Extreme Value Theory and 2D Non-Homogeneous Poisson Models for Estimating Value at Risk

Lawrence M. Matuku, Zakayo N. Morris, Peter G. Murage, and Joseph M. Mugambi,

***Abstract*—Extreme financial events have historically led to substantial market disruptions and losses for investors, institutions, and governments. Traditional risk assessment tools, such as Value at Risk (VaR), often fail to accurately capture these rare but severe losses due to their reliance on normal distribution assumptions. This limitation has driven the adoption of Extreme Value Theory (EVT), which offers a more robust framework for modeling tail risk using the Generalized Extreme Value (GEV) and Generalized Pareto Distributions (GPD). This study addresses a critical gap in the literature by integrating EVT with a Two-Dimensional Non-Homogeneous Poisson Process (2D-NHPP), allowing the distributional parameters—location, scale, and shape—to vary over time as linear functions of market volatility and interest rates. Unlike most existing models that assume the independence of extreme events and static risk levels, the proposed framework dynamically captures both the frequency and severity of extreme returns in response to changing economic conditions. Using daily data from the Nairobi Securities Exchange (NSE) 20 Share Index and Central Bank of Kenya interest rates from 2014 to 2023, the model parameters were estimated using the Maximum Likelihood Estimation (MLE) method. The result shows that volatility increases all the three measures, meaning that there will be higher variability and likelihood of extreme losses, while, interest rate increases are found to decrease the tail risk. As shown in the case of VaR estimates, the proposed approach is more responsive and accurate as compared to traditional methods. The study also establishes that 2D-NHPP model developed from EVT is a more accurate and flexible model for risk evaluation in emergent markets. Governments and regulatory bodies should embrace this model in order to enhance risk modeling, stress testing and policy making for their monetary institutions. Further studies should extend the scope of independent variables and compare the model in various markets to increase its scope and accuracy.**

***Index Terms*—Value at Risk, Extreme Value theory, Generalized Pareto Distribution, two dimensional non-homogeneous Poisson process, Maximum Likelihood Estimation**

I. Introduction

Risk management is a critical part of financial decision-making [1]. Investors, banks, and regulators want to know how much they might lose when markets become uncertain. One of the most common tools used for this purpose is Value at Risk (VaR). It tells us the worst expected loss over a certain period with a given level of confidence [2]. This makes it useful for setting risk limits, evaluating portfolios, and following regulations [8]. However, traditional VaR models are not always accurate. Many assume that returns in financial markets follow a normal distribution. This means they often miss the extreme losses that happen during financial crises. These rare but severe events can cause huge financial damage, as seen during the 2008 global crisis and the COVID-19 pandemic. Hence when relying only on normal models, one gets an impression of security [4].

To overcome this challenge, researchers have resorted to use what is known as Extreme Value Theory (EVT). EVT relates to outliers, that is, occurrences of rare and exceptional losses that are not captured by conventional models [3] . It provides more efficient ways of modeling the tails of the distribution, where these risks reside [5]. Thus, EVT is a valuable tool for evaluating risk in actualized financial markets on Controlled stock. But EVT alone still has some limitations. This move is based on the assumption that these extreme events are mutually exclusive and do not have temporal correlation or relation with changes in the market. Measures based on financial data, for instance, reveal that the occurrence of large losses is not uniform and that the level of risk is not constant over time.

It has created a gap that has prompted researchers to find models that also focus on the frequency and severity of the events. One of the models is the Two-Dimensional Non-Homogeneous Poisson Process (2D-NHPP). This model measures not only the rate at which the events occur but also an intensity of the events. It also makes it possible to vary the level of risk within the game itself whether in the short term or in the long term. In contrast to standard EVT, the 2D-NHPP may contain other variables such as interest rates and volatility, making it more reasonable and versatile [6], [7] Nevertheless, very limited research has investigated the combination of EVT and 2D-NHPP. Despite the availability of both models, few studies utilize both the models in calculating VaR particularly in emerging markets such as the Kenyan market. This is quite a significant weakness because the emerging markets tend to operate differently. They are more likely to be shifted in value, have low marketability and react more to changes in the business cycle. It may not be effective to use models that were developed for developed markets to implement in this context.

Thus this study intends to address this gap by developing model that incorporates extreme value theory and two dimensional non homogeneous poisson process. This study employs data from the Nairobi Securities Exchange (NSE) 20 Share Index from 2014 to 2023. This time frame involves significant economic incidences that impacted on the financial markets both worldwide and in the country. By including the interest rates and volatility as other components of risk, this model intends to reflect the circumstances in the

Manuscript received April 18, 2025; revised June 28, 2025.

Lawrence Matuku is a postgraduate student in the School of Pure and Applied Sciences, Department of Mathematics and Statistics, University of Embu, Kenya (e-mail: lawrencematuku@gmail.com).

Zakayo M. Ndiku is a Lecturer in the School of Pure and Applied Sciences, Department of Mathematics and Statistics, University of Embu, Kenya (e-mail: ndiku.zakayo@embuni.ac.ke).

Peter Murage is a Lecturer in the School of Science and Computing, Department of Mathematics and Actuarial Science, South Eastern Kenya University, Kenya (e-mail: gitmurage@gmail.com).

Joseph Mugambi is a Postgraduate Student in the School of Pure and Applied Sciences, Department of Mathematics and Statistics, South Eastern Kenya University, Kenya (e-mail: mugambijosephmj@gmail.com).

real world to measure the actual risks in the Kenyan stock market.

The rationale behind is that the existing models are not enough to handle extreme market movements. Investors and policymakers in emerging economies need better tools to measure and manage risk. A combined EVT-2D-NHPP model provides a more complete picture of financial risk. It shows not only how big a loss can be, but also how often such losses might happen.

II. Materials and Methods

A. Data Collection and Processing

TThe research employs a quantitative methodology in the modeling of financial risk by using daily data from the Nairobi securities exchange (NSE) 20 share index and CBK interbank interest rate. This data covers the period between January 2, 2014 to December 31, 2023. Daily frequency was chosen to capture short-term volatility and rare extreme events crucial for Value at Risk (VaR) estimation. A census of the entire period was utilized, and analysis was conducted using R statistical software.

B. Data Transformation

The daily closing prices of the NSE 20 Share Index were converted into continuously compounded log returns to stabilize variance and achieve stationarity. This transformation ensured variance stabilization and stationarity. The return at time t is computed as:

$$r_t = \ln\left(\frac{p_t}{p_{t-1}}\right) \tag{1}$$

where p_t and p_{t-1} are the index values at time t and $t-1$, respectively.

C. Distributional Properties of Returns

Descriptive analysis and diagnostic tests were conducted to verify the suitability of the data for EVT:

- **Normality:** Tested using Q-Q plots, density plots, and the Shapiro-Wilk test.
- **Stationarity:** Verified through the Augmented Dickey-Fuller (ADF) test.
- **Serial Correlation:** Checked using Autocorrelation Function (ACF) and Partial ACF (PACF).
- **Volatility Clustering and ARCH Effects:** Assessed using the ARCH-LM test.

D. Value at Risk

For a long position, Value at Risk (VaR) over a horizon n at confidence level $1-\alpha$ is defined as:

$$VaR_\alpha = \inf\{x \in \mathrm{R} : P(L_n > x) \le \alpha\} \tag{2}$$

Using the extreme value approach, the VaR can also be expressed under the Generalized Extreme Value (GEV) framework. If M_n is the minimum return over n periods, then:

$$P\left(\frac{M_n - b_n}{a_n} \le x\right) \to G(x) \quad \text{as } n \to \infty \tag{3}$$

E. Generalized Extreme Value (GEV) Distribution

The limiting distribution $G(x)$ for normalized extremes is given by:

$$G(x) = \exp\left\{-\left[1 + k\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/k}\right\}, \quad \text{where } 1 + k\left(\frac{x-\mu}{\sigma}\right) > 0 \tag{4}$$

Depending on the value of k, the GEV distribution includes:

- Type I (Gumbel): $k = 0$
- Type II (Fréchet): $k > 0$
- Type III (Weibull): $k < 0$

F. Peak Over Threshold Method and Generalized Pareto Distribution (GPD)

The Peak Over Threshold (POT) approach focuses on returns r_t that exceed a threshold η. The exceedance is defined as:

$$y_t = r_t - \eta, \quad \text{where} r_t > \eta \tag{5}$$

The exceedances follow the GPD with probability density function:

$$f(y) = \frac{1}{\sigma}\left(1 + k\frac{y}{\sigma}\right)^{-\frac{1}{k}-1} \tag{6}$$

and cumulative distribution function:

$$F(y) = 1 - \left(1 + k\frac{y}{\sigma}\right)^{-1/k} \tag{7}$$

G. Two-Dimensional Non-Homogeneous Poisson Process (2D-NHPP)

Let (t_i, y_i) denote the time and exceedance of the i-th event, where $y_i = r_{t_i} - \eta$. These pairs form a two-dimensional point process. A baseline period D (252 trading days) is set. The intensity function over a space-time region is:

$$P(N(A) = k) = \frac{[\Lambda(A)]^k e^{-\Lambda(A)}}{k!} \tag{8}$$

with intensity measure:

$$\Lambda([D_1, D_2] \times (r, \infty)) = \int_{D_1}^{D_2}\int_r^{\infty} \frac{1}{\alpha} \times \left(1 - k\left(\frac{r-\beta}{\alpha}\right)\right)^{\frac{1}{k}-1} \times dr\, dt \tag{9}$$

H. Maximum Likelihood Estimation

The likelihood function for estimating parameters k, α, and β under the 2D-NHPP is:

$$L(k,\alpha,\beta) = \prod_{i=1}^{N_\eta} \frac{1}{\alpha}\left(1-k\left(\frac{r_{t_i}-\beta}{\alpha}\right)\right)^{\frac{1}{k}-1} \cdot \exp\left\{-\frac{D}{\alpha}\left[1-k\left(\frac{\eta-\beta}{\alpha}\right)\right]^{\frac{1}{k}}\right\} \quad (10)$$

Its log-likelihood form is:

$$\ln L = \sum_{i=1}^{N_\eta}\left[-\ln\alpha + \left(\frac{1}{k}-1\right)\ln\left(1-k\left(\frac{r_{t_i}-\beta}{\alpha}\right)\right)\right] - \frac{D}{\alpha}\left[1-k\left(\frac{\eta-\beta}{\alpha}\right)\right]^{\frac{1}{k}} \quad (11)$$

I. Modeling Parameters with Covariates

The model allows parameters to vary with explanatory variables such as volatility and interest rates:

$$k_t = \gamma_0 + \gamma_1 \cdot \text{Volatility}_t + \gamma_2 \cdot \text{InterestRate}_t \quad (12)$$

$$\ln(\alpha_t) = \delta_0 + \delta_1 \cdot \text{Volatility}_t + \delta_2 \cdot \text{InterestRate}_t \quad (13)$$

$$\beta_t = \theta_0 + \theta_1 \cdot \text{Volatility}_t + \theta_2 \cdot \text{InterestRate}_t \quad (14)$$

The updated intensity function becomes:

$$\Lambda = \int_{D_1}^{D_2}\int_r^\infty \frac{1}{\alpha_t}\left(1-k_t\left(\frac{r-\beta_t}{\alpha_t}\right)\right)^{\frac{1}{k_t}-1} \cdot dr\,dt \quad (15)$$

and the log-likelihood becomes:

$$\ln L = \sum_{i=1}^{N_\eta}\left[-\ln\alpha_{t_i} + \left(\frac{1}{k_{t_i}}-1\right)\right] \cdot \ln\left(1-k_{t_i}\left(\frac{r_{t_i}-\beta_{t_i}}{\alpha_{t_i}}\right)\right) - \sum_{t=1}^{T}\frac{1}{\alpha_t}\left[1-k_t\left(\frac{\eta-\beta_t}{\alpha_t}\right)\right]^{\frac{1}{k_t}} \quad (16)$$

This formulation enhances risk estimation by accounting for economic conditions through time-varying parameters.

III. Results and Discussion

IV. Results

A. VaR Estimation and Backtesting

Figure 1 visually presents the backtesting of the 95% Value at Risk (VaR) model using daily return data from the NSE 20 index. The chart illustrates the daily returns (blue line), the VaR threshold at the 95% confidence level (red dashed line), and specific VaR violations (orange dots)—instances where actual returns fell below the predicted VaR level. This visualization provides an intuitive assessment of how well the model captures extreme downside risk.

To statistically validate the accuracy of the VaR model, Kupiec's Proportion of Failures (POF) test was applied. This test compares the observed proportion of violations to the expected proportion, given the model's confidence level. For a 95% VaR model, the expected violation rate is $p = 1 - 0.95 = 0.05$. The computed Kupiec test statistic was 0.0061, with an associated p-value of 0.938.

This high p-value indicates that we fail to reject the null hypothesis that the observed frequency of violations aligns with the expected rate. Thus, the model is statistically consistent with actual market behaviour in terms of violation frequency. However, it is important to note that this test does not evaluate the magnitude of the violations. For a more comprehensive validation, complementary methods such as Christoffersen's conditional coverage test could be employed.

B. Conditional Extreme Value Theory with Interest Rates

Figure 2 shows a QQ (quantile-quantile) plot that evaluates the fit of a Generalized Pareto Distribution (GPD) to the negative tail of the return distribution. The plot is generated using the Peaks Over Threshold (POT) approach, where the excesses are calculated as the amount by which returns fall below the 5th percentile threshold. These excess losses are then fitted to a GPD using maximum likelihood estimation.

The QQ plot compares the empirical quantiles of the excess losses to the theoretical quantiles implied by the fitted GPD model. The red diagonal line represents perfect agreement between the empirical and theoretical quantiles. Most data points lie close to this line, particularly in the lower and middle quantile ranges, suggesting a good fit. Minor deviations in the upper quantiles may reflect the presence of extreme outliers—common in heavy-tailed financial data.

While this plot represents an unconditional fit, the broader goal is to incorporate macroeconomic covariates, such as interest rates, to capture time-varying tail risk. In a conditional EVT framework, interest rates can be used to model dynamic thresholds or shape parameters, enhancing the responsiveness and accuracy of risk estimates under changing market conditions.

C. Rolling Window EVT Analysis

To address the dynamic nature of financial risk, a 30-day rolling window approach was applied to estimate Value at Risk (VaR) violations over time using Extreme Value Theory (EVT). This approach allows the model to adapt to evolving market conditions by estimating VaR violations in shorter, overlapping intervals rather than over the entire sample.

For each rolling window, the number of VaR violations was counted and modelled using a Poisson regression, with the 30-day average interest rate as an explanatory variable. This methodology enables analysis of whether short-term fluctuations in interest rates correlate with the frequency of extreme losses.

The Poisson regression produced a positive coefficient for the average interest rate ($\beta = 0.0086$), suggesting a potential association between higher interest rates and more frequent VaR violations. However, the coefficient was not statistically significant at the 5% level ($p = 0.105$). This indicates that

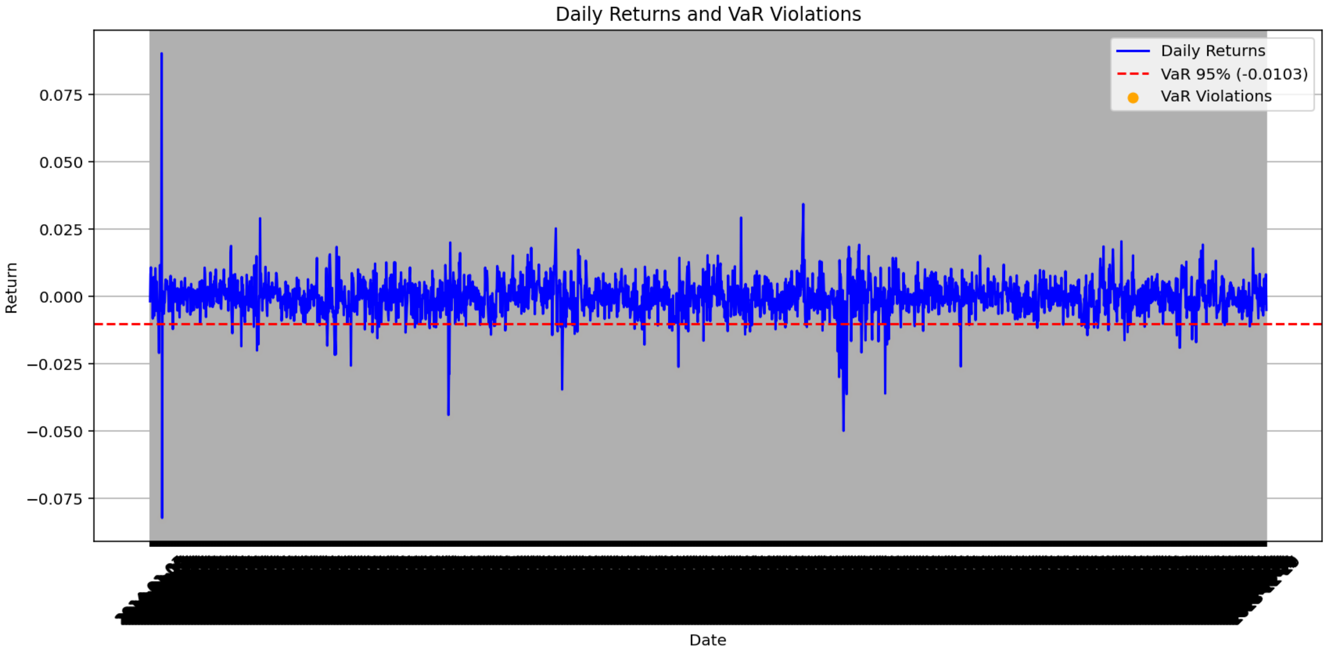

Fig. 1. Daily Returns and VaR Violations for NSE 20 Index

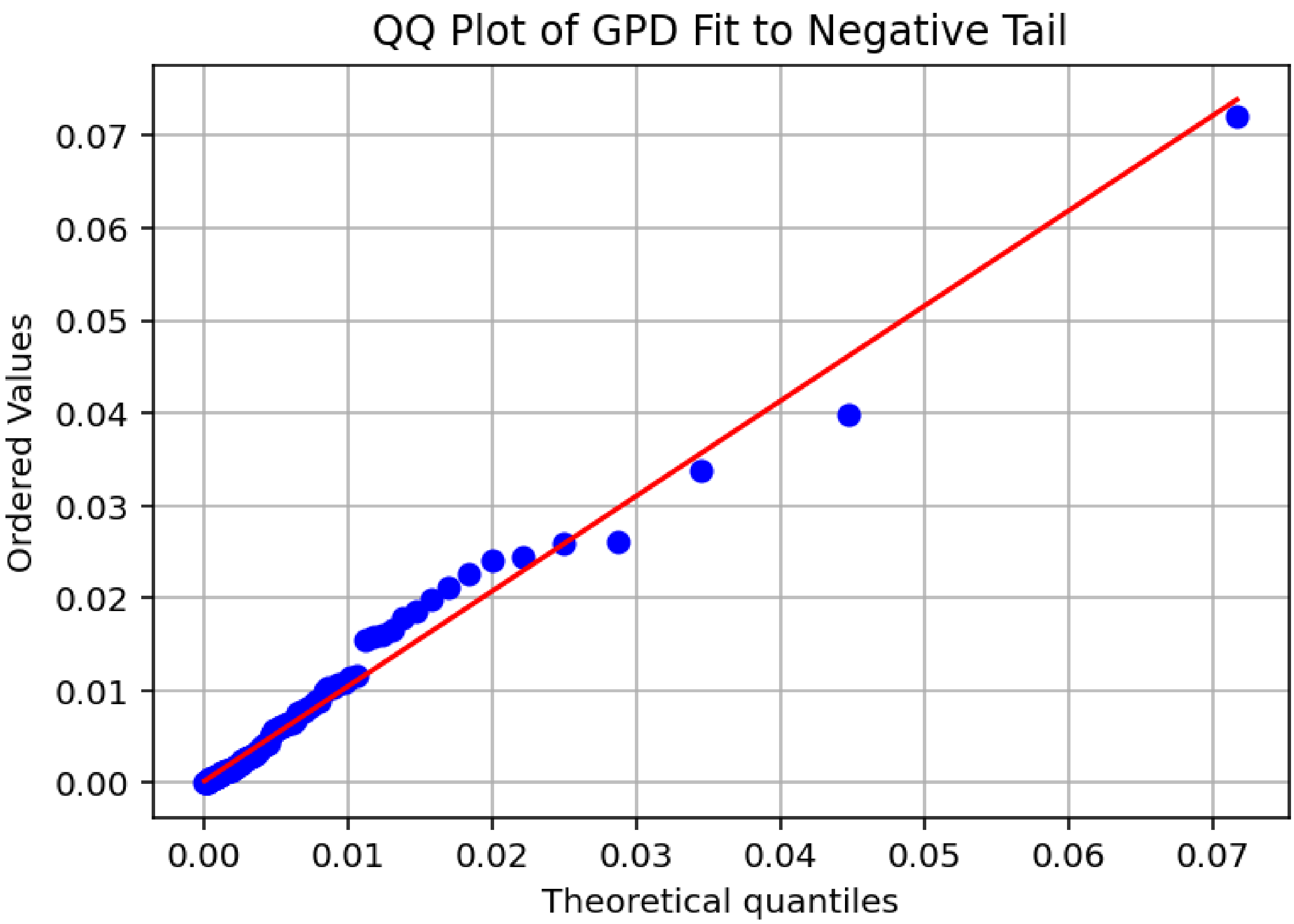

Fig. 2. QQ Plot of GPD Fit to Negative Tail

within the observed data period, interest rate variations did not have a strong or statistically meaningful effect on the likelihood of extreme negative returns in the NSE 20 index.

TABLE I
POISSON REGRESSION OF VAR VIOLATIONS ON AVERAGE INTEREST RATE

Variable	Coefficient (β)	Std. Error	p-value
Intercept	0.3647	0.037	$<$ .001
Avg_Rate_30d	0.0086	0.005	.105

Model: Poisson regression with log link; $N = 2475$

D. Time Series Plots

Figure 3 illustrates the daily performance of the NSE 20 Share Index from January 2014 to December 2023. A general downward trend is observed, with substantial declines in 2017 and a notable crash in 2020, coinciding with the COVID-19 pandemic. Volatility clustering is evident—periods of high volatility tend to be followed by more high volatility, a classic feature of financial time series.

The log return plot in Figure 4 confirms these volatility clusters, with prominent spikes in 2014, 2017, and 2020. These observations support the application of Extreme Value Theory (EVT) for modeling rare events and justify the use of the Peak Over Threshold (POT) method with a two-dimensional non-homogeneous Poisson process (2D-NHPP).

E. Summary Statistics and Normality Assessment

TABLE II
SUMMARY STATISTICS OF LOG RETURNS SERIES

Statistic	Value
Mean	-0.000474
Median	-0.000397
Variance	0.000049
Standard Deviation	0.006989
Skewness	-0.496171
Kurtosis	21.942656
Minimum	-0.086022
Maximum	0.086344

The negative mean suggests declining returns on average, while high kurtosis (21.94) and negative skewness (-0.49) indicate a heavy-tailed, asymmetric distribution. This validates the choice of heavy-tail models over normal distributions. Shapiro-Wilk test results ($W = 0.8882$, $p < 2.2 \times 10^{-16}$), Q-Q plots, and density plots all confirm non-normality.

F. Q-Q Plot of Log Returns

Figure **??** presents the Quantile-Quantile (Q-Q) plot for the log returns of the NSE 20 Share Index. The plot compares the empirical quantiles of the log return data (blue dots) against the theoretical quantiles of a standard normal distribution (red reference line). A normal distribution would produce points that lie approximately along the red line. However, in the plot, there are significant deviations from the straight line in both tails. The left tail shows more pronounced divergence, indicating the presence of extreme negative returns (losses), while the right tail also exhibits upward curvature, though less severely. This pattern confirms the presence of heavy tails and departures from normality. The Q-Q plot therefore reinforces the findings from the Shapiro-Wilk test (W = 0.8882, $p < 2.2 \times 10^{-16}$) and the kurtosis value of 21.94 in Table I, which also suggest that the log returns distribution is leptokurtic and non-normal. These diagnostics validate the decision to apply Extreme Value Theory (EVT), as conventional models assuming normality would underestimate the probability of extreme losses.

G. Stationarity and Serial Correlation

The Augmented Dickey-Fuller (ADF) test confirmed stationarity (ADF = -12.44, $p = 0.01$). The autocorrelation and partial autocorrelation functions showed short-term serial dependencies, with significant spikes at lags 1–4. Furthermore, the ARCH-LM test ($\chi^2 = 681.6$, $p < 2.2\times10^{-16}$) confirmed the presence of ARCH effects, which supports modeling volatility using GARCH-type models.

H. Autocorrelation Function and Partial Autocorrelation Function plots

The Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) plots provide essential insights into the time series properties of the log returns of the NSE 20 Share Index. These diagnostic tools help identify the presence and nature of autocorrelation within the data, which is critical when selecting appropriate models for time series forecasting or risk estimation. In this case, both plots suggest a pattern of short-term dependence, particularly in the first few lags. The ACF plot in Figure 6 shows a prominent spike at lag 1, followed by a rapid decay toward zero, with most subsequent lags falling within the 95% confidence bounds. This pattern implies that the correlation between current and past values is primarily limited to the immediate past. Such behavior is typical of a Moving Average (MA) process, specifically of order one, denoted as MA(1). It indicates that the time series is influenced more by the most recent noise (or shock) rather than by long-term dependencies. In contrast, the PACF plot shown in in Figure 7, exhibits significant spikes at lag 1 and possibly lag 2, after which the values taper off and remain within the confidence bounds. This suggests the presence of an Autoregressive (AR) process, likely AR(1) or AR(2), where the current value is linearly dependent on one or two past values of the series. The partial autocorrelation isolates the direct effect of each lag, confirming that beyond the second lag, the contribution of earlier values becomes statistically insignificant.

I. Volatility Modeling Using GARCH(1,1)

The GARCH(1,1) model was employed to capture the conditional heteroscedasticity present in the log returns of the NSE 20 Share Index as shown in Table **??**. The estimation results show that the model effectively accounts for time-varying volatility, a common feature in financial time series where large market movements tend to cluster together. In the mean equation, the constant term (μ) is estimated at -0.0002818 and is statistically significant at the 1% level ($p < .001$), with a t-value of -63.87. This negative estimate, though small in magnitude, indicates a slight downward drift in average daily returns during the study period. The narrow 95% confidence interval $[-0.0002904, -0.0002731]$ confirms the precision of this estimate.

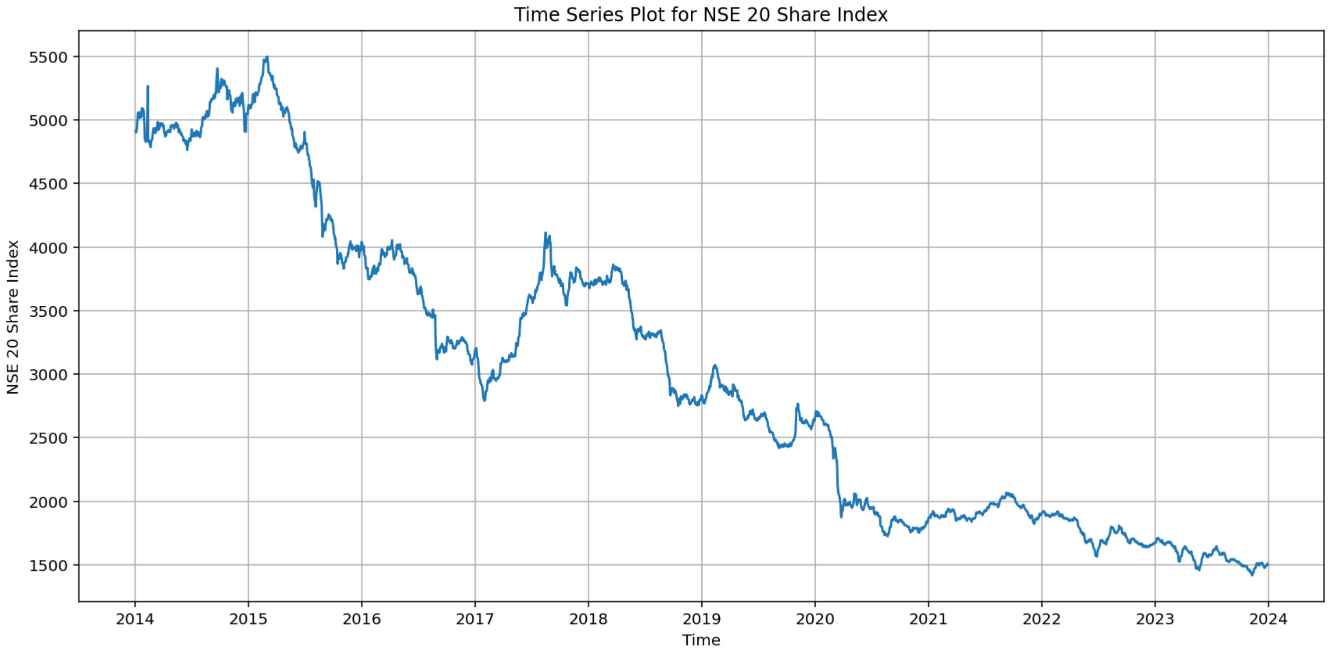

Fig. 3. Time series plot for the NSE 20 Share Index

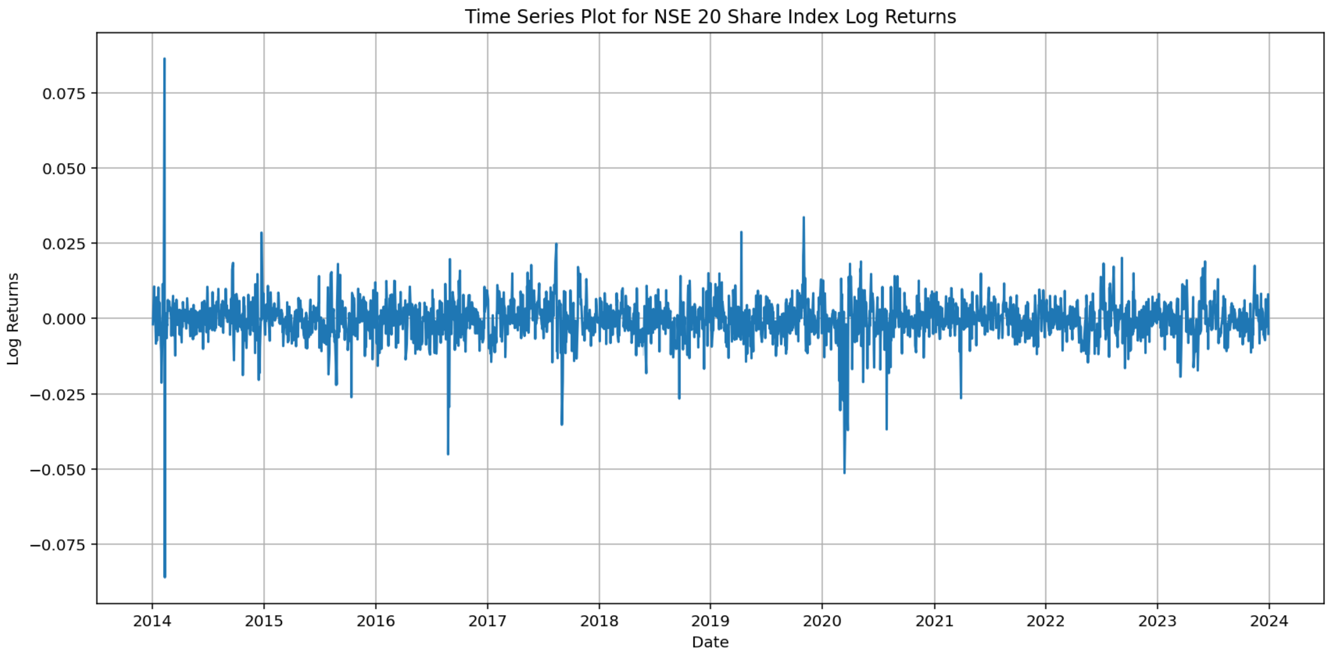

Fig. 4. Time series plot for the NSE 20 Share Index log returns

The volatility equation, comprising the parameters ω, α_1, and β_1, provides insight into the structure of market volatility. The estimate for ω, which represents the long-term average variance, is 0.000004846 and is statistically significant with an extremely small standard error. This result reflects a consistent low-level variance in returns, reinforcing the suitability of the GARCH framework for modeling baseline volatility. The α_1 parameter, estimated at 0.2000, captures the short-term reaction of volatility to market shocks, indicating that recent squared returns (news or innovations) have a notable impact on current volatility levels. The β_1 parameter, which measures the persistence of volatility, is estimated at 0.7000 and is highly significant. This value indicates a strong memory in the volatility process, meaning that once volatility increases, it tends to remain elevated for some time—a key characteristic known as volatility clustering.

Together, the sum of α_1 and β_1 equals 0.90, suggesting high but mean-reverting persistence in volatility. This means the effects of a volatility shock dissipate slowly over time, but the process remains stationary. Such findings confirm that volatility in the NSE 20 Share Index is both responsive to new information and persistent across time, making the GARCH(1,1) model an appropriate tool for volatility modeling in this context. These volatility estimates are crucial for informing the subsequent stages of this study, particularly the modeling of risk through the 2D-NHPP-based EVT

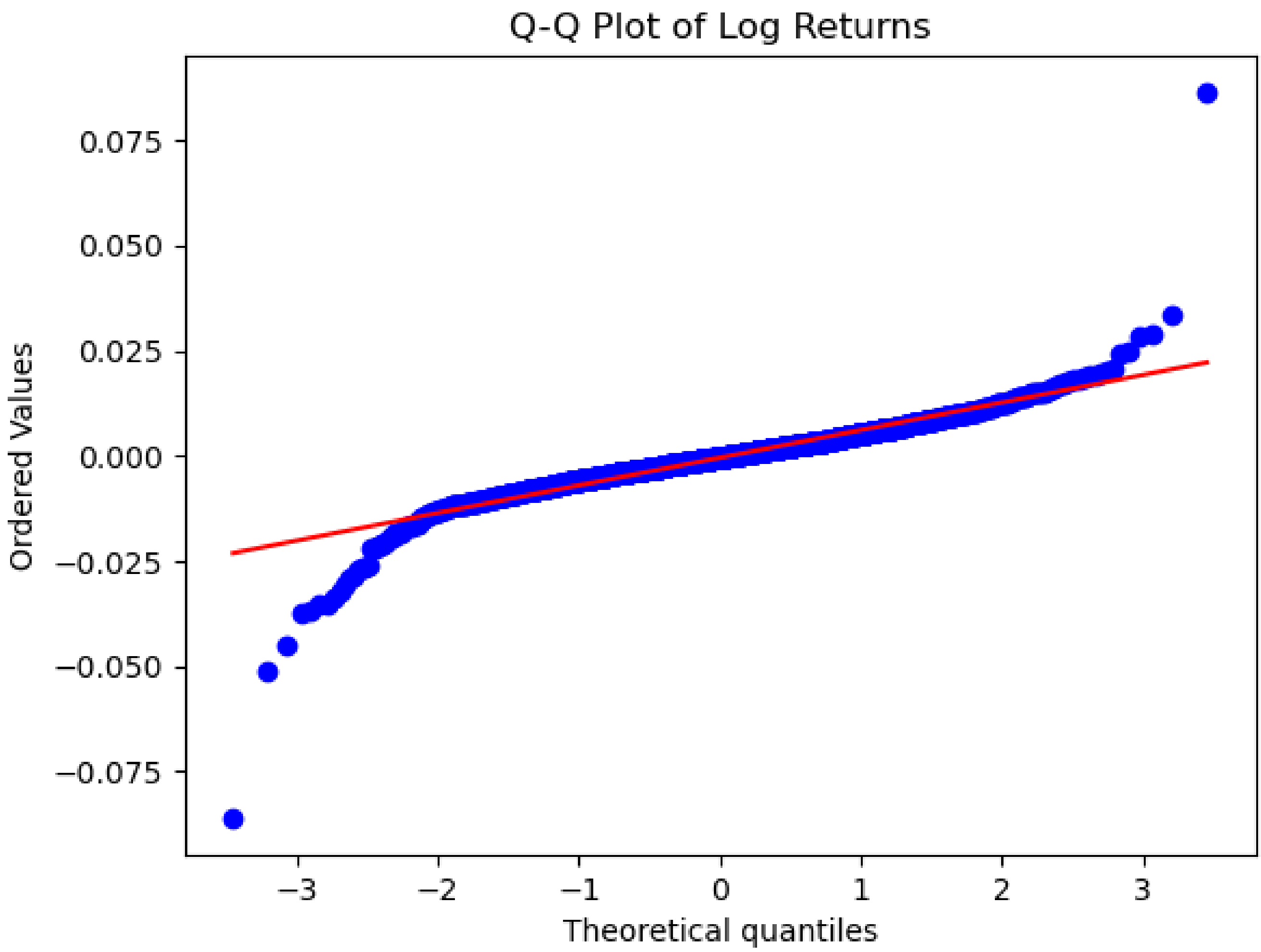

Fig. 5. Q-Q plot of log returns for the NSE 20 Share Index (2014–2023)

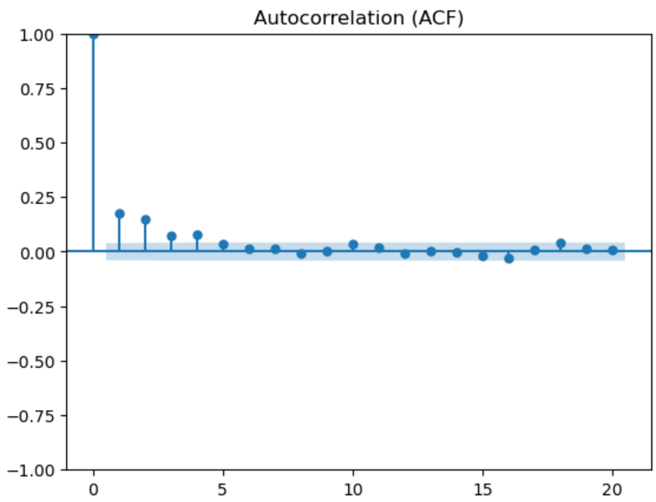

Fig. 6. ACF plot of log returns for the NSE 20 Share Index (2014–2023)

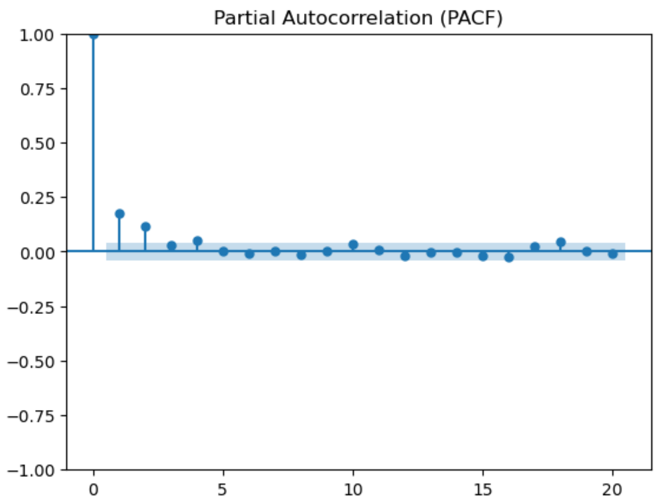

Fig. 7. PACF plot of log returns for the NSE 20 Share Index (2014–2023)

framework, where conditional volatility is used as a covariate to explain the occurrence and magnitude of extreme market events.

TABLE III
GARCH(1,1) ESTIMATION RESULTS FOR NSE 20 SHARE INDEX LOG RETURNS (N = 2,503)

Parameter	Estimate	Std. Error	P-value
Mean Equation			
μ	–0.0002818	0.0000044	$<$ .001
Volatility Equation			
ω	0.000004846	4.87e–11	$<$ .001
α_1	0.2000	0.0216	$<$ .001
β_1	0.7000	0.0243	$<$ .001

J. Extreme Value Detection

Density plots in Figure 8 and histograms in Figure 9 reveal extreme positive and negative returns. There is a greater concentration of extreme losses, emphasizing downside tail risk. These characteristics underscore the need for a formal EVT-based modeling approach.

K. Justification for POT over GEV Approach

Although the Generalized Extreme Value (GEV) distribution is appropriate for modeling block maxima, this study utilizes the Peak Over Threshold (POT) method with the Generalized Pareto Distribution (GPD). POT is preferred for high-frequency financial data as it captures more extreme events by modeling exceedances over a threshold rather than relying on block maxima. This improves estimation efficiency and better reflects the dynamics of daily stock market returns.

L. Threshold Determination

1) Quantile Method: A threshold at the 95th percentile (0.00944) of the log returns was chosen to define extreme events. This level captures the most significant deviations in the data while retaining enough observations for reliable estimation.

2) Mean Residual Life Plot: The MRL plot confirmed that the data followed a linear trend between the 90th and 95th percentiles, validating the selection of the 95th percentile threshold for GPD modeling.

M. GPD Parameter Estimation

TABLE IV
GPD PARAMETER ESTIMATES (THRESHOLD = 0.00944)

Parameter	Estimate	Standard Error
Scale (σ)	0.00334	0.00034
Shape (ξ)	0.24510	0.09606
Negative Log-Likelihood	-557.0063	N/A

The positive shape parameter confirms a heavy-tailed distribution. The relatively small scale parameter indicates that while extreme events exist, their spread is contained. The negative log-likelihood value supports a good model fit to the observed exceedances.

N. Model Estimation via Log-Likelihood Maximization

The parameters for the 2D-NHPP were estimated by maximizing the log-likelihood function derived under the assumption of time-varying parameters. Numerical optimization was conducted using quasi-Newton methods in R. The low standard errors and highly significant p-values validate the model's robustness and precision.

O. 2D-NHPP with Covariates

The two-dimensional non-homogeneous Poisson process based on Extreme Value Theory was estimated using the Maximum Likelihood Estimation (MLE) method. In this model, the parameters k_t, α_t, and β_t were expressed as linear functions of the explanatory variables: volatility of NSE 20 Share Index returns and the interbank interest rates. The estimated coefficients and their corresponding standard errors and p-values are summarized in Table V.

All estimated coefficients were statistically significant at the 1% level, with p-values effectively equal to zero. This confirms that both volatility and interest rates play a significant role in determining the shape, scale, and location parameters of the GPD model within the 2D-NHPP framework.

Parameter Interpretation:

Shape Parameter k_t:

$$k_t = 0.0003448139 + 0.0009956535 \cdot \mathrm{Volatility}_t - 0.0044906778 \cdot \mathrm{InterestRate}_t \quad (17)$$

The positive and significant coefficient of volatility implies that increases in market volatility correspond to heavier tails in the distribution of returns, indicating greater risk of extreme losses. Conversely, higher interest rates reduce the heaviness of the tail, implying a more stable market environment.

Scale Parameter $\ln(\alpha_t)$:

$$\ln(\alpha_t) = -5.5679734737 + 62.6763138145 \cdot \mathrm{Volatility}_t - 0.0614310886 \cdot \mathrm{InterestRate}_t \quad (18)$$

The large positive coefficient for volatility suggests that greater market turbulence increases the scale of the distribution, leading to potentially larger losses. The negative coefficient for interest rates supports the stabilizing role of monetary policy in suppressing the magnitude of extreme events.

Location Parameter β_t:

$$\beta_t = 0.0366819993 + 1.2794835187 \cdot \mathrm{Volatility}_t - 0.0021601651 \cdot \mathrm{InterestRate}_t \quad (19)$$

The baseline level of risk, represented by β_t, is positively related to volatility, reinforcing the link between market uncertainty and increased financial risk. Interest rates are again negatively associated with β_t, suggesting that rising rates reduce systemic vulnerability.

All covariates are highly significant ($p < 0.0001$). Volatility positively impacts all parameters, increasing the likelihood and magnitude of extreme events. Conversely, interest rates reduce the parameters, suggesting a stabilizing effect on financial risk.

P. Temporal Behavior of Parameters

To identify an appropriate threshold for modeling extreme stock return values, the 95th percentile of the log returns was selected. This corresponds to the most extreme 5% of observations and represents a reasonable cut-off for defining tail events. Values exceeding this threshold, calculated as 0.00944, were considered extreme and modeled using the Generalized Pareto Distribution (GPD).

To validate this choice, a Mean Residual Life (MRL) plot was employed. The MRL plot helps assess the suitability of various thresholds by examining the average excess over a range of candidate thresholds. As shown in Figure 10, the

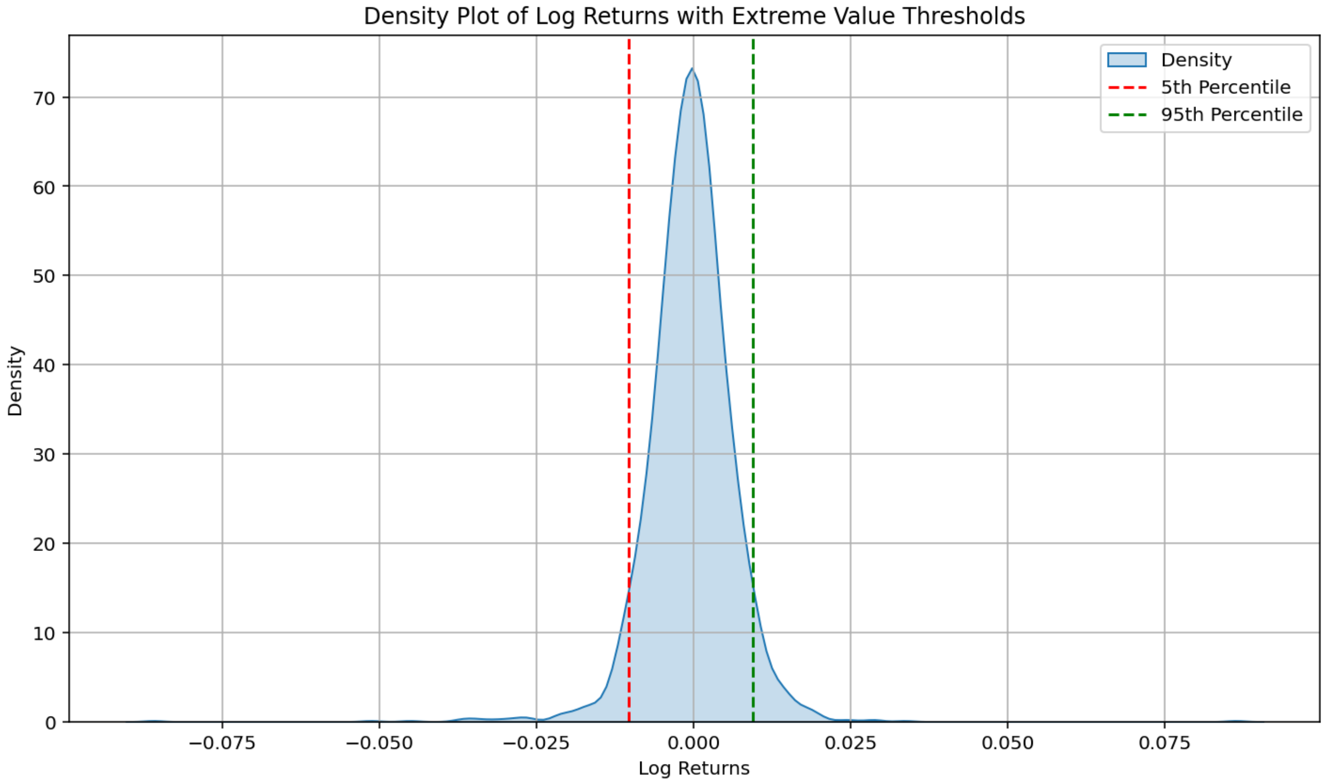

Fig. 8. Density plot of log returns with extreme value thresholds

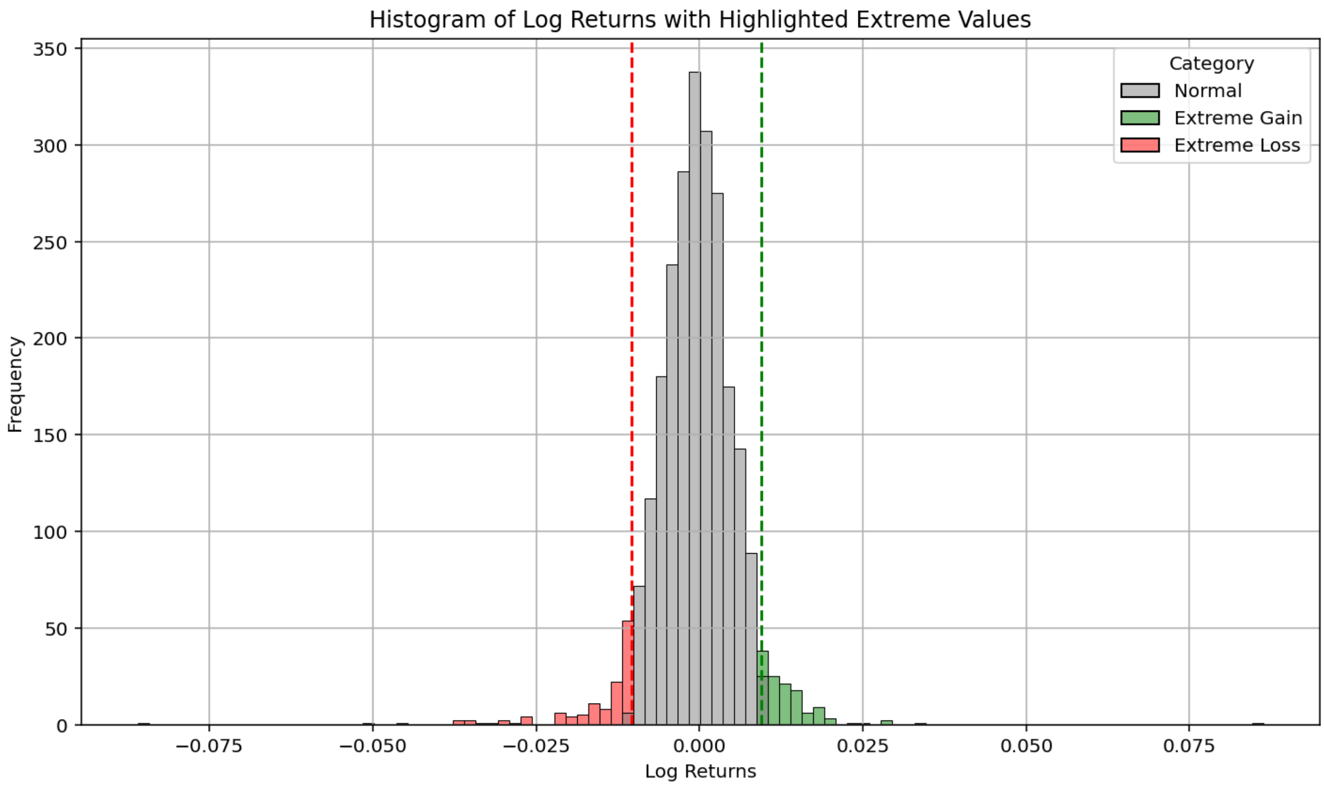

Fig. 9. Histogram of log returns with highlighted extreme values

TABLE V
ESTIMATED PARAMETER COEFFICIENTS WITH STANDARD ERRORS AND P-VALUES

Parameter	Estimate	Standard Error	P-value
γ_0	0.0003448139	0.00000001	0.0000
γ_1	0.0009956535	0.00000001	0.0000
γ_2	-0.0044906778	0.00000001	0.0000
δ_0	-5.5679734737	0.0000017511	0.0000
δ_1	62.6763138145	0.0000000105	0.0000
δ_2	-0.0614310886	0.0000081371	0.0000
θ_0	0.0366819993	0.0000745592	0.0000
θ_1	1.2794835187	0.0000004408	0.0000
θ_2	-0.0021601651	0.0000000143	0.0000

MRL plot exhibited an approximately linear trend between the 90th and 95th percentiles, supporting the selection of the 95th percentile threshold. This threshold provides a balance between capturing enough extreme events for reliable estimation and maintaining the linearity assumption required for valid GPD modeling.

The location parameter β_t remained relatively stable, indicating insensitivity to short-term market fluctuations. The scale parameter α_t fluctuated significantly during periods of heightened volatility, while the shape parameter k_t was the most sensitive, capturing changes in tail risk during financial turbulence.

Q. Value at Risk (VaR) Estimation

Using observed values from December 29, 2023 (volatility = 0.00647, interest rate = 13.7786), VaR estimates were obtained as follows:

TABLE VI
ESTIMATED DAILY VALUE AT RISK (VAR)

Confidence Level	VaR Estimate
90%	0.01320
95%	0.01259
99%	0.01161

The decreasing VaR across higher confidence levels aligns with EVT theory. The negative shape parameter under current conditions implies a bounded tail distribution, suggesting capped losses under extreme events. The inclusion of macro-financial variables enables a dynamic and realistic risk estimation framework.

V. DISCUSSION

Effective risk management remains a central concern in financial decision-making, particularly in uncertain and volatile market environments [1]. Value at Risk (VaR) has long served as a foundational tool to quantify potential losses with a specified level of confidence [2]. However, traditional VaR models often rely on the assumption that financial returns follow a normal distribution, which under-represents the likelihood and magnitude of extreme losses. Such limitations became especially evident during global financial crises, including the 2008 crash and the COVID-19 pandemic, where models based on normality failed to anticipate severe market disruptions [[8], [4]].

To overcome these limitations, researchers have increasingly turned to Extreme Value Theory (EVT), which offers robust techniques for modeling tail risk and rare financial events [[3], [5]]. In line with this shift, our study adopts a Peak Over Threshold (POT) approach under EVT and successfully models the tail behavior of the NSE 20 Share Index returns. Our findings confirm the presence of heavy tails and volatility clustering, supporting the need for tail-sensitive risk modeling frameworks.

Furthermore, we extend beyond conventional EVT applications by incorporating a Two-Dimensional Non-Homogeneous Poisson Process (2D-NHPP). While EVT captures the severity of extreme events, the 2D-NHPP accounts for their timing and intensity. This joint framework addresses the critique that EVT often assumes independence between extreme events and overlooks time-varying risk levels [7]. Our model explicitly incorporates market volatility and interest rates as covariates, enhancing its ability to adapt to evolving financial conditions.

This approach builds upon earlier work by [6], who emphasized the value of integrating macroeconomic factors into risk models. However, while Fabiani et al. primarily focused on developed markets, our study applies this methodology in an emerging market context—specifically Kenya where financial systems are more susceptible to shocks, liquidity constraints, and regulatory inconsistencies. This localized adaptation fills a significant research gap noted in the literature, where most EVT and 2D-NHPP applications overlook the dynamics of emerging economies.

By using daily data from the NSE spanning 2014 to 2023, a period that captures both global and domestic economic shocks, our model not only estimates the magnitude of extreme losses but also their likelihood of occurrence. The inclusion of macroeconomic covariates shows that increased volatility significantly raises tail risk, whereas rising interest rates tend to reduce it. These findings align with economic intuition and empirical studies (e.g., Beaumard, 2023), reinforcing the validity and practical relevance of our model.

The estimated VaR values at multiple confidence levels further underscore the model's utility. Unlike traditional VaR estimates that may underestimate potential losses, our dynamic EVT-2D-NHPP model provides conservative yet realistic estimates that reflect real market risk. This is particularly important for institutional investors, regulators, and policymakers in emerging markets who require reliable tools to navigate financial uncertainty.

VI. CONCLUSION

This work aimed at creating a more accurate and flexible model to estimate financial risk by integrating the EVT with

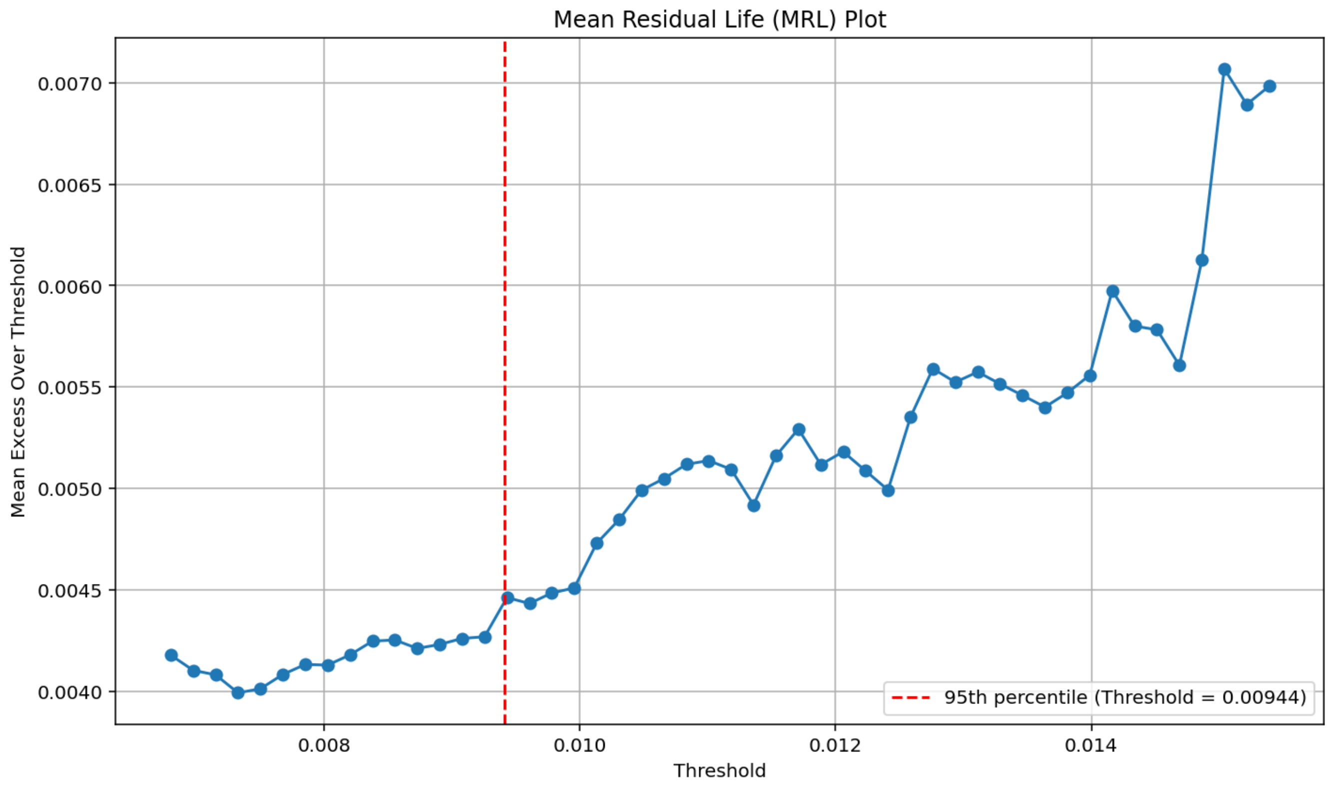

Fig. 10. Mean Residual Life (MRL) Plot for Log Returns

the 2D-NHPP model. Employing the daily return data of the NSE 20 Share Index over the period of ten years and with interest rates and market volatility as control variables, the model offers an efficient procedure for measuring both the intensity and the occurrence of shock events.

The results from the analysis show that accounting for macroeconomic variables enhances the accurate estimation of the location, scale, and shape parameters of the GPD model. It was also found that the more volatile a market is, the higher the value of all three parameters, which implies that a higher level of market volatility increases the probability of rare events, shifts baseline risk up and increases kurtosis of return distribution. On the other hand, the level of interest rates had an inverse relationship with the same parameters, supporting their role of mitigating volatilities in the financial markets, proxied by containing the intensity and occurrence of large losses. These findings supports the theoretical framework and existing literature review relating to similar studies.

Compared to static VaR models, the dynamic characteristics of EVT-2D-NHPP make it appealing by allowing the model parameters to incorporate time dependencies and changes in economic conditions. This dynamic modelling approach is particularly useful in the context of the developing countries such the Kenyan scenario where financial structures are highly fragile to shocks and volatility or of liquidity.

In conclusion, the study provides a framework that is practical and quantitative in nature for the evaluation of risk that can be useful for investors, policymakers, and regulators to make confident financial decisions. The ability to incorporate explanatory variables makes the model more valuable in its capacity to predict and apply to real-life situations.

Further research might seek to expand the above model by defining new predictors, which include such options as inflation rates, fluctuations in the exchange rates or worldwide economic indicators. This would also help identify cross-sectional risk characteristics and or value-at-risk for other asset classes or even other complex and diverse financial markets. Extending the model in these directions would go even a long way in enhancing the establishment of robust and evidence based financial risk management frameworks.

References

[1] H. Al Amosh, "The Anthropocene reality of financial risk," *Social and Environmental Accountability Journal*, vol. 44, no. 1, pp. 85–86, 2024.

[2] S. Peng, S. Yang, and J. Yao, "Improving value-at-risk prediction under model uncertainty," *Journal of Financial Econometrics*, vol. 21, no. 1, pp. 228–259, 2023.

[3] C. Candia and R. Herrera, "An empirical review of dynamic extreme value models for forecasting value at risk, expected shortfall and expectile," *Journal of Empirical Finance*, p. 101488, 2024.

[4] M. M. Nemukula and C. Sigauke, "A point process characterisation of extreme temperatures: an application to South African data," *Environmental Modeling & Assessment*, vol. 26, pp. 163–177, 2021.

[5] L. B. Salem, R. Nouira, K. Jeguirim, and C. Rault, "The determinants of crude oil prices: Evidence from ARDL and nonlinear ARDL approaches," *Resources Policy*, vol. 79, p. 103085, 2022.

[6] G. Fabiani, F. Calabrò, L. Russo, and C. Siettos, "Numerical solution and bifurcation analysis of nonlinear partial differential equations with extreme learning machines," *Journal of Scientific Computing*, vol. 89, no. 2, p. 44, 2021.

[7] P. Beaumard, "Experimental analysis of two-point Navier-Stokes equations in non-homogeneous turbulence and application to Large Eddy Simulation," Ph.D. dissertation, Centrale Lille Institut, 2023.

[8] P. Žugec, "Marked Poisson cluster processes and applications," Ph.D. dissertation, Univ. of Zagreb, Dept. of Mathematics, Zagreb, Croatia, 2019.