



Article A Wavelet-Decomposed WD-ARMA-GARCH-EVT Model Approach to Comparing the Riskiness of the BitCoin and South African Rand Exchange Rates

Thabani Ndlovu * D and Delson Chikobvu

Department of Mathematical Statistics and Actuarial Science, University of the Free State, Bloemfontein 9300, South Africa; chikobvu@ufs.ac.za

* Correspondence: thabsndlovu89@gmail.com

Abstract: In this paper, a hybrid of a Wavelet Decomposition-Generalised Auto-Regressive Conditional Heteroscedasticity-Extreme Value Theory (WD-ARMA-GARCH-EVT) model is applied to estimate the Value at Risk (VaR) of BitCoin (BTC/USD) and the South African Rand (ZAR/USD). The aim is to measure and compare the riskiness of the two currencies. New and improved estimation techniques for VaR have been suggested in the last decade in the aftermath of the global financial crisis of 2008. This paper aims to provide an improved alternative to the already existing statistical tools in estimating a currency VaR empirically. Maximal Overlap Discrete Wavelet Transform (MODWT) and two mother wavelet filters on the returns series are considered in this paper, viz., the Haar and Daubechies (d4). The findings show that BitCoin/USD is riskier than ZAR/USD since it has a higher VaR per unit invested in each currency. At the 99% significance level, BitCoin/USD has average values of VaR of 2.71% and 4.98% for the WD-ARMA-GARCH-GPD and WD-ARMA-GARCH-GEVD models, respectively; and this is slightly higher than the respective 2.69% and 3.59% for the ZAR/USD. The average BitCoin/USD returns of 0.001990 are higher than ZAR/USD returns of -0.000125. These findings are consistent with the mean-variance portfolio theory, which suggests a higher yield for riskier assets. Based on the *p*-values of the Kupiec likelihood ratio test, the hybrid model adequacy is largely accepted, as p-values are greater than 0.05, except for the WD-ARMA-GARCH-GEVD models at a 99% significance level for both currencies. The findings are helpful to financial risk practitioners and forex traders in formulating their diversification and hedging strategies and ascertaining the risk-adjusted capital requirement to be set aside as a cushion in the event of the occurrence of an actual loss.

Keywords: wavelets; Haar; Daubechies (d4); generalised auto-regressive conditional heteroscedasticity; generalised extreme value distribution; generalised Pareto distribution; var; maximal overlap discrete wavelet transform

1. Introduction

Cryptocurrencies are decentralised currencies that are transacted without the regulations of a reserve bank or financial intermediaries. Blockchain technology is used to process transactions. BitCoin is on top of the list of traded cryptocurrencies in terms of traded volume.

Since BitCoin is not backed by any central bank or government, its users and traders are expected to be vulnerable to higher risk (volatility). As with the global trend, cryptocurrency trading, particularly in BitCoin, has gained a lot of traction in South Africa. There is a steady increase in movements of people's savings and investments between the Rand and BitCoin.

Cryptocurrencies are said to be very risky [1]. Developing countries' currencies, including the South African Rand, are equally risky [2]. The purpose of this study is to use the wavelet-decomposed–ARMA–GARCH–Extreme Value Theory (Generalised Pareto Distribution (GPD) and the Generalised Extreme Value Distribution (GEVD))-based Value at Risk (VaR) to compare the riskiness of the two currencies. The Generalised Pareto



Citation: Ndlovu, T.; Chikobvu, D. A Wavelet-Decomposed WD-ARMA-GARCH-EVT Model Approach to Comparing the Riskiness of the BitCoin and South African Rand Exchange Rates. *Data* 2023, *8*, 122. https://doi.org/ 10.3390/data8070122

Academic Editors: Varun Gupta, Chetna Gupta, Leandro Ferreira Pereira, Lawrence Peters and Antonio Ferreras

Received: 13 April 2023 Revised: 10 June 2023 Accepted: 14 July 2023 Published: 24 July 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). Distribution and Generalised Extreme Value Distribution used to describe/characterise extreme residuals from the return series models are preferred because these distributions are good at analysing extreme risk. The two distributions differ in the way they select extreme observations.

The VaR is a statistic that quantifies the riskiness of a financial portfolio of assets. It is the largest value or amount expected to be lost over a specified time horizon, i.e., daily, weekly, or ten days, at a pre-defined statistical confidence level. Ref. [3] defined VaR as the value that "compresses all Greek letters for all the market variables underlying a portfolio into a single number". Investors and practitioners rely heavily on VaR as a risk measure, even though it is not globally sub-additive, and hence not a coherent risk measure. The VaR metric is popular because its practical advantages outweigh its theoretical disadvantages. According to [4], VaR is sub-additive in most practical situations, which is in line with the diversification concept of modern portfolio theory.

Ref. [5] states that the behaviour of the financial time series shows volatility clustering, noise, leptokurtosis, and autocorrelation behaviour. Volatility clustering is when large changes in price tend to be followed by large changes, and small changes tend to be followed by small changes, whilst autocorrelation is the degree of correlation of the same variables between two time intervals. Hence, the use of wavelets transforms to decompose the series into different time horizons based on their volatility regime or cluster. Generalised Auto-Regressive Conditional Heteroscedasticity (GARCH) models capture conditional heteroscedasticity, making residuals independent and identically distributed [6]. Leptokurtosis describes fat tails or excess peaks prevalent in financial time series. Extreme value theory (EVT) models are recommended in efforts to correctly capture the extreme risk in financial assets.

The wavelet is a mathematical function that decomposes (breaks down) the original signal or time series into various components or sub-series [7]. They are very useful in capturing important features of the signal at each component level (resolution level) like long memory volatility features of a financial time series, which, in turn, helps in improving the forecasting capability of a model.

More recently, the use of wavelets in estimating volatility has gained popularity in the fields of hydrology, geophysics, and financial time series analysis [7–9].

Ref. [10] gave traction to the concept of wavelet analysis of a time series by providing the theory behind extracting useful statistical information from a physical science series, such as in hydrology. The wavelet methodology for the prediction of time-series data based on multi-scale decomposition was developed by [11]. Although a number of research papers have been published dealing with various theoretical aspects of wavelets, their application to data is still a difficult task. This is true with respect to financial data that are noisy, non-stationary, with fat tails, and often auto-correlated.

Cryptocurrencies are said to be very risky, and so is the South African Rand. The purpose of this study is to provide a detailed comparison of risks associated with each of the two currencies so as to provide information to the global investors, and foreign currency traders in the Republic of South Africa, and member countries of the Rand Union. In this paper, an improved empirical technique for estimating Value at Risk (VaR) by combining wavelets decomposition (WD), an ARMA model, a Generalised Autoregressive Conditional Heteroscedasticity (GARCH), and Extreme Value Theory (EVT) models is proposed. Will this "hybrid" model add value to the subject of the computation of VaR? This question will be addressed by fitting EVT models to the standardised residuals extracted from the WD-ARMA-GARCH model of the currency's daily log series, and employing backtesting techniques to the resulting VaR.

The rest of the paper is organised as follows: Section 2 presents the literature review. The methodology is in Section 3. Results and discussions are in Section 4, and Section 5 concludes.

2. Literature Review

Value at Risk (VaR) is one of the most commonly used risk measures in finance because of its ability to compress all Greeks to a single value [3]. Typically, VaR is computed by first modelling the entire returns distribution for the asset or a portfolio, then calculating the statistic at the percentile corresponding to the desired confidence level.

Ref. [12] showed that the traditional normal distribution-based VaR is not only incoherent but also fails to precisely estimate the risk of loss when the loss distributions have 'fat tails', unless EVT distributions are used. "This significantly discredits the accuracy of the traditional Normal distribution-based VaR risk measure" according to [13].

The EVT assumes independent and identically distributed (i.i.d) observations. This i.i.d assumption does not always hold for financial time series data. To correct this, [6] proposed a two-stage methodology in the form of a GARCH-EVT model using five index returns in their illustrations of modelling volatility. The first step is to capture the heteroscedasticity (non-constant variation or fluctuations) features by fitting a GARCH model. The second step is to apply the EVT to residuals extracted from a selected GARCH model using the Generalised Pareto Distribution (GPD) or the Generalized Extreme Value Distribution (GEVD). The second part of this modelling process allows one to capture or describe the large fluctuations in prices and returns. The merits of the GARCH-EVT hybrid model lie in its ability to capture conditional heteroscedasticity (changing variation) in the data through the GARCH framework, while, at the same time, modelling the extreme tail (large fluctuations) behaviour through the EVT methods.

While GARCH models have been to a larger extent successful in capturing most of the volatility stylised facts of financial time series [14], they are still lagging behind in detecting many structural changes (which are also prevalent) in the signals [15]. Ref. [9] showed that WD (Maximal Overlap Discrete Wavelet Transform)-GARCH captures well these structural changes and overall performs better than simple GARCH models. They used the daily stock price indices of four African countries' stock markets, Kenya's NSE20, Nigeria's All Share, South Africa's FTSE/JSE100, and Tunisia's TUNINDEX. Data from 2 January 2000 to 31 December 2014 were used.

According to [16], wavelet transforms perform better compared to the traditional Fourier transforms in signal processing. [17] used wavelet decomposition transforms and an ARIMA model to forecast the volatility of daily prices from the Amman Stocks Market (Jordan), from 1993 until 2009. Their findings showed that the approximated series under the wavelet transforms was better than the original series as it provided more stability in variance, and mean and smoothed out outliers. Furthermore, their forecasts using the ARIMA (p,d,q) under the wavelet transformation gave more accurate results than using the original signal.

Ref. [8] compared the performance of a Wavelet-Decomposed–Generalised Auto-Regressive Conditional Heteroscedasticity (WD-GARCH) model with a simple GARCH model in forecasting climate anomalies using the Multivariate ENSO Index (MEI), the global climate data time series index, for the period January 1950 to February 2018. They used the Akaike information criterion, Schwarz criterion, Hannan–Quinn criterion to perform the model selection, and the residual mean square error to assess the goodness of fit. Their results showed that both models fit the MEI data well. The forecast produced by the GARCH (1, 2) model underestimated the observed score, while the newly proposed WD-GARCH (1, 1) model generated more accurate forecasts for the given data. The authors recommend the WD-GARCH (1, 1) be applied to forecasts in the fields reflected by MEI variability.

Ref. [18] compared the performance of the Auto-Regressive Moving Average with eXplanatory variable(s) (ARIMAX), ARIMAX-GARCH, and ARIMAX-GARCH-WAVELET models in modelling volatility of wheat yield in the Kanpur district of Uttar Pradesh, India, from 1972 to 2013. His findings showed that the ARIMAX-GARCH-WAVELET outperformed other GARCH family models in forecasting the volatility of the wheat yield.

Ref. [19] used a statistical test approach in comparing the average returns and volatility of BitCoin against the Indonesian Composite Index, and gold. BitCoin average returns

were significantly higher than the other financial assets studied. BitCoin was expected to be riskier. This would be consistent with mean-variance portfolio theory, which suggests a higher yield for riskier assets [20].

There has been an increase in the amount of research to ascertain whether the stylised facts of cryptocurrency are similar to those of other financial assets. Ref. [1] showed that cryptocurrencies have similar distributional characteristics with Gold and the FTSE/JSE 40, though the cryptocurrency is more volatile. Ref. [21] noted the presence of heavy-tailedness and excess kurtosis in the one-minute return data of BitCoin. Ref. [22] observed a high negative skewness and volatility in BitCoin in comparison to other stock returns. Refs. [23,24] came to a similar conclusion that cryptocurrencies characteristics are nearly indistinguishable from the forex markets in well-established financial markets.

Ref. [25] argued that the shocks that are prevalent in the financial market do not affect BitCoin and gold returns; hence, they can be used for hedging. Conversely, ref. [26] noted relative stability in BitCoin and Ethereum using asymmetric power-law statistical distributions.

Ref. [27] used the RiskMetrics (constrained Integrated GARCH (1,1)) with heavy-tailed error distribution to compare the riskiness of investing in BitCoin against keeping the savings in a developing economies' currency, the South African Rand account. Their findings showed that BitCoin is riskier than the Rand. However, their backtest results suggested that the RiskMetrics model is inadequate at a 10% level of significance. Ref. [6] empirically showed the Generalised Auto-Regressive Conditional Heteroscedasticity–Extreme Value Theory (GARCH-EVT) model is superior (more accurate) to the EVT models in the estimation of VaR. Ref. [28] emphasised the importance of residuals in VaR estimation when she showed that a GARCH model with normal innovations is inferior to the GARCH model with Student's t innovations when data have fat tails (a common feature in the financial series data).

In this research, the interest is in capturing the distributional features like volatility clustering, conditional heteroscedasticity, structural breaks, and fat tails of two exchange rates returns time series, namely, BitCoin (BTC) to United States Dollar (USD), and the South African Rand (ZAR) to the USD, using wavelet-decomposed–ARMA-GARCH and EVT models. The ARMA component model quantifies the behaviour of the mean return. The modelling results can then be used in the estimation of Value at Risk to compare the riskiness of the two currencies.

3. Methodology

The steps involved in the estimation of the wavelet decomposed WD-ARMA-GARCH-EVT Value at Risk (VaR) can be summarized as follows:

- 1. Decompose and filter the log return series' using Maximal Overlap Discrete Wavelet Transform with two mother wavelets, namely, the Haar and Daubechies (d4).
- 2. Fit the Auto-Regressive Moving Average–Generalised Auto-Regressive Conditional Heteroscedasticity (ARMA-GARCH) component model to the wavelet transformed series.
- 3. Extract residuals and fit Extreme Value Theory models (Generalised Pareto Distribution, Generalised Extreme Value Distribution).
- 4. Estimate VaR and confirm model adequacy using Kupiec's backtest technique.

3.1. Wavelets

The wavelet is a mathematical function that decomposes (breaks down) the original signal or time series into various components. They are very useful in capturing long memory volatility features of a time series, which, in turn, helps in improving the forecasting capability of a model. Their advantage is their ability to be localised both in time and in the frequency domain, thus enabling the researcher to observe and analyse data at different scales [29].

"This methodology involves recursively applying a succession of low-pass and highpass filters to the signal (return series). This process allows the separation of high-frequency component from the low-frequency one" [30]. Therefore, improving the model's signal processing ability to capture structural changes including mildly explosive bubbles.

Wavelet transformations have been successfully applied to non-stationary time series data and yielded fairly good results in the fields of geo-sciences, remote sensing, engineering, hydrology, finance, medicine, ecology, renewable energy, chemistry, and history [31]. Its spectro-temporal abilities (ability to decompose the signal into different temporal scales/portions and inspect and carry out a frequency analysis of those portions of the signals) mean that they can be applied to non-stationary time series data with mildly explosive bubbles like BitCoin.

The decomposition of the series can be obtained using wavelet transform that is based on two filters, which are the "mother wavelet"

$$\int \varphi(t)dt = 0,$$

and the "father wavelet"

$$\int \Phi(t)dt = 1.$$

Then, the wavelet decomposition of a time series can be expressed as a linear combination of a wavelet function.

If $r(t) \in L^2(\mathbb{R})$ is a time series function (for t = 1, ..., T), the wavelet decomposed version is

$$r(t) = \sum_{k} s_{J,k} \Phi_{J,k}(t) + \sum_{k} d_{J,k} \varphi_{J,k}(t) + \sum_{k} d_{j-1,k} \varphi_{j-1,k}(t) + \dots + \sum_{k} d_{1,k} \varphi_{1,k}(t), \quad (1)$$

where the orthogonal basis functions $\Phi_{i,k}$ and $\varphi_{i,k}$ are defined as

$$\Phi_{j,k} = 2^{\frac{-j}{2}} \Phi\left(\frac{t-2^{j}k}{2^{j}}\right)$$
$$\varphi_{j,k} = 2^{\frac{-j}{2}} \varphi\left(\frac{t-2^{j}k}{2^{j}}\right),$$

where *j* represents the multiresolution, or scale level, and *k* depicts the number of coefficients in each scale level. $s_{j,k}$ and $d_{j,k}$ are the scaling (or smooth) and detail (or wavelet) coefficients, respectively, and are defined as

$$s_{j,k} = \int r(t)\Phi_{j,k}dt \tag{2}$$

$$d_{j,k} = \int r(t)\varphi_{j,k}dt \qquad \text{for } j = 1, 2, \dots J$$
(3)

The magnitude of these coefficients reflects a measure of the contribution of the corresponding wavelet function to the total signal.

The scale 2^{j} is also called the dilation factor and controls the length of the wavelet (window), whereas the translation parameter $2^{j}k$ refers to the location and indicates the non-zero portion of each wavelet basis vector. Equation (2) presents the long-scale smooth components that are used to generate the scaling coefficients, whereas the differencing coefficients are generated in Equation (3). The resulting multi-scale decomposition of Equation (1) can be simplified as

$$R(t) = S_J + D_J + D_{j-1} + \dots + D_1, \tag{4}$$

where D_j is the *j*th level wavelet and S_j represents the aggregated sum of variations at each detail of the scale. In Equations (1) and (4), the father wavelet reconstructs the smooth and low-frequency parts of the signal, whereas the mother wavelet function describes the detailed and high-frequency parts of the signal.

Therefore, Equation (4) provides a complete reconstruction of the signal decomposed into a set of j frequency components, so that each component corresponds to a particular range of frequencies.

3.2. Types of Wavelet Transforms

The wavelet transforms break the original signal (currency log returns time series) into projections of translated and scaled versions of the original mother wavelet as follows for Continuous Wavelet Transform (CWT).

$$r(\tau,s) = \frac{1}{|s|} \int_{-\infty}^{+\infty} f(t) \cdot \Psi^*\left(\frac{t-\tau}{s}\right) dt,$$
(5)

where r(t) the original signal or time series is, $\Psi^*(.)$ is the wavelet, $s = \frac{1}{frequency}$ is the scale parameter, and $\tau =$ translate wavelet across the signal.

For Discrete Wavelet Transform (DWT), the signal is broken down as follows:

$$G(a,b) = \frac{1}{\sqrt{b}} \sum_{m=0}^{p-1} r(t_m) \cdot \Psi^*\left(\frac{t_m - a}{b}\right), \ (6)$$

where $r(t_m)$ is the original signal or time series, $\Psi^*(.)$ is the wavelet, $b = s = 2^{-j}$ is the scale parameter, $a = \tau = k2^{-j}$ is the translate wavelet across signal, j = scaling index, and k = wavelet transformed signal.

The limitations of the above-mentioned transforms are well documented. The CWT maps the one-dimensional function r(t) to a function $R(\tau, s)$ having continuous real variables τ and s. The coefficients of $R(\tau, s)$ at a particular scale and translation measure how well the original function or signal r(t) matches with the scaled or translated mother wavelet. However, to recover the function, all the coefficients of $R(\tau, s)$ are not required. As a result, CWT gives a redundant way to represent the signal [32]. On the other hand, DWT, G(a, b), takes times that are multiples of 2^j , making it hard to compare with the original series or signal.

The Maximal Overlap Discrete Wavelet Transform (MODWT) is the preferred methodology for this research paper because its decomposition at different scales can easily be compared with the original time series, since it does not only take multiples of 2^j and is less sensitive to the starting point. This is helpful in understanding the patterns at different frequencies, i.e., short-term, medium-term or long-term. According to [10], the Maximal Overlap Discrete Wavelet Transform of a time series r(t), t = 1, 2, ..., N to the j^{th} level is as follows:

Wavelet coefficient
$$\widetilde{d}_{j,t} = \sum_{l=0}^{L_j-1} \widetilde{h}_{j,l} r(t)_{MODN}$$
,
Scale coefficient $\widetilde{s}_{j,t} = \sum_{l=0}^{L_j-1} \widetilde{g}_{j,l} r(t)_{MODN}$,

where

 $h_{j,l}$ = wavelet filter constructed by convolving *j* filters composed of \tilde{g}_l and h_l . It suffices the following conditions:

$$\sum_{l=0}^{L_j-1} \widetilde{h}_l = 0,$$
$$\sum_{l=0}^{L_j-1} \widetilde{h}_l^2 = \frac{1}{2},$$
$$\sum_{l=0}^{\infty} \widetilde{h}_l \widetilde{h}_{l+2n} = 0,$$

 $\tilde{g}_{j,l}$ = scale filter constructed by convolving *j* filters composed of \tilde{g}_l . It suffices the following conditions:

$$\sum_{l=0}^{L_j-1} \widetilde{g}_l = 1,$$
$$\sum_{l=0}^{L_j-1} \widetilde{g}_l^2 = \frac{1}{2},$$
$$\sum_{\to -\infty}^{+\infty} \widetilde{g}_l \widetilde{g}_{l+2n} = 0,$$

for all integers n > 0,

 $\sum_{l\to-\infty}^{+\infty} \widetilde{g}_l \widetilde{h}_{l+2n} = 0, \text{ for all integers } n, L_j = (2^j - 1)(L - 1) + 1.$

1

L is the width of the base level filter. The maximum number of levels depends on the available data points.

Although there are several mother wavelets, only some wavelets are suitable for financial time series analysis [33]. The Haar and Daubechies (d4) mother wavelets capture better the economic and financial time series characteristics, with non-stationary and structural changes [34,35]. These wavelets are well localised in time domain yet dispersed in frequency domain. Therefore, they can be used for analysis of time series with structural breaks and sharp jumps.

3.2.1. Haar Wavelet

Proposed by [36], the mother wavelet function $h_{i,l}$ can be described as follows for the *i*th level:

$$\widetilde{\widetilde{h}}_{j,l} = \begin{cases} 1, & 0 \le l < \frac{1}{2}; \\ -1, & \frac{1}{2} \le l < 1; \\ 0, & \text{otherwise} \end{cases}$$
(7)

and the scaling function can be described as follows:

$$\widetilde{g}_{j,l} = \begin{cases} 1, & 0 \le t < 1\\ 0, & \text{otherwise} \end{cases}$$

The Haar wavelet extracts information about how much difference there is between the two unit scale averages of f(t) bordering on the time t = 0.

3.2.2. Daubechies 4 (d4) Wavelet

The Daubechies 4 (d4) wavelet filter is

$$\widetilde{h}_{j,l} \begin{cases} \frac{1-\sqrt{3}}{4\sqrt{2}}, & l = 0\\ \frac{-3+\sqrt{3}}{4\sqrt{2}}, & l = 1\\ \frac{3+\sqrt{3}}{4\sqrt{2}}, & l = 2\\ \frac{-1-\sqrt{3}}{4\sqrt{2}}, & l = 3 \end{cases}$$

$$(8)$$

3.3. Auto-Regressive Moving Average–Generalised Auto-Regressive Conditional Heteroscedasticity (ARMA-GARCH)

The second step involves using the wavelet transformed series obtained from the multiresolution decomposition to fit the ARMA-GARCH-family models to model the

$$r_t = c + \varepsilon_t + \sum_{i=1}^p \varphi_i r_{t-1} + \sum_{j=1}^q \theta_j \varepsilon_{t-j}$$
(9)

The GARCH(p, q) model is mathematically defined as

$$\sigma_t^2 = w + \sum_{i=1}^q \alpha_i \varepsilon_{t-1}^2 + \sum_{j=1}^p \beta_j \sigma_{t-1}^2$$
(10)

The simplest form of the Generalised Auto-Regressive Conditional Heteroscedasticity model is the GARCH (1,1) with the equation and variance of form:

$$\begin{cases} r_t = \mu_t + \varepsilon_t \\ \varepsilon_t = \sigma_t y_t \\ \sigma_t^2 = w + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{cases}$$
(11)

Assuming residuals follow conditional normality and are i.i.d. Let $\theta = (\mu, w, \alpha_i, \beta_j)'$, the quasi-Gaussian maximum likelihood function is

$$l(\theta) = \frac{1}{\sqrt{2\pi\sigma_t^2(\theta)}} exp\left(-\frac{1}{2}\frac{\varepsilon_t^2(\theta)}{\sigma_t^2(\theta)}\right),\tag{12}$$

The log-likelihood of the above expression is

$$L(\theta) = \frac{1}{T} \sum_{t=1}^{T} log\left(\frac{1}{\sqrt{2\pi\sigma_t^2(\theta)}} exp(-\frac{1}{2}\frac{\varepsilon_t^2(\theta)}{\sigma_t^2(\theta)}\right),\tag{13}$$

the optimal parameters are obtained by maximising (13) with respect to θ .

While [37] found that the i.i.d innovations do not have fourth moments—leading to slower rates and unstable limits—the advantage of each of the Generalised Pareto Distribution and Generalised Extreme Value Distribution models lies in their ability to take a continuous range of possible distributional shapes influenced by the shape parameter also known as Extreme Value Index (EVI) parameter, which includes the bounded and unbounded innovation (tails) distributions as special cases. Each distribution is associated with 3 distributional forms depending on the heaviness of the tails. When the EVI is zero, the distributions are light-tailed, and when it is less than zero, the distributions are short-tailed or bounded. A positive EVI is associated with heavy tailedness. The Generalised Pareto Distribution and Generalised Extreme Value Distribution allow one to "let the data decide" which of these distributions is appropriate within each distribution, instead of having to select a particular form of the distribution function. The end result is a good model fit to the extremes (tails) of the data emanating from this flexibility.

3.4. Extreme Value Theory

The peak over threshold (POT) is used to select data for fitting the Generalized Pareto Distribution (GPD), and is used in our approach to model the standardized residuals emanating from the selected GARCH family model.

The block maxima method is another EVT approach for identifying maximum (extremes) in a data set. The Generalised Extreme Value Distribution (GEVD) is then fitted to the set of block maxima chosen in a given set of data. The data are initially arranged in time sequence and then grouped into non-overlapping blocks.

3.4.1. The Generalised Pareto Distribution (GPD)

The peak over threshold (POT) approach, used in fitting the Generalized Pareto Distribution, is used to model the standardized residuals from the WD-GARCH-family model.

Refs. [38,39] showed that for large enough thresholds, *u*'s, the POT function/exceedances above this threshold, can be estimated by Generalised Pareto Distribution. The Generalised Pareto Distribution is defined as follows:

$$G_{\xi,\beta}(y) = \begin{cases} 1 - \left(1 + \frac{\xi(y-u)}{\beta}\right)^{\frac{-1}{\xi}} \text{if } \xi \neq 0\\ 1 - exp\left(-\left(\frac{y-u}{\beta}\right)\right) \text{if } \xi = 0 \end{cases},$$
(14)

where y > 0 are standardised residuals above the threshold u, ξ is the shape parameter or extreme value index (EVI), and β is the scale parameter. The value of ξ shows how heavy the tail is, with a bigger positive value indicating a heavy tail. When ξ is negative, the tail is short (bounded). $\xi = 0$ gives indicates a light tail.

3.4.2. Parameter Estimation of the Generalised Pareto Distribution

Let *u* be a sufficiently high threshold, assuming *n* observations *y* such that $y_i - u \ge 0$, the subsample $\{y_1 - u, \ldots, y_n - u\}$ has an underlying distribution of a Generalised Pareto Distribution, where $y_i - u \ge 0$ for $\xi \ge 0$, $0 \le y_i - u \le -\frac{\beta}{\xi}$ for $\xi < 0$, then the logarithm of the probability density function of $y_i - u$ is

$$\ln(f(y_i - u)) = \begin{cases} -\ln(\beta) - \frac{1+\xi}{\xi} \ln\left(1 + \xi\left(\frac{y_i - u}{\beta}\right)\right) \text{if } \xi \neq 0\\ -\ln(\beta) - \frac{1}{\beta}(y_i - u) & \text{if } \xi = 0 \end{cases},$$
(15)

Then, the log-likelihood $L(\xi, \beta | y_i - u)$ for the model is the logarithm of the joint density of the *n* observations, i.e.,

$$L(\xi,\beta|y_i-u) = \begin{cases} -n\ln(\beta) - \frac{1+\xi}{\xi} \sum_{i=1}^n \ln\left(1+\xi\left(\frac{y_i-u}{\beta}\right)\right) \text{if } \xi \neq 0\\ -n\ln(\beta) - \frac{1}{\beta} \sum_{i=1}^n (y_i-u) \quad \text{if } \xi = 0 \end{cases},$$
(16)

We obtain the parameters (ξ, β) by maximising the log-likelihood function of the sub-sample under a suitable threshold *u*.

3.4.3. The Generalised Extreme Value Distribution (GEVD)

The Fisher–Tippett–Gnedenko theorem, first proposed by [40] and later revised by [41] is very important in extreme value theory. According to this theorem, the maxima or minima of a sample of observations will converge to the Generalised Extreme Value Distribution (GEVD).

The GEVD is the limiting distribution of the normalised block maxima of a sequence of independent identically distributed random variables. The GEVD is given as follows:

$$G_{\xi,\mu,\sigma}(y) \begin{cases} \exp\left\{-\left(1+\xi\left(\frac{y-\mu}{\sigma}\right)\right)^{-1/\xi}\right\}, \text{ if } \xi \neq 0\\ \exp\left\{-\exp\left(-\left(\frac{y-\mu}{\sigma}\right)\right)\right\}, \text{ if } \xi \to 0 \end{cases}$$
(17)

with $\xi \neq 0, \sigma > 0$ and $1 + \xi \left(\frac{y-\mu}{\sigma}\right) > 0$.

The probability density function, obtained as the derivative of the above distribution function, is given by:

$$g_{\xi,\mu,\sigma}(y) \begin{cases} \frac{1}{\sigma} \left(1 + \xi \left(\frac{y-\mu}{\sigma} \right) \right)^{-1-1/\xi} \exp\left\{ - \left(1 + \xi \left(\frac{y-\mu}{\sigma} \right) \right)^{-1/\xi} \right\} & \text{if } \xi \neq 0 \\ \frac{1}{\sigma} \exp\left\{ - \exp\left(- \left(\frac{y-\mu}{\sigma} \right) \right) \right\} \exp\left(- \left(\frac{y-\mu}{\sigma} \right) \right) & \text{if } \xi \to 0 \end{cases}$$
(18)

with $\xi \neq 0$ and where μ and σ are the location and scale parameters, respectively. The shape parameter ξ is also known as the extreme value index (EVI).

The block maxima method is an EVT approach for identifying maximum (extremes) in a data set and for describing their behaviour. The Generalised Extreme Value Distribution is fitted to a set of maxima chosen in a given set of data. This research will use a weekly block of size 7 days.

3.4.4. Parameter Estimation

Let $Y_1, ..., Y_m$ are independent variables following the Generalised Extreme Value Distribution, the log-likelihood for the parameters, when $\xi \neq 0$, is

$$L = \begin{cases} -m\ln(\sigma) - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{m} \ln\left[1 + \xi\left(\frac{y_i - \mu}{\sigma}\right)\right] - \sum_{i=1}^{m} \left[1 + \xi\left(\frac{y_i - \mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ -m\ln(\sigma) - \sum_{i=1}^{m} \exp\left[-\left(\frac{y_i - \mu}{\sigma}\right)\right] - \sum_{i=1}^{m} \left(\frac{y_i - \mu}{\sigma}\right) & \text{if } \xi \to 0 \end{cases}$$
(19)

Maximisation of the above function with respect to the parameters vector (ξ , μ , σ), leads to the maximum likelihood estimates for the entire Generalised Extreme Value Distribution family [42].

3.5. Value at Risk

The formula to compute Value at Risk, for a small tail probability *p*, and total sample size *n*.

For a Generalised Pareto Distribution with maximum likelihood estimates $(\hat{\beta}, \hat{\sigma}, \hat{\zeta})$, threshold *u* and *N*_{*u*}, the number of exceedances, is given by:

$$\widehat{VaR_p} = \begin{cases} u + \frac{\hat{\beta}}{\hat{\xi}} \left\{ 1 - \left[-n\ln(1-p) \right]^{-\hat{\xi}} \right\} & \text{if } \hat{\xi} \neq 0 \\ u - \hat{\sigma} \ln(-n\ln(1-p)) & \text{if } \hat{\xi} = 0 \end{cases}$$
(20)

for a Generalised Extreme Value Distribution with maximum likelihood estimates $(\hat{\mu}, \hat{\sigma}, \xi)$:

$$\widehat{VaR_p} = \begin{cases} u + \frac{\hat{\sigma}}{\hat{\xi}} \left\{ \left(\frac{n}{N_u} p \right)^{-\hat{\xi}} - 1 \right\} \text{if } \hat{\xi} \neq 0 \\ u - \hat{\beta} \ln\left(\frac{n}{N_u} (1-p) \right) \text{if } \hat{\xi} = 0 \end{cases}$$
(21)

Finally, according to [6], the VaR of the asset is computed using the following formula:

$$VaR_p(r_t) = \mu_t + \sigma_t VaR_p(y_t), \tag{22}$$

where is μ_t is derived from the mean equation (Auto-Regressive Moving Average) and σ_t is estimated from the volatility model (Generalised Auto-Regressive Conditional Heteroscedasticity). VaR_p(y_t) is the *p* percentile of the standardised residuals. Often, μ_t does not vary much and is often predictable. The riskiness of the asset is then expressed through VaR_p(y_t), hence, the modelling of the residuals. This is especially important when modelling extreme risk.

4. Data Analysis

The currency data used in this research were obtained from the finance sector website (www.investing.com/currencies accessed on 1 July 2021). Analyses were conducted using R, RStudio, WaveletGARCH, wavelets, evir, FinTS, PerformanceAnalytics, ismev, and eva statistical packages. The adjusted closing values of daily exchange rates from 1 January 2015 to 30 June 2021 were fitted to the WD-ARMA-GARCH-EVT model. BitCoin is traded every day; hence, there are 2372observations. The Rand is not traded on weekends and South Africa's public holidays resulting in 1694 observations. To align our data for analysis, we replaced missing return values in the Rand exchange rate with zero since there are no

profits or losses realised by the holder of the local currency during the weekend and/or public holidays. The daily log returns were calculated and used for modelling. The formula for log returns used is $y_t = \log \left[\frac{P_t}{P_{t-1}}\right]$, where P_t and P_{t-1} are today and yesterday's closing values of daily prices (exchange rates), respectively.

In Figures 1 and 2, the log returns look stationary, around the zero mean, although volatility is non-constant and clustered, indicating heteroscedasticity, which is common with financial data. Isolated extreme returns are visible; hence, EVT models will be used to capture risks associated with these extremes.

BTC/USD prices

BTC/USD returns



Figure 1. Plot of BitCoin/USD prices (left) and one-day log returns (right).



Figure 2. Plot of ZAR/USD prices (left) and one-day log returns (right).

4.1. Descriptive Statistics

Table 1 below gives the descriptive statistics.

	Observations	Mean	Median	Maximum	Minimum	Skewness	Kurtosis	
Bitcoin/USD	2370	0.001990	0.00177	0.23722	-0.4809	-0.9944	16.154	
ZAR/USD	1694	-0.00125	0.00000	0.04955	-0.0485	-0.2641	4.1216	
		Test for No	ormality, Autocorre	lation, and Heteros	scedasticity			
		Bitcoin/USD			ZAR	/USD		
TEST	Statistic		<i>p</i> -value	Stat	Statistic		<i>p</i> -value	
Jarque-Bera	17,478.40		0.000000	108.4967		0.0000		
Ljung–Box	11.7		0.0006249	0.40	0.40504		0.5245	
ARCH LM Test	52.87		$4.3 imes10^{-7}$	70.789		$2.28 imes 10^{-10}$		
			Test for unit root	t and stationarity				
		Bitcoin/USD			ZAR	/USD		
Unit Root Test	Statis	stic	<i>p</i> -value	Stat	istic	p-va	lue	
ADF Test	-52.20	0130	0.0001	-40.4	47263	0.00	000	
PP Test	-52.10	0963	0.0001	-40.4	47011	0.00	000	
KPSS Test	0.092	067	0.3470	0.09	0747	0.34	170	

Table 1. Descriptive statistics of exchange rate price returns.

Source: ref. [43] (page 7).

In Table 1, the Null Hypothesis of Normality using Jarque–Bera is rejected at the 5% level of significance, meaning the use of symmetric models should not be considered when analysing the above-mentioned return series.

The significant *p*-value of the Ljung–Box test for ZAR/USD returns suggest the failure to reject the Null Hypothesis of no autocorrelation. This means observations can be assumed to be independent and identically distributed (i.i.d). However, for BitCoin/USD returns, this null hypothesis is rejected; hence, a two-stage approach by McNeil and Frey will therefore be used to help deal with the autocorrelation problem. The first stage of fitting WD-ARMA-GARCH will eliminate this autocorrelation.

The stationarity tests (ADF and PP) confirm that, at the 5% level of significance, the Null Hypothesis of a unit root is rejected, and it can be concluded that both exchange rate return series are stationary. The KPSS test results showed that all returns are stationary as well.

4.2. WAVELETS-ARMA-GARCH

Based on the literature and the financial characteristics presented above, using a hybrid of wavelets decomposition, ARMA, GARCH, and EVT models can lead to a better measure of equity risk. The return series are non-normal and have a heteroscedasticity feature, and, as shown in Figure 3, they have long memory volatility. Using the Wavelet-ARMA-GARCH model can aid in capturing these features in the estimation of risk.

Figure 3 shows the wavelet coefficients $(s_{i,k})$ for all eight levels and the scale co-

efficients $(d_{j,k})$ for the eighth level. The Haar and Daubechies wavelets were used for computing the Maximal Overlap Discrete Wavelet Transform coefficients. The wavelet coefficients are used to decompose the signal or series (WD) according to the level (information they have), i.e., high frequency (volatility) or low frequency (volatility). Then, at each level, we fit an Auto-Regressive Moving Average (ARMA)–Generalised Auto-Regressive Conditional Heteroscedasticity model (GARCH), and, finally, combine the values for estimating volatility.

On the left-hand side is the BitCoin/USD plot of Maximal Overlap Discrete Wavelet Transform, and on the right-hand side is the ZAR/USD plot of Maximal Overlap Discrete Wavelet Transform. The wavelet coefficients are smoother at a higher level, representing longer-term volatility. The scale coefficients at the highest level represent the volatility that is not explained by wavelet coefficients. The T^{-i} means that the series of the coefficients is shifted by *i* positions backwards so that all the series are on the same timeline. The WD(Haar)-ARMA-GARCH(1,1) and the WD(d4)-ARMA-GARCH(1,1) was then fitted for



both exchange rates. The GARCH parameters were estimated using the quasi-Gaussian maximum likelihood estimators.

Figure 3. Maximal Overlap Discrete Wavelet Transform Wavelet Coefficients and Scaling Coefficients for BitCoin/USD, on the left, and ZAR/USD, on the right.

Table 2 presents the WD-ARMA-GARCH optimal parameters for BitCoin/USD and ZAR/USD for the Haar wavelet transformed and the Daubechies (d4) transformed series. These are the models that were used to capture volatility clustering and conditional heteroscedasticity. The Haar transformed series resulted in WD(Haar)-ARMA(2,0)-GARCH(1,1) for BitCoin/USD and WD(Haar)-ARMA(1,0)-GARCH(1,1) for the ZAR/USD. The Daubechies (d4) transformed series resulted in WD(d4)-ARMA (2,3)-GARCH(1,1) for BitCoin/USD and WD(d4)-ARMA(1,3)-GARCH(1,1) for ZAR/USD. To capture fat tails, their residuals were extracted, standardized, and used to fit EVT models, which, in turn, were used to estimate VaR, and model adequacy confirmation was performed usingbacktesting techniques.

		Bitcoin/USD				ZAR	/USD	
	Haar		(Daubechies, d4)		Haar		(Daubechies, d4)	
	Parameter	<i>p</i> -Value	Parameter	<i>p</i> -Value	Parameter	<i>p</i> -Value	Parameter	<i>p</i> -Value
constant(c)	-0.001211	< 0.0001	0.000398	< 0.0001	0.00099	< 0.001	0.00018	0.0396
$AR(1)(\varphi_1)$	0.917516	< 0.0001	1.962064	< 0.0001	0.99441	< 0.001	0.99962	< 0.001
AR(2)(φ_2)	0.080384	< 0.0001	-0.962151	< 0.0001				
$MA(1)(\theta_1)$			-0.318567	< 0.0001			0.80351	< 0.001
$MA(2)(\theta_2)$			-0.218019	< 0.0001			0.52842	< 0.001
MA(3)(θ_3)			0.127469	< 0.0001			0.53169	< 0.001
<i>(w)</i>	0.000000	0.98643	0.000000	0.999825	0.00000	0.0058	0.00000	0.9999
(<i>a</i> ₁)	0.051305	< 0.0001	0.025967	0.039547	0.02594	0.0774	0.05005	< 0.001
(<i>a</i> ₂)			0.025967	0.082290	0.02580	0.0974		
(β ₁)	0.896119	< 0.0001	0.894167	< 0.0001			0.89989	< 0.001

Table 2. The optimal parameters for the 8th level of each return series.

4.2.1. BitCoin Returns (BTC/USD)

To fit the Generalised Pareto Distribution model, a threshold *u* must be selected. The mean excess plots determine a suitable threshold, which is necessary for fitting the Generalised Pareto Distribution model. The choice of a threshold should be depicted by linear increases in the mean excess plot. Figures 4 and 5 present the mean excess function (in blue) and Q-Q plots (in red) of BitCoin/USD residuals from WD(Haar)-ARMA(2,0)-GARCH(1,1) and WD(d4)-ARMA(2,3)-GARCH(1,1), respectively. By observing these mean excess functions, a threshold of between 0 and 1 seems to be a reasonable choice. The 80th percentile was selected for all series, and it provided a reasonable choice as it falls within the above range.



Figure 4. Mean Excess for BitCoin/USD residuals from WD (Haar)-ARMA(2,0)-GARCH(1,1).



Figure 5. Mean Excess for BitCoin/USD residuals from WD(d4)-ARMA(2,3)-GARCH(1,1).

The parameters of the Generalised Pareto Distribution were estimated using the Maximum Likelihood method and are presented in the Table 3 below.

Table 3. Parameters of estimates using Generalised Pareto Distribution for BitCoin/USD.

Model	Threshold	Exceeds	ŝ	$\hat{Se(\xi)}$	$\hat{\beta}$	$\hat{Se(\beta)}$
WD (Haar)-ARMA(2,0)-GARCH(1,1)-GPD	0.5676	474	0.04988	0.0505	0.66238	0.04519
WD (d4)-ARMA(2,3)-GARCH(1,1)-GPD	0.5681	474	0.05204	0.0508	0.65925	0.04513

The shape parameter, in Table 3, is positive for all cases. Hence, the extreme returns exhibit heavy tails forall models. This can be loosely interpreted as a suggestion that BitCoin is riskier as heavy tails imply a higher concentration of observations at the extremes.

Figure 6 shows graphical goodness-of-fit plots for BitCoin returns fitted with WD(Haar)-ARMA(2,0)-GARCH(1,1)-GPD. The Probability and Quantile plots are almost linear, confirming a good fit. The return levels are within the confidence bands, as expected. The density plot is also a good estimate for the histogram of the data. The model fits the BitCoin returns data well.



Figure 6. Model diagnostic plots for BitCoin/USD returns fitted with WD (Haar)-GARCH–Generalised Pareto Distribution (WD(Haar)-ARMA(2,0)-GARCH(1,1)-GPD).

Figure 7 shows graphical goodness of fit plots for BitCoin returns fitted with WD(d4)-ARMA (2,3)-GARCH(1,1)-GPD. The Probability and Quantile plots are almost linear, confirming a good fit. The return levels are within the confidence bands as expected. The density plot is also a good estimate for the histogram of the data. The model fits the BitCoin returns data well.

The block maxima of the BitCoin/USD log returns have been fitted to Generalised Extreme Value Distribution with weekly block sizes. Table 4 shows the maximum likelihood estimates of the parameters and their corresponding standard errors (SE). The shape parameters ($\hat{\xi}$'s) are positive, implying that extreme returns follow a heavy tail Fréchet class distribution [44]. These parameter estimates imply that data sets are heavy-tailed.

Figure 8 shows graphical goodness-of-fit plots for BitCoin returns fitted with WD (Haar)-ARMA(2,0)-GARCH(1,1)-GEVD. The Probability and Quantile plot are almost linear,

confirming a good fit. The return levels are within the confidence bands as expected. The density plot is also a good estimate for the histogram of the data. The model fits the BitCoin returns data well.



Figure 7. Model diagnostic plots for BitCoin/USD returns fitted with WD (d4)-GARCH-GPD (WD(d4)-ARMA(2,3)-GARCH(1,1)-GPD.

 Table 4. Parameter estimates using Maximum likelihood for Generalised Extreme Value Distribution (BitCoin/USD).

Model	Maxima	$\hat{\xi}$	$Se(\hat{\xi})$	$\hat{\sigma}$	$Se(\hat{\sigma})$	$\hat{\mu}$	Se(µ)
WD(Haar)-ARMA(2,0)-GARCH(1,1)-GEVD	339	0.1896	0.0556	0.5804	0.0300	0.7270	0.03733
WD(d4)-ARMA(2,3)-GARCH(1,1)-GEVD	339	0.1895	0.0555	0.5796	0.0299	0.7275	0.03728



Figure 8. Model diagnostic plots for BitCoin/USD using WD (Haar)-ARMA-GARCH–Generalised Extreme Value Distribution (WD(Haar)-ARMA(2,0)-GARCH(1,1)-GEVD).

Figure 9 shows graphical goodness-of-fit plots for BitCoin returns fitted with WD (d4)-ARMA(2,3)-GARCH(1,1)-GEVD. The Probability and Quantile plots are almost linear, and the density plot is also a good estimate for the histogram of the data. The (WD(d4)-ARMA (2,3)-GARCH(1,1)-GEVD) model fits the BitCoin returns data well.



Figure 9. Model diagnostic plots for BitCoin/USD using WD (d4)-ARMA-GARCH–Generalised Extreme Value Distribution (WD(d4)-ARMA(2,3)-GARCH(1,1)-GEVD).

4.2.2. South African Rand Returns (ZAR/USD)

Figures 10 and 11 present the mean excess function (in blue) and the Q-Q plots (in red) of ZAR/USD residuals from the WD(Haar)-ARMA(1,0)-GARCH and WD(d4)-ARMA(1,3)-GARCH(1,1) models, respectively. By observing these mean excess functions, a threshold of between 0 and 1 seems to be a reasonable choice. The 80th percentile was selected for all series and it provided a reasonable choice as falls within the above range.



Figure 10. Mean Excess for ZAR/USD residuals from WD (Haar)-ARMA(1.0)-GARCH(1,1).



Figure 11. Mean Excess for ZAR/USD residuals from WD(d4)-ARMA-GARCH–Generalised Pareto Distribution(WD(d4)-ARMA(1,3)-GARCH(1,1)-GPD).

The parameters of the Generalised Pareto Distribution were estimated using the ML method and are presented in the Table 5 below.

Table 5. Parameters of estimates using	Generalized Pareto Distribution for ZAR/USI
--	---

Model	Threshold	Exceeds	$\hat{\boldsymbol{\xi}}$	$\hat{Se(\xi)}$	$\hat{oldsymbol{eta}}$	$Se(\hat{\beta})$
WD(Haar)-ARMA(1,0)-GARCH(1,1)-GPD	0.7684	339	-0.06377	0.04701	0.70598	0.05066
WD(d4)-ARMA(1,3)-GARCH(1,1)-GPD	0.7641	339	-0.06875	0.04624	0.71449	0.05090

In Table 5, the shape parameter $(\hat{\zeta})$ is negative in all cases. Hence, the extreme returns follow a short-tailed Pareto type II family of Generalised Pareto Distribution for all models.

Figure 12 shows graphical goodness-of-fit plots for Rand returns fitted with WD (Haar)-ARMA(1,0)-GARCH(1,1)-GPD. The Probability and Quantile plots are almost linear, confirming a good fit. The return levels are within the confidence bands, as expected. The density plot is also a good estimate for the histogram of the data. The model fits the Rand returns data well.



Figure 12. Model diagnostic plots for ZAR/USD returns fitted with WD(Haar)-ARMA-GARCH–Generalised Pareto Distribution(WD(Haar)-ARMA(1,0)-GARCH(1,1)-GPD).

Figure 13 shows graphical goodness-of-fit plots for Rand returns fitted with WD(d4)-ARMA(1,3)-GARCH(1,1)-GPD. The Probability and Quantile plots are almost linear suggesting a good fit, and the density plot is also a good estimate for the histogram of the data. The model fits the Rand returns data well.



Figure 13. Model diagnostic plots for ZAR/USD returns fitted with WD(d4)-ARMA-GARCH–Generalised Pareto Distribution(WD(d4)-ARMA(1,3)-GARCH(1,1)-GPD).

The block maxima of the extreme returns have been fitted to Generalised Extreme Value Distribution with weekly block size. Table 6 shows the maximum likelihood estimates of the parameters and their corresponding standard errors (SE). The shape parameter estimates $(\hat{\xi})$ are negative for all cases. Hence, the extreme returns follow a negative Weibull class distribution of the Generalised Extreme Value Distribution for all models, implying that the returns are short-tailed or bounded.

Table 6. Parameter estimates using Maximum likelihood for GEVD (ZAR/USD).

Model	Maxima	$\hat{\boldsymbol{\xi}}$	Se(z)	$\hat{\sigma}$	$Se(\hat{\sigma})$	ĥ	Se(µ)
WD(Haar)-ARMA(1,0)-GARCH(1,1)-GEVD	339	-0.0097	0.03659	0.62710	0.02731	0.76964	0.03797
WD(d4)-ARMA(1,3)-GARCH(1,1)-GEVD	339	-0.0090	0.03661	0.62703	0.02732	0.76928	0.03797

Figure 14 shows graphical goodness-of-fit plots for Rand returns fitted with WD (Haar)-ARMA(1,0)-GARCH(1,1)-GEVD. The Probability and Quantile plots are almost linear, confirming a good fit. The return levels are within the confidence bands, as expected. The density plot is also a good estimate for the histogram of the data. The model fits the Rand returns data well.

Figure 15 shows graphical goodness-of-fit plots for Rand returns fitted with WD (d4)-ARMA(1,3)-GARCH(1,1)-GEVD. The Probability and Quantile plots are almost linear, confirming a good fit. The return levels are within the confidence bands, as expected. The density plot is also a good estimate for the histogram of the data. The model fits the Rand returns data fairly well.



Figure 14. Model diagnostic plots for ZAR/USD using WD (Haar)-ARMA-GARCH–Generalised Extreme Value Distribution (WD (Haar)-ARMA(1,0)-GARCH(1,1)-GEVD).



Figure 15. Model diagnostic plots for ZAR/USD using WD (d4)-ARMA-GARCH–Generalised Extreme Value Distribution (WD(d4)-ARMA(1,3)-GARCH(1,1)-GEVD).

4.3. Value at Risk and Back Test Results

The computed VaR figures presented in Table 7, suggest that BitCoin/USD is riskier than the ZAR/USD since it has a higher value at risk per USD invested in each currency. At 99% significance levels, BitCoin/USD has an average of (2.70784 + 2.705287)/2 = 2.71% and (4.989638 + 4.983419)/2 = 4.98% for WD-ARMA-GARCH-GPD and WD-ARMA-GARCH-GEVD, respectively; this is slightly higher than 2.69% and 3.59% for the ZAR/USD, respectively. In monetary terms, at 99% level significance, an investor holding on to BitCoin is likely to lose extremes of almost USD 5.00 per USD 100.00 invested compared to the USD

3.60 likely to be lost by one holding on to the Rand, confirming the high risk associated with BitCoin [1]. The average returns presented in Table 1 show that BitCoin/USD returns are higher than ZAR/USD returns. These findings are consistent with the mean-variance portfolio theory, which suggests a higher yield for riskier assets [20].

Table 7. VaR estimates using fitted.

Model	Bitcoi	n/USD	ZAR/USD		
inoucl	95%	99%	95%	99%	
WD (Haar)-ARMA-GARCH-GPD	1.70549	2.70784	1.518355	2.693945	
WD (Haar)-ARMA-GARCH-GEVD	3.042301	4.989638	2.605691	3.591025	
WD (d4)-ARMA-GARCH-GPD	1.709211	2.705287	1.51578	2.698803	
WD (d4)-ARMA-GARCH-GEVD	3.039408	4.983419	2.606896	3.594547	

Based on the *p*-values presented in Table 8, the Kupiec likelihood ratio tests confirm that the fitted models are well suited to the returns series since the observed *p*-values are greater than 0.05, except for the WD(Haar)-ARMA-GARCH-GEVD and WD(d4)-ARMA-GARCH-GEVD models with99% significance level for both currencies, both at 95%. Model adequacy is largely accepted. The model with the highest *p*-value is considered to be the best-fit model, and hence it is recommended for use by financial risk analysts in estimating currency VaR.

Table 8. *p*-Values of Kupiec likelihood ratio test for VaR estimates using fitted models.

Model	Bitcoi	n/USD	ZAR/USD		
mouch	95%	99%	95%	99%	
WD (Haar)-ARMA-GARCH-GPD	0.8143285	0.3148938	0.9377182	0.3157879	
WD (Haar)-ARMA-GARCH-GEVD	0.4490511	< 0.001	0.6204366	< 0.001	
WD (d4)-ARMA-GARCH-GPD	0.7426442	0.3148938	0.8492139	0.2030873	
WD (d4)-ARMA-GARCH-GEVD	0.4490511	< 0.001	0.6204366	<0.001	

The Kupiec test suggests that the WD (Haar)-ARMA-GARCH-GPD models give the best fit for both the BitCoin/USD and ZAR/USD currencies.

5. Discussions and Conclusions

In this study, the computation and performance of selected wavelet-decomposed (WD)–ARMA-GARCH-EVT-based Value at Risk (VaR) methodologies are explored using BitCoin/USD and ZAR/USD data. The time series are decomposed using the Maximal Overlap Discrete Wavelet Transform technique and filtered using the Haar and Daubechies, d4 wavelets. Table 2 suggests that theHaar transformed series resulted in WD(Haar)-ARMA(2,0)-GARCH(1,1) for BitCoin/USD, while WD(Haar)-ARMA(1,0)-GARCH(1,1) was the result for ZAR/USD. Daubechies (d4) transformed series resulted in WD(d4)-ARMA(2,3)-GARCH(1,1) for BitCoin/USD, while WD(d4)-ARMA(1,3)-GARCH(1,1) was the result for ZAR/USD. The shape parameters ($\hat{\xi}$) for BitCoin/USD, as presented in Tables 3 and 4, are positive, implying that the BitCoin returns follow a heavy-tailed distribution like Pareto and Fréchet. Conversely, the shape parameters ($\hat{\xi}$) for the ZAR/USD, as presented in Tables 5 and 6, are negative, signifying that the Rand returns follow a bounded distribution [44].

The EVT model provided a good fit to the tails of the distribution of the returns. The diagnostic plots showed that the Probability and Quantile plots do not deviate significantly from a straight line, signifying a good fit.

The daily VaR with two percentiles 95% and 99% were estimated and are summarised in Table 7. Both confidence levels reveal that BitCoin/USD has a higher value at risk than the ZAR/USD, leading one to conclude that BitCoin is riskier than the Rand. But, also,

Table 1 shows that the average returns for the BitCoin/USD are higher than those of the ZAR/USD. These findings are consistent with the mean-variance portfolio theory, which suggests a higher yield for riskier assets [20].

Kupiec's likelihood ratio test values presented in Table 8 confirm model adequacy for both series, except WD(Haar)-ARMA-GARCH-GEVD and WD(d4)-GARCH-GEVD for both currencies at 99% as the *p*-values are less than 0.05, rejecting model adequacy at a 5% level of significance. This implies that the capacity of the models partly depends on the level of significance used.

The purpose of this study was to use Wavelet decomposition, ARMA, and GARCH combined with EVT in the estimation of VaR of the daily returns of both BitCoin (BTC) and the South African Rand (ZAR) against the USD, and to compare their riskiness. Both currencies suggest that the WD(Haar)-ARMA-GARCH-GPD model performs fairly well. This could be of great help to global investors and forex market risk managers in South Africa to understand the risk to which they are exposed when they convert their savings from Rand to BitCoin, particularly in choosing the model that gives better estimates in computing VaR, an important risk measure in the estimation of risk-adjusted capital requirements.

This information is useful to local foreign currency traders and investors who need to fully appreciate the tail-related return levels and risk exposure when they convert their savings or investments to BitCoin instead of the South African currency, the Rand. Particularly, when the market enters a turbulent time, BitCoin is riskier than the South African Rand, which is a developing country's currency. These results, though, do not imply that WD(Haar)-ARMA-GARCH(1,1)-GPD will always give a better fit among wavelet filters for every currency data set. As further research, we recommend the consideration of other wavelet filters such as Coiflets, Symlets, etc., and compare their performance.

Author Contributions: Conceptualisation: T.N. and D.C.; data curation: T.N.; formal analysis: T.N. and D.C.; investigation: T.N.; methodology: T.N. and D.C.; project administration: D.C.; resources: D.C.; funding acquisition: D.C.; software: T.N.; supervision: D.C.; validation: T.N.; writing—original draft: T.N.; writing—review and editing: T.N. and D.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Currency data used in this research were obtained from the finance sector website (www.investing.com/currencies, accessed on 1 July 2021).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ARIMAX	Auto-Regressive Moving Average with eXplanatory variable(s)
ARMA	Auto-Regressive Moving Average
CWT	Continuous Wavelet Transform
DWT	Discrete Wavelet Transform
EVT	Extreme Value Theory
GARCH	Generalised Autoregressive Conditional Heteroscedasticity
GEVT	Generalised Extreme Value Distribution
GPD	Generalised Pareto Distribution
MODWT	Maximal Overlap Discrete Wavelet Transform

References

- Kaseke, F.; Ramroop, S.; Mhwambi, H. A Comparison of the Stylised Facts of BitCoin, Ethereum and the JSE Stock Returns. *Afr. Financ. J.* 2021, 23, 50–64.
- Available online: https://assets.ctfassets.net/4cgqlwde6qy0/3UYrVVpyqckCsw802wWoOi/7abfe71c3b60ff521635f713865cad1 6/FX_Risk_in_Development_Primer.pdf (accessed on 18 February 2023).
- 3. Hull, J.C. Risk Management and Financial Institutions, 1st ed.; Prentice Hall: Hoboken, NJ, USA, 2006.
- 4. Daníelsson, J.; Jorgensen, B.N.; Samorodnitsky, G.; Sarma, M.; de Vries, C.G. Fat Tails, Var and Subadditivity. *J. Econom.* **2013**, 172, 283–291. [CrossRef]

- 5. Danielsson, J. Financial Risk Forecasting; Wiley: London, UK, 2011.
- 6. McNeil, A.J.; Frey, R. Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *J. Empir. Financ.* 2000, *7*, 271–300. [CrossRef]
- Karim, S.A.; Ismail, M.T.; Hasan, M.K.; Sulaiman, J. Denoising the temperature data using wavelet transform. *Appl. Math. Sci.* 2013, 7, 5821–5830. [CrossRef]
- Nazmul, A.M.; Abdul, K.M.; Mesbahul, A.M. Modeling via Wavelet GARCH Algorithm on Multivariate ENSO Index. Int. J. Sci. Res. Publ. 2019, 9, 9195. [CrossRef]
- Ismail, M.T.; Audu, B.; Tumala, M.M. Comparison of forecasting performance between Maximal Overlap Discrete Wavelet Transform-Garch(1,1) and Maximal Overlap Discrete Wavelet Transform-Egarch(1,1) models: Evidence from African stock markets. J. Financ. Data Sci. 2016, 2, 254–264. [CrossRef]
- 10. Percival, D.B.; Walden, A.T. Wavelet Methods for Time Series Analysis; Cambridge University Press: New York, NY, USA, 2000.
- 11. Aminghafari, M.; Poggi, J.M. Non-stationary Time Series Forecasting Using Wavelets and Kernel Smoothing. *Commun. Stat. Theory Methods* **2012**, *41*, 485–499. [CrossRef]
- 12. Rockafellar, R.T.; Uryasev, A.S. Conditional value-at-risk for general loss distributions. *J. Bank. Financ.* 2002, 26, 1443–1471. [CrossRef]
- 13. Chen, J.M. On Exactitude in Financial Regulation: Value-at-Risk, Expected Shortfall, and Expectiles. Risks 2018, 6, 61. [CrossRef]
- 14. Dritsaki, C. An Empirical Evaluation in GARCH Volatility Modeling: Evidence from the Stockholm Stock Exchange. *J. Math. Financ.* 2017, *7*, 366–390. [CrossRef]
- 15. Bauwens, L.; Backer, B.; Dufays, A. A Bayesian Method of Change-Point Estimation with Recurrent Regimes: Application to GARCH Models. *J. Empir. Financ.* 2014, 29, 207–229. [CrossRef]
- Khalek, M.A.; Ali, M.A. Comparative Study of Wavelet-SARIMA and Wavelet-NNAR Models for Groundwater Level in Rajshahi District. J. Environ. Sci. Toxicol. Food Technol. IOSRJESTFT 2016, 10, 2319–2399.
- 17. Wadi, A.S.; Ismail, M.T.; Karim, S.A.A. A comparison between the Daubechies wavelet transformation and the fast Fourier transformation in analyzing insurance time series data. *Far East J. Appl. Math* **2010**, *45*, 53–63.
- 18. Paul, R.K. ARIMAX-GARCH-WAVELET Model for forecasting volatile data. Model Assist. Stat. Appl. 2015, 10, 243–252. [CrossRef]
- 19. Dasman, S. Analysis of return and risk of cryptocurrency BitCoin asset as investment instrument. In *Accounting and Finance Innovations;* IntechOpen: Rijeka, Croatia, 2021. [CrossRef]
- 20. Markowitz, H.M. Portfolio Selection: Efficient Diversification of Investments; John Wiley & Sons: New York, NY, USA, 1959.
- 21. Takaishi, T. Statistical properties and multifractality of BitCoin. Phys. A Stat. Mech. Its Appl. 2018, 506, 507–519. [CrossRef]
- 22. Bouri, E.; Moln´ar, P.; Azzi, G.; Roubaud, D.; Hagfors, L.I. On the hedge and safe haven properties of BitCoin: Is it really more than a diversifier? *Financ. Res. Letters.* **2017**, *20*, 192–198. [CrossRef]
- Wątorek, M.; Drożdż, S.; Kwapień, J.; Minati, L.; Oświęcimka, P.; Stanuszek, M. Multiscale characteristics of the emerging global cryptocurrency market. *Phys. Rep.* 2021, 901, 1–82. [CrossRef]
- Drożdż, S.; Kwapień, J.; Watorek, M. What Is Mature and What Is Still Emerging in the Cryptocurrency Market? Entropy 2023, 25, 772. [CrossRef] [PubMed]
- 25. Dyhrberg, A.H. BitCoin, gold and the dollar—A Garch volatility analysis. Financ. Res. Lett. 2016, 16, 85–92. [CrossRef]
- 26. Shanaev, S.; Ghimire, B. A Fitting Return to Fitting Returns: Cryptocurrency Distributions Revisited. 2021. Available online: https://ssrn.com/abstract=3847351 (accessed on 31 March 2023).
- Chikobvu, D.; Ndlovu, T. RiskMetrics method for estimating Value at Risk to compare the riskiness of BitCoin and Rand. *Invest. Manag. Financ. Innov.* 2023, 20, 207–217. [CrossRef]
- 28. Fernandez, V. Extreme value theory and value at risk. Rev. Anál. Econ. 2003, 18, 57–85.
- 29. Schleicher, C. *An Introduction to Wavelets for Economists;* Bank of Canada: Ottawa, ON, Canada, 2002.
- 30. Benhmad, F. Bull or bear markets: A wavelet dynamic correlation perspective. Econ. Model. 2013, 32, 576–591. [CrossRef]
- 31. Rhif, M.; Ben Abbes, A.; Farah, I.R.; Martinez, B.; Sang, Y. Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review. *Appl. Sci.* 2019, *9*, 1345. [CrossRef]
- 32. Khan, A.A.; Shahidehpour, M. One day ahead wind speed forecasting using wavelets. In Proceedings of the Power Systems Conference and Exposition (PSCE'09. IEEE/PES), Seattle, WA, USA, 15–18 March 2009; pp. 1–5.
- Renaud, O.; Starck, J.L.; Murtagh, F. Wavelet-Based Combined Signal Filtering and Prediction. IEEE Trans. Syst. ManCybern. 2005, 35, 1241–1251. [CrossRef]
- 34. Zheng, G.; Starck, J.L.; Campbell, J.G.; Murtagh, F. Multiscale transforms for filtering financial data streams. *J. Comput. Intell. Financ.* **1999**, *7*, 18–35.
- Reis, A.J.R.; da Silva, A.P.A. Feature extraction via multiresolution analysis for short term load forecasting. *IEEE Trans. Power Syst.* 2005, 20, 189–198.
- 36. Haar, A. Zur Theorie der orthogonalen Funktionensysteme. Math. Ann. 1910, 69, 331–371. [CrossRef]
- Straumann, D.; Mikosch, T. Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *Ann. Statist.* 2006, 34, 2449–2495. [CrossRef]
- 38. Balkema, B.; de Haan, L. Residual lifetime at great age. Ann. Probab. 1974, 2, 792–804. [CrossRef]
- 39. Pickands, J. Statistical inference using extreme order statistics. *Ann. Stat.* **1975**, *3*, 119–131.

- 40. Fisher, R.; Tippett, L. Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Camb. Philos. Soc.* **1928**, 24, 180–190. [CrossRef]
- 41. Gnedenko, B.V. Sur la distribution limite du terme maximum of d'unesérie Aléatorie. Ann. Math. 1943, 44, 423–453. [CrossRef]
- 42. Coles, S. An Introduction to Statistical Modeling of Extreme Values; Springer: Berlin/Heidelberg, Germany, 2001. [CrossRef]
- Ndlovu, T.; Chikobvu, D. Comparing riskiness of exchange rates volatility using the Value at Risk and Expected Shortfall methods. *Invest. Manag. Financ. Innov.* 2022, 19, 360–371. [CrossRef]
- 44. Penalva, H.A.; Nunes, S.C.; Neves, M.M. Extreme Value Analysis: A brief overview with an application to flow discharge rate data in a hydrometric station in the north of Portugal. *Revstat Stat. J.* **2016**, *14*, 193–215.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.