STATISTICAL MODELLING OF MALAYSIA TRADING GOLD PRICE USING EXTREME VALUE THEORY APPROACH

N. ALI[1], N.N. ZAIMI, AND N. MOHAMED ALI

ABSTRACT. The monthly maxima of negative log returns of Malaysia gold prices are modelled using Generalized Extreme Value distribution from Extreme Value Theory. The maximum likelihood estimation and L-moments methods are used to estimate the parameters of the GEV model. The extreme risk measure through Value-at-risk for 10%, 5%, and 1% upper quantile is calculated in order to evaluate maximum losses in investing in a gold market. The assessment of the quantile-quantile plot showed that the GEV model fit the data well using the maximum likelihood estimation. The estimates of VaR showed that for the worst case scenario, the expected market value of gold would not lose more than 0.009424%, 0.014211% and 0.028159% with 90%, 95% and 99% confidence respectively.

1. INTRODUCTION

Gold is a well-known valuable metal for investment compared to silver, platinum, and palladium. Gold has been widely used in many industries not only in the manufacturing of jewellery but also in financing and investing, medical and dentistry, manufacture of electronics, components of computers and aerospace. In the financing and investing industry, gold is held in the form of gold bullion. Gold bullion is a valuable precious metal that comes in two main forms which are gold bullion bars and gold bullion coins. Nowadays, many institutions including

---

in Malaysia hold investment of gold in the convenient form of bullion due to its stable value. Gold is a major asset of world investment due to its stability which plays an important role in the global economy. Gold is a good investment, neither in the short-term nor the long-term and it is a great asset that can be converted into paper money.

Malaysia's official gold bullion coin known as Kijang Emas has been launched by the former prime minister of Malaysia, Tun Dr. Mahathir Mohamad on 17th July 2001. Malaysia becomes the 12th country in the world to have its own bullion coins. The design of Malaysia's gold bullion coin is a barking deer (kijang) and the national flower of Malaysia, hibiscus on its reverse side. The price of Kijang Emas is determined by the world gold market price. Gold bullion coins are traded daily throughout the world and their price depends on the world gold price.

The gold price return is the main concern in the finance area of study because the investors are paying attention to profit or losses rather than looking at the price itself. Identifying the gold returns distribution used in financial forecasting models depicts great interest among researchers. There is empirical evidence that the distribution of returns possesses fat or heavy tails [1]. [2] suggested that the Student's t-distribution which has heavier tails than normal is adequate for modelling the returns distribution of the gold index. Other distribution that mostly considered in financial modelling includes the generalized extreme value (GEV) and generalized Pareto (GP) distributions from extreme value theory (EVT). See among others [3], [4] and [5].

A method of maximum likelihood estimation (MLE) and L-moments (LMOM) for parameter estimation has been popularly applied to the GEV model in numerous literature. In the financial field, the MLE has been used in most of the studies [6]. A notable exception is [7], who used probability weighted moments to estimate the parameters of the limiting distribution of extremes in Asian stock markets.

The volatility of the gold price can be measured by the daily returns of the gold prices. Losses or negative returns are the main concerns in the field of financial risk management. Value-at-Risk (VaR) remain popular risk measurement technique in financial among practitioners [8]. VaR can provide some useful information to the investor regarding the prediction of a future risk so that they can control and reduce the investment risk. VaR calculations in the extreme situation can be calculated using the method of EVT.

This study aims to establish if the monthly maxima of negative log return of Malaysia gold price data follow a heavy-tailed distribution. GEV distribution will be applied to model the data. The second objective is to identify the best estimation method for the GEV model using MLE and LMOM before the future risk in investing in gold is assessed through the VaR calculations.

The rest of the paper is organized as follows: Section 2 describes the methodology for data analysis. Section 3 presents the results and discussions. Section 4 gives a few concluding remarks.

2. METHODOLOGY

2.1. **Model Formulation.**

The movement in gold prices can be measured by the daily log returns of the gold prices. Losses or negative log returns are a focused in this study as it is very relevant to a risk measure of financial markets through a VaR calculations. Let R_i be the negative log return of the gold prices between day i and day $i-1$. Define

$$(2.1) \qquad R_i = -(\ln P_i - \ln P_{i-1}),$$

where P_i and P_{i-1} are the gold prices on day i and day $i-1$. Let $R_1, R_2, \ldots, R_n$ be a sequence of independent and identically distributed random variable with distribution function $F(r) = \Pr(R_i \leq r)$. Extreme values are defined as maximum (or minimum) of the n independent and identically distributed random variables $R_1, R_2, \ldots, R_n$. Let maximum negative side movements in the daily log returns of gold prices as $R_{(n)} = \max\{R_1, R_2, \ldots, R_n\}$. The distribution of $R_{(n)}$ can be derived as

$$\Pr\{R_{(n)} \leq x\} = \Pr\{R_1 \leq x, R_2 \leq x, \ldots, R_n \leq x\}$$

$$= \prod_{i=1}^{n} \Pr\{R_i \leq x\}$$

$$= F^n(x).$$

In practice, the parent distribution F is unknown. However, the limiting distribution of $F^n(x)$ can be approximated as $n \to \infty$. According to EVT, the distribution of normalized $R_{(n)}$ is given as such

$$Z_{(n)} = \frac{R_{(n)} - b_n}{a_n}$$

for sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ converge to a non-degenerate distribution function G, as $n \to \infty$ where G is given by

$$(2.2) \qquad G(z) = \exp\left\{ -\left[1 + \xi\left(\frac{z - \mu}{\sigma}\right)\right]^{-1/\xi} \right\},$$

defined on $\left\{ z : 1 + \xi\left(\frac{z-\mu}{\sigma}\right) > 0 \right\}, -\infty < \mu < \infty, \sigma > 0$ and $-\infty < \xi < \infty$. The distribution in (2.2) is known as GEV.

2.2. **Parameter Estimation.**

There are many techniques proposed for parameter estimation in the GEV model. This study used the MLE and LMOM methods to estimate the parameters of the GEV model. Let $Z_1, Z_2, \ldots, Z_n$ are the block maxima that are assumed to be independent variables having the GEV distribution. A random variable Z has GEV distribution function for case where $\xi \neq 0$ as stated in Equation (2.2). The probability distribution function of random variable Z is given by:

$$g(z) = \frac{1}{\sigma}\left[1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi - 1} \exp\left[-\left\{ 1 + \xi\left(\frac{z-\mu}{\sigma}\right)\right\}^{-1/\xi} \right],$$

where $1 + \xi\left(\frac{z-\mu}{\sigma}\right) > 0$. The likelihood function based on data $Z = (Z_1, \ldots, Z_m)$ is given by:

$$\ell(\mu, \sigma, \xi) = -m \, log \, \sigma - \left(1 + \frac{1}{\xi}\right) \, \log \sum_{i=1}^{m}\left[1 + \xi\left(\frac{z_i - \mu}{\sigma}\right)\right]$$

$$(2.3) \qquad\qquad - \sum_{i=1}^{m}\left[1 + \xi\left(\frac{z_i - \mu}{\sigma}\right)\right]^{-1/\xi},$$

provided that $1 + \xi\left(\frac{z_i-\mu}{\sigma}\right) > 0$, for $i = 1, \ldots, m$. Solving the partial derivatives of the Equation (2.3) with respect to μ, σ and ξ yields the parameter estimates of the GEV model.

The method of LMOM is based on a linear combination of order statistics where the first until the fourth order statistics correspond to measures of location, scale, skewness and kurtosis respectively. The rth LMOM, denoted as λ_r is defined as

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} E(Z_{r-k:r}); \qquad r = 1, 2, \ldots,$$

where $Z_{r-k:r}$ is the random variable for $(r-k)$th order statistics.

The assessment on the uncertainty due to the parameter estimates for both methods is conducted using the parametric bootstrap method [9]. In the parametric bootstrap method, the inference about a population from sample data can be modelled by resampling the sample data and performing inference about a sample from resampled data. Sorting the bootstrap means from the lowest to the highest and lowering the 2.5% smallest and 2.5% largest for lower and upper confidence would then form a 95% confidence interval for the mean population.

2.3. **Estimation of Value-at-Risk.**

Some of the common question concerning risk management in finance is extreme quantile estimation. This corresponds to the determination of the value a random variable exceeds with a given probability p. This value, known as VaR, remains to be a popular financial risk measure among practitioners. VaR measures the maximum loss in value of assets over a predetermined time for a given probability p. The estimated VaR measure for a financial risk with negative returns R_i is [10]

$$(2.4) \qquad \widehat{\text{VaR}}_p = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\xi}} \left\{ 1 - [-n\log(1-p)]^{-\hat{\xi}} \right\},$$

where n is the block size and $\{\hat{\mu}, \hat{\sigma}, \hat{\xi}\}$ are the estimated parameters of GEV model.

3. Results and Discussion

This study used one ounce the daily trading gold price of Malaysia's gold bullion coin, Kijang Emas. The trading prices of Kijang Emas (in Ringgit Malaysia) for the years 2002 until 2019 is obtained from Bank Negara Malaysia (BNM) official website www.bnm.gov.my. However, there is no record of trading prices for certain days. This occurrence is common due to the holidays or market closing day.

The scatter plot of gold prices from the year 2002 until 2019 in Figure 1 (left) display a roughly exponential growth over the period. However, the gold price has extreme fluctuations during the year 2011 until the end of the year 2018. The extreme drop is due to frightened market associated with the September 11 terrorist attacks in the U.S. Since our interest is on losses or negative log returns, we focused the analysis on loss returns calculated using Equation (2.1). The time series plot of loss return is shown in Figure 1 (right). In Table 1 some descriptive statistics for the time series are summarized. The largest loss for gold is around

0.112 with a large value of kurtosis indicates the distribution of loss returns has a fat tail.
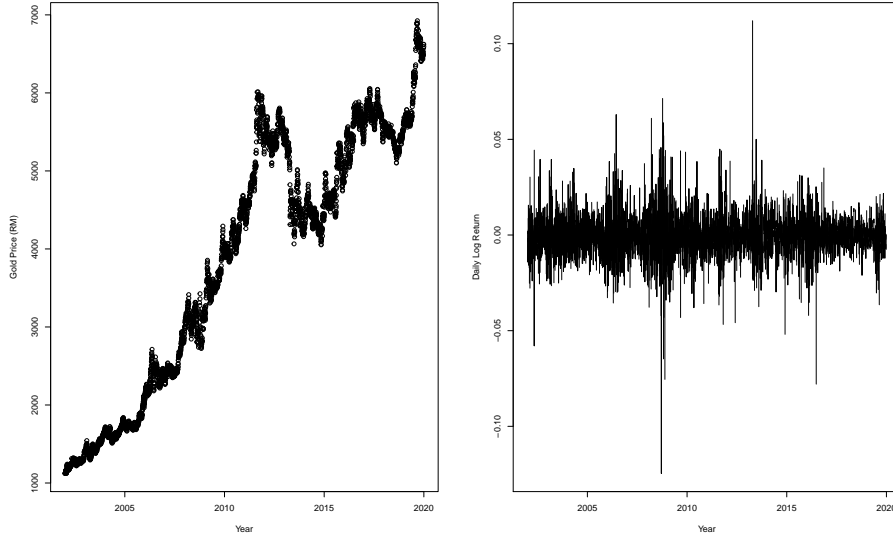

FIGURE 1. Scatter (left) and Time Series Plots (right) of Loss Returns

TABLE 1. Descriptive Statistics for the Loss Return Series

Mean	-0.0003957
Median	-0.0005310
Standard Deviation	0.01114853
Skewness	0.0892172
Kurtosis	13.24764
Minimum	-0.1246452
Maximum	0.1119943

For the analysis of extreme losses, we considered monthly maxima of negative log returns with 21 working days of the block size. The histogram in Figure 2 shows that the distribution of monthly maxima of negative log returns is not symmetric but rather skewed to the right. The decay in frequency on the right side of the histogram suggests that the monthly maxima of negative log returns have a
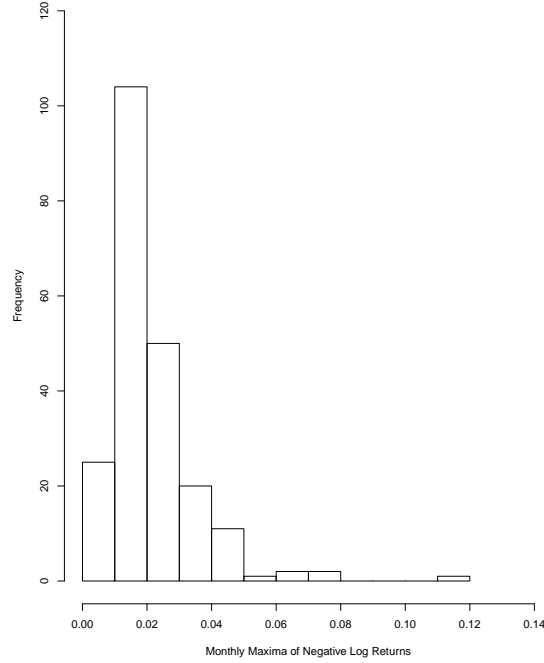
FIGURE 2. Histogram of Monthly Maxima of Negative Log Returns

heavy tails distribution. Since the distribution of data demonstrates a heavy tail, the model of capturing the behaviour in the tails is applied through the EVT approach. Here, we adopted the GEV model to explain the behaviour of those tail. We test for stationary of maximum series prior to the extreme modelling using Augmented Dicker-Fuller (ADF) and Phillips-Perron (PP) with 0.05 significance level. Results are reported in Table 2. The rejection of the null hypothesis of non-stationary using ADF and PP tests ($p-$value < 0.05) concluded that the monthly maxima of negative log return series are stationary.

TABLE 2. Stationary Test for Monthly Maxima of Negative Log Return Series

Stationary Test	p-value
ADF	0.0192
PP	0.0100

We further the analysis by fitting the GEV model to the monthly maxima of negative log return series. The parameter of the GEV model is estimated using MLE and LMOM methods. Table 3 gives the estimated parameter values using both MLE and LMOM methods. The performance of the methods of estimation is compared using the width of confidence interval (CI) of each estimated parameter values. It is found that the MLE is superior to the LMOM since it has narrower CI indicating less uncertainty due to the parameter estimates. Looking at the quantile-quantile (Q-Q) plot in Figure 3 gives cause no doubt the validity of the fitted GEV model to the monthly maxima of negative log returns where the plotted points are near-linear.

TABLE 3. Parameter Estimates of GEV Model Fitted to the Monthly Maxima of Negative Log Returns

Method	Parameter	Estimated Values (CI)	Width of CI
MLE	μ	0.014748 (0.013732, 0.015785)	0.002053
	σ	0.007282 (0.006431, 0.008202)	0.001771
	ξ	0.211220 (0.104309, 0.323892)	0.219583
LMOM	μ	0.014676 (0.013599, 0.015783)	0.002184
	σ	0.007188 (0.006258, 0.008083)	0.001825
	ξ	0.224131 (0.095081, 0.350182)	0.255101

TABLE 4. Estimates of VaR on Monthly Block

Quantiles Level	VaR
90%	0.009424
95%	0.014211
99%	0.028159

Table 4 presents the estimates of VaR at several quantiles level. At upper 10% quantile, the estimated VaR from GEV is 0.009424. It means that we are 90% confidence that the expected maximum loss due to investing in a gold market would not more than 0.009424% for the worse case scenario, within one month duration. Suppose an investor invests in gold with RM1000000, then the maximum
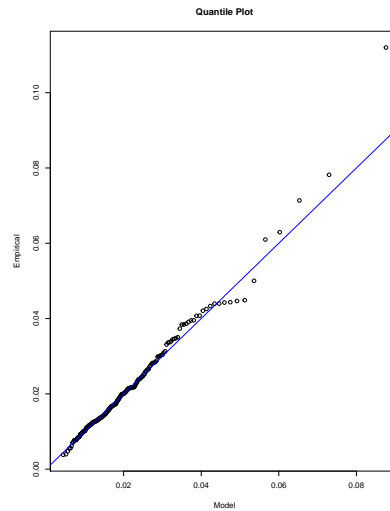
FIGURE 3. Q-Q Plot of Monthly Maxima of Negative Log Returns

loss that may be obtained in the next day is RM94.24. If the probability is 0.05, then the corresponding VaR is RM142.11.

4. CONCLUSION

In this study, the application of EVT in the field of finance has been illustrated to gold price data from the year 2002 until 2019. GEV distribution is used to model the monthly maxima of negative log returns of gold prices. The assessment of the Q-Q plot suggests that GEV distribution provides a good fit for the series of data using the MLE method. In the gold market, the stability of gold prices as the time increase has captured the investors' attention and become a major asset in terms of investment all over the world. The model and risk measure through the VaR calculation provided in this study gives some useful information to the investors or regulators who hold gold as an asset or a hedge against financial market risks. Physical gold bars and coins is a good investment since it is very liquid and easy to buy in one place and sell in another.

REFERENCES

[1] F.M. APARICIO, J. ESTRADA: *Empirical distributions of stock returns: European securities markets*, European Journal of Finance, **7**(1) (2001), 1–21.

[2] S.N. KUMARI, A. TAN: *Asset return distribution and its application to gold future index*, Journal of Mathematics, Statistics and Operations Research, **2**(1) (2013), 72–76.

[3] S.N. NEFTCI: *Value at risk calculations, extreme events, and tail estimation*, The Journal of Derivatives, **7**(3) (2000), 23–37.

[4] Q.L. CHEN: *The Volatility of the Price of Gold: An Application of Extreme Value Theory*, PhD thesis, University of Victoria, 2014.

[5] C.K. HUANG, C.S. HUANG, D. CHIKOBVU: *Extreme risk, value-at-risk and expected shortfall in the gold market*, The International Business & Economics Research Journal, **14**(1) (2015), 107–122.

[6] K. TOLIKAS, A. KOULAKIOTIS, R.A. BROWN: *Extreme risk and value-at-risk in the German stock market*, European Journal of Finance, **13**(4) (2007), 373–395.

[7] A.L. DA SILVA, B.V. DE MELO MENDES: *Value-at-risk and extreme returns in Asian stock markets*, International Journal of Business, **8**(1) (2003), 17–40.

[8] K. CHINHAMU, C.K. HUANG, C.S. HUANG, D. CHIKOBVU: *Extreme risk, value-at-risk and expected shortfall in the gold market*, International Business & Economics Research Journal (IBER), **14**(1) (2015), 107–122.

[9] A.C. DAVISON, D.V. HINKLEY: *Bootstrap Methods and their Application*, University press, Cambridge, 1997.

[10] R.S. TSAY: *Analysis of Financial Time Series, Third Edition*, John Wiley & Sons Inc, New Jersey, 2010.

DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE, UNIVERSITY PUTRA MALAYSIA
43400 UPM SERDANG, SELANGOR, MALAYSIA.
Email address: norhaslinda@upm.edu.my

DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE, UNIVERSITY PUTRA MALAYSIA
43400 UPM SERDANG, SELANGOR, MALAYSIA.
Email address: nadhirahzaimi@gmail.com

DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE, UNIVERSITY PUTRA MALAYSIA
43400 UPM SERDANG, SELANGOR, MALAYSIA.
Email address: nazihahma@upm.edu.my