Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/120315

Investigating risk within South African Financial markets using Extreme Value Mixture Models

Kimera Naradh, Knowledge Chinhamu and Retius Chifurira*

School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban, South Africa

Received: 4 May 2022, Revised: 7 Feb. 2023, Accepted: 9 Feb. 2023
Published online: 1 Sep. 2023

Abstract: In the past decade, analysts, statisticians and researchers have become more interested in the research and applications of extreme value mixture models in the stock market and insurance as well as medical industries. This study aims to evaluate the fit of two extreme value mixture models namely GPD-Normal-GPD (GNG) and GPD-KDE-GPD (GKG), where KDE represents the Kernel density estimator, for three FTSE/JSE indices namely All Share Index (ALSI), Banks Index and Mining Index and the USD/ZAR currency exchange rate. Value at Risk (VaR) assesses market risk and many financial corporations often seek reliable VaR estimates. VaR estimates and the Kupiec likelihood backtesting procedure are calculated to evaluate the tail behaviour of the fitted GNG models. Results highlight the robustness of the GNG and GKG mixed model for each daily returns when compared to the traditional Normal model that is commonly applied model in financial literature. Financial practitioners looking to curb losses and explore alternatives for financial modeling in the South African financial industry using an extreme value mixed model approach may gain the most by implementing the GNG or GKG model

Keywords: GNG, Kernel Density Estimator, Value at Risk, Kupiec likelihood ratio test, Extreme Value Theory, Mixture models

1 Introduction

Pre COVID-19, South Africa was already experiencing economic despair with a myriad of socio-economic dilemmas. The onslaught of the pandemic in March 2020 and subsequent three-week imposed lockdown drastically reduced economic activity causing mass job losses. More recently, with the conflict between Russia and Ukraine escalating, impacting global markets, South African markets too will not be spared. Despite the bleak economic outlook, the South African government remains positive on the implementation and progress of economic recovery and reform strategies [1]. The need for dependable models that track the movement of volatile indices and exchange rates during globally disruptive events is of great importance in managing risk and executing the relevant structural changes and governance required for financial stability. Extreme value theory is often considered when dealing with unpredicted and rare events and is used to develop sound models that are useful for extracting valuable insights. In literature, it is very well known that financial data exhibits heavy tails and skewness and one way of dealing with this issue is to propose a volatility model that adequately describes the characteristics of financial data. Notable literature by [2] developed a two-stage model, where the first stage fits a Generalised Autoregressive Conditional Heteroskedasticity (GARCH) model to capture volatility clustering and the second stage fits a Generalised Pareto Distribution (GPD) to the tails. This method is extended by [3] where the proposed extreme value mixed model fits the GPD at the tails and the Normal distribution is fitted as the main model also known as the bulk model between the two tails. This paper aims to investigate the fit of the GPD-Normal-GPD (GNG) and GPD-KDE-GPD, where KDE represents the Kernel density estimator, models to three FTSE/JSE stock market indices and the United States of American Dollar to the South African Rand. Distribution density plots of the fitted GNG and GKG models validate the adequacy of the models on each of the return series investigated in this study. The Kupiec likelihood ratio test is applied as a backtesting procedure to the VaR estimates with the aim of evaluating the model adequacy of each fitted GNG and GKG models. In essence, this study aims to contribute and shed light on the usefulness of extreme value mixture models particularly the GNG and GKG distribution as an alternative to fitting large sets of economic data that

* Corresponding author e-mail: kimmi.naradh@gmail.com

possess heavy tails and skewness as well as attain a holistic knowledge base of the health of the South African Financial segment.

2 Literature Review

Research by [4] utilized the truncated gamma as the parametric model to fit the bulk distribution of an extreme value mixed model with the GPD model fitted at both tails. The threshold is estimated as a parameter that splits both distributions at this estimated threshold point. This paper aimed to analyse extreme events using a Bayesian approach. [5] mentions an advantage of applying the proposed mixed model by [4] is the relative ease and flexibility to work with. The most attractive property of the suggested mixed model is the ability to consider threshold estimation. A possible setback of this model is the dependence and sensitivity of the tail model on the bulk distribution. For example, if the bulk distribution is mis specified, then the threshold is affected and subsequently the tail distribution.

[6] introduced a dynamically weighted mixture model by combining GPD (for the tail model) with the Weibull distribution (for the bulk model). It is assumed that the bulk model is light-tailed so the GPD dominates the tail model particularly the upper tail. The maximum likelihood method is used to estimate parameters and this model was applied to simulated data and the Danish fire loss data. The approach in this work can be valuable in unsupervised tail estimation in heavy tailed situation.

Work by [7] proposed a mixture model where the bulk distribution is the Normal distribution and the GPD model is fitted on the lower and upper tails. For estimation there is a specific procedure outlined by [5] which implies the data is standardised and ordered where the GPD parameters in the mixed model are estimated using the L-moments method.

The study by [3] applies a model like [7] however with a different method of parameter estimation. Bayesian inference is used where the two tails are GPD and the bulk is the Normal distribution. The breakthrough with this approach is the absence of influence of the bulk distribution on tail model fitting. A possible drawback pointed out by [5] deals with the misfit of the model, more especially the bulk model.

[8] evaluated the fit to four stock indices, namely S&P, Euroxx, FSTE100 and Nikkei using extreme value mixture models including the GNG mixture model. VaR estimates were calculated and the Kupiec proportion of failure test was used to gauge model adequacy. The GNG model provided good results for the data.

Work by [9] investigated a hybrid (GARCH)-type model with GPD and Nolan's $\mathbf{S_0}-$parameterization Stable Distribution (SD) to the returns of three FTSE/JSE indices and the South African Rand to USD exchange rate (All Share Index (ALSI), Banks Index and Mining Index, as well as the daily closing prices of the US dollar against the South African rand exchange rate. VaR values were assessed and back-tested using the Kupiec likelihood ratio test. Results from this study imply that the GARCH (1,1)-SD model is a better choice for VaR robust modelling of financial returns. This study provides noticeable results for individuals with a vested interested in minimizing losses or gaining insights of the South African stock market.

[10] proposed a flexible extreme value mixture model that combines a non-parametric kernel density estimator for the bulk model and the GPD distribution as the fitted tail model. The performance of the proposed mixture model was evaluated by empirical analysis and simulation study for determining normal physiological measurements for pre-mature infants.

There is limited research on the topic of modelling FTSE/JSE indices and the USD/ZAR exchange rate to the GNG mixture model proposed by [3] and GKG mixture model applied by [10] to the best of our literature knowledge. The key value-add of this study is to investigate the appropriateness of the GNG and GKG mixture model in estimating VaR values for the JSE financial data and USD/ZAR currency exchange rate that exhibit heavy tails and asymmetry.

3 Research and Methodology

The work of [5] notes that extreme value modelling is used to implement models that remedy problems faced in the real world particularly those that are in connection with unusual or disruptive events.

Extreme value mixture models generally have two components. The first component is a model that describes non-extremal data below are certain threshold known as the bulk model and the second component, known as the tail model, refers to the data above the threshold. Typically, the bulk model is a parametric or non-parametric distribution and the tail model is a commonly an adaptable threshold model like the GPD.

3.1 Generalized Pareto Distribution (GPD)

The paper by [10] notes the GPD by the distribution as:

$$G(x|u,\sigma_u,\xi) = \begin{cases} 1-[1+\xi(\frac{x-u}{\sigma_u})]_+^{\frac{1}{\xi}} & \xi \neq 0 \\ 1-exp\left[-\left(\frac{x-u}{\sigma_u}\right)\right] & \xi = 0. \end{cases} \tag{1}$$

where, $x > u,\ \sigma_u > 0,\ \left[1 + \xi\left(\frac{x-u}{\sigma_u}\right)\right] > 0$

3.2 Threshold choice

Threshold values are selected as a benchmark that separates extreme and non-extreme data points.

The study by [5] highlights threshold selection as an area of continuous research of great importance and [11] emphasizes selection of the threshold is a compromise between bias and variance. The asymptotic properties of the GPD model are violated if the selected threshold is too low. The work of [12] points out that a too high threshold selection may fit well however data points beyond the threshold value will return parameter estimates with large standard errors. Normally, the threshold is chosen prior to the model estimation, however there all exists another method whereby parameter estimates are evaluated by fitting models over a range of threshold values. Both approaches are shown in the work by [11] and there is also mention that there are several diagnostic plots namely the mean residual life plot, the parameter stability plot and model diagnostics plots. [8] states that regardless of the approach used in threshold selection, there is often a reliance on intuition.

3.3 GPD-Normal-GPD

This study investigates the fit of the GPD-Normal-GPD (GNG) on the South African currency and stocks. The GNG mixture model is a two-tailed model where the Normal distribution is the bulk model and the GPD is fitted to the tails.

[3] shows the GNG distribution function as:

$$G(f|\theta) = \begin{cases} \phi(u_l|u,\beta)[1-G(-y|\eta_l,\sigma_l,-u_l)], & y \leq u_l \\ \phi(u_l|u,\beta) & u_l < y < u_r \\ \phi(u_r|u,\beta)+1-\phi(u_r|u,\beta)\,G(-y|\eta_r,\sigma_r,-u_r) & y \geq u_r \end{cases} \tag{2}$$

where,$\theta = \left(u_l,\sigma_{u_l},\eta_l,\mu,\beta,u_r,\sigma_{u_r},\eta_r\right), G(\cdot|\eta,\sigma,u)$ are the GPD distribution function for the upper (shown by subscript r) and lower tail (shown by subscript l) with η the shape parameter and σ the scale parameter and threshold u. $\phi(\cdot|\mu,\beta)$ is the Normal distribution with mean μ and β as standard deviation.

This study considers the parameterized tail fraction approach specified by [5].

3.3.1 Parameterized Tail Fraction Approach

Define $\phi_{u_l} = P(Y < u_l)$ and $\phi_{u_r} = P(Y > u_r)$ where the distribution function is defined as:

$$G(f|\theta) = \begin{cases} \phi_{u_l}[1-G(-y|\eta_l,\sigma_l,-u_l)], & y \leq u_l, \\ \phi_{u_l} + \left(1-\phi_{u_l}-\phi_{u_r}\right)\frac{\phi(y|\mu,\beta)-\phi(u_l|\mu,\beta)}{\phi(u_r|\mu,\beta)-\phi(u_l|\mu,\beta)} & u_l < y < u_r, \\ \left(1-\phi_{u_r}\right)+\phi_{u_r}G(y|\eta_r,\sigma_r,-u_r) & y \geq u_r \end{cases} \tag{3}$$

where, $\theta = \left(u_l,\sigma_{u_l},\varepsilon_l,\mu,\beta,u_r,\sigma_{u_r},\eta_r\right)$.

3.3.2 GNG parameter estimation

This work explores the *evmix* package available in R by [13] along with the *evir* and *ismev* packages. In this study the *evmix* package estimates GNG and GKG mixed model parameters where the threshold is considered a parameter and is also estimated.

The GNG parameter vector is described by $\theta = \left(\mu, \beta, u_l, \sigma_{u_l}, \eta_l, u_r, \sigma_{u_r}, \eta_r\right)$ where μ is the Normal mean, β is the Normal standard deviation, u_l and u_r are the lower and upper tail threshold respectively, σ_{u_l} and σ_{u_r} define the lower tail and upper tail GPD scale parameter respectively, η_l and η_r specify the lower and upper tail GDP shape parameter.

3.4 GPD-KDE-GPD

This paper also evaluates the fit of the GPD-KDE-GPD extreme value mixture model to South African financial data where KDE is the kernel density estimate for the bulk model between a lower and upper threshold and the GPD model beyond these thresholds.

The *evmix* package in R by [13] describe the cumulative distribution function (cdf) with three components. The lower tail model with trail fraction ϕ_{ul} is described by the KDE model up and till the lower threshold $y < u_l$:

$$F(y) = H(u_l)\left[1 - G_l(y)\right] \tag{4}$$

where, $H(y)$ represents the kernel density estimator of the cdf and $G_l(Y)$ is the conditional GPD with negated threshold and y value.

Between the thresholds $u_l \leq y \leq u_r$, the KDE bulk model is defined as:

$$F(y) = H(y) \tag{5}$$

Above the upper threshold u_r, that is, where $y > u_r$ is the conditional GPD denoted as:

$$F(y) = H(u_r) + \left[1 - H(u_r)\right] G_r(y) \tag{6}$$

where, $G_r(y)$ is the cdf of the GPD tail model.

GKG parameter estimates are discussed in detail in the results of this study.

3.5 Value-at-Risk and Backtesting

The Basel Committee on Banking Supervision implements Value-at-Risk (VaR) as the standard benchmark measure for evaluating market risk. The capital requirements of financial establishments are based on VaR estimates, therefore, tests for assessing the out-of-sample forecast accuracy of VaR models through backtesting procedures have become a practical necessity [14]. VaR intends to estimate the highest possible loss for a portfolio over a specified period, where VaR estimations focus on the tails of a distribution and robustness testing procedures for a model. For a random variable Z, which is usually the log-return of a risky financial instrument with distribution function F over a specified period, VaR at given probability p is defined as the $p-th$ quantile of F, that is,

$$\mathrm{VaR_p} = F^{-1}(1-p) \tag{7}$$

where, F^{-1} is the quantile function.

To examine the usefulness and competence of VaR estimates, various backtesting procedures are employed. The Kupiec likelihood ratio test and Christoffersen conditional coverage test infer more formal conclusions on model robustness [15].

4 Empirical data and analysis

4.1 Data Selection

The data sets in this study are the daily FTSE/JSE All-Share Index, FTSE/JSE Banks Index, FTSE/JSE Mining Index and USD/ZAR prices obtained from McGregor BFA and were recorded over the period from 13 August 2010 to 14 August 2020. This data includes the lockdown period during the COVID-19 global pandemic. The return series for each index is calculated as the first backward differences of the index values' natural logarithm. For day t, the daily log return r_t is defined as:

$$r_t = \ln(P_t) - \ln(P_{t-1}), \tag{8}$$

where, P_t is the price at day t.

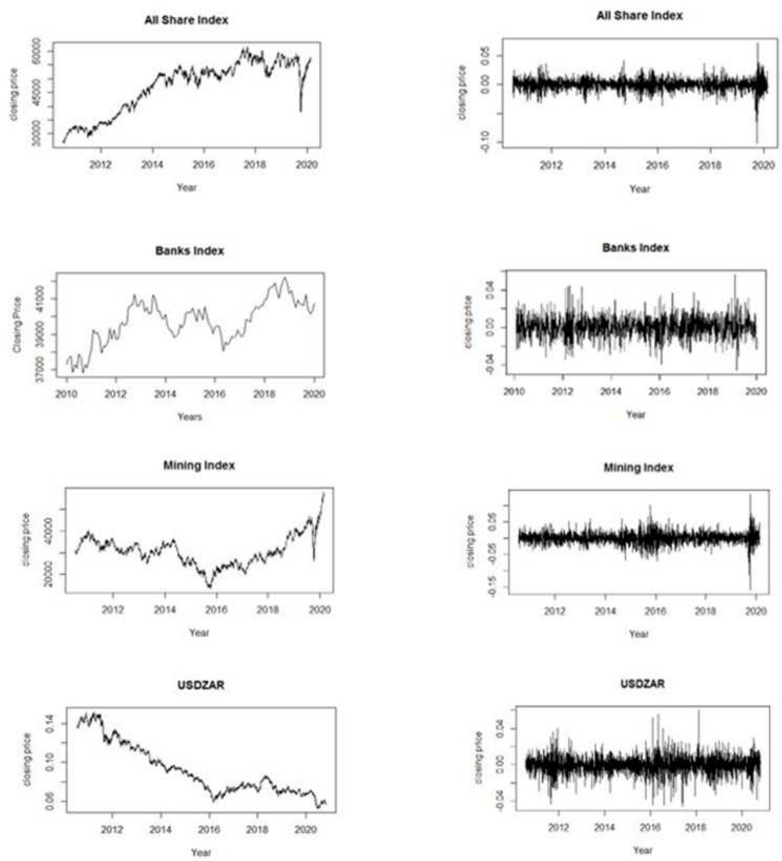

Fig. 1: Time series plot of JSE Indices and USDZAR exchange rate (left) and one day returns (right)

In Figure 1, the plots specify numerous trends in mean and variance over time indicating non-stationarity. The variance varies over time indicating heteroscedasticity and volatility clustering which is plausible when dealing with financial data. Secluded extreme returns caused by shocks to financial markets are evident, such as the 2015 stock market crash and the 2019-2020 global COVID-19 pandemic.

Table 1: Descriptive statistics of financial market indices and exchange rate price returns

	ALSI		Banks Index		Mining Index		USD/ZAR	
Panel A:Descriptive Statistics								
Number of observations	2499.00		2499.00		2499.00		2675.00	
Minimum	-0.1023		-2.3021		-0.1589		0-0.0460	
Maximum	0.0726		0.0991		0.1346		0.0603	
Mean	0.0003		-0.0008		0.0002		-0.0003	
Skewness	-0.7310		-41.0875		-0.1605		-0.1671	
Excess Kurtosis	8.8822		1919.8771		6.0443		2.7766	
Panel B:Testing for unit root and stationarity								
	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
Jarque-Bera	8455.2667	$< 0.0001^{***}$	385117132	$< 0.0001^{***}$	3823.93	$< 0.0001^{***}$	874.3992	$< 0.0001^{***}$
Ljung Box	67.2900	$< 0.0001^{***}$	7.0024	0.9967	47.62	0.0005	11.7751	0.9236
ARCH LM Test	936.0966	$< 0.0001^{***}$	0.00 0.99	50.03	0.00	15.41	0.00	
Panel B:Testing for unit root and stationarity								
	Statistic	p-value	Statistic	p-value	Statistic	p-value	Statistic	p-value
ADF Test	-13.6259	0.01	-12.9576	0.01	-13.5509	0.01	-14.7402	0.01
PP Test	-2586.908	0.01	-2548.281	0.01	-2430.247	0.01	-2688.028	0.01
KPSS Test	0.1303	0.10	0.0914	0.10	0.2394	0.10	0.0431	0.10
*** symbolizes a very small p-value								

Descriptive Statistics for the daily closing prices of the FTSE/JSE financial stock indices returns and USD/ZAR are shown in Table 1 on Panel A. The excess kurtosis value indicates the leptokurtic behaviour of these return series. This means that the empirical distribution of the daily returns is much heavier than the well-known normal distribution. Large values for skewness and excess kurtosis are noticed for the Banks Index returns and possibly describe the inadequate performance of the South African economy as well as the devastating rate of unemployment and the government bail out of the state power utility (Eskom Holdings SOC Ltd) has proven to have negative effects for South African Bank Index stocks [16]. The Jarque-Bera test for normality gives a p-value less than 0.0001 for all four returns, thus rejecting the normality assumption at all levels of significance. Panel B displays tests for normality, autocorrelation and heteroscedasticity are shown. The null hypothesis of normality for the Jarque-Bera test is rejected at 5% level of significance for all stock and currency returns. This presumes considering the use of heavy tailed models when analyzing the returns series.

The Ljung box test provides mixed results. The significant p-values of the Ljung box test for ALSI and the Mining Index suggest rejecting the null hypothesis of no autocorrelation. On the contrary, the null hypothesis for FTSE/JSE Banks Index and USD/ZAR exchange rate is rejected implying that the return series show serial correlation.

Observing the results from Panel C, it can be deduced at a 5% level of significance, the null hypothesis of a unit root is rejected, and it can be decided that all return series are stationary. The KPSS test showed that all returns are stationary since all p-values are 0.1 which is greater than 0.05 therefore the null hypothesis of stationarity is rejected.

4.2 Results and discussion

4.2.1 Parameter Estimation

GNG parameter estimates are shown in Table 2 where $\hat{\mu}$ denotes the estimated Normal mean, $\hat{\beta}$ is the estimated Normal standard deviation of the Normal bulk model.$\widehat{u_l}, \sigma_{u_l}, \widehat{\varepsilon_l}$ represents the lower GPD tail estimated parameters for respective threshold, scale and shape parameters. Likewise, $\widehat{u_r}, \sigma_{u_r}, \widehat{\varepsilon_r}$ describes the upper GPD tail estimated parameters for the respective threshold, scale and shape parameters.

Table 2: ML estimates of the GNG mixture model

GNG Parameter Estimates	Financial Stock return			
	ALSI	**Banks Index**	**Mining Index**	**USD/ZAR**
$\widehat{\mu}$	0,0007	0.0003	0.0003	-0.0001
$\widehat{\beta}$	0.0084	0.0138	0.0149	0.0086
$\widehat{u_l}$	-0.0033	-0.0151	0.0110	0.0086
$\widehat{\sigma_{u_l}}$	0.0070	0.0087	0.0096	0.0101
$\widehat{\eta_l}$	0.0600	0.3304	0.1829	0.0069
$\widehat{u_r}$	0.0106	0.0040	0.0212	0.0170
$\widehat{\sigma_{u_r}}$	0.0054	0.0106	0.0116	0.0115
$\widehat{\eta_r}$	0.1473	0.0336	0.1224	0.1463

GPD-KDE-GPD parameter estimates are shown in Table 3 where $\hat{\lambda}$ denotes the estimated bandwidth of the kernel. $\widehat{u_l}, \sigma_{u_l}, \widehat{\eta_l}$ represents the respective threshold, scale and shape parameters for the lower GPD tail estimates. Likewise, $\widehat{u_r}, \sigma_{u_r}, \widehat{\eta_r}$ describes the upper GPD tail estimated parameters for respective threshold, scale and shape parameters.

Table 3: ML estimates of GKG mixture model

GPD-KDE-GPD Parameter Estimates	Financial Stock return			
	ALSI	**Banks Index**	**Mining Index**	**USD/ZAR**
$\hat{\lambda}$	0.0015	0.0023	0.0022	-0.0007
$\widehat{u_l}$	-0.0116	-0.0175	-0.0200	-0.0118
$\widehat{\sigma_{u_l}}$	0.0064	0.0092	0.0111	0.0068
$\widehat{\eta_l}$	0.1744	0.1790	0.1224	0.0170
$\widehat{u_r}$	0.0115	0.0197	0.0203	0.0109
$\widehat{u_r}$	0.0106	0.0040	0.0212	0.0170
$\widehat{\sigma_{u_r}}$	0.0057	0.0080	0.0098	0.0048
$\widehat{\eta_r}$	0.1426	0.3349	0.1829	0.1463

4.2.2 Fitted models

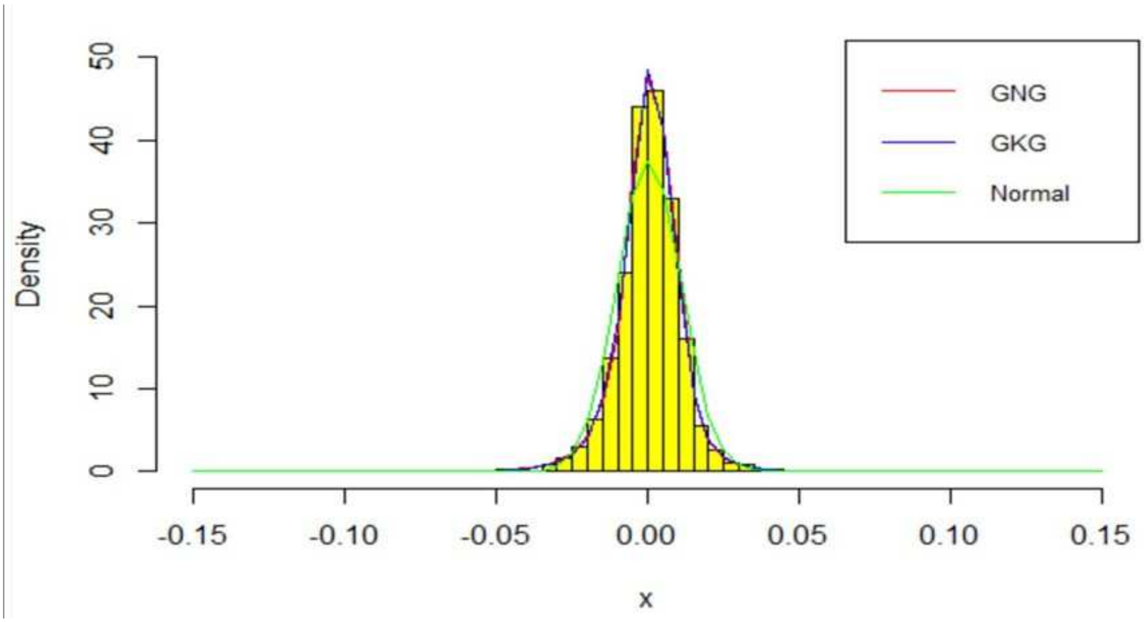

Fig. 2: Fitted GNG and GKG mixture model FTSE/ALSI

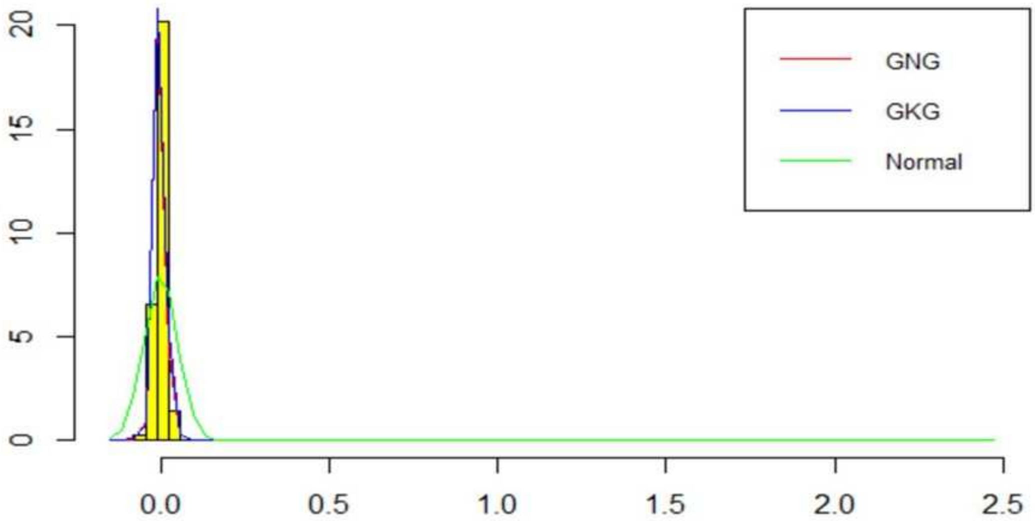

Fig. 3: Fitted GNG and GKG mixture model FTSE/Banks Index

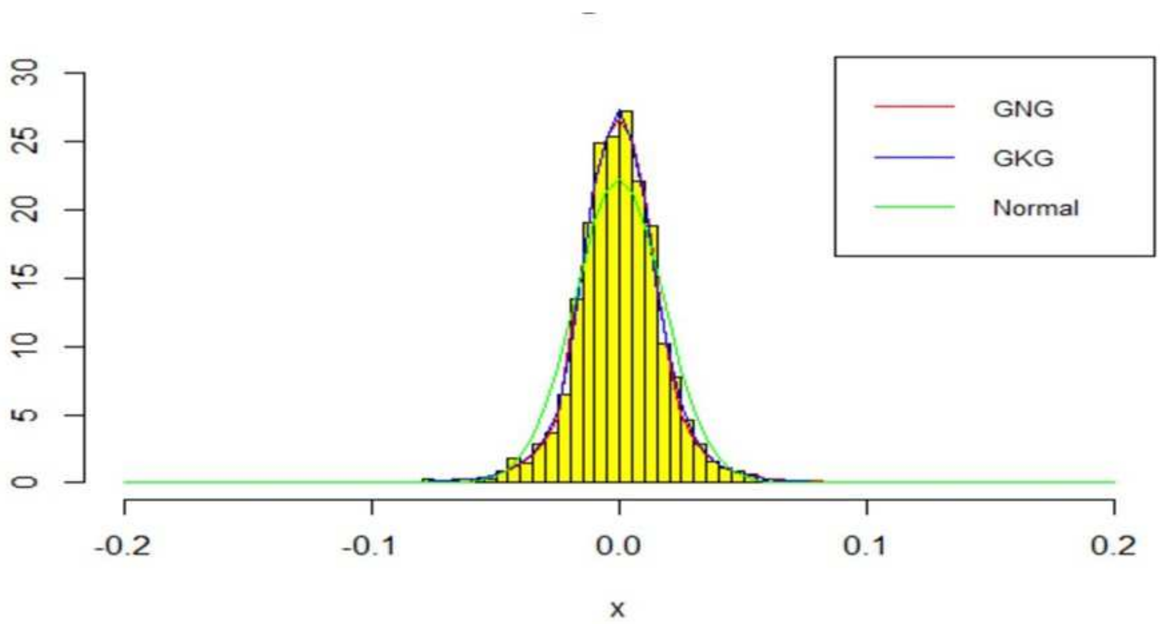

Fig. 4: Fitted GNG and GKG mixture model of the FTSE/Mining Index

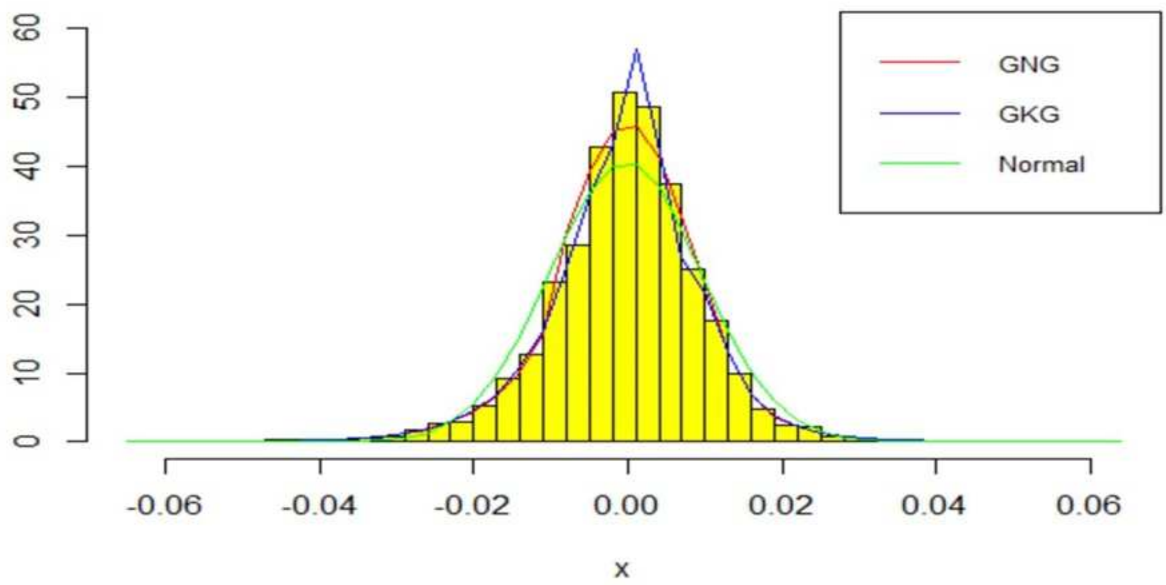

Fig. 5: Fitted GNG and GKG model USD/ZAR

The fitted GNG and GKG model distribution plots are shown in the figures above for each stock indices and South African exchange rate. Evidently, the two-tailed GNG and GKG model is a good fit across the entire range of data as compared to the commonly used Normal distribution. The plot for the Banks Index is visually compressed due to return series having extreme values, despite this, the GNG and GKG model outperforms the Normal model. For the USD/ZAR exchange rate the GKG model outperforms the GNG and Normal model over the entire return series. Observations from the plots lead to the need for more formal model adequacy testing. Application of VaR estimation and Kupiec likelihood test may provide a more formal method of checking for model robustness.

4.2.3 VaR and Backtesting

Table 4: VaR estimates of financial market indices and exchange rate price returns using fitted GNG and GKG model

VaR level							
		Short position			Long position		
		1%	2.5%	5%	95%	97.5%	99%
Fitted GNG model	ALSI	2.1880	2.7394	1.0394	0.1328	0.0360	0.1591
	Banks Index	2.1880	2.3034	1.0011	2.5317	3.2157	¡0.0002
	Mining Index	1.5811	0.3367	0.0076	0.2122	0.2004	0.1645
	USD/ZAR	0.0214	0.0696	0.0599	50.1711	66.2705	34.5848
Fitted GKG model	ALSI	-0.4199	-0.2601	-0.1392	0.0159	0.0205	0.0274
	Banks Index	-0.0454	-0.0334	-0.0255	0.0251	0.0328	0.0461
	Mining Index	-0.0491	-0.0364	-0.0276	0.0273	0.0355	0.0480
	USD/ZAR	-0.0278	-0.0214	-0.0166	0.0145	0.0184	0.0242 b

Table 5: VaR estimates of financial market indices and exchange rate price returns using fitted GNG and GKG model

VaR level							
		Short position			Long position		
		1%	2.5%	5%	95%	97.5%	99%
Fitted GNG model	ALSI	0.1391	0.0980	0.3080	0.7156	0.8495	0.6900
	Banks Index	0.1391	0.1290	0.3170	0.1116	0.0729	0.9984
	Mining Index	0.2086	0.5617	0.9304	0.6451	0.6544	0.6851
	USD/ZAR	0.8836	0.7935	0.8066	¡0.0002	$<$ 0.0002	$<$ 0.0002
Fitted GKG model	ALSI	-	-	-	0.3080	0.7496	0.6900
	Banks Index	0.8401	0.2665	0.7858	0.3170	0.3985	0.6851
	Mining Index	0.5395	0.8457	0.9963	0.9233	0.9514	0.6851
	USD/ZAR	0.8836	0.8895	0.4877	0.4877	0.5411	0.4553

VaR estimates for the FTSE/JSE Indices and USD/ZAR and associated Kupiec p-values are displayed in Table 4 and Table 5 respectively. From Table 5, it is observed that at a 5% level of significance, the Kupiec test indicates that the fitted GNG model is an apt fit at almost all VaR levels for each of the returns since the p-values are greater than 0.05. Thus, the null hypothesis of model adequacy is not rejected. The p-values for the long position on the USD/ZAR exchange rate returns are less than 0.05 and indicate a model misfit, however, a substantial fit is observed in the short position for the return series. Inconclusive results are seen for the short position of the fitted GKG model, however, for all return series, p-values are greater than 0.05 therefore it can be concluded at 5% level of significance the null hypothesis of model adequacy is accepted thus highlighting the robustness of the fitted GKG model.

5 Discussion

The Gaussian framework is frequently used in applications within the financial industry due to many favorable and practical attributes. A few reasons why the Normal distribution is widely adopted in the financial industry is firstly, for the simplicity in the application of numerical methods. Secondly, the Central Limit Theorem address complex problems by assuming and working with models that are approximately Normal. Lastly, Normally distributed random variables assume values around the central mean whereas the odds of deviation exponentially decrease as one deviates from the mean. The article by [17] rejects the normality assumption of the returns on 30 stock market indices. The fitting of alternative models is applied to better capture the leptokurtic behavior observed in the financial return series. [18] evaluated the daily stock returns for 13 European securities. The normality assumption was tested and visibly rejected when compared to four alternative empirical models. There are many studies with empirical evidence to advise that the Normal distribution fails to sufficiently capture properties often seen in financial market stock indies and exchange rates to remedy the limitations and inadequacies of the traditional Gaussian approach, modelling financial asset returns using alternatives is suggested.

Initial descriptive statistics tests were completed to determine the underlying nature of each return series. Each of the daily log returns of the three FTSE/JSE indices and the Dollar/South African exchange rate has shown to be non-stationary, with evidence of heteroscedasticity and volatility clustering which is in line with the characteristics frequently seen in financial data.

The normality assumption is rejected by the Jarque-Bera test thus inferring the shortcomings of the Gaussian approach and leads to the subsequent fitting of alternative models such as the GNG proposed by [13] and the flexible GKG model evaluated by [10]. The evmix package in R was used to compute GNG estimates using a Bayesian inference where the threshold is estimated as a parameter. VaR estimates were calculated and the Kupiec likelihood results show robustness of the GNG and GKG at most VaR level thus emphasizing the use of fitting extreme value mixed models when describing South African financial data where heavy-tails and asymmetry are observed.

6 Conclusion

The day-to-day log returns of 3 FTSE/JSE indices and the Dollar/South African exchange rate (FTSE/JSE All Share Index, FTSE/JSE Banks Index, FTSE/JSE Mining Index and USD/ZAR) were examined using the fitted GPD-Normal-GPD (GNG) and GPD-KDE-GPD mixture model. Analysis in this paper brings to light the proposed two-tailed model by [13] remedies the threshold selection predicament on the traditional GPD fitting approach, which is usually based of intuition, due to the threshold being estimated as a parameter in GNG estimation. This study points out the robustness of the fitted GNG and GKG models as compared to the traditional Normal distribution which is commonly applied in the financial industry. Backtesting results for the South African rand to United States of American Dollar are mixed but overall, relatively good and suggest that the GNG model in the short position as compared to the long position. Favourable results from the Kupiec Likelihood ratio test emphasize the robustness of the GKG model. This work may provide useful insights to researchers or financial analysts who are concerned with extreme value mixture modeling within the South African financial landscape as an alternative to traditional models which often fail to capture the empirical characteristics of financial data such as heavy tails and skewness. As further research, developing alternative choices for the bulk model should be considered where simulation study is carried out to compare the performance to the GNG model which in essence promotes the use of extreme value mixture models in attaining a further understanding of the South African financial industry.

References

[1] BusinessTech. South Africa 'won't be spared' from the Russia-ukraine war: Ramaphosa. Retrieved March 20, 2022, from https://businesstech.co.za/news/government/568890/south-africa-wont-be-spared-from-the-russia-ukraine-war-ramaphosa/(2022).

[2] McNeil, A., & Frey, R. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. Journal of Empirical Finance, 7(3-4), 271-300,(2000). https://doi.org/10.1016/S0927- 5398(00)00012-8

[3] Zhao, X., Scarrott, C., Oxley, L., & Reale, M. Extreme value modelling for forecasting market crisis impacts. In The Global Financial Crisis (pp. 67-76). Routledge(2014).

[4] Behrens, C. N., Lopes, H. F., & Gamerman, D. Bayesian analysis of extreme events with threshold estimation. Statistical Modelling, 4(3), 227-244(2004).

[5] Hu, Y. Extreme value mixture modelling with simulation study and applications in finance and insurance,(2013).

[6] Frigessi, A., Haug, O., & Rue, H. A dynamic mixture model for unsupervised tail estimation without threshold selection. Extremes, 5(3), 219-235(2002).

[7] de Melo Mendes, B. V., & Lopes, H. F. Data driven estimates for mixtures. Computational statistics & data analysis, 47(3), 583-598(2004).

[8] Akakpo, R. M. Value-at-risk (VaR) estimation with extreme mixture model and backtesting diagnostics,(2016).

[9] Naradh, K., Chinhamu, K. and Chifurira, R. Estimating the value-at-risk of JSE indices and South African exchange rate with Generalized Pareto and stable distributions. Investment Management and Financial Innovations, 18(3), 151-165(2021). doi:10.21511/imfi.18(3).2021.14

[10] MacDonald, A. E. Extreme value mixture modelling with medical and industrial applications,(2011).

[11] Coles, S. An introduction to statistical modeling of extreme values. Springer Series in Statistics. Springer-Verlag, London (2001).

[12] Reich, B.J., & Porter, M. D. Discussion of "Estimating the Historical and Future Probabilities of Large Terrorist Events" by Aaron Clauset and Ryan Woodward". Annals of Applied Statistics. 7(4), 1871-1875(2013).

[13] Hu, Y., & Scarrott, C. evmix: An R package for extreme value mixture modeling, threshold estimation and boundary corrected kernel density estimation,(2018).

[14] Escanciano, J., & Olmo, J. Backtesing parametric valueat-risk with estimation risk. Journal of Business & Economic Statistics, 28(1), 36-51 (2010). https://doi. org/10.1198/jbes.2009.07063.
[15] Kupiec, P. Techniques for verifying the accuracy of risk measurement models. The Journal of Derivatives, 3(2), 73-84(1995). https:// doi.org/10.3905/jod.1995.407942
[16] Changole, A. Bad news about economy proves too much for Sa's bank stocks. Fin24. Available at:https://www.news24.com/fin24/companies/bad-news-about-economy-proves-too-much-for-sas-bank-stocks-20190806 (2019). [Accessed August 24, 2021].
[17] Peká r, J.,& Pčolár, M. Empirical distribution of daily stock returns of selected developing and emerging markets with application to financial risk management. Central European Journal of Operations Research, 1-33(2021).
[18] Aparicio, F. M., & Estrada, J. Empirical distributions of stock returns: European securities markets, 1990-95. The European Journal of Finance, 7(1), 1-21(2001).