time series data shows the high volatility as depicted in Figure 4.10. The next subsection, 4.6.2, illustrates the diagnostic tests.

4.6.2 Diagnostic tests of the SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ − GARCH (1,1) model

This section presents the diagnostic tests of the SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ − GARCH (1, 1) model. The results of the Ljung-Box (R) test and Ljung-Box(R-squared) are presented in the following Table 4.16.

Table 4.16: Heteroscedasticity test of residuals for SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ − GARCH (1,1) model

Ljung-Box test (R)			**Ljung-Box test (R-squared)**		
Lag	**statistic**	**p-value**	**Lag**	**statistic**	**p-value**
[1]	29.380	5.963e-08	[1]	1.626	0.202
[2*(p+q)+(p+q)-1][2]	32.630	2.173e-09	[2*(p+q)+(p+q)-1][5]	4.096	0.242
[2*(p+q)+(p+q)-1][5]	35.330	9.812e-10	[2*(p+q)+(p +q)-1][9]	5.383	0.375

Table 4.16 presents the Ljung-Box test of residuals of the hybrid model. The results of the R^2 shows that there is no evidence of autocorrelation in the residuals. This implies that the residuals behave as a white noise process. Since the null hypothesis of Box-Ljung test is accepted, it concludes that the fitted hybrid SARIMA- GARCH model does not shows any lack of fit. The ARCH effects were also computed, and the results are presented in Table 4.17.

Table 4.17: ARCH effect of the hybrid SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ − GARCH (1,1) model

ARCH LM tests	**Statistic**	**Shape**	**Scale**	**p-value**
ARCH lag [3]	2.237	0.500	2.000	0.135
ARCH lag [5]	3.042	1.440	1.667	0.284
ARCH lag [7]	3.412	2.315	1.543	0.439

Table 4.17 presents the ARCH effect in the fitted hybrid model. The results presented in Table 4.16 revealed that all the p-values are greater than 5% significance level, which implies that there are no ARCH errors. Therefore, it is evident that there is no serial correlation in squared residuals for the fitted SARIMA (1, 1, 2) (0, 1, $1)_{12} -$ GARCH $(1, 1)$ model. The goodness-of-fit tests were also computed to confirm the results, and the results are summarised in Table 4.18.

Table 4.18: Goodness-of-fit test for the SARIMA (1, 1, 2) (0, 1, $1)_{12} -$ GARCH $(1, 1)$model

Adjusted Pearson Goodness-of-fit Test			
Obs	**Group**	**Statistic**	**p-value(g-1)**
1	20	20.58	0.361
2	30	25.05	0.676
3	40	42.53	0.322
4	50	53.02	0.322

The results of the goodness-of-fit test are presented in Table 4.18. It is evident from the results presented in Table 4.17 that all the p-values are greater than 0.05 significance level. Therefore, the fitted residuals of the model are normally distributed. The SARIMA (1, 1, 2) (0, 1, $1)_{12} -$ GARCH $(1, 1)$ model under the skewed student–t distribution appears to be adequate. Since the hybrid model has passed all the diagnostic tests and appears to be a good fit, the model can be used to forecast the future values of electricity consumption. The next subsection, 4.6.3, presents the forecast of the SARIMA (1, 1, 2) (0, 1, $1)_{12} -$ GARCH $(1, 1)$ model.

4.6.3 Forecast of the SARIMA (1, 1, 2) (0, 1, $1)_{12} - GARCH$ $(1, 1)$ model

Figure 4.11 presents the series plot of the fitted mean and volatility forecast values of electricity consumption in South Africa.

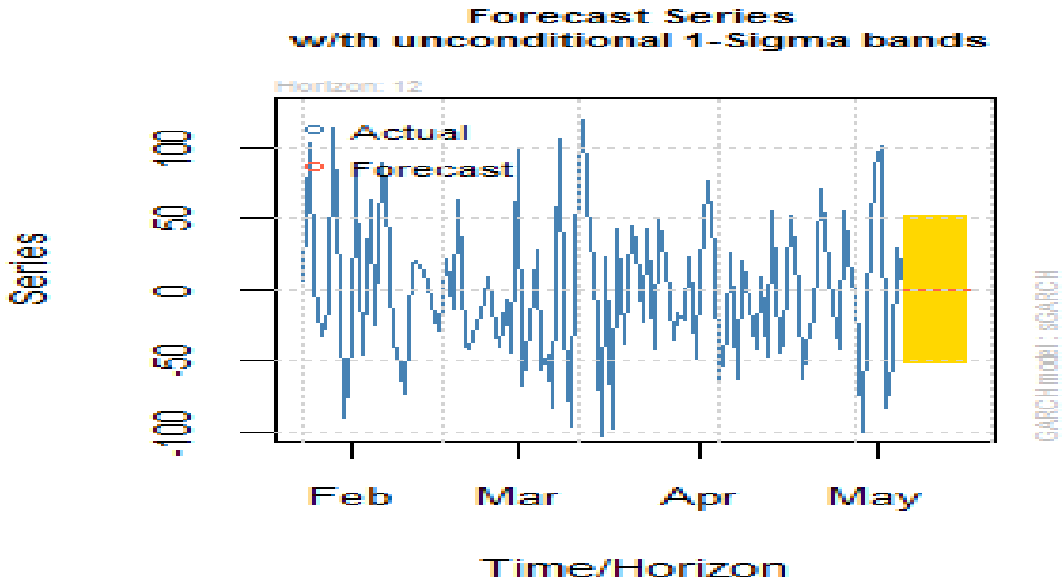


Figure 4.11: The volatility forecast plot obtained from Hybrid SARIMA (1, 1, 2) (0, 1, 1)$_{12}$ − GARCH (1, 1) model

Figure 4.11 illustrates the volatility forecast plot of the monthly data of electricity consumption for the next 5 years, and the data seem to be constant. The plot also shows that the volatility forecasts are within the 95% confidence limit. The next section, 4.7, is the summary of the chapter.

4.7 CHAPTER SUMMARY

The chapter presented the analysis of the monthly time series data for electricity consumption in South Africa. The three models used to analyse the electricity consumption are the SARIMA, GARCH, and hybrid of the SARIMA-GARCH. The ADF and PP tests were used to determine whether the data is stationary or not. The information criteria were used to determine the best distribution to be used when modelling the GARCH and the SARIMA-GARCH models. The discussion of results, conclusions and recommendations are presented next in Chapter 5.

CHAPTER 5

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

5.1 INTRODUCTION

This chapter presents a summary of key findings of the study based on the data analysis and interpretation of results obtained in chapter 4. The discussions as well as the conclusions are drawn in relation to the objectives of the study, and recommendations are proposed to give direction on the areas that need further research to be conducted.

The rest of the chapter is presented as follows; 5.2 presents key findings of the study based on each objective; in 5.3, conclusions are drawn, in 5.4 recommendations are presented and suggestions for future research. Finally, 5.5 presents a chapter summary of the dissertation

5.2 KEY FINDINGS OF THE STUDY

The main objective of the study was to model the South African electricity consumption data and use the best-fitted volatility forecasting model to predict future values. The volatility forecasting models employed in the study were SARIMA model and GARCH model, and residuals for the SARIMA model were fitted into the GARCH model to make a hybrid of SARIMA-GARCH model. The objective was successfully achieved to determine the best-performing model and the model was used to forecast electricity consumption in South Africa.

The descriptive statistics showed that the electricity consumed in South Africa ranges between 865.00 and 1739.000 Gigawatt-hour, with the mean of 1436.508 and the standard deviation of 147.613. The results showed that the median is greater than the mean, which implies that electricity consumed is negatively skewed. It was concluded that electricity consumption is leptokurtic, since sample kurtosis for electricity consumed was above three. This implies that the data were not distributed normally and had a heavier tail. This means that the electricity consumption is not normally distributed.

The visual examination revealed that the electricity consumption series showed roughly similar trends increasing over time and also revealed a seasonality component according to the change of seasons. The ACF plot showed irregular fluctuations in the electricity consumption data.

This implies that the original series was not stationary at level but stationary at first difference. The ADF and PP unit root tests were employed to confirm the visual examination. The ADF and PP test proved that the electricity consumption series is stationary at first difference.

5.2.1. The results from determining the seasonal effect on electricity consumption.

The model was estimated using MA and AR observed from ACF and PACF respectively. The ACF plot had spikes and then declines in an oscillating way, while the PACF cuts off and dies quickly. The suitable models for the data were indicated. Furthermore, the ACF and PACF plots illustrated a strong presence of seasonality on electricity consumption series that was not affected by differencing, which led to seasonal differencing performed on the differenced data.

5.2.2.The results from modelling SARIMA using electricity consumption in South Africa

The results from the SARIMA model revealed that there were four competing SARIMA models. The four models were SARIMA (3, 1, 3)(1, 1, $1)_{12}$, SARIMA $(1, 1, 2)$ $(1, 1, 1)_{12}$, SARIMA (1, 1, 2) (0, 1, $1)_{12}$ and SARIMA (3, 1, 3)×(0, 1, $1)_{12}$ as suggested by ACF and PACF plots. The four models were fitted and estimated to determine the best model. According to the BIC, SARIMA (1, 1, 2)(0, 1, $1)_{12}$ model was found to be the best performing model amongst the four. The parameters of the SARIMA (1, 1, 2)(0, 1, $1)_{12}$ model were found to be statistically significant at 5% level of significance. Diagnostic tests were also performed and confirmed that the SARIMA (1, 1, 2)×(0, 1, $1)_{12}$ model is adequate to predict South Africa's electricity consumption. The model was used to forecast the future values of electricity consumption in South Africa for the next 5 years. The results are supported by the studies by Etuk, Uchendu and Victoredema (2012), Etuk *et al.* (2014), and Boran (2014).

5.2.3. The results from modelling ARCH and GARCH models

The results of the study showed that the estimated parameters of the ARCH (1) model were statistically significant. This implies that the mean equation can be fitted to the GARCH variance equation. The results obtained from the LM test and the Box-Ljung test further revealed that there is existence of heteroscedasticity in the residuals of the ARCH model. Therefore, the GARCH model was then computed using the normal distribution. The parameter estimation of the fitted GARCH (1, 1) model showed that β_1 is the only significant parameter, with a p-value that is statistically significant at 0.05 level of significance. The results of the GARCH (1, 1) model also revealed that the unconditional volatility for series data is found to be finite. Furthermore, the results revealed that the electricity consumption series has the highest volatility persistence value. The diagnostic tests revealed that the GARCH (1, 1) model is adequate and was used for further analysis. The results of the study are in line with the study by Francq and Sucarrat (2018).

5.2.4.The results from modelling hybrid of SARIMA-SARIMA-GARCH

It was observed that the SARIMA (1, 1, 2)(0, 1, 1)$_{12}$ model and the GARCH (1, 1) models are the best models for modelling the electricity consumption. The study then fitted residuals of SARIMA model into GARCH (1, 1) to make a hybrid of SARIMA-GARCH model. The study wanted to determine whether the hybrid model could produce similar or better results. The SARIMA (1, 1, 2)(0, 1, 1)$_{12}$-GARCH (1, 1) model was fitted using skewed student–t distribution, which proved to be the best distribution with the smallest AIC value. The parameter estimation of the hybrid model revealed that β_1 was the only significant parameter. The results are similar to the results obtained from the GARCH (1, 1) model. Similar to the GARCH (1, 1) model, the results of the hybrid model showed that electricity consumption series has the highest volatility persistence value, and unconditional volatility for the series is finite. The diagnostic tests revealed that the SARIMA (1, 1, 2)(0, 1, 1)$_{12}$-GARCH (1, 1) model is adequate and was used for further analysis. The volatility forecasts plot of the monthly data of electricity consumption was illustrated and the plot showed that the mean forecasts fall within the 95% confidence limit. The findings are supported by the study by Nkwatoh (2012) and Sigauke and Chikibvu (2011).

5.3 DISCUSSION

The study modelled South African electricity consumption using the SARIMA model, GARCH model and a hybrid of the SARIMA-GARCH model to determine the best performing model. The study used monthly time series data for electricity consumption in South Africa obtained from Quantec Easy Data. The SARIMA (1, 1, 2)(0, 1, 1)$_{12}$ model was found to be adequate to model South African electricity consumption. Looking at literature studies by Etuk *et al.* (2012), Osarumwense (2013), Mao *et al*. (2018), and Mwanga *et al.* (2017) also confirmed that SARIMA models are good in time series with seasonal patterns. However, the series was found to have the presence of an ARCH effect, which led to the modelling of a GARCH model. The GARCH (1, 1) was also modelled and confirmed to be a good fit. Furthermore, since both the SARIMA (1, 1, 2) $(0, 1, 1)_{12}$ and the GARCH (1,1) were confirmed to be adequate fit, the study further fitted the residual of the SARIMA model into the GARCH model to make it a hybrid of SARIMA (1, 1, 2)$(0, 1, 1)_{12}$-GARCH (1,1). Previous studies by Tendai and Chikobvu (2017), Sigauke *et al.* (2011), Dritsakis *et al.* (2018), and Molebatsi and Raboloko (2016) revealed that hybrid models provide better predictions. The study by Molebatsi and Raboloko (2016) confirmed that the hybrid models can improve results. Therefore, the current study looked at application of hybrid models and confirmed that hybrid model gives the best fit and a more accurate forecast, unlike when each is modelled individually (Christodoulos *et al*., 2010). Therefore, the study successfully modelled monthly electricity consumption data and predicted future values using hybrid SARIMA (1, 1, 2)$(0, 1, 1)_{12}$-GARCH (1,1)

5.4 SUGGESTIONS FOR FUTHER STUDY

The study recommends the following areas for future research:

- Although the three volatility forecasting models proved to produce good results, there can still be more hybrid models in the forecasting process that can be developed to perfectly suit the data, more especially a hybrid with other GARCH families. The study therefore for future inquiry suggests the usage of the hybrid models such as SARIMA-TARCH (1,1), SARIMA-EGARCH, SARIMA-PARCH and SARIMA-AGARCH for South African electricity consumption.
- The current study modelled SARIMA and fitted residuals of SARIMA into GARCH model to make a hybrid SARIMA-GARCH model, and it was the best model for predicting electricity consumption in South Africa. Similar future studies can be conducted using the residuals of the GARCH model to form the hybrid of SARIMA-GARCH model. The results of the study can be compared with the results obtained from this study to determine the best hybrid model.
- The current study forecasts are accurate with minimum error and can therefore be vital to decision- and policy-makers in the energy industry.

5.5 SUMMARY OF THE DISSERTATION

The study was arranged into 5 chapters in order to address the research objectives as mentioned in chapter 1. Chapter 1 of the study introduced the research topic and provided some background of the study. It further established the problem statement, aim and objectives of the study and the research questions. Furthermore, the significance of the study, limitations of the study, and definition of terms were outlined. Literature based on the previous studies related to the current study was thoroughly reviewed in Chapter 2. In Chapter 3, the research methodology was presented. The data analysis and interpretation of results were presented in Chapter 4 of the study. Chapter 5 discussed and summarized the key findings of the study; furthermore, conclusions were drawn in line with the research objectives, and lastly recommendations were suggested as well as areas that need further research in the future.

REFERENCES

Adair-Rohani, H., Zukor, K., Bonjour, S., Wilburn, S., Kuesel, A.C., Hebert, R. and Fletcher, E.R. 2013. Limited electricity access in health facilities of sub-Saharan Africa: a systematic review of data on electricity access, sources, and reliability. *Global Health: Science and Practice*, 1(2), pp.249-261.

Adhikari, R. and Agrawal, R.K., 2013. An introductory study on time series modelling and forecasting. *arXiv preprint arXiv:1302.6613*.

Adhikari, S. 2013. *Structural dynamic analysis with generalised damping models: analysis*. John Wiley & Sons.

Alberg, D., Shalit, H. and Yosef, R. 2008. Estimating stock market volatility using asymmetric GARCH models. *Applied Financial Economics, 18*(15), pp.1201-1208.

Ampaw, E.M., Akuffo, B., Larbi, S.O. and Lartey, S. 2013. Time series modelling of rainfall in new Juaben municipality of the eastern region of Ghana. *International Journal of Business and Social Science*, *4*(8).

Asgharian, H., Christiansen, C. and Hou, A.J. 2016. Macro-finance determinants of the long-run stock–bond correlation: The DCC-MIDAS specification. *Journal of Financial Econometrics*, *14*(3), pp.617-642.

Awartani, B.M. and Corradi, V. 2005. Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries. *International Journal of Forecasting*, *21*(1), pp.167-183.

Baruník, J., Kočenda, E. and Vacha, L. 2014. Wavelet-based correlation analysis of the key traded assets. In *Wavelet Applications in Economics and Finance* (pp. 157-183). Springer, Cham.

Beveridge, S. and Oickle, C. 1994. A Comparison of Box-Jenkins and objective methods for determining the order of a non-seasonal ARMA Model. *Journal of Forecasting*, *13*(5), pp.419-434.

Bollerslev, T., 1986. Generalised autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), pp.307-327.

Bopulas, B. 2011. *Forecasting the Gross Domestic Product (GDP) of Malaysia* (Doctoral dissertation, Universiti Malaysia Sarawak).

Boran, K. 2014. The Box-Jenkins approach to forecast net electricity consumption in Turkey. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, *36*(5), pp.515-524.

Boshnakov, G.N., 2011. On first and second order stationarity of random coefficient models. *Linear Algebra and its Applications*, *434*(2), pp.415-423.

Bowerman, B.L., O'Connell, R.T. and Koehler, A.B., 2005. *Forecasting, time series, and regression: an applied approach* (Vol. 4). South-Western Pub.

Box, J. 1994. Reinsel, editor. Time Series Analysis, Forecasting and Control. Englewood Clifs.

Bulmer, M.G. 1979. *Principles of Statistics*. Courier Corporation.

Butgereit, L. 2015, September. An algorithm for measuring relative anger at Eskom during load-shedding using Twitter. In *AFRICON 2015* (pp. 1-5). IEEE.

Caner, M. and Kilian, L. 2001. Size distortions of tests of the null hypothesis of stationarity: evidence and implications for the PPP debate. *Journal of International Money and Finance*, *20*(5), pp.639-657.

Chakhchoukh, Y., Panciatici, P. and Mili, L. 2010. Electric load forecasting based on statistical robust methods. *IEEE Transactions on Power Systems*, *26*(3), pp.982-991.

Chen, H., Zhang, J., Tao, Y. and Tan, F. 2019. Asymmetric GARCH type models for asymmetric volatility characteristics analysis and wind power forecasting. *Protection and Control of Modern Power Systems*, *4*(1), pp.1-11.

Christodoulos, C., Michalakelis, C. and Varoutas, D. 2010. Forecasting with limited data: Combining ARIMA and diffusion models. *Technological Forecasting and Social change*, *77*(4), pp.558-565.

Cifter, A. 2012. Volatility forecasting with asymmetric normal mixture GARCH model: evidence from South Africa.

Conejo, A.J., Carrión, M. and Morales, J.M. 2010. *Decision making under uncertainty in electricity markets* (Vol. 1). New York: Springer.

Cox, D.R. and Miller, H.D. 1965. Stochastic processes. *Methuen & Co., London*.

Dahlvid, C. and Granberg, P. 2017. The Leverage Effect – Uncovering the true nature of US asymmetric volatility.

Dergiades, T. and Tsoulfidis, L. 2008. Estimating residential demand for electricity in the United States, 1965–2006. *Energy Economics*, 30(5), pp.2722-2730.

Dickey, D.A. and Fuller, W.A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a), pp.427-431

Dritsakis, N. and Klazoglou, P. 2018. Forecasting unemployment rates in USA using Box-Jenkins methodology. *International Journal of Economics and Financial Issues*, *8*(1), p.9.

Dürre, A., Fried, R. and Liboschik, T. 2015. Robust estimation of (partial) autocorrelation. Wiley Interdisciplinary Reviews: *Computational Statistics*, 7(3), pp.205-222.

Easydata. https://www.easydata.co.za/dataset/P4141/timeseries/P4141-ELEKTR22/

Engle, R.F. and Bollerslev, T. 1986. Modelling the persistence of conditional variances. *Econometric Reviews*, *5*(1), pp.1-50.

Engle, R.F. and Patton, A.J. 2007. What good is a volatility model?. In *Forecasting Volatility in the Financial Markets* (pp. 47-63). Butterworth-Heinemann.

Engle, R.F. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, pp.987-1007.

Engle, R.F., Lilien, D.M. and Robins, R.P. 1987. Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica: Journal of the Econometric Society*, pp.391-407.

Eni, D. and Adesola, A.W. 2013. Sarima modelling of passenger flow at Cross Line Limited, Nigeria. *Journal of Emerging Trends in Economics and Management Sciences*, *4*(4), pp.427-432.

Etuk, E.H. and Igbudu, R.C. 2013. A Sarima Fit to Monthly Nigerian Naira-British Pound Exchange Rates. *Journal of Computations & Modelling*, *3*(1), pp.133-144.

Etuk, E.H. and Ojekudo, N. 2014. Another look at the SARIMA modelling of the number of dengue cases in Campinas, State of Sao Paulo, Brazil. *International Journal of Natural Sciences Research*, *2*(9), pp.156-164.

Etuk, E.H., Elenga, B.N., Adonijah, M.A., Allen, D.S.A. and Etuk, F.E. 2014. A SARIMA fit to monthly Nigerian import commodity price indices. *Journal of Empirical Economics*, *3*(5), pp.306-313.

Etuk, E.H., Uchendu, B. and Victoredema, U.A. 2012. Forecasting Nigerian inflation rates by a seasonal arima model. *Canadian Journal of Pure and Applied Sciences*, p.2179.

Fang, T. and Lahdelma, R. 2016. Evaluation of a multiple linear regression model and SARIMA model in forecasting heat demand for district heating system. *Applied Energy*, 179, pp.544-552.

Fischer, C. 2008. Feedback on household electricity consumption: a tool for saving energy?. *Energy efficiency*, 1(1), pp.79-104.

Francq, C. and Sucarrat, G. 2018. An exponential chi-squared QMLE for log-GARCH models via the ARMA representation. *Journal of Financial Econometrics*, *16*(1), pp.129-154.

Fung, M.C., Peters, G.W. and Shevchenko, P.V. 2017. A unified approach to mortality modelling using state-space framework: characterisation, identification, estimation and forecasting. *Annals of Actuarial Science*, *11*(2), pp.343-389.

George, J.F., 2004. The theory of planned behavior and Internet purchasing. *Internet Research*.

Ghahramani, M. and Thavaneswaran, A. 2008. A note on GARCH model identification. *Computers & Mathematics with Applications*, *55*(11), pp.2469-2475.

Glosten, L.R., Jagannathan, R. and Runkle, D.E. 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), pp.1779-1801.

Gökbulut, R.I. and Pekkaya, M. 2014. Estimating and forecasting volatility of financial markets using asymmetric GARCH models: An application on Turkish financial markets. *International Journal of Economics and Finance*, *6*(4), pp.23-35.

Gouriéroux, C. 2012. *ARCH models and financial applications*. Springer Science & Business Media.

Granger, C.W. and Jeon, Y. 2004. Thick modelling. Economic Modelling, 21(2), pp.323-343.

Hanke, J.E. and Wichern, D.W. 2005. *Business forecasting*. Pearson Educación.

Helman, K. 2011. SARIMA models for temperature and precipitation time series in the Czech Republic for the period 1961–2008. *Aplimat-Journal of Applied Mathematics*, *4*(3), pp.281-290

Higgins, M.L. and Bera, A.K. 1992. A class of nonlinear ARCH models. *International Economic Review*, pp.137-158.

Hondroyiannis, G. 2004. Estimating residential demand for electricity in Greece. *Energy Economics*, *26*(3), pp.319-334.

Hou, A. and Suardi, S. 2012. A nonparametric GARCH model of crude oil price return volatility. *Energy Economics*, 34(2), pp.618-626.

Hyndman, R.J. and Athanasopoulos, G. 2014. Optimally reconciling forecasts in a hierarchy. *Foresight: The International Journal of Applied Forecasting*, (35), pp.42-48.

Inglesi, R. and Pouris, A. 2010. Forecasting electricity demand in South Africa: A critique of Eskom's projections. *South African Journal of Science*, *106*(1), pp.50-53.

Jarque, C.M. and Bera, A.K. 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, *6*(3), pp.255-259.

Jarque, C.M. and Bera, A.K. 1987. A test for normality of observations and regression residuals. *International Statistical Review/Revue Internationale de Statistique*, pp.163-172.

Jere, S. and Siyanga, M. 2016. Forecasting inflation rate of Zambia using Holt's exponential smoothing. *Open Journal of Statistics*, *6*(2), pp.363-372.

Jin, Y.H., Williams, B.D., Tokar, T. and Waller, M.A. 2015. Forecasting with temporally aggregated demand signals in a retail supply chain. *Journal of Business Logistics*, *36*(2), pp.199-211.

Johannesen, N.J., Kolhe, M. and Goodwin, M. 2019. Relative evaluation of regression tools for urban area electrical energy demand forecasting. *Journal of Cleaner Production*, *218*, pp.555-564.

Johansson, A. and Sowa, V. 2013. A comparison of GARCH models for VaR estimation in three different markets.

Karmakar, M. 2005. Modelling conditional volatility of the Indian stock markets.

Khan, M.S., Khan, K.I., Mahmood, S. and Sheeraz, M. 2019. Symmetric and asymmetric volatility clustering via GARCH family models: An evidence from religion dominant countries. *Paradigms*, *13*(1), pp.20-25.

Kipiński, L., König, R., Sielużycki, C. and Kordecki, W. 2011. Application of modern tests for stationarity to single-trial MEG data. *Biological Cybernetics*, *105*(3), pp.183-195.

Kwiatkowski, D., Phillips, P.C., Schmidt, P. and Shin, Y. 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?. *Journal of Econometrics*, *54*(1-3), pp.159-178.

Lee, A.H., WK, L. and van Hui, Y. 1996. On the Empirical Influence Function of the Portmanteau Statistic in AR (1) Process. *Journal of the Japan Statistical Society*, *26*(1), pp.83-90.

Lee, J. and Wong, D.W. 2001. *Statistical analysis with ArcView GIS*. John Wiley & Sons.

Lewis, C. D. 1982. International and Business Forecasting Methods. Butterworths, London.

Leybourne, S.J. and McCabe, B.P. 1994. A consistent test for a unit root. *Journal of Business & Economic Statistics*, *12*(2), pp.157-166.

Lindgren, G., Rootzén, H. and Sandsten, M. 2013. Stationary stochastic processes. *Theory and Applications; Texts in Statistical Science Series; CRC Press: Boca Raton, FL, USA*, p.347.

Ljung, G.M., Ledolter, J. and Abraham, B. 2014. George Box's contributions to time series analysis and forecasting. *Applied Stochastic Models in Business and Industry*, 30(1), pp.25-35.

Luo, C., Seco, L.A., Wang, H. and Wu, D.D. 2010. Risk modelling in crude oil market: a comparison of Markov switching and GARCH models. *Kybernetes*.

Lütkepohl, H. and Xu, F. 2012. The role of the log transformation in forecasting economic variables. *Empirical Economics*, *42*(3), pp.619-638.

Mahmud, I., Bari, S.H. and Rahman, M. 2017. Monthly rainfall forecast of Bangladesh using autoregressive integrated moving average method. *Environmental Engineering Research*, *22*(2), pp.162-168.

Mao, J., Yu, F., Zhang, Y., An, J., Wang, L., Zheng, J., Yao, L., Luo, G., Ma, W., Yu, Q. and Huang, C. 2018. High-resolution modelling of gaseous methylamines over a polluted region in China: source-dependent emissions and implications of spatial variations. *Atmospheric Chemistry and Physics*, *18*(11), pp.7933-7950.

Marwala, L. and Twala, B., 2014, July. Forecasting electricity consumption in South Africa: ARMA, neural networks and neuro-fuzzy systems. In *2014 International Joint Conference on Neural Networks (IJCNN)* (pp. 3049-3055). IEEE.

Miswan, N.H., Ping, P.Y. and Ahmad, M.H. 2013. On parameter estimation for Malaysian gold prices modelling and forecasting. *International Journal of Mathematical Analysis*, *7*(21-24), pp.1059-1068.

Molebatsi, K. and Raboloko, M. 2016. Time series modelling of inflation in Botswana using monthly consumer price indices. *International Journal of Economics and Finance*, *8*(3), p.15.

Mukaram, M.Z. and Yusof, F. 2017. Solar radiation forecast using hybrid SARIMA and ANN model: A case study at several locations in Peninsular Malaysia. *Malaysian Journal of Fundamental and Applied Sciences Special Issue on Some Advances in Industrial and Applied Mathematics*, *13,* pp.346-350.

Mwanga, D., Ong'ala, J. and Orwa, G. 2017. Modelling sugarcane yields in the Kenya sugar industry: A SARIMA model forecasting approach. *International Journal of Statistics and Applications*, *7*(6), pp.280-288.

Myers, D.E. 1989. To be or not to be... stationary? That is the question. *Mathematical Geology*, *21*(3), pp.347-362.

Nakiyingi, W. 2016. *Forecasting electricity demand using univariate time series volatility forecasting models: a case study of Uganda and South Africa* (Doctoral dissertation).University of kwaZulu-Natal.

Nelson, D.B., 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pp.347-370.

Nelson, D.B. and Cao, C.Q., 1992. Inequality constraints in the univariate GARCH model. *Journal of Business & Economic Statistics*, *10*(2), pp.229-235.

Ngailo, E., Luvanda, E. and Massawe, E.S. 2014. Time series modelling with application to Tanzania Inflation Data. *Journal of Data Analysis and Information Processing*, *2014*.

Nkwatoh, L.S. 2012. Forecasting unemployment rates in Nigeria using univariate time series models. *International Journal of Business and Commerce*, *1*(12), pp.33-46.

Nugroho, D.B., Kurniawati, D., Panjaitan, L.P., Kholil, Z., Susanto, B. and Sasongko, L.R. 2019, August. Empirical performance of GARCH, GARCH-M, GJR-GARCH and log-GARCH models for returns volatility. In *Journal of Physics: Conference Series* (Vol. 1307, No. 1, p. 012003). IOP Publishing.

Nyoni, T. and Nathaniel, S.P. 2018. Modelling rates of inflation in Nigeria: an application of ARMA, ARIMA and GARCH models.

Okafor, E.E. 2008. Development crisis of power supply and implications for industrial sector in Nigeria. *Studies of Tribes and Tribals*, 6(2), pp.83-92.

Osarumwense, O. and Waziri, E.I. 2013. Modelling monthly inflation rate, using Generalised Autoregressive Conditionally Heteroskedastic (GARCH) models: evidence from Nigeria. *Australian Journal of Basic and Applied Sciences*, *7*(7), pp.991-998.

Osarumwense, O.I. 2013. Applicability of Box-Jenkins SARIMA model in rainfall forecasting: A case study of Port-Harcourt South-South Nigeria. *Canadian Journal on Computing in Mathematics, Natural Sciences, Engineering and Medicine*, *4*(1), pp.1-4.

Otero, J. and Baum, C.F. 2017. Response surface models for the Elliott, Rothenberg, and Stock unit-root test. *The Stata Journal*, *17*(4), pp.985-1002.

Pandey, P.K., Tripura, H. and Pandey V., 2019. Improving Prediction Accuracy of Rainfall Time Series By Hybrid SARIMA–GARCH Modelling. *Natural Resources Research*, 28(3), pp.1125-1138.

Pepple, S.U. and Harrison, E.E. 2017. Comparative performance of GARCH and SARIMA techniques in the modelling of Nigerian board money. *CARD International. Journal Social Science Conflict. Management 2*, pp.258-270.

Phillips, P.C. and Perron, P. 1988. Testing for a unit root in time series regression. *Biometrika*, *75*(2), pp.335-346.

Rastogi, S., Don, J. and Nithya, V., 2018. Volatility estimation using GARCH family of models: Comparison with option pricing. *Pacific Business Review International*, *10*(8), pp.54-60.

Russo, E., Foster-McGregor, N. and Verspagen, B. 2019. Characterizing growth instability: new evidence on unit roots and structural breaks in long run time series (No. 026). United Nations University-Maastricht Economic and Social Research Institute on Innovation and Technology (MERIT).

Shabri, A. and Samsudin, R. 2014. A Hybrid GMDH and Box-Jenkins Models in Time Series Forecasting. *Applied Mathematical Sciences*, 8(62), pp.3051-3062.

Shongwe, S.C., Malela-Majika, J.C. and Molahloe, T. 2019. One-sided runs-rules schemes to monitor autocorrelated time series data using a first-order autoregressive model with skip sampling strategies. *Quality and Reliability Engineering International*, *35*(6), pp.1973-1997.

Shumway, R.H., Stoffer, D.S. and Stoffer, D.S. 2000. *Time series analysis and its applications* (Vol. 3). New York: Springer.

Sigauke, C. and Chikobvu, D. 2011. Prediction of daily peak electricity demand in South Africa using volatility forecasting models. *Energy Economics*, *33*(5), pp.882-888.

Ssekuma, R. 2011. *A study of cointegration models with applications* (Doctoral dissertation), pp.19 University of South Africa, South Africa

Stockhammar, P. and Öller, L.E. 2012. A simple heteroscedasticity removing filter. *Communications in Statistics-Theory and Methods*, *41*(2), pp.281-299.

Su, C. 2010. Application of EGARCH model to estimate financial volatility of daily returns: The empirical case of China.

Tang, J., Sriboonchitta, S. and Yuan, X. 2015. Forecasting inbound tourism demand to China using time series models and belief functions. In *Econometrics of Risk* (pp. 329-341). Springer, Cham.

Taylor, S.J. 2011. *Asset price dynamics, volatility, and prediction*. Princeton University Press.

Tendai, M. and Chikobvu, D., 2017. Modelling international tourist arrivals and volatility to the Victoria Falls Rainforest, Zimbabwe: Application of the GARCH family of models. *African Journal of Hospitality, Tourism and Leisure*, *6*(4), pp.1-16.

Thadewald, T. and Büning, H. 2007. Jarque–Bera test and its competitors for testing normality–a power comparison. *Journal of Applied Statistics*, *34*(1), pp.87-105.

Tseng, F.M., Yu, H.C. and Tzeng, G.H. 2002. Combining neural network model with seasonal time series ARIMA model. *Technological Forecasting and Social Change*, *69*(1), pp.71-87.

Tsoku, J.T., Phokontsi, N. and Metsileng, D., 2015, September. Forecasting South African Gold Sales: The Box-Jenkins Methodology. In *Proceedings of International Academic Conferences* (No. 2704589). International Institute of Social and Economic Sciences.

Vidalis, K., Markakis, G. and Tsimenides, N. 1997. Discrimination between populations of picarel (*Spicara smaris* L., 1758) in the Aegean Sea, using multivariate analysis of phenetic characters. Fisheries Research, 30(3), pp.191-197.

Warant, P., 2006. A comparison of forecasting method of daily jewellery gold price: Holt's forecast method, Box-Jenkins method and combined forecast method. *Naresuan University Journal*.

Yaziz, S.R., Azizan, N.A., Zakaria, R. and Ahmad, M.H. 2013, December. The performance of hybrid ARIMA-GARCH modelling in forecasting gold price. In *20th International Congress on modelling and simulation, Adelaide* (pp. 1-6).

Zhang, J., Xia, X. and Alexander, D. 2008, September. Demand side optimal strategy for voluntary load shedding. In *The Second IASTED Africa Conference on power and energy systems*.

Zhang, S. and Zhao, C. 2017. Stationarity test and Bayesian monitoring strategy for fault detection in nonlinear multimode processes. *Chemometrics and Intelligent Laboratory Systems*, 168, pp.45-61.

Zhou, B., He, D. and Sun, Z. 2006. Traffic modelling and prediction using ARIMA/GARCH model. In *Modelling and Simulation Tools for Emerging Telecommunication Networks* (pp. 101-121). Springer, Boston, MA.

Ziramba, E., 2008. The demand for residential electricity in South Africa. *Energy policy*, 36(9), pp.3460-3466.

APPENDIX

Lag Length: 12 (Automatic - based on SIC, maxlag=14)

	t-Statistic	Prob.*

Augmented Dickey-Fuller test statistic		-2.200454	0.2069
Test critical values:	1% level	-3.461783	
	5% level	-2.875262	
	10% level	-2.574161	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(ELECTRICITY)
Method: Least Squares
Date: 09/17/18 Time: 12:49
Sample (adjusted): 2001M02 2018M05
Included observations: 208 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ELECTRICITY(-1)	-0.053996	0.024539	-2.200454	0.0290
D(ELECTRICITY(-1))	-0.271128	0.057259	-4.735141	0.0000
D(ELECTRICITY(-2))	-0.096016	0.059469	-1.614540	0.1080
D(ELECTRICITY(-3))	-0.153941	0.059467	-2.588662	0.0104
D(ELECTRICITY(-4))	-0.184043	0.059840	-3.075597	0.0024
D(ELECTRICITY(-5))	-0.129855	0.060032	-2.163112	0.0318
D(ELECTRICITY(-6))	-0.229726	0.057316	-4.008075	0.0001
D(ELECTRICITY(-7))	-0.276267	0.056764	-4.866903	0.0000
D(ELECTRICITY(-8))	-0.198541	0.059105	-3.359121	0.0009
D(ELECTRICITY(-9))	-0.145773	0.059357	-2.455877	0.0149
D(ELECTRICITY(-10))	-0.102327	0.059413	-1.722288	0.0866
D(ELECTRICITY(-11))	-0.186895	0.059158	-3.159269	0.0018
D(ELECTRICITY(-12))	0.584951	0.056943	10.27267	0.0000
C	1158.746	509.8425	2.272753	0.0241

R-squared	0.842885	Mean dependent var	23.45673
Adjusted R-squared	0.832356	S.D. dependent var	1061.915
S.E. of regression	434.7935	Akaike info criterion	15.05255
Sum squared resid	36674801	Schwarz criterion	15.27720
Log likelihood	-1551.466	Hannan-Quinn criter.	15.14339
F-statistic	80.05869	Durbin-Watson stat	2.002448
Prob(F-statistic)	0.000000		

Null Hypothesis: D(ELECTRICITY) has a unit root
Exogenous: Constant
Lag Length: 11 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.772739	0.0001
Test critical values:	1% level	-3.461783	
	5% level	-2.875262	
	10% level	-2.574161	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation

Dependent Variable: D(ELECTRICITY,2)
Method: Least Squares
Date: 09/17/18 Time: 12:51
Sample (adjusted): 2001M02 2018M05
Included observations: 208 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(ELECTRICITY(-1))	-2.399091	0.502665	-4.772739	0.0000
D(ELECTRICITY(-1),2)	1.103726	0.465748	2.369794	0.0188
D(ELECTRICITY(-2),2)	0.989237	0.426663	2.318541	0.0215
D(ELECTRICITY(-3),2)	0.819643	0.385554	2.125882	0.0348
D(ELECTRICITY(-4),2)	0.624857	0.342166	1.826183	0.0694
D(ELECTRICITY(-5),2)	0.488050	0.297752	1.639116	0.1028
D(ELECTRICITY(-6),2)	0.252640	0.257085	0.982710	0.3270
D(ELECTRICITY(-7),2)	-0.024000	0.219088	-0.109546	0.9129
D(ELECTRICITY(-8),2)	-0.215591	0.180109	-1.196999	0.2328
D(ELECTRICITY(-9),2)	-0.349392	0.142142	-2.458055	0.0148
D(ELECTRICITY(-10),2)	-0.436973	0.103351	-4.228033	0.0000
D(ELECTRICITY(-11),2)	-0.606331	0.056658	-10.70168	0.0000
C	39.00155	31.78663	1.226980	0.2213

R-squared	0.947457	Mean dependent var	4.793269
Adjusted R-squared	0.944224	S.D. dependent var	1859.072
S.E. of regression	439.0559	Akaike info criterion	15.06759
Sum squared resid	37590158	Schwarz criterion	15.27619
Log likelihood	-1554.030	Hannan-Quinn criter.	15.15194
F-statistic	293.0229	Durbin-Watson stat	2.013391
Prob(F-statistic)	0.000000		

Null Hypothesis: ELECTRICITY has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 12 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.536954	0.8138
Test critical values:	1% level	-4.002786	
	5% level	-3.431576	
	10% level	-3.139475	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(ELECTRICITY)
Method: Least Squares
Date: 09/17/18 Time: 12:51
Sample (adjusted): 2001M02 2018M05
Included observations: 208 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ELECTRICITY(-1)	-0.051363	0.033419	-1.536954	0.1259
D(ELECTRICITY(-1))	-0.274482	0.064230	-4.273400	0.0000
D(ELECTRICITY(-2))	-0.099224	0.065684	-1.510629	0.1325
D(ELECTRICITY(-3))	-0.157052	0.065338	-2.403686	0.0172
D(ELECTRICITY(-4))	-0.186956	0.065004	-2.876059	0.0045
D(ELECTRICITY(-5))	-0.132545	0.064469	-2.055949	0.0411
D(ELECTRICITY(-6))	-0.232089	0.060942	-3.808376	0.0002

D(ELECTRICITY(-7))	-0.278377	0.059726	-4.660896	0.0000
D(ELECTRICITY(-8))	-0.200527	0.061665	-3.251892	0.0014
D(ELECTRICITY(-9))	-0.147532	0.061398	-2.402869	0.0172
D(ELECTRICITY(-10))	-0.103859	0.061003	-1.702540	0.0903
D(ELECTRICITY(-11))	-0.188270	0.060474	-3.113238	0.0021
D(ELECTRICITY(-12))	0.583880	0.057825	10.09729	0.0000
C	1114.444	637.2821	1.748745	0.0819
@TREND("2000M01")	-0.084165	0.723069	-0.116400	0.9075

R-squared	0.842896	Mean dependent var	23.45673
Adjusted R-squared	0.831500	S.D. dependent var	1061.915
S.E. of regression	435.9031	Akaike info criterion	15.06210
Sum squared resid	36672227	Schwarz criterion	15.30279
Log likelihood	-1551.458	Hannan-Quinn criter.	15.15942
F-statistic	73.96317	Durbin-Watson stat	2.001066
Prob(F-statistic)	0.000000		

Null Hypothesis: D(ELECTRICITY) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 11 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.039363	0.0002
Test critical values:	1% level	-4.002786	
	5% level	-3.431576	
	10% level	-3.139475	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(ELECTRICITY,2)
Method: Least Squares
Date: 09/17/18 Time: 12:54
Sample (adjusted): 2001M02 2018M05
Included observations: 208 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(ELECTRICITY(-1))	-2.667490	0.529331	-5.039363	0.0000
D(ELECTRICITY(-1),2)	1.350541	0.490069	2.755819	0.0064
D(ELECTRICITY(-2),2)	1.213122	0.448490	2.704903	0.0074
D(ELECTRICITY(-3),2)	1.020193	0.404920	2.519491	0.0126
D(ELECTRICITY(-4),2)	0.801662	0.359115	2.232325	0.0267
D(ELECTRICITY(-5),2)	0.641494	0.312419	2.053313	0.0414
D(ELECTRICITY(-6),2)	0.385361	0.269796	1.428342	0.1548
D(ELECTRICITY(-7),2)	0.087936	0.229686	0.382854	0.7022
D(ELECTRICITY(-8),2)	-0.126763	0.188198	-0.673561	0.5014
D(ELECTRICITY(-9),2)	-0.283850	0.147673	-1.922152	0.0561
D(ELECTRICITY(-10),2)	-0.393803	0.106596	-3.694341	0.0003
D(ELECTRICITY(-11),2)	-0.585324	0.058020	-10.08825	0.0000
C	141.3047	72.60952	1.946090	0.0531
@TREND("2000M01")	-0.836333	0.534151	-1.565722	0.1190

R-squared	0.948113	Mean dependent var	4.793269
Adjusted R-squared	0.944636	S.D. dependent var	1859.072
S.E. of regression	437.4309	Akaike info criterion	15.06465
Sum squared resid	37121077	Schwarz criterion	15.28929
Log likelihood	-1552.724	Hannan-Quinn criter.	15.15548

F-statistic	272.6846	Durbin-Watson stat	1.994208
Prob(F-statistic)	0.000000		

Null Hypothesis: ELECTRICITY has a unit root
Exogenous: Constant
Bandwidth: 2 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-4.956489	0.0000
Test critical values:	1% level	-3.460035	
	5% level	-2.874495	
	10% level	-2.573751	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	983500.2
HAC corrected variance (Bartlett kernel)	768740.4

Phillips-Perron Test Equation
Dependent Variable: D(ELECTRICITY)
Method: Least Squares
Date: 09/17/18 Time: 12:54
Sample (adjusted): 2000M02 2018M05
Included observations: 220 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ELECTRICITY(-1)	-0.218244	0.040745	-5.356287	0.0000
C	4512.967	840.1924	5.371349	0.0000

R-squared	0.116299	Mean dependent var	27.05909
Adjusted R-squared	0.112245	S.D. dependent var	1057.362
S.E. of regression	996.2546	Akaike info criterion	16.65493
Sum squared resid	2.16E+08	Schwarz criterion	16.68578
Log likelihood	-1830.043	Hannan-Quinn criter.	16.66739
F-statistic	28.68981	Durbin-Watson stat	2.759894
Prob(F-statistic)	0.000000		

Null Hypothesis: D(ELECTRICITY) has a unit root
Exogenous: Constant
Bandwidth: 7 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-29.18952	0.0000
Test critical values:	1% level	-3.460173	
	5% level	-2.874556	
	10% level	-2.573784	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	793519.9
HAC corrected variance (Bartlett kernel)	590191.3

Phillips-Perron Test Equation
Dependent Variable: D(ELECTRICITY,2)
Method: Least Squares
Date: 09/17/18 Time: 12:55
Sample (adjusted): 2000M03 2018M05
Included observations: 219 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(ELECTRICITY(-1))	-1.542184	0.057583	-26.78181	0.0000
C	38.40278	60.48090	0.634957	0.5261

R-squared	0.767732	Mean dependent var	9.488584
Adjusted R-squared	0.766662	S.D. dependent var	1852.584
S.E. of regression	894.8930	Akaike info criterion	16.44038
Sum squared resid	1.74E+08	Schwarz criterion	16.47133
Log likelihood	-1798.221	Hannan-Quinn criter.	16.45288
F-statistic	717.2656	Durbin-Watson stat	1.784065
Prob(F-statistic)	0.000000		

Null Hypothesis: ELECTRICITY has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 4 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-6.598726	0.0000
Test critical values:	1% level	-4.000316	
	5% level	-3.430383	
	10% level	-3.138772	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	927418.2
HAC corrected variance (Bartlett kernel)	935618.9

Phillips-Perron Test Equation
Dependent Variable: D(ELECTRICITY)
Method: Least Squares
Date: 09/17/18 Time: 12:56
Sample (adjusted): 2000M02 2018M05
Included observations: 220 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ELECTRICITY(-1)	-0.323344	0.049138	-6.580368	0.0000
C	6162.706	936.0256	6.583908	0.0000
@TREND("2000M01")	4.620313	1.275462	3.622462	0.0004

R-squared	0.166690	Mean dependent var	27.05909
Adjusted R-squared	0.159010	S.D. dependent var	1057.362
S.E. of regression	969.6596	Akaike info criterion	16.60531
Sum squared resid	2.04E+08	Schwarz criterion	16.65159
Log likelihood	-1823.584	Hannan-Quinn criter.	16.62400

F-statistic	21.70369	Durbin-Watson stat	2.610371
Prob(F-statistic)	0.000000		

Null Hypothesis: ELECTRICITY has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 4 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-6.598726	0.0000
Test critical values:	1% level	-4.000316	
	5% level	-3.430383	
	10% level	-3.138772	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	927418.2
HAC corrected variance (Bartlett kernel)	935618.9

Phillips-Perron Test Equation
Dependent Variable: D(ELECTRICITY)
Method: Least Squares
Date: 09/17/18 Time: 12:56
Sample (adjusted): 2000M02 2018M05
Included observations: 220 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
ELECTRICITY(-1)	-0.323344	0.049138	-6.580368	0.0000
C	6162.706	936.0256	6.583908	0.0000
@TREND("2000M01")	4.620313	1.275462	3.622462	0.0004

R-squared	0.166690	Mean dependent var	27.05909
Adjusted R-squared	0.159010	S.D. dependent var	1057.362
S.E. of regression	969.6596	Akaike info criterion	16.60531
Sum squared resid	2.04E+08	Schwarz criterion	16.65159
Log likelihood	-1823.584	Hannan-Quinn criter.	16.62400
F-statistic	21.70369	Durbin-Watson stat	2.610371
Prob(F-statistic)	0.000000		

```
>
Phillips-Perron Unit Root Test

data:  Data1
Dickey-Fuller = -5.451, Truncation lag parameter = 4, p-value = 0.01


Phillips-Perron Unit Root Test

data:  diff(Data1)
Dickey-Fuller = -16.681, Truncation lag parameter = 4, p-value = 0.01
```

```
> kpss.test(Data1, null = c("Level", "Trend"), lshort = TRUE)

        KPSS Test for Level Stationarity

data:  Data1
KPSS Level = 2.7541, Truncation lag parameter = 2, p-value = 0.01

Warning message:
In kpss.test(Data1, null = c("Level", "Trend"), lshort = TRUE) :
  p-value smaller than printed p-value
> kpss.test(diff(Data1), null = c("Level", "Trend"), lshort = TRUE)

        KPSS Test for Level Stationarity

data:  diff(Data1)
KPSS Level = 0.031487, Truncation lag parameter = 2, p-value = 0.1

Warning message:
In kpss.test(diff(Data1), null = c("Level", "Trend"), lshort = TRUE) :
  p-value greater than printed p-value
> fit1.ARCH1<-garch(Data1,order=c(0,1))

 ***** ESTIMATION WITH ANALYTICAL GRADIENT *****


     I     INITIAL X(I)        D(I)

     1     3.292225e+06     1.000e+00
     2     5.000000e-02     1.000e+00

    IT   NF      F         RELDF     PRELDF     RELDX    STPPAR   D*STEP   NP
RELDF
     0    1  1.914e+03
     1    2  1.164e+03  3.92e-01  8.11e+00  1.5e-07  1.6e+04  1.0e+00  6.2
9e+04
     2    3  1.164e+03  3.14e-07  1.57e-07  2.3e-11  0.0e+00  1.5e-04  1.5
7e-07

 ***** X-CONVERGENCE *****

 FUNCTION     1.164253e+03   RELDX        2.332e-11
 FUNC. EVALS       3         GRAD. EVALS       3
 PRELDF       1.572e-07      NPRELDF      1.572e-07

     I      FINAL X(I)        D(I)          G(I)

     1    3.292225e+06     1.000e+00     5.298e-09
     2    1.049846e+00     1.000e+00     2.377e+00

Warning messages:
1: In garch(Data1, order = c(0, 1)) : singular information
2: In sqrt(pred$e) : NaNs produced
> attributes(fit1.ARCH1)
$names
 [1] "order"         "coef"          "n.likeli"      "n.used"        "resi
duals"
 [6] "fitted.values" "series"        "frequency"     "call"          "vcov
"

$class
[1] "garch"
```

```
> Coefficients<-fit1.ARCH1$coef
> Coefficients
          a0           a1 
3.292225e+06 1.049846e+00 
> alpha0=Coefficients[1]
> alpha1=Coefficients[2]
> summary(fit1.ARCH1)

Call:
garch(x = Data1, order = c(0, 1))

Model:
GARCH(0,1)

Residuals:
   Min     1Q Median     3Q    Max 
0.8866 0.9288 0.9605 1.0150 1.0906 

Coefficient(s):
    Estimate  Std. Error  t value Pr(>|t|)
a0 3.292e+06          NA       NA       NA
a1 1.050e+00          NA       NA       NA

Diagnostic Tests:
        Jarque Bera Test

data:  Residuals 
X-squared = 8.055, df = 2, p-value = 0.01782


        Box-Ljung test

data:  Squared.Residuals
X-squared = 28.7, df = 1, p-value = 8.453e-08
```