AN ALGORITHM FOR CONDITIONAL EXTREME VALUE THEORY GARCH-EVT TECHNIQUE FOR ESTIMATING VALUE AT RISK

K.M. Sakthivel[1] and V. Nandhini[2]

•

[1]Professor, Department of Statistics, Bharathiar University, Coimbatore, Tamilnadu, India
[2]Research Scholar, Department of Statistics, Bharathiar University, Coimbatore, Tamilnadu, India
[1]sakthithebest@buc.edu.in, [2]nandhinivtvt@gmail.com

Abstract

Extreme events in financial time series are characterized by their low probability yet high impact and they pose significant challenges in financial risk management. This study aims to model and forecast extreme events, with a particular emphasis on Value at Risk (VaR) estimation. It explores the concept of conditional Extreme Value Theory (EVT) for modeling volatility series to estimate VaR by integrating Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models with EVT, forming the GARCH-EVT approach. An automated algorithm was developed to optimize both model selection and threshold determination, ensuring accurate estimation of VaR. This automated procedure enhances the model selection process by identifying the optimal GARCH model and the most appropriate EVT threshold, addressing the complexities inherent in modeling extreme events. The comprehensive backtesting procedures are used to assess the effectiveness and precision of the algorithm in forecasting VaR, along with a simulation that evaluates both in-sample and out-of-sample performance of the model and candidate thresholds across various methods. The automated GARCH-EVT approach demonstrates effectiveness in accurately estimating VaR, providing a reliable and efficient method for extreme risk assessment in financial markets. This method streamlines the process of model selection and threshold optimization, contributing to improved risk management in financial markets.

Keywords: Extreme events, Value at Risk (VaR), GARCH models, Threshold selection, Backtesting, Risk management.

I. Introduction

Extreme events in financial time series, such as sudden market crashes or dramatic price movements, pose considerable challenges for risk management strategies. These events are often rare but have significant financial consequences. To effectively manage such risks, accurate Value at Risk (VaR) estimation is critical. VaR is a standard tool for risk management, adopted by financial institutions like banks, investment funds, and corporations worldwide. VaR is determined by the quantile of the gain and loss distribution of the financial positions and it is defined as the maximum possible loss over a time horizon with a given confidence level [22]. Specifically, VaR has emerged as one of the most popular risk management methods. This may also be utilized to estimate the tail probability. The literature also emphasizes the significance of fat tails in calculating and predicting VaR [8], [28]. However, traditional VaR models, which often rely on normal distribution assumptions, may underestimate the likelihood and impact of extreme events. The limitation of this approach is evident as the assumption of normality for the underlying distribution is unrealistic. In practice, the financial data exhibit the properties of asymmetry and heavy tails. Consequently, there has been growing interest in alternative methods for VaR estimation, particularly for capturing extreme tail behavior and volatility clustering. An alternative way is a non-parametric historical simulation (HS) approach that calculates empirical quantiles from past data without assuming a specific distribution. Parametric models, such as those in the GARCH type model, the entire return distribution under conditional normality, capturing volatility dynamics. On the other hand, the extreme value approach based on VaR estimation is superior to traditional parametric and non-parametric methods in identifying extreme risk [2]. The conventional time series models often assume constant volatility, which fails to adequately account for periods of varying volatility in financial returns. This limitation can lead to misleading conclusions and ineffective risk management strategies.

To address these shortcomings, Engle [15] introduced the Autoregressive Conditional Heteroskedasticity (ARCH) model, which was later extended by Bollerslev [7] into the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model. GARCH models effectively capture essential properties of financial time series, such as volatility clustering, where large price changes tend to occur in clusters, reflecting the time-varying nature of risk. However, while GARCH models allow for dynamic volatility forecasting, they often assume symmetric responses to shocks. This limits their ability to fully capture the asymmetry typically observed in financial returns, where negative shocks have a more significant impact on volatility than positive ones known as the leverage effect. As a result, while GARCH models provide valuable insights into volatility dynamics, their limitations necessitate the exploration of more advanced models that can accommodate asymmetrical volatility behavior and better reflect the complexities of financial markets. The GARCH models with alternative distributions, such as the Student-*t* or skewed-*t*, can offer some improvement, as shown by Giot and Lauren [21]. Nevertheless, these models may still struggle to capture extreme tail events. Recently, EVT has been widely used in VaR estimation for capturing the effect of market behavior under extreme circumstances. EVT has gained popularity in risk management due to its ability to model extreme tail events, which are critical for assessing financial risk. The financial crises of the 1990s and beyond have improved interest in modeling extreme events [18]. Embrechts et al. [14], and Reiss and Thompson [30] provide a theoretical framework for EVT in the context of finance and risk management to model the behavior of extreme events. Beirlant et al. [6] discuss how extreme value models are used to capture tail behavior, while Gilli and Kellezi [19] applied EVT to stock market indices for calculating VaR. Bali [4] demonstrated that EVT outperforms traditional models, such as those based on normal and skewed-*t* distributions, in accurately estimating the VaR of financial assets. However, EVT has two key limitations: it typically assumes independent and identically distributed data, and it does not account for time-varying volatility.

McNeil and Frey [26] proposed the GARCH-EVT approach, or conditional EVT to overcome these limitations, which combines the strengths of both GARCH and EVT models. This two-stage procedure effectively captures both time-varying volatility and tail behavior. In the first stage, GARCH models are used to estimate the conditional volatility and obtain standardized residuals. In the second stage, EVT is applied to the residuals to model extreme tail events. Several studies have demonstrated the superiority of conditional EVT for VaR estimation. Bali and Neftci [3] showed that conditional EVT outperforms GARCH models with skewed distributions when applied to U.S. short-term interest rates. Marimoutou et al. [25] explore the daily Brent oil price and compare the performance of unconditional and conditional EVT models with the conventional GARCH model and historical simulation. Allen et al. [1] found that conditional EVT produced fewer violations in out-of-sample backtesting using stock indices. Karmakar and Shukla [23] confirmed the effectiveness of conditional EVT for estimating VaR for daily stock indices in six countries. By integrating time-varying volatility with extreme tail modeling, the GARCH-EVT approach offers a more accurate and robust measure of risk compared to traditional methods. Zhang et al. [33] utilized extreme value analysis to investigate the tail risk behavior of the high-frequency returns of the four most popular cryptocurrencies estimating VaR and expected shortfall with varying thresholds.

This study proposes an automated framework for Value at Risk forecasting with conditional extreme value theory. The algorithm automates key steps, including stationarity checks, ARCH effect testing, GARCH model fitting, residual distribution analysis, threshold selection for EVT, and VaR forecasting. Various GARCH models are considered to capture volatility dynamics, while EVT is applied to model extreme tail behavior. A novel dual-phase threshold (DPT) selection technique is introduced to enhance the accuracy of EVT threshold estimation. The framework generates in-sample and out-of-sample VaR forecasts, and performance is validated through backtesting using unconditional and conditional coverage tests. This automated approach provides a robust, data-driven solution for risk management by addressing both volatility clustering and extreme events. The paper is organized as follows: section 2 presents a theoretical framework of conditional extreme value theory, section 3 describes the proposed algorithmic approach for the GARCH-EVT framework, section 4 describes the data analysis of cryptocurrencies, section 5 shows the simulation results, and section 6 provides the summary and conclusion of the study.

II. Methodologies

I. Volatility Models

Volatility models are used to estimate and forecast the variance or volatility of a time series, especially in financial data like stock returns, interest rates, exchange rates, etc. Volatility is a measure of how much the price of an asset fluctuates over time and is commonly used to assess risk. Higher volatility often indicates higher risk, as it increases the likelihood of significant price changes either upward or downward. The Autoregressive Conditional Heteroskedasticity (ARCH) model is designed for modeling time-varying volatility in financial time series. It assumes that the variance of the error term (or the residuals) at time t depends on the squared values of previous error terms. This is particularly useful for capturing volatility clustering, where periods of high volatility are followed by more high volatility, and periods of low volatility are followed by more low volatility. The ARCH model is defined as $r_t = \mu + \epsilon_t$; where, r_t is the observed returns at time t, μ is the constant mean, ϵ_t is the error term or innovation. The conditional variance σ_t^2 at time t depends on past squared residuals ϵ_{t-i}^2 for $i = 1,2, \dots, q$,

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2; \ \epsilon_t \sim N(0, \sigma_t^2) \tag{1}$$

where q is the order of the ARCH model, $\omega > 0$ is the constant or intercept, $\alpha_i \geq 0$ are the ARCH coefficients, concerning the current volatility to post residuals.

The Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model extends ARCH model by including lagged conditional variances in the variance equation. It is used to analyze time-series data where the variance of the error term is assumed to be serially auto-correlated. The GARCH models are utilized when the variance of the error term changes, indicating the presence of heteroskedasticity. Let r_t be the return series, μ is the mean and ϵ_t the innovation or error term. The GARCH (p, q) model can be specified in terms of the mean and variance equation as follows

$$r_t = \mu + \epsilon_t, \ \epsilon_t = \sigma_t z_t$$

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_j \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2;\ \epsilon_t \sim N(0, \sigma_t^2) \tag{2}$$

where, $\omega > 0$ is the constant or intercept term, $\alpha_i \geq 0$ for $i = 1,2,\dots,q$ are the ARCH parameters that measure the impact of past squared innovations, $\beta_j \geq 0$ for $j = 1,2,\dots,p$ are the GARCH parameters that measure the impact of past conditional variances, and σ_t^2 is the conditional variance at time t, which is updated based on both the previous squared innovations and lagged variances. In this study, several GARCH-type specifications are considered namely the standard GARCH (SGARCH) by Bollerslev [7], Integrated GARCH (IGARCH) by Engle and Bollerslev [16], Exponentiated GARCH (EGARCH) by Nelson [27], GJR-GARCH by Glosten et al. [20], and Asymmetric Power ARCH (APARCH) by Ding et al., [13] to model the time-varying volatility.

Let r_t be the return at time t and $\epsilon_t = r_t - \mu$, where μ is the conditional mean. The standard GARCH (1,1) model is described as follows

$$r_t = \mu + \sigma_t z_t$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{3}$$

where, $\omega > 0$, $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \beta < 1$ to ensure stationarity, z_t the innovations are *iid* random variables with zero mean and unit variance, σ_t^2 is the conditional variance at time t representing the time-varying volatility, α measures the impact of past residuals ϵ_{t-1}^2 on current volatility, β measures the persistence of volatility from one period to the next. The GARCH (1,1) models tend to be more flexible, efficient, and significant than higher-order models in the out-of-sample analysis. The GARCH model converges to the Integrated GARCH model, where the long-term volatility bears an infinite process.

The IGARCH model is the special version of the SGARCH (1,1) model where the persistence parameter $\alpha + \beta = 1$, implying that volatility follows a unit root GARCH process. Thus, the conditional variance in the IGARCH (1,1) is

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + (1-\alpha)\sigma_{t-1}^2 \tag{4}$$

by taking $\beta = 1 - \alpha$ in (3) with parameter restriction $\omega > 0$, $\alpha \geq 0$, $1 - \alpha \geq 0$ respectively.

Both the SGARCH and IGARCH models assume that positive and negative shocks affect the conditional variance symmetrically. These models impose non-negative constraints on all coefficients, limiting their ability to account for the negative correlation often observed between returns and volatility. To address these limitations, certain long-memory GARCH-type models have been developed. These models are designed to capture key characteristics such as asymmetry and fat tails in return distributions, which enhance their ability to model volatility and improve the accuracy of Value-at-Risk calculations.

The Exponential GARCH (EGARCH) model allows for asymmetric effects of positive and negative shocks on volatility. The conditional variance equation is logarithmic, ensuring non-negativity without imposing parameter restriction.

$$\ln(\sigma_t^2) = \omega + \alpha\frac{\epsilon_{t-1}}{\sigma_{t-1}} + \gamma\left(\left|\frac{\epsilon_{t-1}}{\sigma_{t-1}}\right| - E\left[\frac{\epsilon_{t-1}}{\sigma_{t-1}}\right]\right) + \beta\ln(\sigma_{t-1}^2) \tag{5}$$

where, γ captures the asymmetric effect of positive and negative shocks on volatility. If $\gamma \neq 0$, then positive and negative shocks have different impacts on volatility.

The Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) model captures leverage effects, where negative shocks increase volatility more than positive shocks of the same magnitude.

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \gamma\epsilon_{t-1}^2 I(\epsilon_{t-1} < 0) + \beta\sigma_{t-1}^2 \tag{6}$$

where, $I(\epsilon_{t-1} < 0)$ is an indicator function that takes the value 1 when ϵ_{t-1} is negative and 0 otherwise; γ represents the additional impact of negative shocks on volatility.

The Asymmetric Power ARCH (APARCH) model generalizes GARCH by allowing for power transformations of the conditional standard deviations and incorporating asymmetry.

$$\sigma_t^\delta = \omega + \alpha(|\epsilon_{t-1}| - \gamma\epsilon_{t-1})^\delta + \beta\sigma_{t-1}^\delta \tag{7}$$

where, δ controls the power transformation of volatility and γ captures the asymmetry of positive shocks and negative shocks.

For every GARCH-type model, the innovation process z_t can follow one of several distributions: symmetric, skewed, or heavy-tailed distributions to better capture the characteristics of financial returns, such as symmetry, asymmetry, and fat tails. These distributions include: normal, Student's *t* distribution, skewed normal, skewed Student's t, generalized error, and skewed generalized error distribution. The parameters for all GARCH-type models can be estimated using maximum likelihood, as it is a reliable and efficient method that produces valid asymptotic standard errors in spite of non-normality. Model selection is performed using information criteria, specifically the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

II. Extreme value theory

Extreme value theory is the statistical framework for analyzing and modeling extreme events in the tail of the probability distributions. The two main approaches in EVT are block maxima and peaks over threshold approaches. In block maxima, the data is divided into non-overlapping blocks or periods of equal sizes and select the maximum value of each block, which is then modeled using generalized extreme value (GEV) distribution. The peaks over threshold (POT) approach focuses on values that exceed a specified learning threshold and then modeled using a generalized Pareto (GP) distribution. The main challenge in this framework is to select an appropriate threshold for effectively identifying extreme values. The POT method is widely recognized for its effectiveness in characterizing extreme events in the dataset. The cumulative distribution function of the GP distribution with shape parameter ξ and scale parameter σ has the following representation.

$$G_{\xi,\sigma}(y) = \begin{cases} 1 - \left(1 + \xi\left(\frac{y}{\sigma}\right)\right)^{-1/\xi} & ;\ if\ \xi \neq 0 \\ 1 - e^{-\left(\frac{y}{\sigma}\right)} & ;\ if\ \xi = 0 \end{cases} \tag{8}$$

where, i) $y \geq 0$ when $\xi \geq 0$ and $0 \leq y \leq -\sigma/\xi$ when $\xi < 0$ and ii) $\sigma > 0$ when $\xi = 0$.

The parameter ξ plays a crucial role in characterizing the tail behavior of the distribution. When $\xi = 0$, the distribution simplifies to the exponential distribution (light tail). When $\xi > 0$, the distribution follows the ordinary Pareto distribution (heavy tail). When $\xi < 0$, the distribution is characterized as a short-tailed Pareto distribution.

Let $Y_1, Y_2, \dots, Y_n$ be the excesses above the sufficient large threshold u, where $Y_i = X_i - u$. Balkema and de Haan [5] and Pickands [29] justify that $F_u(y) \approx G_{\xi,\sigma}(y)$ provided that for large u. By setting $x = u + y$, an approximation of $F(x)$, for $x > u$, can be obtained as

$$F(x) = \big(1 - F(u)\big)G_{\xi,\sigma}(y) + F(u) \tag{9}$$

and here $F(u) = \frac{n-N_u}{n}$; where n is the total number of observations, and N_u the number of observations above the threshold. By using (4) in (5), we get the tail estimator.

$$\hat{F}(x) = 1 - \frac{n}{N_u}\left(1 + \hat{\xi}\left(\frac{x-\hat{u}}{\hat{\sigma}}\right)\right)^{-1/\hat{\xi}} \tag{10}$$

where, $\hat{\xi}$ and $\hat{\sigma}$ are the estimated values obtained using the MLE.

The Value at Risk is calculated by using the (6), we get

$$\widehat{VaR}_p = u + \frac{\hat{\sigma}}{\hat{\xi}}\left(\left[\frac{n}{N_u}(1-p)\right]^{-\hat{\xi}} - 1\right) \tag{11}$$

where, u is the threshold, $\hat{\xi}$ is the estimated shape parameter and $\hat{\sigma}$ is the estimated scale parameter.

The main difficulty of modeling with the POT method is setting the right threshold. It is important to find a good balance in setting the threshold to obtain a suitable balance between the variance and the bias of the model. A high threshold reduces sample size while also increasing uncertainty. At the same time, selecting a small truncation level increases both the sample size and the bias of the results [6].

Method 1: (Threshold or Parameter Stability Method**)** The parameter stability plot, also called the threshold stability plot discussed by Coles [10] is a graphical method to study the stability of the parameter in GP distribution. This method is based on the stability property of the GP distribution. The scale parameter for a GP distribution over a threshold v where $v \geq u$ is specified as $\sigma_v = \sigma_u + \xi(v-u)$, where σ_u is the scale parameter at threshold u, and ξ is the shape parameter. If $\xi \neq 0$, the scale parameter changes as the threshold v varies. To remove the scale parameter dependence on v, it is re-parameterized as $\sigma^* = \sigma_v + \xi u$. In practice, estimates of ξ and σ^* are plotted against different thresholds v, typically with symmetric confidence intervals. The resulting plot is defined by the locus of points: $\{(u,\sigma^*); u < x_{max}\}$ and $\{(u,\xi_u); u < x_{max}\}$. The different thresholds result in different samples of peak magnitudes and times of occurrence. The threshold should be set to the lowest value for which the parameter estimates are approximately stable or constant. The parameter stability plot shows how the shape and modified scale parameters of the GP change over a range of threshold values.

Method 2: (Minimization of Asymptotic Mean Squared Error Method) The minimization of an asymptotic mean squared error (DAMSE) method is an algorithm developed by Cariro and Gomes [9] to identify the tail in data by minimizing the asymptotic mean squared error (AMSE) criterion concerning upper-order statistic k. The optimal number, k_0 corresponds to the unknown threshold u for the tail index in relation to k. The procedure works as follows: Given the observed returns $r_1, \dots, r_n$, for the tuning parameters $\tau = 0$ and $\tau = 1$, the values of $\hat{\rho}_\tau(k)$ are calculated as:

$$\hat{\rho}_\tau(k) := -\left|\frac{3\left(W_{k,n}^{(\tau)}-1\right)}{\left(W_{k,n}^{(\tau)}-1\right)}\right|, \tag{12}$$

which depend on the statistic:

$$W_{k,n}^{(\tau)} := \begin{cases} \dfrac{\left(M_{k,n}^{(1)}\right)^{\tau} - \left(M_{k,n}^{(2)}/2\right)^{\tau/2}}{\left(M_{k,n}^{(2)}/2\right)^{\tau/2} - \left(M_{k,n}^{(3)}/6\right)^{\tau/3}} & \text{if } \tau \neq 0 \\ \dfrac{ln\left(M_{k,n}^{(1)}\right)^{\tau} - ln\left(M_{k,n}^{(2)}/2\right)^{\tau/2}}{(1/2)ln\left(M_{k,n}^{(2)}/2\right) - (1/3)ln\left(M_{k,n}^{(3)}/6\right)} & \text{if } \tau = 0 \end{cases}$$

Here, $M_{k,n}^{(j)}$ is defined as: $M_{k,n}^{(j)} = \frac{1}{k}\sum_{i=1}^{k}(\log r_{n-i+1:n} - \log r_{n-k:n})^j,\ j = 1,2,3.$

To compute the optimal tail parameters:

i. Consider $K = (n^{(0.995)}, n^{(0.999)})$ and compute the median of $\hat{\rho}_\tau(k)$ denoted as K_τ,
ii. Compute $I_\tau = \sum_{k\in K}(\hat{\rho}_\tau(k) - K_\tau)^2$ for $\tau = 0,1$.
iii. Select the tuning parameter, $\tau^* = 0$, if $I_0 \le I_1$, otherwise, select $\tau^* = 1$.

Next, work with $\hat{\rho} = \hat{\rho}_{\tau^*}(k) = \hat{\rho}_{\tau^*}(k_{01})$ and $\hat{\beta} = \hat{\beta}_{\tau^*}(k) = \hat{\beta}_{\tau^*}(k_{01})$ for $k_{01} = n^{0.999}$ and the estimator $\hat{\beta}_{\hat{\rho}}(k)$ is computed as

$$\hat{\beta}_{\hat{\rho}}(k) = \left(\frac{k}{n}\right)^{\hat{\rho}} \frac{d_k(\hat{\rho})D_k(0) - D_k(\hat{\rho})}{d_k(\hat{\rho})D_k(\hat{\rho}) - D_k(2\hat{\rho})} \tag{13}$$

where, $d_k(\alpha) = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha}$, $D_k(\alpha) = \frac{1}{k}\sum_{i=1}^{k}\left(\frac{i}{k}\right)^{-\alpha} U_i$ for any $\alpha \le 0$, with the scaled line spacing or thresholds,

$$U_i = i\sum_{i=1}^{k}(\log r_{n-i+1:n} - \log r_{n-k:n}),\ 1 \le i \le k < n, n^{0.999}. \tag{14}$$

Finally, based on the estimators $\hat{\beta}$ and $\hat{\rho}$ compute: $\hat{k}_0 = \left(\frac{(1-\hat{\rho})^2 n^{-2\hat{\rho}}}{-2\hat{\rho}\hat{\beta}^2}\right)^{\frac{1}{1-2p}}$ and estimate the shape parameter $\hat{\xi} = \hat{\xi}_{k_0,n}$.

Method 3: (Dual-Phase Threshold Selection - A Proposed Method) The dual-phase threshold (DPT) method can be used to find the optimum threshold based on the two-phase procedure (Sakthivel and Nandhini, [31] and [32]). The procedure is described as follows:

Phase 1: Let $X_1, X_2, \dots, X_n$ be an independent and identically distributed random sample of size n. The non-extremes are trimmed from X and sequential testing of the hypothesis is used to select the most appropriate threshold. The null hypothesis is: $H_0^{(i)}$: The distribution of exceedances n_i above the chosen threshold follows the GP distribution. The sequence of the null hypothesis $H_0^{(1)}, H_0^{(2)}, \dots, H_0^{(k)}$ is tested using goodness of fit tests. For instance, the Kolmogorov-Smirnov (K-S) test and Cramer von Mises (CvM) test with significance level $\alpha = 0.05$ have been performed for this case. The test statistic ω_{ij} and its p-values $p_{ij} \in [0,1]$ for $i \in 1,2,\dots,k$, $j \in 1,2,\dots,l$ denotes the k hypothesis and l test criteria are evaluated. If the p-value $p_{ij} > \alpha$, then $H_0^{(i)}$ is accepted. Otherwise, it is rejected for any $p_{ij} < \alpha$ can be represented as $H_0^{(r)}; r \in 1,2,\dots,k-1$ correspond to the threshold. If $H_0^{(r)}$ is rejected, then the threshold u_r is excluded and the values below u_r are considered to be non-extremes. The refined threshold sequence $u_{r+1} < u_{r+2} < \cdots < u_k$ is tested iteratively until all the null hypotheses are accepted, indicating the exceedances follow the GP distribution. To remove the non-extremes, if both the KS and CvM test yield, $p_{ij} < \alpha$ at different thresholds, the trimming point δ is set as $\delta = \{u_i; \max((p_{CvM}, p_{KS}) < \alpha)\}$. The values $X_i < \delta$ are excluded, and only $X_i > \delta$ are used for selecting an appropriate threshold in the next phase.

Phase 2: Consider a set of threshold values, starting from the trimming point δ obtained in phase 1, as the initial threshold and evaluated up to the 99th percentile with 0.01 increments. For each threshold A_i, where $k = 1,2,\dots,m$, there exists an n_k exceedances, and the p-value for each threshold is calculated based on multiple test criteria. The decision matrix D is created from the p-values of the test criteria evaluated across the threshold range. The matrix $D = (d_{ij})_{m\times n}$ represents the performance values d_{ij} of the ith threshold against the jth criterion, where m is the number of thresholds A_i, and l is the number of test criteria C_j. The matrix D is defined as:

$$D = \begin{array}{c} \\ A_1 \\ A_2 \\ \vdots \\ A_m \end{array} \begin{array}{c} \begin{array}{cccc} C_1 & C_2 & \dots & C_l \end{array} \\ \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1l} \\ d_{21} & d_{22} & \dots & d_{2l} \\ \vdots & \vdots & & \vdots \\ d_{m1} & d_{m2} & \dots & d_{ml} \end{bmatrix} \end{array} \tag{15}$$

Here, A_j represents the threshold and C_j represents the criteria for $i = 1,2,\dots,m$ and $j = 1,2,\dots,l$. In multiple tests, the p-values can be smoothed to control the overall fluctuation rate of different test criteria. The normalized values are calculated as

$$p_{ij} = \frac{d_{ij}}{m + \sum_{i=1}^{m}\left(d_{ij}\right)^2}$$

where d_{ij} is the value of the j^{th} criterion for the i^{th} threshold, and m is the number of thresholds. The normalized decision matrix is $p = \left(p_{ij}\right)_{m\times n}$. The entropy values for each criterion can be calculated with cross-entropy defined as

$$E_j = -\sum_{i=1}^{m}\left(p_{ij}log\left(p_{ij}\right)\right) - \left(1 - \sum_{i=1}^{l} p_{ij}\right)\left(log\left[1 - \sum_{i=1}^{l} p_{ij}\right]\right) \tag{16}$$

The relative significance of each criterion is given by

$$w_j = \frac{1 - E_j}{\sum_{j=1}^{m}\left(1 - E_j\right)}$$

This is the reasonable expression of normalized weighted value, $\sum_{j=1}^{m} w_j = 1$, for $w_j \in [0,1]$. The evaluation indicator (V) can be calculated as

$$V_j = \sum_{j=1}^{m} w_j\, d_{ij} \tag{17}$$

where, w_j is the weight of each criterion d_{ij}. The best threshold is chosen as $u^* = max\left(V_j\right)$. This threshold u^* is considered to be optimal, with exceedances above it modeled using the generalized Pareto distribution. The DPT method tests the multiple thresholds, adjusts p-values to control the error rate, and selects the most appropriate threshold.

III. Conditional Extreme Value Theory

The conditional extreme value theory called GARCH-EVT was proposed by McNeil and Frey [26] integrates GARCH and EVT to estimate Value at Risk. By filtering the returns with a GARCH model, it produces an *i.i.d* suitable for the EVT technique, and it captures both conditional heteroskedasticity and extreme tail behavior. The steps for GARCH-EVT VaR estimation:

Step 1: Fit the GARCH-type model to return data by quasi-maximum likelihood. Estimate the one-step ahead forecast of μ_{t+1} and σ_{t+1} from a fitted model and extract the standardized residuals z_t.

Step 2: Consider the standardized residuals computed in step 1, and estimate the tail quantiles of the innovations using EVT. Then construct VaR: The one-step ahead VaR measures for the dynamic volatility model described earlier can be formulated as:

$$VaR_{t+1} = \mu_{t+1} + \sigma_{t+1} VaR_t(z). \tag{18}$$

The backtesting is employed to rigorously evaluate the predictive performance of the GARCH-EVT model used for VaR forecasting. To quantitatively assess the performance of the model, a series of rigorous statistical tests are employed, including the Kupiec Unconditional Coverage (UC) test, and the Christoffersen Conditional Coverage (CC) test.

IV. Rolling Window Method

In the rolling window method, the dataset is divided into overlapping segments, with each segment containing an in-sample and an out-of-sample portion. Initially, the model is trained on the in-sample data, which consists of a fixed number of observations, and the remaining data is used for out-of-sample forecasting. In this study, 80% of the data might be used for training called in-sample, and the next 20% for testing called out-of-sample. After fitting a GARCH-type model to the in-sample data, it produces one-step-ahead volatility forecasts and VaR estimates for the out-of-sample segment. Then, the window shifts forward by a set number of observations (e.g., one day), removing the earliest observations and adding new ones. The model is re-estimated with the updated in-sample data, and fresh forecasts are made for the new out-of-sample period. This process is repeated continuously, ensuring each forecast is based on previously unseen data. The rolling window approach is effective for evaluating model performance over time, as it mimics real-world forecasting scenarios and prevents over-fitting, leading to more reliable out-of-sample predictions.

III. Automated GARCH-EVT Algorithm

The automated algorithm for GARCH-EVT forecasts Value at Risk by combining GARCH-type models with various advanced threshold selection methods. The procedure is as follows:

Step 1: *Data*: Let Y_t, be the values of time series at the time $= 1,2, \dots, n$.

Step 2: *Test for Normality:* The Jarque-Bera test checks whether a time series follows a normal distribution by measuring skewness and kurtosis. A low p-value suggests non-normality, signaling potential risk from extreme events.

Step 3: *Calculate returns:* The log return series, r_t at time t is log $(r_t) = \log\left(\frac{P_t}{P_{t-1}}\right)$; where, P_t is the price at time t.

Step 4: *Stationarity Check:* The Augmented Dickey-Fuller (ADF) and Kwiatkowski-Phillips-Schmidt-Shin test (KPSS) test are used to check for stationarity in the series. If it shows stationarity then move on to step 5. Otherwise, transform the data and repeat this process.

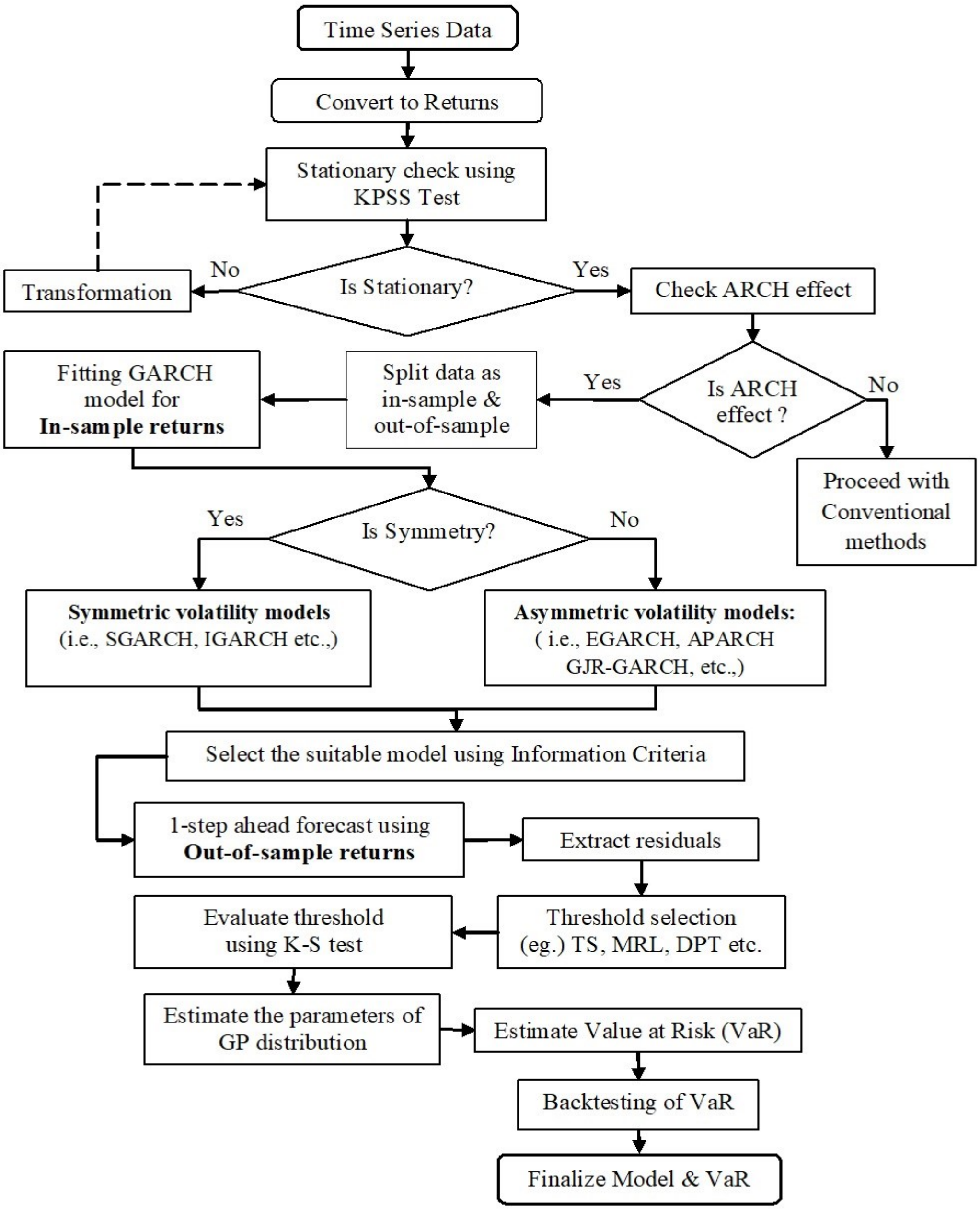
Step 5: *Check for ARCH Effect:* The ARCH-Lagrange Multiplier (ARCH-LM) test is used for testing the auto-correlation in the time series data. If there exists the ARCH effect in the series we proceed to step 6. Otherwise, end this process and proceed with conventional methods.

Step 6: *In-sample and Out-of-sample*: Fixing of In-sample and Out-of-sample proportion for rolling window procedure to obtain better model and VaR forecasting.

$$\text{In-sample: } R_{in} = r_t[1: \lfloor p.k \rfloor]$$
$$\text{Out-of-sample: } R_{out} = r_t[(\lfloor p.k \rfloor + 1): n]$$

where, p is the proportion of the data.

Schematic Representation of GARCH-EVT Algorithm for Volatility Series



Step 7: *Fitting of In-sample returns:* Set In-sample returns as $R_{in} = \{r_1, r_2, \ldots, r_k\}$. The iterative procedure through model types and residual distributions is as follows:

For each GARCH model type $M = \{m_1, m_2, \ldots, m_i\}$; $m_i \in M$ with each residual distribution is $D = \{d_1, d_2, \ldots, d_j\}$; $d_j \in D$, we implement the following procedure for optimal selection.

(i) *Specify the GARCH model:* Create the GARCH specification S_{ij} with the variance model m_i, mean model ARMA(0,0) and distribution d_j Respectively.

(ii) *Fit the GARCH model:* Fit the S_{ij} to the data Y to obtain the best-fitted model F_{ij}. Calculate AIC for F_{ij} to update the best model that is, $F_{best} = arg \min_{F_{ij}} \left(AIC(F_{ij})\right)$. If the fit fails, continue the iterative process until selecting the more suitable model.

Step 8: *Out-of-sample forecast:* The rolling window forecast W_i for $i = 1,2, \ldots, n_{out}$.

$$W_i = \{r_j | \, j = i, i+1, \dots, n_{in} + (i-1)\}$$

Fit the Out-of-sample R_{out} returns using the selected best GARCH model from step 7. Then extract residuals e_t, and conditional volatility σ_t.

Step 9: *Threshold Selection:* The threshold selection methods are $u_i = \{u_1, u_2, \dots, u_n\}$; for $i = 1, 2, \dots, n$. Fit the GP distribution to the residuals of u_i and to estimate the parameters. The CvM and K-S test can be used to evaluate the threshold-based estimates and choose the best suitable threshold selection method among u_i. The threshold selection methods used in this study are Threshold stability, DAMSE, DPT, and empirical thresholds like 90th percentile, 95th percentile.

Step 10: *Value at Risk Forecast:* The Value at Risk for one step ahead forecast from out-of-sample is defined as

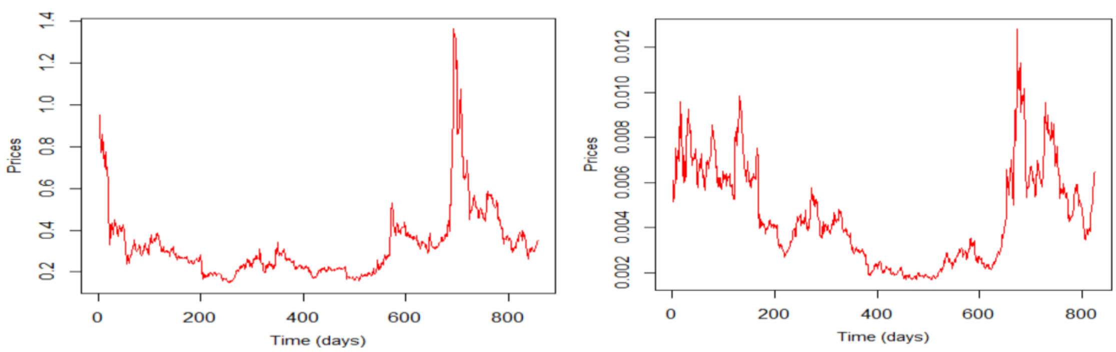
$$VaR_{t+1} = \mu_{t+1} + \sigma_{t+1} VaR_t(z_\alpha);$$

where μ_{t+1} forecasted mean returns and σ_{t+1} forecasted volatility, z_α be the quantile of GP distribution, α is the significance level.

Step 11: *Backtesting:* The Kupiec and Christoffersen test can be used for VaR backtesting. If the *p*-value of the chosen model VaR forecast is greater than the level of significance $\alpha = 0.05$ or 0.01, then finalize the GARCH EVT model. Otherwise, conventional GARCH and EVT techniques can be suitable.

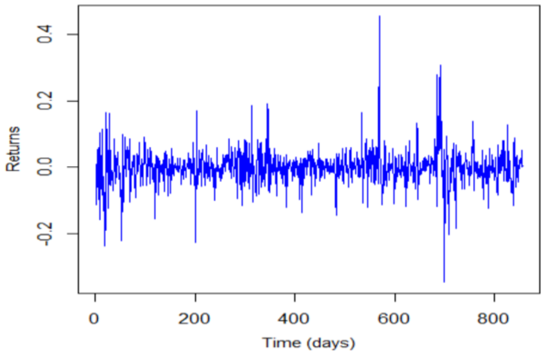
IV. Data Analysis on Real-Time Applications

In this study, the dataset consists of daily closing prices (in dollars) of two cryptocurrencies ZRX token and RSR token from 24 May 2022 to 25 August 2024 (825 observations). The data are available online at marketcap.com and the Kaggle website. Figure 1 shows the time series plots for the daily trading prices of cryptocurrencies. The sample period covers both stable and volatile phases, as well as price fluctuations and extreme jumps. The datasets of cryptocurrencies exhibit clear volatility clustering over time. A data adjustment process is used to achieve stationarity in the cryptocurrency return series, accounting for heteroskedasticity. Figure 2 shows the dynamic behavior of the log returns for all cryptocurrencies, highlighting the characteristic leptokurtosis resulting from time-varying volatility clustering, where high-volatility periods are followed by further high volatility and low-volatility periods are followed by low volatility.

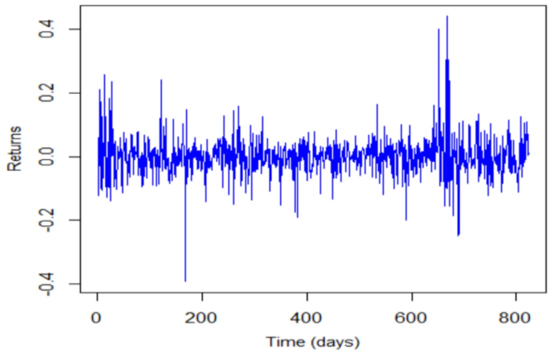


(a) *ZRX Token* (b) *RSR Token*

Figure 1: *Time series plot for the cryptocurrency dataset*



(a) *ZRX Token*



(b) *RSR Token*

Figure 2: *Return series for the cryptocurrency dataset*

Table 1 presents summary statistics for the cryptocurrencies and the results of statistical tests. The series shows excess kurtosis, indicating fat tails and non-normal distributions. Table 2 shows the JB test confirms that none of the cryptocurrencies follow a normal distribution. To assess stationarity, the KPSS test was applied, and the results rejected the null hypothesis, indicating that all return series are non-stationary at all levels. Additionally, the presence of significant ARCH effects was confirmed by using the ARCH-LM test and Box-Pierce test in cryptocurrency datasets. The results from these tests confirm the existence of significant ARCH effects in the analyzed datasets, highlighting the importance of using models that account for changing volatility in cryptocurrency datasets.

Table 1: *Descriptive statistics*

Data	Min	Q1	Median	Mean	Q3	Max	Skewness	Kurtosis
ZRX	0.1476	0.2185	0.2887	0.3289	0.3715	1.3634	2.57713	9.0602
RSR	0.0017	0.0026	0.0041	0.0045	0.0061	0.0128	0.5655	7.7823

Table 2: *Preliminary Tests*

Data	JB Test		KPSS test		ARCH-LM test		Box-Pierce test	
	χ^2	*p*-value	KPSS	*p*-value	χ^2	*p*-value	χ^2	*p*-value
ZRX	3884.4	<0.05	2.1584	<0.05	803.44	<0.05	820.28	<0.05
RSR	62.375	<0.05	2.2168	<0.05	766.85	<0.05	797.48	<0.05

The results from the estimated GARCH-type models are presented in this section. The sample period is divided into two sub-sample periods called the in-sample period; it takes 80% from the starting point and the out-of-sample period covers the last 20% of the dataset. In-sample returns are used to estimate the parameters of the selected models, subject to the assumptions and constraints of each model. The calculated in-sample parameters are applied to forecast the volatilities for both in-sample and out-of-sample periods. We first estimate the SGARCH, EGARCH, GJR-GARCH, APARCH, and IGARCH models for our dataset. Table 3 presents the AIC values of the fitted GARCH type specifications under different types of error distributions such as normal, Student's *t*, generalized error (GE), skew-normal, skew-*t*, and skew-generalized error (skew-GE) distribution. The student's *t* distribution is suitable for both datasets based on the AIC values for all the GARCH-type models. The student *t* distribution accounts for heavy tails, which allows it to capture the extreme values effectively. The estimated results of GARCH-type models with the selected innovation student's *t* distribution are presented in Table 4. The diagnostic results like minimum AIC, and BIC reveal that the IGARCH specifications for the ZRX dataset and APARCH specifications for the RSR dataset filter the serial autocorrelation, conditional volatility dynamics, and leverage effects in return series. Therefore we can apply the EVT methods to the *iid* residual series. For the ZRX dataset, we took the IGARCH-EVT approach and for the RSR dataset, we took the APARCH-EVT approach to compute the one-step-ahead Value at Risk forecast for these cryptocurrencies. The forecast performance of these types of models should be evaluated for the out-of-sample period and using more accurate performance criteria. In this study, optimal POT thresholds are obtained by evaluating the five different threshold methods as 90th percentile, 95th percentile, threshold stability (TS) method, minimization of an asymptotic mean squared error (DAMSE) method, and the proposed dual phase threshold (DPT) selection method and to estimate the GP distribution parameters for both the left and right tails.

Table 3: *In-Sample Estimated Results and Model Selection*

Models		Normal	*t*	GE	Skew-Normal	Skew-*t*	Skew-GE
Data 1: ZRX Token							
SGARCH	AIC	-3.1327	-3.3519	-3.3175	-3.1373	-3.3497	-3.3160
	BIC	-3.1063	-3.3188	-3.2844	-3.1043	-3.3100	-3.2763
EGARCH	AIC	-3.1505	-3.3491	-3.3172	-3.1531	-3.3471	-3.3159
	BIC	-3.1174	-3.3094	-3.2775	-3.1135	-3.3008	-3.2696
GJR-GARCH	AIC	-3.1299	-3.3499	-3.3149	-3.1350	-3.3475	-3.3132
	BIC	-3.0969	-3.3102	-3.2752	-3.0954	-3.3012	-3.2669
APARCH	AIC	-3.1446	-3.3473	-3.3136	-3.1453	-3.3451	-3.3122
	BIC	-3.1049	-3.3010	-3.2673	-3.0990	-3.2922	-3.2593
IGARCH	AIC	-3.1291	-3.3536	-3.3174	-3.1331	-3.3514	-3.3161
	BIC	-3.1093	-3.3272	-3.2909	-3.1067	-3.3183	-3.2830
Data 2: RSR Token							
SGARCH	AIC	-2.8775	-3.0855	-3.0593	-2.8747	-3.0852	-3.0591
	BIC	-2.8503	-3.0514	-3.0252	-2.8406	-3.0446	-3.0182
EGARCH	AIC	-2.9452	-3.0912	-3.0649	-2.9425	-3.0897	-3.0633
	BIC	-2.9111	-3.0503	-3.0240	-2.9016	-3.0420	-3.0156
GJR-GARCH	AIC	-2.9188	-3.0926	-3.0677	-2.9162	-3.0915	-3.0665
	BIC	-2.8848	-3.0517	-3.0268	-2.8754	-3.0438	-3.0188
APARCH	AIC	-2.9125	-3.0941	-3.0697	-2.9105	-3.0938	-3.0689
	BIC	-2.8716	-3.0464	-3.0220	-2.8628	-3.0392	-3.0144
IGARCH	AIC	-2.8595	-3.0824	-3.0505	-2.8578	-3.0821	-3.0508
	BIC	-2.8390	-3.0551	-3.0232	-2.8306	-3.0481	-3.0167

To evaluate the out-of-sample performance of the VaR forecast models using the EVT approach, we implemented a rolling window scheme where 80% of the data was used for in-sample fitting of the GARCH-type model, while the remaining 20% was reserved for out-of-sample forecasting. Within each rolling window, we fitted the chosen best GARCH-type model from in-sample analysis and to extract residuals based on evaluating the AIC. This selection process allowed us to extract the residuals, ensuring that the thresholds for EVT analysis were derived from the most accurate representation of the underlying volatility dynamics. The one-step-ahead VaR is calculated at 95% and 99% confidence levels, which are essential for evaluating the performance of the GARCH-EVT approach in forecasting VaR. We consider both the left and the right tail of the return distribution. The reason is that the left tail represents losses for an investor with a long position on the index, whereas the right tail represents losses for an investor being short on the index.

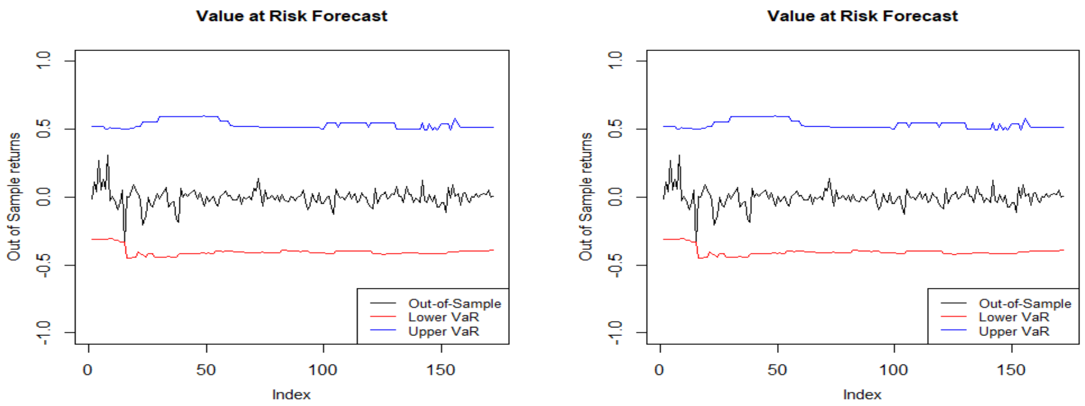
Table 4: *In-Sample: Estimated Values of the Selected Models*

Data 1: ZRX Token- Student t distribution					
Parameters	SGARCH	EGARCH	GJR-GARCH	APARCH	IGARCH
μ	0.0006 (0.0017)	0.0008 (0.0014)	0.0009 (0.0014)	0.0009 (0.0014)	0.0006 (0.0013)
ω	0.0005 (0.0001)	-0.7151 (0.2652)	0.0003 (0.0001)	0.0006 (0.0010)	0.0003 (0.0001)
α_1	0.3011 (0.0639)	0.0207 (0.0495)	0.3451 (0.1137)	0.2886 (0.0759)	0.3454 (0.0756)
β_1	0.5875 (0.0685)	0.8818 (0.0435)	0.6575 (0.0755)	0.6769 (0.0807)	0.6545 (0.0000)
γ	-	0.4267 (0.0835)	-0.0938 (0.1173)	-0.0804 (0.1023)	-
δ	-	-	-	1.7277 (0.5162)	-
Shape		4.0116 (0.5893)	3.9221 (0.5667)	3.9463 (0.5741)	3.6521 (0.4106)
log L	1076.95	1153.05	1153.34	1153.45	1153.61
AIC	-3.1327	-3.3491	-3.3499	-3.3473	-3.3536
BIC	-3.1063	-3.3094	-3.3102	-3.3010	-3.3272
$Q(5)$ (p-value)	0.8911 (0.8838)	0.7278 (0.9175)	0.7517 (0.9128)	0.7591 (0.9113)	0.7927 (0.9045)
$Q^2(5)$ (p-value)	0.2218 (0.9909)	0.2890 (0.9848)	0.3192 (0.9817)	0.3116 (0.9825)	0.3907 (0.9732)
Data 2: RSR Token - Student t distribution					
Parameters	SGARCH	EGARCH	GJR-GARCH	APARCH	IGARCH
μ	0.0011 (0.0017)	0.0018 (0.0019)	0.0017 (0.0017)	0.0015 (0.0017)	0.0013 (0.0016)
ω	0.0002 (0.0001)	-0.1385 (0.0255)	0.0001 (0.0001)	0.000001 (0.000001)	0.0001 (0.0001)
α_1	0.0798 (0.0310)	0.0753 (0.0259)	0.1013 (0.0368)	0.0067 (0.0042)	0.1095 (0.0415)
β_1	0.8623 (0.0533)	0.9759 (0.0045)	0.9164 (0.0309)	0.9307 (0.0180)	0.8904 (0.0000)
γ	-	0.1235 (0.0522)	-0.0862 (0.0345)	-0.4294 (0.1752)	-
δ	-	-	-	3.4999 (0.1193)	-
Shape	3.8293 (0.5575)	3.9752 (0.5136)	3.9360 (0.5719)	4.3291 (0.6649)	3.1781 (0.3396)
log L	1021.66	1024.55	1025.01	1026.52	1019.65
AIC	-3.0855	-3.0912	-3.0926	-3.0941	-3.0824
BIC	-3.0514	-3.0503	-3.0517	-3.0464	-3.0551
$Q(5)$ (p-value)	1.1773 (0.8184)	1.4867 (0.7432)	1.1772 (0.8184)	1.4892 (0.7425)	1.1252 (0.8307)
$Q^2(5)$ (p-value)	0.9295 (0.7540)	2.549 (0.4956)	0.9816 (0.8638)	1.3000 (0.7889)	0.7477 (0.9136)

Table 5: *Parameter estimates of the GP distribution for the selected threshold of returns*

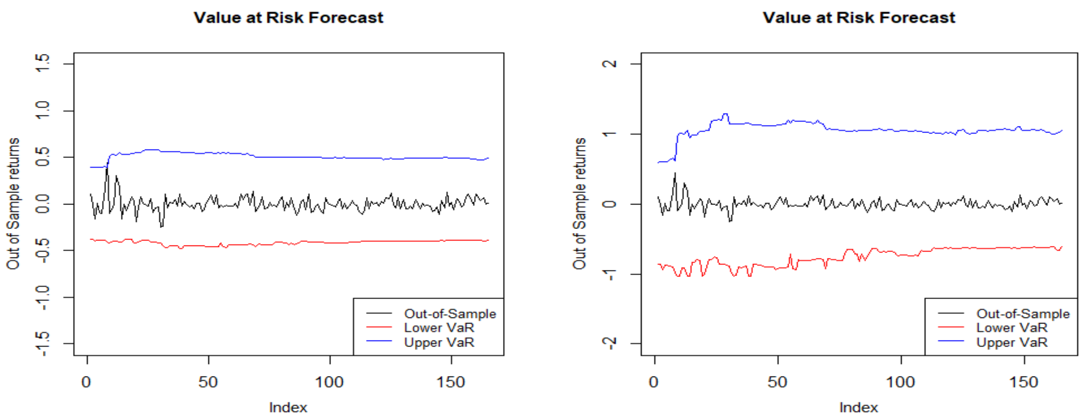
Method	Threshold (Excess)	Estimates		CvM		KS	
		Shape	Scale	Statistic	p-value	Statistic	p-value
Data 1: Left Tail							
90th Percentile	0.056 (69)	0.2346 (0.1667)	0.0307 (0.0062)	0.0849	0.6650	0.0826	0.7023
95th Percentile	0.081 (35)	0.1466 (0.1899)	0.0386 (0.0097)	0.0561	0.8419	0.1121	0.7296
TS	0.083 (34)	0.1838 (0.2043)	0.0362 (0.0096)	0.0629	0.7989	0.1151	0.6997
DAMSE	0.072 (40)	0.0696 (0.1568)	0.0445 (0.0098)	0.0612	0.8090	0.1089	0.6752
DPT	0.092 (29)	0.2990 (0.2566)	0.0304 (0.0094)	0.0327	0.9686	0.0964	0.9266
Data 1: Right Tail							
90th Percentile	0.053 (69)	0.5141 (0.1872)	0.0232 (0.0049)	0.0611	0.8086	0.0804	0.7328
95th Percentile	0.075 (35)	0.8237 (0.3523)	0.0212 (0.0077)	0.0962	0.6063	0.1233	0.6174
TS	0.074 (36)	0.7726 (0.3315)	0.0223 (0.0077)	0.0303	0.9786	0.10315	0.9848
DAMSE	0.037 (113)	0.3228 (0.1162)	0.0268 (0.0039)	0.0904	0.6346	0.0757	0.5303
DPT	0.087 (18)	0.0064 (0.2848)	0.0913 (0.0337)	0.0275	0.986	0.0963	0.9903
Data 2: Left Tail							
90th Percentile	0.062 (66)	0.1236 (0.1354)	0.0425 (0.0077)	0.0468	0.8967	0.0705	0.8756
95th Percentile	0.090 (31)	0.0476 (0.0129)	0.1308 (0.2119)	0.0306	0.9759	0.0875	0.9430
TS	0.09 (33)	0.1218 (0.2081)	0.0484 (0.0131)	0.0297	0.9784	0.0726	0.958
DAMSE	0.084 (39)	0.1681 (0.2038)	0.0433 (0.0112)	0.0479	0.8915	0.0988	0.7938
DPT	0.033 (162)	0.1903 (0.0945)	0.0318 (0.0038)	0.0164	0.9993	0.0321	0.9963
Data 2: Right Tail							
90th Percentile	0.057 (63)	0.2532 (0.1292)	0.0361 (0.0066)	0.0517	0.8673	0.0666	0.9127
95th Percentile	0.084 (33)	0.0411 (0.0107)	0.2982 (0.1999)	0.1252	0.4768	0.1609	0.3248
TS	0.084 (33)	0.3080 (0.2025)	0.0403 (0.0105)	0.0322	0.9705	0.0869	0.9518
DAMSE	0.077 (40)	0.2948 (0.1876)	0.0393 (0.0095)	0.0798	0.6953	0.1156	0.6170
DPT	0.092 (32)	0.5680 (0.2841)	0.0254 (0.0081)	0.0322	0.9705	0.0869	0.9518

Table 5 presents the estimated parameters of the GP distribution, along with standard errors and goodness-of-fit results, including the CvM and KS tests with their p-values. It displays the threshold values and excesses above the threshold for each method. The evaluation of the CvM and K-S test results shows that the excess values from the DPT threshold method yield the best fit for the GP distribution compared to alternative methods like the 95th percentile, 99th percentile, TS method, and DAMSE. Additionally, the positive shape parameter, significantly different from zero for both datasets, indicates a heavy-tailed distribution with finite variance, confirming that the tail distribution of this cryptocurrency data belongs to the Fréchet class.



(a) *95% VaR* **(b)** *99% VaR*

Figure 3: *The graph of VaR for IGARCH for the ZRS Token dataset*



(a) *95% VaR* **(b)** *99% VaR*

Figure 4: *The graph of VaR for APARCH for the RSR Token dataset*

Table 6: *Backtesting: Kupiec and Christoffersen test Results*

Level of Significance	α = 0.05 (95%)		α = 0.01 (99%)	
Tails	Left Tail	Right Tail	Left Tail	Right Tail
Data 1	IGARCH- DPT-VaR			
UC: Statistics	0.3584	3.4573	1.8685	0.3442
UC: p- value	0.5494	0.0929	0.1716	0.5574
CC: Statistics	0.3702	3.2357	1.8803	0.3528
CC: p- value	0.8310	0.1775	0.3906	0.8418
Data 2	APARCH- DPT-VaR			
UC: Statistics	0.3010	3.3166	1.8685	0.3302
UC: p- value	0.5832	0.0686	0.1716	0.5656
CC: Statistics	0.3133	3.3256	1.8803	0.3103
CC: p- value	0.8549	0.1905	0.3906	0.8478

The graphical representation of the out-of-sample alongside calculated VaR for the return series of the two datasets is in Figures 3 and 4. The x-axis represents the period over which the returns and VaR are measured and the y-axis represents the out-of-sample returns. The black middle line denotes the actual out-of-sample returns of the cryptocurrencies and the fluctuation indicates the performance of the price over time. The red line represents the lower tail VaR indicating that the value below which a certain percent of the returns are expected to fall. The blue line represents the upper tail VaR, indicating that values above which a certain percentage of returns are expected to rise. The results in Table 6 show the performance of the unconditional coverage (UC) and conditional coverage (CC) tests for both IGARCH-EVT for the ZRX dataset and APARCH-EVT for the RSR dataset at the level of significance $\alpha = 0.05$ and $\alpha = 0.01$ indicate that the models perform well in terms of VaR estimation. For both models, the UC test p-values are greater than their significance level, suggesting that the null hypothesis of correct unconditional coverage cannot be rejected. In terms of CC tests, both models yield high p-values, confirming that the model accurately captures the dynamics of the return distributions. Overall, both models corresponding to its datasets demonstrate the performance in estimating VaR concerning both UC and CC across both left and right tails.

V. Simulation Study

The simulation of returns with time-varying volatility is crucial for understanding financial dynamics, particularly in assessing risk. This process allows for the modeling of more realistic return behaviors that account for fluctuations in market conditions. We have set the parameters, the mean return $\mu = 0$, and $\sigma_0 = 1$ is the initial standard deviation. Let n be the number of observations and we have a time index $t = 1,2, \dots, n$, representing each point in time. To introduce time-varying volatility, the standard deviation is calculated at each time step is defined as

$$\sigma_t = \sigma_0 \times \left(1 + 0.5 \sin\left(\frac{2\pi t}{100}\right)\right).$$

This equation can be used to generate a standard deviation that fluctuates over time. The random returns at each time step r_t are then generated from the normal distribution, represented as $r_t \sim N(0, \sigma_t)$. In this case, the mean return $\mu = 0$, and the standard deviation σ_t changes at each time point according to the sinusoidal function. The cumulative returns $R(t)$, representing the sum of returns over time, are calculated as

$$R(t) = \sum_{i=1}^{t} r_i \,.$$

This cumulative process allows us to observe the total gain or loss of the simulated series over time. By simulating random returns with time-varying volatility, we gain insights into volatility clustering in financial markets, where large price movements tend to be followed by similar movements. This simulation is crucial for risk management and financial modeling, as it accurately reflects market behavior compared to constant-volatility models.

We generated two different samples of size n=3000, 5000 respectively. In this simulation of returns, the rolling window procedure of in-sample and out-of-sample techniques was employed to find the best VaR forecast and determine the adequacy and efficiency of the proposed automated GARCH EVT algorithm.

Table 7: *In-Sample Estimated Results and Model Selection*

Models		Normal	*t*	GED	Skew-Normal	Skew-*t*	Skew-GED
Case 1: *n*=3000							
SGARCH	AIC	-2.1057	-2.2148	-2.1938	-2.1095	-2.2169	-2.1968
	BIC	-2.0941	-2.2004	-2.1793	-2.0950	-2.1995	-2.1794
EGARCH	AIC	-2.1450	-2.2543	-2.2246	-2.1464	-2.2579	-2.2290
	BIC	-2.1305	-2.2370	-2.2072	-2.1291	-2.2377	-2.2087
GJR-GARCH	AIC	-2.1429	-2.2394	-2.2177	-2.1477	-2.2429	-2.2227
	BIC	-2.1285	-2.2220	-2.2003	-2.1304	-2.2226	-2.2025
APARCH	AIC	-2.1426	-2.2437	-2.2182	-2.1449	-2.2468	-2.2214
	BIC	-2.1253	-2.2234	-2.1980	-2.1246	-2.2237	-2.1982
IGARCH	AIC	-2.1073	-2.2160	-2.1950	-2.1110	-2.2180	-2.1980
	BIC	-2.0986	-2.2044	-2.1834	-2.0994	-2.2035	-2.1835
Case 2: *n*=5000							
SGARCH	AIC	-2.8250	-2.9005	-2.8883	-2.8361	-2.9041	-2.8926
	BIC	-2.8175	-2.8910	-2.8788	-2.8266	-2.8928	-2.8813
EGARCH	AIC	-2.8622	-2.9225	-2.9116	-2.8717	-2.9278	-2.9173
	BIC	-2.8527	-2.9111	-2.9003	-2.8603	-2.9145	-2.9040
GJR-GARCH	AIC	-2.8567	-2.9158	-2.9062	-2.8672	-2.9211	-2.9123
	BIC	-2.8473	-2.9044	-2.8948	-2.8558	-2.9078	-2.8990
APARCH	AIC	-2.8561	-2.9193	-2.9075	-2.8633	-2.9241	-2.9114
	BIC	-2.8447	-2.9193	-2.8942	-2.8500	-2.9089	-2.8962
IGARCH	AIC	-2.8262	-2.9013	-2.8892	-2.8372	-2.9050	-2.8935
	BIC	-2.8205	-2.8937	-2.8816	-2.8296	-2.8955	-2.8840

Table 7 presents the AIC values of the fitted GARCH-type specifications under different types of error distributions. The skewed student's *t* distribution is suitable for both cases based on the AIC values for all the GARCH-type models. The skewed student *t* distribution accounts for asymmetry and heavy tails, which allows it to capture the extreme values effectively. The estimated results of GARCH-type models with the selected innovation skewed student's t distribution are presented in Table 8. The residuals of the selected models are approximately *iid's* which is the requirement for the further process of applying EVT. For simulated returns, we select the EGARCH-EVT approach to compute the one-step-ahead Value at Risk forecast. The forecast performance of these types of models should be evaluated for the out-of-sample period and using more accurate performance criteria.

The estimated values of parameters of the GP distribution, including their standard errors and the results of goodness-of-fit tests, specifically the CvM and KS tests, along with their *p*-values are shown in Table 9. Our analysis of the CvM and KS test results indicates that the excess values derived from the DPT threshold method yield the best fit for the GP distribution compared to alternative methods. Furthermore, the positive shape parameter indicates that the distribution is heavy-tailed. This means that there is a higher chance of observing extreme values (very large or very small). Heavy-tailed distributions are crucial in risk assessment, particularly in finance and insurance, as they can more accurately reflect the occurrence of rare but significant events.

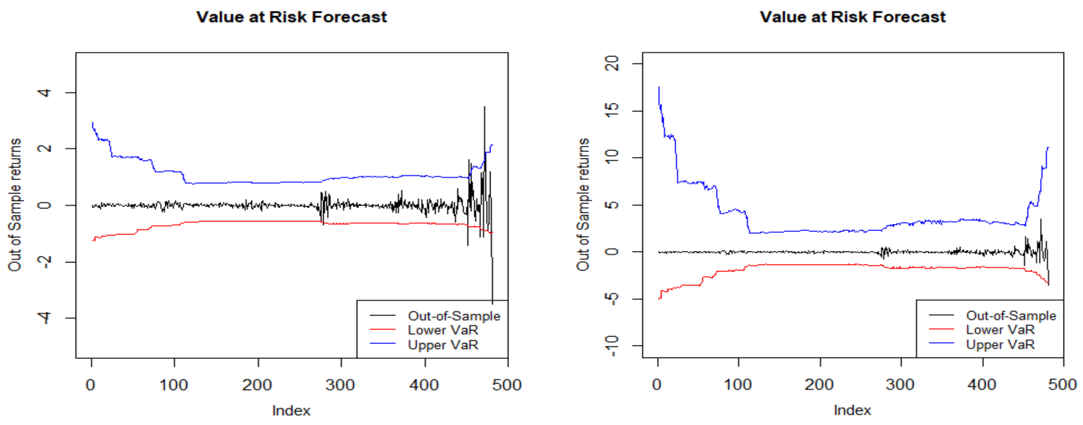
Table 8: *In-Sample: Estimated Values of the Selected Models*

Case 1: n=3000					
Parameters	SGARCH	EGARCH	GJR-GARCH	APARCH	IGARCH
μ	-0.0002 (0.0012)	-0.0029 (0.0013)	-0.0015 (0.0013)	-0.0025 (0.0013)	-0.0002 (0.0012)
ω	0.0002 (0.0001)	-0.1883 (0.0317)	0.0002 (0.0001)	0.0008 (0.0005)	0.0002 (0.00003)
α_1	0.2567 (0.0273)	-0.1584 (0.0201)	0.0832 (0.0222)	0.2383 (0.0265)	0.2577 (0.0215)
β_1	0.7422 (0.0215)	0.9624 (0.0062)	0.7739 (0.0192)	0.7924 (0.0195)	0.7422 (0.0000)
γ	-	0.3913 (0.0349)	0.2943 (0.0421)	0.4116 (0.0641)	-
δ	-	-	-	1.4110 (0.1847)	-
Skew	0.9285 (0.0282)	0.9066 (0.0296)	0.9122 (0.0284)	0.9123 (0.0295)	0.9285 (0.0283)
Shape	5.6529 (0.6078)	5.5375 (0.6304)	6.0268 (0.6845)	5.4114 (0.6179)	5.6431 (0.5857)
log L	2137.51	2178.01	2163.52	2168.31	2137.60
AIC	-2.2169	-2.2579	-2.2429	-2.2468	-2.2180
BIC	-2.1995	-2.2377	-2.2226	-2.2237	-2.2035
$Q(5)$ (p-value)	2.614 (0.4819)	2.886 (0.4282)	3.267 (0.3604)	4.206 (0.2295)	2.620 (0.4809)
$Q^2(5)$ (p-value)	1.2656 (0.7972)	5.0203 (0.1515)	4.8132 (0.1686)	42.898 (6.868e-12)	1.2502 (0.8009)
Case 2: n=5000					
Parameters	SGARCH	EGARCH	GJR-GARCH	APARCH	IGARCH
μ	-0.0002 (0.0006)	-0.0012 (0.0005)	-0.0008 (0.0006)	-0.0012 (0.0005)	-0.0002 (0.0006)
ω	0.00003 (0.00001)	-0.0975 (0.0163)	0.0001 (0.00001)	0.0003 (0.0002)	0.0001 (0.00001)
α_1	0.2159 (0.0149)	-0.1410 (0.0166)	0.1029 (0.0162)	0.2166 (0.0157)	0.2169 (0.0124)
β_1	0.7831 (0.0149)	0.9832 (0.0028)	0.7960 (0.0114)	0.8162 (0.0132)	0.7830 (0.0000)
γ	-	0.3663 (0.0232)	0.2081 (0.0279)	0.3485 (0.0547)	-
δ	-	-	-	1.4081 (0.1810)	-
Skew	0.9173 (0.0215)	0.8986 (0.0222)	0.9025 (0.0214)	0.9039 (0.0222)	0.9172 (0.0215)
Shape	6.9561 (0.6768)	6.5902 (0.7198)	7.3866 (0.7791)	6.6791 (0.7063)	6.9376 (0.6579)
log L	4651.17	4689.94	4679.25	4685.02	4651.53
AIC	-2.9041	-2.9278	-2.9211	-2.9241	-2.9050
BIC	-2.8928	-2.9145	-2.9078	-2.9089	-2.8955
$Q(5)$ (p-value)	3.8152 (0.2780)	4.4863 (0.1993)	0.4004 (0.2536)	4.4171 (0.2064)	3.8176 (0.2777)
$Q^2(5)$ (p-value)	1.5662 (0.7236)	1.6791 (0.6959)	1.6405 (0.7053)	0.9894 (0.8621)	1.5529 (0.7269)

Table 9: *Parameter estimates of the GP distribution for a selected threshold of simulated returns*

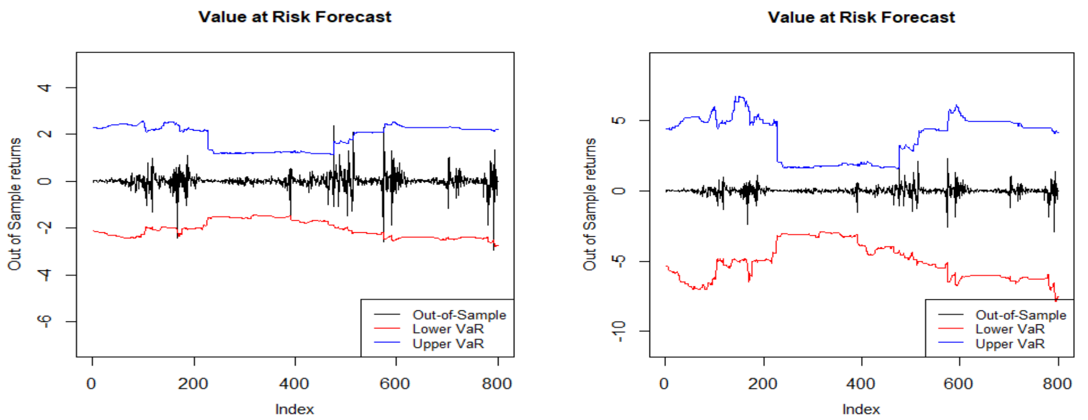
Method	Threshold (Excess)	Estimates		CVM		KS	
		Shape	Scale	Statistic	p-value	Statistic	p-value
Case 1: Left Tail							
90th Percentile	0.15 (92)	0.3934 (0.1518)	0.1104 (0.0189)	0.0672	0.7702	0.0788	0.5893
95th Percentile	0.24 (46)	0.6624 (0.2878)	0.1033 (0.0318)	0.0988	0.7224	0.3728	0.8746
TS	0.16 (90)	0.4254 (0.1494)	0.1054 (0.1868)	0.0539	0.8536	0.0747	0.6683
DAMSE	0.18 (76)	0.4169 (0.1615)	0.1153 (0.0222)	0.0758	0.7177	0.0935	0.4821
DPT	0.06 (342)	0.5004 (0.0827)	0.0508 (0.0048)	0.0181	0.9985	0.0213	0.9978
Case 1: Right Tail							
90th Percentile	0.14 (101)	0.6917 (0.1663)	0.0835 (0.0151)	0.0355	0.9555	0.0480	0.9740
95th Percentile	0.21 (51)	0.1319 (0.0362)	0.7204 (0.2565)	0.0476	0.8924	0.0809	0.8649
TS	0.15 (95)	0.6819 (0.1693)	0.0886 (0.0164)	0.0443	0.9106	0.0527	0.9415
DAMSE	0.17 (92)	0.7839 (0.2034)	0.0836 (0.0179)	0.0349	0.9583	0.0684	0.8006
DPT	0.11 (173)	0.7229 (0.1319)	0.0543 (0.0077)	0.0156	0.9995	0.0280	0.9992
Case 2: Left Tail							
90th Percentile	0.17 (156)	0.4495 (0.1235)	0.1743 (0.0248)	0.0257	0.9884	0.0325	0.9906
95th Percentile	0.31 (78)	0.3273 (0.1615)	0.2778 (0.0538)	0.0501	0.8774	0.0745	0.7509
TS	0.18 (152)	0.4516 (0.1258)	0.1757 (0.0255)	0.0279	0.9826	0.0342	0.99
DAMSE	0.23 (115)	0.3732 (0.1342)	0.2237 (0.0357)	0.0327	0.9672	0.0481	0.9511
DPT	0.14 (189)	0.4409 (0.1097)	0.1631 (0.0208)	0.0217	0.9952	0.0317	0.9912
Case 2: Right Tail							
90th Percentile	0.15 (165)	0.3371 (0.1166)	0.1869 (0.0257)	0.0970	0.6003	0.0620	0.5494
95th Percentile	0.30 (83)	0.2106 (0.1407)	0.2685 (0.0475)	0.0434	0.8929	0.0618	0.7936
TS	0.16 (160)	0.3258 (0.1170)	0.1922 (0.0267)	0.0941	0.6155	0.0642	0.5242
DAMSE	0.21 (118)	0.2254 (0.1205)	0.2460 (0.0369)	0.0569	0.8339	0.0726	0.5572
DPT	0.22 (113)	0.2062 (0.1195)	0.2567 (0.0387)	0.0345	0.9597	0.0585	0.8377

The graphical representation of the out-of-sample returns and corresponding Value at Risk for the two simulated returns series are shown in Figures 5 and 6. The black line shows that the returns exhibit some volatility, with notable fluctuations around the mean. This behavior is typical in financial markets, where returns can vary significantly over time. The red and blue lines illustrate the estimated Value at Risk levels. The area between these lines indicates the range of potential losses and gains that are considered acceptable within the specified confidence levels (lower and upper VaR). If the black line (out-of-sample returns) crosses below the red line (lower VaR), it indicates a loss exceeding the expected threshold, suggesting that the portfolio is experiencing a significant risk event. Conversely, if the black line crosses above the blue line (upper VaR), it suggests extremely positive returns, indicating potential gains exceeding expectations.



(a) 95% *VaR* **(b)** 99% *VaR*

Figure 5: *The graph of VaR for EGARCH-EVT for n=3000*



(a) 95% *VaR* **(b)** 99% *VaR*

Figure 6: *The graph of VaR for EGARCH-EVT for n=5000*

Table 10: *Backtesting: Kupiec and Christoffersen test Results*

Level of Significance	$\alpha = 0.05$ (95%)		$\alpha = 0.01$ (99%)	
Tails	Left Tail	Right Tail	Left Tail	Right Tail
Case 1: n=3000	Model: EGARCH-EVT-VaR			
UC: Statistics	0.2757	0.7944	0.4263	0.9624
UC: P- value	0.5995	0.3728	0.5138	0.3265
CC: Statistics	0.4022	0.8459	0.4263	0.9626
CC: P- value	0.8179	0.8321	0.8080	0.6180
Case 2: n=5000	Model: EGARCH-EVT-VaR			
UC: Statistics	2.4749	4.1465	1.6008	0.0463
UC: P- value	0.1156	0.0517	0.2057	0.8296
CC: Statistics	2.5152	4.1691	1.6218	0.0488
CC: P- value	0.2843	0.1244	0.4492	0.9758

The UC and CC test results are displayed in Table 10 for the EGARCH-EVT model applied to simulated returns with sample sizes of n=3000 and n=5000 at significance levels of $\alpha = 0.05$ and $\alpha = 0.01$. Specifically, the p-values from the UC tests exceed the significance levels for both sample sizes, indicating that we cannot reject the null hypothesis of correct unconditional coverage which suggests the model accurately estimates VaR. Similarly, the CC tests also yield high p-values, demonstrating that the models effectively capture the dynamics of the return distributions without overestimating or underestimating the risk. Overall, the EGARCH-EVT models show strong reliability and stability in estimating VaR, as evidenced by the favorable outcomes of both UC and CC tests across the left and right tails in the simulated datasets. We observe that the conditional EVT-based models give the best one-step-ahead VaR forecast according to the backtesting results.

VI. Conclusion

This paper developed an algorithm for the GARCH-EVT approach that allows us to model the tails of the time-varying conditional return distribution. In this study, we provide a framework to estimate and forecast the long position as well as short position VaR using this GARCH-EVT algorithm. Modeling the tail behavior of the returns is of utmost importance for both investors and policymakers. The GARCH-EVT approach is implemented in modeling the tail distribution of cryptocurrency returns and forecasting out-of-sample VaR. By employing a rolling window approach, we identified the best GARCH model through in-sample fitting, allowing us to extract reliable residuals for EVT analysis. The DPT method proved to be an effective strategy for selecting appropriate thresholds, significantly improving the fit of the GP distribution to the excess values. The evaluation of goodness-of-fit tests, such as the CvM and KS tests, further confirmed the superiority of the DPT method over alternative threshold selection approaches. Additionally, the positive shape parameter observed in the GP distribution analysis indicates the presence of heavy-tailed behavior, underscoring the potential for extreme events. The backtesting results demonstrate the suitability of the heavy-tailed GARCH EVT models in forecasting out-of-sample VaR. The dual-phase threshold selection procedure is more adaptable in threshold selection for conditional EVT, which has been proved in this paper. Our application and simulation captures the heavy-tailed behavior in daily returns and the asymmetric characteristics in distributions; we treat positive and negative returns separately. Overall, the GARCH EVT with DPT threshold provides a significant improvement in forecasting Value at Risk.

References

[1] Allen, D.E., Singh, A.K., and Powell, R. J., (2013). EVT and tail-risk modelling: Evidence from market indices and volatility series. *The North American Journal of Economics and Finance*, 26, 355-369.

[2] Aragones, J. R., Dowd, K., and Blanco, C., (2000). Extreme Value VaR. *Derivatives Week*, 7-8.

[3] Bali, T.G., and Neftci, S. N. (2003). Disturbing extremal behavior of spot rate dynamics. *Journal of Empirical Finance*, 10(4), 455-477.

[4] Bali, T.G., (2007). A generalized extreme value approach to financial risk measurement. *Journal of Money, Credit and Banking*, 39(7), 1613-1649.

[5] Balkema, A.A., and De Haan, L., (1974). Residual life time at great age. *The Annals of Probability*, 2(5), 792-804.

[6] Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. L., (2006). *Statistics of Extremes: Theory and Applications*. John Wiley and Sons.

[7] Bollerslev, T., (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307-327.

[8] Bollerslev, T., Chou, R.Y., and Kroner, K.F., (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, 52(1-2), 5-59.

[9] Caeiro, F., and Gomes, M. I., (2015). Threshold selection in extreme value analysis. *Extreme Value Modeling and Risk Analysis: Methods and Applications*, 69-87.

[10] Coles, S., (2000). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, Landon.

[11] Christoffersen, P.F., (1998). Evaluating interval forecasts. *International Economic Review*, 841-862.

[12] Davison, A.C., and Smith, R.L., (1990). Models for exceedances over high thresholds. *Journal of Royal Statistical Society Series B*, 52 (3), 393-442.

[13] Ding, Z., Granger, C.W., and Engle, R.F., (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1), 83-106.

[14] Embrechts, P., Kluppelberg, C., and Mikosch, T., (1997). *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin.

[15] Engle, R.F., (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007.

[16] Engle, R.F., and Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, *5*(1), 1-50.

[17] Fisher, R.A., Tippett, L.H.C., (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24, 180-190.

[18] Gencay, R., Selcuk, F., and Ulugulyagci, A., (2003). High volatility, thick tail and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics*, 33, 337-356.

[19] Gilli, M., and Kellezi, E., (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, 27, 207-228.

[20] Glosten, L.R., Jagannathan, R., and Runkle, D.E., (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, *48*(5), 1779-1801.

[21] Giot, P., and Laurent, S., (2004). Modelling daily value-at-risk using realized volatility and ARCH type models. *Journal of Empirical Finance*, 11(3), 379-398.

[22] Jorion, P., (2007). *Value at risk: The New Benchmark for Managing Financial Risk*. McGraw-Hill.

[23] Karmakar, M., and Shukla, G.K., (2015). Managing extreme risk in some major stock markets: An extreme value approach. *International Review of Economics & Finance*, 35, 1-25.

[24] Kupiec P., (1995). Techniques for Verifying the Accuracy of Risk Management Models. *Journal of Derivatives*, 3, 73-84.

[25] Marimoutou, V., Raggad, B., and Trabelsi, A., (2009). Extreme value theory and value at risk: application to oil market. *Energy Economics*, 31(4), 519-530.

[26] McNeil, A. J., and Frey, R., (2000). Estimation of tail related risk measure for heteroskedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7, 271-300.

[27] Nelson, D. B., (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370.

[28] Pagan, A., (1996). The econometrics of financial markets. Journal of empirical finance, 3(1), 15-102.

[29] Pickands, J., (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 3, 119-131.

[30] Reiss, R. D., and Thomas, M., (1997). *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields*. Birkhauser Verlag, Basel.

[31] Sakthivel, K. M., and Nandhini, V., (2024). An Entropy-Based Validation of Threshold Selection Technique for Extreme Value Analysis and Risk Assessment. *Lobachevskii Journal of Mathematics*, 45(4), 1633-1651.

[32] Sakthivel, K. M., and Nandhini, V. (2024). Modeling Extreme Values of Non-Stationary Precipitation Data with Effects of Covariates. *Indian Journal of Science and Technology*, 17(22), 2283-2295.

[33] Zhang, Y., Chan, S., and Nadarajah, S., (2019). Extreme value analysis of high-frequency cryptocurrencies. *High Frequency*, 2(1), 61-69.