MODELING CLAIMS OF NIGERIA OIL & GAS INSURANCE CLASS OF BUSINESS WITH EXTREME VALUE THEORY

Agunbiade Dawud Adebayo, Adesina Olumide Sunday

Department of Mathematical Sciences, Olabisi Onabanjo UniversityAgo-Iwoye, Nigeria

Corresponding author: Olumide Sunday Adesina, olumidestats@gmail.com

ABSTRACT: Oil and gas sector has been faced with huge losses in Nigeria as a result; insurance companies pay high premium due to these loses and Extreme Value Theory (EVT) is found to be suitable in modeling the extreme losses. Claims resulting from Nigeria Oil, and Gas insurance class of business for five insurance companies were modeled with EVT to estimate Value-at-Risk (VaR), where VaR measures the minimal anticipated loss over a period with a given probability and under exceptional market conditions. The mean excess plot was obtained which helps in determining the threshold value to be chosen and gives the shape of the distribution in the tail. Q-Q plot being linear and curved tail plots reveals that parametric model fits the data well. VaR based on EVT-Generalized Parameter (GPD) was carried using the chosen thresholds at 5% confidence interval. Results obtained were compared with VaR based on Historical and Gaussian method; it was established that Extreme VaR is most suitable to calculate VaR as against the Historical and Gaussian method.

KEYWORDS: Extreme Value Theory, Value-at-Risk, GPD model, Peak-Over-threshold, Insurance.

1. INTRODUCTION

Oil is considered major source of energy in Nigeria and, it is a sector that contributes most to the economic development as shown in the study by ([Odu07]). The risk of large losses and consequently large insurance claims can be modeled with Pareto, Gamma, and Lognormal distributions for deciding on deductible and premium levels as outlined by ([NEH87]) and Value-at-Risk (VaR) based on Extreme value theory which was adopted by ([WW14, AAO16]) to model returns of fire insurance and motor insurance respectively was found effective.

VaR measures the worst anticipated loss over a period for a given probability and under normal market conditions. It can also be said to measure the minimal anticipated loss over a period with a given probability and under exceptional market conditions ([Lon99]). The advantage of estimating VaR using GPD method is that this method can estimate VaR outside the sampling interval. The main attraction of Value-at-Risk as a risk indicator is that it is able to compress all market factors into a single number ([RG11]).

VaR is estimated using (i) Parametric- assumes that returns have a normal or Gaussian distribution ([LZ96]) (ii) Non-parametric-employs the empirical distribution of returns from the historical sample ([AT02a]); and (iii) Semi-parametric- Specifically applying Extreme Value Theory ([DV98; MF00]), [EM01]) critique of (i) that financial time series data are hardly normally distributed. Seminal works of ([Man63; Fam65]) provide evidence that financial data are negatively skewed and it is leptokurtic with heavier tails that the normal distribution.

With VaR methods, no assumptions are made about the nature of the original distribution of all the observations. Some EVT techniques can be used to solve for very high quantiles, which is very useful for predicting crashes and extreme-loss situations. On the other hand, the historical approach in calculating Value-at-Risk provides advantages of simplicity, and does not assume any particular distribution, ([NR08]).

According to ([Dow05]), the extreme events are generally classified into two: (i) High Frequency Low Severity (HFLS) - risky events that occur frequently but have a minimal impact, and (ii) Low Frequency High Severity (LFHS) - extreme events that occur infrequently but have a high severity.

The second category poses a greater problem, because the high impact on the firm is hidden under the guise of low probability. In the light of these, Extreme Value Theory is a modern and more efficient statistical techniques developed to for analyzing extreme events. Through EVT, the distribution of the tails of a random variable is obtained by modelling extreme observations.

Some studies have also employed Value-at Risk estimation on Treasury Yields ([Bal03]), Future Contracts ([B+05]), Electricity Spot Prices (Chan & Gray, 2006) and Empirical Comparison ([RG11]). These studies show that Extreme VaR dominates the other approaches in forecasting Value-at-Risk, especially in estimating quantiles at the extreme tails. In this study the efficiency of VaR is further

investigated using Nigeria Oil and Gas Insurance class of business.
The remaining part of this study is as follows: in section 2, extreme value is discussed, materials and method adopted is discussed in section 3, while in section 4; results obtained are discussed. Finally in section 5, results were summarized and conclusion is drawn.

2. EXTREME VALUE THEORY

Extreme value theory is developed on the assumption that there is a sequence of independent and identical distribution observation $X_1,\ldots\ldots,X_n$ which is a sequence of independent random variable with population distribution function F and M_n represents the maxima (minima) of the process over a block size n ([FT28]) and ([Gne43]).
The distribution of EVT follows the Gumbel, Frechet, and the Weibull distribution as outlined below

Gumbel: $\Lambda(x) = \exp\{-e^{-x}\}, x \in \mathbb{R}$

$$\text{Frechet: } \Phi_\alpha(x) = \begin{cases} 0, & x \le 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases} \quad (1)$$

$$\text{Weibull: } \psi_\alpha(x) = \begin{cases} \exp\{-(-x^{-\alpha}\}, & x \le 0 \\ & \alpha > 0 \\ 1, & x > 0 \end{cases}$$

2.1 Generalized Extreme Value Distribution (GEV)

The three types of distributions introduced in (1) may be combined into a single distribution using the point process theory with the cdf as follows:

$$G(x) = \exp\left[-\left(1+\xi\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\xi}\right] \quad (2)$$

The shape parameter ξ is crucial in determining the class of GEV distribution. When ξ>0 the distribution is known as Frechet which is characterized by heavy tailed distribution, when ξ=0 it is a Gumbel distribution, and it is known for thin tailed distribution and when ξ<0 it is a Weibull distribution characterized by finite endpoint distribution. The classic extreme value theory can be divided into two groups; (Block maximum and Peak over threshold). In modeling excess distribution; the interest is estimating the distribution function F_u of values of x above a high threshold u. The distribution F_u is called the conditional excess distribution function F of a random variable X as:

$$F_u(y) = P(X - u \le y \mid X > u), \quad (3)$$
$$0 \le y \le x_F - u$$

$y = x - u$ are the right endpoint of F. We verify that F_u can be terms as F,

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)}$$
$$= \frac{F(x) - F(u)}{1 - F(u)}, 0 \le X \le u \quad (4)$$

From GPD, the tail index ξ gives an indication of heaviness or lightness of the tail, the larger ξ the heavier the tail. Only distributions with tail parameter $\xi \ge 0$ are fit to model returns.
Various authors have shown that there the two major ways of modelling extremes of stochastic variable using the extreme value models. One approach is by dividing the sample into blocks and then obtains the maximum from each block, which is referred to as the *block maxima* method. The distribution of the block maxima can be modelled by fitting it into Generalized Extreme Value (GEV) model ([Upp13]). The limitations observed in GEV model necessitate the use of an alternative and preferred approach to modeling time series data, the peak-over-threshold (POT), it take observations which exceed a certain threshold u and the POT models are generally considered to be the most useful for practical applications ([RT02]). The distribution of the exceedances is obtained by employing Generalized Pareto distribution (GPD). ([RG11]) pointed out that the advantage of estimating VaR using GPD method is that the method can estimate VaR outside the sampling interval.

3. MATERIALS AND METHOD

3.1 The Generalized Pareto Distribution (GPD)

The model adopted in this study is generalized Pareto distribution. The cumulative density function is:

$$G_{\xi,\beta} = 1 - \left(1 + \xi \frac{x}{\beta}\right)^{\frac{-1}{\xi}}, \xi \neq 0 \quad (5)$$
$$G_{\xi,\beta} = 1 - \exp(-x/\beta), \xi = 0$$

where $\beta > 0, x \geq 0$ when $\xi \geq 0$ and $0 \leq x \leq -\frac{\beta}{\xi}$ when $\xi < 0$.

ξ is the important shape parameter of the distribution and β is an additional scaling parameter. If $\xi > 0$, the $G_{\xi,\beta}$ is re-parametrized version of the ordinary Pareto distribution use for large losses, $\xi = 0$ corresponds to the exponential distribution and $\xi < 0$ is known as a Pareto of type II distribution.

The case where $\xi > 0$ is the most relevant for risk management. Whereas normal distribution has moments of all orders, a heavy-tailed distribution does not possess a complete set of moments.

3.2 Maximum Likelihood Estimation for GPD

Maximum Likelihood Estimate (MLE) is a statistical technique for estimating model parameters. It helps to determine model parameters that are most likely to characterize a given set of data. After selecting a model for a given data, the model must have one or more (unknown) parameters. As the name implies, MLE proceeds to maximize a likelihood function, which in turn maximizes the agreement between the model and the data, in other word, the values of these parameters maximizes the sample likelihood. With $G_{\xi,\beta}$ for the density of the GPD, the log-likelihood may be calculated to be:

$$L(\xi,\beta;Y_j) = -N_u In\beta + \left(\frac{1}{\xi} - 1\right)\sum_{j=1}^{N_u}\left(1 - \xi\frac{Y_j}{\beta}\right), \xi \neq 0 \quad (6)$$

$$L(\xi,\beta;Y_j) = -N_u In\beta - \frac{1}{\beta}\sum_{j=1}^{N_u} Y_j, \ \xi = 0 \quad (7)$$

Further working reveals that:

$$\hat{\xi}_{MLE} = -\left(\frac{1}{N_u}\right)\sum_{j=1}^{N_u} In\left(1 - \hat{\theta}_{MLE} Y_j\right)$$

and

$$\hat{\beta}_{MLE} = \frac{\hat{\xi}_{MLE}}{\hat{\theta}_{MLE}} \quad (8)$$

3.3 Distribution of Exceedances

Excess distribution is modeled using peak over threshold (POT) method for distribution of exceedances, considering an unknown distribution function F of a random variable X. The interest is estimating the distribution function F_u of values of x above certain threshold u.

The distribution F_u is called the conditional excess distribution function and is defined as:

$$F_u(y) = P(X - u \leq y \mid X > u),$$
$$0 \leq y \leq x_F - u \quad (9)$$

Where X is a random variable, u is a given threshold, $y = x - u$ are the right endpoint of F. We verify that F_u can be terms as F,

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)}$$

i.e

$$= \frac{F(x) - F(u)}{1 - F(u)}, 0 \leq X \leq u \quad (10)$$

from GPD, the tail index ξ gives an indication of heaviness of the tail, the larger ξ, the heavier the tail. Only distributions with tail parameter $\xi \geq 0$ are suitable to model financial returns.

3.4 Value-at-Risk

Value at Risk is a threshold loss value, such that the probability that the loss on a given portfolio over a given time horizon exceeds an estimated value. Value at Risk is the capital sufficient to cover, in most instances, loss from a portfolio over a holding period of a fixed number of days.

The Value at Risk of a random variable X with continuous distribution function F that models losses or negative return on a certain financial instrument over a certain time horizon is given by:

$$VaR_p = F^{-1}(1-p)$$

Where VaR_p is the p^{th} quantile of the distribution F.

Where F^{-1} the so called quantile is function, and is the inverse of the distribution function F.

The VaR_p the tail distribution of a GPD function can be defined as a function of GPD parameters.

Expected shortfall is defined as the excess of a loss that exceeds VaR_p. Isolating $F(x)$ from (10).

$$F(x) = (1 - F(u))F_u(y) + F(u)$$

And replacing F_u by the GPD and $F(u)$ by the estimate $\left({n - N_u}/{n} \right)$, where n is the total number of observations and N_u the number of observations above the threshold u, we obtain

$$\hat{F}(x) = \frac{N_u}{n}(1 - (1 + \frac{\hat{\xi}}{\hat{\beta}}(x - u)^{-\frac{1}{\hat{\xi}}} + \left(1 - \frac{N_u}{n}\right) \tag{11}$$

Simplifying we have:

$$\hat{F}(x) = 1 - \frac{N_u}{n}((1 + \frac{\hat{\xi}}{\hat{\beta}}(x - u)^{-\frac{1}{\hat{\xi}}} \tag{12}$$

Inverting (11) for a given probability

$$V\hat{a}R_p = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{n}{N_u} p \right)^{\hat{\xi}} - 1 \right) \tag{13}$$

Expected shortfall can be written as:

$$ES_p = V\hat{a}R_p + E(X - V\hat{a}R_p \mid X > V\hat{a}R_p) \tag{14}$$

Where the second term on the right is the expected value of the exceedances over VaR_p. The excess function for the GPD with parameter $\xi < 1$ is:

$$e(w) = E(X - w \mid X > w) = \frac{\sigma + \xi w}{1 - \xi}, \tag{15}$$

$\sigma + \xi w > 0$

The function gives average of the excesses of X over varying values of threshold w.

Similarly, from (14), for $z = VaR_p - u$ and X representing the excesses y over u we obtain

$$\hat{E}S_p = V\hat{a}R_p + \frac{\hat{\sigma} + \hat{\xi}(V\hat{a}R_p - u)}{1 - \hat{\xi}}$$

$$= \frac{V\hat{a}R_p}{1 - \hat{\xi}} + \frac{\hat{\sigma} - \hat{\xi}}{1 - \hat{\xi}} \tag{16}$$

The computational formulas for the three approaches used in this paper are as follow:

(i) Gaussian VaR

$$VaR(\alpha) = u + (N^{-1}(1 - \alpha))\beta$$

(ii) Historical VaR

$$VaR(\alpha) = ((1 - \alpha) \times N)^{th}$$

(iii) Extreme VaR

$$V\hat{a}R_p(\alpha) = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{n}{N_u} p \right)^{\hat{\xi}} - 1 \right)$$

Where $n =$ number of observations in the parent distributions, N_u = number of tail observations with parameters β and ξ substituted with the maximum likelihood estimates.

The data set for this study was collected from annual publication of Nigeria Insurance digest, where claims that were made over a period of 2 years; (2011 and 2012) was obtained. Data of five Insurance companies for non-life was collected for the analysis and estimation. The companies include; Custodian, Leadway, Mansard, Zenith and Aiico insurance. The presence of extreme values was tested in the data using the Grubb's test and quartile test and presence of extreme (outliers) were found in each dataset.

In order to implement computation outlined in this study software package by ([RCT17]) is used. The data set was fitted into a GPD Model with maximum likelihood estimate, fExtremes package in R by ([W+13]) was used for Extreme VaR and PerformanceAnalytics package was used for Historical and Gaussian VaR.

4. RESULTS

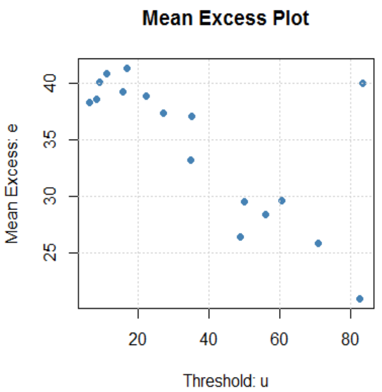
Table 1: Descriptive Statistics of Oil and Gas Class of Business Claims

Company	Descriptive			
	MEAN	**STDEV**	**SKEW**	**KURT**
Leadway	42.37101	35.56762	0.806973	-0.21832
Aiico	8.4009	6.218003	0.370206	-2.33333
Zentih	24.51326	22.32984	0.025004	-2.25008
Mansard	19.56729	14.99951	0.423351	-1.88089
Custodian	68.16553	60.82756	0.670910	-0.46239

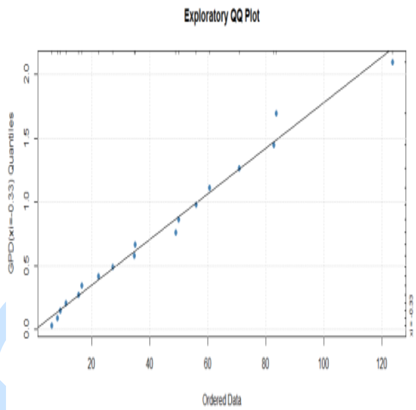
Table 2: Grubb's test for extreme Value, H_A : There is outlier in the data set

Company	*G*	*U*	*p*-value	Outlying Value
Leadway	2.4946	0.6124	0.0521	123.6134
Aiico	1.1498	0.0085	0.0882	15.55020
Zenith	1.1206	0.6986	0.7777	49.53673
Mansard	1.4551	0.3383	0.2349	49.53673
Custodian	2.3223	0.5533	0.0628	209.4480

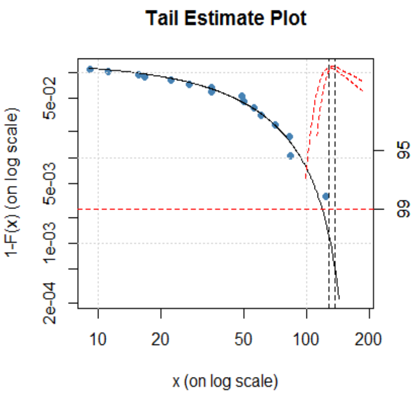
The above test statistics reveals that there are outlying value(s) in the each data set.



(a) Leadway Mean Excess

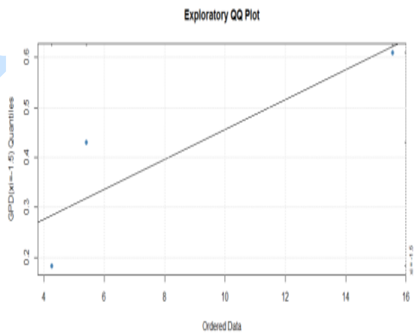


(b) Leadway exploratory QQ plot

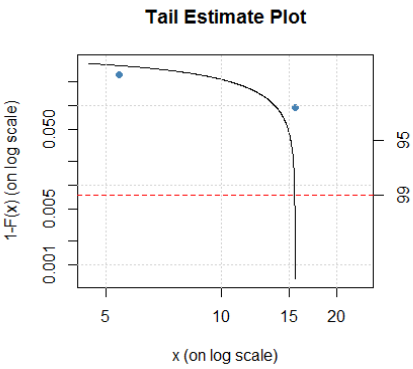


(c) Leadway tail plot

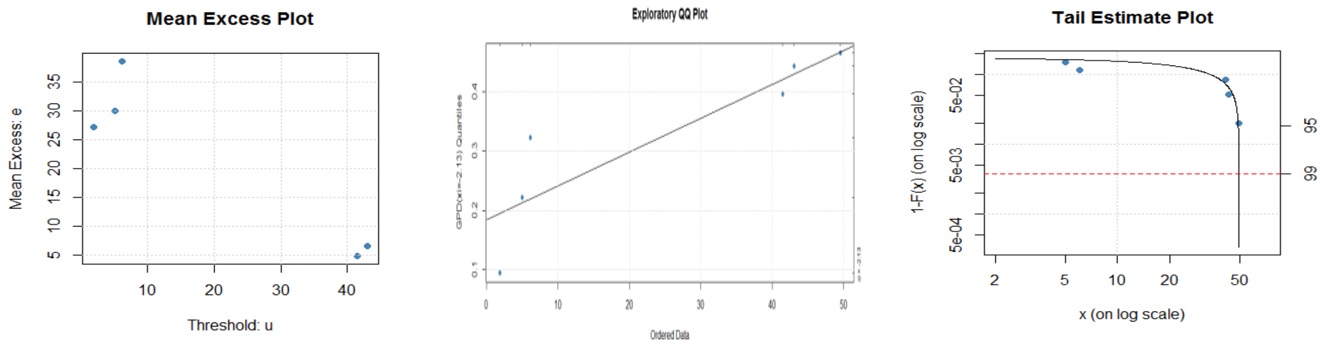
(d) Aiico Mean Excess



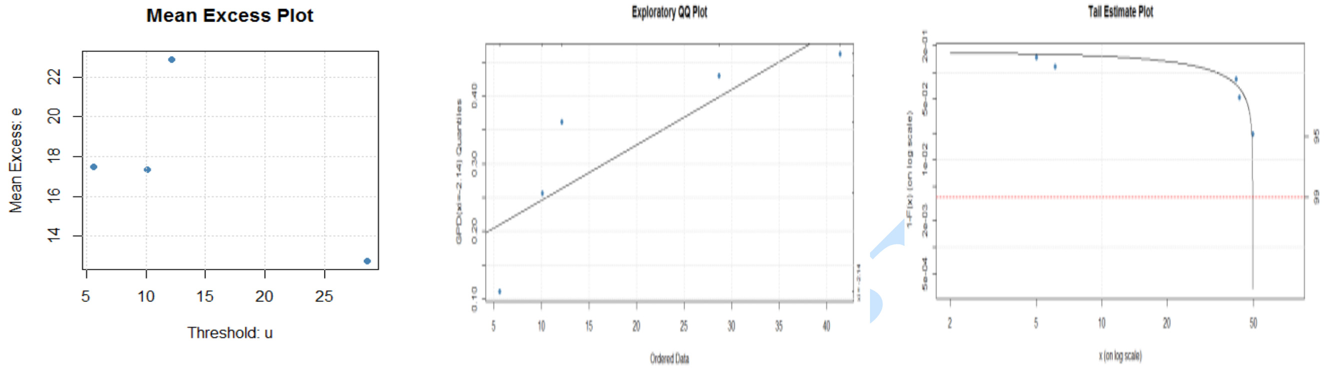
(e) Aiico exploratory QQ



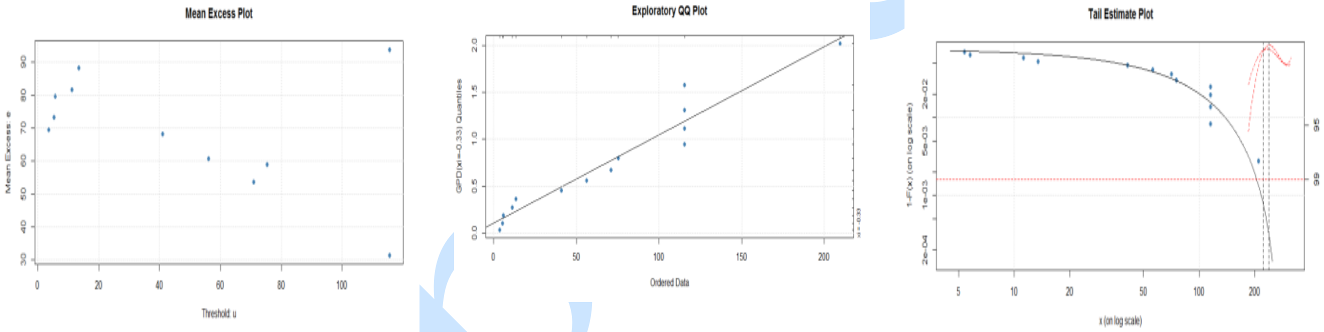
(f) Aiico tail plot



(g) Zenith Mean Excess **(h) Zenith exploratory QQ plot** **(i) Zenith tail plot**



(j) Mansard Mean Excess **(k) Mansard exploratory QQ plot** **(l) Mansard tail plot**



(m) Custodian Mean Excess **(n) Custodian exploratory QQ plot** **(o) Custodian tail plot**

Figure 1: Graphic representation of data in Table 1

Table 3: Value at risk, Threshold Selection and parameter of estimates

Value-at Risk estimates at α =5%				Threshold (u) Selected and GPD Parameters		
Company	**Historical**	**Gaussian**	**Extreme VaR**	u	ξ	β
Leadway	-89.62	9.69	6.2102	9	-0.3325	50.876
Aiico	-14.54	-0.05	4.2525	4.5	-1.5050	16.625
Zenith	-47.92	9.01	1.8491	2	-2.1361	101.54
Mansard	-38.94	2.50	5.5758	5.5	-2.1361	101.54
Custodian	-148.8	28.24	3.5998	4.5	-0.3283	92.480

An empirical analysis was carried out to model the Oil and Gas insurance claims of the Nigerian Insurance market. Five insurance companies' claims that were captured were modeled and discussed. From table 1, all the five company's data for Oil & Gas insurance class of business are positively skewed, they also all have fat tails with kurtosis greater than that of normal distribution.

The threshold u, used to fit the GPD model for the five companies are stated in columns 5 in table 3. The Oil and Gas industrial insurance claim data was fitted in GPD model and discussed. Comparison of the three types of VaR estimate was equally made in table 3, as given above.

Value at Risk (VaR) for Leadway with 5% level of confidence was #6.2104 million, is a 1-in-20. For twenty four months period, it implies that the coming one month six day's loss for the company would exceed #6.2104 million. Value at Risk (VaR) for Aiico with 5% level of confidence was #4.2525 million, is a 1-in-20. For twenty four months period, it implies that the coming one month six day's loss for the company would exceed #4.2525 million. The Value at Risk (VaR) for Zenith with 5% level of confidence was #1.849182 million, is a 1-in-20. For twenty four months period, it implies that the coming one month six day's loss for the company would exceed #1.849182 million.

The Value at Risk (VaR) for Mansard with 5% level of confidence was #5.575801 million, is a 1-in-20. For twenty four months period, it implies that the coming one month six day's loss for the company would exceed #5.575801 million. The Value at Risk (VaR) for Custodian with 5% level of confidence was #3.599844 million, is a 1-in-20. For twenty four months period, it implies that the coming one month six day's loss for the company would exceed #3.599844 million. Then precautions can be taken to mitigate against it.

5. SUMMARY AND CONCLUSION

Detailed analysis is carried out to model Oil and Gas industrial insurance returns of the Nigerian insurance companies. The presence of extreme values was tested using Grubb's test, the datasets were found to contain extreme values. The POT-EVT along with Historical and Gaussian approach was used to estimate Value-at-Risk. Estimates of Value-at-Risk using Historical and Gaussian methods gives negative values; which is not acceptable. Findings in this study reveal that Extreme Value Theory method of calculating VaR outweighs other methods of estimation as it is known for its ability to model the tail area of the distribution much well. Other methods of estimation may be suggested for large datasets.

REFERENCES

[AT02a] **Acerbi C., Tasche D.** - *Expected Shortfall: A natural coherent alternative to value at risk.* Economic Notes 31, 379-388. 2002.

[AT02b] **Acerbi C., Tasche D.** - *On the coherence of expected shortfall*, Banking and Finance, 26: 1487-1503, 2002.

[AAO16] **Adesina O. S., Adeleke I., Oladeji T. F.** - *Using Extreme Value Theory to Model Insurance Risk of Nigeria Motor Industrial Class of Business*. The Journal of Risk Management and Insurance, Assumption University, Vol 20, No 1, pp 40-51. 2016.

[Bal03] **Bali T. G.** - *An Extreme Value Approach to Estimating Volatility and Value-at-Risk*, Journal of Business Vol. 76, No. 1, 2003, pp. 83-108. doi:10.1086/344669.

[BH74] **Balkema A. A., De Haan L.** - *Residual life time at great age*. Annals of probability, (2):792-804, 1974.

[B+05] **Brooks C., Clare A., Molle J., Persand G.** - *A comparison of extreme value theory approaches for determining value at risk*. Journal of Empirical Finance 12, 339-353. 6-11, 2005.

[CG06] **Chan K., Gray, P.** - *Using extreme value theory to measure value-at-risk for daily electricity spot prices*. International Journal of Forecasting 22, 283-300. 2006.

[Dow05] **Dowd K.** - *Measuring market risk* (2nd ed.). West Sussex, England: Wiley. 2005.

[DV98] **Danielsson J., de Vries C.** – *Beyond the Sample: Extreme quantile and probability estimation* (Discussion Paper 298), London School of Economics. 1998.

[EM01] **Engle R., Manganelli S.** - *Value-at-Risk Models in Finance*, ECB Working Paper No. 75, European Central Bank Working Paper Series. 2001.

[Fam65] **Fama E.** - *The behavior of stock market prices*, Journal of Business 38, 34-105. 1965.

[FT28] **Fisher R., Tippet L. H. C.** – *Limiting forms of the Frequency distribution of largest or smallest members of a sample*. Proceeding of the Cambridge philosophical society, 24: 180-190. 1928.

[Gne43] **Gnedenko B. V.** - *Sur la distribution limite du term d'une series aleatoire*. Annals of mathematics, 44: 423-453, 1943.

[Lon99] **Longuin M. F.** - *From Value at Risk to Stress Testing: The Extreme Value Approach*. Center for Economic Policy Research, Discussion Paper No. 2161, 1999.

[LZ96] **Longerstaey J., Zangari P.** - *RiskMetrics* – Technical Document// J.P. Morgan, New York: Fourth edition, 1996.

[Man63] **Mandelbrot B.** - *The Variation of Certain Speculative Prices*, Journal of Business 36, 394-419, 1963.

[MF00] **McNeil A., Frey R.** – *Estimation of tail-related risk measures for heteroscedastic financial time series*: *an extreme value approach*, Journal of Empirical Finance, Vol. 7, pp. 271-300. 2000.

[NR08] **Nieto M., Ruiz E.** - *Measuring Financial Risk: Comparison of Alternative Procedures to estimate VaR and ES,* Statistics and Econometrics Working Paper – University Carlos III, 45 pages. 2008.

[NEH87] **Nigm A. M., El-Habashi M. H., Hamdy H. I.** - *The effect of risk parameters on decision making, Insurance, Mathematics and economics*, 6, 237-244. 1987.

[Odu07] **Odularo G. O.** - *Crude Oil and the Nigerian Economic Performance*, Oil and Gas Business,1-29, 2007.

[RG11] **Rufino C. C., Guia E.** - *Emprirical Comparison of Extreme Value Theory Vis-à-vis Other Methods of VaR Estimation Ysing ASEAN+3 Exchange Rates*. DLSU Business & Economics Review 20.2 pp.9-22, 2011.

[RT02] **Reiss R. D., Thomas M.** - *Statistical Analysis of Extreme Values with Application to Insurance, Finance, Hydrology and other fields* (Third Edition). 2002.

[RCT17] **R Core Team** - *R: A language and environment for statistical computing. R Foundation for Statistical Computing*, Vienna, Austria. URL https://www.R-project.org/. 2017.

[Upp13] **Uppal Y. J.** - *Measures of Extreme Loss Risk - An Assessment of Performance During the Global Financial Crisis*: Journal of Accounting & Finance. Vol 13, No. 3, 2013.

[WW14] Wainnaina H. W., Waititu A. G. - *Modelling Inurance Returns with Extreme Value Theory (A Case study for Kenya's Fire Industrial Class of Business)*. Journal of Mathematical Theory and Modeling. pp. 48-53, 2014.

[W+13] **Wuertz D. et al.** - *fExtremes: Rmetrics* – 2013, on: https://cran.r-project.org/web/packages/fExtremes/index.html.