Predicting Extreme Returns of Bitcoin: Extreme Value Theory Approach

Saiful Izzuan Hussain¹, Nurulkamal Masseran², Nadiah Ruza³ and Muhammad Aslam Mohd Safari⁴

¹,²,³ Department of Mathematical Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia
⁴Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, Serdang 43400, Selangor, Malaysia

1E-mail: sih@ukm.edu.my

Abstract. Extreme value theory (EVT) has been used to study the frequency and probability related to extreme situations in finance. This approach focuses on the extreme values and able to provide a better estimation for risk models. In this study, Generalized Pareto Distribution (GPD) is employed to model daily extreme returns in the Bitcoin market from 2017 to 2019. These periods have witnessed three phases of extreme volatility for the cryptocurrency market. The returns level for the Bitcoin range between 17.011 and 18.746. The results demonstrate heavy tail and finite tail distribution characteristics for the tails. The findings provide a better understanding of the tails' behaviour in the cryptocurrency market and help investors to make a financial decision.

Keywords: Bitcoin, Extreme Value Theory, cryptocurrency

1. Introduction

Many finance literature is focused on the premise of the normal distribution framework [1-3]. While the use of standard distribution assumptions to assess risks is acceptable in most cases, it is quite ineffective in severe situations. Financial data is showed tend to follow right-skewed or left-skewed distributions in most cases. A study in [4] shows that the tail of the market possesses useful information to understand the extreme volatility movement. This brings the application of Extreme Value Theory (EVT) method to describe extreme characteristics since it focuses more on extreme data. EVT is widely used in different fields. This approach is first introduced by [5] and it was applied in finance by [6]. Fat tails on the distribution of returns to the market are being used to determine the likelihood of a market crash. This could provide insightful information to the financial institution because it could help to forecast the state of the economy. Several recent studies about EVT could be found in [7-10].

The cryptocurrency industry is booming. According to 20 February 2021 market capitalization info, there are around USD860 billion in total market capitalization [11]. These numbers have caused an increase of interest in the public, media, investment and government. For this study, we will
concentrate on Bitcoin, which is the most common and the largest cryptocurrency by market capitalization and trading volume. Bitcoin is the first and currently largest cryptocurrency. It was launched in 2009 via a proposed protocol developed by [12]. As a percentage, approximately 29% of the market capitalization is accounted for by trade volume and market capitalization for Bitcoin. Trade in bitcoin started on the Mt. Gox exchange in 2010 and it was normal to see bitcoin exchanged in various exchanges across the world after that.

Over the past few years, the market for Bitcoin has seen tremendous growth. Bitcoin has seen an extreme upward movement as a USD 1,000 investment in July 2010 for 7 years could have generated over USD 81 million [13]. Other people might lose confidence in the major currencies and choose Bitcoin as an alternative. Bitcoin does not have a central authority and there are no intrinsic attributes that set it apart from other digital currencies. There is a finite supply of Bitcoin and it cannot be recreated. All of this runs on a state-of-the-art protocol, thus being completely decentralized.

This research is motivated by the extreme behaviour found in Bitcoin’s price movements. Based on [14], Bitcoin seems to face scrutiny as being speculative. According to [15], the Bitcoin market is still in its infancy and inefficient. The previous study shows the empirical data tend to exhibit heavy tail, long memory, dynamic volatility and leverage. This paper seeks to remedy this by expanding our study to cover the period during 2017 and 2019 which consider a very volatile period for cryptocurrency. This paper attempts to understand the behaviour of extreme returns of Bitcoin.

It's speculated that Bitcoin might be a good match for conservative investors and asset diversification during volatile times. The results of a study conducted by [16] indicate that Bitcoin can be used to hedge the value of the US dollar for the short term. It has been found that Bitcoin is uncorrelated with other financial assets in general, during both a financial crisis and regular time [17]. According to the results of the report, the financial characteristics of Bitcoin are entirely different from traditional financial assets. Additional research that details Bitcoin's behaviours and predictability can be found in [18] and [19].

The characteristics of Bitcoin under the extreme condition are not yet well described and still has a gap to explore [20]. The assessment of risks and their return period is important for the risk management process to assess and monitor the possible negative implications of extreme phenomena. Until now, not many extensive works have been performed on the extreme volatility of Bitcoin. Most empirical studies in this area have concentrated on modelling and estimating Bitcoin's general volatility. Existing accounts tend to concentrate on the entire dataset rather than the unusual events at the extremes and only a few have covered Bitcoin's tail [21].

Another problem in risk management is comprehending the extreme tail dependency in Bitcoin. This research assesses the extreme returns in Bitcoin and understands the negative returns of Bitcoin's. Taking a central role in recognizing and understanding risk situations is a great aid. This research aims to discover and understand the traits of severe returns in Bitcoin, especially using the application of EVT. Generalized Pareto Distribution (GPD) is used to suit the values of returns that exceed a given threshold in the current data. We gave particular attention to the severe characteristics to better understand these characteristics. In addition, the return-level probabilities for the extreme returns are determined based on the fitted GPD model. The probability of the return level can be used to deduce the probability of unpredictable returns in Bitcoin at such events.

2. Extreme Value Theory

The EVT has two methods to describe the extreme data. The first method is related to POT and the other is the block maxima method (BMM). Extreme events are characterised as the maximum
(minimum) value within each sub-period or block under the block maxima method. This sub-period may be part of a sample of daily, weekly, or annual. However, the BMM approach has its limits. BMM disregards an enormous amount of data that might occur at the same period calculation and chooses only the highest or lowest value for that study. In contrast, the POT method is focused on sorting clustered phenomena based on a threshold, which can give weight to information that is commonly contained in data. The POT method is used to accommodate the observed excesses over the threshold \( v \) using GPD distribution. This research will concentrate on the POT approach to model the extreme returns for Bitcoin.

### 2.1 Generalize Pareto Distribution (GPD) model

For this paper, the returns for the Bitcoin that exceed the threshold value at the \( u = 90\% \) percentile of negative returns is used. The returns that are larger than a value of \( u \) consider as extreme data for Bitcoin’s series. Let \( x = [x_1, x_2, \ldots, x_n] \) be the vector of data of the returns. The equation to describe the conditional exceedance density function \( F^\circ \) can be shown as

\[
F^\circ(x) = P(X \leq x | X > u) = \frac{F(x) - F(u)}{1 - F(u)}, \quad x \geq u
\]  

(1)

For this case, \( x_i > u \) are the exceedances of \( y_i \) over the threshold \( u \). Next, the equation distribution function \( \hat{F}_y(y) \) based on the vector data \( y = (y_1, y_2, \ldots, y_k) \) can be written as below

\[
\hat{F}_y(y) = \left(\hat{F}_y(y)\right)^{|y|}
\]  

(2)

Equation (2) suggests equivalent knowledge of the exceedance distribution function, a report in [22]. This article utilizes the Generalized Pareto Distribution which known as one of the parametric methods to understand and explain the distribution of the exceedances of extreme values at the 90\% percentile.

### 2.2 GPD parameters estimation

Study in [23,24] explains the GPD is a limiting distribution of the standardized excesses over a threshold, as the threshold approaches the endpoint of the variable. The event of exceedance over the threshold is explained using the random variable as \( Y = X - u \) for a given random variable \( X \). The distribution function for GPD is implied as in Equation (3).

\[
F(x) = P(X \leq x | X > u) = \begin{cases} 1 - \left(1 + \frac{x - u}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - e^{-\left(\frac{x - u}{\sigma}\right)} & \text{if } \xi = 0 \end{cases}
\]  

(3)

This can be shown as,
\[ G(y) = p(Y \leq y) = \begin{cases} 
1 - \left(1 + \frac{\xi y}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\
1 - e^{-\frac{y}{\sigma}} & \text{if } \xi = 0 
\end{cases} \] (4)

where \( y \geq 0, \ 1 + \frac{\xi y}{\sigma} > 0, \) and \( \sigma \) are the scale parameters, and \( \xi \) is the shape parameter [25]. The finding of GPD may be related to the distribution known as the exponential, which exists in the form of the medium-sized tail distribution properties in the results. The shape of a normal Pareto is also considered to have the following properties where parameter \( \xi = 0 \). GPD has associated a Pareto type-II model if the parameter \( \xi < 0 \) is related to long tail behaviour. GPD will follow short tail distribution if \( \xi > 0 \).

Several studies indicate that a method exists to estimate GPD parameters, using various approaches [26-28]. According to [29], the most recently proposed estimator is based on the maximum probability and goodness of fit method. The set of parameters for GPD is \((\theta, \xi)\) by assuming \( \xi = \frac{\tilde{\xi}}{\sigma} \). The log-likelihood function for the \( y_1, y_2, \ldots, y_k \) can be described using equation (5).

\[
\log(L) = \sum_{i=1}^{k} \left( \log\left(\frac{\theta}{\xi}\right) - \left(1 + \frac{1}{\xi}\right) \log(1 + \theta y_i) \right)
\] (5)

Equation (5) is applied to estimate the parameters involve according to the study in [29]:

\[
\frac{\sum_{i=1}^{k} \hat{H}(y_i)}{k} = \frac{1}{2}
\] (6)

wherewith the condition that \( \hat{H}(y_i) \) that known distribution distributed on \([0, 1]\). Next, the estimated parameters \((\hat{\theta}, \hat{\xi})\) are referred to as below.

\[
\hat{\xi} = \frac{\sum_{i=1}^{k} \log\left(1 + \hat{\theta} y_i\right)}{k}
\] (7)

To measure the parameters for \((\theta, \xi)\), we refer to Equation (8) where \( \hat{\sigma} = \frac{\tilde{\xi}}{\bar{\theta}} \).

\[
\frac{\sum_{i=1}^{k} \left(1 + \hat{\theta} y_i\right)^{-\frac{1}{\hat{\xi}}}}{k} = \frac{1}{2}
\] (8)

The process of estimators \((\hat{\theta}, \hat{\xi})\) can be quantified using the iterative procedure based on the criteria of Equations (7) and (8). The value of suitable threshold \( \theta \) is a critical factor as it is considering a bias-variance trade-off.

2.3 Estimation of Return Period and Return Level

For this case, a return period is related to a recurrence interval, that is used to estimate the likelihood that future returns that are greater than the 90\text{th} percentile will happen. Average recurrence interval over a long period can be used to determine this value. This provides investors and risk managers with
an idea of how long real returns data will last if it does not reach a percentile. In a sense, the return duration is used to characterize the potential for extreme return levels in the future.

The return period for an extreme event is determined by using the annual maximum series and a single random variable. The distribution of the random variables will affect the value of the return period based on the formula [30].

\[
P(X > x | X > u) = \left[ 1 + \xi \left( \frac{x - u}{\sigma} \right) \right]^{\frac{1}{\xi}}
\]  

which can be written as,

\[
P(X > x) = \xi_u \left[ 1 + \xi \left( \frac{x - u}{\sigma} \right) \right]^{\frac{1}{\xi}}
\]  

where \( \xi = p(x > u) = \frac{k}{n} \) for \( k \) amount of data with the excess over threshold \((y_1, y_2, \ldots, y_k)\). Thus, the level returns \( (x_u) \) exceeding the \( u = 90\% \) percentile on average one every \( m \) observation is the solution of,

\[
\frac{1}{m} = \xi_u \left[ 1 + \xi \left( \frac{x_u - u}{\sigma} \right) \right]^{\frac{1}{\xi}}
\]  

This can be simplified as,

\[
x_{u=90} = u + \frac{\sigma}{\xi} \left( (m\xi_u)^{\frac{1}{\xi}} - 1 \right)
\]  

The returns concentration at \( u = 90\% \), the return level based on the return period can be described as,

\[
P_x(x_u) = \frac{1}{1 - G(y)}
\]  

where \( P_x(x_u) \) is the return period of a returns values more than \( u > 90\% \) percentile. This information could translate that the possible dataset will not more than the optimal threshold during a time series.

3. Data and Descriptive Statistics

This paper used daily data from January 1, 2017 to December 31, 2019 with a total of 1095 daily data in the US Dollar (USD). The equation of daily return for the Bitcoins is calculated as \( R_t = \ln(P_t / P_{t-1}) \), where \( R_t \) is the return on the Bitcoin for period \( t \), \( P_t \) is the Bitcoin at the end of period \( t \), and \( P_{t-1} \) is the Bitcoin at the end of the period \( t-1 \).

| Table 1. Bitcoin Daily Returns |
|-----------------------------|
| **Mean**      | **Std** | **Skew** | **Kurtosis** | **Max** | **Min** |
| **BTC** | 0.1808 | 4.3253 | -0.0580 | 3.0611 | 22.7602 | -18.6939 |

Table 1 depicts the daily rate returns for Bitcoin. Some interesting conclusions can be drawn about this table. The mean is positive with a value of 0.1808. Bitcoin's lowest daily return is -18.6939% and the Bitcoin's highest daily return is 22.7602%. The daily return values for Bitcoin are extremely distorted (-0.0508 for skewness, 3.0611 for kurtosis). Overall, this insightful description demonstrates that daily return is the potential to display fat-tailed behaviour.
Our aim in this study is to examine the movement of negative returns, which is related to the downside risk. To make it easier to understand, we transform the negative returns to positive ones. This study will use daily returns values more than the 90th percentile of the sample to assess extreme negative returns. 90th percentile daily returns for the study would result in an intense movement in Bitcoin.

4. Results and Discussion

This analysis used the POT model which is based on the GPD model to classify the extreme characteristics of the Bitcoin series. Using the quantile method, daily returns above 90% quantile are characterized as a dataset of extreme negative movement. However, the cut-off points for GPD need to be investigated properly. When choosing an appropriate threshold, there is a balance between two major criteria that need to be considered.

To ensure that the underlying GPD approximation is accurate, the asymptotic EVT method must have a reasonably high threshold. Nevertheless, the high threshold could serve to restrict the number of observations, raise the predicted parameter's variance, and suggest a large standard error. As the variance decreases, the asymptotic approximation occurs less often. Choosing a threshold involves balancing for both bias and accuracy trade-offs. There should be a rational relation between the threshold setting and the study problem. Any research outcomes would be of little use if the selection of the threshold is either too high or too low.

There is no objective way to pick the optimal threshold value. Optimal thresholds can be more than one, depending on how the tails behave. There is a complex trade-off to be made here between a high warning threshold and the number of events. Setting a high threshold would ensure that the GPD approximation is preserved but the total number of events will be too small to be used for parameter estimation. The threshold is set at the suitable quantile to make sure enough data gained. Thus, we must determine whether this attribute is capable of being used in POT analysis when dealing with the dataset. As stated before, if the selection threshold of the quantile is calculated to be high, the results of POT are not likely to describe informative information. We would lose interest in the analysis if the percentile threshold were too low since much of the data could only be viewed as a case of exceedance and cannot be linked to extreme movements. As a result, the POT analysis could introduce the potential of bias.

The threshold is selected using the mean of the excess and the mean residual life plot (MRL). Based on the MRL approach, a threshold $u$ is plotted against the mean of the excess over the threshold. As many studies indicate, the MRL plot should be linear above the threshold to get a better representation of GPD. The mean excess above the threshold is given by Equation (14).

$$
E[X - u | X > u] = M(u) = \frac{\sum_{i=1}^{\infty} (X_i - u) I_{[x_i > u]}}{\sum_{i=1}^{\infty} I_{[x_i > u]}}, \quad u \geq 0
$$

(14)

The MRL function can be described as,

$$
E[X - u | X > u] = \frac{\sigma}{1 - \xi} \cdot \frac{\xi u}{1 - \xi} - u
$$

(15)

where

$0 \leq u < \infty$ if $0 \leq \xi < 1$ and $0 \leq u < -\frac{\sigma}{\xi}$ if $\xi < 0$.
Figure 1. Mean residual life for Bitcoin

Figure 1 illustrates the MRL for the extreme returns in Bitcoin to valid the use of the threshold value at $u = 90$ quantile. The threshold value $u = 90$ quantile is well fitted to be used as a suitable model based on the POT analysis for Bitcoin. Assessing this type of data using the POT analysis is acceptable and will include details on that exceeding.

Figure 2. GPD model for the Bitcoin

Figure 2 describes the results of the GPD model, which was used to model returns for all Bitcoin cryptocurrencies. Furthermore, the figure depicts the density plot associated with the probability plot, showing that the GPD can produce an accurate model prediction for extreme data. Thus, it can be said that GPD offers a good approximate model.

The estimated return amounts and return times for the events are measured. The results from the plot of the return period versus the return level are visualised in figure 3. GPD's parameter estimates are shown in Table 2. All the returns in the sequence exhibit short-tailed properties, following a Pareto-type-II model.

| Currency | Shape ($\xi$) | Scale ($\sigma$) |
|----------|--------------|-----------------|

Table 2. Parameter estimates of GPD for the data of each currency return.
Figure 2 and table 2 demonstrate the parameter estimates based on GPD and depicts the frequency of extreme events at various magnitudes. Generally, when the shape parameter is positive, the tail gets increasingly shorter as a polynomial. The form parameter had a negative value, which meant that the distribution had a finite tail. Based on table 2, it can be deduced that the extreme returns in Bitcoin have a finite tail. The detailed information on the distribution's spreads is obtained from the scale values. Since the scale is high, the uncertainty of the returns is high, but the frequency for the scale is used is low. Scalability is a significant advantage of a large number of scales. However, this number must be carefully reported, as the information is based on $\xi$, and citation is required. Based on these findings, it can be inferred that Bitcoin has a high value of scale, with a value of 3.8374. Figure 3 presents the return level with return periods for Bitcoin.

![Figure 3. Return level for the Bitcoin](image)

**Table 3. Return level for various return periods**

|           | Return Level Estimates on Daily Loses |
|-----------|---------------------------------------|
|           | 2-year period | 3-year period | 4-year period | 5-year period |
| BTC       | 17.011        | 17.808        | 18.345        | 18.746        |

* Data for 1-year period is too small.

Applied to estimate the extreme return level in table 3, the parameter estimations in the table are referred to. For Bitcoin, the returns level for the 2 years is 17.011, which indicates that the potential maximum of losses daily will not exceed 17.011. In other words, within this period, the maximum extreme return will not exceed 17.011. A high level of returns indicates that the Bitcoin returns have a high magnitude for the interval period observed. Table 3 also presents that Bitcoin has the highest return level for 5 years periods. As we can see, this result is consistent and not far between all period. These results provide some insights indicating that the magnitude of extreme returns level for the Bitcoin range between 17.011 and 18.746. This indicates that the daily negative returns are between 17.011 and 18.746 for most of the cases. Overall, these results show that the extreme behaviours on Bitcoin is different and can be explained using EVT.

5. Conclusions

Understanding the risk probability is of primary importance when making financial decisions. Many different approaches have been suggested, but they only seem to be applicable when the normality statement is valid. In comparison, EVT provides a better way to better comprehend tail returns.
Extreme events, including volcanic eruptions, tsunamis, and earthquakes, can be explained using EVT. As a result, EVT can provide insightful information on these events, as opposed to other approaches. This study focuses on the downside risk of extreme returns.

Up now, there have only been a few studies of extreme returns for Bitcoin and the majority concentrate on the entire distribution instead of the tail distribution. For the past research, these values tend to be distinguished based on the drastic price fluctuations or the tails. The empirical investigation carried out in the current research used EVT for financial risk modelling for Bitcoin.

This research aims to provide insightful knowledge on the EVT approach and generate opportunities for expanding tail dependency modelling. The EVT model was applied to the frequent extreme returns of Bitcoin. The results provide a decent representation of the observed data, as well as sufficient agreement with the theoretical models. Furthermore, our findings show that informative knowledge obtained from the model is realistic in the financial sectors.

A few important findings arose from the observational studies in terms of high-magnitude extreme returns and uncertainty of Bitcoin. As well, we check this through return levels, which show that Bitcoin has the highest return level over all periods. However, the value is not significantly different between the periods.

The results of this investigation also show that investors and policymakers should be cognizant of the riskier asset. Currently, the study has been focused on modelling extreme returns, but it also opens the door for additional discussions. This discovery expands our comprehension of the extreme characteristics of Bitcoin. However, the results have left other questions unanswered, such as the explanations for the observations that were found. More research is needed to establish the relationship between this feature and its significance. Future research should invest more time and effort to include more information and specifics to strengthen the EVT as an extreme calculation in risk management.

Acknowledgements
This research would not be possible without sponsorship from the Universiti Kebangsaan Malaysia (grant number GGPM-2018-070).

References
[1] H Markowitz 1952 The Journal of Finance vol. 7 no. 1 pp. 77-91.
[2] R C 1973 Merton Econometrica vol. 41, no. 5, pp. 867-87.
[3] W F Sharpe 1964 The Journal of Finance vol. 19, no. 3, pp. 425-42.
[4] M Parkinson 1980 Journal of Business vol. 53, no. 1, pp. 61-5.
[5] R Fisher and L H Tippett 1928 In Mathematical Proceedings of the Cambridge Philosophical Society Vol. 24, No. 2, pp. 180-190. Cambridge University Press.
[6] F M Longin 1996 The Journal of Business vol. 69, no. 3, pp. 383-408.
[7] S I Hussain and S Li 2018 Financial Markets and Portfolio Management May;32(2):207-33.
[8] S I Hussain and S Li 2018 International Review of Economics & Finance. Jul 1;56:421-37.
[9] S I Hussain and S Li 2021 World Econ. March :1-19
[10] R S Tsay 2020 Journal of Econometrics. May 1;216(1):106-17.
[11] https://coinmarketcap.com . Retrieved on 20 Feb 2021.
[12] S Nakamoto 2008 Bitcoin: A peer-to-peer electronic cash system.
[13] P. Nakamoto 2017 Bitcoin CreateSpace Independent Publishing Platform.
[14] E. T. Cheah and J. Fry 2015 Economics Letters, 130, 32-36.
[15] A. Urquhart 2016 *Economics Letters*, 148, 80-82.
[16] A. H. Dyhrberg 2016 *Finance Research Letters*, 16, 85-92.
[17] D. G. Baur, K. Hong and A. D. Lee 2018 *Journal of International Financial Markets, Institutions and Money*, 54, 177-189.
[18] L. Catania, S. Grassi and F. Ravazzolo 2018 *Mathematical and Statistical Methods for Actuarial Sciences and Finance* (pp. 203-207). Springer, Cham.
[19] J. Osterrieder and J. Lorenz 2017 *Annals of Financial Economics*, 12(01), 1750003.
[20] M S Ismail, S I Hussain and M S Noorani MS 2020 *IEEE Access*. Nov 6;8:202042-57.
[21] S I Hussain, N Ruza, N Masseran and M.A.M Safari 2020 AIP Conference Proceedings 2266(1):030002.
[22] R. D. Reiss, M. Thomas and R.D. Reiss 2007 *Statistical analysis of extreme values* Vol. 2. Basel: Birkhäuser.
[23] J. Pickands 1975 *Annals of Statistics* 3, 119–131.
[24] M. Ribatet 2007 *R News*, 7, 33-36.
[25] H. Southworth and J.E. Heffernan 2014 R package version 2.1.
[26] J. R. Hosking and J. R. Wallis 1987 *Technometrics*, 29, 339–349.
[27] V.P. Singh and H. Guo 1995 *Hydrological Sciences-Journal-des Sciences Hydrologiques* 40, 165-181.
[28] R.L. Smith 1985 *Biometrika* 72, 67–90.
[29] J. Husler D. Li and M. Raschke 2011 *Communications in Statistics-Theory and Methods* 40, 2500-2510.
[30] S. Coles 2001 London: Springer.