Menemui Matematik (Discovering Mathematics)
Vol. 44, No. 2: 109-120 (2022)

Value at Risk measure on Oil Price by using Extreme Value Theory Approach

Putra Nor Hakimi Bin Jajiman[1], Wendy Ling Shinyie[1,a) and Nur Amirah Binti Buliah[1]

[1]*Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, 43400 Serdang, Malaysia*

[a)]Corresponding author: sy_ling@upm.edu.my

ABSTRACT

The oil market is well known for its unpredictable market trend and volatility, which makes the trading risk in this market high and can lead to huge losses. The purpose of this study is to measure the risk of extreme returns in the oil market. Effective risk management can avoid investors from suffering huge losses. The data used in this study is 10-year daily return futures price of the two most traded commodities on the oil market, West Texas Intermediate (WTI) and Brent Crude Oil. This study utilises Value at Risk (VaR) as a measure of risk. Extreme Value Theory (EVT) is used to improve the reliability of the risk measurement. As the extreme values in the data series are limited, the peaks over threshold (POT) method is used to extract all values above a certain threshold, which is considered as the limit value in the process of parameter estimation and modelling. By fitting the excesses to Generalized Pareto distribution (GPD), we can obtain the estimation of the risk measure. The model is then assessed by backtesting to identify whether the estimated risk measure can capture the risk accurately. The backtesting methods used in this study are the Christoffersen test and the Basel backtesting. The findings show that Extreme-VaR captured the risk perfectly. Therefore, this approach can be used by investors as a risk management tool in the portfolio management.

Keywords: Value at risk, oil market, extreme value theory, Christoffersen backtesting

INTRODUCTION

According to the United States Energy Information Administration (EIA), crude oil is formed from the decomposed marine animals and plants that lived millions of years ago. In the early days of this discovery, paraffin was the only crude oil product used to light lamps and heat oil. Over the past few decades, people have invented many technologies to completely transform oil into a more practical product, such as gasoline, diesel and jet fuel. Therefore, oil has become the most important commodity in the world, because gasoline is still the main fuel source for most vehicles. As oil becomes the most important commodity in the world, many countries are competing to produce and reserve a large amount of oil to gain economic advantages. Countries with large reserves can decide the oil supply, and then the oil price. Therefore, if the world economy experiences a recession, they can turn the situation into a favourable situation for themselves. According to the BP Statistical Review of World Energy (2019), at the end of year 1998, the total proven oil reserves in Saudi Arabia were 261.5 billion barrels, ranking first in the world, followed by Russia with 113.1 billion barrels. At the end of year 2008, Saudi Arabia still ranked first with its reserves of 264.1 billion barrels, and Canada ranked second with its reserves of 176.3 billion barrels. But by the end of 2017, Venezuela had become the country with the largest oil reserves with 302.8 billion barrels, and Saudi Arabia ranked second with 296 billion barrels.

According to the annual review of the World Gold Council in 2018, the global daily output is about 94 million barrels per day, and the size of the oil market is expected to be 1.7 trillion US dollars per year. By contrast, other important commodities markets are the metal market. It is worth noting that the size of the gold market is about USD170 billion per year. Other commodities such as iron and copper are only about USD115 billion per year and USD91 billion per year respectively. Objectively speaking, the crude oil market is bigger than the sum of all the metal markets and ten times larger than the gold market. In view of the scale of crude oil market, the crude oil market is naturally an important market for investors to analyse.

The oil market is very unstable because it is greatly affected by the relationship between supply and demand, and its supply is controlled to some extent by its main supplier, the Organization of the Petroleum Exporting Countries (OPEC), and other European countries including the United States and Russia. For oil market participants, high volatility is accompanied by great risks. Extreme risk exists in all areas of risk management. For analysts, there are many different types of risk that can be used to measure whether we pay attention to market, credit, operational or insurance risk. One of the biggest challenges faced by risk manager is to implement risk management models, which allows rare but destructive events to occur and its consequences to be measured.

In the past decade, WTI's price change percentage of one trading day was as high as 14%, and it was as low as -12% between 2009 and 2018. At the same time, in the same period, the price change percentage of Brent crude oil was 11% at the highest and -9% at the lowest. In particular, on 7th January 2009, the price dropped by 12% at USD42.63 per barrel for WTI, which was the biggest one-day drop in this period. However, this is not the lowest oil price ever recorded for WTI. On 11th February 11 2016, WTI hit the lowest price per barrel in 10 years, at USD26.21 per barrel. On the same day of the year, Brent also recorded its biggest 9% price drop through that period which closed at USD45.86 per barrel. However, during this period, the lowest price record of Brent crude oil was on 20th January 2016, at USD 27.88 per barrel. Compared with the average daily negative return of WTI and Brent crude oil of -1.6% and -1.4% respectively during this period, these cases provide examples of extreme events. If investors and risk management organizations failed to foresee the risks in these markets, the losses could be huge. The high volatility of crude oil prices and the significant influence of this volatility prompt us to study the modelling of oil price volatility and providing an effective tool to measure the risk of energy price.

This study aims to measure risks of extreme returns on oil prices. Return on investment can be defined as profit-and-loss as a percentage of the previous investment. Measuring the risk of extreme losses can prevent investors from losing money when they invest. To measure the risk of oil price returns, we will fulfill the specific objectives of this study which are: (i) to determine the exceedances over a threshold at high quantiles and fit the exceedances to the generalized Pareto distribution (GPD); (ii) to estimate parameters scale and shape parameter of GPD by using maximum likelihood estimation (MLE); and (iii) to estimate the value at risk and test the procedures by using Christoffersen backtesting procedure.

The remainder of the article is organized as follows. In next section, we briefly review the relative literature. The data and methods used are introduced in third section, followed by the empirical results and discussion. Finally, we will present the conclusion and policy implication.

LITERATURE REVIEW

Risk in finance is defined as the probability that the actual return of investors is different from the expected return. In other words, the risk is the possibility that investors will suffer losses from their initial investment. The relationship between risk and return is that the higher the risk, the higher the expected return of investors. Therefore, to get the expected return, investors need to evaluate how much risk they need to take. Some tools to measure the risk are Value at Risk (VaR) and Capital Asset Pricing Model (CAPM). However, different risks need different instruments to measure them, so it is important to recognize the tools that can be used to measure the risks involved.

The traditional VaR approach has become the focus of some criticisms due to its incompetence in measuring the VaR. For instance, the non-parametric approach suffers from several problems such as the heavy reliance on data and the problems of assigning equal weight to each of the observations (Bensalah, 2000; Abad et al., 2014). The parametric approach works under the assumption of financial returns following a normal distribution, which explains the reason this model tends to fail when estimating the left tail of the return distribution because common financial data have a fat-tail series (Bensalah, 2000; Odening & Hinrichs, 2002; Abad et al., 2014). The risk measured by VaR under this assumption is usually underestimated especially for the financial data which are common with the fat-tailed series (Bensalah, 2000). Gencay and Selcuk (2004) found that common financial data are constricted to a fat-tailed distribution, while the traditional VaR underestimated the risk measure because the traditional distribution failed to consider the extreme situation at the tail part of the distributions. Jondeau and Rockinger (1999) and Neftci (2000) further discussed the study of the tail behaviour of financial series.

Bensalah (2000) applied the EVT approach in measuring VaR to a series of daily exchange rates of Canadian and U.S. dollars over a 5-year period (1995-2000) to study whether EVT will improve the VaR estimation. Bensalah (2000) concluded that the EVT approach is a useful addition to measuring VaR because EVT focuses on the extreme observations that lie at the tail part of the distribution. By modelling the tail part of the distribution, EVT greatly complements traditional VaR. Gencay and Selcuk (2004) studied the performance of different VaR approaches on the daily stock market returns of nine different emerging markets. The approaches used in their study are the historical simulation, variance-covariance method and EVT approach, which are used to estimate the daily returns at high quantiles of 0.999 and 0.95. EVT approach makes better predictions than other models (such as t-distribution, ARCH, and GARCH family models) because these models assume symmetric in the data distribution while EVT allows asymmetry in the data distribution because EVT only deals with the tail of the distribution. Gencay and Selcuk (2004) further concluded that EVT method models from GPD accurately estimated VaR, which was an important part of risk values, especially in risk management of emerging markets.

Mcneil (1999) provided a general review of EVT in risk management. To measure extreme risks, Mcneil focused on two measures to describe the tail of the loss distribution, which are VaR and Expected Shortfall (ES). McNeil (1999) used two type of POT models to model extreme values, namely (i) Hill estimator and its relatives proposed by Beirlant et al. (1996) and Danielsson et al. (1998) and (ii) parametric approach based on maximum likelihood estimator (MLE) of Generalized Pareto Distribution proposed by Embrechts et al. (1997). McNeil (1999) found that methods based on assumptions of normality and historical simulation are very likely to underestimate the tail risk and provide inaccurate estimates of tail risk. Therefore, he concluded that EVT is the best approach whenever the tail of the distribution is of interest.

To determine whether the VaR estimates are reliable, a testing method that is called backtesting is required to check the adequacy of the developed model. Halilbegovic and Vehabovic (2016) defined backtesting as the process of evaluating the difference in the value of the predicted VaR with the actual gains and losses. Brown (2008) stated that the backtest of VaR is just as important as the modelling method of VaR. According to Christoffersen (2000), backtesting in finance is defined in two ways; either (i) test on the proposed trading plan on the theoretical historical performance or (ii) test on the financial risk model using historical data on risk predictions and loss-profit realization.

Christoffersen (1998) defines a good VaR model, which should satisfy two properties, namely unconditional coverage and independent coverage. The unconditional coverage property sets an interval for the frequency of expected VaR violations. However, it does not provide any information about whether the violations occur independently. Therefore, the independence test is necessary to test the independence between observed violations in the data series. Campbell (2005) highlights that it is important to understand that both independent property and unconditional coverage property are different and distinct, and an accurate VaR model should satisfy both properties. VaR backtesting is an essential process of comparing the percentage of actual return exceeding the estimate of risk value with the confidence level used when estimating risk value.

MATERIALS AND METHODS

Data Used

This study only focuses on the two main traded commodities in the oil market. The two most traded commodities are West Texas Intermediate (WTI), where its future contracts are traded on the New York Mercantile Exchange (NYMEX), and Brent Crude Oil (Brent), which is traded on the Intercontinental Exchange (ICE).

Generalized Pareto Distributions (GPD)

According to the Pickands-Balkema-de Haan Theorem, the peaks over threshold (POT) method considers the distribution of exceedances over a certain threshold. The distribution function of the exceedances over a certain threshold is called the conditional excess distribution function, F_u. Let $(r_1, r_2, r_3, \dots, r_n)$be the financial returns and $\left(y_1, y_2, y_3, \dots, y_{N_u}\right)$ be all the financial returns exceeding a threshold u, where every exceedance y is defined as $y_i = r_i - u$, and N_u is the number of exceedances that is greater than the threshold$u$. The conditional excess distribution function is defined as follows.

$$F_u(y) = \Pr(r - u \leq y | r > u) = \frac{F(y+u) - F(u)}{1 - F(u)}, \quad 0 \leq y \leq x_F - u$$

where y represents the exceedances above the threshold$u$, and $x_F \leq \infty$ is the right endpoint. By assuming the distribution of the conditional excess distribution function follows a generalised Pareto distribution (GPD) for a certain threshold u, the conditional excess distribution function F_u can be written as,

$$F_u(y) = G_{\xi,\sigma}(y) = 1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-\frac{1}{k}}$$

where ξ is the shape parameter, and σ is the scale parameter. From the definition of exceedances y, $y_i = r_i - u$, financial return r can be defined as $r_i = y_i + u$. Hence, the distribution function of returns is written as,

$$F(r) = F(y+u) = \left(1 - F(u)\right)F_u(y) + F(u) = \left(1 - F(u)\right)G_{\xi,\sigma}(y) + F(u)$$

The tail estimator can be constructed by estimating the $F(u)$, where it can be estimated by defining $F(u) = 1 - \frac{N_u}{n}$; N_u is the total number of observations exceeds the threshold u, and n is the total number of sample observations. The tail estimator is obtained by substituting GPD and $F(u)$ estimator into the distribution function of the return, $F(r)$. Hence,

$$\hat{F}(r) = \frac{N_u}{n}\left(1 - \left(1 + \frac{\hat{\xi}}{\hat{\sigma}}(r-u)^{-\frac{1}{\hat{\xi}}}\right)\right) + \left(1 - \frac{N_u}{n}\right) = 1 - \frac{N_u}{n}\left(1 + \frac{\hat{\xi}}{\hat{\sigma}}(r-u)\right)^{-\frac{1}{\hat{\xi}}}$$

Parameter Estimation

Maximum Likelihood Estimation (MLE) is used in estimating the parameters σ and ξ in the GPD. MLE is used because the estimator obtained from MLE is asymptotically normal, which can be used to calculate standard errors and confidence intervals. After obtaining a suitable threshold u with the distribution of GPD, the individual probability density function in log form derived from the GPD is

$$\log f(r_i) = \begin{cases} -\log(\sigma) - \dfrac{1+\xi}{\xi}\log\left(1 + \dfrac{\xi}{\sigma}(r_i - u)\right), & \xi \neq 0 \\ -\log(\sigma) - \dfrac{1}{\sigma}(r_i - u), & \xi = 0 \end{cases}$$

The log-likelihood function for GPD is the log of the joint density of the n observations.

$$L(\sigma, \xi | r_i - u) = \begin{cases} -n\log(\sigma) - \dfrac{1+\xi}{\xi}\displaystyle\sum_{i=1}^{n}\log\left(1 + \dfrac{\xi}{\sigma}(r_i - u)\right), & \xi \neq 0 \\ -n\log(\sigma) - \dfrac{1}{\sigma}\displaystyle\sum_{i=1}^{n}(r_i - u), & \xi = 0 \end{cases}$$

The estimates of the shape parameter ξ and scale parameter σ can be estimated by taking the maximum log-likelihood function for the sample corresponding to a suitable threshold u.

The VaR estimate for the significance level α is obtained by inverting the tail estimator. Hence,

$$\widehat{VaR}(\alpha) = u + \frac{\hat{\xi}}{\hat{\sigma}}\left(\left(\frac{n}{N_u}(1-\alpha)^{-\hat{\xi}} - 1\right)\right)$$

Backtesting

Once the VaR estimates are obtained, it is necessary to test the adequacy of the model if the model captures the risk accurately. In this study, the Christoffersen backtesting method and the Basel backtesting method are used to testing the model adequacy of the VaR models.

1. Christoffersen Test

The Christoffersen test was developed by Christoffersen in 1998 as an extended study to Kupiec's unconditional coverage test that was proposed in 1995. The Christoffersen test was also known as the Christoffersen mix test is the combination of tests on the unconditional coverage property and the independence coverage property. The hit sequence of VaR violations is defined as,

$$V_{t+1} = \begin{cases} 1, & PL_{t+1} > VaR_{t+1}^{p} \\ 0, & PL_{t+1} \leq VaR_{t+1}^{p} \end{cases}$$

where VaR_{t+1}^{p} is a VaR value forecasted at time t for time $t+1$, and p is the probability of observing loss larger than the VaR_{t+1}^{p} at time $t+1$. The violation in the hit sequence is equal to 0 if the loss at time $t+1$ does not exceed the forecasted VaR_{t+1}^{p}. Otherwise, the violations can be observed as 1 if the forecasted VaR is violated.

The Christoffersen mix test is a test that includes testing for both properties that are required for a good VaR model; (i) the unconditional coverage property, and (ii) independence property. To test the unconditional coverage property, the model needs to satisfy the relationship $\Pr(V_{t+1}=1)=p$. If the relationship is not satisfied, the proposed model does not satisfy the unconditional coverage property and is said to be an underestimate or overestimate model.

(i) Unconditional Coverage Likelihood Ratio Test

Unconditional coverage test utilizes a likelihood ratio test to check whether the expected violations, p is equal to the actual violations in the series. Hence, the tested hypothesis is,

$$H_0: p=\pi$$
$$H_1: p\neq\pi$$

The likelihood ratio test for the unconditional coverage test is defined as,

$$LR_{UC}=-2\ln\left(\frac{L(p)}{L(\pi)}\right)\sim\chi^2(1)$$

where the null hypothesis, $L(p)=(1-p)^{t_0}p^{t_1}$; p is the expected number of violations in the series, t_0 is the number of non-violations in the series, t_1 is the number of violations in the series. The alternative hypothesis, $L(\pi)=(1-\pi)^{t_0}\pi^{t_1}$; π is the actual probability the violations observed in the series and defined as $\pi=\frac{t_1}{T}$.

If the null hypothesis is failed to be rejected, then the conclusion is the number of actual violations, π is not statistically different than the expected number of violations, p. Hence, the model satisfied the unconditional coverage property.

(ii) Independence Coverage Likelihood Ratio Test

The independence coverage test also utilizes the likelihood ratio test to check if $\Pr_t(V_{t+1}=1)=p$. According to Christoffersen, the violations that occurred in the data series should be independent. The likelihood ratio test for the independence coverage test is defined as,

$$LR_{IND}=-2\ln\left(\frac{L(\pi)}{L(\sqcap_1)}\right)\sim\chi^2(1)$$

where the null hypothesis, $L(\pi)=(1-\pi)^{t_0}\pi^{t_1}$ is exactly defined as the alternative hypothesis in the unconditional coverage test. The alternative hypothesis, $L(\sqcap_1)=(1-\pi_{01})^{T_{00}}\pi_{01}^{T_{01}}(1-\pi_{11})^{T_{10}}\pi_{11}^{T_{11}}$; T_{00} is the number of observations of non-violation followed by non-violation, T_{01} is the number of observations of non-violations followed by violations, and accordingly. The estimates for π_{01} and π_{11} are defined as,

$$\hat{\pi}_{01}=\frac{T_{01}}{T_{00}+T_{01}},\quad \hat{\pi}_{11}=\frac{T_{11}}{T_{10}+T_{11}}$$

The Christoffersen mix test will then utilises these two likelihood ratio tests to produce a mixed test to validate the adequacy of the model according to the Christoffersen definition.

$$LR_{CM}=LR_{UC}+LR_{IND}\sim\chi^2(2)$$

2. Basel Backtest

Basel backtest, also known as The Basel Committee's Traffic Light Coverage Test is a test that was initially created by the Basel Committee for any financial institutions to test the risk measure. The assessment is done by taking the most recent 250 days of data. From the 250 days of data, the number of violations from the data is observed. For a one-day 99% VaR, a total of 2.5 violations is expected to be observed from the model.

Table 1. Basel Accord Penalty Zones

Basel Accord Penalty Zone	Number of Violations	Increase in scaling, k
Green Zone	0 1 2 3 4	0.00 0.00 0.00 0.00 0.00
Yellow Zone	5 6 7 8 9	0.40 0.50 0.65 0.75 0.85
Red Zone	10 or more	1.00

For a one-day 99% VaR, the Basel backtest allows up to 4 violations for the model to be accepted in the green region. According to the Basel Committee, the VaR model that fall in the green zone is a valid model and can be used by bank institutions to measure their capital requirement. However, those models that fall in the yellow zone are models that require supervision. If the model is to be used, the capital requirement is advised to increase the capital requirement in scaling factor k as suggested by the Basel Accord and shown in Table 1. Meanwhile, models that fall in the red zone will require a higher increasing factor as the model is inadequate to predict the risk. Nevertheless, the Basel backtest used in this study is not used to gain information on which model is adequate to predict the capital requirement, rather it is used to give information whether this model is adequate according to Basel Accord or otherwise.

RESULTS AND DISCUSSIONS

1. Data Background

The data used in this study is the day-to-day future price observations data of two main traded commodities in the oil market; West Texas Intermediate (WTI) that is traded on New York Mercantile Exchange (NYMEX), and Brent Crude Oil (Brent) that is traded on Intercontinental Exchange (ICE). All the data are sourced from Investing.com. The sampled period for the study extends from 2 January 2009 to 31 December 2018. The total sample period is divided into two parts: the period from 2 January to 31 December 2016 is the in-sample period and that from 1 January 2017 to 31 December 2018 is the out-of-sample period. The in-sample period will be used to estimate the VaR, while the out-of-sample period will be used to test the adequacy of the model for the estimated risk value.

2. Exploratory Data Analysis

Before begin analyzing, the raw daily observations underwent a preliminary treatment by taking the differences in the natural logarithms of the price as $R_t = \log(P_t) - \log(P_{t-1}) = \log\left(\frac{P_t}{P_{t-1}}\right)$, where P_t is the future price at time t, and P_{t-1} is the future price at time $t-1$.

The tests of normality table of Shapiro-Wilk for WTI and Brent are presented in Table 2. From the normality test table, the results show that the basic distribution of daily yield of WTI and Brent crude oil does not conform to normal distribution. The significant value under the Shapiro-Wilk column is 0.000 for both the WTI and Brent Log-Return which is less than any significance value. Hence, we have significant evidence to reject the null hypothesis and conclude that the distribution of daily returns for both WTI and Brent does not follow the normal distribution.

Table 2. Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
WTI Log-Return	.062	2578	.000	.957	2578	.000
Brent Log-Return	.068	2578	.000	.962	2578	.000

a. Lilliefors Significance Correction

3. Distribution of Exceedances and Parameter Estimation

The distribution of extreme returns is independent of the main distribution. The EVT approach models only the extreme tail portion of the distribution which generally has a GPD distribution. This study focuses on the EVT approach based on the GPD model with the POT method. When implementing EVT model to simulate the tail of distribution based on POT framework, it is an important aspect to determine the appropriate threshold level. Pickands-Balkema-de Haan Theorem states that the distribution of excesses over a high threshold converges to GPD. In order to determine the optimal threshold, we have implemented the single bootstrap procedure proposed in Hall (1990) for selecting the optimal sample fraction in tail index estimation using the Hill estimator. The mean-excess plot is an alternative graphical approach to the quantile determination under which the GPD distribution is appropriate. After obtaining an appropriate threshold value for both WTI and Brent, the parameter estimation involving the shape parameter ξ and the scale parameter σ is carried out by using maximum likelihood estimation. The estimated $\hat{\xi}$ and $\hat{\sigma}$ along with their respective standard errors are presented in Table 3. The mean-excess plot at the upper tail for WTI and Brent is presented in Figure 1.

The number of excesses and estimated scale parameter for both datasets are quite similar despite the value for threshold being slightly different. However, the estimated shape parameters of WTI are much higher than Brent's, which means that WTI is heavier than Brent's at the tail distribution. By observing Figure 1, we can find that the thresholds of these two data sets are appropriate, because they are located at the stable area. As the increase of threshold, the mean excess becomes unstable, and there is an irregular trend until the end of the two data.

Table 3. Threshold value and parameter estimates

	WTI	**Brent**
Threshold, u	0.03936	0.03599
Number of excesses, k	81	86
Estimated shape, $\hat{\xi}$	0.11952	0.00340
Standard error of shape, $SE(\hat{\xi})$	0.14633	0.12377
Estimated scale, $\hat{\sigma}$	0.01671	0.01515
Standard error of scale, $SE(\hat{\sigma})$	0.00303	0.00245

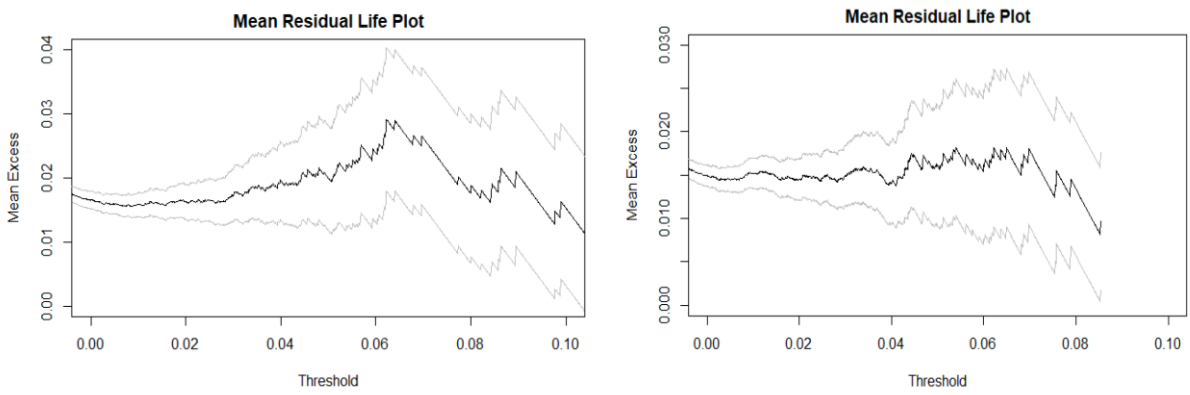

Figure 1. Mean excess plot for WTI (left) and Brent (right)

Extreme Value Theory (EVT) implies is that the distribution of exceedances is based on the Generalized Pareto Distribution (GPD). It is important to test the goodness-of-fit of the modelled distribution to see if it fits the distribution of the GPD as proposed by EVT. The diagnostic plot for both the WTI and Brent is presented in Figure 2 and Figure 3.

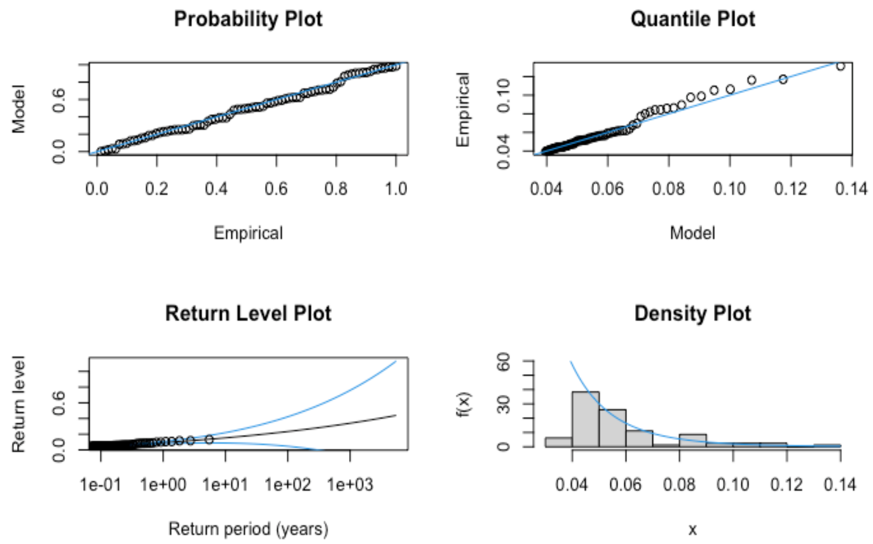

Figure 2. Diagnostic plots for WTI

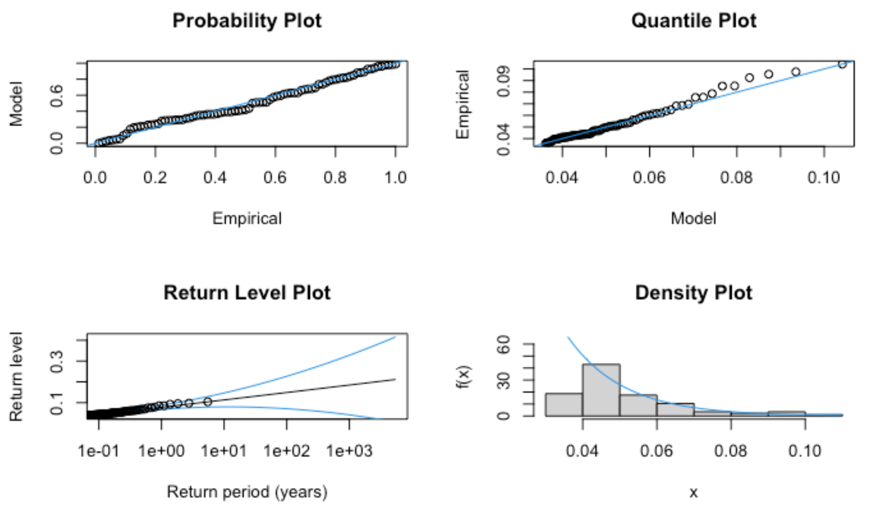

Figure 3. Diagnostic plots for Brent

In order for the model to be suitable for the GPD distribution, the data points are required to be approximately centred on a linear trend on the theoretical line for the probability plot, the quantile plot and the return level plot. For WTI, the data points show a reasonably linear pattern and nearly all data points fall on the theoretical line for all the diagnostic plots. Brent diagnostic plots also have the similar trend as WTI where nearly all the data points are on the theoretical line for all the diagnostic plots. This indicates that the GPD distribution fits the underlying distribution for exceedances above the selected threshold.

4. Risk Measure

The 1-day VaR estimates for the p-quantile or α significance level can be calculated after obtaining the estimated parameters $\hat{\xi}$ and $\hat{\sigma}$, and the appropriate threshold u. The 99% 1-day VaR estimate for both WTI and Brent is presented in Table 4. When measuring potential profit, VaR can be defined as the probability of profit exceeding the VaR over a holding period. This is particularly useful for day-traders who adopt day-to-day strategies to derive profit from market shifts for a given commodity. This study is interested in measuring the potential profit for WTI and Brent exceeding the VaR over the 1-day holding period.

Table 4 shows that the VaR estimate for WTI is 0.0392 at the 99th percentile for the right tail. This suggests that under normal market condition, we expect a daily change in the value of WTI futures contract that is traded on the NYMEX would not increase by more than 3.92%. To put it into perspective, the futures price, with 1% probability, would be expected to gain by USD2,744 or more in one day if the traded value is USD70,000 (smallest contract size of 1000 U.S. barrels with price per barrel of USD70) in the market. Meanwhile, the VaR estimate for Brent is 0.03587 at the 99th percentile. This implies that the largest profit in the market value will be projected to reach 3.59% or more. In other words, if the traded value is USD70,000 in the Brent on the ICE, there is a 1% probability that the profit will be USD2,513 or more during one trade day.

Table 4. VaR estimates for WTI and Brent

	WTI	Brent
Quantile	99%-level	99%-level
Threshold	0.03936	0.03599
Estimate	0.03920	0.03587

5. Backtesting Procedures

This study utilises two backtesting approaches to assess the model adequacy of the estimated VaR and whether it captures the risk accurately. The first test is the Christoffersen test that is used to test the unconditional and independent coverage property of the model. A VaR model is considered good when it fulfills the unconditional property where the actual violation is equal to the expected violations, and when it satisfies the independence property where the violations occurred in the series should be independent. The second test is the Basel backtest that is used to verify whether this model is adequate in compliance with the Basel Accord or otherwise.

Assuming 250 trading days in a year, the total out-of-sample continuous forecast period is 500 observations. With $\alpha = 0.01$, the expected violation hit for a correctly specified model is 5 violations in the series. To evaluate this, the Christoffersen test is applied to see if the produced hit series is statistically different from the expected violation. The output for the Christoffersen test is presented in Table 5.

Table 5. Christoffersen's test results

	WTI	Brent
Expected Violations	5	5
Actual Violations	4	4
H_0	Correct Exceedances	Correct Exceedances
p-value	0.6414	0.6414
Decision	Fail to reject H_0	Fail to reject H_0
H_0	Correct Exceedances & Independent	Correct Exceedances & Independent
p-value	0.8687	0.0585
Decision	Fail to reject H_0	Fail to reject H_0

Table 5 shows that the estimated VaR model for WTI and Brent captures the risk accurately. For WTI, there were 4 violations in total, which was lower than the expected number of violation. The p-value for both the unconditional and independence test is larger than the critical level of $\alpha = 0.01$. This suggests that the estimated VaR model for WTI has correct exceedances and independent. For Brent, a total of 4 violations can be seen in the series. Similar to WTI, the p-value for Brent for the unconditional and independence test is larger than the critical larger. Hence, the estimated model for Brent is said to be unconditional and independent.

The Basel Test only evaluates the last 250 trading days. Therefore, it is estimated that there will be 2.5 violations in this series of comparisons. However, according to the Basel Accord, a maximum of 4 violations are allowed for the model to be accepted in the green region. The values of these tests are presented in Table 6. According to Table 6, the estimated risk value model of WTI and Brent crude oil accurately capture risks. This can be seen where the actual violations for both WTI and Brent Crude oil are equal and lower than the maximum allowed violations. Therefore, according to Basel Accord, this model is considered acceptable.

Table 6. Basel backtest results

	WTI	Brent
Allowed violations	4	4
Actual violations	3	4
Zone	Green	Green

CONCLUSION

As the price of the oil market fluctuates greatly, the risk of investing in the oil market is high. This research implements the EVT to model the tail-related risk by applying it to the daily returns of the WTI and Brent futures contracts traded on NYMEX and ICE. This implementation effectively captures the fat-tailed behaviour that appears in the distribution of returns. POT method provides a convenient and direct means to determine threshold and estimate parameters. After modelling the distribution that exceeds the selected threshold, we evaluate the adequacy of tail modelling through backtesting. In conclusion, the results show that the estimated risk measures for WTI and Brent crude oil are valid. This implies that the modelling approach with just extreme values is accurate in capturing the risk adequately and can also be used to measure the risk of other extreme events. This Extreme-VaR approach used in this research provides quantitative statistics on the degree of risk that exists in the commodity market, especially in the crude oil market. These strategies have been proved to be effective, and any company or independent traders can apply this method to their portfolio risk management techniques. Meanwhile, traders and investors who have capital on the crude oil market, especially WTI and Brent crude oil, can use these findings as their benchmark for decision-making whether to hold or sell the contract.

ACKNOWLEDGEMENT

We acknowledge the financial support from the Ministry of Education Malaysia for the fundamental research grant scheme (FRGS/1/2019/STG06/UPM/02/7).

REFERENCES

Abad, P., Benito, S. & Lopez, C. (2014). A Comprehensive Review of Value at Risk Methodologies. *The Spanish Review of Financial Econometrics*, 12(1), 15-32.

Beirlant, J., Vynckier, P. & Teugels, J. L. (1996). Tail Index Estimation, Pareto Quantile Plots Regression Diagnostics. *Journal of the American Statistical Association*, 91(436), 1659-1667.

Bensalah, Y. (2000). Steps in Applying Extreme Value Theory to Finance: A Review. *Staff Working Papers 00-20*, Bank of Canada.

BP. (2019). *Statistical Review of World Energy 2019 68th edition*. UK: London.

Brown, A. (2008). Counterpoint: Capital inadequacy. *Global Association of Risk Professionals*, vol. 42, 19-26.

Campbell, S. D. (2005). A Review of Backtesting and Backtesting Procedures. *Finance and Economics Discussion Series*, 200521.

Christoffersen, P. F. (1998). Evaluating Interval Forecasts. *International Economic Review*, 39, 841-862.

Christoffersen, P. F. & Diebold, F. X. (2000). How Relevant is Volatility Forecasting for Financial Risk Management? *Review of Economics and Statistics*, 82, 1-11.

Danielsson, J., Hartmann, P. & de Vries, C. G. (1998). The Cost of Conservatism. *Risk*, 11, 101-103.

EIA. (2022). *Oil and petroleum products explained*. Retrieved July 20, 2022 from https://www.eia.gov/energyexplained/oil-and-petroleum-products/

Embrechts, P., Klüppelberg, C. & Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer, Berlin.

Gencay, R. & Selcuk, F. (2004). Extreme value theory and value-at-risk: relative performance in emerging markets. *International Journal of Forecasting*, 20(2), 287-303.

Halilbegovic, S., & Vehabovic, M. (2016). Backtesting Value at Risk Forecast: The Case of Kupiec Pof-Test. *European Journal of Economic Studies*, 17, 393-404.

Hall, P. G. (1990). Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of Multivariate Analysis*, 32, 177-203.

Jondeau, E., & Rockinger, M. (1999). The tail behavior of stock returns: Emerging versus mature markets. *Documents de Travail*, 66: Banque de France.

McNeil, A. J. (1999). Extreme Value Theory for Risk Managers *Internal Modelling and CAD* (Vol. II, pp. 93-113): RISK Books.

Neftci, S. N. (2000). Value at risk calculations, extreme events, and tail estimation. *Journal of Derivatives*, 23-37.

Odening, M. & Hinrichs, J. (2002). Using extreme value theory to estimate value-at-risk. *Agricultural Finance Review*, 63(1), 55-73.