Oil Price Volatility: GARCH, SVR-GARCH and EVT APPROACH

B Akomaning

orcid.org/0000-0003-3714-1803

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 E Munapo

Co-supervisor: Mr M Chanza

Graduation ceremony: October 2019

Student number: 22629092

DECLARATION

I Bridget Akomaning, hereby declare that the thesis titled “Oil Price Volatility: GARCH, SVR-GARCH and EVT Approach” was composed by me. All the sources have been referenced and acknowledged. The thesis is being submitted for Master of Commerce in Statistics with Business Statistics. Its original research has not been submitted anywhere.

....

Signature

ACKNOWLEDGEMENTS

I thank God for giving me the strength, wisdom and guidance throughout the writing of this document. I would like to thank my family for their support and encouragement for this dissertation. To my supervisors Professor Munapo and Mr Martin Chanza, thank you for your guidance and support throughout this research.

I would like to thank Edward Ocansey, Ishmeal Rapoo and Katlego Makatjane for their assistance. I would like to acknowledge NWU for the financial assistance for this studies.

ABSTRACT

Oil prices have been volatile over the past few years. Several models have been developed to describe volatility but the frequently used models are the ARCH and GARCH models. Research on GARCH and SVR-GARCH models have received little attention for studies on volatility especially in South Africa. This research seeks to assess the effectiveness of GARCH and SVR-GARCH models in modelling oil price volatility in South Africa. The study further employed EVT to fit and model the tails of oil prices. Daily data was collected from the JSE covering the period 7th August 2008 - 7th August 2018. The period was selected to cover the most recent trends of oil prices for the past 10 years. The study applied GARCH (1,1), FIGARCH(1,d,1) , EGARCH(1,1) and GJR-GARCH(1,1) and in comparison with SVR-GARCH(1,1), SVR-EGARCH(1,1) ,SVR-GJR-GARCH(1,1), SVR-FIGARCH(1,d,1) to model Brent Crude oil Prices in South Africa.

Preliminary data analysis was conducted before the actual analysis to quantify the behaviour of oil prices. The results indicated that Brent crude oil prices are heteroscedastic and auto correlated; hence the GARCH models are applicable. A detailed analysis of GARCH and SVR-GARCH was given. The study found SVR-EGARCH (1, 1) superior to the GARCH models. For the GARCH models, EGARCH (1, 1) was the best. EVT was used to fit the tails of the returns. The study fitted EGARCH (1, 1) and SVR-EGARCH (1, 1). The POT (Peak over threshold) method was employed in evaluating the GPD exceedances. The results showed that GPD fits adequately well and is sufficient in estimating tail risks.

The study recommends the use of SVR-EGARCH (1, 1) model as it is superior to EGARCH (1, 1). Multivariate data sets should be used for future studies. In addition, Stochastic Volatility models should be compared with the Support Vector Regression-GARCH models.

Keywords: GARCH, SVR-GARCH, Extreme Value theory, GPD, Volatility, POT

ABBREVIATIONS

ACF	Autocorrelation Function
ADF	Augmented Dickler Fuller
ARCH	Autoregressive Conditional Heteroscedascity
ES	Expected Shortfall
EVT	Extreme value theory
GARCH	Generalized Autoregressive Conditional Heteroscedascity
GEV	Generalized Extreme Value
GJR-GARCH	Glosten Jagannathan and Runkle GARCH
GPD	Generalized Pareto distribution
iid	Independent and identically distributed
JSE	Johannesburg Stock Exchange
MAE	Mean Absolute Error
MSE	Mean Squared Error
PACF	Partial Autocorrelation Function
POT	Peak over Threshold
PP	Philips and Perron
QQ	Quantile Quantile
RMSE	Root Mean squared error
SVM	Support Vector Machines
SVR	Support Vector Regression
VaR	Value at Risk
LM	Lagrange Multiplier
LB	Ljung Box

DEFINITION OF TERMS

I. Autocorrelation: The correlation between observations at different lags

II. Heteroscedascity: Uneven error variances

III. Homoscedastic : Even error variances

IV. Volatility : The fluctuations of a variable at different time intervals

V. Value at Risk: Probable loss of a portfolio at a given time frame (Francq & Zakoian, 2010).

VI. Expected Shortfall: “Expected loss of a financial position after a catastrophic event ” (Tsay, 2013) .

VII. Time series: A sequence of data points that is measured at different time intervals (Rhoda, 2013).

VIII. Mean excess plot: A graphical plot used to determine a threshold (Gilli & Këllezi, 2006).

IX. Support Vector Machine: Machine learning procedure used for classification and regression analysis.

X. Residuals: The variance between the actual and the predicted value.

XI. Extreme value theory deals with rare extreme events. It can be used to describe fat tails of profit or loss distribution (Rhoda, 2013).

LIST OF FIGURES

LIST OF TABLES

Table of Contents

CHAPTER ONE

ORIENTATION OF THE STUDY

1.1 BACKGROUND OF THE STUDY

The study assessed the effectiveness of GARCH and SVR-GARCH as models for studying volatility of oil prices in South Africa. Extreme value theory was applied to account for extreme events present in the data. The analysis was done by assessing the forecasting performance of each model in selection of the best model based on accuracy measures. Below gives a brief overview of GARCH, SVM/SVR-GARCH models and Extreme value theory (EVT). Section 1.2 provides the problem statement followed by the objectives and research design.

Modelling and forecasting volatility has sparked an interest amongst scholars, practitioners and researchers over the past few years. This is mainly motivated by its significance in the financial markets as it is being used to measure risk (Knight & Satchell, 2007). High frequency data of financial returns often displays characteristics such as volatility clustering, excess kurtosis and negative skewness. These are collectively known as stylized facts (Ugurlu *et al.*, 2014). Factoring stylized facts, numerous volatility models have been developed (Knight & Satchell, 2007). The commonly used models are the ARCH and the GARCH models (Bouseba & Zeghdoudi, 2015; Lim & Sek, 2013; Onwukwe *et al.*, 2014; Salisu & Fasanya, 2012; Shabani *et al.*, 2016). The ARCH was first established by Engle (1982). It was easy to use but many parameters were required to describe the volatility. This resulted to the introduction of the GARCH models revised in 1986 by Bollerslev. The weaknesses of the ARCH model was also found in the GARCH model as both models reacts to negative and positive shocks equally (Tsay, 2010; Wennstrom, 2014). Since then there have been several variants of the GARCH models such as IGARCH, GARCH-M, EGARCH, TGARCH, FIGARCH (Lim & Sek, 2013). These models helped improved GARCH as GARCH could not capture leverage effect.

Boser, Guyon and Vapnik (1992) developed a machine learning procedure used to analyse data through classification and regression analysis. This procedure was called Support Vector Machine (SVM). The SVM has found its way through many applications such as weather and stock predictions, speaker recognition and handwriting identification etc. The disadvantages of the SVM are it suffers from “slow training convergence when dealing with large datasets” because “the storage of variable requires a lot of memory and computational time” (Awad & Khanna, 2015; Wang, 2013).

When the SVM accounts for “linear and non-linear models” it is known as Support vector regression (Bezerra and Albuquerque, 2017). The Support vector Regression (SVR) has an exceptional predicting accuracy and robust to outliers. The major drawback is that it is subtle to corrupt data and retraining is required if there are changes made in the model (Awad and Khanna, 2015). Empirical studies have found the GARCH models to have low forecasting performance (Bildirici & Ersin, 2013; Geng & Liang, 2011; Lai & Liu, 2014; Ou & Wang, 2010). To overcome this weakness, volatility models based on the SVM/SVR-GARCH have been suggested in literature (Bezerra and Albuquerque, 2017). The performance of the SVR model is based on the kernel selected. Some of the most commonly used kernels are the linear, polynomial, radial basis (Awad and Khanna, 2015). Chung and Zhang (2017) found the radial base kernel to work best when the dataset is high with kurtosis. Qu and Zhang (2016) argued that using a kernel based on its non-linear dynamic enhances forecasting accuracy. A study by Huang *et al.*, (2014) found hybrid kernels to outperform single kernels. Li (2014) compared Gaussian and wavelet kernels to APARCH type models. The study found wavelet superior to the Gaussian kernels. Bezerra and Albuquerque (2017) compared a mixture of Gaussian kernels with GARCH type models and mortlet kernels; the study found a mixture of the Gaussian kernels superior to the other models.

Financial time series data often displays characteristics such as heavy tails or fat tails. A method that can be used to estimate and predict extremes in a data is known as Extreme value theory (Rhoda, 2013). According to Kiragu and Kyalo (2016) a careful analysis needs to be done when differentiating between extreme values and outliers. The peak over threshold (POT) is one of the methods that fall under extreme value theory (EVT) in addition to the block maxima method (BMM). The BMM method only examines extreme values in specified blocks (Sowdagur and Narsoo, 2017). The POT method is preferred over the BMM method as it is fully parametric and easy to extrapolate (Li, 2015). In addition, one major drawback of the POT method is the selection of a suitable threshold (Susan and Waititu, 2015). The Mean excess plot has been used by Frad & Zouari, 2014; Sowdagur & Narsoo, 2017; Susan & Waititu, 2015 in finding the suitable threshold selection. One major drawback of the EVT is that in the short term risk, managers are more concerned with loss in present times but the EVT is incapable of revealing time varying volatility (Li, 2015). The purpose of the study was to compare GARCH with SVR-GARCH models and also examine the performance of EVT method on oil prices in South Africa.

1.2 PROBLEM STATEMENT

Previous empirical studies have successfully used GARCH, SVR-GARCH and Extreme value theory in modelling volatility and extreme risk. For the SVR-GARCH models, many studies only compared SVR-GARCH (1, 1) with asymmetric GARCH models. Little or no study has compared asymmetric GARCH models with asymmetric SVR-GARCH models. This study attempts to deviate from previous studies by filling this gap. In modelling extreme risk, the GARCH models have been used, especially the GARCH (1, 1). A thorough search of literature revealed only little research has been done using SVR-GARCH in modelling EVT.

1.3 OBJECTIVES

The study assessed the effectiveness of GARCH and SVR-GARCH as models for studying volatility of oil prices in South Africa.

1. To model symmetric and asymmetric GARCH and SVR-GARCH models.
2. Evaluate GARCH and SVR-GARCH based on accuracy measures and make predictions.
3. Use the best GARCH and SVR-GARCH to model extreme risk of Brent crude oil prices.

1.4 DATA COLLECTION AND VARIABLES

The study used daily closing prices from the Johannesburg Stock Exchange (JSE) for Brent Crude oil from 7th August 2008 to 7th August 2018. The period was selected to cover the most recent trends of oil prices.

1.5 RESEARCH METHODS AND TESTS

The study used GARCH and Support Vector Regression-GARCH to model and forecast oil price volatility in South Africa. EVT was further employed to fit the tails of the returns. In testing for stationarity, the study employed formal and informal methods. For the informal methods graphical plots were used, the formal tests included the ADF and PP. To test for heteroscedascity, the study used Breusch Pagan and the LM test. For serial correlation, the Ljung box test was used. In executing the analysis, EViews 10 and RStudio 3.5 were used.

1.6 SIGNIFICANCE OF THE STUDY

Due to the current fluctuations of Brent Crude oil prices, this study is worthy. The outcome of the results will help practitioners, researchers and risk managers to adopt alternative methods for forecasting. The study serves as a guide for the Department of Energy in selecting the model that best describes volatility amongst GARCH and Support Vector Regression-GARCH.

1.7 CONTRIBUTION OF THE STUDY

GARCH and SVR-GARCH models were evaluated. Depending on the best model selected, the study used that model for forecasting. The study also reviewed EVT to fit the tails of the returns of oil prices. The results of the study will provide new evidence to policy makers about the volatility of future oil prices and extreme risk of oil prices. The study will be used as a reference for other researchers.

1.8 DELIMITATIONS

No multivariate data sets were used because the aim of the study was to compare one variable using GARCH and SVR-GARCH models. The quantitative study was conducted from the 7th August 2008 – 7th August 2018 to cover the most recent trends of Brent crude oil prices in South Africa. Increasing the study period to include periods prior to 2008 instead of the ten years reported in this study will result in different conclusions as a result of price volatility over the extended time period. The study is in the context of South Africa and similar works on GARCH, SVR-GARCH and EVT are limited.

1.9 ETHICAL CONSIDERATIONS

Ethical clearance was obtained from NWU (NWU - 0 0 1 3 4 - 1 9 - A 4).

1.10 ORGANISATION OF THE STUDY

Chapter one reviewed a general background, the problem statement, objectives, methodology and limitations of the study. **Chapter two** provides the empirical studies of previous works. **Chapter three** underpins the methods and procedures that was used for the analysis. **Chapter four** discusses the empirical findings. Conclusions and recommendations are discussed in **Chapter five.**

1.11 CHAPTER SUMMARY

The Chapter reviewed the introduction, background, problem statement, objectives, research methods and tests. Ethical clearances was obtained from the NWU. The significance of conducting the study are outlined in the Chapter. Limitations and delimitations are also provided. The next Chapter presents the literature review of GARCH, SVR-GARCH and Extreme Value theory.

CHAPTER TWO

LITERATURE REVIEW

2.1 INTRODUCTION

Literature on GARCH, SVM/SVR-GARCH and EVT are reviewed. The Chapter seeks to provide empirical literature of previous studies in order to identify gaps. The Chapter consists of three parts; the first part deals with empirical literature on GARCH, the second part presents a review of SVR-GARCH followed by Extreme value theory.

2.1.1 GARCH MODELS

In recent times, there has been a vast amount of literature on models such as ARCH and GARCH in modelling volatility all over the world. These models have become prevalent due to their capability to capturing stylized facts. One study to note was by Yaziz *et al.*, (2011), in their studies, they investigated the use of GARCH and Box Jenkins models to analyse oil prices in West Texas intermediate from 1986-2009. The study used ARIMA (1, 2, 1) and GARCH (1, 1) models. Based on the forecasting performance of MSFE, MSE, RMSE, MAE and Theil U; GARCH (1, 1) displayed a small error forecast compared to ARIMA (1, 2, 1). This was due to its "ability to capture the volatility of non-constant of conditional variance". The authors recommended further studies to use the hybrid model Box Jenkins-GARCH and also investigate the use of IGARCH and EGARCH models. A study by Aamir and Shabri (2016) confirmed the use of ARIMA-GARCH model by forecasting crude oil in Pakistan.

For symmetric models the GARCH (1, 1) model has been recommended by many studies. For instance, a study by Bouseba and Zeghdoudi (2015) used univariate GARCH models to model daily oil prices from 2009-2014. The study used the following models: GARCH(1,1), GARCH(1,2), GARCH(2,1), GARCH(2,2), GARCH(1,3). The results indicated that GARCH (1,1) is simple and easy to use. Similarly, Daddikar and Raygopal , 2016; Shabani *et al.*, 2016 also endorsed the use of GARCH (1,1) model. Another study to note was by Cheteni (2016), who compared the relationship between stock returns in South Africa and China. The study only used GARCH (1, 1) to measure volatility. It didn't try to measure volatility using EGARCH and IGARCH models as its primary objective was to estimate volatility and to examine the existence of dependence in the returns. The results of the study exhibited that the GARCH (1, 1) is adequate to capture volatility clustering and leptokurtosis. The study recommended that future studies should use multivariate methods in identifying the relationship between China and South Africa and also include data from other countries such as India and Russia in comparison of stock markets to the South African markets.

In comparing GARCH models, the asymmetric models out performs the symmetric GARCH models (Maqsood *et al.*, 2017; Onwukwe *et al.*, 2014; Salisu & Fasanya, 2012). A study by Salisu and Fasanya (2012) considered symmetric (GARCH (1, 1), GARCH-M (1, 1)) and asymmetric (EGARCH (1, 1) and TGARCH (1, 1)) models to measure volatility. The authors found oil price to be most volatile during financial crisis. The asymmetric models appeared to be superior over the symmetric models.

Lim and Sek (2013) conducted empirical analysis to model Malaysian stock market using GARCH models for the period 1990-2010. The study used GARCH (symmetric) and asymmetric (TGARCH, EGARCH) models. The results showed that the symmetric and asymmetric models perform differently at specified times. The GARCH and TGARCH models performed well in the pre-crisis, GARCH performed best in the crisis period and TGARCH worked well in post crisis.

Cristiani-d'Ornano *et al.*, (2010) evaluated GARCH, IGARCH and FIGARCH in quest of finding the model that can capture persistence in volatility. The study found FIGARCH to be the best as it generated accurate forecasts.

Onwukwe *et al.*, (2014) used symmetric GARCH (1,1), ARCH(1), ARCH(2) in comparison with asymmetric EGARCH(1,1) and TGARCH(1,1) models to forecast the returns of 15 Nigerian banks stocks. The results displayed that EGARCH (1,1) is superior.

Maqsood *et al.*, (2017) examined GARCH family of models in modelling Nairobi securities. The study estimated GARCH(1,1), EGARCH(1,1) ,GARCH-M(1,1), TGARCH(1,1) and PGARCH(1,1). The study found TGARCH (1, 1) superior due to its "significant effects of leverage".

Atoi (2014) estimated GARCH models using three error distributions such as normal, the student *t* and the GED. The volatility models used were GARCH, EGARCH, TGARCH and PGARCH models. The outcome of the results revealed the existence of leverage effect as volatility responds to bad news more than good. The out-of-sample estimates exhibited that the PGARCH model is the best. The study recommended future studies to use different distributions to achieve a robust forecasting model.

Daddikar and Raygopal (2016) analysed crude oil prices volatility patterns by employing GARCH models. The results indicated that GARCH (1, 1) and E-GARCH (1, 1) with *t* distribution were found to better explain asymmetric volatility. The authors recommended that future studies should undertake modelling crude oil prices with intra-day frequency and also the impact of external factors such as foreign exchange, gold prices on oil volatility.

Klein and Walther (2016) proposed a mixture memory GARCH (MM GARCH) in comparison with Risk metrics, EGARCH, FIGARCH, GARCH, HYGARCH, FIAPACH and APARCH. The results indicated that MMGARCH out performs the other models "due to its dynamic approach in varying volatility".

In evaluating the performance of the best asymmetric model, Daddikar and Raygopal (2016) found the EGARCH (1, 1) model to best explain asymmetric volatility. Similar results was obtained by Onwukwe *et al.*, (2014). However in the studies of Atoi (2014), the PGARCH model gave accurate forecasts based on the out-of-sample prediction. Klein and Walther (2016) then proposed a hybrid model mixture memory GARCH (MMGARCH) and found that it outperformed the other discrete GARCH models.

The distribution of error also has an impact on the results. Kosapattarapim *et al.*, (2012) examined the forecasting capabilities of GARCH models using six distributions of error namely normal, skewed normal, student t, GED and skewed GED. The outcome of the results revealed that skewed error distributions is superior to the other distributions. Ahmed and Shabri (2013) fitted GARCH models using the prices Brent and WTI. The authors fitted GARCH-t, GARCH-N, and GARCH-G to forecast oil prices. The results indicated that GARCH-N is good for Brent and GARCH-G is adequate for WTI. Atoi (2014) compared three distributions of error namely student *t* distribution, normal and GED. The study found EGARCH, PGARCH and GARCH well fitted by the student *t* distribution and TGARCH fits the GED distribution.

In comparison of the best estimation technique Shabani *et al.*, (2016) compared GARCH models with MLE and GMM to evaluate the performance. The study used GARCH (1, 1), GARCH (2, 2), GARCH (3, 1), GARCH (3, 3). GARCH (1, 1) was found to be the best model when compared to the competing models. The study compared GARCH (1, 1) with MLE and GMM. The results obtained indicated that MLE is better than GMM when estimating the parameters of GARCH. The authors recommended that future studies on multivariate and bivariate GARCH should use different estimation parameters. The next Section presents the SVR-GARCH models.

2.1.2 SVR-GARCH MODELS

The SVM models are machine learning models used for regression and classification. Due to the low forecasting performance of GARCH models, literature has proposed a hybrid model SVM/SVR-GARCH models to improve forecasting performances. In comparing the GARCH and SVR/SVM-GARCH models, the SVM/SVR-GARCH models provides accurate results as compared to the GARCH models (Chung & Zhang, 2017; Geng & Zhang, 2015; Ou & Wang, 2010). Studies by Lai & Liu, 2014; Lux *et al.*, (2018) have found hybrid models of the SVR models to outperform the standard SVR models. Below provides a summary of empirical findings.

Ou and Wang (2010) compared LSSVM (least square support vector machines) with GARCH models. The study evaluated GARCH (1, 1), GJR-GARCH (1, 1) and EGARCH (1, 1) with GARCH-LSSVM, EGARCH-LSSVM and GJR-LSSVM. Results suggested that the hybrid models outperforms the parametric models.

Geng and Liang (2011) compared GARCH, GM-GARCH and SVRGM-GARCH in quest of finding the best model. The results proved that SVRGM-GARCH is superior to the other models. In evaluation of the GM-GARCH and the GARCH model, the GM-GARCH outperforms the GARCH.

Bildirici and Ersin (2013) evaluated GARCH, SVR-GARCH and MLP-GARCH models in quest of finding the best model. Results showed that SVR-GARCH and MLP-GARCH models outperformed the GARCH. The study further compared SVR-GARCH and MLP-GARCH; the SVR-GARCH provided better forecasts than the MLP-GARCH model.

Lai and Liu (2014) explored SVM and least square support vector machines (LS-SVM) in predicting stock prices. The study combined GARCH, SVR and LSSVM with wavelet kernels in forecasting stock prices. The results indicated that the wavelet model is not as great as the LS-SVM model. It was then concluded that the LS-SVM is the best model.

Li (2014) evaluated the performance of support vector machines and Quasi maximum likelihood (QML) using APARCH type models. The study used GJR, TS-GARCH, GARCH and TGARCH models. The results indicated that the SVM models outperform the QML, the study further investigated the kernels of SVM. The author compared Gaussian and wavelet kernels; results showed that the wavelet kernel outperformed the Gaussian by producing accurate forecasts as fewer vectors are needed which in turn improves prediction ability.

Yekkehkhany *et al.*, (2014) compared the performance of SVM using 3 different kernels namely linear, radial and polynomial to classify crops. The results showed that the radial kernel is more suitable than the linear and polynomial kernel functions due to its speed of convergence.

Qu and Zhang (2016) proposed a new kernel in comparison with radial basis and sigmoid kernels for predicting stock returns in China. The results indicated that the new kernel was the best amongst the radial basis and sigmoid kernels. The study recommended the development of a new kernel for SVR forecasting and that future studies should apply the new kernels in energy markets.

Chung and Zhang (2017) examined six different foreign exchange rates using GARCH and SVM-GARCH models. The study evaluated the performance of SVM-GARCH (1,1) with parametric models : GARCH (1,1), GJR-GARCH(1,1) and EGARCH(1,1) models. The study used three kernels namely: polynomial, radial and linear. The outcome of the results proved that the SVM-GARCH model displays accurate results than the parametric models. The polynomial kernel gave better performance in 4/6 datasets followed by the linear kernel with 5/6. The radial kernel gave a better performance in all the six data sets. The study found radial kernel to work best if the data is high in kurtosis.

Bezerra and Albuquerque (2017) assessed the performance of SVR-GARCH (1, 2, 3, 4) Gaussian kernels with SVR-GARCH mortlet, GARCH, EGARCH and GJR models. The study used skew-student *t*, student *t*, Gaussian and GED innovations in estimating the models. The results exhibited that SVR-GARCH when mixed with Gaussian kernels improves volatility forecasting.

Lux *et al.*, (2018) proposed SVR-GARCH-KDE and compared it with GARCH, EGARCH, TGARCH using three error distributions such as normal, *t* distribution and skewed *t* distribution. The outcome of the results proved SVR-GARCH-KDE competitive to the competing models. The authors recommended that the hybrid model could be enhanced by the use of asymmetric models such as TGARCH with skewed t distribution as it displayed great results.

Nanda *et al.*, (2018) used the linear, radial basis function (RBF), the sigmoid and the polynomial kernel functions to detect termites' acoustic signal. The study used the AUC known as the area under the curve to evaluate and enhance the performance of the results. The polynomial kernel function with the accuracy of 0.19188 outperformed the other kernels.

The Section reviewed empirical studies on SVM/SVR-GARCH models in modelling volatility. The next Section reviews empirical literature on EVT.

2.1.3 EXTREME VALUE THEORY

EVT handles events that are extreme in the financial markets. To model for extreme values, the data needs to be independent and identically distributed (i.i.d). Studies by Kiragu & Kyalo, 2016; Susan & Waititu, 2015 used GARCH models to eliminate ARCH effects. In comparison of the best fitting distribution Alam *et al.,*(2018) compared Pearson type 3, GEV and log pearson type 3. The study found GEV as the best distribution. In comparison of EVT, Jammazi and Nguyen (2017) proposed a wavelet EVT in contrast of the standard EVT. The authors found wavelet EVT to produce more accurate results than the standard EVT. Below reviews empirical literature on Extreme value theory.

Yu and Shih (2007) used probability distributions to explore the returns and volatility of gold, dollar and crude oil. The study estimated the parameters using Gaussian distribution for the returns and log-normal distribution for volatility. The outcome of the results showed that crude oil market displayed the greatest return followed by gold and British market. For volatility, crude oil was the most unstable market with gold and pound in succession.

Tolikas and Gettinby (2009) investigated the performance of generalized pareto (GP), generalized logistic (GL) and generalized extreme value (GEV). The study found GL distributions superior to the other models as it fitted the data best.

Yu and Shih (2011) used probability distributions to find the effect of the weekend on oil and gold markets. The study found that the weekend does not have an effect on gold and oil markets as Friday and Monday did not show the highest and lowest returns respectively. The study found Wednesday and Thursday had effects on oil and gold markets. The results indicate that the trading behaviour of investors and their beliefs change with time.

Rösch and Harald (2012) investigated the impact of OPEC announcement on the tail behaviour of crude oil prices. The study fitted the tail behaviour using the Generalized Pareto distribution. The study found the tail behaviour to be heavy during pre-announcement and relaxed after the announcement is made. The lower tail reacted in the opposite way.

Zin *et al.*, (2014) investigated the suitable models to model extreme share returns in Malaysia. The study evaluated models such as: Gumbel, GEV, GPA, GNO and Pearson (PE3) distributions. The study found Generalised Pareto and Pearson distribution to best fit the data.

Susan and Waititu (2015) modelled oil price risk using two oil benchmarks namely WTI and Brent. The GARCH-EVT approach was employed. The study used the GJR-GARCH and GARCH to model volatility. The study found the GARCH (1, 1) adequately more fitting than the asymmetric GARCH GJR based on the AIC. The peak over threshold was used in analysing the Generalized Pareto distribution (GPD). The study found oil prices to be highly volatile, fat tailed and heteroscedastic; WTI yielded higher risk than the Brent. The authors recommended further studies to consider methods of threshold selection.

Nortey *et al.*, (2015) applied EVT to fit stock returns in Ghana. The study employed the POT method to fit the GPD. In estimating the GPD, the study employed the maximum likelihood estimator (MLE) and the probability weighted moment (PWM). The MLE showed more accurate estimates compared to the PWM. The authors concluded that GPD fits the left tail as compared to the right.

Kiragu and Kyalo (2016) used the POT method to model the tail behaviour of the Nairobi security exchange (NSE) index. The study fitted ARMA (1, 1)-GARCH (1, 1) in order to account for heteroscedascity and autocorrelation that might be present in the residuals. The threshold selection was based on the hill plot, shape parameter and mean excess function. The results showed that the right tail is 1.15 and the left is 0.84. The VAR and expected shortfall were carried out the results designated that when investing in the NSE, the probability of losses is smaller than gaining.

Halder *et al.*, (2016) evaluated the performance of generalized pareto distribution estimation methods with an application to stock data. The study compared the maximum likelihood (MLE), Method of moments (MOM), PWM (probability weighted moments) estimators and the maximum penalized likelihood (MPLE). The study found the PWM method efficient when the data is positively skewed, MLE with large sample size when the data is not skewed. The MOM estimator performed well, but the PWMU provided good estimates. The study found no method uniformly best based on the stock data but the MOM performed well compared to the MLE estimator.

Jammazi and Nguyen (2017) proposed a wavelet extreme value theory (W-EVT) in comparison with the standard EVT using exchange rate and crude oil data. The empirical results suggested that the W-EVT executes accurate forecasts than the standard EVT model.

Marsani *et al.*, (2017) evaluated the performance of GLD, GPA, GLO and pearson (PE3). The study estimated the parameters using L-moments; the results revealed that the GLD model outperformed other models.

Noshkov and Demirtas (2017) examined the market risk of different energy commodities such as WTI, natural gas and coal to come up with a good model that for risk management. The study calculated the value at risk (VAR) for value weighted historical simulations (VWHS), Extreme value theory conditional peaks over threshold together with GARCH (1, 1), EGARCH (1, 1) and TGARCH (1, 1) using the student *t* distribution as the data was not normal. The results exhibited that the EGARCH model is superior amongst the competing GARCH models, the VWHS underperformed the VAR models and the EVT conditional POT model provided great results.

Li, (2017) employed the GARCH and EVT in quest of finding the best model. The study evaluated the unconditional GARCH models (GARCH, TGARCH and EGARCH) with the conditional EVT models (GARCH-EVT, TGARCH-EVT and EGARCH-EVT) using the GED. The results proved that the EGARCH was the best model for predicting VAR.

Alam *et al.*, (2018) evaluated three models in finding the best probability distribution model. The study compared GEV, pearson and log-pearson type 3. The L moments were used to estimate the parameters. The results proved that the generalized extreme value was the best fit as it yielded 36% followed by Pearson and log-pearson type 3 with 26% of the stations.

The Section reviewed various empirical literature on extreme value theory. The next Section provides a summary of empirical literature on GARCH, SVM/SVR-GARCH and Extreme value theory.

2.1.4 SUMMARY OF EMPIRICAL LITERATURE: TABLE 2.1

Authors	**Techniques**	**Models**	**Results**	**Conclusions and Recommendations**
Cristiani-d'Ornano *et al.*, (2010)	GARCH	IGARCH,GARCH,FIGARCH	FIGARCH outperformed the other two models	FIGARCH
Yaziz *et al.*, (2011)	GARCH, BOX JENKINS	ARMA(1,2,1), GARCH(1,1)	GARCH(1,1) outperformed ARIMA	IGARCH, EGARCH
Atoi (2014)	GARCH	GARCH, TGARCH, EGARCH and PGARCH	PGARCH is the best amongst the three models	Use different error distributions to achieve a robust forecasting model.
Bildirici and Ersin (2013)	GARCH/SVR-GARCH	GARCH, SVR-GARCH and MLP-GARCH	SVR-GARCH provided better forecasts than MLP-GARCH and GARCH	SVR-GARCH
Chung and Zhang (2017)	GARCH and SVM-GARCH models	SVM-GARCH (1,1), GARCH(1,1), EGARCH(1,1), GJR-GARCH(1,1)	SVM-GARCH outperformed GARCH(1,1), EGARCH(1,1), GJR-GARCH(1,1)	The study found radial kernel to work best if the data is high in kurtosis.
Bezerra and Albuquerque (2017)	GARCH, SVR-GARCH	SVR-GARCH (1, 2, 3, 4) Gaussian kernels with SVR-GARCH mortlet, GARCH, EGARCH and GJR models	SVR-GARCH with a mixture of Gaussian kernels improves volatility forecasting.	SVR-GARCH with a mixture of Gaussian kernels
Zin *et al.*, (2014)	EVT	Gumbel, GEV, GPA, GNO and Pearson (PE3) distributions	Generalised Pareto and Pearson distribution	Generalised Pareto and Pearson distribution
Nortey *et al.*, (2015)	EVT	POT (GPD)	GPD of the left tail fits better	POT method fits the GPD and efficient in modelling extreme events
Kiragu and Kyalo, (2016)	EVT	POT (GPD)	GPD provides a great fit for the data	The probability of losses is lesser than gaining if investing in the NSE

2.1.5 RESEARCH GAPS

Table 2.2 reviews the research gaps of previous empirical studies

TABLE 2.2

Authors	**techniques**	**Models**	**Results**	**Remarks**
Cristiani-d'Ornano *et al.*, (2010)	GARCH	IGARCH,FIGARCH,GARCH	FIGARCH	Didn't include asymmetric GARCH models
Atoi (2014)	GARCH	GARCH, TGARCH, EGARCH and PGARCH	PGARCH	Didn't include skewed student t distribution
Chung and Zhang (2017)	GARCH and SVM-GARCH models	SVM-GARCH (1,1), GARCH(1,1), EGARCH(1,1), GJR-GARCH(1,1	SVM-GARCH(1,1)	Only compared symmetric and asymmetric GARCH with SVM GARCH(1,1)
Bezerra and Albuquerque (2017)	GARCH and SVM-GARCH models	SVR-GARCH(1,1),GARCH(1,1), EGARCH(1,1), GJR-GARCH(1,1	SVR-GARCH(1,1)	Only compared symmetric and asymmetric GARCH with SVM GARCH(1,1)
Susan and Waititu (2015)	Extreme value theory	GARCH(1,1), GJR-GARCH(1,1)	GARCH (1,1)	Only compared GARCH(1,1) with GJR-GARCH(1,1)
Kiragu and Kyalo (2016)	Extreme value theory	GARCH(1,1) and POT method	The GPD is adequate in modelling extreme values	Only used GARCH (1,1)

2.2 SUMMARY AND CONCLUSIONS

Throughout the review of empirical literature, it is quite evident that Brent crude oil prices can be modelled with GARCH, SVR-GARCH and Extreme value theory. The frequently used GARCH models are the GARCH (1,1), EGARCH(1,1), PGARCH(1,1,1) and TGARCH(1,1); but the FIGARCH model has not been given much attention. This study fills a gap by including FIGARCH model to account for long memory volatility. Atoi (2014) highlighted that the type of distribution often has an impact on the results. Motivated by Atoi (2014), this study seeks to compare different error distribution in selection of the best error distribution. SVR-GARCH models have been reported to be the best amongst the GARCH models. In estimating the GARCH models using SVR, many studies only used GARCH (1, 1) framework (Bezerra & Albuquerque, 2017; Chung & Zhang, 2017). Little or very few studies have included asymmetric GARCH models. This study will deviate from previous studies by including asymmetric SVR-GARCH models. In modelling the tail behaviour using EVT, the data needs to be i.i.d. Previous studies have used the GARCH models to account for i.i.d. A study by Kiragu and Kyalo (2016) found ARMA(1,1)-GARCH(1,1) model to be best fitting. Similarily Susan and Waititu (2015) also found the GARCH(1,1) adequate.This study will use SVR-GARCH and GARCH in modelling the tails of oil prices using EVT. The next Chapter presents the methodology of GARCH, SVR-GARCH and EVT.

CHAPTER THREE

METHODOLOGY

3.1 INTRODUCTION

This Chapter focuses on the techniques that were employed to attain the objectives set out for this study. The Chapter reviews data descriptions, methods and techniques that were used in analysing the data. The study adopts GARCH, SVR-GARCH and EVT as they are adequate in modelling volatility and extreme values.

3.1.1 RESEARCH PROCEDURE

Figure 3.1 illustrates the procedure of the analysis. The study uses GARCH, SVR-GARCH models and EVT which consists of the Generalized Pareto distributions.

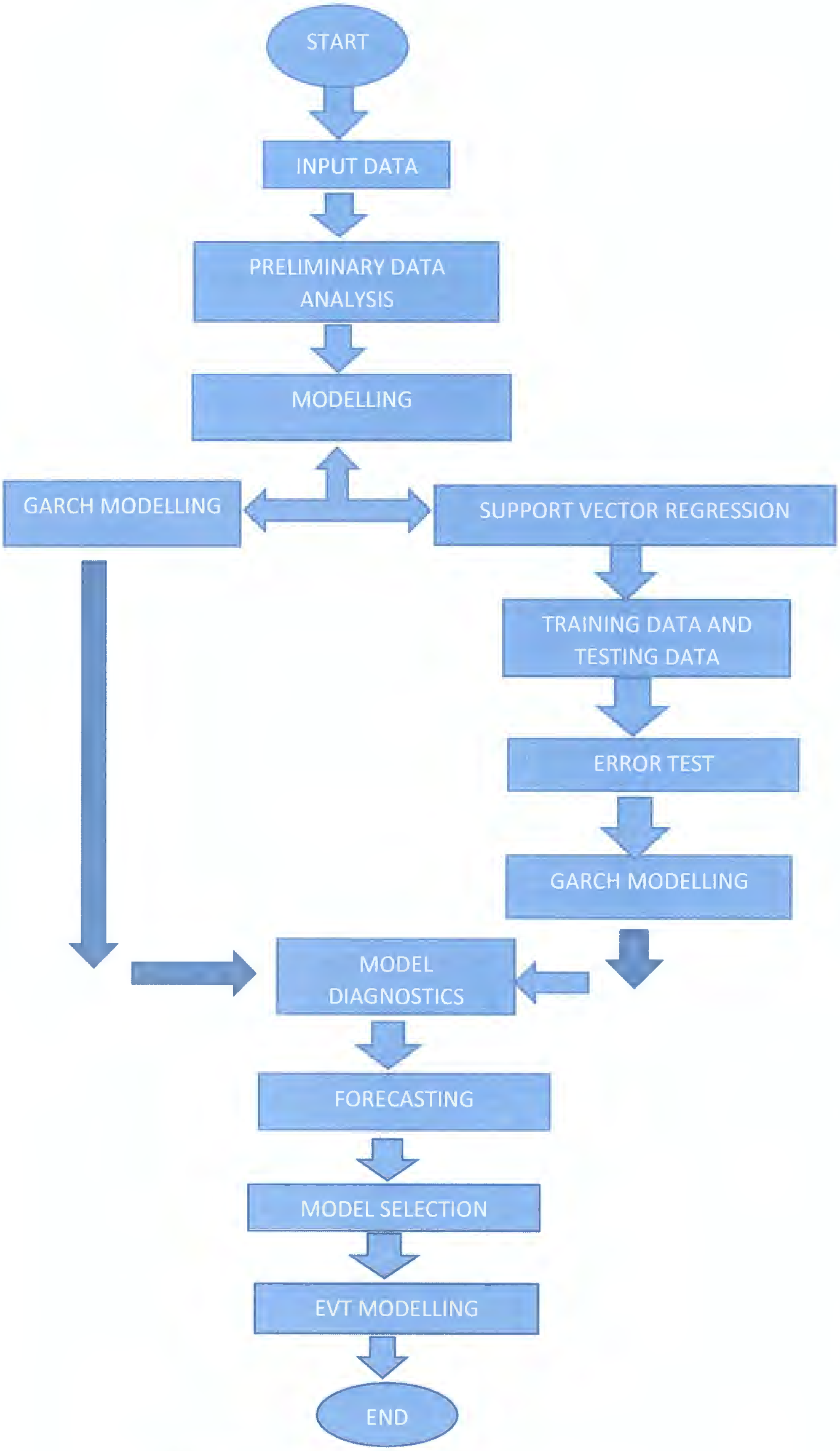


Figure 3.1: Research Procedure

3.1.2 DATA DESCRIPTION AND SOURCES

The study used quantitative data as time series is based on estimating outcomes. Secondary data were obtained from the JSE from 7th August 2008 - 7th August 2018 resulting a total of 2525 observations excluding weekends. The period was selected to cover the most recent trends of Brent crude oil prices in South Africa. The study used daily closing prices for the past ten years. Gaston (2016) stated that the return series has greater statistical properties than the actual prices. Following Gaston (2016), the study used the return series of Brent crude oil prices.

3.1.3 ETHICAL ISSUES

The study was approved by the university's ethics committee.

3.1.4 STATISTICAL SOFTWARE PACKAGES

Various software packages such as SAS, Stata, RStudio and EViews to mention a few can also be used in analysing data. In modelling for volatility and Extreme value theory analysis authors like Tsay (2010) used RStudio because it is an open source, easy to use and its graphical displays are exceptional. This study used EViews 10 and RStudio 3.5 packages to execute the analysis. The study used EViews 10 mainly for unit root tests as it displays comprehensive results compared to RStudio. RStudio was mainly used for the analysis because EViews cannot perform Support vector regression and Extreme value theory analysis.

3.1.5 PRELIMINARY DATA ANALYSIS

The main purpose of a preliminary data analysis is to edit, summarize and describe the data features before the actual analysis can be conducted (Blischke *et al.,* 2011). This can be done by using formal and informal tests. Graphical plots and unit root tests were used in this study to identify the behaviour of the random variable. Descriptive statistics such as skewness and kurtosis are provided.

3.2 NORMALITY TESTS

A normality test is used to identify normal and non-normal data distributions. In identifying the distribution of the data set, the study used Skewness, Kurtosis, Jarque Bera and Shapiro Wilks tests.

3.2.1 SKEWNESS

Wegner (2012) defined skewness as the shape of a unimodal distribution of a random variable. Skewness can follow a symmetrical, positive and negative distribution. When a distribution is symmetrical its mean, mode and median are the same. If the distribution is positive, its mean is larger than the median and if it is negative the mean value is less than the median.

Following Wegner (2012) the Pearson skewness can be computed as follows:

$$Sk_p = \frac{n\sum(x_i - \bar{x})^3}{(n-1)(n-2)s^3} \tag{3.1}$$

Where n = number of values,

x_i = ith data value of x,

$\bar{x}$ = the mean value and s is the standard deviation.

According to Wegnor (2007) a marginal skewness is present when the coefficient lies between -0.5 and +0.5, it is moderate when it lies between -1 and +1 and excess skewness is detected when the coefficients lies outside the range (< -1 and > +1).

3.2.2 KURTOSIS

Kurtosis measures how peaked a distribution is. Following Rhoda (2013) the value of kurtosis is 3 if it is normally distributed, if it surpasses three it means that the distribution has heavy tails and it points close to the mean.

The kurtosis is expressed as follows:

$$K = \frac{E(x-\mu)^4}{\sigma^4} \tag{3.2}$$

Where $\mu = mean, \sigma\ is\ the\ standard\ deviation\ of\ X$ and E is the expectation operator.

3.2.3 JARQUE BERA

The JB test is a combination of kurtosis and skewness. The JB test is written as:

$$JB = \left[\frac{n}{6}\right]\left[s^2 + (k-3)^2/4\right] \tag{3.3}$$

The standardization is based on normality since skewness (S) is zero and Kurtosis (K) is 3 for a normal distribution. Their asymptotic variances are $\frac{6}{n}$ and $\frac{24}{n}$. The JB test uses a chi-square distribution with 2 df. One may reject the null hypothesis (H_0) of normal distribution if JB is less than the significance level (Das & Imon, 2016; Tsay, 2013).

H_0: The residuals are normal distribution

H_1: The residuals are not normal distribution

3.2.4 SHAPIRO-WILK'S TEST

The Shapiro Wilk's test is widely used in literature due to its powerful properties. Let's consider a random sample $y_1 < y_2 < \cdots y_n$ the Shapiro Wilks test is written as:

$$W = \frac{(\sum_{i=1}^{n} a_i y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{3.4}$$

Where y_i= i^{th} order

$\bar{y}$= mean sample

$$a_i = (a_1 \ldots . a_n) = \frac{m^T V^{-1}}{(m^T V^{-1} V^{-1} m)^{1/2}} \text{ and m = } (m_1, \ldots . m_n)^T$$

$V = covariance\ matrix$

The $W\ lies\ between\ 0\ and\ 1$ (Razali & Wah, 2011). The null hypothesis of normality is rejected if the probability value is smaller than the chosen alpha level.

3.2.5 HISTOGRAM AND KERNEL DISTRIBUTION FUNCTION

A histogram graphically displays the distribution of data. It can also be used to identify if the data is normally distributed by providing an insight to the skewness, outliers and fat tails. A kernel density estimation (KDE) is a non-parametric way of approximating a probability density function (pdf). The kernel density estimator allows you to have a smooth feel of how the data is distributed unlike the histogram. Consider a series of random variables $X_1, X_2 \ldots\ldots X_n$

The KDE at point x is computed as follows:

$$f_h(x) = \frac{1}{nh} \sum_{i=1}^{n} k \left(\frac{x - X_i}{n}\right) \tag{3.5}$$

n is represented by the sample size and $f\ is\ the\ density$

Where kernel k satiates $\int_{\infty}^{\infty} k(x)dx = 1,$ (3.6)

The smoothing *h* is known as a bandwidth; equation (3.6) is due to the division of the sum by *nh* (Rhoda, 2013).

3.3 STATIONARITY PROCESS

A process is said to be stochastic if the random variable (RV) is ordered in time. When stationarity occurs, the mean, variance and autocorrelation are constant with time. To illustrate a weak stationarity process, let Y_t denote a stochastic process with these properties:

$$E(Y_t) = \mu \tag{3.7}$$

$$Var(Y_t) = E[(Y_t - \mu)^2] = \delta^2 \tag{3.8}$$

$$\gamma_k = E[(Y_{t,} - \mu)(Y_{t+k} - \mu)] \tag{3.9}$$

Where γ_k is the covariance at lag k between $Y_{t,}$ andY_{t+k}. If K= 0, γ_0 is obtained which is the variance of Y $(= \sigma^2)$ (Gujarati and Porter, 2008).

Time series is non-stationary if its variance and mean are not steady with time; this might produce spurious results. Spuriosity occurs when the results of the data displays a high R-squared (R^2) with a low Durbin-Watson value. This can be regarded as good results but are in fact of no use (Jan van Greunen *et al.*, 2014). Datta and Mukhopadhyay (2011) highlighted that in order to avoid spuriosity in the data, a unit root test must be conducted. In conducting unit root tests, numerous studies have used the ADF and Philips and PP tests (Abdalla, 2012; Ahmed & Shabri, 2013; Arce *et al.*, 2015; Kristjanpoller & Minutolo, 2016). The DF, PP and ADF unit root tests are discussed in section 3.3.1 and 3.3.2.

3.3.1 DICKEY FULLER AND AUGMENTED DICKEY FULLER TEST

Dickey fuller (1979) is a unit root test on the basis of AR (1) model. It can be represented as follows:

$$Y_t = \emptyset_1 y_{t-1} + \varepsilon_t \tag{3.10}$$

Where t=1,....T

$\emptyset_1$= AR parameter and ε_t meets the characteristics of a white noise process.

The null hypothesis H_0: $\emptyset_1$= 1 denotes non-stationarity or unit root. The alternative hypothesis H_1 :$|\emptyset|$ <1 denotes stationarity or no unit root (Arltová & Fedorová, 2016).

Arltová and Fedorová (2016) further highlighted that when calculating the test statistic for DF, when y_{t-1} equation (3.10) is removed from both sides:

$$\Delta y_t = \beta y_{t-1} + \varepsilon_t \tag{3.11}$$

Where $\beta = \emptyset_1$-1

The DF test statistic as defined as

$$t_{DF} = \frac{\widehat{\emptyset}_1 - 1}{S\,\widehat{\emptyset}_1} \tag{3.12}$$

Where $\widehat{\emptyset}_1$ is the least square estimate of $\emptyset_1$ and the standard error is $S\,\widehat{\emptyset}_1$

Equation (3.10) can be extended with a constant or linear trend

$$y_t = \beta_0 + \emptyset_1 y_{t-1} + \varepsilon_t \tag{3.13}$$

$$y_t = \beta_0 + \beta_1 t + \emptyset_1 y_{t-1} + \varepsilon_t \tag{3.14}$$

As soon as a non-systematic element in DF is auto correlated, the ADF test is created and converted as:

$$y_t = \emptyset_1 y_{t-i} + \sum_{i=1}^{p-1} \gamma_1 \Delta y_{t-i} + \varepsilon_t \tag{3.15}$$

The ADF test can then be calculated as follows:

$$\Delta y_t = (\emptyset_1 - 1) y_{t-i} + \sum_{i=1}^{p-1} \gamma_1 \Delta y_{t-i} + \varepsilon_t \tag{3.16}$$

A major drawback of this test is the selection of lag p (Arltová & Fedorová, 2016). Çevik *et al.*, (2013) pointed out that the ADF test lacks power. The calculated value of the ADF test is compared with the critical value at a certain level of significance (1%, 5%, and 10 %). If the value calculated surpasses the critical value the null hypothesis is rejected as the data is stationary.

3.3.2 PHILIPS PERRON

Philips and Perron test is an alternate to the ADF test. The difference between both tests is by how they deal with heteroscedascity and autocorrelation. The ADF test uses parametric methods whiles the PP unit root test uses non parametric methods. In a regression model, Philips and Perron ignores any serial correlation present. No lag length needs to be specified when using the PP unit root which serves as an advantage (Zivot & Wang, 2006).

The PP unit root test is written as:

$$\Delta Y_t = \beta' D_t + \pi y_{t-1} + u_t \tag{3.17}$$

Where u_t $I(0)$ might be heteroskedastic, so the test improves it for any autocorrelation and heteroskedasticity in the errors of u_t.

The critical value is compared with the calculated value at a certain level of significance (1%, 5%, and 10%). If the calculated value surpasses the critical value the null hypothesis is rejected as the data is stationary.

3.4 MODEL BUILDING

To fit a good model, the study adopted the Box and Jenkins methodology. The Box and Jenkins methodology follows the following iterative steps:

- **Identification**: To find the p, d, q values. Plots such as the ACF and the PACF can be used in finding the values.
- **Parameter estimating**: Having identified the values of the p, d, q the next stage is to estimate the parameters.
- **Diagnostic checks:** Having chosen and estimated the parameters, it is essential to assess the adequacy of the model. A way to do this is to confirm whether or not the residuals are i.i.d. If not then the process is restarted.
- **Forecasting:** If the residuals are i.i.d then the model can be used for prediction (Gujarati and Porter, 2008).

3.4.1 AUTOCORRELATION (ACF) AND PARTIAL AUTOCORRELATION FUNCTION (PACF)

A measure of correlation between a variable and its lagged value at different lags is known as the ACF.

The ACF at lag k is given as

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{covariance\ at\ lag\ k}{variance} \qquad (3.18)$$

If $k = 0, \rho_0 = 1$

ρ_k lies between -1 and +1

If ρ_k is plotted against k the graph obtained is called a population correlogram. The PACF measures the relationship between Y_t and Y_{t-k} after removing the influence of intermediate Ys (Gujarati and Porter, 2008).

Following Katchova (2013) the partial autocorrelation function is given as:

$$\rho_k = Corr[y_t - E^*(y_t | y_{t-1}, \ldots, y_{t-k+1}), y_{t-k}] \qquad (3.19)$$

Where $E^*(y_t | y_{t-1}, \ldots, y_{t-k+1})$ is the minimum MSE predictor of y_t by $y_{t-1}, \ldots . y_{t-k+1}$.

Katchova (2013) further detailed the ACF and PACF properties in Table 3.1 as:

Table 3.1: ACF and PACF

	AR(p)	MA(q)	ARMA(p,q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off at lag p	Tails off	Tails off

Source: Katchova (2013)

3.4.2 ARCH EFFECTS

When modelling for volatility using the ARCH and GARCH models, it is important to test for ARCH effects in the residuals. This allows you to identify autocorrelation or heteroscedasticity. For serial correlation the Ljung Box (LB) and Lagrange multiplier test (LM) were used. The Breusch Pagan was used to test for heteroscedascity. Initially we discuss the LB test. The LB-test (Q_m) for autocorrelation is applied to the ϵ_t^2 series where H_0 the first m lags of the ACF of ϵ_t^2 is 0 (Gaston, 2016).

The Ljung Box test expressed as:

$$Q(m) = N(N+2)\sum_{i=1}^{m}\frac{\hat{\rho}_i^2}{N-i} \tag{3.20}$$

Where the sample size is represented by N and m denotes the lags. $\hat{\rho}_i^2$ is the estimate of the ACF squared residuals.

$$\rho_i = \frac{\sum_{i+1}^{n}(c-\mu)(\varepsilon_{r-i}^2-\mu^2)}{\sum_{r=1}^{N}(\varepsilon_{r-i}^2-\mu)^2} \tag{3.21}$$

Where $\hat{\mu}$ is the sample mean specified as $\hat{\mu} = \frac{1}{N}\sum_{r=1}^{N}\epsilon_r^2$, the null hypothesis is rejected if $Q(m) > X_m^2(\alpha)$ (Gaston, 2016).

When testing for ARCH effects present in the residuals ξ_t using the LM; the null hypothesis is such that $H_0: \pi_i = 0$ meaning that no ARCH effect is present up to order q at 5% significance level illustrated in equation (3.22)

$$\xi_t^2 = \psi_0 + \left(\sum_{i=1}^{q}\pi_i\varepsilon_{t-I}^2\right) + \mu_t \tag{3.22}$$

Where ψ_0 is a constant and the error term is denoted by μ_t. For the GARCH models to be applicable in this study the error terms need to be heteroscedastic thus the acceptance of the alternative hypothesis (Atoi, 2014).

Breusch Pagan test

As noted by Halunga *et al.*, (2015) the Breusch Pagan test is expressed as:

$$BP_T = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \rho_{ij,T} \tag{3.23}$$

Where

$$\rho_{ij,T} = \frac{\frac{1}{\sqrt{T}} \sum_{t=1}^{T} u_{it} \hat{u}_{jt}}{\sqrt{\left\{ \frac{1}{T} \sum_{t=1}^{T} u_{it}^{2} \right\} \left\{ \frac{1}{T} \sum_{t=1}^{T} u_{jt}^{2} \right\}}}$$

Following Gujarati and Porter (2008) to illustrate the Breusch Pagan test, let's consider a k-variable linear regression model:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \cdots \beta_k X_{ki} + u_i \tag{3.24}$$

Assuming that the error variance is expressed as:

$$\sigma_i^2 = f(\alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi}) \tag{3.25}$$

Thus, σ_i^2 is a certain function of non-random Z variables; some of the X can serve as Z.

Let's assume that

$$\sigma_i^2 = \alpha_1 + \alpha_2 Z_{2i} + \cdots + \alpha_m Z_{mi} \tag{3.26}$$

Equation 3.26 designate that σ_i^2 is a linear function of Zs. If $\alpha_2 = \alpha_3 \dots \alpha_m = 0$, $\sigma_i^2 = \alpha_1$, which is constant. In order to test if σ_i^2 is homoscedastic, the null hypothesis can be tested as $\alpha_2 = \alpha_3 \dots \alpha_m = 0$.

The hypothesis are as follows:

$H_0: homoscedastic$

$H_1: Heteroscedastic$

The next section presents the AR, MA and ARMA models.

3.5 AUTOREGRESSIVE (AR MODELS)

The ARCH and GARCH models are a combination of simple models which both need to be well understood. Below are representations which illustrate why the ARCH and GARCH models only relate to a time series data (Wong, 2014).

The AR (p) models utilize the p lag variables and it is given as:

$$Y_t = c + \sum_{i=1}^{p} \phi_i Y_{t-i} + \epsilon_t \tag{3.27}$$

The notion of the AR model is that the response variable is a "linear function of its previous values as lag variables".

Consider an AR (1) model: $Y_t = c + \phi Y_{t-1} + \epsilon_t$ (3.28)

Where ϕ_1 is constant, ϵ_t is the error term with t as time which is considered a white noise process. The white noise process has a finite variance and the mean = 0. The Gaussian white noise distribution is frequently used (Wong, 2014).

3.6 MOVING AVERAGE (MA)

The MA uses q lag error terms and is computed as

$$Y_t = d + \epsilon_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} \tag{3.29}$$

The MA model uses past errors for prediction. The MA (1) process is expressed as

$$Y_t = d + \epsilon_t + \theta_1 \epsilon_{t-1} \tag{3.30}$$

Where the constant term is denoted by θ_1 and ϵ is a white noise process (Wong, 2014).

3.7 AUTOREGRESSIVE MOVING AVERAGE (ARMA)

Wong (2014) further illustrated the ARMA model which consist of the AR and MA process represented in equation (3.27) and (3.29) respectively. The ARMA model is written as

$$Y_t = c \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_i \epsilon_{t-j} \tag{3.31}$$

The ARMA (1, 1) is represented as

$$Y_t = c + \epsilon_t + \phi_1 Y_{t-1} + \theta_1 \epsilon_{t-1} \tag{3.32}$$

Equation (3.32) incorporates AR (1) and MA (1) models. The next section provides Stylized facts and volatility measurement followed by the ARCH/GARCH models.

3.8 VOLATILITY: STYLIZED FACTS

In modelling volatility in financial markets, the data sometimes displays patterns that are critical for estimation, precise model specification and prediction. The stylized facts are as follows:

- Fat tails: A normal distribution's kurtosis is 3, if it surpasses 3, it means that the data exhibits excess kurtosis meaning that they have fat tails.
- Volatility clustering: Sometimes bigger movements are followed by massive movements; this signals shock persistence. Box Ljung and correlograms can be used to test and identify the existence of correlations.
- Leverage effects: A negative relationship between price movement and volatility.
- Long memory: When handling high frequency data, volatility is incessant. This led to two proposals to account for persistence. These include the unit root and long memory. The stochastic volatility and ARCH models can be used to model persistence (Knight and Satchell, 2007).

3.9 VOLATILITY MEASUREMENT

When a variable is inconsistent with time, it is known as volatility. Following Poon (2005) volatility is given by:

$$\hat{\sigma}=\sqrt{\frac{1}{T-1}\sum_{t-1}^{T}(r_t-\mu)^2} \qquad (3.33)$$

Where r_t is the return of day t, μ is the average return over $T\ period.$

The returns are computed as follows:

$$r_t = \log\left(\frac{P_a}{P_{a-1}}\right) \qquad (3.34)$$

Where P_a is the current closing price of Brent crude oil and P_{a-1} is the price of the previous day. The next section presents the ARCH and GARCH models.

3.10 AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY (ARCH) MODELS

The ARCH model developed in 1982 by Engle was designed to capture volatility persistence. The ARCH model (q) is expressed as:

$$h_t = \omega + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 \tag{3.35}$$

With $\omega > 0$ and $\alpha_j \geq 0$ to ensure that h_t always has a positive variance. q is of high order as a result of volatility persistence (Poon, 2005).

3.10.1 PROPERTIES OF THE ARCH MODEL

Considering an ARCH (1) process:

$$a_t = \delta_t \epsilon_t \qquad \sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 \tag{3.36}$$

Where $\alpha_0 > 0$, $\alpha_1 \geq 0$.

- The unconditional mean of a_t residues zero because
  $E(a_t) = E(E(a_t|F_{t-1}) = E[\sigma_t E(\epsilon_t)] = 0$

- Unconditional variance of a_t can be designated as:

$Var(a_t) = E(a_t^2) = E[E(a_t^2|F_{t-1})] = E(\alpha_0 + \alpha_1 a_{t-1}^2) = \alpha_0 + \alpha_1 E(a_{t-1}^2)$ since a_t is a stationary process with $E(a_t) = 0, Var(a_t) = Var(a_{t-1}) = E(a_{t-1}^2)$.

So the $Var(a_t) = \alpha_0 + \alpha_1 Var(a_t)$ and $Var(a_t) = \frac{\alpha_0}{(1-\alpha_1)}$. Since the variance a_t must be positive $0 \leq \alpha_1 < 1$ is required (Tsay, 2010).

3.10.2 WEAKNESS OF THE ARCH MODEL

Tsay (2010) outlined the ARCH weaknesses as:

- The positive and negative shocks responds equally.
- A restriction is imposed on the model; the α_1^2 of an ARCH (1) should be in an interval of $[0,\frac{1}{3}]$.
- The model needs more parameters to describe volatility.

To overcome these weaknesses, Bollerslev (1986) introduced the GARCH models. The next section presents the GARCH models.

3.11 GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSCEDASCITY (GARCH) MODELS

The ARCH model was simple to use but it often required more parameters to estimate volatility. Bollerslev (1986) then proposed the GARCH models to overcome the weaknesses. Following Francq and Zakoian (2010) a process is called GARCH (p, q) only if it meets these two conditions:

i) $E(\epsilon_t|\epsilon_u, u<t) = 0, t\epsilon Z$

ii) The constants $\omega, \alpha_i, i = 1, \ldots q \ and \ \beta_j, j = 1, \ldots p$ such that:

$$\sigma_t^2 = Var(\epsilon_t|\epsilon_u, u < t) = \omega + \sum_{i=1}^{q} \alpha_i \, \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \, \sigma_{t-j}^2, t\epsilon \ Z \tag{3.37}$$

Alternatively equation (3.37) can be written as follows:

$\sigma_t^2 = \omega + \alpha(B)\epsilon_t^2 + \beta(B)\sigma_t^2, \ t\epsilon Z,$ where B is the backshift operator ($B^i \ \epsilon_t^2 = \epsilon_{t-i}^2 \ and \ B^i\sigma_t^2 = \sigma_{t-i}^2$ for integer i) such that $\alpha \ and \ \beta$ are polynomials of $q \ and \ p$.

$$\alpha(B) = \sum_{i=1}^{q} \alpha_i \, B^i, \ \beta(B) = \sum_{i=1}^{p} B_j \, B^j$$

The positive and drawbacks of the GARCH model is evident in the GARCH (1, 1) process.

The simplest GARCH (1, 1) can be modelled as:

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \qquad 0 \leq \alpha_1, \beta_1 \leq 1, (\alpha_1 + \beta_1) < 1 \tag{3.38}$$

Where a big a_{t-1}^2 /σ_{t-1}^2 contributes to a large σ_t^2; it implies that a large a_{t-1}^2 leads to a larger a_t^2. In addition, when comparing normal distribution, the GARCH (1,1) tail is heavier (Tsay, 2010). A disadvantage of the traditional GARCH models is that the normal error distribution can't capture "stylized facts" (Sahoo *et al.*, 2018). Many variants of the GARCH models have been introduced to "mimic stylized facts". These extensions are able to detect long memory properties, asymmetry and non-linearity that is present in volatility process (Ardia, 2008). Among these models are the GARCH-M, EGARCH, GJR-GARCH, IGARCH, FIGARCH, PGARCH, and AGARCH. The next section presents the symmetric and asymmetric GARCH models.

3.11.1 SYMMETRIC GARCH MODELS

This section presents the GARCH (p,q) and GARCH-M models

3.11.1.1 GARCH (p, q)

Following Ødegaard (2018)The simple GARCH (p, q) is expressed as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \beta_i \sigma_{t-i}^2 \quad (3.39)$$

3.11.1.2 GARCH IN MEAN (GARCH-M)

The GARCH-M model can be expressed as follows:

$$y_t = \mu + \delta\sigma_{t-1} + u_t, u_t \sim N(0, \sigma_t^2) \quad (3.40)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_t^2$$

If δ is positive and significant, it means that if the conditional variance rises it leads to a higher return of the mean (Brooks, 2008).

3.11.2 ASYMMETRIC GARCH MODELS

This section presents the EGARCH, GJR-GARCH and PGARCH models.

3.11.2.1 EXPONENTIAL GARCH (EGARCH) MODEL

Nelson (1991) introduced the EGARCH model designed to capture leverage effect. It has quite a few advantages over the standard GARCH. The conditional variance is modelled in logarithmic form ($\log(\sigma_t^2)$) so that with negative parameters, σ_t^2 will be positive and inflicting non-negativity restraints on the parameters will not be required (Brooks, 2008; Gaston, 2016).

The conditional of EGARCH (p,q) is written as:

$$\log(\sigma_t^2) = \beta_0 + \sum_{i=1}^{q} \left\{ \alpha_i \left| \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right| + \gamma_i \left(\frac{\varepsilon_{t-i}}{\sigma_{t-i}} \right) \right\} + \sum_{j=1}^{p} \beta_j \log(\sigma_{t-j}^2) \quad (3.41)$$

$\varepsilon_{t-i} > 0\ and \varepsilon_{t-i} < 0$ designates positive and negative news with their total effects as $(1 + \gamma_i)| \varepsilon_{t-i}|$ and $(1 - \gamma_i)| \varepsilon_{t-i}|$ respectively. Once $\gamma_i < 0$ negative news will have a great impact on volatility. Covariance stationarity is achieved by the model when $\sum_{j=1}^{p} \beta_j < 1$. γ_1 = 0 when H_0 is rejected this indicates the presence of leverage effect (Atoi, 2014).

3.11.2.2 GJR-GARCH

The GJR-GARCH model detects leverage effect. The conditional variance is as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma u_{t-1}^2 I_{t-1} \tag{3.42}$$

Where $I_{t-1} = 1$ if $u_{t-1} < 0$

= 0 , otherwise

For leverage effect $\gamma > 0$ and for non-negativity the condition is that $\alpha_0 > 0, \beta \geq 0$ *and* $\alpha_1 + \gamma \geq 0$. The model is still acceptable if $\gamma < 0$ provided $\alpha_1 + \gamma \geq 0$ (Brooks, 2008).

3.11.2.3 POWER GARCH (PGARCH)

Another variant that takes leverage effect into account is the PGARCH model developed by Ding *et al.*, (1993). The PGARCH model is as follows:

$$\sigma_t^d = \omega + \sum_{i=1}^{p} \alpha_i (u_{t-i} + \beta u_{t-i})^d + \sum_{j=1}^{q} \beta_i \, \sigma_{t-j}^d \tag{3.43}$$

Where d and γ_1 denote a positive exponent and leverage effect coefficient respectively (Sahoo *et al.*, 2018). The long memory model FIGARCH is presented in 3.11.3.

3.11.3 LONG MEMORY VOLATILITY (FRACTIONAL INTERGRATED GARCH)

Baillie *et al.*, (1996) developed a FIGARCH model to factor for long memory property that might be present in financial returns. The FIGARCH (1, d, 1) is written as follows:

$$\sigma_t^2 = \omega + [1 - \beta(L) - \emptyset(L)\,(1 - L)^d\,]\,x_t^2 + \beta\sigma_{t-1}^2 \tag{3.44}$$

Where $\omega > 0, \emptyset < 1, \beta < 1, 0 \leq d \leq 1$ (Lux *et al.*, 2015).

The lag operator is denoted by L and d is *the parameter of fractional differentiation*. The parameters have to meet the following criteria:

$$\beta - d \leq \emptyset \leq \frac{(2-d)}{3} \text{ and d } [\emptyset - \frac{(1-d)}{2}] \leq \beta(d - \beta + \emptyset) \tag{3.45}$$

Equation (3.44) can be rewritten as follows:

$$\sigma_t^2 = \omega(1-\beta)^{-1} + [1 - (1-\beta)\emptyset\,(L)(1-L)^d]x_t^2$$

$$= \omega(1 - \beta)^{-1} + \eta(L)\,x_t^2 \tag{3.46}$$

Where $\eta(L) = \eta_1 L + \eta_2 L^2 + \ldots\ldots, \eta_j \geq 0$

For j= 1,2,.....

$\eta(L)$ *can be illustrated as follows:*

$$\left\{\begin{array}{llll} \eta_1 = & \hat{\phi} - & \hat{\beta} + & \hat{d} \\ \vdots & \vdots & . & . \\ \eta_j = \hat{\beta}_{\eta_{j-1}} + [(j-1-\hat{d})j^{-1} - \hat{\phi} & &] & \pi_{j-1} \end{array}\right\}$$

Where π_j=$\pi_{j-1}(j-1-\hat{d})_j^{-1}$ are the coefficients of fractional differencing operator $(1-L)^d$ (Lux *et al.*, 2015).

In selection of the models to be estimated, this study considered the GARCH (1,1), EGARCH(1,1) , GJR-GARCH(1,1) to account for symmetric and asymmetric volatility as they have been highly recommended in literature (Bouseba & Zeghdoudi, 2015; Cheteni, 2016; Onwukwe *et al.*, 2014). The FIGARCH model was endorsed by Cristiani-d'Ornano *et al.*, (2010) to account for long memory volatility.

3.11.4 ESTIMATING THE GARCH MODELS

In estimating the parameters of the GARCH models, Shabani *et al.*, (2016) compared the maximum likelihood estimator (MLE) with the generalized method of moments (GMM). The authors recommended the use of MLE as it outperformed the GMM estimator.

The MLE model is expressed as follows:

$$L^*=\prod_{n=1}^{N} f(y_n|\, y_{n-1}, y_{n-2}, \dots\dots y_{1,}\, \theta_1, \theta_{2\dots\dots}\theta_k) \tag{3.47}$$

Where L is the likelihood function,$f()$ is the probability density function

y_n is the time series value at time *n*

$y_{n-1}, y_{n-2}, \dots\dots y_{1,}$ is the time series values at time *n*

$\theta_1, \theta_2, \dots . \theta_k$ are the parameters of the model.

The study estimated the GARCH, EGARCH, GJR-GARCH, FIGARCH models using the maximum likelihood estimator in equation 3.47.

3.12 ERROR DISTRIBUTION

A normal distribution is described by its variance and mean. The density function is given by:

$$f(x) = \frac{e^{\frac{-0.5(x-\mu)^2}{\sigma^2}}}{\sigma\sqrt{2\pi}} \tag{3.48}$$

Where x is a random variable, the mean and variance are denoted by $\mu \; and \; \sigma^2$

Following a whitening process the residual ε, standized by σ yields the standard normal density expressed as:

$$f\left(\frac{x-\mu}{\sigma}\right)=\frac{1}{\sigma}f(z)=\frac{1}{\sigma}\left(\frac{e^{-0.5z^2}}{\sqrt{2\pi}}\right) \tag{3.49}$$

With a normal distribution, skewness = 0 with no excess kurtosis (Ghalanos, 2018).

The student *t* is an alternate to the normal distribution. It is designated by the parameter v. For normalisation, its 3 parameter is given as follows:

$$f(x)=\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{\beta v}\Gamma\left(\frac{v}{2}\right)}\left(1+\frac{(x-\alpha)^2}{\beta_v}\right)^{-\left(\frac{v+1}{2}\right)} \tag{3.50}$$

Where α =location, $\beta =$ scale and $v =$ the shape parameter. The Gamma function is symbolized as Γ.

The variance is written as:

$$Var(x)=\frac{\beta_v}{(v-2)} \tag{3.51}$$

For nominalization process it is required that:

$$Var(x)=\frac{\beta_v}{(v-2)}=1 \tag{3.52}$$

$$\therefore \beta=\frac{v-2}{v}$$

Replacing $\left(\frac{v-2}{v}\right)$ into equation 3.51 we have:

$$f\left(\frac{x-\mu}{\sigma}\right)=\frac{1}{\sigma}f(z)=\frac{1}{\sigma}\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{(v-2)\pi\Gamma\left(\frac{v}{2}\right)}}\left(1+\frac{z^2}{(v-2)}\right)^{-\left(\frac{v+1}{2}\right)} \tag{3.53}$$

With the student t distribution, the skewness is zero with excess kurtosis = $6/(v-4)$ for v>4 (Ghalanos, 2018).

The skewed student t distribution is written as:

$$f(z;\mu,\sigma,v,\lambda)=\begin{cases} bc(1+\frac{1}{v-2}\left(b\left(\frac{z-\mu}{\sigma}\right)+a\right)2)^{\frac{-v+1}{2}}, if\ z<\frac{-a}{b} \\ bc(1+\frac{1}{v-2}\left(\frac{b\left(\frac{z-\mu}{\sigma}\right)}{1+\lambda}2\right)\frac{-v+1}{2}, if\ z\geq\frac{-a}{b}, \end{cases} \tag{3.54}$$

Where the shape parameter is designated by $v\ with\ 2<v<\infty$ and the skewness parameter is denoted by $\lambda\ with -1<\lambda<1$. The constants (a, b and c) are expressed as follows:

$$a = 4\lambda c\left(\frac{v-2}{v-1}\right), b = 1 + 3\lambda^2 - a^2, c$$

$$=\Gamma\left(\frac{\frac{v+1}{2}}{\sqrt{\pi(v-2)\Gamma(\frac{v}{2})}}\right) \tag{3.55}$$

The σ^2 and μ are represented by the variance and mean respectively of the skewed student *t* distribution (Kosapattarapim *et al.*, 2012).

Kosapattarapim *et al.*, (2012) further detailed the generalized error distribution as:

$$f(z,\mu,\sigma,v)=\frac{\sigma^{-1}ve^{(-0.5|(\frac{z-\mu}{\sigma})/\lambda|^{v})}}{\lambda2^{\left(1+(\frac{1}{v})\right)}\Gamma(\frac{1}{4})}, 1 < z < \infty \tag{3.56}$$

$v > 0$ is the tail thickness parameter

And $\lambda = \sqrt{2^{\left(-\frac{2}{v}\right)}\Gamma(\frac{1}{v})\Gamma(\frac{3}{4})}$

If the value of $v = 2$ the GED yields a normal distribution, if $v > 2$ it has thin tails. If $v < 1$ it has thicker tails (Kosapattarapim *et al.*, 2012).

Atoi (2014) highlighted that it is important to use alternative distribution of error as this helps in achieving a robust forecasting model. In view of this, the study compared four error distributions namely skewed student *t*, normal, student *t* and the GED to find the best distribution that adequately fits the data.

3.13 MODEL SELECTION CRITERIA

The study used the Akaike Information Criterion (AIC) and Schwarz Information criterion (SIC) to select the best models amongst the candidate models.

The AIC and SIC are discussed below:

The AIC criterion is written as:

$$AIC = e^{2k/n}\frac{\sum \hat{u}_i^2}{n} = e^{2k/n}\frac{RSS}{n} \tag{3.57}$$

Where the number of regressors is designated by *k* and the number of observations is represented by *n*.

Equation 3.57 can be expressed as:

$$\ln AIC = \left(\frac{2^k}{n}\right) + \ln\left(\frac{RSS}{n}\right) \quad (3.58)$$

Where $2k/n$ is the penalty factor and $\ln AIC$ represents the natural log of AIC. When evaluating the performance of two models, the smallest AIC is desired. The Schwarz criterion (SIC) is similar to the AIC.

The SIC is written as:

$$SIC = n^{k/n} \frac{\sum \hat{u}^2}{n} = n^{k/n} \frac{RSS}{n} \quad (3.59)$$

Or in logarithmic form as:

$$\ln SIC = \frac{k}{n} \ln n + \ln\left(\frac{RSS}{n}\right) \quad (3.60)$$

Where the penalty factor is denoted by [(k/n)$\ln n$]. The SIC inflicts a stricter penalty as compared to the AIC and this is quite evident in equation (3.58) and (3.60). The advantages of using the AIC and SIC is that they can both be used for in and out of sample forecasting (Gujarati and Porter, 2008).

3.14 ACCURACY MEASURES

When comparing models in quest of finding the best model, accuracy measures such as the MSE, RMSE, MAE, MAPE have been commonly used in literature (Ahmed & Shabri, 2014; Będowska-Sójka & Kliber, 2010; Bezerra & Albuquerque, 2017).

3.14.1 MEAN SQUARE ERROR (MSE)

The MSE computes the mean square errors. It is used to assess the performance of the estimator. It is also useful in bias, precision and accuracy of the statistical estimator (SAS Institute Inc., 2011).

Following Poon (2005) the MSE is expressed as:

$$\text{MSE} = \frac{1}{n}\sum_{t=1}^{n} \varepsilon_t^2 \quad (3.61)$$

The number of observations is represented by n, ε is the error term and σ is represents the variance.

3.14.2 MEAN ABSOLUTE ERROR (MAE)

The MAE computes the arithmetic mean of absolute errors. One disadvantage of using the MAE is that it produces biased results when outliers are present in the data set (Chen *et al.*, 2017).

Following Chokri Slim (2015) the MAE is written as:

$$\text{MAE} \ = \frac{1}{n}\sum_{i=1}^{n} | a_i - p_i | \frac{1}{n}\sum_{i=1}^{n} | a_i - p_i | \qquad (3.62)$$

Where a_i is the actual value and p_i is the predicted value.

3.14.3 MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)

The MAPE calculates the average sum of percentage errors. Chen *et al.*, (2017) stated that when the target time series is zero or has values that are close to zero, it would result in excessively large percentage errors.

As noted by Swanson *et al.*, (2011) the MAPE is computed as:

$$\text{MAPE} = \frac{(\sum_{t=1}^{n} | PE |)}{n} \qquad (3.63)$$

Where PE is the percentage error.

3.14.4 ROOT MEAN SQUARE ERROR (RMSE)

The RMSE computes the square root of mean square errors. The RMSE measures the uncertainty in prediction. If the RMSE value is close to zero it simply means that the model has a good forecasting performance (Kavuncuoglu *et al.*, 2018).

Lim and Sek, (2013) expressed the RMSE is as:

$$\text{RMSE} = \sqrt{\sum_{t}^{n} \frac{\varepsilon_t^2}{n}} \qquad (3.64)$$

In selection of the best model, the study used the MSE and RMSE to check the precision and accuracy of the model. The study did not use the MAPE and MAE due to the drawbacks discussed above.

3.15 FORECASTING

In predicting the conditional variance of GARCH, EGARCH, GJR-GARCH and FIGARCH models, the k step ahead forecasts are presented in equation (3.65), (3.66), (3.67) and (3.68) respectively. They are written as follows:

GARCH (p,q)

$$\sigma_{h+k}^2 = \alpha_0 + \alpha_1 a_h^2 + \beta_1 \sigma_h^2 \tag{3.65}$$

EGARCH (p,q)

$$\log(\sigma_{t+k}^2) = \beta_0 + \sum_{i=1}^{q} \{\alpha_i |Z| + \gamma_i (Z)\} + \sum_{j=1}^{p} \beta_j \log(\sigma_{t-j}^2) \tag{3.66}$$

GJR-GARCH (p,q)

$$\sigma_{t+k}^d = \omega + \sum_{i=1}^{p} \alpha_i (u_{t-i} + \beta u_{t-i})^d + \sum_{j=1}^{q} \beta_i \sigma_{t-j}^d \tag{3.67}$$

FIGARCH(1,d,1)

$$\sigma_{t+k}^2 = \omega + [1 - \beta(L) - \emptyset(L)\,(1 - L)^d\,]\, x_t^2 + \beta \sigma_{t-1}^2 \tag{3.68}$$

The next Section presents the Support vector regression.

3.16 SUPPORT VECTOR REGRESSION (SVR)

Vapnik and Chervonenkis (1974) developed a machine learning procedure used for regression and classification known as the support vector machines. When the SVM is applied to linear and nonlinear regression it is known as a SVR (Bezerra & Albuquerque, 2017; Chokri Slim, 2015). Following Bezerra and Albuquerque (2017) The non-linear SVR is as follows: Consider a training data $(x_1, y_1), \dots, (x_n, y_n)$ where the input vector is $x_i \epsilon\, X \subseteq \mathbb{R}$ and the output scaler is $y_i \epsilon Y \subseteq \mathbb{R}$. The main objective of support vector regression is to locate an $f(x)$ function that can estimate the output scalar y_i smaller than a prediction error. In order to obtain it, the SVR nonlinear plats the $\mathbb{R}^n$ input vector space into a much bigger dimension feature space denoted by $(\mathcal{F})$, where the input space of the non-linear relations is estimated by using a linear function:

$$f(x) = \omega^T \phi(x) + b, \text{ with } \phi: \mathbb{R}^n \rightarrow \mathcal{F}, \omega \in \mathcal{F} \tag{3.69}$$

Where $\omega \ and\ b$ denote the regression parametrs, $\phi\,(.)$ denotes the non-linear map function, in which when the equation (3.69) is defined, it projects the input vector in to a bigger dimensional space. The ϵ insensitive loss function is denoted by L_ϵ it measures variance between the actual and the forecast values. The $\epsilon - SVR's$ objective is to locate a function with the minimum ϵ deviance from y_i. The b and w vector can be identified by curtailing the R (C) function:

$$Minimize: R\ (C) = \frac{1}{2} \|w\|^2 + \frac{C}{n} \textstyle\sum_{i=1}^{n} (L_\epsilon(f(y_i, f(x_i)); \tag{3.70}$$

Where:

$$L_\epsilon(y, f(x)) = \begin{cases} |y_i - f(x_i)| - \epsilon, & \text{if } |y_i - f(x_i)| > \epsilon \\ 0, & otherwise \end{cases} \quad ,\epsilon \geq 0 \tag{3.71}$$

is the $\epsilon -$ insensitive loss function (L_ϵ). The observations that are found on or outside the $\epsilon -$ insensitive are the only ones that function as support vectors to create $(f(x))$ function. To specify the errors that are exterior the $\epsilon -$ *insensitive*, slack variables $(\xi_i \xi_i^*)$, where $i = 1,2, \dots . n)$ are introduced. The original problem of SVR is given as follows:

$$Minimize : \frac{1}{2} \|w\|^2 + C \textstyle\sum_{i=1}^{n} (\xi_i + \xi_i^*), \tag{3.72}$$

$$\text{s.t}\begin{cases} y - w^T\phi(x) - b \leq \epsilon + \ \xi_i \\ w^T\phi(x) + b - y \leq \epsilon + \xi_i^*, \\ \qquad \xi_i\xi_i^* \geq 0 \end{cases}$$

The term $\frac{1}{2}||w||^2$ denotes model intricacy. The parameter C is the adjustment between the training error and the function complexity $\sum_{i=1}^{n}(\xi_i + \xi_i^*)$. ϵ parameter regulates the width of the $\epsilon -$ insensitive. If ϵ is greater it requires less support vectors. C and ϵ are the SVR parameters determined by cross validation. Equation (3.70) can be simplified using the Karush-Kuhn-Tucker conditions and Lagrangian multipliers; by converting equation (3.70) into a dual problem we have:

$$Maximize : L = -\frac{1}{2}\sum_{i=1}^{n}(\alpha_i - \alpha_i^*)(\alpha_j^* - \alpha_j)\langle\phi(x_i), \phi(x)\rangle + \sum_{i=1}^{n} y_i(\alpha_i - \alpha_1^*) - \epsilon\sum_{i=1}^{n}(\alpha_i + \alpha_i^*) \tag{3.73}$$

$$\text{s.t}\begin{cases} \sum_{i=1}^{n}(\alpha_i^* - \alpha_i) = 0, \\ 0 \leq \alpha_i \leq C, i = 1, \dots \dots . n \\ 0 \leq \alpha_i^* \leq C = 1, \dots \dots . n \end{cases}$$

From the solution of equation (3.73) the $\epsilon - SVR$ is given as:

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\langle\phi(x_i), \phi(x) + b \tag{3.74}$$

Where $\langle\phi(x_i), \phi(x)$ denotes the dot feature in $\mathcal{F}$. the kernel function can be replaced by the dot function as:

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)K(x_i, x) + b \tag{3.75}$$

With kernel function $K(x.x') = \langle\phi(x')\rangle$ is critical to the performance of the SVR (Bezerra and Albuquerque, 2017).

3.16.1 KERNEL FUNCTIONS

One major problem in the classification procedure is that when the data is scattered, it becomes difficult to linearly detach them (Nanda *et al.*, 2018). A kernel function converts data into a greater dimensional space known as the kernel space; after the conversion, the data can now be linearly detachable. A linear hyperplane can be used to separate dissimilar classes in a kernel space. A kernel function is given as:

$$K(u, x) = \sum_r \varphi_r(x)\,\varphi_r(u), \tag{3.76}$$

Where $\varphi(x)$ belongs to the Hilbert space. Alternatively $\iint K(x,u)g(x)g(u)dxdu \geq 0 \, \forall_g(x) where \int g^2(x)dx < +\infty$ (Awad & Khanna, 2015).

Awad and Khanna (2015) outlined some of the commonly used kernels as:

Table 3.2: Kernels

Linear kernel	$K(x,u) = x^T.u$
Polynomial function	$K(x,u) = (ax^Tu + c)^q, q > 0$
Hyperbolic tangent (sigmoid)	$K(x,u) = \tanh(\beta x^T u + \gamma)$
Gaussian radial basis function	$K(x,u) = \exp\left(-\frac{\lVert x-u \rVert^2}{\sigma^2}\right)$
Laplacian radial basis	$K(x,u) = exp\left(-\frac{\lVert x-u \rVert}{\sigma}\right)$
Randomized blocks analysis of variance	$K(x,u) = \sum_{k=1}^{n} \exp(-\sigma(x^k - u^k)^2)^d$
Linear spline kernel in 1D	$K(x,u) = 1 + x.u.\min(x,u) - \frac{x+u}{2}(\min(x,u)^2 + \frac{1}{3}\min(x,u)^3)$

The selection of kernel is reliant on the data specifics. The simplest kernel is the linear kernel, it is very suitable in big sparse data vectors. For image processing the polynomial kernel is suitable. The hyperbolic tangent (sigmoid) is ideal for neural networks (NN). The Randomized blocks analysis of variance (ANOVA) is for regression analysis. The Gaussian kernel comprises of the sigmoid and the linear kernel by imposing limits on the penal parameter. When a kernel matrix is diagonal, it means that the feature space is redundant and that a new kernel should be used after reduction of feature (Awad & Khanna, 2015; Li, 2014).

3.16.2 SVR-GARCH MODELLING

Figure 3.2 illustrates the SVR-GARCH modelling process employed in the study. The hybrid SVR-GARCH models have been found to yield accurate forecasts as compared to the GARCH (Chung & Zhang, 2017; Geng & Zhang, 2015; Ou & Wang, 2010). This study used four GARCH modes for the Support Vector regression namely GARCH(1,1), GJR-GARCH(1,1), EGARCH(1,1) and FIGARCH(1,d,1) to account for symmetric, assymetric and long memory volatility respectively. The next section presents Extreme value theory.

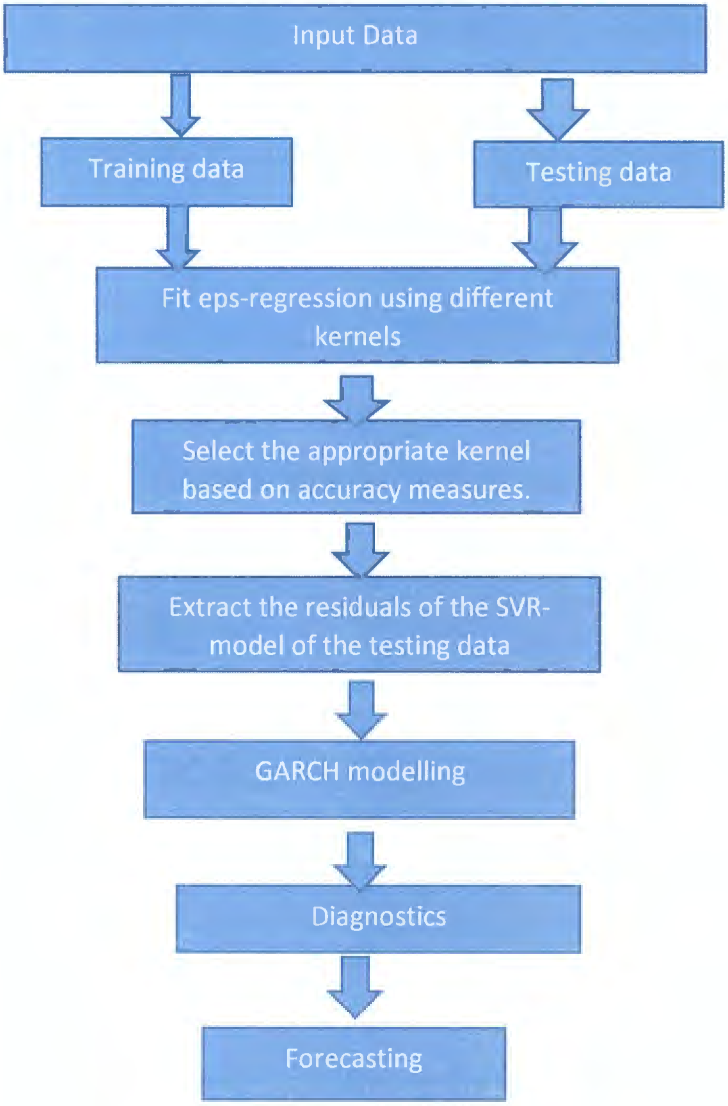


Figure 3.2: SVR-GARCH PROCESS

3.17 EXTREME VALUE THEORY

This Section presents EVT using the POT method as shown in Figure 3.3. Firstly the residuals of the best GARCH and SVR-GARCH models will be extracted, standardized and then modelled using Extreme value theory.

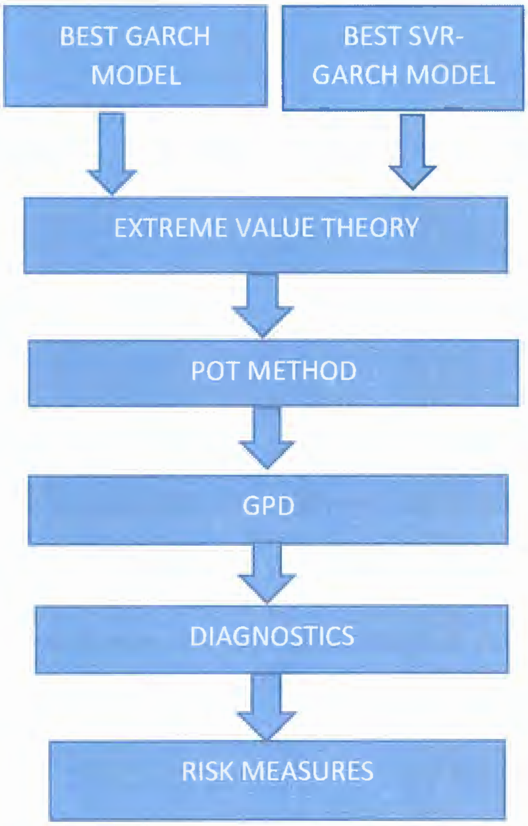


Figure 3.3: EVT Process

EVT handles rare extreme events. It can be used to describe fat tails of profit or loss distribution (Rhoda, 2013). Let's assume x_t represents losses and it is serially correlated with a cumulative distribution function (CDF) F(x) and the return range of x_t is $[l, u]$. The CDF of $x_{(n)}$ designated by $F_{n,n}(x)$ is written as:

$$F_{n,n}(x) = \Pr[x_{(n)} \leq x]$$

$$= \Pr(x_1 \leq x, x_2 \leq x, \dots \dots x_n \leq x)$$

$$= \prod_{j=1}^{n} \Pr(x_j \leq x)$$

$$= \prod_{j=1}^{n} f(x) = [F(x)]^n \tag{3.77}$$

The CDF $F(x)$ is not known hence $F_{n,n}(x)$ of $x_{(n)}$ is unknown. If n increases to infinity, $F_{n,n}(x)$ of $x_{(n)}$ becomes degenerated: $F_{n,n}(x) \rightarrow 0$ if $x < u$ and $F_{n,n}(x) \rightarrow 1\ if\ x \geq u$ as $n \rightarrow \infty$. The degenerated CDF is of no use, therefore the main purpose of the EVT is to find two sequences $\{\mu_n\}$ and $\{\sigma_n\}$ where $\sigma_n > 0$ so that the distribution of $x^*_{(n)} \equiv (x_{(n)} - \mu_n)/\sigma_n$ converges to a nondegenerate distribution as $n \rightarrow \infty$. $\{\mu_n\}$ and $\{\sigma_n\}$ are location and scaling series. The limiting distribution of $r^*_{(n)}$ is as follows:

$$F_*(x)=\begin{cases} \exp[-(1+\xi x)^{-\frac{1}{\xi}}] & \text{if } \xi \neq 0 \\ \exp[-\exp(-x)] & \text{if } \xi = 0 \end{cases} \tag{3.78}$$

For $x < -1/\xi$ if $\xi < 0$ and for $x > -\frac{1}{\xi} if\ \xi > 0$, where normalized maximum is denoted by * . when $\xi \rightarrow 0$, $\xi = 0$ is taken as the limit. ξ controls the tail behaviour of the limiting distribution. Equation 3.78 is called the generalized extreme value. It includes three limiting distributions namely the Gumbel, Frechet and Weibull family. The distributions are listed in equation 3.79, 3.80 and 3.81.

Type I: Gumbel: $\xi = 0$, with a CDF function of

$$F_*(x) = \exp[-\exp(-x)], \ -\infty < x < \infty. \tag{3.79}$$

Type II: Frechet: $\xi > 0$, with a CDF of

$$F_*(x) = \begin{cases} \exp[-(1+\xi x)^{-1/\xi} & if\ x > -\frac{1}{\xi}, \\ 0 & otherwise \end{cases} \tag{3.80}$$

Type III: Weibull: $\xi < 0$, The CDF function is as follows:

$$F_*(x) = \begin{cases} \exp[-(1+\xi x)^{-1/\xi} & if\ x < -\frac{1}{\xi}, \\ 1 & otherwise \end{cases} \tag{3.81}$$

The tail behaviour of $f(x)$ governs the limiting distribution of $F_*(x)$. The right tail of the Gumbel (includes the normal and log normal distributions) decrease exponentially, it is finite with the Weibull and a power law for the Frechet (Tsay, 2013).

3.17.1 THE PEAK OVER THRESHOLD METHOD

There are two methods that fall under EVT namely the BMM and the POT method. The BMM does not make sufficient use of information provided by the data whereas the POT method makes use of data more proficiently (Dicks *et al.*, 2014). In addition, when dealing with a univariate time series data, specifically with stylized facts; the volatility clustering of the data points are selected as BMM while they are not, now this leads to estimation biasness (Pfaff, 2016). The POT method produces conclusive results unlike the BMM. Due to the drawbacks of the BMM method, the study modelled the tails using the POT method.

The Peak over threshold is given by:

$$P\{X > u + y|X > u\} = \frac{1-F(u+y)}{1-f(u)}, y > 0 \tag{3.82}$$

When F is known, the expression can be solved for the exceedances $y > u$.

The main purpose of the of the POT is to construct a tail estimator (Li, 2015). A challenge of the POT method is the threshold selection; following Susan and Waititu (2015) the study will use the mean excess plot to find the appropriate threshold.

3.17.2 GENERALIZED PARETO DISTRIBUTION

Following Halder *et al.*, (2016) let's consider X_1,X_2, ... X_n to be i.i.d random variable with a marginal distribution function F *and* M_n = $\max\{X_1, \dots, X_n\}$. The extreme event X_i that exceeds the threshold is designated by u. The random behaviour of extreme events is represented by the conditional probability

$$P(X > x + u|X > u)=\frac{1-F(x+u)}{1-F(u)}, \text{ where } x > 0, \tag{3.83}$$

If any random term in the X_i is X with F distribution function for large sample size n

$$P(M_n) \leq x) \approx G(x)^n \tag{3.84}$$

Where $G(x) = exp\left(-(1 + k\frac{x-\mu}{\sigma})^{-1/k}\right)$ (3.85)

For some $\mu, \sigma > 0$ and k, $\mathit{for\ large\ threshold\ value\ u, the\ distribution\ function\ of\ (X - u), conditional\ on\ X > u}$, is $\mathit{described\ by\ the\ generalized\ pareto\ family}$. The Generalized Pareto has three parameters namely the location, shape and scale parameter denoted by μ, k, σ respectively. The CDF of a Generalized Pareto of a RV is written as:

$$G(x) = \begin{cases} 1 - (1 + \frac{k(x-\mu)^{-\frac{1}{k}}}{\sigma}, & k \neq 0, \sigma > 0 \\ 1 - \exp - \left(\frac{x-\mu}{\sigma}\right), & k = 0, \sigma > 0 \end{cases} \tag{3.86}$$

Equation (3.83) decreases to 2 parameter GPD for $\mu = 0$ and in most cases a 2 GPD parameter seems more suitable than a 3 GPD. The shape parameter of the GPD produces different distributions: when $k < 0$ it displays a heavy pareto-type upper tail, for $k = 0$, it yields an exponential distribution with mean σ, and when $k = 0.5$ *and* $k = 1$ it shows a triangular and uniform distribution. When $k \le -\frac{1}{2}, Var(X) = \infty\ the\ rth\ central\ exists\ if\ and\ only\ if\ k > -\frac{1}{r}$ (Halder *et al.*, 2016).

3.17.3 MEAN EXCESS PLOT

The mean excess plot helps in selecting the threshold. It is also used to decide the suitability of the GPD model. The mean excess function is expressed as:

$$M(u) = E(X - u|X > u) \tag{3.87}$$

And the mean excess function (MEF) of the GPD is as follows:

$$M(u) = \frac{\sigma(u)}{1-k} = \frac{\sigma + ku}{1-k} \tag{3.88}$$

Where $0 \le u < \infty\ when\ 0 \le k \le 1\ and\ 0 \le u \le -\frac{\sigma}{k}\ when\ k < 0$. The $M(u)$ is non-existent when $k > 1$ (Halder *et al.*, 2016).

3.17.4 MAXIMUM LIKELIHOOD ESTIMATOR (MLE)

In estimating the GPD, Nortey *et al.*, (2015) compared the MLE with the probability weighted moments (PWM). The study found the MLE to produce more accurate results than the PWM. Based on the results, the current study used the MLE to estimate the parameters.

In the MLE method, the probability density function (pdf) of $f(x)$ is often not known but it is assumed that the joint density function emanates from a known distribution. The sample size n from an iid joint density function is written as:

$$f(x_1, x_2, \dots . x_n|\theta) = f(x_1|\theta).f(x_2|\theta) \dots \dots \cdot f((x_n|\theta) \tag{3.89}$$

Where the model parameters are denoted by θs and x_i represents the observed variables. The parameters of θ are to be estimated and the observed variables x_i are known.

The likelihood function is expressed as:

$$L(\theta|x_1 \dots \dots x_n) = f(x_1, x_2, \dots . x_n|\theta) = \prod_{i=1}^{n} f(x_1|\theta) \tag{3.90}$$

The log-likelihood function is given as:

$$ln\ \alpha(\vartheta|x_1, \dots . x_n) = \sum_{i=1}^{n} \ln f(x_i|\theta) \tag{3.91}$$

(Kiragu & Kyalo, 2016).

3.18 VALUE AT RISK AND EXPECTED SHORTFALL

Market and credit risks in addition to operational risks are some of the risks found in financial markets. The VaR is primarily concerned with market risk. The VaR is used in financial markets to ensure that after a cataclysmic event occurs, financial markets can still be in business (Tsay, 2010). Following Aboura (2014) the VaR is written as follows:

$$VaR_q = u + \frac{\beta}{\xi}\left[\left(\frac{n_u}{n}(1-q)\right)^{-\xi} - 1\right] \tag{3.92}$$

The VaR model has been carped due to its fractional shortfall. The VaR can be disingenuous during volatile phases. It also disdains any loss beyond VaR level. The Expected shortfall is centred on the loss above VaR level. It also takes for the size of the tail loss into consideration by assessing the expected loss size given the VaR is surpassed. The ES is as follows:

$$ES_q = VaR_q + E[X - VaR_q|X > VaR_q] \tag{3.93}$$

(Aboura, 2014).

3.19 CONCLUDING REMARKS

The Chapter reviewed various techniques such as GARCH, SVR-GARCH and EVT employed in the study. The Chapter discussed the data sources, methods and the software packages that were used to execute the analysis. Preliminary data analysis methods such as normality and unit root tests were reviewed. Some of the Normality tests discussed were skewness, Kurtosis and Shapiro Wilks test. The normality test's aim is to identify if the data is normally distributed or not. For the unit root tests, a detailed description of ADF and PP tests were discussed. These tests aim to test stationarity present in the data as well as to avoid spuriority that may be present in the data set. The study employed four models in modelling volatility namely GARCH (1, 1), FIGARCH (1,d,1), EGARCH(1,1) and GJR-GARCH(1,1). The choice of the models were selected through the help of literature as discussed in Chapter 2. A comprehensive review of GARCH, SVR-GARCH and Extreme value theory was revised. Diagnostic tests were reviewed in order to examine the adequacy of the models. Model selection methods such as AIC, SIC, MSE and RMSE were reviewed with the aim of selecting the best candidate model. The results are presented in the next Chapter.

CHAPTER FOUR

DATA ANALYSIS

4.1 INTRODUCTION

This Chapter presents and analyses the Brent Crude oil prices data obtained from the JSE. The aim of the study was to assess the effectiveness of GARCH and SVR-GARCH models. EVT was further employed fit the tails of the returns. The analysis took into consideration the objectives in Chapter one and the techniques discussed in Chapter three. The analysis was executed using EViews 10 and RStudio 3.5 software packages.

4.2 GRAPHICAL PLOT OF BRENT CRUDE OIL PRICES

Figure 4.1 illustrates the plot of Brent Crude oil prices from the 7th August 2008-7th August 2018 resulting in a total of 2525 observations. There was a substantial increase in crude oil prices in 2011, ever since then the prices have been fluctuating. By visual inspection, it is evident that the prices are not stationary with time; this implies that the variance and mean are not consistent with time. Some transformation is needed to make the prices stationary.

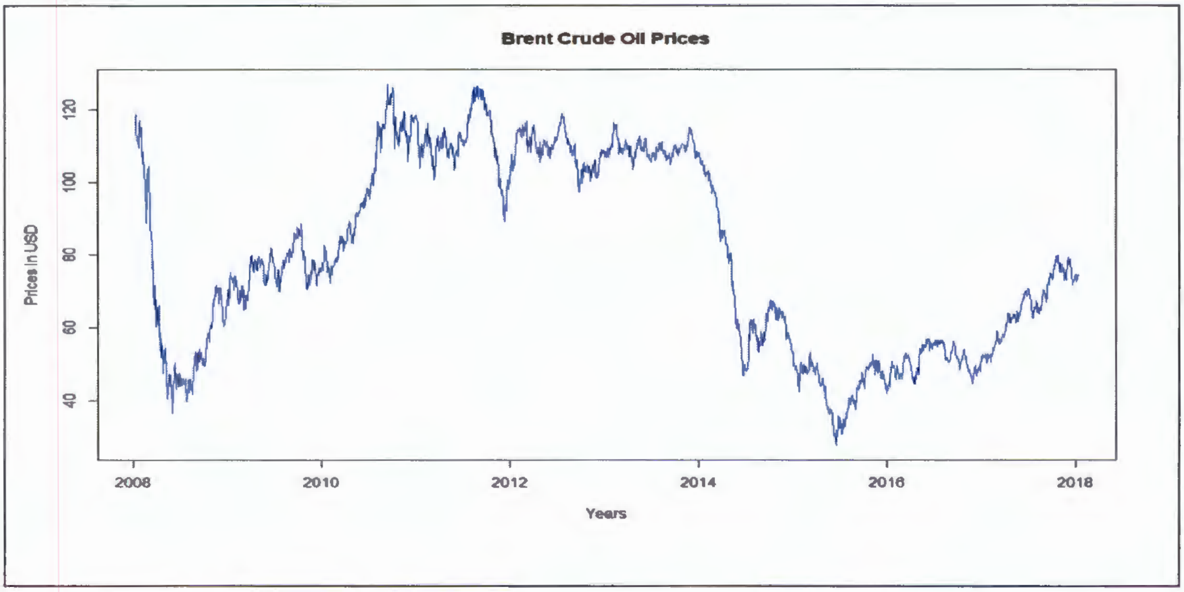


Figure 4.1: Plot of Brent Crude oil Prices

4.2.1 GRAPHICAL PLOT OF THE RETURNS

Figure 4.2 illustrates the returns of Brent crude oil prices. The returns were computed as Log $(\frac{P_a}{P_{a-1}})$ thus the natural log of the difference between the Brent crude oil prices. Where P_a is denoted by the current price and P_{a-1} is the preceding price. The log returns were used for the analysis because they have more 'tractable statistical properties" and the data is multiplicative (Gaston, 2016). By visual inspection it evident that the mean and variance of the return series are constant as it is fluctuating around zero. A positive and a negative spike designate a gain and a loss. Figure 4.2 further displays volatility clustering, as large variations in oil prices cluster together. This confirms the volatility of Brent crude oil prices in South Africa. This is in accord with the studies of Atoi (2014).

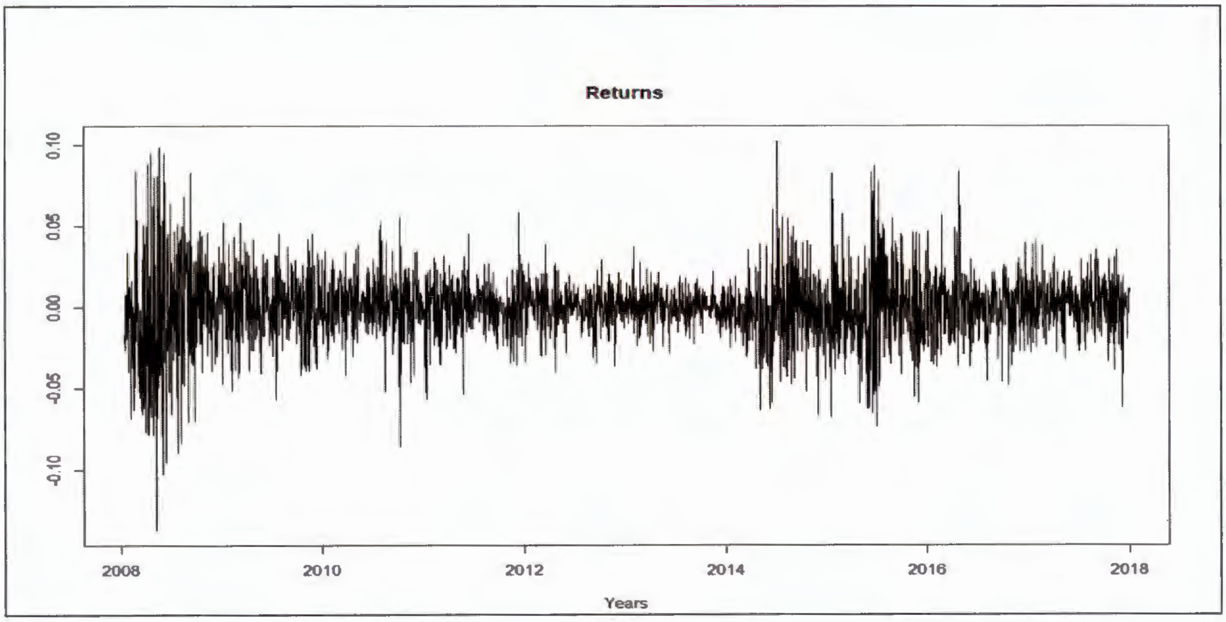


Figure 4.2: Returns plot

4.3 DESCRIPTIVE STATISTICS

The number of observations for the Brent crude oil prices were 2525 (Table 4.1). The average mean of the prices is 80.01 USD. The standard deviation is 26.47 which shows how the observations are from the sample mean. The minimum and maximum prices are 27.88 and 126.9 USD respectively. To check for normality, skewness, kurtosis, Jarque-Bera and Shapiro Wilks test were considered in this study. Skewness and kurtosis of a normal distribution is 0 and 3 respectively. For the JB and Shapiro Wilks test, if the probability value is smaller than 5%, the null hypothesis for normality is rejected. The value of skewness is 0.03, which is very close to 0, and a kurtosis of -1.45 this implies the non-normality of the data. The number of observation for the returns are 2524 (see Table 4.1) because one observation was lost due to the log transformations. The average mean and median was 0. The minimum and maximum values are -0.14 and 0.1 respectively. The skewness value of -0.08, indicates that the distribution is skewed to the left. The value of kurtosis is 3.32; meaning that the distribution is peaked. The JB test and the Shapiro-Wilk's test statistic further confirms non-normality as the p-value was less than 5%.

Table 4.1: Descriptive Statistics

Measure	**Brent Crude oil Prices**	**Returns**
n	2525	2524
mean	80.01	0
Standard deviation	26.47	0.02
median	76.72	0
min	27.88	-0.14
max	126.9	0.1
Range	99.02	0.24
Skewness	0.03	-0.08
Kurtosis	-1.45	3.32
Standard error	0.53	0
Jarque Bera	220.11(0.000000)	1167.8(0.000000)
Shapiro- Wilks	0.91798 (0.0000)	0.95692(0.00000)

4.4 UNIT ROOT TEST

Formal unit root tests were carried out using the ADF test and the PP test for the Brent Crude oil and the returns (see Table 4.2). At levels the ADF and the PP test statistic for Brent crude oil was -1.780349 and -1.766435 and their critical values were -2.862485 and -2.862485 respectively. Both the ADF and PP test statistics were greater than the critical values at 5% significance level. The probability value further confirms non stationarity as it is insignificant with the value of 0.3907 and 0.3976 for the ADF and PP tests respectively. This leads to the acceptance of the null hypothesis at levels that Brent crude oil has a unit root. At first difference, Brent crude oil prices was found to be stationary for the ADF and PP as the test statistic -58.82850 and -52.82850 respectively were less than the critical values with -2.862485 and -2.862485. The probability value further confirms it with a value of 0.0001. The null hypothesis is then rejected as Brent crude oil is found stationary at first difference. The returns test statistic at levels for the ADF and PP were -52.58429 and -52.55190 with a critical value of -2.862485 and -2.862485 at 5%. The test statistic for the ADF and PP are less than the critical values. The probability values are 0.0001 leads to the rejection of the null hypothesis as the returns are stationary at levels.

Table 4.2: ADF and PP unit root tests

Variable	Unit Root test	Level statistic	Critical value	P-value	First difference statistic	Critical value	P-value	Decision
Brent Crude oil	ADF	-1.780349	-2,862485	0.3907	-52.82850	-2,8625**	0.0001	I(1)
	PP	-1.766435	-2,862485	0.3976	-52.82850	-2,8625**	0.0001	
Returns	ADF	-52.58429	-2,862485**	0.0001				I(0)
	PP	-52.55190	-2,862485**	0.0001				

Note ** indicates at significant 5%

4.5 NORMALITY PLOTS

This Section presents the Histogram, QQ plots and the Kernel density function of Brent crude oil and the returns.

4.5.1 HISTOGRAM

The histograms of Brent crude oil prices and the Returns are exhibited in Figure 4.3(a) and 4.3(b). The histogram of the Brent crude oil prices is under dispersed confirming the non-normality of the data. For the returns (Figure 4.3(b)) the distribution is over dispersed thus having an increase in outliers; the distribution is peaked confirming the non-normality of the data. This is due to the uncertainty of price changes.

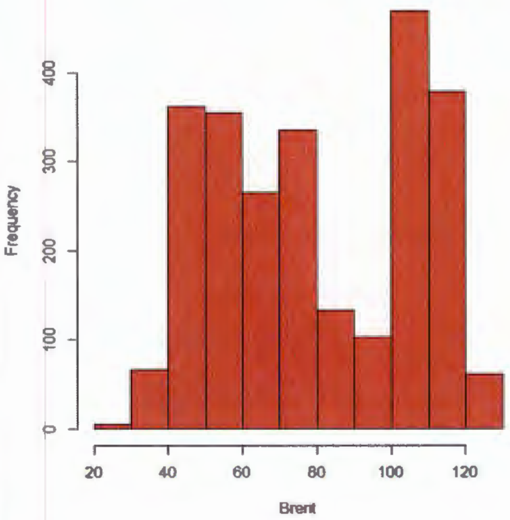


Figure 4.3(a) Brent Crude Oil Prices

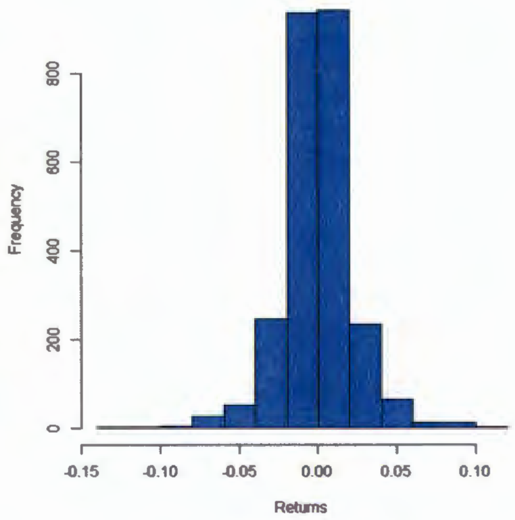


Figure 4.3(b) Returns

4.5.2 QUANTILE-QUANTILE (Q-Q) PLOTS

The Q-Q plots for Brent crude oil prices and the returns are shown in Figure 4.3(c) and 4.3(d). By visual inspection it is evident that the QQ plot of Figure 4.3(c) is under dispersed; and the returns (Figure 4.3(d)) are over dispersed. This confirms the non-normality of the data.

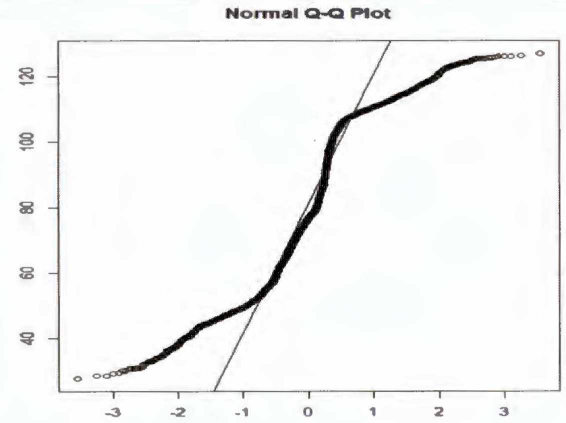


Figure 4.3(c) Brent Crude oil

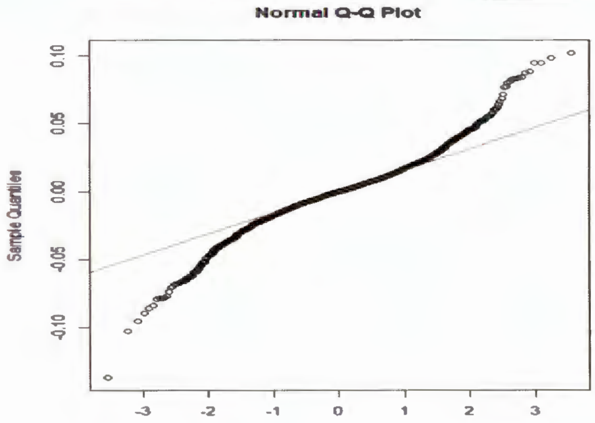


Figure 4.3(d) Returns

4.5.3 KERNEL DENSITY FUNCTION

The Kernel density function plots are displayed in Figure 4.4. The kernel density of Brent crude oil and the returns further confirms non-normality.

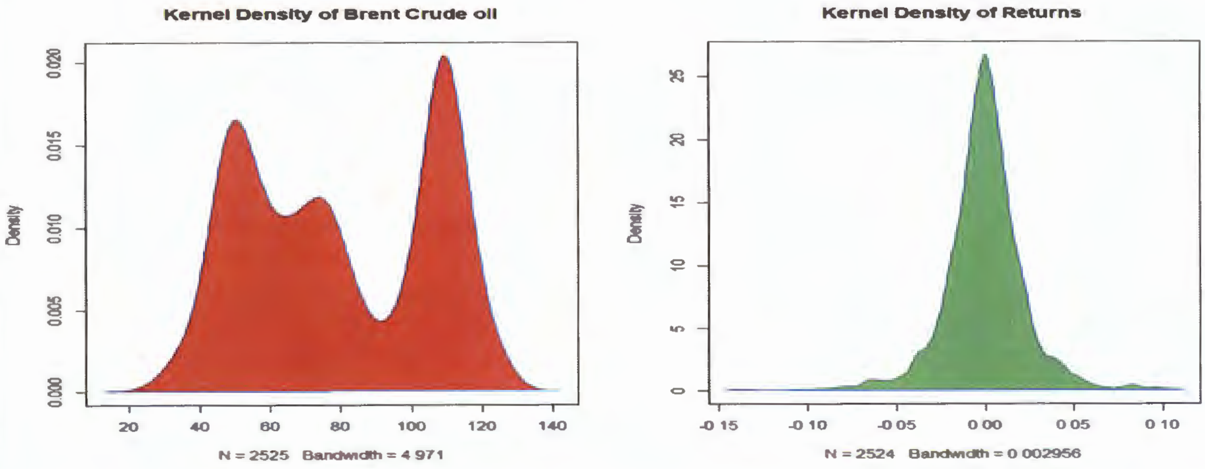


Figure 4.4 Kernel density Plots

4.6 MODEL IDENTIFICATION

The ACF and PACF for the returns are displayed in Figure 4.5(a) and 4.5(b). They both give a clear indication of the model. For the ACF (see Figure 4.5a) at lag 1 only displays one significant spike. This is a clear indication of an MA (1) model. Table 4.3 displays the results for the returns. The chosen model is MA (1) as the coefficient is less than 1. The PACF has a significant spike at lag 1, 10, 12, 26 indicating ARCH effects.

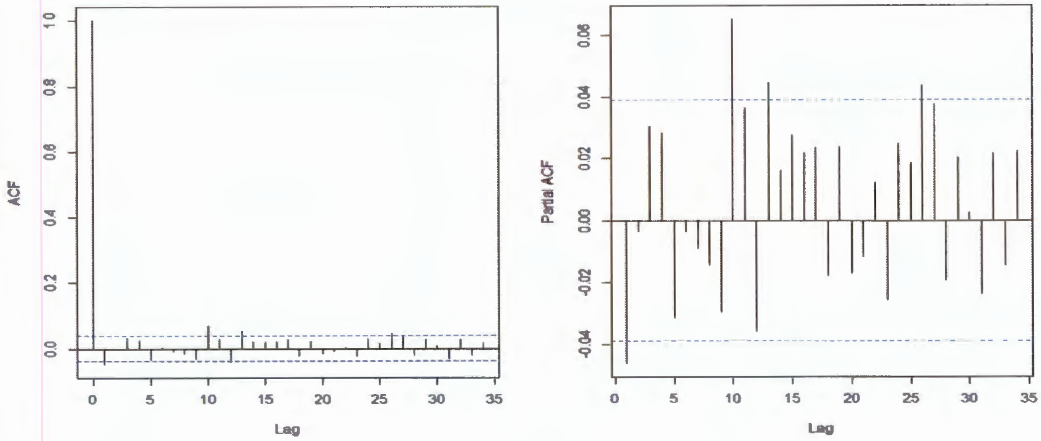


Figure 4.5(a)

Figure 4.5(b)

Figure 4.5: ACF and PACF Plots

Table 4.3: Model Identification

Model	Estimates
MA(1)	-0.0460
Standard error	0.0198
BIC	-12234.52

4.6.1 ARCH EFFECTS

This Section presents the LM and the LB test to identify existence of autocorrelation and heteroscedascity in the residuals of the returns.

4.6.1.1 RESIDUALS

The residual plot of the returns are displayed in Figure 4.6. By visual inspection it can be seen that the error terms are dispersed. This is a clear indication of non-constant error variance.

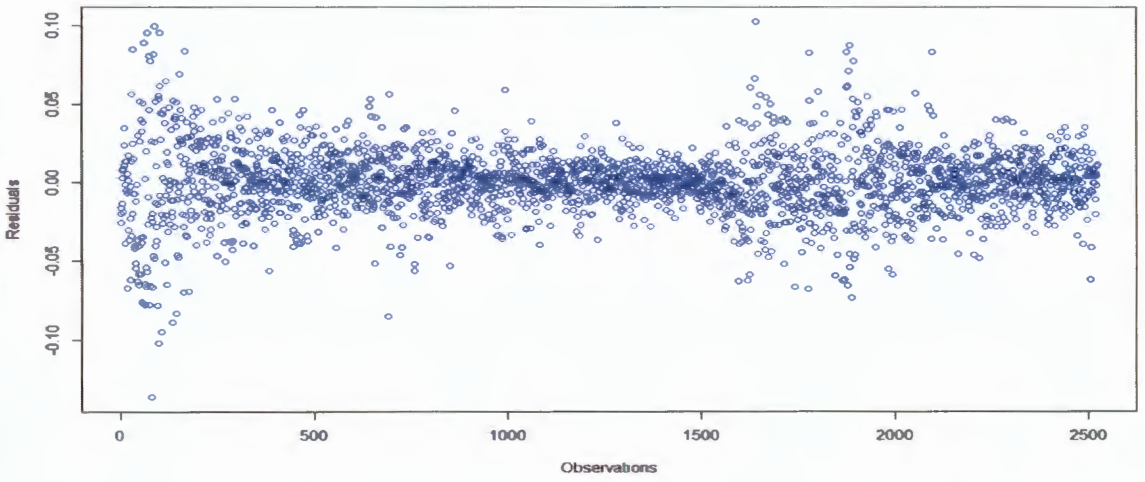


Figure 4.6: Residuals

4.6.2 LJUNG-BOX TEST

The Q statistic for randomness at lag 5 is 12.279 and its probability value is 0.031 as shown in Table 4.4. The $X_{\alpha,n}^2$ =$X_{0.05,5}^2$ = 11.071. The Q_{LB}= 12.279 > 11.071. All the probability values from lag 5-35 are less than 5%. This leads to the rejection of H_0 of no autocorrelation. The alternative hypothesis is accepted as there is autocorrelation present in the returns.

Hypothesis

H_0: No autocorrelation

H_1: Autocorrelation

Table 4.4: Ljung Box test

Lag	AC	PAC	Q-stat	Probability value
5	-0.034	-0.031	12.279	0.031
10	0.068	0.065	26.993	0.003
15	0.021	0.027	42.006	0.000
20	-0.015	-0.017	47.872	0.0001
25	0.016	0.019	52.029	0.0001
30	0.009	0.003	63.938	0.0000
35	0.009	0.009	70.097	0.0000

4.6.3 LAGRANGE MULTIPLIER

Table 4.5 displays the ARCH test results of the residuals of the returns. The large values of the F-statistic and the significant probability values of the chi squared leads to the rejection of the null hypothesis of no ARCH effects at 5% as the probability values are less than 5%.

Hypothesis

H_0: Homoscedasctic

H_1: Heteroskedasticity

Table 4.5: Heteroskedasticity Test

Heteroskedasticity Test: ARCH			
F-statistic	**131.6376**	**Prob.F(1,2521)**	**0.0000**
Obs*R-squared	**125.2043**	**Prob. Chi-Square(1)**	**0.0000**

4.7 GARCH MODELS

The heteroskedascity test (see Table 4.5) indicated that the returns are heteroscedastic as the null hypothesis was rejected at 5% because the probability values were less than 5%. This signifies the presence of ARCH effects hence GARCH models are applicable. The GARCH type models used in this study consist of the asymmetric, symmetric and long memory volatility models. The study uses GARCH (1,1) model because it is simple to use and has been widely recommended in literature (Bouseba and Zeghdoudi, 2015). This study uses GARCH (1, 1) to account for volatility clustering; EGARCH (1, 1), GJR-GARCH (1, 1) to capture leverage effect and FIGARCH (1, d, 1) to capture long memory property of the returns. The GARCH type models are used with MA (1) because in Table 4.3 the generated model displayed an MA (1) model hence this study used an MA (1) model.

4.7.1 DISTRIBUTION OF ERROR

Table 4.6 displays the Student *t*, Normal, Skewed student *t* and the GED distribution in comparison of the best distribution that fits the data. The results of Table 4.6 are based on the AIC because it displayed the least error. The skewed student t distribution best fits the GARCH (1,1) , EGARCH(1,1) , FIGARCH(1,d,1) and GJR-GARCH (1,1). This further confirms that the data is not normally distributed. Specification of these volatility models in a normal distribution will lead to misspecification. The estimation of the volatility models will be based on the skewed student *t* distribution as it displayed the least error.

Table 4.6: Error distributions

	Normal		Std		sStd		GED	
Model	AIC	SIC	AIC	SIC	AIC	SIC	AIC	SIC
GARCH(1,1)	-5,1802	-5,1687	-5,2065	-5,1926	-5,208	-5,1918	-5,2054	-5,1915
GJR-GARCH(1,1)	-5,1945	-5,1806	-5,2183	-5,2021	-5,2197	-5,2012	-5,2161	-5,2
E-GARCH(1,1)	-5,2087	-5,1948	-5,2276	-5,2115	-5,229	-5,2105	-5,226	-5,2098
FIGARCH(1,d,1)	-5,179	-5,1652	-5,2059	-5,1897	-5,2075	-5,189	-5,2046	-5,1884

4.7.2 GARCH (1, 1)

Table 4.7a displays the GARCH (1, 1) model results. Let Ma (1) denote θ mean equation is as follows $rt = -0.000026 - 0.056023$ as displayed in Table 4.7(a). The probability value of $\theta(0.004966\,)$ is significant at 5% further supporting the use of an MA (1) model for the returns. The probability value of β_1 and alpha are significant at 5%. $\hat{\beta}_1$ = 0.939482 indicating persistent volatility clustering. Volatility persistence is the summation of ARCH and GARCH coefficients. The volatility persistence is 0.998102 which is close to 1.

Table 4.7(a): GARCH (1, 1)

		Sstd			
		GARCH(1,1)			
	Estimate	Std. Error	t value	Pr(>|t|)	Parameters significant at 5%
mu	-0.000026	0.000292	-0.088437	0.929529	not significant
ma1	-0.056023	0.019943	-2.809212	0.004966	significant
omega	0.000001	0.000002	0.932890	0.350877	not significant
alpha1	0.058620	0.011574	5.064801	0.000000	significant
beta1	0.939482	0.011939	78.688102	0.000000	significant
skew	0.935079	0.026200	35.690417	0.000000	significant
shape	8.206637	1.170993	7.008272	0.000000	significant

4.7.3 E-GARCH (1, 1)

The skewness statistics in the returns displayed asymmetry in the data. The study compared the E-GARCH (1, 1) and the GJR-GARCH (1, 1) to account for asymmetry. Table 4.7(b) displays the EGARCH (1, 1) model. Gamma1 is positive with the value of 0.087047 and significant at 5%. This implies that good news has more influence on future volatility. Gamma1≠ zero meaning that the news impact is asymmetric. Beta1 (β_1) measures the persistence of conditional volatility; the estimate of β_1 is 0.994 and significant at 5%. The sum of α_1 and β_1 is 0.933762 for volatility persistence.

Table 4.7(b) : EGARCH (1,1)

		EGARCH(1,1)			
	Estimate	Std. Error	t value	Pr(>|t|)	Parameters significant at 5%
mu	-0.000284	0.000280	-1.0145	0.310321	not significant
ma1	-0.047533	0.019894	-2.3893	0.016879	significant
omega	-0.041833	0.001590	-26.3169	0.000000	significant
alpha1	-0.061016	0.008072	-7.5592	0.000000	significant
beta1	0.994778	0.000124	7996.3379	0.000000	significant
gamma1	0.087047	0.002716	32.0473	0.000000	significant
skew	0.935972	0.026508	35.3088	0.000000	significant
shape	9.644532	1.433235	6.7292	0.000000	significant

4.7.4 GJR-GARCH (1, 1)

The GJR-GARCH (1, 1) is presented in Table 4.7(c). The value of gamma1 (γ) ≠ 0 meaning that the news impact is asymmetric. The parameter of γ is significant and positive at 5% this indicates leverage effect. All the parameters are significant except for mu, omega and alpha. The volatility persistence is 0.967002.

Table 4.7(c): GJR-GARCH (1, 1)

	GJR-GARCH(1,1)				
	Estimate	Std. Error	t value	Pr(>|t|)	Parameters significant at 5%
mu	-0.000218	0.000285	-0.76409	0.444812	not significant
ma1	-0.055251	0.018443	-2.99579	0.002737	significant
omega	0.000001	0.000005	0.23107	0.817261	not significant
alpha1	0.016222	0.018629	0.87077	0.383880	not significant
beta1	0.950780	0.030125	31.56105	0.000000	significant
gamma1	0.062337	0.018279	3.41027	0.000649	significant
skew	0.935506	0.026348	35.50555	0.000000	significant
shape	8.806508	1.579956	5.57389	0.000000	significant

4.7.5 FIGARCH (1, d, 1)

Table 4.7(d) displays the FIGARCH (1, d, 1) model. The delta estimate is positive and significant at 5%. All the parameters are significant except for alpha and mu. Volatility persistence is 0.97032.

Table 4.7(d): FIGARCH (1, d, 1)

	FIGARCH(1,d,1)				
	Estimate	Std. Error	t value	Pr(>|t|)	Parameters significant at 5%
mu	-0.000030	0.000292	-0.10164	0.919046	not significant
ma1	-0.057144	0.020333	-2.81044	0.004947	significant
omega	0.000001	0.000000	4.48499	0.000007	significant
alpha1	0.027352	0.039519	0.69213	0.488857	not significant
beta1	0.942968	0.005797	162.67584	0.000000	significant
delta	0.999999	0.025045	39.92830	0.000000	significant
skew	0.933698	0.026387	35.38455	0.000000	significant
shape	7.984035	1.133255	7.04522	0.000000	significant

4.8 DIAGNOSTICS

This Section provides the diagnostics of GARCH (1, 1), GJR-GARCH (1, 1), EGARCH (1, 1) and FIGARCH (1, d, 1). The probability values of the weighted Ljung box standardized residuals and squared residuals (see Table 4.8) are all more than 5% indicating the absence of autocorrelation. This implies that the estimated model for the variance and conditional mean are satisfactory for removing autocorrelation from the data. Hence the null hypothesis of no autocorrelation is accepted. The probability values for the ARCH test were all greater than 5% indicating that the error terms are homoscedastic. The null hypothesis of homoscedastic is then accepted.

4.9 NORMALITY PLOTS

The normality plots are displayed in Figure 4.7(a-d). For the QQ-plots of the residuals of all the four models, only a few points fell outside the line. The empirical distribution plots shows how the error distributions fitted the data. The News impact curve (NIC) of GARCH (1, 1) is symmetrical meaning that good and bad news reacts the same. This is in accord with Atoi (2014). For the NIC, the asymmetric models (EGARCH and GJR-GARCH), the upward trend indicates the level of certitude in oil Market in South Africa.

Table 4.8: DIAGNOSTICS

GARCH(1,1)		
Weighted Ljung-Box test standized residuals		
Lags	Statistic	Pvalue
1	2,082	0,14904
2	2,517	0,08812
5	3,317	0,35578
Weighted Ljung- Box test squared residuals		
Lags	Statistic	Pvalue
1	0,1874	0,665
5	1,7289	0,6837
9	3,0389	0,7525
ARCH test		
Lags	Statistic	Pvalue
3	0,03358	0,8546
5	1,01197	0,7294
7	1,77304	0,7652
EGARCH(1,1)		
Weighted Ljung-Box test standized residuals		
Lags	Statistic	Pvalue
1	0,5708	0,45
2	0,7569	0,8723
5	1,1432	0,9232
Weighted Ljung- Box test squared residuals		
Lags	Statistic	Pvalue
1	0,1206	0,7284
5	5,7079	0,1053
9	7,5916	0,1053
ARCH test		
Lags	Statistic	Pvalue
3	0,2136	0,6439
5	1,9045	0,4927
7	2,5777	0,5966

Table 4.8: Diagnostics continued

GJR-GARCH(1,1)		
Weighted Ljung-Box test standized residuals		
Lags	Statistic	Pvalue
1	1,236	0,2663
2	1,537	0,4173
5	2,041	0,699
Weighted Ljung- Box test squared residuals		
Lags	Statistic	Pvalue
1	0,04975	0,8235
5	3,12475	0,3847
9	4,67961	0,4779
ARCH test		
Lags	Statistic	Pvalue
3	0,2119	0,6453
5	1,9572	0,4808
7	2,5328	0,6057
FIGARCH(1,d,1)		
Weighted Ljung-Box test standized residuals		
Lags	Statistic	Pvalue
1	2,045	0,15275
2	2,542	0,08429
5	3,377	0,34274
Weighted Ljung- Box test squared residuals		
Lags	Statistic	Pvalue
1	0,1676	0,6822
5	1,8026	0,6657
9	2,9678	0,7643
ARCH test		
Lags	Statistic	Pvalue
3	0,07313	0,7868
5	0,87673	0,7698
7	1,50222	0,8209

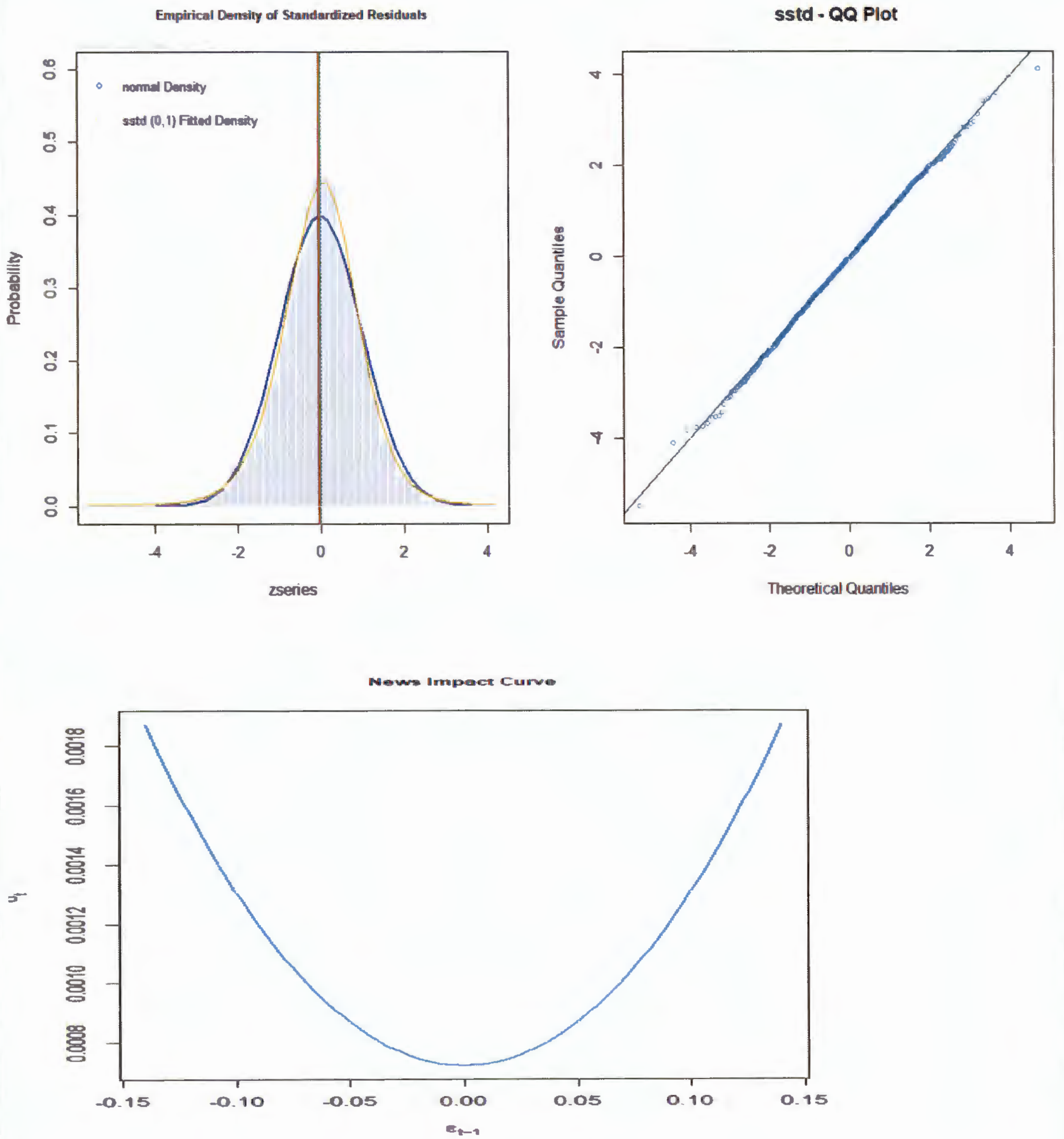


Figure 4.7 (a) GARCH (1,1)

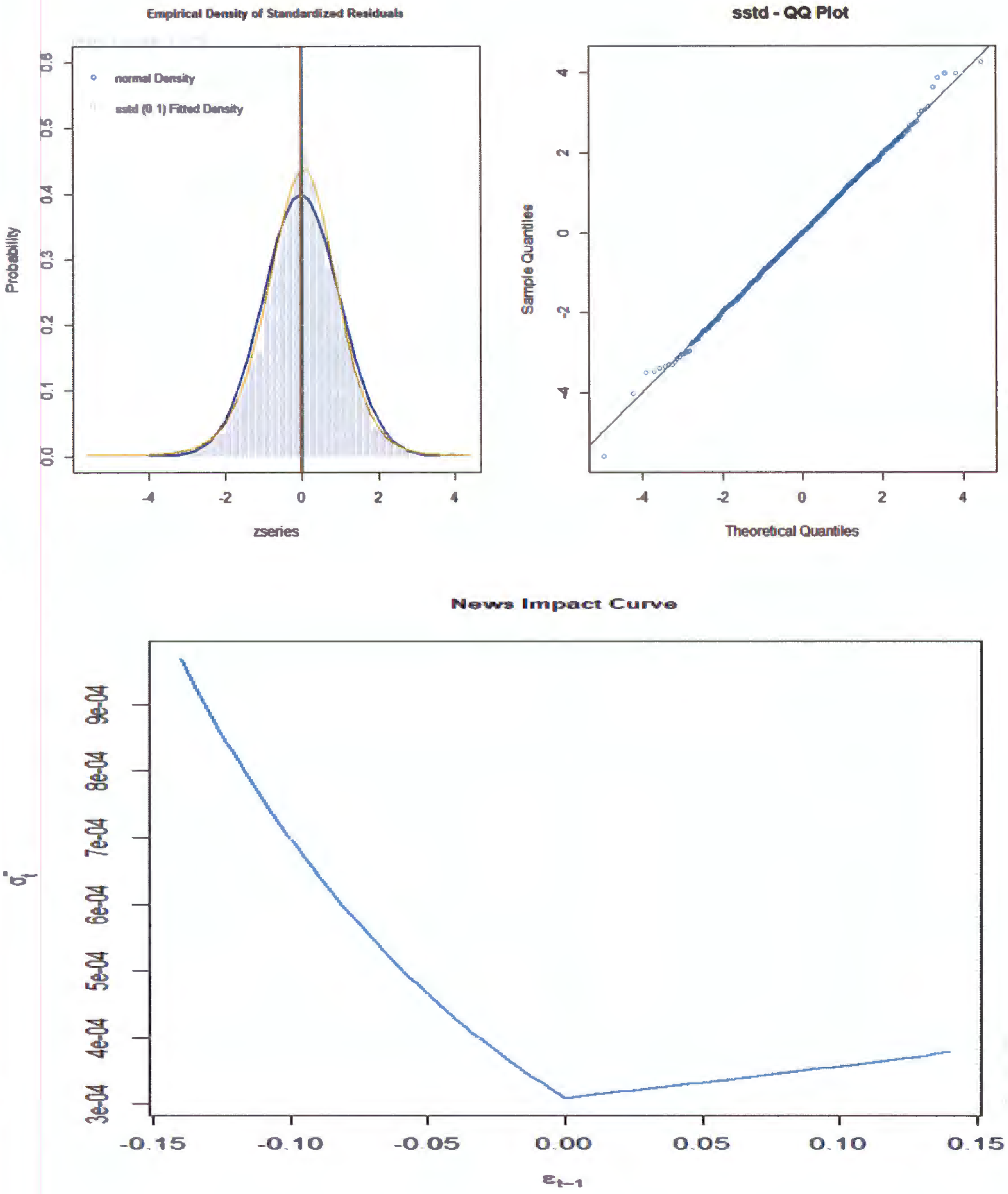


Figure 4.7 (b) EGARCH (1,1)

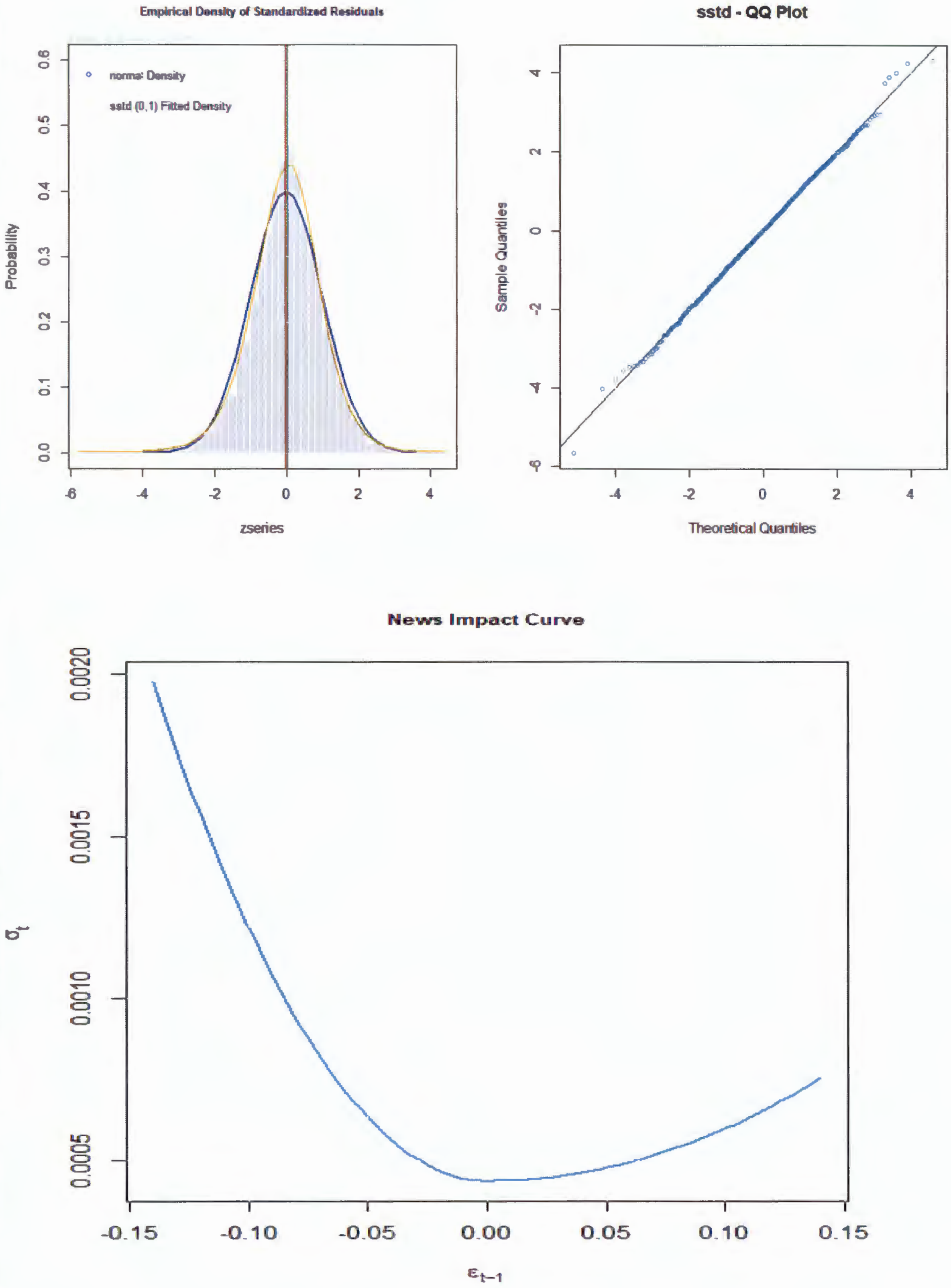


Figure 4.7 (c) GJR-GARCH (1,1)

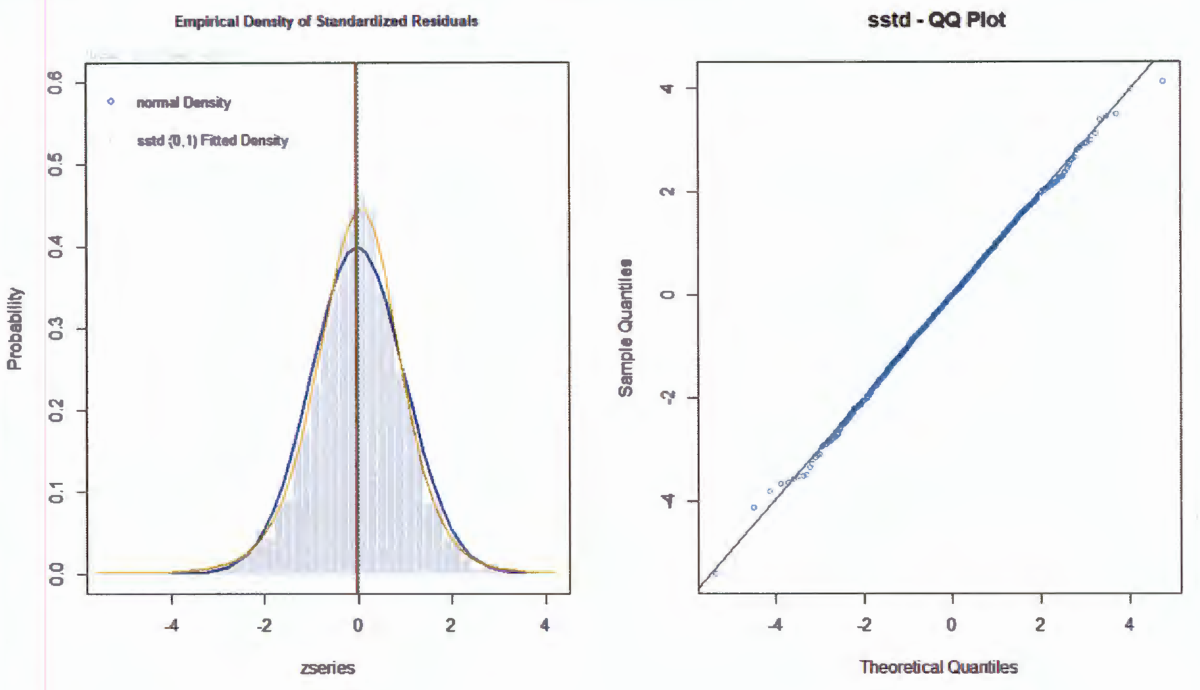


Figure 4.7 (d) FIGARCH (1, d, 1)

**** Note the FiGARCH model does not have a news impact curve**

4.10 VOLATILITY PERSISTENCE SUMMARY TABLE

The volatility persistence of the models are displayed in Table 4.9. The sum of GARCH and ARCH coefficients in the first order of GARCH, GJR-GARCH, EGARCH and FIGARCH are 0.998102, 0.967002, 0.933762 and 0.97032 respectively. All the volatility persistence of the models are greater than 0.5 and close to 1. This means that the volatility persistence are persistent. The GARCH (1,1) model has the highest volatility persistence with 0.998102 .

Table 4.9: Volatility persistence table

	EGARCH(1,1)	**GJR-GARCH(1,1)**	**FIGARCH(1,d,1)**	**GARCH(1,1)**
Volatility Persistence	**0.933762**	**0.967002**	**0.97032**	**0.998102**

4.11 MODEL SELECTION

The study evaluated GARCH(1,1), GJR-GARCH(1,1) , EGARCH(1,1) and FIGARCH(1,d,1) using RMSE, MSE and AIC. Out of the four models, the EGARCH (1, 1) (see Table 4.10) outperformed the other models with an MSE value of 0.0004567327, RMSE value of 0.02137131 and AIC value of -5.229. This is similar to the findings of Onwukwe, Samson and Lipcsey, 2014 as well as Okeyo, Ngare and Ivivi, 2016.

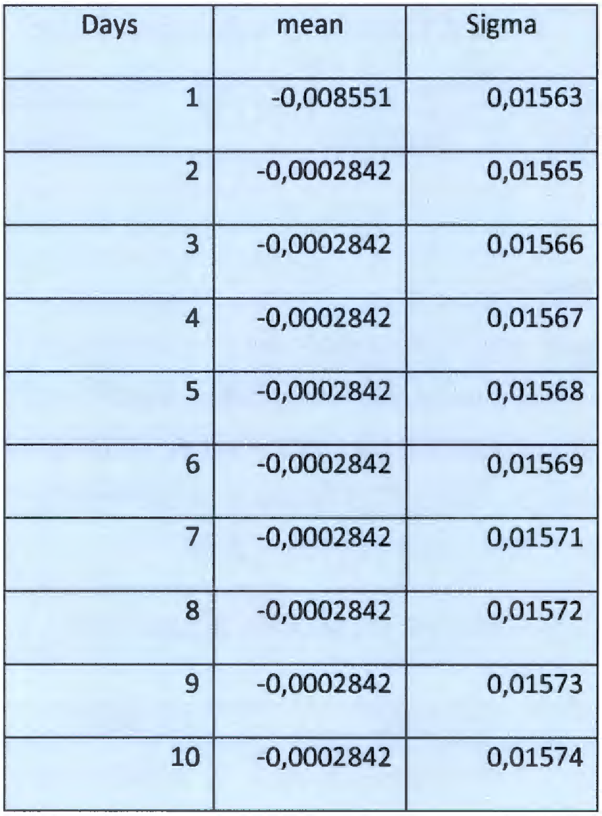
Table 4.10: MODEL SELECTION

	MSE	RMSE	AIC
GARCH(1,1)	0.0004567945	0.02137275	-5.208
EGARCH(1,1)	0.0004567327	0.02137131	-5.229
GJR-GARCH(1,1)	0.0004567608	0.02137196	-5.2197
FIGARCH(1,d,1)	0.0004568041	0.02137298	-5.2075

4.12 FORECASTING THE EGARCH MODEL

Table 4.11 display the forecasts of the EGARCH (1, 1) model for 10 days. Figure 4.8(a) and Figure 4.8(b) displays the graphical plot of the actual vs predicted volatility forecasts. The forecast series converges with the forecast volatility (unconditional variance).

Table 4.11: 10-days ahead forecasts



Days	mean	Sigma
1	-0,008551	0,01563
2	-0,0002842	0,01565
3	-0,0002842	0,01566
4	-0,0002842	0,01567
5	-0,0002842	0,01568
6	-0,0002842	0,01569
7	-0,0002842	0,01571
8	-0,0002842	0,01572
9	-0,0002842	0,01573
10	-0,0002842	0,01574

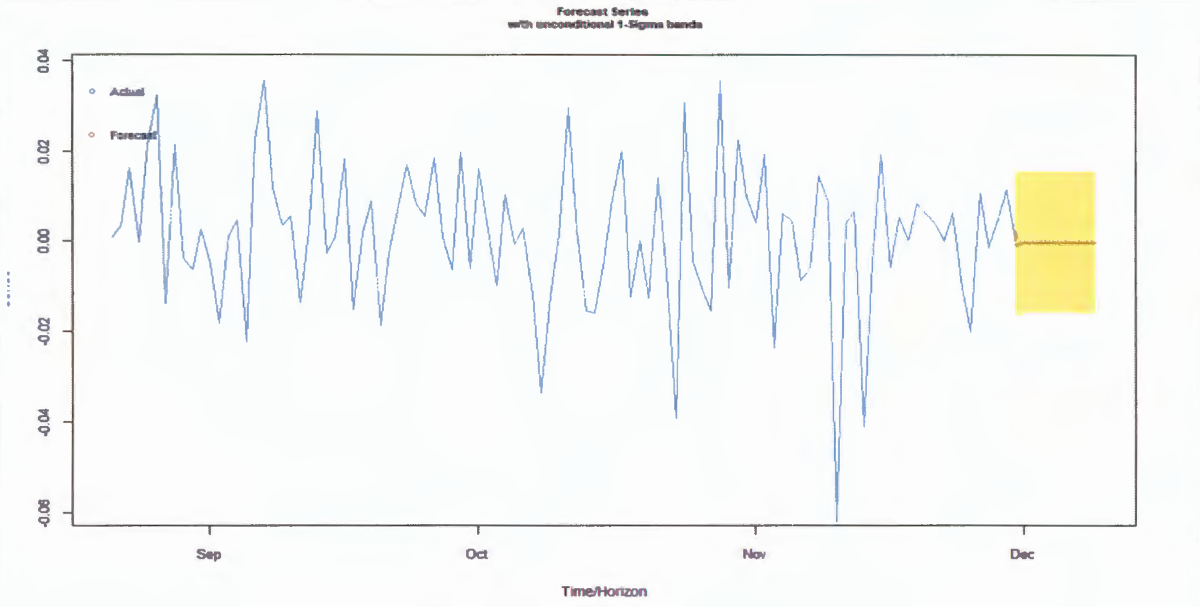


Figure 4.8(a): Forecast series

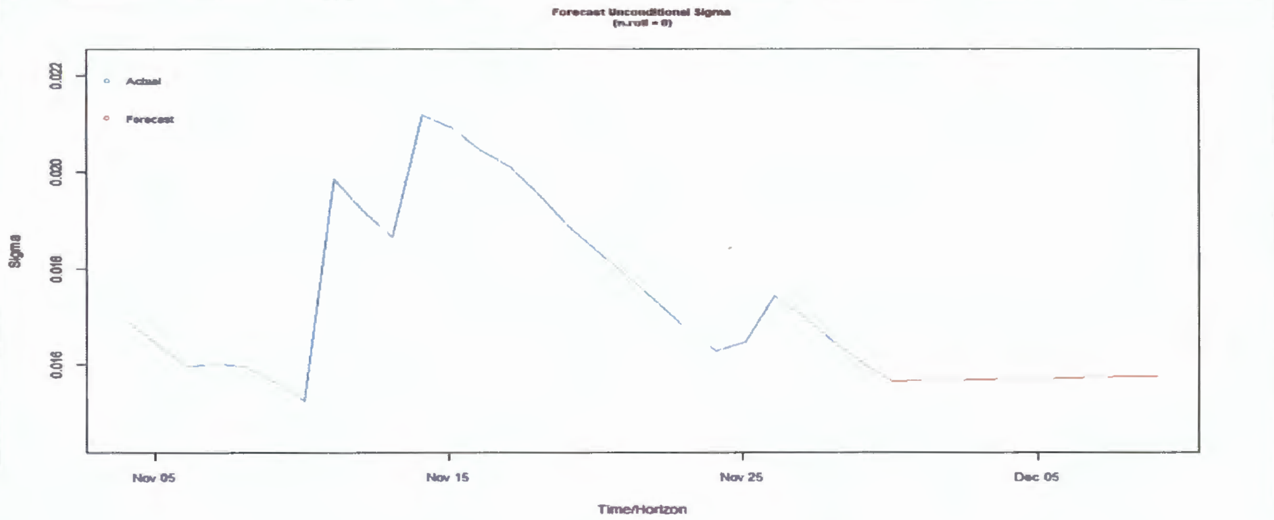


Figure 4.8(b): Forecast volatility

The next Section presents the Support vector regression

4.13 SUPPORT VECTOR REGRESSION-GARCH

This Section presents the analysis of SVR-GARCH models. In the process of Support vector regression, the data is segmented into a training and testing set. About 70% of the data was trained and 30% was tested (see Figure 4.9).

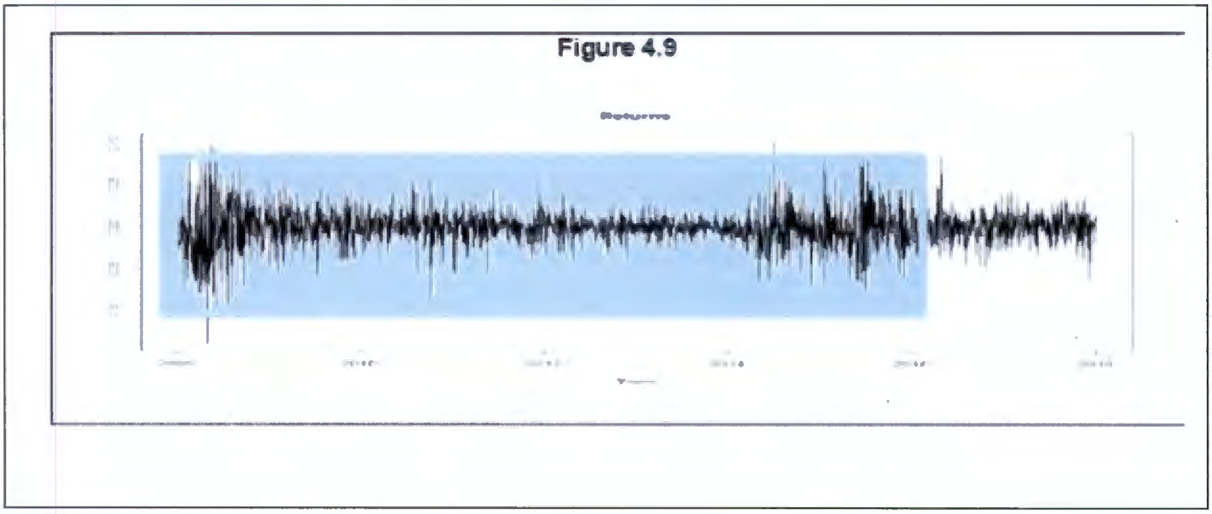


Figure 4.9: Returns

4.13.1 KERNEL SELECTION

Before selecting a kernel, the distribution of the data must be known. A summary of the data displayed a kurtosis of 3.32 confirming the non-normality of the returns as a kurtosis of a normal distribution is 3. In kernel selection, Chung and Zhang (2017) recommended the use of a radial kernel if the data is high in kurtosis; this was then validated by the results of the training and testing data (see Table 4.12). Based on the results, the study used the radial kernel, as it displayed the least error in both training and testing.

Table: 4.12: Training and Testing

Training			
	MSE	RMSE	Number of support vectors
Linear	0.0004587975	0.02141956	1553
Polynomial	0.000458629	0.02141563	1551
Sigmoid	1.909339	1.381788	1767
Radial	0.0004578275	0.0213969	1546
Testing			
	MSE	RMSE	Number of support vectors
Linear	0.0004542153	0.02131233	663
Polynomial	0.0004527015	0.02127678	664
Sigmoid	0.2998114	0.5475503	753
Radial	0.0004411292	0.02100307	678

4.13.2 TESTING DATA SET

In comparing the predictive error of the training and testing data, the testing data displayed a minimum MSE and RMSE (0.00044 and 0.0210 respectively). According to Yu and Yang, (2013) the testing data is used for GARCH modelling.

Table 4.13: SVR-test data

SVM-type	Eps-regression
SVM Kernel	Radial
Gamma	1
Number of support vectors	678
Cost	1
epsilon	0.1

4.13.3 RESIDUALS OF THE SVR-TESTING DATA

The residuals of the testing dataset are displayed in Figure 4.10. By visual inspection it evident that the points don't have a constant variance, as the data points are dispersed. This means that the error terms are heteroscedastic.

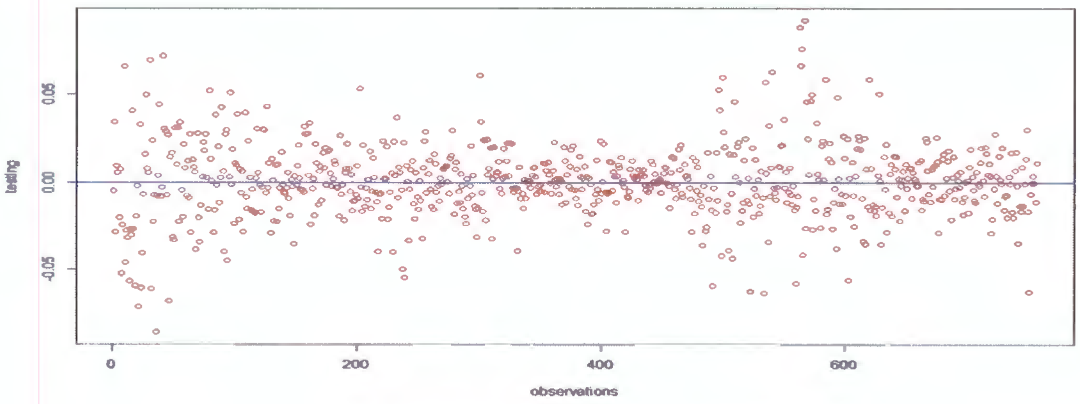


Figure 4.10 : SVR test error

4.13.4 THE ACF AND PACF PLOTS

The ACF and PACF of the residuals are displayed in Figure 4.11. The ACF of the testing data at lag 1, only shows one significant spike. The PACF displays a significant spike at lag 14 and lag 20 this indicates of ARCH effects in the testing data.

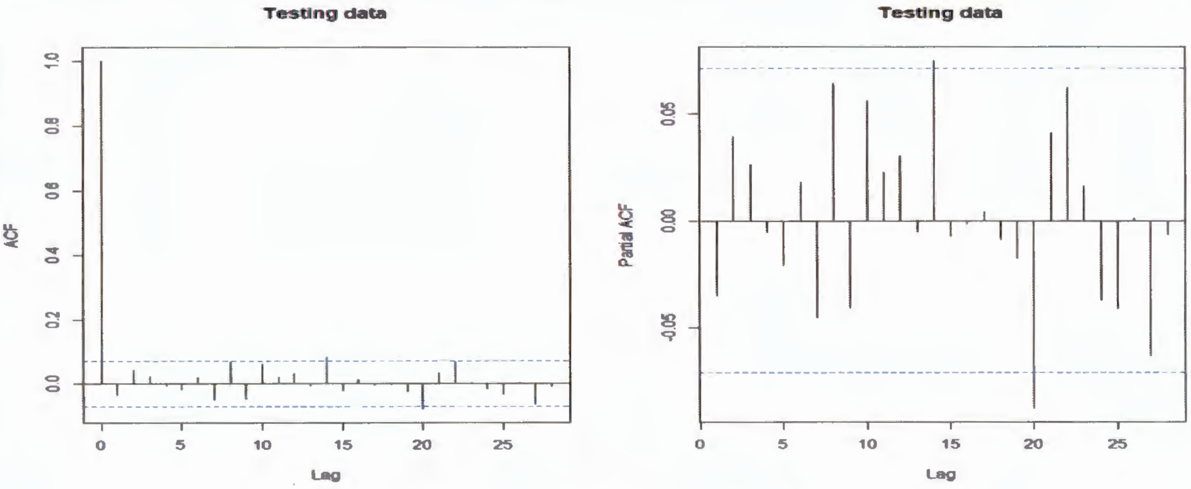


Figure 4.11: ACF and PACF plots of SVR- testing data

4.13.5 THE LJUNG BOX (LB) TEST

To test for autocorrelation in the residuals the LB test was used. Table 4.14 displays the LB test results. The probability values at lag 1 and 2 are smaller than 5%. The null hypothesis is rejected for no autocorrelation.

Hypothesis

H_0: No autocorrelation

H_1: Autocorrelation

Table 4.14: Ljung Box test

Box-Ljung test		
X-squared	df	p-value
27.711	1	1.409e-07
67.672	2	1.998e-15

4.13.6 BREUSCH PAGAN TEST

To test for heteroscedascity in the residuals, the study used Breusch Pagan test. Table 4.15 displays the Breusch Pagan test results. The probability value of the residuals is 0.005582 which is less than 5%. This indicates that the residuals are heteroscedasctic. The null hypothesis of no heteroscedacity is rejected as the residuals are heteroscedastic.

Hypothesis

H_0: Homoscedascity

H_1: Heteroscedascity

Table 4.15: Breusch Pagan test

Test	Degrees of freedom	P-value
Studentized Breusch Pagan	1	0.005582

4.14 SVR-GARCH MODELLING

From Table 4.15 it is quite evident that the error terms of the testing data are heteroscedastic as the null hypothesis of no heteroscedascity was rejected because the p-value is less than 5%. In the studies of Bezerra & Albuquerque, 2017; Chung & Zhang, 2017; Ou & Wang, 2010 they only used SVM/SVR-GARCH(1,1) framework. This study attempted to deviate from previous studies by including EGARCH (1, 1), GJR-GARCH (1, 1) and FIGARCH (1, d, 1) to account for long memory volatility and asymmetry.

4.15 DISTRIBUTION OF ERROR

Table 4.16 displays four different error distributions in finding the right fit of the data. The analysis was based on the AIC as it displayed the least error compared to the SIC. The study found GARCH(1,1) and FIGARCH(1,d,1) belonging to the student t distribution; EGARCH(1,1) and GJR-GARCH(1,1) belonged to the GED distribution.

Table 4.16: Error distribution

	Normal		Std		sStd		GED	
Model	AIC	SIC	AIC	SIC	AIC	SIC	AIC	SIC
GARCH(1,1)	-5,1384	-5,1017	-5,1706	-5,1278	-5,1682	-5,1193	-5,1701	-5,1272
EGARCH(1,1)	-5,1495	-5,1067	-5,1704	-5,1214	-5,1683	-5,1133	-5,1731	-5,1242
GJRGARCH(1,1)	-5,1425	-5,0997	-5,1694	-5,1204	-5,167	-5,1119	-5,1703	-5,1214
FIGARCH(1,d,1)	-5,1377	-5,0949	-5,1705	-5,1216	-5,1683	-5,1132	-5,1696	-5,1207

4.15.1 SVR-GARCH (1, 1)-Std

The SVR-GARCH (1, 1) output is displayed in Table 4.17(a). Let ar1 denote ϕ and Ma (1) denote θ. The probability values of $\phi \ and \ \theta$ are significant at 5% supporting the use of ARMA (1, 1) model. The probability value of β_1 and alpha are significant at 5%. Volatility persistence is the summation of ARCH and GARCH coefficients. The volatility persistence is 0.993983 which is almost 1; this designates the persistence of volatility in oil prices.

Table 4.17(a): SVR-GARCH (1, 1)

		std			
		SVR-GARCH(1,1)			
	Estimate	Std. Error	t value	Pr(>|t|)	Parameters significant at 5%
mu	0.000555	0.000534	1.0398	0.298433	not significant
ar1	-0.961066	0.068891	-13.9505	0.000000	significant
ma1	0.948146	0.079031	11.9972	0.000000	significant
omega	0.000005	0.000004	1.2440	0.213486	not significant
alpha1	0.114918	0.027448	4.1868	0.000028	significant
beta1	0.879065	0.025004	35.1566	0.000000	significant
shape	6.615753	1.567111	4.2216	0.000024	significant

4.15.2 SVR-EGARCH (1, 1)-GED

To account for asymmetry, this study used the SVR-EGARCH (1, 1) and SVR-GJR-GARCH (1, 1). Table 4.17(b) displays the SVR-EGARCH (1, 1) model. Gamma1$\neq$ zero meaning that the news impact is asymmetric. Beta1 (β_1) measures the persistence of conditional volatility; the estimate of β_1 is 0.979736 and significant at 5%. The sum of α_1 and β_1 is 0.9249 for volatility persistence.

Table 4.17(b): SVR-EGARCH (1, 1)

		SVR-EGARCH(1,1)			
	Estimate	Std. Error	t value	Pr(>|t|)	Parameters significant at 5%
mu	0.000463	0.000467	0.99093	0.321718	not significant
ar1	-0.957569	0.133235	-7.18704	0.000000	significant
ma1	0.940917	0.157250	5.98359	0.000000	significant
omega	-0.163850	0.054923	-2.98326	0.002852	significant
alpha1	-0.054836	0.026955	-2.03434	0.041917	significant
beta1	0.979736	0.006788	144.32385	0.000000	significant
gamma1	0.221794	0.044111	5.02814	0.000000	significant
shape	1.448784	0.103653	13.97721	0.000000	significant

4.15.3 SVR-GJR-GARCH (1, 1)-GED

The SVR-GJR-GARCH (1, 1) is presented in Table 4.17 (c). Let gamma1 denote (γ). The parameter of γ is positive and insignificant at 5%. The volatility persistence is 0.96394; this implies that volatility in the oil market is still persistent. All the parameters are significant except for mu, omega and gamma1.

Table 4.17(c): SVR-GJR-GARCH (1, 1)

		SVR-GJR-GARCH(1,1)			
	Estimate	Std. Error	t value	Pr(>|t|)	Parameters significant at 5%
mu	0.000347	0.000567	0.61134	0.54097	not significant
ar1	-0.957242	0.031719	-30.17873	0.00000	significant
ma1	0.941031	0.037918	24.81758	0.00000	significant
omega	0.000005	0.000004	1.15580	0.24776	not significant
alpha1	0.075281	0.028216	2.66803	0.00763	significant
beta1	0.888659	0.024110	36.85846	0.00000	significant
gamma1	0.057185	0.039308	1.45480	0.14572	not significant
shape	1.424255	0.100371	14.18985	0.00000	significant

4.15.4 SVR-FIGARCH (1, d, 1)

Table 4.17 (d) displays the estimate of delta is positive and significant at 5%. Volatility persistence is 0.803975. All the parameters are significant except for mu, omega and alpha.

Table 4.17(d): SVR-FIGARCH (1, d, 1)

			SVR-FIGARCH(1,d,1)		
	Estimate	Std. Error	t value	Pr(>|t|)	Parameters significant at 5%
mu	0.000569	0.000528	1.0770	0.281483	not significant
ar1	-0.962068	0.096130	-10.0080	0.000000	significant
ma1	0.947732	0.113464	8.3527	0.000000	significant
omega	0.000007	0.000006	1.2824	0.199695	not significant
alpha1	0.052120	0.114652	0.4546	0.649400	not significant
beta1	0.751855	0.127683	5.8885	0.000000	significant
delta	0.768730	0.217173	3.5397	0.000401	significant
shape	6.456927	1.435445	4.4982	0.000007	significant

4.16 DIAGNOSTICS

This Section provides the diagnostics of SVR (GARCH (1, 1), EGARCH (1, 1), GJR-GARCH (1, 1) and FIGARCH (1, d, 1)). The p-values of the weighted Ljung box standardized residuals and squared residuals (see Table 4.18) are all more than 5% indicating the absence of autocorrelation. This means that the estimated model for the variance and conditional mean are satisfactory for removing autocorrelation from the data. Hence the acceptance of the null hypothesis of no autocorrelation. The probability values for the ARCH test were all greater than 5% indicating that the error terms are homoscedastic. The null hypothesis is accepted for no heteroscedascity present in the error terms.

4.17 NORMALITY PLOTS

The normality plots are displayed in Figures 4.12(a-d). For the QQ-plots of the residuals of all the four models only a few points fell outside the line. The empirical distribution plots show how the error distributions fitted the data. The News impact curve of SVR-GARCH (1, 1) was symmetrical meaning that good and bad news react the same. For the News impact curve, the asymmetric models SVR-EGARCH, SVR-GJR-GARCH models the upward trend indicates the level of certitude in oil Market in South Africa.

Table 4.18: Diagnostics

SVR-GARCH(1,1)		
Weighted Ljung-Box test standized residuals		
Lags	Statistic	Pvalue
1	1,044	0,3069
5	2,478	0,7901
9	4,399	0,5957
Weighted Ljung- Box test squared residuals		
Lags	Statistic	Pvalue
1	0,8973	0,3435
5	2,0108	0,6159
9	3,5297	0,6693
ARCH test		
Lags	Statistic	Pvalue
3	1,369	0,242
5	1,845	0,5065
7	2,862	0,5399
SVR-EGARCH(1,1)		
Weighted Ljung-Box test standized residuals		
Lags	Statistic	Pvalue
1	1,007	0,3157
5	2,616	0,7136
9	4,449	0,5837
Weighted Ljung- Box test squared residuals		
Lags	Statistic	Pvalue
1	0,3068	0,5796
5	1,7517	0,6781
9	3,1727	0,7301
ARCH test		
Lags	Statistic	Pvalue
3	1,874	0,171
5	2,16	0,437
7	3,169	0,4822

Table 4.18 Diagnostics continued

SVR-GJR-GARCH(1,1)		
Weighted Ljung-Box test standized residuals		
Lags	Statistic	Pvalue
1	1,076	0,2995
5	2,705	0,6599
9	4,676	0,529
Weighted Ljung- Box test squared residuals		
Lags	Statistic	Pvalue
1	0,9152	0,3387
5	1,8991	0,6425
9	3,3904	0,6931
ARCH test		
Lags	Statistic	Pvalue
3	1,15	0,2835
5	1,51	0,5898
7	2,615	0,589
SVR-FIGARCH(1,d,1)		
Weighted Ljung-Box test standized residuals		
Lags	Statistic	Pvalue
1	1,298	0,2546
5	2,844	0,5718
9	4,799	0,4999
Weighted Ljung- Box test squared residuals		
Lags	Statistic	Pvalue
1	0,01832	0,8923
5	0,94145	0,8728
9	2,5396	0,8318
ARCH test		
Lags	Statistic	Pvalue
3	0,9929	0,319
5	1,4973	0,593
7	2,783	0,5554

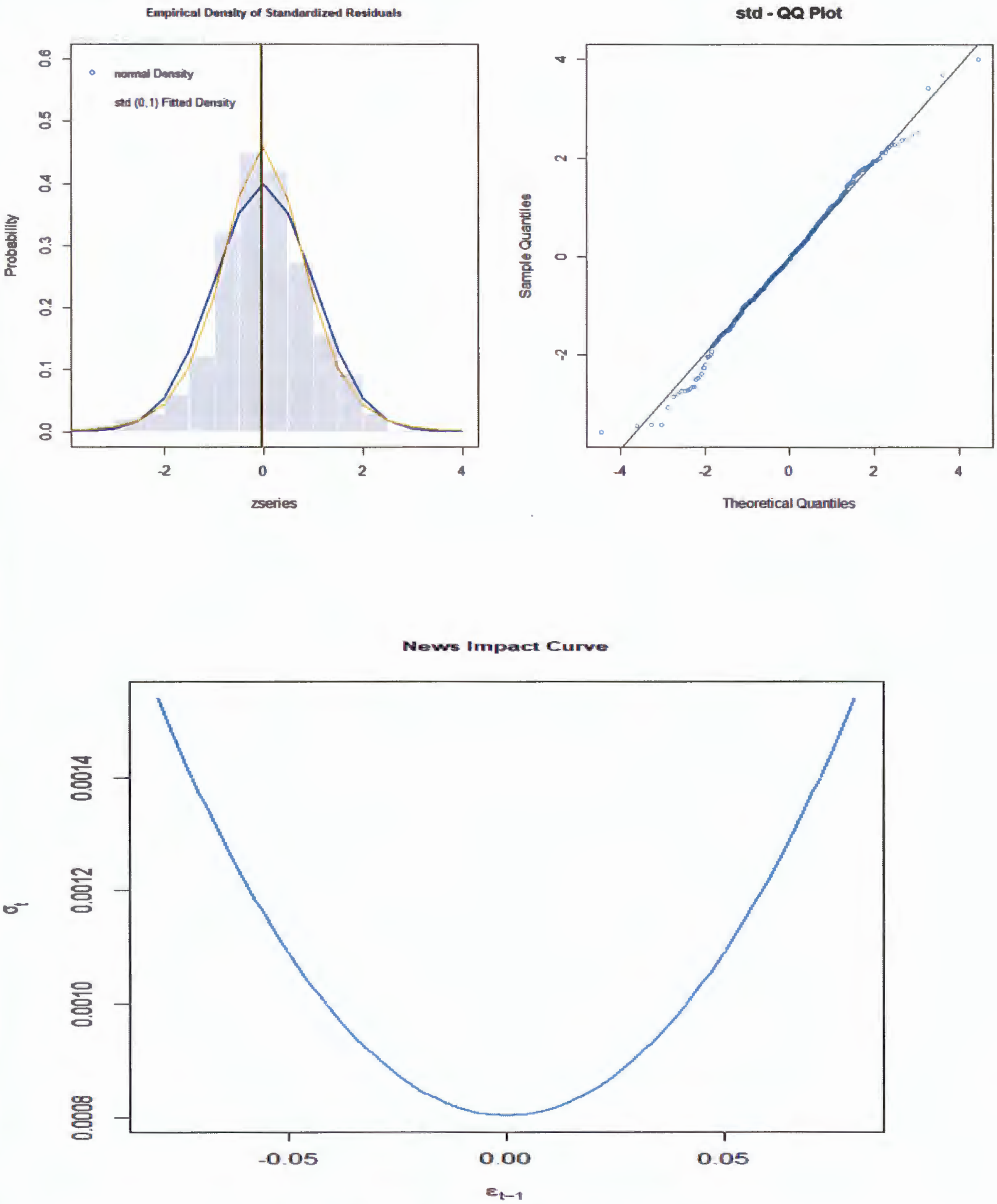


Figure 4.12(a) SVR-GARCH(1,1)

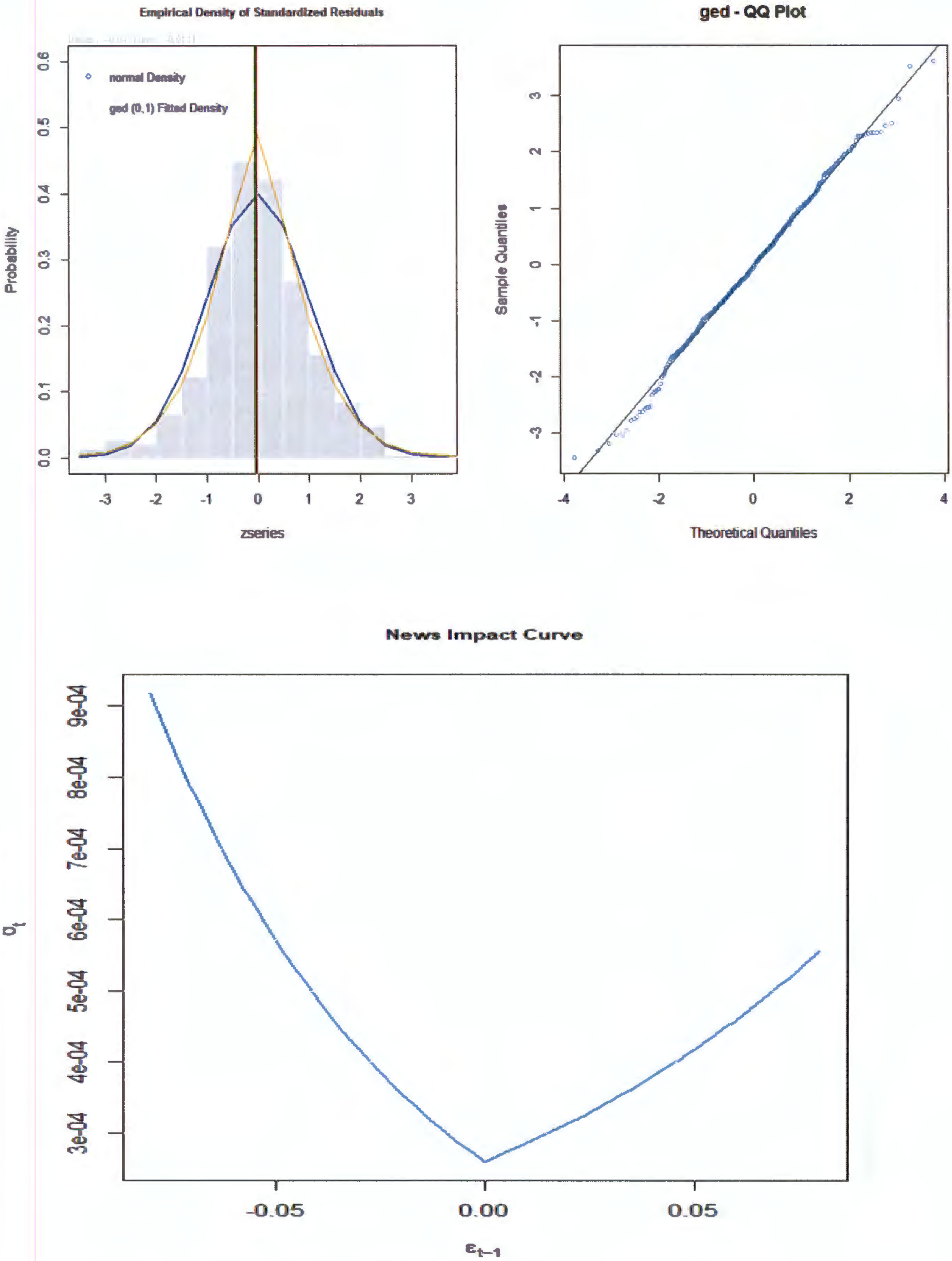


Figure 4.12(b) SVR-EGARCH(1,1)

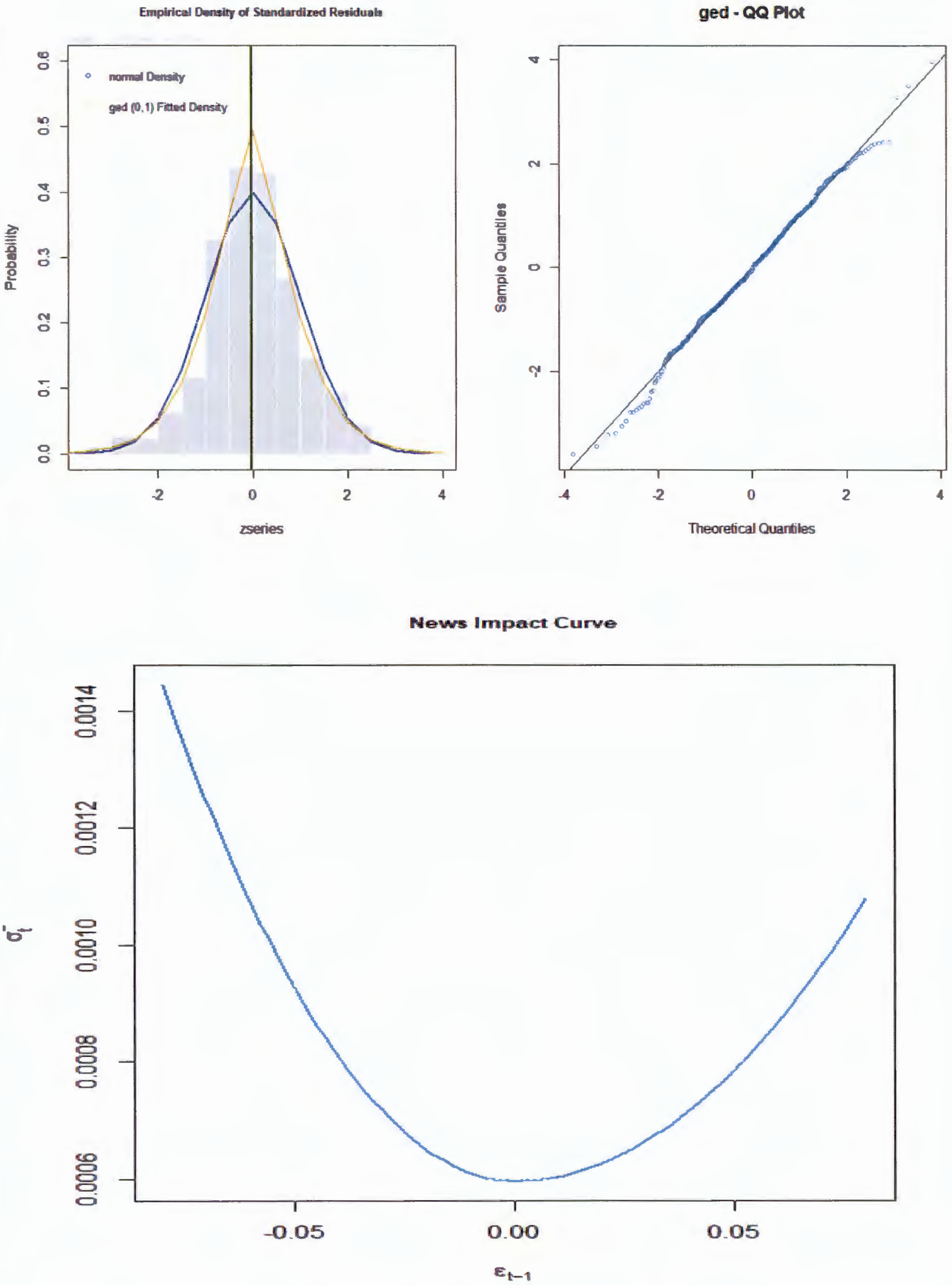


Figure 4.12 (c) SVR-GJR-GARCH(1,1)

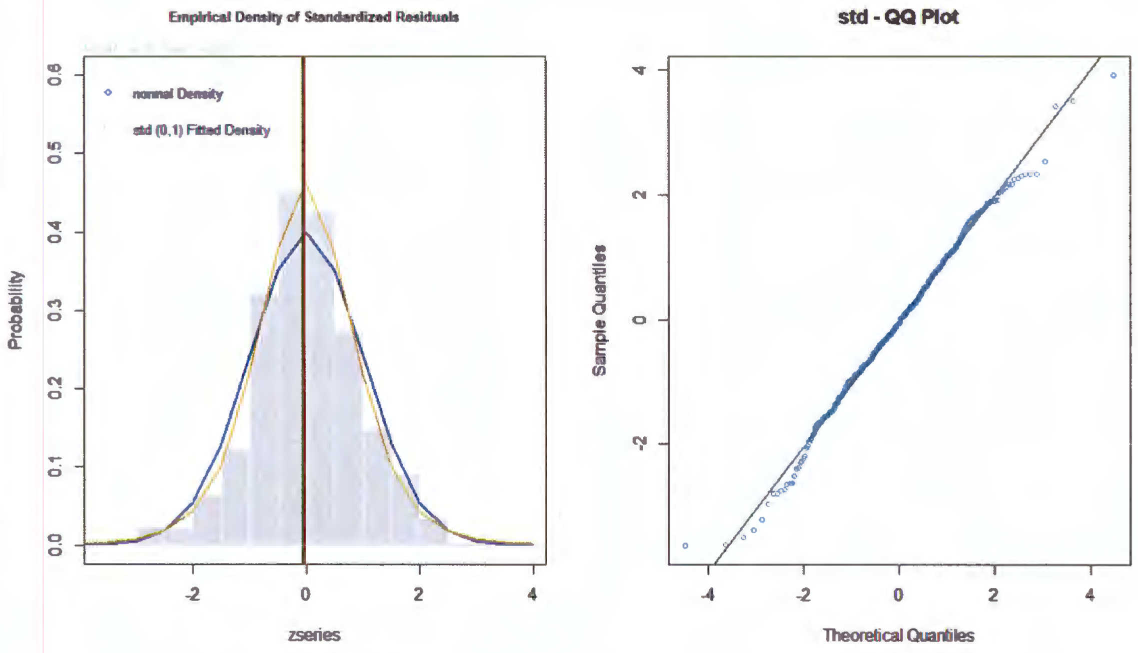


Figure 4.12 (d) SVR-FIGARCH (1,d,1)

*****Note SVR-figarch does not have a news impact curve**

4.18 VOLATILITY PERSISTENCE

The volatility persistence of the models are displayed in Table 4.19. The sum of GARCH and ARCH coefficients in the first order of SVR (GARCH, GJR-GARCH, EGARCH and FIGARCH) are 0.993983, 0.96394, 0.9249 and 0.803975 respectively. All the volatility persistence of the models are greater than 0.5 and close to one meaning that the volatility persistence are persistent. The SVR-GARCH (1, 1) model has the highest volatility persistence with 0.993983

Table 4.19: VOLATILITY PERSISTENCE

	SVR-EGARCH(1,1)	**SVR-GJR-GARCH(1,1)**	**SVR-FIGARCH(1,d,1)**	**SVR-GARCH(1,1)**
Volatility Persistence	**0.9249**	**0.96394**	**0.803975**	**0.993983**

4.19 MODEL SELECTION

The study compared the performance of SVR-GARCH(1,1), SVR-EGARCH(1,1),SVR-GJR-GARCH(1,1) and SVR-FIGARCH(1,d,1) using AIC, MSE and RMSE as displayed in Table 4.20. The results indicated that SVR-EGARCH (1, 1) was superior amongst the other three models.

Table 4.20: Model Selection

	MSE	RMSE	AIC
SVR-GARCH(1,1)	0.0004394751	0.02096366	-5.1706
SVR-EGARCH(1,1)	0.0004391812	0.02095665	-5.1731
SVR-GJR-GARCH(1,1)	0.0004391891	0.02095684	-5.1703
SVR-FIGARCH(1,d,1)	0.0004394272	0.02096252	-5.1705

4.20 COMPARISON OF GARCH AND SVR-GARCH

In comparing the GARCH with the SVR-GARCH models, the SVR-GARCH models outperformed the GARCH models as shown in Table 4.21 and this was in line with the results of Chung & Zhang, 2017; Geng & Zhang, 2015; Ou & Wang, 2010. The SVR-EGARCH (1,1) model was superior amongst the competing models. The EGARCH model was the best amongst GARCH, GJR-GARCH and FIGARCH models.

Table 4.21: Comparison of SVR-GARCH and GARCH

	GARCH(1,1)	**EGARCH(1,1)**	**GJR-GARCH(1,1)**	**FIGARCH(1,1)**
MSE	0.0004567945	0.0004567327	0.0004567608	0.0004568041
RMSE	0.02137275	0.02137131	0.02137196	0.02137298
	SVR-GARCH(1,1)	**SVR-EGARCH(1,1)**	**SVR-GJR-GARCH(1,1)**	**SVR-FIGARCH(1,1)**
MSE	0.0004394751	0.0004391812	0.0004391891	0.0004394272
RMSE	0.02096366	0.02095665	0.02095684	0.02096252

4.21 FORECASTING SVR-EGARCH (1, 1)

The study found SVR-EGARCH (1, 1) model the best as it displayed low forecasting MSE and RMSE. Table 4.22 displays a 10 day forecast for the model. Figure 4.13 and 4.14 further display a graphical plot of the actual and predicted forecasts. The volatility forecasts converge with the forecast series.

Table 4.22:10- days ahead forecasts

Time	**Sigma**
1	**0.01951**
2	**0.01947**
3	**0.01943**
4	**0.01939**
5	**0.01935**
6	**0.01931**
7	**0.01927**
8	**0.01924**
9	**0.01920**
10	**0.01917**

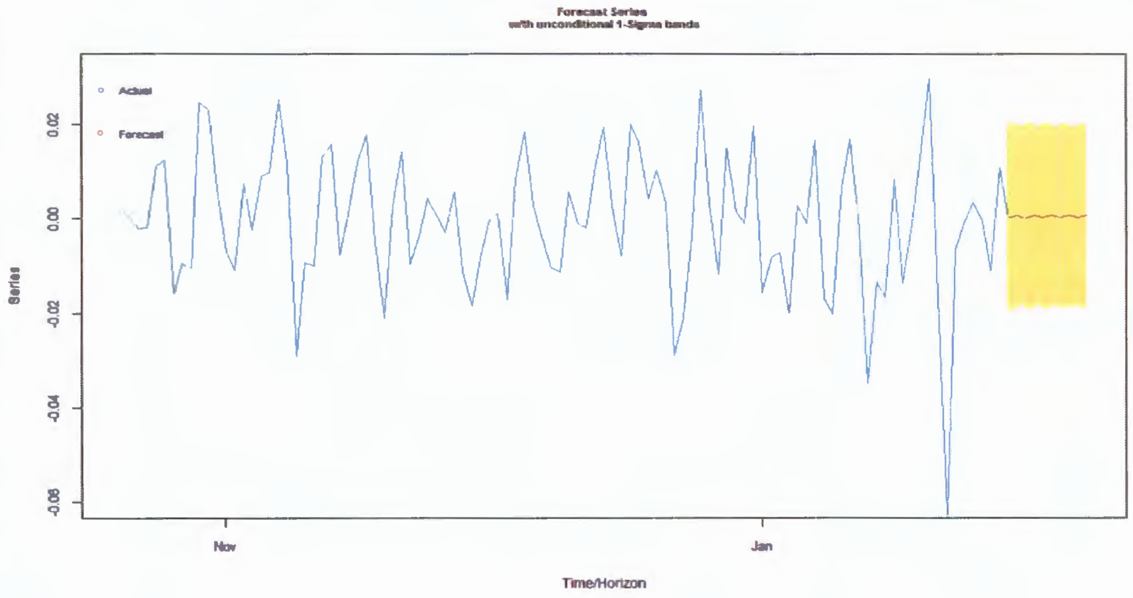


Figure 4.13: Forecast Series

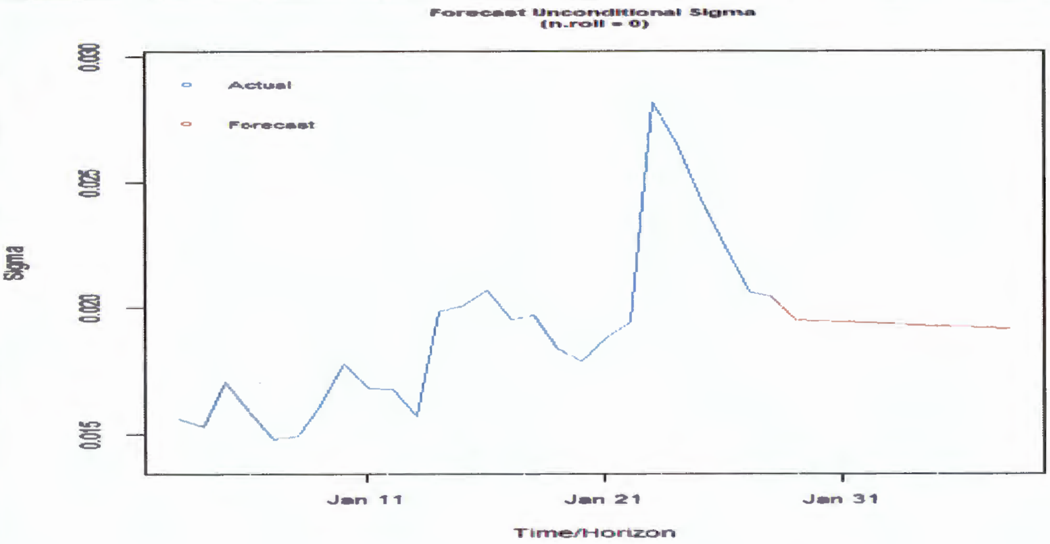


Figure 4.14: Forecast volatility

The next Section presents Extreme value theory

4.22 EXTREME VALUE THEORY (EVT)

In order to model for EVT, the data needs to be i.i.d. As discussed above, the EGARCH (1, 1) model had the least error with no autocorrelation and heteroscedascity. The EGARCH (1, 1) model is then suitable for modelling extreme value theory.

4.22.1 Tail quantiles

The GEV estimates and their standard errors are displayed in Table 4.23. The estimate of $\hat{\xi}$ (-0.1965) suggests that the returns can be modelled using a Weibull distribution since $\hat{\xi} < 0$.

Table 4.23: GEV estimates

Parameter	Estimate	Standard Error
Location ($\hat{\mu}$)	-0.3964	0.022635
Scale ($\hat{\sigma}$)	1.0746	0.013779
Shape ($\hat{\xi}$)	-0.1965	0.003532

4.22.2 THRESHOLD DETERMINATION

The very first step under the POT method is to select the appropriate threshold. This was done graphically by using the mean excess plot as displayed in Figure 4.15. It is evident that the mean excess plot is trending upwards for the oil prices; this is a clear indication of heavy tails. The plot is a straight line with a positive gradient. According to Susan and Waititu (2015), one may choose a threshold where the plot is positive and linear. In selection of the threshold, the points are linear around 1. The study selected a threshold of 1.

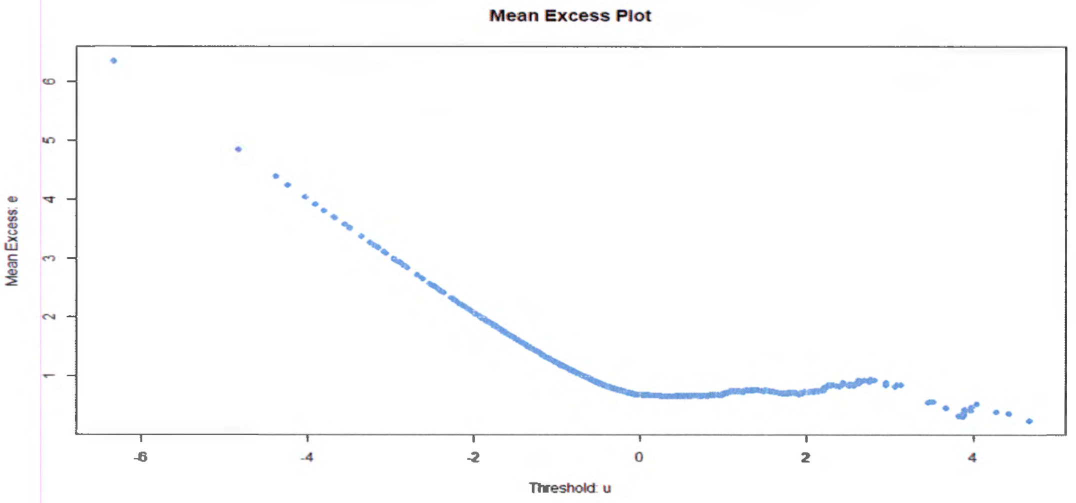


Figure 4.15: Mean excess plot

4.22.3 Fitting the GPD distribution

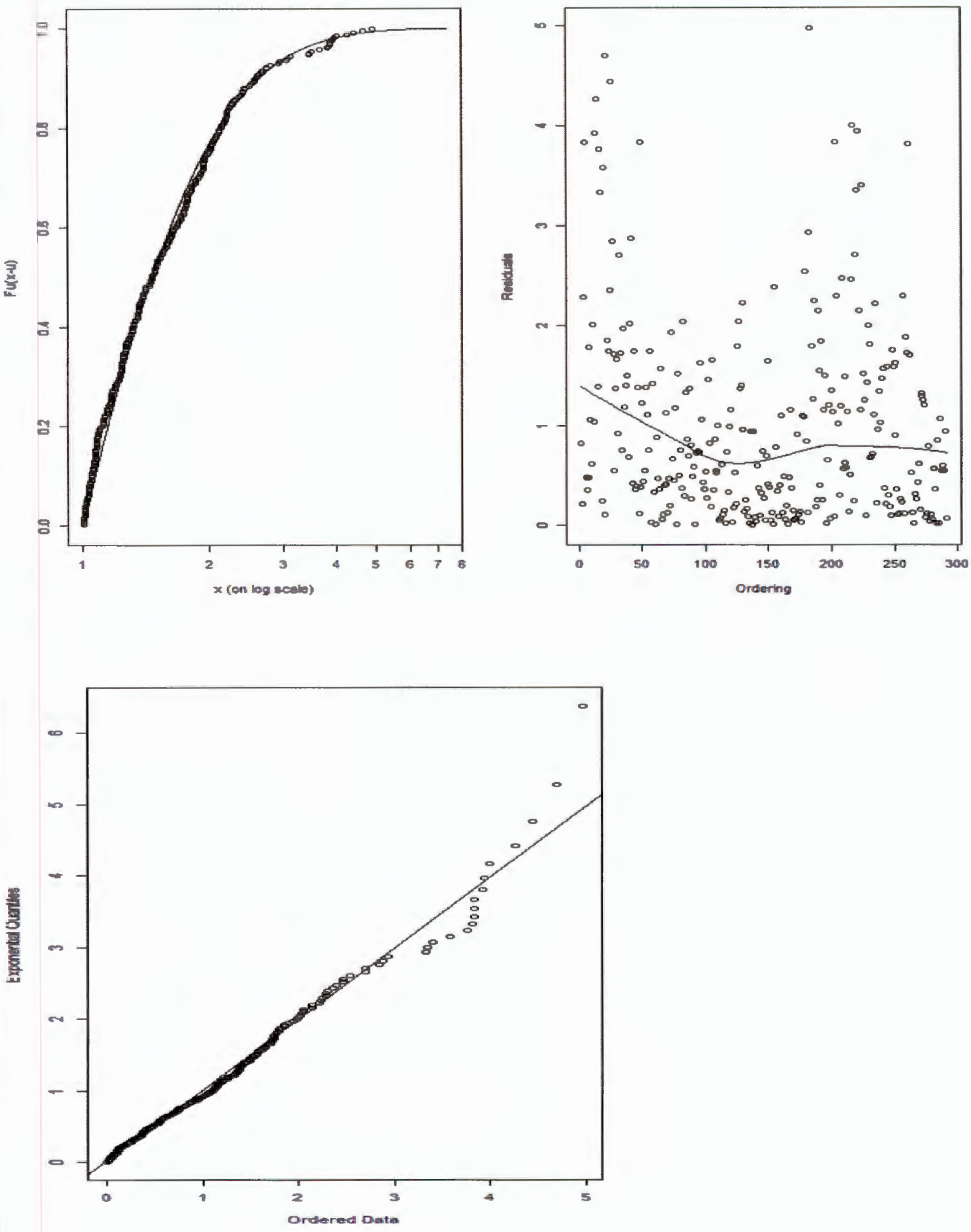
After selecting the threshold value, it is used to determine the GPD parameters. Table 4.24 displays the number of exceedances and the GPD parameter estimates. X_i symbolises the shape parameter, β denotes the scale parameter and μ represents the location parameter. The number of exceedance is 291.

Table 4.24: Estimates of the GPD

	X_i	σ	μ	β	Number of exceedances
Left tail	0.07600268	0.55072056	-0.29270533	0.64896964	291
Standard error					
	0.0728969	0.1294812	0.2075962		

4.22.4 MODEL CHECKING

Figure 4.16 is used to assess the goodness of fit of the GPD model. For the excess distribution plot, all the points' fit perfectly on the curve. An inference can be drawn that the GPD model provides a great fit for extreme values. For the QQ plot, the points on the QQ-line do not depart significantly from the straight line. In addition, there is no visible trend of the scatter plot in the residuals. This then validates the use of Generalized Pareto distribution. The diagnostic plots are displayed in Figure 4.17. All the points on the probability plot line fit perfectly, the points on the QQ-line do not depart significantly from the line. The density plots validate the use of GPD as it provides a great fit for the exceedances. The return level plot is adequate as there are no significant deviations from the plot. This is in accord with the studies of Kiragu & Kyalo, 2016; Susan & Waititu, 2015.



Excess distribution

Residual Plot

QQ-plot

Figure 4.16: Model Checking

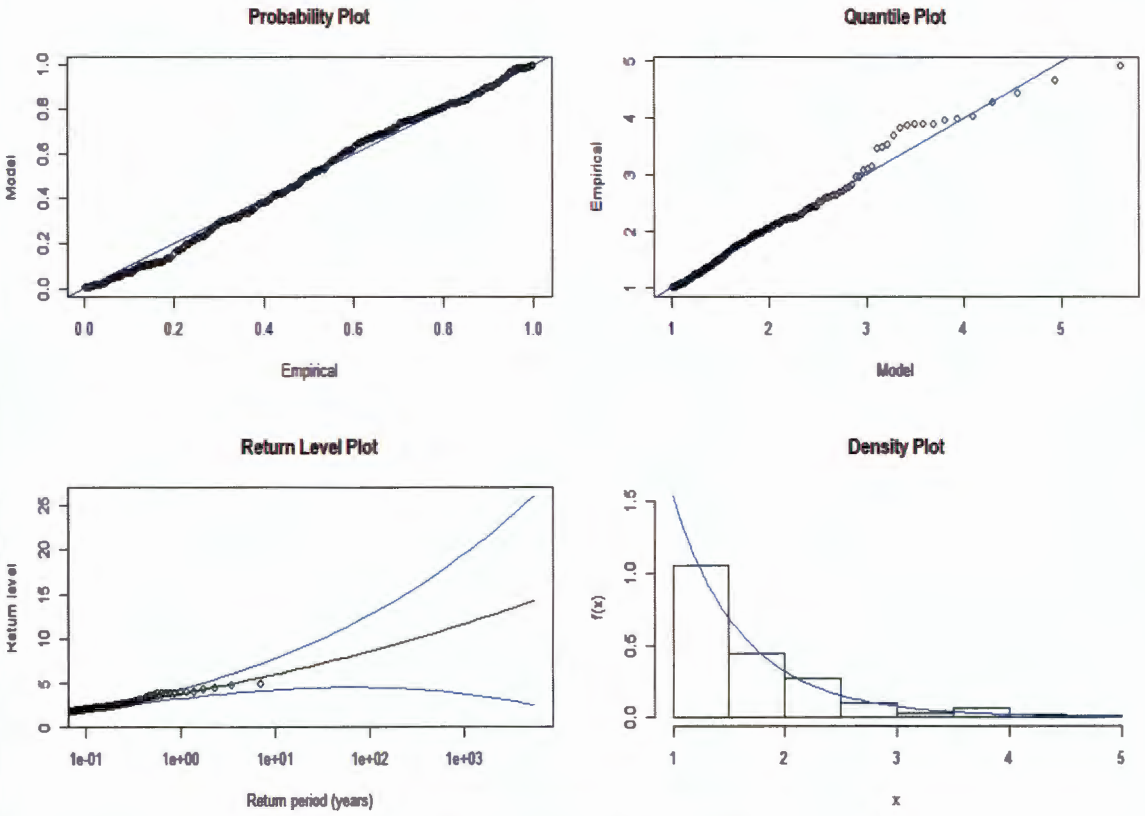


Figure 4.17: Diagnostics

4.23 RISK MEASURES

The risk measures of the left tail for the Generalized Pareto distribution are displayed in Table 4.25. The results designate that with probability 0.01, thus 99.0% confidence level, the daily loss will not go beyond 2.74% and even if it does surpass, the probable loss is 3.58% in a day. For a 99.5% (0.9950) confidence interval, the market loss will not exceed 3.29%. If it exceeds 3.26% the expected loss is 4.19%. For upper quantiles with 99.95% confidence interval the expected market losses will not surpass 5.37% if it exceeds it is expected to be 6.43%.

Table 4.25: Risk Measures

Measure of left tail		
Probability	**VaR**	**Expected Shortfall**
0.9900	**2.743267**	**3.588620**
0.9950	**3.299252**	**4.190257**
0.9990	**4.708773**	**5.715516**
0.9995	**5.370903**	**6.432013**
0.9999	**7.049521**	**8.248464**

The next Section presents SVR-EGARCH-EVT

4.24 EXTREME VALUE THEORY

The SVR-EGARCH (1, 1) had the least error with no autocorrelation and heteroscedascity. The SVR-EGARCH (1, 1) is then suitable to model extreme value theory.

4.24.1 TAIL QUANTILES

The GEV estimates and their standard errors are displayed in Table 4.26. The shape $\hat{\xi}$ (-0.1965) suggests that the returns can be modelled using a Weibull distribution since $\hat{\xi} < 0$.

Table 4.26: GEV Estimates

Parameter	Estimate	Standard Error
Location ($\hat{\mu}$)	-0.3906	0.03960
Scale ($\hat{\sigma}$)	1.0219	0.02518
Shape ($\hat{\xi}$)	-0.1965	0.01067

4.24.2 THRESHOLD DETERMINATION

The very first step under the POT method is to select the appropriate threshold. This was done graphically by using the mean excess plot as displayed in Figure 4.18. It is evident that the mean excess plot is trending upwards for the oil prices; this is a clear indication of heavy tails. The plot is a straight line with a positive gradient. In selection of the threshold, the points are linear around 0.5. The study selected a threshold value of 0.5.

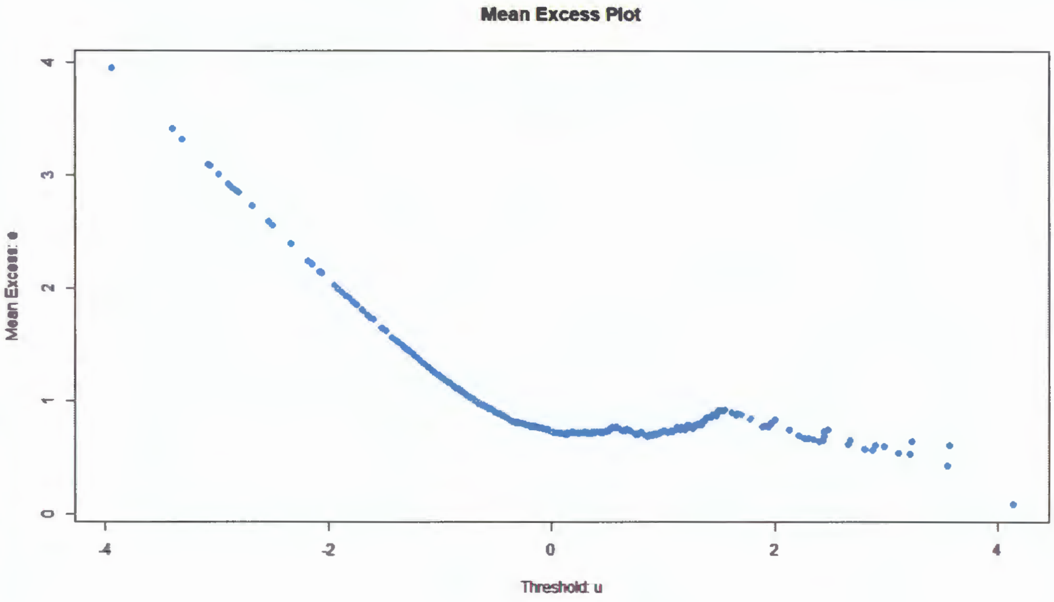


Figure 4.18

4.24.3 FITTING THE GPD

A 0.5 threshold value was selected to estimate the GPD. The GPD parameter estimates are shown in Table 4.27. The location, scale and shape are represented by μ ,X_i and β respectively.

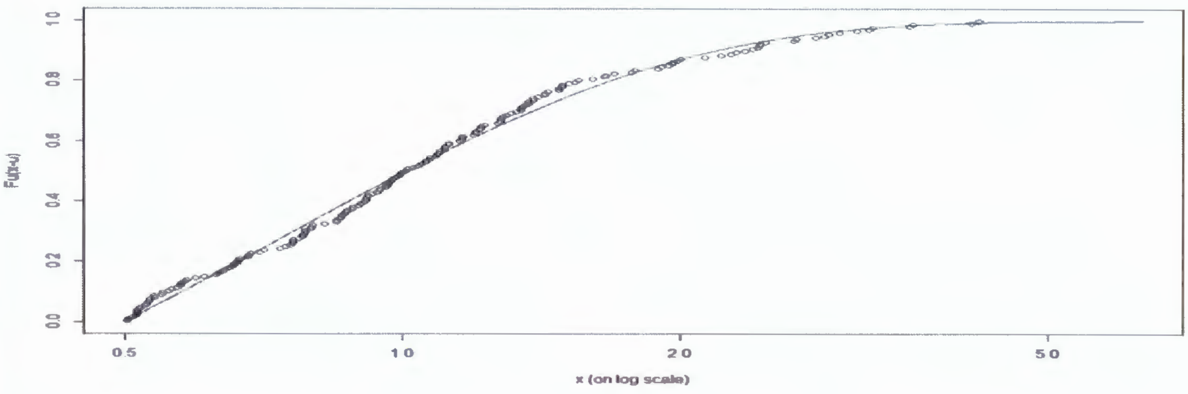
Table 4.27: Estimates of the GPD

	X_i	sigma	μ	β	Number of exceedances
Left tail	-0.00376834	0.74158161	-0.54918157	0.73762794	183
Standard error					
	0.08103158	0.15543626	0.17373784		

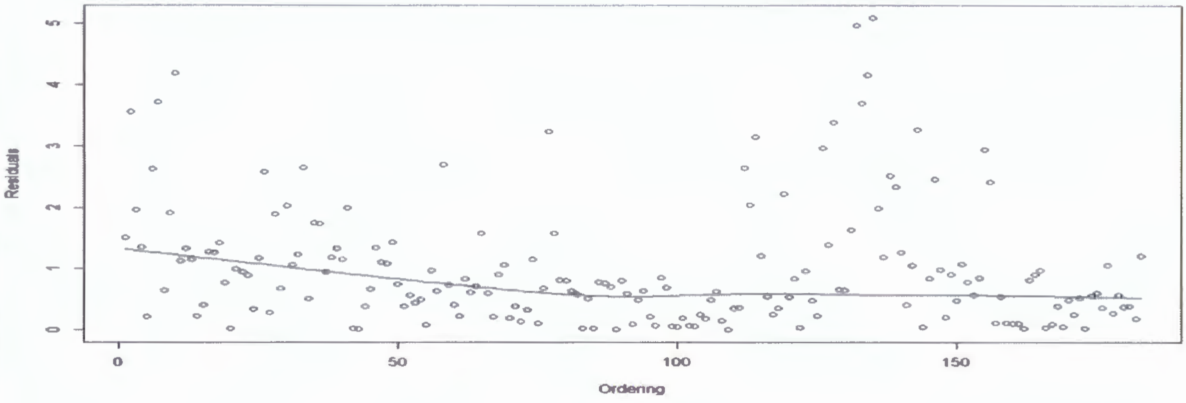
4.25 MODEL CHECKING

The goodness of fit for the GPD is displayed in Figure 4.19. All the points' fit on the curve of the excess distribution. An inference can be drawn that the GPD model provides a great fit to extreme values. For the QQ plot, the points on the QQ-line do not depart significantly from the straight line. In addition, there is no visible trend of the scatter plot of the residuals. This then validates the use of Generalized Pareto distribution. The diagnostic plots are displayed in Figure 4.20, all the points on the probability plot line fits perfectly and the points on the QQ-line do not depart from the line significantly. The density plots validate the use of GDP as it provides a great fit for the exceedances. The return level plot is adequate as there are no significant deviations from the plot.

Excess distribution



Scatter Plot of Residuals



QQ-Plot

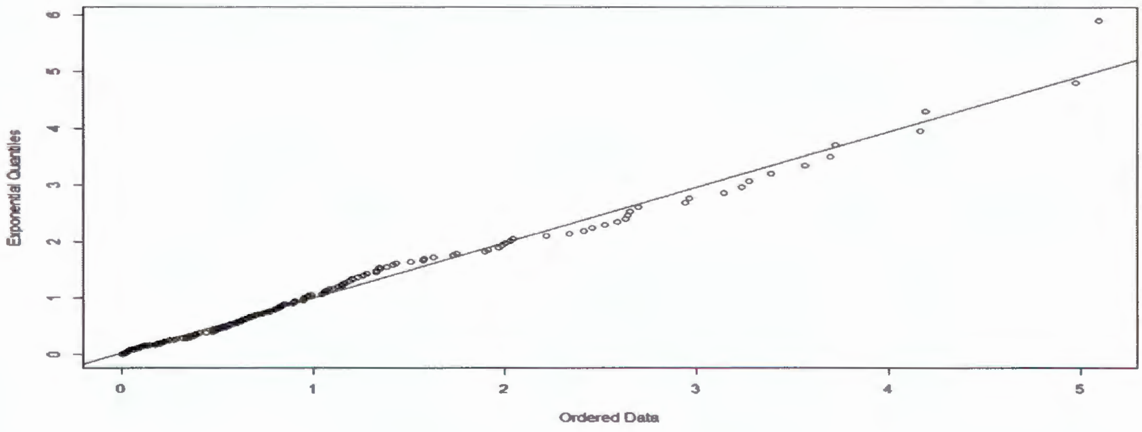


Figure 4.19: Model Checking

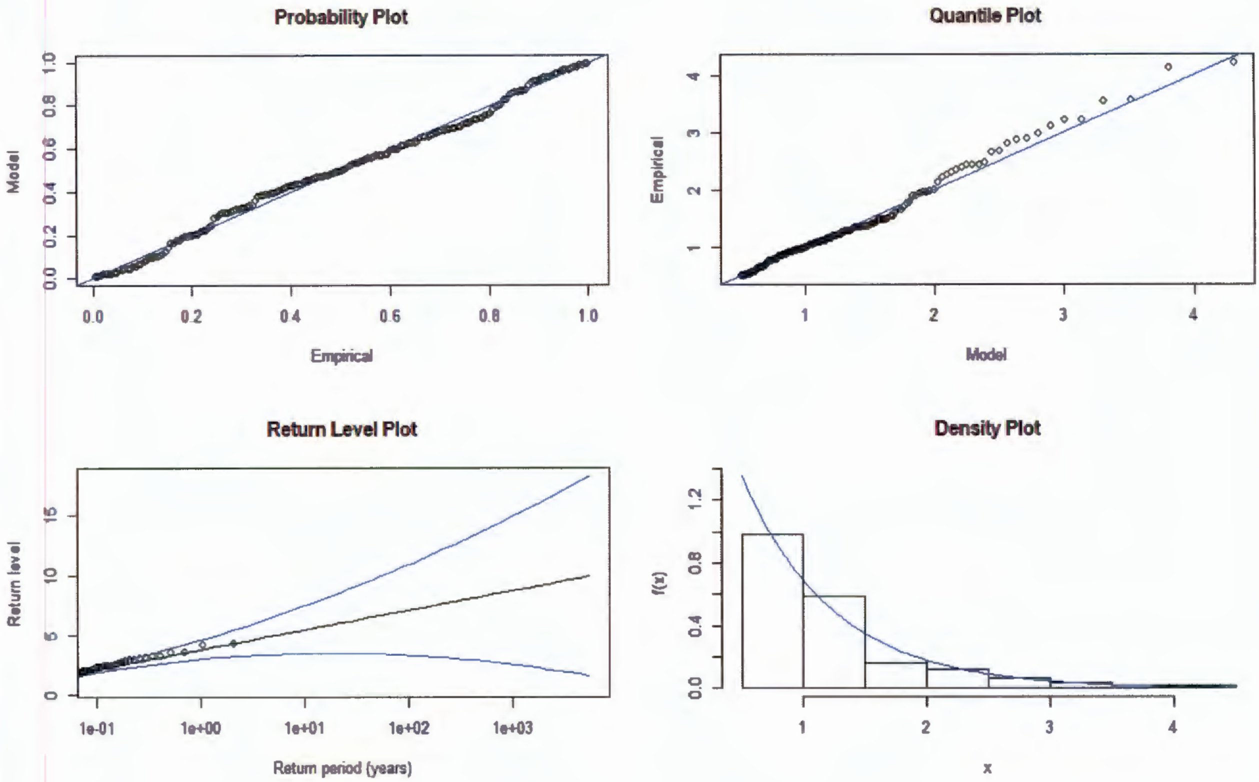


Figure 4.20: Diagnostics

4.26 RISK MEASURES

Table 4.28 displays the Expected Shortfall and Value at Risk values. The results indicate that at 99% confidence level, the daily loss will not go beyond 2.83%, even if it does, the expected loss is 3.56% in a day. For a 99.5% (0.9950) confidence interval, the market loss will not exceed 3.34%. If it exceeds 3.34% the expected loss is 4.06%. For upper quantiles with 99.95% confidence interval, the expected market losses will not surpass 5.0%. If it exceeds, it is expected to be 5.72%.

Table 4.28: RISK MEASURES

Measure of left tail		
Probability	**VaR**	**Expected Shortfall**
0.9900	2.835899	3.562528
0.9950	3.340725	4.065571
0.9990	4.508122	5.228843
0.9995	5.008843	5.727796
0.9999	6.166748	6.881610

The next Section presents the Chapter summary.

4.27 CHAPTER SUMMARY

Chapter Four displayed the results of the study using EViews 10 and Rstudio 3.5. The study assessed the effectiveness of GARCH and SVR-GARCH models to fit the tails of oil prices using extreme value theory. The study used GARCH (1,1), EGARCH(1,1), GJR-GARCH(1,1) and FIGARCH (1,d,1) in comparison with SVR-GARCH(1,1), SVR-EGARCH(1,1), SVR-GJR-GARCH(1,1) and SVR-FIGARCH(1,d,1).

Preliminary data analysis was conducted before the actual analysis. The descriptive statistics displayed negative skewness and excess kurtosis. The conclusion was that the data is not normally distributed as the skewness of the returns was -0.08 with a kurtosis of 3.32. The study conducted stationarity test using the PP and the ADF test. Both tests gave the same results as the returns were found stationary at levels. To identify the presence of autocorrelation and heteroscedascity, the ARCH test was used. The results proved that the returns were autocorrelated and heteroscedastic hence the use GARCH models were applicable. The study used four GARCH models namely GARCH (1,1), GJR-GARCH(1,1), EGARCH(1,1) and FIGARCH(1,d,1). Since the data was non-normal, the study fitted four error distributions namely student *t*, normal, skew student *t* and GED. In selection of the error distribution that fitted the data, the study compared the AIC with the SIC. Based on the minimum error, the AIC was selected and it was found that the returns belonged to the skewed student *t* distribution. The study modelled the returns using the skewed student *t* distribution using GARCH (1, 1), EGARCH (1, 1), GJR-GARCH (1, 1) and FIGARCH (1, d, 1). The results revealed that the EGARCH (1, 1) was the best performing model based on the minimum value of MSE (0.000456) and RMSE (0.02137131) as compared to the other three models.

The study further estimated the GARCH models using Support Vector Regression (SVR). The estimated models were SVR-GARCH(1,1), SVR-EGARCH(1,1), SVR-GJR-GARCH(1,1) and SVR-FIGARCH(1,d,1). Four error distributions were compared namely normal, student t, skewed student t and GED. In selection of the error distribution, the student *t* distribution fitted GARCH (1, 1) and FIGARCH (1, d, 1). For the asymmetric models, EGARCH (1, 1) and GJR-GARCH (1, 1) belonged to the GED distribution based on the smallest AIC. The results of the study indicated that SVR-EGARCH (1, 1) displayed the smallest MSE and RMSE with a value of 0.0004391812 and 0.02095665 respectively. In comparison of GARCH models and SVR-GARCH models the SVR-GARCH outperformed the GARCH models and this is in accord with the studies of Bezerra & Albuquerque, 2017; Chung & Zhang, 2017; Ou & Wang, 2010.

EVT was applied to fit the tails of the returns. The study used EGARCH (1,1) and SVR-EGARCH(1,1) to model for extreme values as they were the best two competing models for GARCH and SVR-GARCH models. The study modelled the left tail as the descriptive statistics displayed a negative skewness of -0.08. The mean excess plot was used to select the threshold. A threshold value of 1 was selected for EGARCH (1,1) and a value of 0.5 was selected for SVR-EGARCH(1,1). The tail quantiles of the models displayed a Weibull distribution as $\xi < 0$. The number of exceedances for EGARCH (1, 1) displayed 291 and for SVR-EGARCH (1, 1) indicated 183. The study then used GPD to model the tail behaviour. The QQ-plot and exceedance plots for both models displayed a great fit for the data. This is in agreement with the studies of Kiragu & Kyalo, 2016; Nortey *et al.*, 2015; Susan & Waititu, 2015. The GPD models was used to compute the VaR and Expected shortfall at lower and upper quantiles. It was found that for both models the VaR and expected shortfall increased as the quantiles increased. This is in accord with the studies of Nortey *et al.*, (2015). The conclusions are presented in the next Chapter.

CHAPTER FIVE

CONCLUSIONS

5.1 INTRODUCTION

The conclusions of the objectives discussed in Chapter one are reviewed. The outcome of the results are compared with previous studies discussed in Chapter two in order to find correspondences and dissimilarities.

5.2 OBJECTIVES AND SUMMARY OF FINDINGS

Objective 1: To model symmetric and asymmetric GARCH and SVR-GARCH models.

Findings

Firstly preliminary analysis was conducted in order to have a feel of the data. Descriptive statistics such as mean, kurtosis and skewness were displayed. The results showed that the returns of Brent Crude oil prices were negatively skewed (-0.08) and high in kurtosis with a value of 3.32. This implies non-normality of the data; this was further validated by the Shapiro Wilks test and JB test. The p-values were all less than 5%. Stationarity tests such as the ADF and the PP test were used. The results proved that the returns were stationary at levels. To test for serial correlation, the study used the LB test. The residuals were auto correlated. To check for the presence of heteroscedascity, the study used the LM test and Breusch Pagan test. They both indicated the presence of heteroscedascity in the residuals. Hence it is necessary to use GARCH models to remove serial correlation and heteroscedascity.

In modelling the GARCH models the study fitted GARCH(1,1), EGARCH(1,1), GJR-GARCH(1,1) and FIGARCH(1,d,1). The study found EGARCH (1, 1) model superior based on its smallest MSE and RMSE values of 0.0004567327 and 0.02137131 respectively. This is in accord with Onwukwe *et al.*, (2014). A study by Cristiani-d'Ornano *et al.*, (2010) compared IGARCH, FIGARCH and GARCH models and found FIGARCH model best fitting. The current study is different from Cristiani-d'Ornano *et al.*, (2010) as it included asymmetric models (EGARCH(1,1) and GJR-GARCH(1,1)).

For the Support Vector Regression the study used SVR-GARCH(1,1), SVR-EGARCH(1,1), SVR-GJR-GARCH(1,1) and SVR-FIGARCH(1,d,1). The study showed that SVR-EGARCH (1, 1) was superior based on MSE and RMSE values of 0.0004391812 and 0.02095665 respectively. The authors Chung and Zhang (2017) examined six different foreign exchange rates using GARCH and SVM-GARCH models. Their study compared GARCH (1, 1), EGARCH (1, 1) and GJR-GARCH (1, 1), with SVM-GARCH (1, 1). The current study is different from Chung and Zhang (2017) because it includes SVR-EGARCH, SVR-GJR-GARCH and SVR-FIGARCH models.

Objective 2: Evaluate GARCH and SVR-GARCH based on accuracy measures and make predictions

Findings

In evaluation of the SVR-GARCH with GARCH models, the SVR-GARCH models outperformed the GARCH models. This is in line with the studies of Bezerra & Albuquerque, 2017; Bildirici & Ersin, 2013; Chung & Zhang, 2017; Li, 2014. The forecasting performance of the SVR-EGARCH (1, 1) was estimated for 10 days (short term prediction) the forecast series converged with the forecast volatility.

Objective 3: To use the best GARCH and SVR-GARCH to model extreme risk of Brent crude oil prices

Findings

The EGARCH (1, 1) and SVR-EGARCH (1, 1) were the best models. No autocorrelation and heteroscedascity existed among these models, hence the models were considered suitable for extreme value analysis. The study followed the POT method; the GPD was used to fit the tails of the returns. The study found that the GPD model provides a good fit for extreme values. Similar results were obtained by Kiragu & Kyalo, 2016; Susan & Waititu, 2015. It was observed that the VaR and Expected shortfall for both EGARCH (1, 1) and SVR-EGARCH (1, 1) for the left tail increased as the quantiles increased. In the studies of Kiragu & Kyalo, 2016 as well as Susan & Waititu, 2015 the authors modelled Extreme value theory using the GARCH(1,1). The current study is different from them as it modelled extreme value theory using EGARCH (1, 1) and SVR-EGARCH (1, 1) models.

5.3 LIMITATIONS

The quantitative study was conducted from 7th August 2008 – 7th August 2018. The study used univariate GARCH type models, SVR-GARCH and Extreme value theory. Secondary data was sourced from the JSE. The researcher has no control over how the data collection process was carried out. For instance, the accuracy of the capturing process. This may have an impact on the results. The study is in the context of South Africa.

5.4 FUTURE STUDIES

Future studies should use Support Vector regression-GARCH type models in modelling volatility in the financial markets as it produces accurate results compared to the standard GARCH models. Multivariate datasets should be used as this will help in comparison of different volatility models. Stochastic Volatility models should be compared with the Support Vector Regression models.

BIBLIOGRAPHY

Aamir, M. & Shabri, A. 2016. Modelling and forecasting monthly crude oil price of Pakistan: A comparative study of ARIMA, GARCH and ARIMA Kalman model. *AIP Conf. Proc.* 1750(3):2365–2371.

Abdalla, S.Z.S. 2012. Modelling Exchange Rate Volatility using GARCH Models: Empirical Evidence from Arab Countries. *Int. J. Econ. Financ.* 4(3):216–229.

Aboura, S. 2014. When the U.S. Stock Market Becomes Extreme? *Risks.* 2(2):211–225.

Ahmed, R. & Shabri, A. 2013. Fitting GARCH Models to Crude Oil Spot Price Data. *Life Sci. J.* 10(4):654–661.

Ahmed, R.A. & Shabri, A. Bin. 2014. Daily crude oil price forecasting model using arima, generalized autoregressive conditional heteroscedastic and Support Vector Machines. *Am. J. Appl. Sci.* 11(3):425–432.

Alam, M., Emura, K., Farnham, C., & Yuan, J. 2018. Best-Fit Probability Distributions and Return Periods for Maximum Monthly Rainfall in Bangladesh. *Climate.* 6(1):9.

Arce, P., Antognini, J., Kristjanpoller, W., & Salinas, L. 2015. An online vector error correction model for exchange rates forecasting. *ICPRAM 2015 - 4th Int. Conf. Pattern Recognit. Appl. Methods, Proc.* 2(January):193–200.

Ardia, D. 2008. Financial Risk Management with Bayesian Estimation of GARCH Models: Theory and Applications. , Lecture Notes in Economics and Mathematical Systems.

Arltová, M. & Fedorová, D. 2016. Selection of Unit Root Test on the Basis of Length of the Time Series and Value of AR(1) Parameter. *Statistika.* 96(3):47–64.

Atoi, N. V. 2014. Testing Volatility in Nigeria Stock Market Using GARCH Models. Central Bank of Nigeria. *J. Appl. Stat.* 5(2):65–93.

Awad, M. & Khanna, R. 2015. Efficient learning machines: Theories, concepts, and applications for engineers and system designers. *Effic. Learn. Mach. Theor. Concepts, Appl. Eng. Syst. Des.* :1–248.

Będowska-Sójka, B. & Kliber, A. 2010. Realized Volatility Versus GARCH and Stochastic Volatility Models. The Evidence from the WIG20 index and the EUR/PLN Foreign Exchange Market. *Prz. Stat.* 4(z.4):105–127.

Bezerra, P.C.S. & Albuquerque, P.H.M. 2017. Volatility forecasting via SVR–GARCH with mixture of Gaussian kernels. *Comput. Manag. Sci. Springer Berlin Heidelberg.* 14(2):179–196.

Bildirici, M. & Ersin, O.O. 2013. Support Vector Machine GARCH and Neural Network GARCH Models in Modeling Conditional Volatility: An Application to Turkish Financial Markets. *Ssrn.*

Blischke W.R., Rezaul Karim M., Prabhakar Murthy D.N. (2011) Preliminary Data Analysis. In: Warranty Data Collection and Analysis. Springer Series in Reliability Engineering. Springer, London

Bouseba, F.Z. & Zeghdoudi, H. 2015. Use of the garch models to energy markets: Oil price volatility. *Glob. J. Pure Appl. Math.* 11(6).

Brooks, C. 2008. Introductory Economics for finance.

Çevik, E.I., Atukeren, E., & Korkmaz, T. 2013. Nonlinearity and nonstationarity in international art market prices: Evidence from Markov-switching ADF unit root tests. *Empir. Econ.* 45(2):675–695.

Chen, C., Twycross, J., & Garibaldi, J.M. 2017. A new accuracy measure based on bounded relative error for time series forecasting. :1–24.

Cheteni, P. 2016. Stock market volatility using GARCH models: Evidence from South Africa and China stock markets. *Munich Pers. RePEc Arch.* (77355).

Chokri Slim. 2015. Improved Crude Oil Price Forecasting With Statistical Learning Methods. *J. Mod. Account. Audit.* 11(1):51–62.

Chung, S.S. & Zhang, S. 2017. Volatility estimation using support vector machine: Applications to major foreign exchange rates. *Electron. J. Appl. Stat. Anal.* 10(2):499–511.

Cristiani-d'Ornano, I., Ksaier, A., & Etienne Carlotti, J. 2010. CRUDE OIL PRICE: HOW TO ANTICIPATE ITS FUTURE TRAJECTORY? A specific phenomenon of volatility clustering Isabelle Cristiani-d'Ornano 1 , Ahmed Ksaier 1 , Jean-Etienne Carlotti 2. 22(1):1–10.

Daddikar, P. and Rajgopal, M. (2016). SYMMETRIC AND ASYMMETRIC VOLATILITY MODELLING FOR CRUDE OIL PRICES IN INDIA. *ELK ASIA PACIFIC JOURNAL OF FINANCE AND RISK MANAGEMENT*, [online] 7(3). Available at: https://www.elkjournals.com/MasterAdmin/UploadFolder/SYMMETRIC%20AND%20ASYMMETRIC%20VOLATILITY%20MODELLING%20FOR%20CRUDE%20OIL%20PRICES%20IN%20INDIA/SYMMETRIC%20AND%20ASYMMETRIC%20VOLATILITY%20MODELLING%20FOR%20CRUDE%20OIL%20PRICES%20IN%20INDIA.pdf [Accessed 15 May 2018].

Datta & Mukhopadhyay, C.K. 2011. Relationship between Inflation and Economic Growth in Malaysia - An Econometric Review. *2011 Int. Conf. Econ. Financ. Res.* 4(1):415–419.

Dicks, A., Conradie, W.J., & de Wet, T. 2014. Value At Risk and Expected Shortfall : Traditional Measures and Extreme Value Theory Enhancements. 38(3).

Frad, H. & Zouari, E. 2014. Estimation of value-at-risk measures in the Islamic stock market : Approach based on Extreme Value Theory (EVT). 3(2):15–20.

Francq, C. & Zakoian, J.-M. 2010. GARCH Models.

Gaston, R.T. 2016. Modelling of Volatility in the South African Mining Sector : Application of ARCH and GARCH Models By.

Geng, L. & Zhang, Z. 2015. Forecast of Stock Index Volatility Using Grey GARCH-Type Models. (1):93–98.

Geng, L.Y. & Liang, Y.G. 2011. Prediction on Fund Volatility Based on SVRGM-GARCH Model. *Adv. Mater. Res.* 403–408:3763–3768.

Ghalanos, A. 2018. Introduction to the rugarch package . :1–36.

Gilli, M. & Këllezi, E. 2006. An application of extreme value theory for measuring financial risk. *Comput. Econ.* 27(2–3):207–228.

Halder, S.C., Das, K., Jones, D., & Average, I. 2016. Understanding extreme stock trading volume by generalized Pareto distribution. 2:45–60.

Halunga, A.G., Orme, C.D., & Yamagata, T. 2015. A heteroskedasticity robust Breusch–Pagan test for Contemporaneous correlation in dynamic panel data models. *J. Econom.* 198(2):209–230.

Huang, C., Gao, F., & Jiang, H. 2014. Combination of biorthogonal wavelet hybrid kernel OCSVM with feature weighted approach based on EVA and GRA in financial distress prediction. *Math. Probl. Eng.* 2014.

Jammazi, R. & Nguyen, D.K. 2017. Estimating and forecasting portfolio's Value-at-Risk with wavelet-based extreme value theory: Evidence from crude oil prices and US exchange rates. *J. Oper. Res. Soc. Palgrave Macmillan UK.* 68(11):1352–1362.

Jan van Greunen, André Heymans, Chris van Heerden, & Gary van Vuuren. 2014. The prominence of stationarity in time series forecasting. *J. Stud. Econ. Econom.* 38(1):16.

Katchova, A. 2013. Time Series ARIMA Models. *Time Ser. ARIMA Model.* :1–22.

Kavuncuoglu, H., Kavuncuoglu, E., Merve, S., Benli, B., & Sagdic, O. 2018. Prediction of the antimicrobial activity of walnut (Juglans regia L .) kernel aqueous extracts using artificial neural network and multiple linear regression Prediction of the antimicrobial activity of walnut (Juglans regia L .) kernel aqueous extract. *J. Microbiol. Methods. Elsevier.* 148(April):78–86.

Kiragu, K.A. & Kyalo, J. 2016. Extreme Values Modelling of Nairobi Securities Exchange Index. *Am. J. Theor. Appl. Stat.* 5(4):234–241.

Klein, T. & Walther, T. 2016. Oil price volatility forecast with mixture memory GARCH. *Energy Econ.* 58(October 2017):46–58.

Knight, J. & Satchell, S. 2007. Forecasting Volatility in the Financial Markets. , University Business.

Kosapattarapim, C., Lin, Y.-X., & Mccrae, M. 2012. Evaluating the volatility forecasting performance of best fitting GARCH models in emerging Asian stock markets Evaluating the volatility forecasting performance of best fitting GARCH models in emerging Asian stock markets Evaluating the Volatility Forecas. *Int. J. Math. Stat.* 12(2):1–15.

Kristjanpoller, W. & Minutolo, M.C. 2016. Forecasting volatility of oil price using an artificial neural network-GARCH model. *Expert Syst. Appl. Elsevier Ltd.* 65:233–241.

Lai, L. & Liu, J. 2014. Support Vector Machine and Least Square Support Vector Machine Stock Forecasting Models. *Comput. Sci. Inf. Technol.* 2(1):30–39.

Li, L. 2015. A Comparative Study of GARCH and EVT models in Modeling Value-at-Risk (VaR).

Li, L. 2017. A Comparative Study of GARCH and EVT Model in Modeling Value-at-Risk. *MPRA.* (6461).

Li, Y. 2014. Estimating and Forecasting APARCH-Skew- *t* Model by Wavelet Support Vector Machines. *J. Forecast.* 33(4):259–269.

Lim, C.M. & Sek, S.K. 2013. Comparing the Performances of GARCH-type Models in Capturing the Stock Market Volatility in Malaysia. *Procedia Econ. Financ. Elsevier B.V.* 5(13):478–487.

Lux, M., Härdle, W.K., Lessmann, S., Lux, M., & H, W.K. 2018. Data Driven Value-at-Risk Forecasting using a SVR-GARCH-KDE Hybrid International Research Training Group 1792 Data Driven Value-at-Risk Forecasting using a SVR-GARCH-KDE Hybrid.

Lux, T., Segnon, M., & Gupta, R. 2015. Modeling and Forecasting Crude Oil Price Volatility: Evidence from Historical and Recent Data. *Work. Pap.* (March):5–42.

Maqsood, A., Safdar, S., Shafi, R., & Lelit, N.J. 2017. Modeling Stock Market Volatility Using GARCH Models: A Case Study of Nairobi Securities Exchange (NSE). *Open J. Stat.* 07(02):369–381.

Marsani, M.F., Shabri, A., & Jan, N.A.M. 2017. Examine generalized lambda distribution fitting performance: An application to extreme share return in Malaysia. *Malaysian J. Fundam. Appl. Sci.* 13(3):229–236.

N.Gujarati, D. & Dawn C.Porter. 2008. BASIC ECONOMETRICS.

Nanda, M.A., Seminar, K.B., Nandika, D., & Maddu, A. 2018. A comparison study of kernel functions in the support vector machine and its application for termite detection. *Inf.* 9(1).

Nortey, E.N.N., Asare, K., & Mettle, F.O. 2015. Extreme value modelling of Ghana stock exchange index. *Springerplus. Springer International Publishing.* 4(1):1–17.

Noshkov, A. & Demirtas, Z. 2017. Managing Risk with Energy Commodities using Value - at - Risk and Extreme Value Theory.

Ødegaard, B.A. 2018. Time series : Variance modelling SP500 Returns SP500 Squared Residuals. (November):1–9.

Okeyo, J., Ngare, P., & Ivivi, M. 2016. Modelling Inflation Rate Volatility in Kenya Using Arch-Type Model Family Modelling Inflation Rate Volatility in Kenya Using Arch -Type Model Family.

Onwukwe, C.E., Samson, T.K., & Lipcsey, Z. 2014. Modelling and Forecasting Daily Returns Volatility of Nigerian Banks. 10(15):449–467.

Ou, P. & Wang, H. 2010. Financial Volatility Forecasting by Least Square Support Vector Machine Based on GARCH , EGARCH and GJR Models : Evidence from ASEAN Stock Markets. *Int. J. Econ. Financ.* 2(1):51–64.

Pfaff, B. 2016. Financial Risk Modelling and Portfolio Optimization with R.

Poon, S.-H. 2005. A Practical Guide to Forecasting Financial Market Volatility Library of Congress Cataloging-in-Publication Data.

Qu, H. & Zhang, Y. 2016. A new kernel of support vector regression for forecasting high-frequency stock returns. *Math. Probl. Eng.* 2016.

Rani Das, K. & Imon, A.H.M.R. 2016. A Brief Review of Tests for Normality. *Am. J. Theor. Appl. Stat.* 5(1):5.

Razali, N.M. & Wah, Y.B. 2011. Power comparisons of Shapiro-Wilk , Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *J. Stat. Model. Anal.* 2(1):21–33.

Rhoda, M. 2013. Modelling volatility and financial market risk of series on the Johannesburg Stock Exchange. *J. Econom. Soc.* 53(9):1689–1699.

Rösch, A. & Harald, S. 2012. Does Opec Announcement Policy Matter in the Creation of Crude Oil Price Volatility ? :1–8.

Sahoo, A., Garg, A., Jangid, J., & Chandra, A. 2018. Improving the Volatility Forecasts of GARCH Family Models with the Recurrent Neural Networks.

Salisu, a a & Fasanya, I.O. 2012. Comparative performance of volatility models for oil price. *Int. J. Energy Econ. Policy.* 2(3):167–183.

SAS Institute Inc. 2011. SAS/STAT® 9.3 User's Guide. Cary, NC: SAS Institute Inc.

Shabani, M., Gharneh, N.S., & Esfahanipour, A. 2016. Comparing the performance of GARCH (p,, q) models with different methods of estimation for forecasting crude oil market volatility. *J. Ind. Syst. Eng.* 9(4):0.

Sowdagur, V & Narsoo, J. 2017. Forecasting Value-at-Risk using GARCH and Extreme-

Value-Theory Approaches for Daily Returns. *Int. J. Stat. Appl.* 7(2):137–151.

Susan, M. & Waititu, A.G. 2015. Modelling Oil Price Risk. 4(6):539–546.

Swanson, D.A., Tayman, J., & Bryan, T.M. 2011. MAPE-R: A rescaled measure of accuracy for cross-sectional subnational population forecasts. *J. Popul. Res.* 28(2–3):225–243.

Tolikas, K. & Gettinby, G.D. 2009. Modelling the distribution of the extreme share returns in Singapore. *J. Empir. Financ. Elsevier B.V.* 16(2):254–263.

Tsay, R. 2013. An introduction to Analysis of Financial Data with R.

Tsay, R.S. 2010. Analysis of Financial Time Series: Third Edition. , Analysis of Financial Time Series: Third Edition.

Ugurlu, E., Thalassinos, E., & Muratoglu, Y. 2014. Modeling Volatility in the Stock Markets using GARCH Models : European Emerging Economies and Turkey. II(3):72–87.

Wang, Y. 2013. Stock price direction prediction by directly using prices data: an empirical study on the KOSPI and HSI.

Wegner, T. 2012. Applied Business Statistics: Methods and Excel-Based Applications. :420.

Wennstrom, A. 2014. Volatility Forecasting Performance: Evaluation of GARCH type volatility models on Nordic equity indices. *Spring.* :2–61.

Wong, B. 2014. Introduction to (Generalized) Autoregressive Conditional Heteroskedasticity Models in Time Series Econometrics. (June):1–12.

Yaziz, S.R., Ahmad, M.H., Lee, C.N., & Muhammad, N. 2011. A Comparative Study on Box-Jenkins and GARCH Models in Forecasting Crude Oil Prices. *J. Appl. Sci.* 11(7):1129–1135.

Yekkehkhany, B., Safari, A., Homayouni, S., & Hasanlou, M. 2014. A comparison study of different kernel functions for SVM-based classification of multi-temporal polarimetry SAR data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.* 40(2W3):281–285.

Yu, B. & Yang, S. 2013. Forecasting Modeling for Short-Termed Electricity Load Based on GARCH-SVR Method. *Int. J. Adv. Comput. Technol.* 5(5):559–566.

Yu, H.-C. & Shih, T.-L. 2007. PROBABILITY DISTRIBUTION OF RETURNS AND VOLATILITY AMONG CRUDE OIL, GOLD, AND DOLLAR/ BRITISH POUND MARKETS. (February).

Yu, H.C. & Shih, T.L. 2011. Gold, crude oil and the weekend effect: A probability distribution approach. *Invest. Manag. Financ. Innov.* 8(2):39–51.

Zin, W.Z.W., Safari, M.A.M., Jaaman, S.H., & Yie, W.L.S. 2014. Probability distribution of extreme share returns in Malaysia. *AIP Conf. Proc.* 1613(Soric 2013):325–333.

Zivot, E. & Wang, J. 2006. Modelling Financial Time Series with S-PLUS. :1020.

APPENDIX 1

Tests	
Normality tests	Jarque Bera, Shapiro Wilks, Skewness and Kurtosis
Unit root	ADF and Philips Perron
ARCH tests	Ljung Box test, LM test, Breusch Pagan test
Accuracy Measures	MSE, RMSE
Error Distributions	Normal, Student t, Skewed student t, GED
Kernel Selection	Radial, polynomial kernel, sigmoid and linear kernels
Models	GARCH(1,1) ,EGARCH(1,1), GJR-GARCH(1,1) and FIGARCH(1,d,1)
Risk Measures	Value at Risk and Expected Shortfall

APPENDIX 2: OBJECTIVES

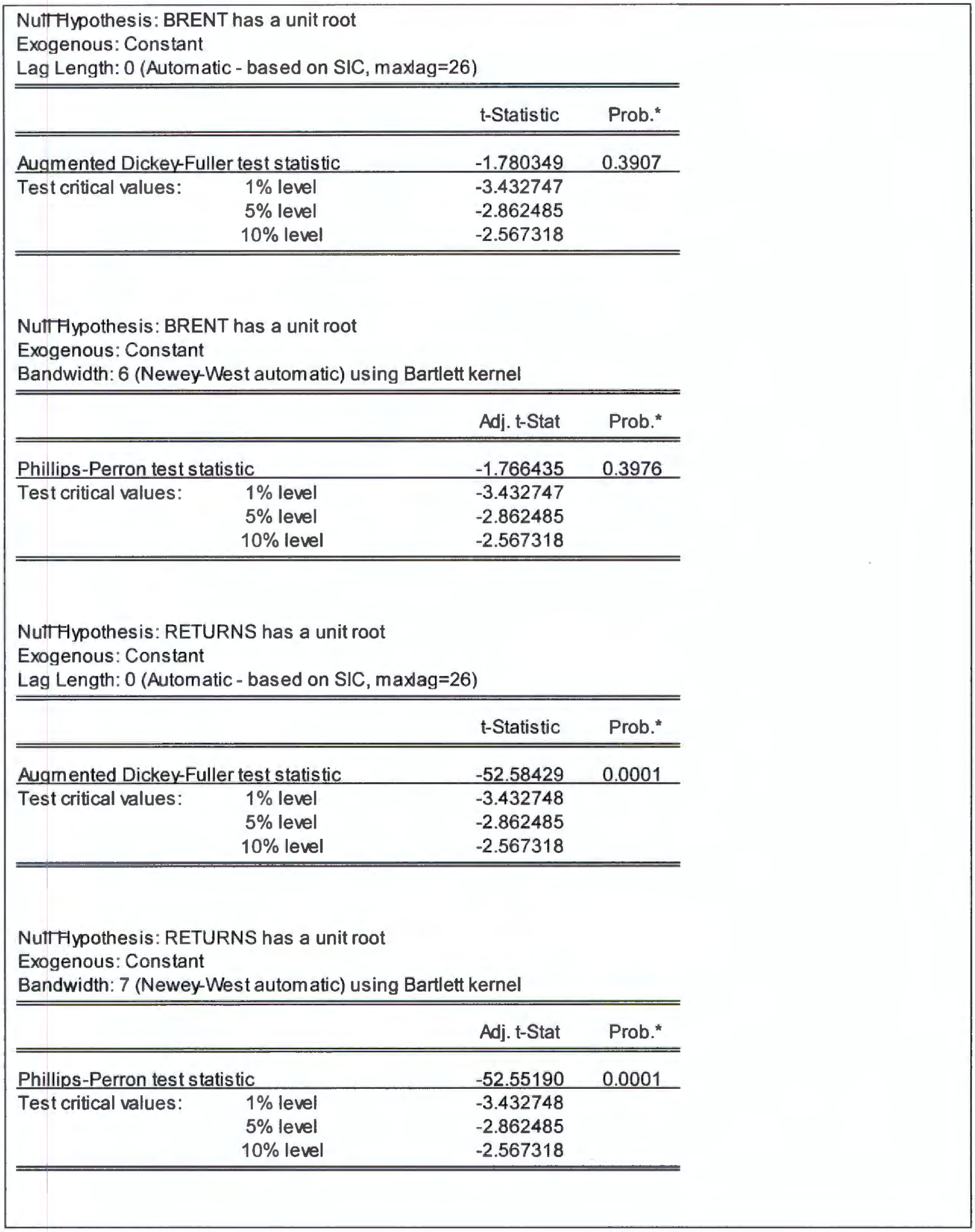
- To model symmetric and asymmetric GARCH and SVR-GARCH models.
  The study modeled GARCH(1,1), EGARCH(1,1), GJR-GARCH(1,1), FIGARCH(1,d,1) and SVR-GARCH(1,1), SVR- EGARCH(1,1), SVR- GJR-GARCH(1,1) and SVR-FIGARCH(1,d,1).

- Evaluate the performance of GARCH and SVR-GARCH based on accuracy measures and make predictions.
  SVR-EGARCH (1, 1) outperformed all the GARCH models. For the GARCH models EGARCH (1, 1) outperformed GARCH(1,1), GJR-GARCH(1,1) and FIGARCH(1,d,1).

- Use the best GARCH and SVR-GARCH to model extreme risk of Brent crude oil prices.
  The study used EGARCH (1, 1) and SVR-EGARCH (1, 1) to model extreme risk of the oil prices. The study used the POT method to fit the Generalized Pareto Distribution (GPD). The GPD adequately fitted the tails of the returns.

APPENDIX 3: RESULTS

ADF and PP unit root tests of Brent and the Returns

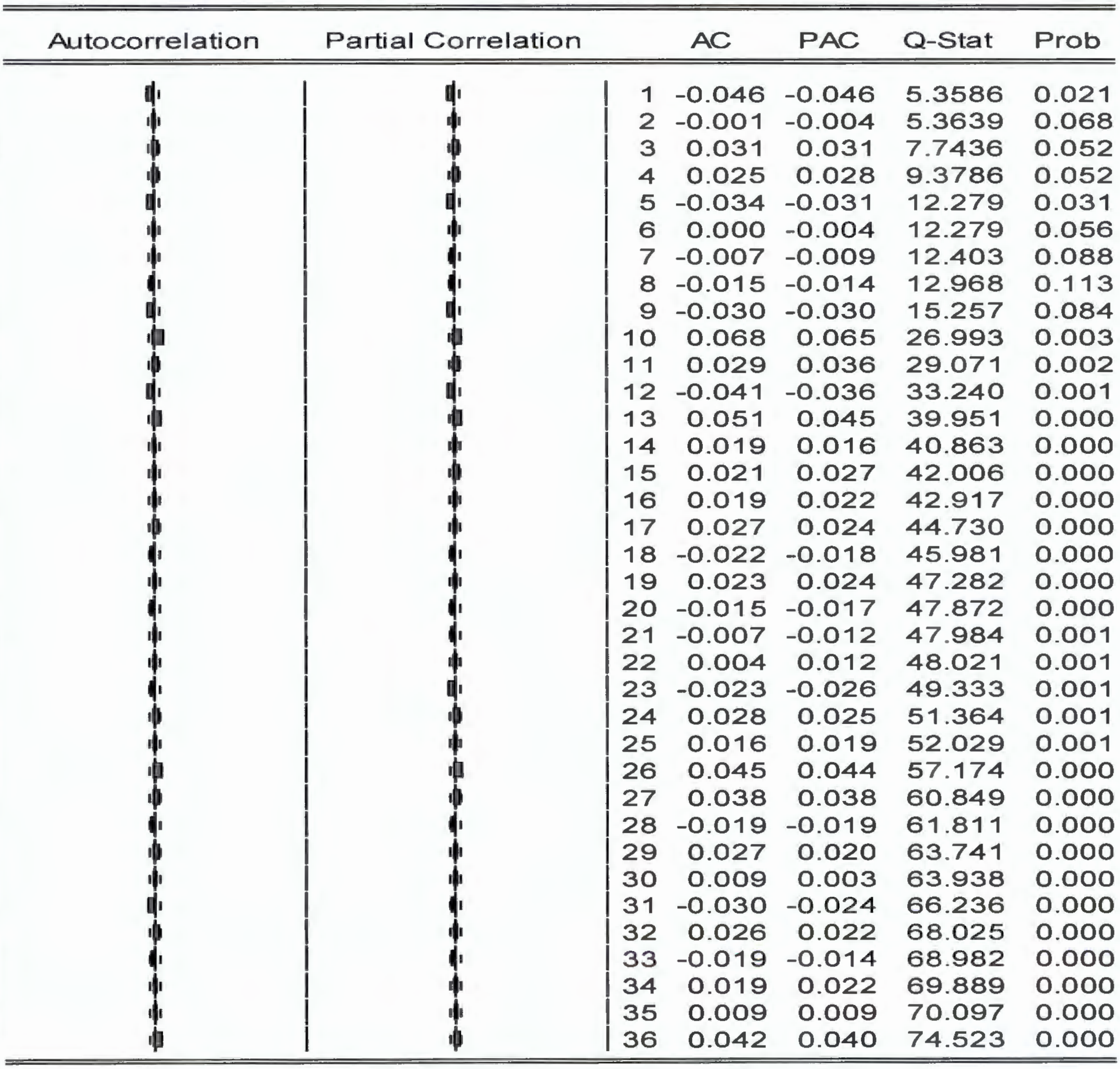


```
Null Hypothesis: BRENT has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=26)
======================================================================
                                              t-Statistic    Prob.*
======================================================================
Augmented Dickey-Fuller test statistic         -1.780349     0.3907
Test critical values:        1% level          -3.432747
                             5% level          -2.862485
                            10% level          -2.567318
======================================================================


Null Hypothesis: BRENT has a unit root
Exogenous: Constant
Bandwidth: 6 (Newey-West automatic) using Bartlett kernel
======================================================================
                                              Adj. t-Stat    Prob.*
======================================================================
Phillips-Perron test statistic                 -1.766435     0.3976
Test critical values:        1% level          -3.432747
                             5% level          -2.862485
                            10% level          -2.567318
======================================================================


Null Hypothesis: RETURNS has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=26)
======================================================================
                                              t-Statistic    Prob.*
======================================================================
Augmented Dickey-Fuller test statistic         -52.58429     0.0001
Test critical values:        1% level          -3.432748
                             5% level          -2.862485
                            10% level          -2.567318
======================================================================


Null Hypothesis: RETURNS has a unit root
Exogenous: Constant
Bandwidth: 7 (Newey-West automatic) using Bartlett kernel
======================================================================
                                              Adj. t-Stat    Prob.*
======================================================================
Phillips-Perron test statistic                 -52.55190     0.0001
Test critical values:        1% level          -3.432748
                             5% level          -2.862485
                            10% level          -2.567318
======================================================================
```

Returns: Heteroscedascity using ARCH

```
Heteroskedasticity Test: ARCH

F-statistic          131.6376   Prob. F(1,2521)          0.0000
Obs*R-squared        125.2043   Prob. Chi-Square(1)      0.0000
```

ACF and PACF of the returns



```
Date: 10/11/18   Time: 13:02
Sample: 8/07/2008 8/07/2018
Included observations: 2524

 Autocorrelation    Partial Correlation        AC      PAC     Q-Stat   Prob

                                         1  -0.046  -0.046   5.3586   0.021
                                         2  -0.001  -0.004   5.3639   0.068
                                         3   0.031   0.031   7.7436   0.052
                                         4   0.025   0.028   9.3786   0.052
                                         5  -0.034  -0.031   12.279   0.031
                                         6   0.000  -0.004   12.279   0.056
                                         7  -0.007  -0.009   12.403   0.088
                                         8  -0.015  -0.014   12.968   0.113
                                         9  -0.030  -0.030   15.257   0.084
                                        10   0.068   0.065   26.993   0.003
                                        11   0.029   0.036   29.071   0.002
                                        12  -0.041  -0.036   33.240   0.001
                                        13   0.051   0.045   39.951   0.000
                                        14   0.019   0.016   40.863   0.000
                                        15   0.021   0.027   42.006   0.000
                                        16   0.019   0.022   42.917   0.000
                                        17   0.027   0.024   44.730   0.000
                                        18  -0.022  -0.018   45.981   0.000
                                        19   0.023   0.024   47.282   0.000
                                        20  -0.015  -0.017   47.872   0.000
                                        21  -0.007  -0.012   47.984   0.001
                                        22   0.004   0.012   48.021   0.001
                                        23  -0.023  -0.026   49.333   0.001
                                        24   0.028   0.025   51.364   0.001
                                        25   0.016   0.019   52.029   0.001
                                        26   0.045   0.044   57.174   0.000
                                        27   0.038   0.038   60.849   0.000
                                        28  -0.019  -0.019   61.811   0.000
                                        29   0.027   0.020   63.741   0.000
                                        30   0.009   0.003   63.938   0.000
                                        31  -0.030  -0.024   66.236   0.000
                                        32   0.026   0.022   68.025   0.000
                                        33  -0.019  -0.014   68.982   0.000
                                        34   0.019   0.022   69.889   0.000
                                        35   0.009   0.009   70.097   0.000
                                        36   0.042   0.040   74.523   0.000
```

GARCH MODELLING

GARCH (1,1) - SSTD

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : sstd

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.000026    0.000292 -0.088437 0.929529
ma1    -0.056023    0.019943 -2.809212 0.004966
omega   0.000001    0.000002  0.932890 0.350877
alpha1  0.058620    0.011574  5.064801 0.000000
beta1   0.939482    0.011939 78.688102 0.000000
skew    0.935079    0.026200 35.690417 0.000000
shape   8.206637    1.170993  7.008272 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.000026    0.000326 -0.07925 0.936834
ma1    -0.056023    0.020048 -2.79438 0.005200
omega   0.000001    0.000005  0.30068 0.763656
alpha1  0.058620    0.042597  1.37617 0.168770
beta1   0.939482    0.042831 21.93487 0.000000
skew    0.935079    0.026241 35.63478 0.000000
shape   8.206637    1.954764  4.19827 0.000027

LogLikelihood : 6579.458

Information Criteria
------------------------------------

Akaike       -5.2080
Bayes        -5.1918
Shibata      -5.2080
Hannan-Quinn -5.2021

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      2.082 0.14904
Lag[2*(p+q)+(p+q)-1][2]     2.517 0.08812
Lag[4*(p+q)+(p+q)-1][5]     3.317 0.35578
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1874  0.6650
Lag[2*(p+q)+(p+q)-1][5]    1.7289  0.6837
```

```
Lag[4*(p+q)+(p+q)-1][9]    3.0389  0.7525
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
             Statistic Shape Scale P-Value
ARCH Lag[3]    0.03358 0.500 2.000  0.8546
ARCH Lag[5]    1.01197 1.440 1.667  0.7294
ARCH Lag[7]    1.77304 2.315 1.543  0.7652

Nyblom stability test
------------------------------------
Joint Statistic:  57.673
Individual Statistics:
mu      0.2426
ma1     0.1451
omega  12.7781
alpha1  0.1521
beta1   0.1252
skew    0.2005
shape   0.4492

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.69 1.9 2.35
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           1.5406 0.12355
Negative Sign Bias  0.2958 0.76738
Positive Sign Bias  0.5794 0.56238
Joint Effect        8.2089 0.04189  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     19.82       0.4055
2    30     29.45       0.4420
3    40     40.53       0.4026
4    50     37.01       0.8957


Elapsed time : 2.776649
```

GARCH (1,1)- STD

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000173    0.000282  0.61152 0.540854
ma1    -0.051483    0.019858 -2.59258 0.009526
omega   0.000002    0.000002  0.97617 0.328980
alpha1  0.058677    0.011745  4.99603 0.000001
beta1   0.939241    0.012136 77.39341 0.000000
shape   7.988133    1.102792  7.24356 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000173    0.000320  0.53919 0.589757
ma1    -0.051483    0.019531 -2.63600 0.008389
omega   0.000002    0.000005  0.31118 0.755661
alpha1  0.058677    0.043883  1.33713 0.181178
beta1   0.939241    0.044066 21.31419 0.000000
shape   7.988133    1.944384  4.10831 0.000040

LogLikelihood : 6576.58

Information Criteria
------------------------------------

Akaike       -5.2065
Bayes        -5.1926
Shibata      -5.2065
Hannan-Quinn -5.2014

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.536  0.2152
Lag[2*(p+q)+(p+q)-1][2]     1.947  0.2317
Lag[4*(p+q)+(p+q)-1][5]     2.726  0.5034
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1814  0.6701
Lag[2*(p+q)+(p+q)-1][5]    1.8347  0.6580
Lag[4*(p+q)+(p+q)-1][9]    3.1618  0.7320
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]   0.04001 0.500 2.000  0.8415
```

```
ARCH Lag[5]    1.04726 1.440 1.667  0.7190
ARCH Lag[7]    1.79377 2.315 1.543  0.7609

Nyblom stability test
------------------------------------
Joint Statistic:  52.8076
Individual Statistics:
mu       0.2474
ma1      0.1434
omega   11.6811
alpha1   0.1600
beta1    0.1273
shape    0.4428

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.49 1.68 2.12
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           1.6528 0.09850   *
Negative Sign Bias  0.2676 0.78902
Positive Sign Bias  0.4726 0.63656
Joint Effect        8.4494 0.03758  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     25.35      0.14933
2    30     39.81      0.08707
3    40     44.43      0.25358
4    50     60.03      0.13423
```

GARCH (1,1)- NORMAL DISTRIBUTION

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000040    0.000301  0.13282  0.89433
ma1    -0.038023    0.021943 -1.73281  0.08313
omega   0.000002    0.000005  0.35103  0.72557
alpha1  0.053929    0.033124  1.62812  0.10350
beta1   0.942329    0.034925 26.98118  0.00000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000040     0.00034  0.117675 0.906325
ma1    -0.038023     0.11078 -0.343233 0.731423
omega   0.000002     0.00007  0.025646 0.979539
alpha1  0.053929     0.44494  0.121205 0.903529
beta1   0.942329     0.47081  2.001516 0.045337

LogLikelihood : 6542.475

Information Criteria
------------------------------------

Akaike       -5.1802
Bayes        -5.1687
Shibata      -5.1803
Hannan-Quinn -5.1761

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.3192  0.5721
Lag[2*(p+q)+(p+q)-1][2]    0.6710  0.9097
Lag[4*(p+q)+(p+q)-1][5]    1.4135  0.8672
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.2908  0.5897
Lag[2*(p+q)+(p+q)-1][5]    2.2857  0.5526
Lag[4*(p+q)+(p+q)-1][9]    3.3682  0.6969
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.1997 0.500 2.000  0.6550
ARCH Lag[5]    0.8658 1.440 1.667  0.7732
ARCH Lag[7]    1.3894 2.315 1.543  0.8433
```

```
Nyblom stability test
------------------------------------
Joint Statistic:  49.96
Individual Statistics:
mu     0.15903
ma1    0.18761
omega  8.61557
alpha1 0.06976
beta1  0.05996

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.28 1.47 1.88
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           1.8386 0.06609   *
Negative Sign Bias  0.4135 0.67925
Positive Sign Bias  0.1672 0.86723
Joint Effect        9.2299 0.02638  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     53.48    3.961e-05
2    30     65.77    1.123e-04
3    40     76.00    3.565e-04
4    50     96.64    5.839e-05
```

GARCH (1,1) –GED

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : ged

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000155    0.000275  0.56425 0.572587
ma1    -0.054175    0.018918 -2.86375 0.004187
omega   0.000002    0.000002  0.82730 0.408070
alpha1  0.055342    0.013199  4.19286 0.000028
beta1   0.941744    0.013998 67.27568 0.000000
shape   1.458942    0.052492 27.79337 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000155    0.000308  0.50345 0.614647
ma1    -0.054175    0.018075 -2.99722 0.002725
omega   0.000002    0.000008  0.21191 0.832175
alpha1  0.055342    0.059098  0.93644 0.349049
beta1   0.941744    0.061164 15.39703 0.000000
shape   1.458942    0.108743 13.41639 0.000000

LogLikelihood : 6575.178

Information Criteria
------------------------------------

Akaike       -5.2054
Bayes        -5.1915
Shibata      -5.2054
Hannan-Quinn -5.2003

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.811  0.1784
Lag[2*(p+q)+(p+q)-1][2]     2.229  0.1464
Lag[4*(p+q)+(p+q)-1][5]     3.026  0.4247
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.2891  0.5908
Lag[2*(p+q)+(p+q)-1][5]    2.1105  0.5925
Lag[4*(p+q)+(p+q)-1][9]    3.2959  0.7093
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]   0.09436 0.500 2.000  0.7587
```

```
ARCH Lag[5]    0.94062 1.440 1.667  0.7506
ARCH Lag[7]    1.55960 2.315 1.543  0.8093

Nyblom stability test
------------------------------------
Joint Statistic:  57.3188
Individual Statistics:
mu      0.25089
ma1     0.16530
omega  11.49402
alpha1  0.10999
beta1   0.08791
shape   0.29220

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.49 1.68 2.12
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           1.5496 0.12136
Negative Sign Bias  0.4417 0.65872
Positive Sign Bias  0.4124 0.68007
Joint Effect        8.1430 0.04315  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     18.85       0.4663
2    30     33.82       0.2459
3    40     38.82       0.4780
4    50     56.39       0.2182


Elapsed time : 2.418206
```

EGARCH (1,1) SSTD

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : eGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : sstd

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu     -0.000284    0.000280   -1.0145 0.310321
ma1    -0.047533    0.019894   -2.3893 0.016879
omega  -0.041833    0.001590  -26.3169 0.000000
alpha1 -0.061016    0.008072   -7.5592 0.000000
beta1   0.994778    0.000124 7996.3379 0.000000
gamma1  0.087047    0.002716   32.0473 0.000000
skew    0.935972    0.026508   35.3088 0.000000
shape   9.644532    1.433235    6.7292 0.000000

Robust Standard Errors:
        Estimate  Std. Error    t value Pr(>|t|)
mu     -0.000284    0.000294   -0.96688 0.333602
ma1    -0.047533    0.019746   -2.40726 0.016073
omega  -0.041833    0.002141  -19.54169 0.000000
alpha1 -0.061016    0.008768   -6.95882 0.000000
beta1   0.994778    0.000169 5899.51372 0.000000
gamma1  0.087047    0.007406   11.75358 0.000000
skew    0.935972    0.027226   34.37742 0.000000
shape   9.644532    1.404798    6.86542 0.000000

LogLikelihood : 6607.042

Information Criteria
------------------------------------

Akaike       -5.2290
Bayes        -5.2105
Shibata      -5.2291
Hannan-Quinn -5.2223

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.5708  0.4500
Lag[2*(p+q)+(p+q)-1][2]    0.7569  0.8723
Lag[4*(p+q)+(p+q)-1][5]    1.1432  0.9232
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1206  0.7284
Lag[2*(p+q)+(p+q)-1][5]    5.7079  0.1053
Lag[4*(p+q)+(p+q)-1][9]    7.5916  0.1545
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.2136 0.500 2.000  0.6439
ARCH Lag[5]    1.9045 1.440 1.667  0.4927
ARCH Lag[7]    2.5777 2.315 1.543  0.5966
```

```
Nyblom stability test
------------------------------------
Joint Statistic:  1.3351
Individual Statistics:
mu     0.4247
ma1    0.1422
omega  0.2314
alpha1 0.1092
beta1  0.2519
gamma1 0.1238
skew   0.2332
shape  0.2417

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.89 2.11 2.59
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias           1.2292 0.2191
Negative Sign Bias  0.5548 0.5791
Positive Sign Bias  0.2140 0.8306
Joint Effect        2.5695 0.4629


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     27.92      0.08503
2    30     37.93      0.12382
3    40     52.04      0.07904
4    50     66.21      0.05107


Elapsed time : 4.90531
```

EGARCH(1,1) -STD

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : eGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error    t value Pr(>|t|)
mu     -0.000110    0.000271   -0.40613 0.684646
ma1    -0.046370    0.020558   -2.25557 0.024098
omega  -0.041844    0.001489  -28.10274 0.000000
alpha1 -0.060988    0.008039   -7.58615 0.000000
beta1   0.994868    0.000127 7806.88520 0.000000
gamma1  0.085776    0.002507   34.21143 0.000000
shape   9.509176    1.370681    6.93756 0.000000

Robust Standard Errors:
        Estimate  Std. Error    t value Pr(>|t|)
mu     -0.000110    0.000279   -0.39424 0.693407
ma1    -0.046370    0.021143   -2.19315 0.028296
omega  -0.041844    0.002126  -19.68027 0.000000
alpha1 -0.060988    0.008635   -7.06271 0.000000
beta1   0.994868    0.000174 5720.03544 0.000000
gamma1  0.085776    0.007449   11.51571 0.000000
shape   9.509176    1.367383    6.95429 0.000000

LogLikelihood : 6604.288

Information Criteria
------------------------------------

Akaike       -5.2276
Bayes        -5.2115
Shibata      -5.2277
Hannan-Quinn -5.2218

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.4721  0.4920
Lag[2*(p+q)+(p+q)-1][2]    0.6412  0.9212
Lag[4*(p+q)+(p+q)-1][5]    1.0042  0.9462
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1377  0.7106
Lag[2*(p+q)+(p+q)-1][5]    5.8738  0.0963
Lag[4*(p+q)+(p+q)-1][9]    7.7495  0.1441
d.o.f=2

Weighted ARCH LM Tests
```

```
------------------------------------
              Statistic Shape Scale P-Value
ARCH Lag[3]     0.1881 0.500 2.000  0.6645
ARCH Lag[5]     1.8769 1.440 1.667  0.4991
ARCH Lag[7]     2.5210 2.315 1.543  0.6081

Nyblom stability test
------------------------------------
Joint Statistic:  1.1775
Individual Statistics:
mu     0.4584
ma1    0.1471
omega  0.2291
alpha1 0.1063
beta1  0.2471
gamma1 0.1394
shape  0.2409

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:          1.69 1.9 2.35
Individual Statistic:     0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias           1.2447 0.2134
Negative Sign Bias  0.5306 0.5958
Positive Sign Bias  0.1629 0.8706
Joint Effect        2.5420 0.4677


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     26.84      0.10844
2    30     40.60      0.07466
3    40     49.50      0.12088
4    50     72.00      0.01785

Elapsed time : 4.368148
```

EGARCH(1,1) –NORMAL DISTRIBUTION

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : eGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu     -0.000272    0.000288   -0.9448 0.344760
ma1    -0.031453    0.019118   -1.6452 0.099936
omega  -0.054680    0.000592  -92.3482 0.000000
alpha1 -0.060252    0.004300  -14.0130 0.000000
beta1   0.992925    0.000226 4401.1138 0.000000
gamma1  0.088637    0.010492    8.4483 0.000000

Robust Standard Errors:
        Estimate  Std. Error    t value Pr(>|t|)
mu     -0.000272    0.000338   -0.80457 0.421070
ma1    -0.031453    0.017861   -1.76095 0.078246
omega  -0.054680    0.003532  -15.48006 0.000000
alpha1 -0.060252    0.005299  -11.37067 0.000000
beta1   0.992925    0.000356 2790.26524 0.000000
gamma1  0.088637    0.015440    5.74076 0.000000

LogLikelihood : 6579.366

Information Criteria
------------------------------------

Akaike       -5.2087
Bayes        -5.1948
Shibata      -5.2087
Hannan-Quinn -5.2037

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                   0.000543  0.9814
Lag[2*(p+q)+(p+q)-1][2]  0.147271  0.9997
Lag[4*(p+q)+(p+q)-1][5]  0.509004  0.9933
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                    0.07229  0.7880
Lag[2*(p+q)+(p+q)-1][5]   4.76785  0.1726
Lag[4*(p+q)+(p+q)-1][9]   6.47000  0.2477
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]     0.120 0.500 2.000  0.7290
```

```
ARCH Lag[5]     1.637 1.440 1.667  0.5570
ARCH Lag[7]     2.300 2.315 1.543  0.6539

Nyblom stability test
------------------------------------
Joint Statistic:  0.8868
Individual Statistics:
mu     0.3633
ma1    0.1341
omega  0.1768
alpha1 0.1312
beta1  0.1943
gamma1 0.1658

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.49 1.68 2.12
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias           1.4587 0.1448    
Negative Sign Bias  0.4054 0.6852    
Positive Sign Bias  0.1440 0.8855    
Joint Effect        2.9796 0.3948    


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     40.44    2.866e-03
2    30     67.70    6.224e-05
3    40     68.04    2.714e-03
4    50     75.17    9.506e-03


Elapsed time : 2.063766
```

EGARCH (1,1)- GED

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : eGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : ged

Optimal Parameters
------------------------------------
        Estimate  Std. Error    t value Pr(>|t|)
mu     -0.000073    0.000266   -0.27271 0.785073
ma1    -0.048960    0.018936   -2.58550 0.009724
omega  -0.046895    0.001537  -30.50206 0.000000
alpha1 -0.058717    0.007555   -7.77187 0.000000
beta1   0.994289    0.000344 2894.01092 0.000000
gamma1  0.086712    0.011777    7.36266 0.000000
shape   1.526918    0.059984   25.45551 0.000000

Robust Standard Errors:
        Estimate  Std. Error    t value Pr(>|t|)
mu     -0.000073    0.000296   -0.24488 0.806553
ma1    -0.048960    0.018295   -2.67610 0.007448
omega  -0.046895    0.004388  -10.68697 0.000000
alpha1 -0.058717    0.008829   -6.65052 0.000000
beta1   0.994289    0.000550 1809.10243 0.000000
gamma1  0.086712    0.015695    5.52466 0.000000
shape   1.526918    0.073310   20.82825 0.000000

LogLikelihood : 6602.236

Information Criteria
------------------------------------

Akaike       -5.2260
Bayes        -5.2098
Shibata      -5.2260
Hannan-Quinn -5.2201

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.6995  0.4030
Lag[2*(p+q)+(p+q)-1][2]    0.8723  0.8135
Lag[4*(p+q)+(p+q)-1][5]    1.2436  0.9040
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1289  0.7196
Lag[2*(p+q)+(p+q)-1][5]    5.3654  0.1264
Lag[4*(p+q)+(p+q)-1][9]    7.2116  0.1821
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
```

```
              Statistic Shape Scale P-Value
ARCH Lag[3]      0.1631 0.500 2.000  0.6863
ARCH Lag[5]      1.8495 1.440 1.667  0.5055
ARCH Lag[7]      2.5370 2.315 1.543  0.6049

Nyblom stability test
------------------------------------
Joint Statistic:  1.265
Individual Statistics:
mu     0.4444
ma1    0.1594
omega  0.2019
alpha1 0.1178
beta1  0.2197
gamma1 0.1843
shape  0.1782

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.69 1.9 2.35
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias          1.23091 0.2185    
Negative Sign Bias 0.40088 0.6885    
Positive Sign Bias 0.08609 0.9314    
Joint Effect       2.45874 0.4828    


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     28.73      0.07042
2    30     39.05      0.10070
3    40     49.06      0.12974
4    50     69.74      0.02734


Elapsed time : 5.068053
```

GJR-GARCH(1,1) - SSTD

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : gjrGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : sstd

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.000218    0.000285 -0.76409 0.444812
ma1    -0.055251    0.018443 -2.99579 0.002737
omega   0.000001    0.000005  0.23107 0.817261
alpha1  0.016222    0.018629  0.87077 0.383880
beta1   0.950780    0.030125 31.56105 0.000000
gamma1  0.062337    0.018279  3.41027 0.000649
skew    0.935506    0.026348 35.50555 0.000000
shape   8.806508    1.579956  5.57389 0.000000

Robust Standard Errors:
        Estimate  Std. Error   t value Pr(>|t|)
mu     -0.000218    0.000705 -0.309165  0.75720
ma1    -0.055251    0.109775 -0.503315  0.61474
omega   0.000001    0.000070  0.016708  0.98667
alpha1  0.016222    0.270066  0.060066  0.95210
beta1   0.950780    0.422601  2.249827  0.02446
gamma1  0.062337    0.300083  0.207731  0.83544
skew    0.935506    0.032943 28.398119  0.00000
shape   8.806508   29.099643  0.302633  0.76217

LogLikelihood : 6595.261

Information Criteria
------------------------------------

Akaike       -5.2197
Bayes        -5.2012
Shibata      -5.2197
Hannan-Quinn -5.2130

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.236  0.2663
Lag[2*(p+q)+(p+q)-1][2]     1.537  0.4173
Lag[4*(p+q)+(p+q)-1][5]     2.041  0.6994
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                    0.04975  0.8235
Lag[2*(p+q)+(p+q)-1][5]   3.12475  0.3847
Lag[4*(p+q)+(p+q)-1][9]   4.67961  0.4779
d.o.f=2
```

```
Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.2119 0.500 2.000  0.6453
ARCH Lag[5]    1.9572 1.440 1.667  0.4808
ARCH Lag[7]    2.5328 2.315 1.543  0.6057

Nyblom stability test
------------------------------------
Joint Statistic:  91.2456
Individual Statistics:
mu      0.3671
ma1     0.1417
omega  19.9004
alpha1  0.1596
beta1   0.1858
gamma1  0.1766
skew    0.2227
shape   0.3372

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.89 2.11 2.59
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias           1.5296 0.1262
Negative Sign Bias  0.4501 0.6526
Positive Sign Bias  0.1125 0.9104
Joint Effect        3.8934 0.2732


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     23.92       0.1991
2    30     32.75       0.2879
3    40     38.09       0.5111
4    50     51.24       0.3860


Elapsed time : 3.96678
```

GJR-GARCH(1,1) -STD

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : gjrGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.000039    0.000312 -0.12471 0.900753
ma1    -0.052760    0.022502 -2.34471 0.019042
omega   0.000001    0.000008  0.14190 0.887158
alpha1  0.015256    0.030757  0.49601 0.619887
beta1   0.951136    0.047959 19.83227 0.000000
gamma1  0.062410    0.035417  1.76212 0.078049
shape   8.648954    3.795430  2.27878 0.022680

Robust Standard Errors:
        Estimate  Std. Error   t value Pr(>|t|)
mu     -0.000039    0.003051 -0.012759  0.98982
ma1    -0.052760    0.235685 -0.223860  0.82287
omega   0.000001    0.000189  0.006288  0.99498
alpha1  0.015256    0.681468  0.022387  0.98214
beta1   0.951136    1.076210  0.883783  0.37681
gamma1  0.062410    0.755995  0.082553  0.93421
shape   8.648954   80.097852  0.107980  0.91401

LogLikelihood : 6592.442

Information Criteria
------------------------------------

Akaike       -5.2183
Bayes        -5.2021
Shibata      -5.2183
Hannan-Quinn -5.2124

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.9614  0.3268
Lag[2*(p+q)+(p+q)-1][2]    1.2414  0.5905
Lag[4*(p+q)+(p+q)-1][5]    1.7209  0.7896
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                    0.06083  0.8052
Lag[2*(p+q)+(p+q)-1][5]   3.29037  0.3566
Lag[4*(p+q)+(p+q)-1][9]   4.82449  0.4556
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
```

```
               Statistic Shape Scale P-Value
ARCH Lag[3]       0.1794 0.500 2.000  0.6719
ARCH Lag[5]       1.9003 1.440 1.667  0.4937
ARCH Lag[7]       2.4383 2.315 1.543  0.6252

Nyblom stability test
------------------------------------
Joint Statistic:  89.8604
Individual Statistics:
mu       0.3813
ma1      0.1425
omega   19.5833
alpha1   0.1555
beta1    0.1835
gamma1   0.1778
shape    0.3440

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.69 1.9 2.35
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias          1.48465 0.1378
Negative Sign Bias 0.38798 0.6981
Positive Sign Bias 0.07826 0.9376
Joint Effect       3.67791 0.2984


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     29.06      0.06507
2    30     40.19      0.08085
3    40     52.04      0.07904
4    50     62.73      0.08998


Elapsed time : 2.928403
```

GJR-GARCH (1,1) NORMAL

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : gjrGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.000193    0.000300 -0.64282 0.520338
ma1    -0.040225    0.020895 -1.92512 0.054214
omega   0.000002    0.000002  0.83698 0.402602
alpha1  0.017093    0.008179  2.08993 0.036624
beta1   0.949805    0.011908 79.76054 0.000000
gamma1  0.058224    0.012641  4.60614 0.000004

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.000193    0.000345 -0.55817  0.57673
ma1    -0.040225    0.032717 -1.22950  0.21888
omega   0.000002    0.000014  0.11946  0.90491
alpha1  0.017093    0.045567  0.37513  0.70756
beta1   0.949805    0.076333 12.44292  0.00000
gamma1  0.058224    0.052482  1.10940  0.26726

LogLikelihood : 6561.459

Information Criteria
------------------------------------

Akaike       -5.1945
Bayes        -5.1806
Shibata      -5.1945
Hannan-Quinn -5.1895

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1823  0.6694
Lag[2*(p+q)+(p+q)-1][2]    0.4275  0.9791
Lag[4*(p+q)+(p+q)-1][5]    0.9001  0.9607
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                    0.05648  0.8122
Lag[2*(p+q)+(p+q)-1][5]   2.83122  0.4387
Lag[4*(p+q)+(p+q)-1][9]   4.20745  0.5541
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]   0.07665 0.500 2.000  0.7819
```

```
ARCH Lag[5]    1.60040 1.440 1.667  0.5663
ARCH Lag[7]    2.11941 2.315 1.543  0.6920

Nyblom stability test
------------------------------------
Joint Statistic:  68.3238
Individual Statistics:
mu       0.3161
ma1      0.1554
omega   10.6468
alpha1   0.1016
beta1    0.1180
gamma1   0.1091

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:          1.49 1.68 2.12
Individual Statistic:     0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias           1.6448 0.1001
Negative Sign Bias  0.2326 0.8161
Positive Sign Bias  0.1272 0.8988
Joint Effect        4.3488 0.2262


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     39.63    0.0036593
2    30     58.12    0.0010534
3    40     72.10    0.0009881
4    50     86.94    0.0006849


Elapsed time : 2.089723
```

GJR-GARCH - GED

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : gjrGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : ged

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.000013    0.000276 -0.04551 0.963701
ma1    -0.055641    0.019924 -2.79260 0.005229
omega   0.000001    0.000002  0.55211 0.580875
alpha1  0.016296    0.009413  1.73122 0.083412
beta1   0.951390    0.013100 72.62265 0.000000
gamma1  0.058122    0.013770  4.22091 0.000024
shape   1.490193    0.068545 21.74024 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.000013    0.000350 -0.035911 0.971353
ma1    -0.055641    0.023264 -2.391756 0.016768
omega   0.000001    0.000016  0.082951 0.933890
alpha1  0.016296    0.050459  0.322962 0.746724
beta1   0.951390    0.080825 11.770964 0.000000
gamma1  0.058122    0.053807  1.080189 0.280058
shape   1.490193    0.254903  5.846113 0.000000

LogLikelihood : 6589.77

Information Criteria
------------------------------------

Akaike       -5.2161
Bayes        -5.2000
Shibata      -5.2162
Hannan-Quinn -5.2103

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.279  0.2581
Lag[2*(p+q)+(p+q)-1][2]     1.566  0.4018
Lag[4*(p+q)+(p+q)-1][5]     2.060  0.6941
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.0799  0.7774
Lag[2*(p+q)+(p+q)-1][5]    3.1452  0.3812
Lag[4*(p+q)+(p+q)-1][9]    4.6352  0.4849
d.o.f=2

Weighted ARCH LM Tests
```

```
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.1306 0.500 2.000  0.7178
ARCH Lag[5]    1.8042 1.440 1.667  0.5162
ARCH Lag[7]    2.3453 2.315 1.543  0.6445

Nyblom stability test
------------------------------------
Joint Statistic:  88.8959
Individual Statistics:
mu      0.3740
ma1     0.1577
omega  17.0227
alpha1  0.1289
beta1   0.1510
gamma1  0.1455
shape   0.2614

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.69 1.9 2.35
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value   prob sig
Sign Bias          1.55168 0.1209
Negative Sign Bias 0.28579 0.7751
Positive Sign Bias 0.02221 0.9823
Joint Effect       3.97444 0.2642


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     30.41      0.04686
2    30     31.32      0.35027
3    40     55.59      0.04125
4    50     57.42      0.19146


Elapsed time : 3.642287
```

FIGARCH(1,d,1) SSTD

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : fiGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : sstd

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu     -0.000030    0.000292  -0.10164 0.919046
ma1    -0.057144    0.020333  -2.81044 0.004947
omega   0.000001    0.000000   4.48499 0.000007
alpha1  0.027352    0.039519   0.69213 0.488857
beta1   0.942968    0.005797 162.67584 0.000000
delta   0.999999    0.025045  39.92830 0.000000
skew    0.933698    0.026387  35.38455 0.000000
shape   7.984035    1.133255   7.04522 0.000000

Robust Standard Errors:
        Estimate  Std. Error   t value Pr(>|t|)
mu     -0.000030    0.000315 -0.093939 0.925158
ma1    -0.057144    0.019377 -2.949051 0.003188
omega   0.000001    0.000000  2.941537 0.003266
alpha1  0.027352    0.052394  0.522039 0.601643
beta1   0.942968    0.012164 77.518637 0.000000
delta   0.999999    0.050382 19.848386 0.000000
skew    0.933698    0.026417 35.344188 0.000000
shape   7.984035    1.182152  6.753815 0.000000

LogLikelihood : 6579.811

Information Criteria
------------------------------------

Akaike       -5.2075
Bayes        -5.1890
Shibata      -5.2075
Hannan-Quinn -5.2007

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      2.045 0.15275
Lag[2*(p+q)+(p+q)-1][2]     2.542 0.08429
Lag[4*(p+q)+(p+q)-1][5]     3.377 0.34274
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1676  0.6822
Lag[2*(p+q)+(p+q)-1][5]    1.8026  0.6657
Lag[4*(p+q)+(p+q)-1][9]    2.9678  0.7643
d.o.f=2
```

```
Weighted ARCH LM Tests
------------------------------------
             Statistic Shape Scale P-Value
ARCH Lag[3]    0.07313 0.500 2.000  0.7868
ARCH Lag[5]    0.87673 1.440 1.667  0.7698
ARCH Lag[7]    1.50222 2.315 1.543  0.8209

Nyblom stability test
------------------------------------
Joint Statistic:  41.1549
Individual Statistics:
mu       0.2533
ma1      0.1586
omega   16.8104
alpha1   0.2341
beta1    0.1112
delta    0.2177
skew     0.1936
shape    0.4027

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.89 2.11 2.59
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias            1.516 0.12966
Negative Sign Bias   0.231 0.81730
Positive Sign Bias   1.086 0.27768
Joint Effect         8.719 0.03326  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     18.19       0.5100
2    30     26.52       0.5974
3    40     38.44       0.4952
4    50     38.40       0.8623


Elapsed time : 4.867369
```

FIGARCH (1,1)-STD

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : fiGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu      0.000176    0.000282   0.62491 0.532033
ma1    -0.052491    0.020214  -2.59680 0.009410
omega   0.000001    0.000000   4.52424 0.000006
alpha1  0.026009    0.040097   0.64865 0.516562
beta1   0.942665    0.005894 159.93000 0.000000
delta   0.999999    0.025494  39.22503 0.000000
shape   7.733444    1.059491   7.29921 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000176    0.000301  0.58584 0.557984
ma1    -0.052491    0.019153 -2.74061 0.006132
omega   0.000001    0.000000  3.03499 0.002405
alpha1  0.026009    0.053046  0.49030 0.623918
beta1   0.942665    0.012488 75.48628 0.000000
delta   0.999999    0.051880 19.27505 0.000000
shape   7.733444    1.100973  7.02419 0.000000

LogLikelihood : 6576.852

Information Criteria
------------------------------------

Akaike       -5.2059
Bayes        -5.1897
Shibata      -5.2059
Hannan-Quinn -5.2000

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.506  0.2198
Lag[2*(p+q)+(p+q)-1][2]     1.976  0.2216
Lag[4*(p+q)+(p+q)-1][5]     2.787  0.4868
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1414  0.7069
Lag[2*(p+q)+(p+q)-1][5]    1.8909  0.6444
Lag[4*(p+q)+(p+q)-1][9]    3.0947  0.7432
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
```

```
               Statistic Shape Scale P-Value
ARCH Lag[3]    0.07318 0.500 2.000  0.7868
ARCH Lag[5]    0.92488 1.440 1.667  0.7554
ARCH Lag[7]    1.55236 2.315 1.543  0.8108

Nyblom stability test
------------------------------------
Joint Statistic:  39.434
Individual Statistics:
mu      0.2604
ma1     0.1582
omega  15.5624
alpha1  0.2352
beta1   0.1254
delta   0.2322
shape   0.4035

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.69 1.9 2.35
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           1.9603 0.05007   *
Negative Sign Bias  0.4133 0.67939
Positive Sign Bias  0.7905 0.42930
Joint Effect       10.0379 0.01825  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     19.26      0.43998
2    30     34.15      0.23368
3    40     36.00      0.60747
4    50     64.91      0.06355


Elapsed time : 3.803041
```

FIGARCH(1,1) - NORMAL

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : fiGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu      0.000039    0.000300   0.13040 0.896252
ma1    -0.039246    0.020883  -1.87933 0.060199
omega   0.000001    0.000000   3.41384 0.000641
alpha1  0.020951    0.033080   0.63336 0.526500
beta1   0.945251    0.005603 168.71258 0.000000
delta   1.000000    0.018627  53.68568 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000039    0.000318  0.12279 0.902271
ma1    -0.039246    0.019467 -2.01600 0.043800
omega   0.000001    0.000000  2.58440 0.009755
alpha1  0.020951    0.059511  0.35206 0.724797
beta1   0.945251    0.015162 62.34292 0.000000
delta   1.000000    0.058959 16.96082 0.000000

LogLikelihood : 6541.951

Information Criteria
------------------------------------

Akaike       -5.1790
Bayes        -5.1652
Shibata      -5.1791
Hannan-Quinn -5.1740

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.3333  0.5637
Lag[2*(p+q)+(p+q)-1][2]    0.7547  0.8734
Lag[4*(p+q)+(p+q)-1][5]    1.5340  0.8383
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                    0.03236  0.8572
Lag[2*(p+q)+(p+q)-1][5]   2.03796  0.6095
Lag[4*(p+q)+(p+q)-1][9]   3.12751  0.7377
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.1509 0.500 2.000  0.6977
```

```
ARCH Lag[5]    0.8038 1.440 1.667  0.7919
ARCH Lag[7]    1.3231 2.315 1.543  0.8560

Nyblom stability test
------------------------------------
Joint Statistic:  38.4692
Individual Statistics:
mu      0.17203
ma1     0.21220
omega  14.77337
alpha1  0.23051
beta1   0.07051
delta   0.22156

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.49 1.68 2.12
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           1.7809 0.07505   *
Negative Sign Bias  0.1414 0.88756
Positive Sign Bias  0.7291 0.46604
Joint Effect        9.1367 0.02753  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     53.02    4.648e-05
2    30     64.61    1.597e-04
3    40     80.69    9.875e-05
4    50     89.47    3.684e-04


Elapsed time : 2.637866
```

FIGARCH(1,1)-GED

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : fiGARCH(1,1)
Mean Model      : ARFIMA(0,0,1)
Distribution    : ged

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu      0.000158    0.000276   0.57436 0.565723
ma1    -0.055385    0.019790  -2.79864 0.005132
omega   0.000001    0.000000   3.75770 0.000171
alpha1  0.026050    0.039429   0.66069 0.508812
beta1   0.945013    0.005819 162.40135 0.000000
delta   0.999999    0.024450  40.89927 0.000000
shape   1.450352    0.057044  25.42512 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000158    0.000297  0.53383 0.593462
ma1    -0.055385    0.018613 -2.97569 0.002923
omega   0.000001    0.000000  2.90020 0.003729
alpha1  0.026050    0.053584  0.48616 0.626853
beta1   0.945013    0.013024 72.55689 0.000000
delta   0.999999    0.052748 18.95796 0.000000
shape   1.450352    0.063070 22.99588 0.000000

LogLikelihood : 6575.222

Information Criteria
------------------------------------

Akaike       -5.2046
Bayes        -5.1884
Shibata      -5.2046
Hannan-Quinn -5.1987

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.788  0.1812
Lag[2*(p+q)+(p+q)-1][2]     2.279  0.1343
Lag[4*(p+q)+(p+q)-1][5]     3.117  0.4023
d.o.f=1
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                    0.09618  0.7565
Lag[2*(p+q)+(p+q)-1][5]   1.99521  0.6196
Lag[4*(p+q)+(p+q)-1][9]   3.11991  0.7390
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
```

```
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.1042 0.500 2.000  0.7468
ARCH Lag[5]    0.8565 1.440 1.667  0.7759
ARCH Lag[7]    1.4064 2.315 1.543  0.8399

Nyblom stability test
------------------------------------
Joint Statistic:  39.931
Individual Statistics:
mu      0.2640
ma1     0.1807
omega  17.1203
alpha1  0.2317
beta1   0.1052
delta   0.2416
shape   0.2662

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.69 1.9 2.35
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           1.8497 0.06447   *
Negative Sign Bias  0.3140 0.75352
Positive Sign Bias  0.8089 0.41866
Joint Effect        9.5357 0.02295  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     20.75       0.3505
2    30     31.09       0.3613
3    40     40.09       0.4217
4    50     51.71       0.3683


Elapsed time : 3.856946
```

SUPPORT VECTOR REGRESSION

TRAINING and TESTING DATA

Load the package caret in R. Divide the data into two sets (Training (70%) and Testing (30%)). Set a sample space to accommodate the 70% of the data that needs to be trained. Train and test the data accordingly. Fit eps regression using the desired kernels and cross validate it.

```
set.seed(2000)

2524*0.7
2524*0.3
```

Train and test the data accordingly by using the below codes.

```
train =sample(1:2524,1767,replace=FALSE)
traindata4=svmm[train,]
testdata4=svmm[-train,]
```

To fit eps regression use the below codes for training and testing.

```
fitepslinear1=svm(Returns~.,data=traindata4,type="eps-regression",kernel="linear",cross=1767)

fitepspoly1=svm(Returns~.,data=traindata4,type="epsregression",kernel="polynomial",cross=1767)

fitepssigmoid1=svm(Returns~.,data=traindata4,type="epsregression",kernel="sigmoid",cross=1767)

fitepsrad1=svm(Returns~.,data=traindata4,type="eps-regression",kernel="radial",cross=1767)
```

SVR-GARCH(1,1) STD

```
Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(1,0,1)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000555    0.000534   1.0398 0.298433
ar1    -0.961066    0.068891 -13.9505 0.000000
ma1     0.948146    0.079031  11.9972 0.000000
omega   0.000005    0.000004   1.2440 0.213486
alpha1  0.114918    0.027448   4.1868 0.000028
beta1   0.879065    0.025004  35.1566 0.000000
shape   6.615753    1.567111   4.2216 0.000024

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000555    0.000528   1.05211 0.292750
ar1    -0.961066    0.067857 -14.16303 0.000000
ma1     0.948146    0.078078  12.14353 0.000000
omega   0.000005    0.000007   0.75335 0.451243
alpha1  0.114918    0.029030   3.95853 0.000075
beta1   0.879065    0.024720  35.56043 0.000000
shape   6.615753    1.364175   4.84964 0.000001

LogLikelihood : 1964.06

Information Criteria
------------------------------------

Akaike       -5.1706
Bayes        -5.1278
Shibata      -5.1707
Hannan-Quinn -5.1541

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.044  0.3069
Lag[2*(p+q)+(p+q)-1][5]     2.478  0.7901
Lag[4*(p+q)+(p+q)-1][9]     4.399  0.5957
d.o.f=2
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.8973  0.3435
Lag[2*(p+q)+(p+q)--1][5]   2.0108  0.6159
Lag[4*(p+q)+(p+q)--1][9]   3.5297  0.6693
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]     1.369 0.500 2.000  0.2420
```

```
ARCH Lag[5]      1.845 1.440 1.667  0.5065
ARCH Lag[7]      2.862 2.315 1.543  0.5399

Nyblom stability test
------------------------------------
Joint Statistic:  1.1331
Individual Statistics:
mu     0.03528
ar1    0.10058
ma1    0.10236
omega  0.10906
alpha1 0.09983
beta1  0.12016
shape  0.32350

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:        1.69 1.9 2.35
Individual Statistic:   0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           2.5457 0.01110  **
Negative Sign Bias  1.5086 0.13181
Positive Sign Bias  0.1047 0.91668
Joint Effect        9.0671 0.02841  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     16.29       0.63790
2    30     44.52       0.03275
3    40     50.24       0.10723
4    50     60.24       0.13034


Elapsed time : 1.052252
```

SVR-EGARCH(1,1)-GED

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : eGARCH(1,1)
Mean Model      : ARFIMA(1,0,1)
Distribution    : ged

Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu      0.000463    0.000467   0.99093 0.321718
ar1    -0.957569    0.133235  -7.18704 0.000000
ma1     0.940917    0.157250   5.98359 0.000000
omega  -0.163850    0.054923  -2.98326 0.002852
alpha1 -0.054836    0.026955  -2.03434 0.041917
beta1   0.979736    0.006788 144.32385 0.000000
gamma1  0.221794    0.044111   5.02814 0.000000
shape   1.448784    0.103653  13.97721 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000463    0.000549   0.8419 0.399843
ar1    -0.957569    0.305688  -3.1325 0.001733
ma1     0.940917    0.358546   2.6243 0.008684
omega  -0.163850    0.093727  -1.7482 0.080437
alpha1 -0.054836    0.025546  -2.1466 0.031828
beta1   0.979736    0.011810  82.9573 0.000000
gamma1  0.221794    0.060947   3.6391 0.000274
shape   1.448784    0.096671  14.9868 0.000000

LogLikelihood : 1966.029

Information Criteria
------------------------------------

Akaike       -5.1731
Bayes        -5.1242
Shibata      -5.1733
Hannan-Quinn -5.1543

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.007  0.3157
Lag[2*(p+q)+(p+q)-1][5]     2.616  0.7136
Lag[4*(p+q)+(p+q)-1][9]     4.449  0.5837
d.o.f=2
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.3068  0.5796
Lag[2*(p+q)+(p+q)-1][5]    1.7517  0.6781
Lag[4*(p+q)+(p+q)-1][9]    3.1727  0.7301
```

```
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]     1.874 0.500 2.000  0.1710
ARCH Lag[5]     2.160 1.440 1.667  0.4370
ARCH Lag[7]     3.169 2.315 1.543  0.4822

Nyblom stability test
------------------------------------
Joint Statistic:  1.2711
Individual Statistics:
mu     0.02793
ar1    0.07156
ma1    0.09224
omega  0.07793
alpha1 0.09989
beta1  0.07381
gamma1 0.20593
shape  0.17012

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.89 2.11 2.59
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           2.5590 0.01069  **
Negative Sign Bias  1.7685 0.07738   *
Positive Sign Bias  0.6815 0.49576
Joint Effect        7.1155 0.06831   *


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     26.96       0.1055
2    30     31.36       0.3486
3    40     43.26       0.2942
4    50     49.94       0.4360
```

SVR-GJR-GARCH-GED

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : gjrGARCH(1,1)
Mean Model      : ARFIMA(1,0,1)
Distribution    : ged

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000347    0.000567  0.61134  0.54097
ar1    -0.957242    0.031719 -30.17873  0.00000
ma1     0.941031    0.037918  24.81758  0.00000
omega   0.000005    0.000004   1.15580  0.24776
alpha1  0.075281    0.028216   2.66803  0.00763
beta1   0.888659    0.024110  36.85846  0.00000
gamma1  0.057185    0.039308   1.45480  0.14572
shape   1.424255    0.100371  14.18985  0.00000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000347    0.000589   0.58873  0.55604
ar1    -0.957242    0.019707 -48.57382  0.00000
ma1     0.941031    0.020778  45.28916  0.00000
omega   0.000005    0.000008   0.65801  0.51053
alpha1  0.075281    0.030700   2.45215  0.01420
beta1   0.888659    0.025882  34.33478  0.00000
gamma1  0.057185    0.042849   1.33458  0.18201
shape   1.424255    0.091984  15.48379  0.00000

LogLikelihood : 1964.964

Information Criteria
------------------------------------

Akaike       -5.1703
Bayes        -5.1214
Shibata      -5.1705
Hannan-Quinn -5.1515

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.076  0.2995
Lag[2*(p+q)+(p+q)-1][5]     2.705  0.6599
Lag[4*(p+q)+(p+q)-1][9]     4.676  0.5290
d.o.f=2
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.9152  0.3387
Lag[2*(p+q)+(p+q)-1][5]    1.8991  0.6425
Lag[4*(p+q)+(p+q)-1][9]    3.3904  0.6931
d.o.f=2
```

```
Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]     1.150 0.500 2.000  0.2835
ARCH Lag[5]     1.510 1.440 1.667  0.5898
ARCH Lag[7]     2.615 2.315 1.543  0.5890

Nyblom stability test
------------------------------------
Joint Statistic:  1.2171
Individual Statistics:
mu     0.03336
ar1    0.06813
ma1    0.07797
omega  0.11794
alpha1 0.05467
beta1  0.07593
gamma1 0.07596
shape  0.20051

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.89 2.11 2.59
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value    prob sig
Sign Bias           2.5308 0.01158  **
Negative Sign Bias  1.6279 0.10397
Positive Sign Bias  0.2648 0.79127
Joint Effect        7.7091 0.05242   *


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     19.83       0.4049
2    30     32.23       0.3097
3    40     40.83       0.3898
4    50     49.41       0.4569
```

SVR-FIGARCH(1,d,1) -STD

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : fiGARCH(1,1)
Mean Model      : ARFIMA(1,0,1)
Distribution    : std

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000569    0.000528   1.0770 0.281483
ar1    -0.962068    0.096130 -10.0080 0.000000
ma1     0.947732    0.113464   8.3527 0.000000
omega   0.000007    0.000006   1.2824 0.199695
alpha1  0.052120    0.114652   0.4546 0.649400
beta1   0.751855    0.127683   5.8885 0.000000
delta   0.768730    0.217173   3.5397 0.000401
shape   6.456927    1.435445   4.4982 0.000007

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.000569    0.000528  1.07673 0.281600
ar1    -0.962068    0.154153 -6.24101 0.000000
ma1     0.947732    0.180648  5.24629 0.000000
omega   0.000007    0.000007  1.01624 0.309517
alpha1  0.052120    0.118635  0.43933 0.660420
beta1   0.751855    0.112452  6.68602 0.000000
delta   0.768730    0.217092  3.54104 0.000399
shape   6.456927    1.315720  4.90752 0.000001

LogLikelihood : 1965.047

Information Criteria
------------------------------------

Akaike       -5.1705
Bayes        -5.1216
Shibata      -5.1708
Hannan-Quinn -5.1517

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                      1.298  0.2546
Lag[2*(p+q)+(p+q)-1][5]     2.844  0.5718
Lag[4*(p+q)+(p+q)-1][9]     4.799  0.4999
d.o.f=2
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                    0.01832  0.8923
Lag[2*(p+q)+(p+q)-1][5]   0.94145  0.8728
Lag[4*(p+q)+(p+q)-1][9]   2.53986  0.8318
d.o.f=2
```

```
Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.9929 0.500 2.000  0.3190
ARCH Lag[5]    1.4973 1.440 1.667  0.5930
ARCH Lag[7]    2.7830 2.315 1.543  0.5554

Nyblom stability test
------------------------------------
Joint Statistic:  1.2854
Individual Statistics:
mu     0.04004
ar1    0.09355
ma1    0.10087
omega  0.07294
alpha1 0.22313
beta1  0.13694
delta  0.19466
shape  0.32726

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.89 2.11 2.59
Individual Statistic:    0.35 0.47 0.75

Sign Bias Test
------------------------------------
                   t-value     prob sig
Sign Bias           2.5981 0.009556 ***
Negative Sign Bias  1.1299 0.258869
Positive Sign Bias  0.3668 0.713863
Joint Effect        8.2979 0.040239  **


Adjusted Pearson Goodness-of-Fit Test:
------------------------------------
  group statistic p-value(g-1)
1    20     19.09      0.45109
2    30     38.42      0.11336
3    40     41.78      0.35077
4    50     65.92      0.05368
```

EXTREME VALUE THEORY

EGARCH(1,1)

```
$threshold
[1] 1

$p.less.thresh
[1] 0.8847068

$n.exceed
[1] 291

$run
[1] NA

$par.ests
         xi       sigma          mu        beta
 0.07600268  0.55072056 -0.29270533  0.64896964

$par.ses
       xi     sigma        mu
0.0728969 0.1294812 0.2075962

$varcov
             [,1]         [,2]        [,3]
[1,]  0.005313958 -0.009092274  0.01372113
[2,] -0.009092274  0.016765374 -0.02622010
[3,]  0.013721127 -0.026220099  0.04309617

$intensity
[1] 0.1153389

$nllh.final
[1] 1106.79

$converged
[1] 0

attr(,"class")
[1] "potd"
```

RISK MEASURES

```
          p quantile    sfall
[1,] 0.9900 2.743267 3.588620
[2,] 0.9950 3.299252 4.190257
[3,] 0.9990 4.708773 5.715516
[4,] 0.9995 5.370903 6.432013
[5,] 0.9999 7.049521 8.248464
```

SVR-EGARCH (1,1)

```
$span
[1] 756

$threshold
[1] 0.5

$p.less.thresh
[1] 0.7582563

$n.exceed
[1] 183

$run
[1] NA

$par.ests
         xi       sigma          mu        beta
-0.00376834  0.74158161 -0.54918157  0.73762794

$par.ses
        xi      sigma         mu
0.08103158 0.15543626 0.17373784

$varcov
             [,1]        [,2]        [,3]
[1,]  0.006566117 -0.01179239  0.01180139
[2,] -0.011792392  0.02416043 -0.02542478
[3,]  0.011801388 -0.02542478  0.03018484

$intensity
[1] 0.2420635

$nllh.final
[1] 569.2324

$converged
[1] 0

attr(,"class")
[1] "potd"
```

RISK MEASURES

```
  p quantile    sfall
[1,] 0.9900 2.835899 3.562528
[2,] 0.9950 3.340725 4.065571
[3,] 0.9990 4.508122 5.228843
[4,] 0.9995 5.008843 5.727796
[5,] 0.9999 6.166748 6.881610
```