Research Article

Risk Analysis of Gold Prices in Pakistan Using Extreme Value Theory

Ghulam Raza Khan [ID],[1] **Alanazi Talal Abdulrahman** [ID],[2] **Osama Alamri** [ID],[3] **Zahid Iqbal** [ID],[4] **and Maqsood Ahmad** [ID][5]

[1]*Department of Statistics, University of Wah, Wah Cantt., Rawalpindi, Pakistan*
[2]*Department of Mathematics, University of Ha'il, Ha'il, Saudi Arabia*
[3]*Department of Statistics, University of Tabuk, Tabuk, Saudi Arabia*
[4]*Department of Statistics, Allama Iqbal Open University, Islamabad, Pakistan*
[5]*Department of Statistics, University of Okara, Okara, Pakistan*

Correspondence should be addressed to Ghulam Raza Khan; ghulam.raza@uow.edu.pk

Extreme value theory (EVT) is useful for modeling the impact of crashes or situations of extreme stress on investor portfolios. EVT is mostly utilized in financial modeling, risk management, insurance, and hydrology. The price of gold fluctuates considerably over time, and this introduces a risk on its own. The goal of this study is to analyze the risk of gold investment by applying the EVT to historical daily data for extreme daily losses and gains in the price of gold. We used daily gold prices in the Pakistan Bullion Market from August 1, 2011 to July 30, 2021. This paper covers two methods such as Block Maxima (BM) and Peak Over Threshold (POT) modeling. The risk measures which are adopted in this paper are Value at Risk (VaR) and Expected Shortfall (ES). The point and interval estimates of VaR and ES are obtained by fitting the Generalized Pareto (GPA) distribution. Moreover, in this paper, return-level forecasting is also included for the next 5 and 10 years by analyzing the Generalized Extreme Value (GEV) distribution.

1. Introduction

Gold is a familiar valuable metal for investment as compared to silver, platinum, and palladium. Gold has been widely used in many industries not only in the manufacturing of jewelry but also in financing and investing, medical and dentistry, manufacture of electronics, components of computers, and aerospace. In the financing and investing industry, gold is held in the form of gold bullion. Gold bullion is a valuable precious metal that is usually in two main forms: gold bullion bars and gold bullion coins. Nowadays, many institutions are including gold as a major asset of the world's investment due to their stability which plays an important role in the global economy. Gold is a good investment, neither in the short-term nor the long-term, and it is a great asset that can be converted into paper money [1].

Gold is one of the oldest and most widely-held commodities used as a hedge against the risk of disruptions in financial markets, but gold prices fluctuated substantially over time: slowly and gradually, and this introduces the risk on its own. The gold prices in Pakistan are influenced by many factors, such as the value of the US dollar, gold production, consumer demand, devaluation of Pakistani currency, and inflation. These factors which affect the gold prices are also discussed by Chaithep et al. [2].

In order to disclose the nature of the risks, under extreme situations, and finally, to avoid the risks to the most degree, we need certain risk measures. Extreme value theory (EVT) is for assessing the asymptotic probability distribution of extreme values for modeling the tail part of the distribution where the risk exists. EVT is playing an important role in dealing with modeling rare events. EVT applications are found in many fields where extreme events may occur,

including hydrology [3], insurance [4], and finance [5]. EVT provides a solid framework for estimating the behavior of extreme events. The performance of EVT is better than the counterparts for predicting unexpected extreme movements because of focusing on the tails of the distribution [6].

Mainly, there are two broad methods of applying EVT; Block Maxima (BM) approach is based on the Generalized Extreme Value (GEV), while Peak Over Threshold (POT) approach is based on the Generalized Pareto (GPA) distribution.

The BM method is the primary and the oldest method of estimating extreme events in EVT. This method was given by Gumbal [7]. The BM method considers the distribution of the maximum order statistic in a GEV distribution that best fitted to the series of extremes observations. In this method, a single-highest value is taken from a block. The second highest value in this block may be larger than the highest of another block but this value is not accounted for. This is the conventional way of modeling extreme events and is used mostly in hydrology and other engineering applications. Further, the BM method is not suited for financial time series because of volatility clustering. Due to this phenomenon, extreme events tend to follow one another; the POT method is efficient as compare to the BM method because it uses the data more efficiently (it consider several large values instead of only the largest one). This approach consists of using observations that are above a given threshold, called exceedances. Therefore, the POT method has become the method of choice in financial applications. This method can be modeled by the Generalized Pareto (GPA) distribution [8].

In Figure 1(a), the observations X_2, X_5, X_7, and X_{11} show that the BM for four periods with three observations each. This is a traditional way of modeling extreme events and is used effectually in hydrology and other engineering applications. In Figure 1(b), the observations $X_1, X_2, X_7, X_8, X_9,$ and X_{11} exceed the threshold (u), which represents extreme events. The distribution of exceedances above a threshold is based on a theory developed by Fisher and Tippett [9] that gives the limiting distribution of sample maxima provided the series which have certain mathematical properties [10].

The risk measures used in this paper are Value at Risk (VaR) and Expected Shortfall (ES). The VaR is a statistical technique that estimates the risk of an investment. In other words, VaR is the maximum possible loss or gain that could happen in investment over a given time period. ES is the average amount of all losses which are greater or equal to VaR. There are two approaches, parametric and nonparametric, which are used to compute VaR and ES. In this paper, we used EVT that is a parametric approach for the estimation of VaR and ES.

Ren and Giles [11] applied the EVT in daily returns of crude oil prices in the spot market from years 1998 to 2006. They concluded that the POT method by analyzing GPA distribution provides an effective risk measure such as VaR and ES. Jang [12] used the daily gold prices data from 1985 to 2006 to clear the main idea of EVT and discussed the tail behavior. He computed VaR and ES by using GPA

distribution and normal distribution, and results show that the GPA distribution is superior to the normal distribution for VaR and ES estimations. Chaithep et al. [2] examined the EVT approaches to daily gold prices from the years 1985 to 2011 in the U.S. dollar. He concluded that VaR estimation by the GPA distribution model is more efficient than GEV distribution. Chen [13] analyzed the gold investment risk by using the EVT to historical daily data from years 1968 to 2014 in the London market. She compared the positive and negative return results of VaR and ES of different metals such as gold, silver, and platinum.

Chinhamu et al. [14] focused on the importance of the EVT approach for tail-related risk measures taking data from years 1969 to 2012. They concluded that GPA distribution gives best results as compared to normal and Student's t distribution to estimate tail risk measures such as VaR and ES. In [15], the authors explained that the classical extreme value statistics consists of two fundamental techniques: the BM and the POT approach. It seems to be a general consensus among researchers in the field that the POT method makes use of extreme observations more efficiently than the BM method. POT is preferable to quintile estimation, while BM is preferable to return level estimation. In [16], the authors presented a study to assess the quality of VaR forecasts in various states of the economic situation. Two methods based on the EVT were compared, BM and POT. Forecasts were made on the daily closing prices of 10 major indices in European countries. Karahanoglu [17] analyzed the VaR method for one single portfolio or for all similar portfolios, and it hampers the opportunity for comparison in the finance market. The dataset is chosen as the period from January 2, 2014, to April 20, 2020, when all returns are recorded daily. The results show for similar portfolios, and different VaR methodologies and different back testing processes must be applied for the best fit. In [18], the author analyzed the VaR which has long been the standard risk measure in financial risk management. In this paper, he shows how EVT can be applied in estimating ES using an unconditional POT method. He measures the VaR as the standard measure of financial risk. In terms of ES, we find that cryptocurrencies are the high-risk asset and fixed income the shield. In [19], the authors explained EVT to model the tail behavior of the J580 Index for the right tail and the left tail. They used the POT method to estimate the threshold using the Pareto Q-Q plot, Hills plot, and mean residual live plot. The MLE method was used to estimate the parameters. Extreme tail risk measures were estimated at relatively high quantile at the 95%, 99%, and 99.5% confidence intervals. The VAR and ES are used to examine the extent of the potential extreme downside and upside risk of the J580 Index. The results indicate the importance of the GPA distribution in fitting suitably the tails of financial time series data characterized by extreme events.

The rest of the paper is organized as follows: In Section 2, we briefly describe the GPA and GEV distribution, Block Maxima, and Peaks Over Threshold method for establishing extremes. Also, it describes the measures of extreme risks VaR and ES. Empirical results obtained in GPA and GEV distribution are discussed in Section 3. Finally, Section 4 concludes the study.
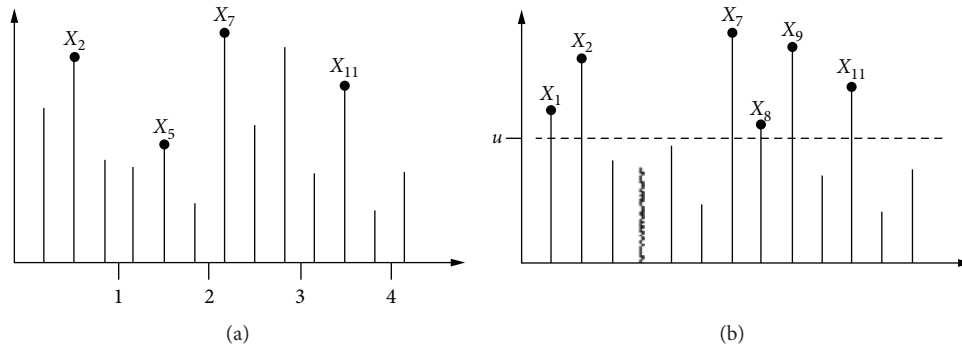
FIGURE 1: (a) Block Maxima method. (b) Peak-Over-Threshold Method.

2. Methodology

2.1. Basic Fundamental Assumptions. Basically, it is compulsory to check the basic assumptions of data before conducting the analysis in modeling of any annual maximum series approach in the field of statistical hydrology. The basic and fundamental assumptions are stationary, randomness, homogeneity, and independence.

2.1.1. The Homogeneity Assumption. The problem of homogeneity largely rises due to unexpected events in the observed sample data. This is occurring due to changes that occurred in the trend of the series. If a heterogeneity problem appears in the observed data series and frequency analysis results become unsuitable, then take out a particularly statistical investigation. In this situation, it is necessary to check the assumption of homogeneity before starting the statistical analysis of a given dataset. For checking the homogeneity assumption, we apply the Mann–Whitney U test. It is a nonparametric test that can also be used in the emplacement of an unpaired t-test and in the extreme type of data (such as gold price, rainfall, and flood data) that present variability that damaged the property of homogeneity.

2.1.2. The Independence Assumption. The word independence means that the observed value of the sample is not the nonexistence of any other sample value. Puricelli [20] also discussed the historical hydrology data series such as the sequence of high and low, catchment flows, ARMS, annual maxima, or minima. The dependence is expected on the daily series, while no dependence is expected for periodic or annual series. The independence assumption is formerly tested before performing the statistical analysis for hydrological variables. To check the independences' assumption, Wald and Wolfowitz [21] explained a test to check the independence assumption for the observed data. WW test is a nonparametric family test. The WW test is used to find whether the observed sample value is not affected by the existent of remaining sample data elements.

2.1.3. The Assumption of Stationarity. The stationary assumption is important on sample data for their probability distribution, and related parameters are fixed with respect to time. In nonstationarity conditions, monotonic and non-monotonic trends (unexpected changes at some point in time), we examined whether the sample data are stationary with respect to monotonic nonlinear or linear trends. They help us in nonparametric tests based on Spearman's rank correlation coefficient. The British psychologist Charles Spearman from 1863 to 1945 performed a nonparametric estimate of the statistical dependence between two random variables. It is shown by and named as in the Spearman rank correlation. ρ is statistical dependence, and it compresses a linear dependence as in usual Pearson's linear correlation coefficient.

2.1.4. The Assumption of Randomness. NERC tests are used to check the randomness and find whether a dataset has an appreciable pattern and also whether the process that generated it is significantly random. If the sample of a hydrology variable can appear no randomness, then statistical dependence among its element has shown no homogeneity, and nonstationary causes can be also natural or man-made. Flood, climate fluctuations, and earthquakes are the natural causes that are mostly allied to where evolution causes are connected with land-use changes and building of big pool dams upstream, and humans make climate change.

2.2. Generalized Extreme Value Distribution. The Generalized Extreme Value (GEV) distribution is three-parameter flexible distributions. It consists of Gumbal, Frechet, and Weibull distributions. The probability density function (PDF) of GEV distribution is given as

$$f(x) = \begin{cases} \dfrac{1}{\alpha}\left(1 + \xi\dfrac{x-\mu}{\alpha}\right)^{-(1/\xi+1)} \exp\left[-\left(1 + \xi\dfrac{x-\mu}{\alpha}\right)^{-1/\xi}\right], & \text{for } \xi \neq 0, \\[3mm] \dfrac{1}{\alpha} \exp\left(-\dfrac{x-\mu}{\alpha}\right) \exp\left(-e^{x-\mu/\alpha}\right), & \text{for } \xi = 0, \end{cases}$$

(1)

where ξ is the shape parameter, μ is the location parameter, and α is the scale parameter. The shape parameter always shows the thinkness of the tail. So, if the shape parameter gives large value (means give positive value), then tail of the distribution is the thicker.

2.3. Generalized Pareto Distribution. According to Pickands [22], the probability distributions of exceedances in an event are approximated to the Generalized Pareto (GPA) distribution. By introducing the shape parameter α and location parameter β, the two parameter GPA distribution has the following representation [23].

The PDF of GPA distribution is given below as

$$f_{\alpha,\mu,\beta}(x) = \begin{cases} 1 - \left\{ 1 + \alpha\left(\dfrac{x}{\beta}\right) \right\}^{-1/\alpha}, & \text{if } \alpha \neq 0, \\ \\ 1 - e^{-(x/\beta)}, & \text{if } \alpha = 0, \end{cases} \quad (2)$$

where

(i) $x > 0$, when $\alpha \geq 0$, and $0 \leq x \leq -\beta/\alpha$ when $\alpha < 0$

(ii) $\beta > 0$

Further, if the value of the shape parameter is high, then tail follows heavy-tailed distribution.

2.4. Excess Distribution. For a random variable X, the excess distribution function F_u above a certain threshold u is defined as

$$F_u(x) = \Pr[X - u \leq x \mid X > u], \quad (3)$$

where x represents the size of exceedances over threshold u. Further, if we denote F as the distribution function for X, then we may write

$$F_u(x) = \frac{F(x + u) - F(u)}{1 - F(u)}. \quad (4)$$

A fundamental theorem in EVT, by Pickands [22], identifies the asymptotic behavior of these exceedances with Generalized Pareto (GPA) distribution. The excess distribution function F_u can be well approximated by GPA distribution for large enough u:

2.5. Peak over Threshold Method. We fit a Generalized Pareto (GPA) distribution to our dataset; we adopt the Peak Over Threshold (POT) method that focuses on the distribution of exceedances above some high threshold. For $x - u \geq 0$, we can rewrite the excess distribution function (4) as

$$F_u(x - u) = \frac{F(x) - F(u)}{1 - F(u)},$$
$$F(x) = (1 - F(u))F_u(x - u) + F(u), \quad (5)$$

which allows us to apply the POT method:

There are two steps in applying the POT method. Firstly, we need to choose an appropriate threshold. Secondly, fit the GPA distribution function to data. Given the choice of a sufficiently high threshold, we may estimate $F(u)$ by $1 - N_u/n$ where n is the total sample size and N_n is the amount of observations above the chosen threshold. And, $F_u(x - u)$ can be estimated by a GPA distribution using maximum likelihood estimation; we then can obtain the following tail estimator [11]:

$$\widehat{F}(x) = 1 - \frac{N_u}{n}\left(1 + \frac{\widehat{\alpha}}{\widehat{\beta}}(x - u) \right)^{-1/\widehat{\alpha}}. \quad (6)$$

2.6. Selection of Threshold. The selection of the threshold level is an issue of adjusting bias and variance. In [24], the authors brought up that if a threshold is too low, it is more likely to violate the asymptotic property of the model and cause bias; if we take a threshold too high, it will generate few exceedances with which the model can be estimated and results in high variance. A basic procedure is to select a threshold as low as possible, as long as the limiting approximation of the model can give a reasonable result. So far, there is no algorithm-based method available for the selection of the threshold (u). Many researchers including McNeil [4], Danielsson and De Vries [25], and others have utilized this data-analytic issue, but none has provided a convincing solution.

In this paper, we used graphical methods for selection of the threshold level. In graphical methods, the most widely used plots are mean excess (ME) plot and shape parameter plot. These plots are also used by Ren and Giles [11], Chen [13], Chinhamu et al. [14], Jakata and Chikobvu [15], and many other researchers for the selection of the appropriate threshold level.

2.6.1. Simple Mean Excess Plot. Davidson and Smith in [26] introduced the simple mean excess (ME) plot. The first approach for threshold selection utilizes the ME plot. A mean excess function is the mean of the exceedances over a certain threshold (**u**). For a random variable X with right end point F_x, the mean excess function is defined as

$$e(u) = E(X - u \mid X > u). \quad (7)$$

For $u < F_x$, if the underlying distribution of $X > u$ has a GPA distribution, then the corresponding mean excess is

$$e(u) = \frac{\beta + \alpha u}{1 - \alpha}, \quad \beta + \alpha u > 0. \quad (8)$$

In equation (8), we can clearly see that the mean excess function must be linear in u. More precisely, it follows a GPA distribution if and only if the mean excess function is linear in u [24]. This gives us a way of selecting an appropriate threshold.

Given the data, we define the empirical mean excess function as

$$e_n(u) = \frac{\sum_i^n (X_i - u)I_{(X_i > u)}}{\sum_i^n I_{(X_i > u)}}, \quad (9)$$

where n is the sample size. The empirical excess plot is a graphical representation of the locus of $(u, e_n(u))$, and we can examine this plot to choose the threshold **u** such that $e_n(u)$ is approximately linear for $X > u$.

We use *R* software to make the plot. More discussion will be presented in Section 2.6.

2.7. Parameter Estimation. There are several techniques for estimating the parameters such as maximum likelihood estimation (MLE), method of moments, and method of probability-weighted moments.

In this paper, the MLE method is used for estimating the parameter of GEV and GPA distribution because the MLE is asymptotically normal and allows simple approximations for standard errors and confidence intervals.

Suppose that block maxima $X_1, X_2, X_3, \ldots, X_k$ are independent from a GEV distribution, and the log-likelihood function for the GEV distribution from the equation (1) is given as

$$\ln L = \begin{cases} -k \ln \sigma - \left(1 + \dfrac{1}{\xi}\right) \sum_{i=1}^{k} \ln \left\{1 + \xi \dfrac{(X_i - u)}{\sigma}\right\} - \sum_{i=1}^{k} \ln \left\{1 + \xi \dfrac{(X_i - u)}{\sigma}\right\}^{-1/\xi}, & \text{for } \xi \neq 0, \\[4mm] -k \ln \sigma - \sum_{i=1}^{k} \dfrac{(X_i - u)}{\sigma} - \sum_{i=1}^{k} \exp\left\{\dfrac{(X_i - u)}{\sigma}\right\}, & \text{for } \xi = 0. \end{cases} \tag{10}$$

Smith [27] declared that, for $\xi > 0.5$, the MLE for $\xi, u,$ and σ satisfies the regular condition and therefore having asymptotic and consistent properties. The number of blocks k and the block size form a crucial trade-off between variance and bias of parameter estimation [2].

Given that we have a adequately high threshold u and assuming there are m observations with $x_i - u \geq 0$, the subsample $(x_1 - u, \ldots, x_m - u)$ has an underlying distribution of GPA distribution, where $x_i - u \geq 0$ for $\alpha \geq 0$ and $0 \leq x_i - u \leq -\beta/\alpha$; then, the logarithm of the PDF of x_i can be derived from equation (2) as

$$\ln f(x_i - u) = \begin{cases} -\ln \beta - \dfrac{1 + \alpha}{\alpha} \ln \left\{1 + \alpha \left(\dfrac{x_i - u}{\beta}\right)\right\}, & \text{for } \alpha \neq 0, \\[4mm] -\ln \beta - \dfrac{1}{\beta}(x_i - u), & \text{for } \alpha = 0. \end{cases} \tag{11}$$

Hence, the log-likelihood function $L(\alpha, \beta \mid x_i - u)$ for the GPA distribution is the logarithm of the joint density of the m observations is as

$$L(\alpha, \beta \mid x_i - u) = \begin{cases} -n \ln \beta - \dfrac{1 + \alpha}{\alpha} \sum_{i=1}^{m} \ln \left\{1 + \alpha \left(\dfrac{x_i - u}{\beta}\right)\right\}, & \text{for } \alpha \neq 0, \\[4mm] -n \ln \beta - \dfrac{1}{\beta} \sum_{i=1}^{m} (x_i - u), & \text{for } \alpha = 0. \end{cases} \tag{12}$$

We can obtain the estimates for α and β by maximizing the log-likelihood function of the subsample under a suitable threshold (u) [14].

2.8. Assumption of Poisson Distribution. The POT approach has many features in the modeling of exceedances in which the observations exceed or occur over a given period of time. In the POT method, the general and simple rule is that the numbers of exceedances follow a Poisson process which is however not necessary for modeling the BM approach [28]. However, Rosbjerg et al. [29] considered this assumption compulsory for modeling of the POT method. A Poisson distribution is given in the equation below:

$$P = \frac{e^{-\mu} \mu^x}{x!}, \tag{13}$$

where e is the base of natural log, μ is the number of success mean, and x is the successes number.

In this assumption, the occurrences must be independent and the probability of occurrence of an extreme event at a time "$t + s$" does not depend on another extreme event that occurs at a time "t." This assumption may not fulfill in the case of natural processes (i.e., outbreak of a disease, floods, hurricanes, and draught) as they occur in clusters. In this condition, the decluttering method is normally used.

The property of Poisson distribution is that it has same mean and variance. However, the gold price dataset usually contain some outliers, which may result in violating the Poisson process assumption. Cunnane [28] suggested a criterion on the bases of confidence interval to test this assumption; the test is called as "Fisher Dispersion Index" (FDI). The test statistics proposed by Cunnane is based on approximating the Poisson variants through normal variety. The number of peaks that occur in the ith year is expressed as n_i and is normally distributed with mean $\widehat{\lambda}$ and standard deviation $\widehat{\lambda}$; then, the test statistics:

$$R = \sum_{i}^{N} \frac{(n_i - \lambda)^2}{\lambda}, \quad (14)$$

where n_i is the number of exceedances that happen in the ith year and λ is the mean of the numbers that exceed the threshold level.

Following the chi-square distribution with degree of freedom $\mathbf{b} = N-1$, where N is total number of years, this is valid for N greater than 5 ($N > 5$). The test statistics that perform at "α" significance level under the null hypothesis H_0 is accepted if the value of R lies between

$$\chi^2_{\alpha/2}(\mathbf{b} - 1) < R < \chi^2_{1-\alpha/2}(\mathbf{b} - 1). \quad (15)$$

2.9. Risk Measures' Estimation. In this paper, we employ two standard choices of risk measurement, which are Value at Risk (VaR) and Expected Shortfall (ES). These two risk measures involved extreme quantile estimation.

2.9.1. Value at Risk. Value at Risk (VaR) is a statistical technique, which measures the maximum possible loss for a given time period. It calculates how much a set of expenditures might lose. The VaR is used for measuring the market risk starting in early 1990 [30].

For a random variable, X is usually the return in some risky financial market with distribution function F over a specified period, and the VaR for a given probability p can be defined as the pth quartiles of F is as

$$\text{VaR}_p = F^{-1}(1 - p), \quad (16)$$

where F^{-1} is the quartile function. VaR is a common measure of extreme risks, and we used GPA distribution to approximate this measure. In particular, using (6), we obtain VaR as

$$\widehat{\text{VaR}}_p = \begin{cases} u + \dfrac{\widehat{\beta}}{\widehat{\alpha}} \left\{ \left(\left(\dfrac{n}{N_u} p \right) \right)^{-\widehat{\alpha}} - 1 \right\}, & \widehat{\alpha} \neq 0. \\ \\ u - \widehat{\beta} \log\left(\dfrac{n}{N_u}(1 - p) \right), & \widehat{\alpha} = 0, \end{cases} \quad (17)$$

where $\widehat{\beta}$ and $\widehat{\alpha}$ are the maximum likelihood estimates of the GPA distribution parameter.

2.9.2. Expected Shortfall. Expected Shortfall (ES) is proposed by Artzner et al. [31], and it is the average amount of loss, or in other words, ES is the average of all losses which are equal to or greater than the VaR level.

In contrast to VaR, ES measures the risk of the market by considering both the size and likelihood of the losses above a particle threshold. ES gives the expected size of return that exceeds VaR for a probability level p:

$$\text{ES}_p = E(X \mid X > \text{VaR}_p), \quad (18)$$

and equivalently,

$$\text{ES}_p = \text{VaR}_p + E(X - \text{VaR}_p \mid X > \text{VaR}_p), \quad (19)$$

where the second term above represents the excess distribution $F_{\text{VaR}_p}(x)$, treating VaR_p as the threshold level, proceeding as before. If the threshold VaR_p is sufficiently large, then $F_{\text{VaR}_p}(x)$ is GPA distribution i.e.,

$$F_{\text{VaR}_p}(x) = G_{\alpha, \beta + \xi(\text{VaR}_p - u)}(x). \quad (20)$$

Thus, the mean of the excess distribution $F_{\text{VaR}_p}(x)$ can be calculated as

$$\frac{\beta + \alpha(\text{VaR}_p - u)}{1 - \alpha}, \quad (21)$$

where $\alpha < 1$ and substring into equation (19) will yield

$$\widehat{\text{ES}_p} = \frac{\widehat{\text{VaR}}_p}{1 - \widehat{\alpha}} + \frac{\widehat{\beta} - \widehat{\alpha}u}{1 - \widehat{\alpha}}. \quad (22)$$

2.10. Return Level Estimation. The most widely useful application of the block maxima approaches is estimating the return levels based on the Generalized Extreme Value (GEV) distribution. It is the quantile points of the GEV distribution:

$$R_n^k = M^{-1}\left(1 - \frac{1}{k}\right), \quad (23)$$

where M is the distribution of maxima observation of equal length over a successive nonoverlapping period.

The expected level to be exceeded in one out of k periods of length n where k is the period and n is the length of this period. The return level can be used as a measure of the maximum loss of a portfolio, a rather more conservative measure than the Value at Risk [32].

3. Results and Discussion

3.1. Data Description. The data used in this paper are the daily gold prices in the Pakistan Bullion market over the time period of August 1, 2011 to July 30, 2021, and the data is taken from the following website: https://www.bullion-rates.com/gold/PKR/2007-2-history.htm. The daily gold returns are calculated by taking log differences of the daily gold prices:

$$\text{Daily Return}_t = \log P_i - \log P_{i-1}. \quad (24)$$

The total number of observations is 2707, and there are 2706 gold returns, among which 1304 are negative returns and 1402 are positive returns. We separated the positive returns and the negative returns and then take the absolute values of the negative returns because Extreme Value Theory (EVT) is only defined for nonnegative values [14].

Table 1 provides a summary of descriptive statistics for the considered return series.

In Table 1, we observed that the mean of daily return is 0.00012 which is a positive indication that the overall gold prices are increasing during the considered time period. The standard deviation (SD) is 0.0128, the maximum value is

TABLE 1: Descriptive summary statistic of gold returns.

Obs.	Mean	Median	Maximum	Minimum	SD	J-B test	Skewness	Kurtosis
2706	0.00012	0.00013	0.03858	−0.03784	0.00477	6061.756 (0.001)	−0.09688	10.3284

0.0871, and the minimum value is −0.1172 in daily return. The magnitude of the average return is very small as compared to the standard deviation. Further, the large kurtosis of 10.3284 indicates the distribution of returns has a fat tail. The Jarque–Bera statistic has zero probability; it indicates that we can reject the null hypothesis in favor of the alternative hypothesis, so the distribution is nonnormal [13].

The daily gold prices in the Pakistani rupee are shown in Figure 2, and daily returns are shown in Figure 3. The plot of the daily gold prices (Figure 2) shows a substantial increase since 2011 with lots of fluctuations, and the graph of daily returns (Figure 3) shows a large number of extreme values in the dataset. According to [12], the graph of daily return confirms the volatility of the Pakistan Bullion Market. The descriptive statistic and time series plots are obtained with the help of Eview 8 software.

The autocorrelations (ACF) pattern helps us to show the unpredictability in returns and volatility clustering. Autocorrelations investigate the correlations between today's value and value in the past time series. Figure 4 gives ACF of return series, and Figure 5 gives the ACF plot of squared returns series. The return autocorrelations are almost all insignificant while the squared returns have significant autocorrelations. Furthermore, the squared return autocorrelations are all positive which is highly unlikely to occur by chance. The figures give significant evidence for both the unpredictability of returns and volatility clustering [14].

Further, to check the stationary of the daily return of gold prices, we fit Augmented Dickey–Fuller (ADF) and Philips Peron (PP) Unit Root tests' results, which are given in Table 2.

In Table 2, we are applying the ADF and the P-P unit root tests on the return series of the gold prices to show the stationarity. In the ADF test, we used Schwartz Information Criterion to set at lag 0, and Bartlett Kernel spectral estimation method was used in the P-P test. The null hypothesis of unit root is rejecting in both tests ($p \leq 0.001$); in this situation, we accept the alternative hypothesis and conclude that the daily return of gold prices series are stationary.

3.2. Basic Fundamental Assumptions. It is compulsory to check the basic assumptions of data before conducting the analysis; the basic and fundamental assumptions are stationarity, randomness, homogeneity, and independence. The details of these assumptions are available in Section 2.1.

In Table 3, the results are satisfying the basic and fundamental assumptions of stationarity, randomness, homogeneity, and independence. We checked all fundamental assumptions for monthly, quarterly, and yearly Block Maxima (BM) of negative return. The test statistic value for each BM of a negative return is small and $p \geq 0.050$. Therefore, we concluded that the monthly, quarterly, and

yearly BM of a negative return is satisfying the fundamental assumptions of stationarity, randomness, homogeneity, and independence.

3.3. Determination of Best-Fitted Distribution. In this paper, first, we used the Block Maxima (BM) approach in which the Generalized Extreme Value (GEV) distribution is fitted. The same distribution was used by Pratiwi et al. [33]. In order to find the best fitting of GEV distribution, both graphical and formal tests of goodness-of-fit are applied. In graphical, to find the best-fitted distribution for BM of negative return, we used crude residual plot and Quantile-Quantile (Q-Q) plot.

In Figure 6(a), crude residual and Q-Q plot are for monthly BM of negative return, Figure 6(b) for quarterly BM of negative return, and Figure 6(c) for yearly BM of negative return.

The crude residual plot for monthly, quarterly, and yearly BM of negative return shows the concave pattern from the straight line, and it indicates that the data series come to a heavy-tailed distribution. The Q-Q plot looks linear to show that the GEV distribution is best-fitted for BM of negative return.

After graphical view, we concluded that GEV distribution is best fitted for monthly, quarterly, and yearly BM of negative return.

Now, Kolmogorov–Smirnov (K–S), Anderson–Darling (A–D), and chi-square tests are applied to find the best-fitted distribution for BM of negative returns.

In Table 4, some formal tests K-S, A-D, and Chi-Square tests' goodness-of-fit test are applied to find the best-fitted distribution. The test-statistics value of these goodness-of-fit tests is calculated by the R software. Table 4 shows the test-statistic value of GEV distributions and their p value. According to the K-S, A-D, and the Chi-Square test, all the p values are greater than 0.05, so GEV distribution is the best-fitted distribution for monthly, quarterly, and yearly BM of negative return.

We are focusing on GEV distribution which is a best-fitted distribution for BM of negative return. Next, we calculate the parameters of the selected distribution.

3.4. Parametric Maximum Likelihood Estimates. The Block Maxima (BM) of negative returns has been fitted to the Generalized Extreme Value (GEV) distribution with three different block sizes (yearly, quarterly, and monthly). We estimate the parameter by using the Maximum Likelihood (MLE) method. Refer to Section 2.7, equation (11) for the maximum likelihood function. The R software is used for modeling the extreme events.

Table 5 shows the the estimates of three-parameter location, scale, and shape with their standard error of GEV distribution for three different blocks size (yearly, quarterly, and monthly). According to Jang [12], all block numbers show
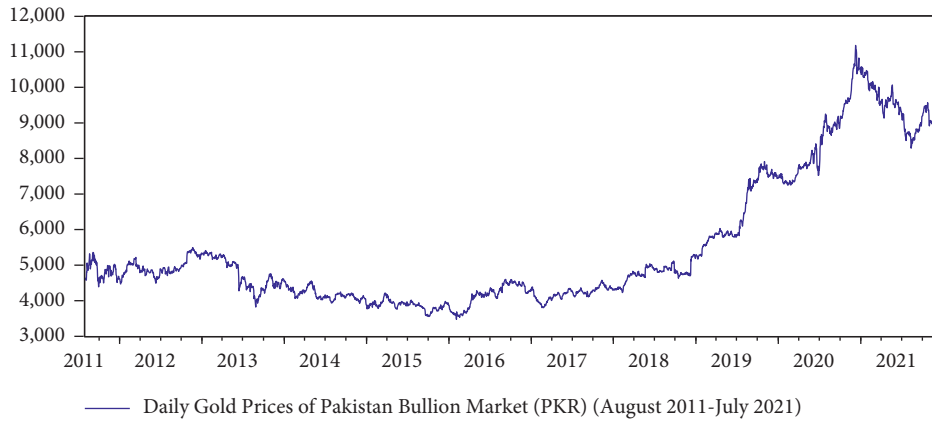
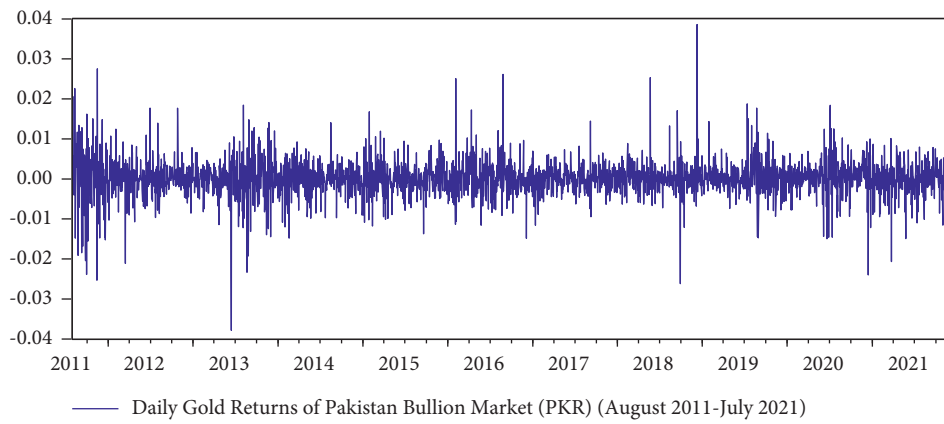Figure 2: Time series plot of daily gold prices from the Pakistan Bullion Market.

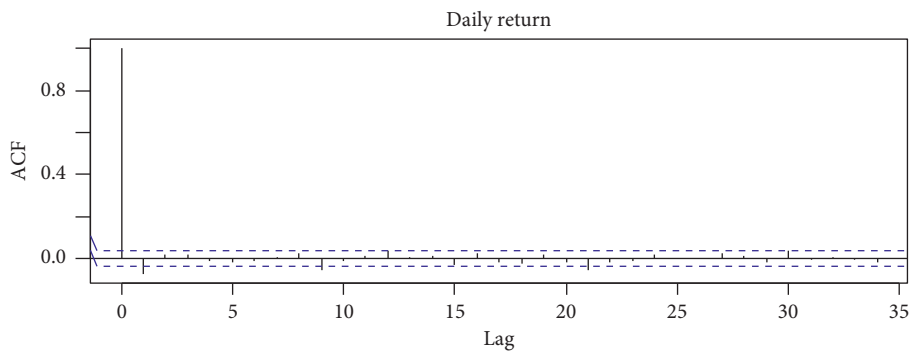Figure 3: Time series plot of daily gold returns from the Pakistan Bullion Market.

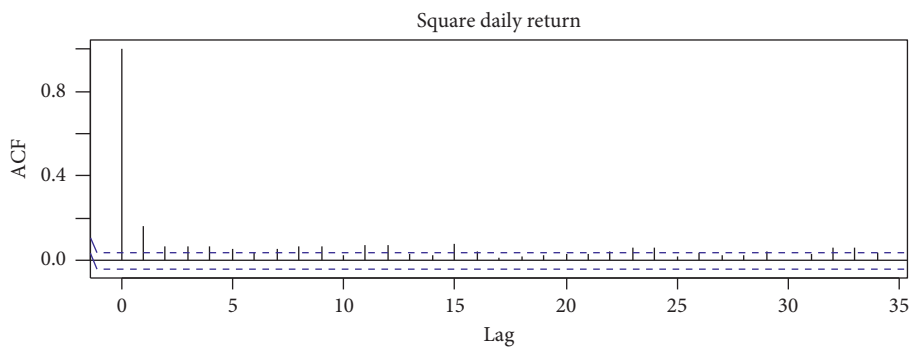Figure 4: Autocorrelation function plot (ACF) of daily gold returns.

Figure 5: Autocorrelation function plot (ACF) of square daily gold returns.

TABLE 2: Results of ADF and P-P unit root tests of gold return series.

Unit root test	Test statistic	p value
ADF test	−55.7761	(0.001)
P-P test	−55.7518	(0.001)

TABLE 3: The basic fundamental assumptions of block maxima of negative return series.

-	Randomness (NERC 1975)		Independence (Wald e Wolfowitz)		Homogeneity (Mann–Whitney test)		Stationary (Spearman test)	
Tests								
Values	p value	Statistic	p value	Statistic	p value	Statistic	p value	Statistic
Month	0.4633	0.7334	0.0501	1.6441	0.0512	−2.7293	0.0498	−2.2678
Quarter	0.0682	1.4891	0.3869	0.2875	0.3167	−0.4770	0.3344	−0.4277
Year	0.2903	0.5526	0.1593	0.9972	0.3008	−0.5222	0.3786	−0.3091

that the shape parameter is positive. All the parameters have decreasing estimated standard errors as the number of blocks increased. In general, the BM of the negative returns follows a Frechet family of GEV distribution for yearly, quarterly, and monthly block frames. The Frechet type of GEV distribution confirms that the original series has a fat tail.

3.5. Estimation of Return Level.

In this paper, we divide the n sample block into monthly, quarterly, and yearly blocks by taking the highest value in each block. The point estimate is estimated for the sample data (obtained through the BM approach) based on the MLE method using the probabilities of nonexceedances for different return periods. By putting the nonexceedances probability to the distribution's, the certain return period is computed. Refer to Section 2.8 for the maximum likelihood function. The R software was used for modeling the extreme values.

In Table 6, we estimate the point and interval estimates for the next five years and ten-year return period. We have monthly BM of the negative returns, and for the next five years' returns period, the point estimate is 0.0293, the corresponding probability of nonexceedances is 0.80, and the probability of exceedances is 0.20. It means that there are 80% chances that monthly maximum loss will not exceed 2.93% once in the next five years on average and a 20% chance it will be exceeded.

For the next ten-year return period, the point estimate is 0.0363 and the corresponding probability of nonexceedances is 0.90, and the probability of exceedances is 0.10. It means that there are 90% chances that the monthly maximum loss will not exceed the value of 3.63% once on average in the next ten years and a 10% chance it will exceed.

According to Jang [12], the point estimate is increased with decreasing block sizes. With the increasing time interval, the point estimate increases, and the confidence interval is wider. It reveals that time is a risk factor.

Similarly, we interpret the return level for quarterly and yearly BM of the negative returns.

3.6. Selection of Threshold.

As we have discussed in Section 2.6, there is no algorithm-based method available for the selection of the threshold; therefore, we used a graphical method for the selection of the threshold level. Two plots could help with the determination of the threshold level. Figure 7 presents the Mean Excess (ME) plot, and Figure 8 presents the estimate of the shape parameter for positive returns, and Figure 9 presents the estimate of the shape parameter for negative returns. These plots are obtained with the help of R software.

The ME plot is helpful in detecting graphically the quintile above which Pareto's relationship is valid. In Section 2.5, the ME plots are approximately linear in the threshold (u) given that the underlying distribution of sample data is a Generalized Pareto (GPA) distribution. More specifically, the ME plot of the data can be used to distinguish between light and heavy-tailed models; the plot of a heavy-tailed distribution shows an upward trend, a medium tail shows a horizontal line, and the plot is downward-sloped for light-tailed data. Common ground in our sample data is that both the ME plots of positive and negative returns have an upward trend followed by an irregular the portion in the far end.

Figure 7 shows ME plots for positive and negative returns. According to Chen [13], we take the threshold (u), where the ME plots are linear and upward-sloping. In this paper, the ME plots' upward-sloping is started from 0.001 to 0.015 thresholds' level. Therefore, we take the range of threshold level where the plots look like linear and upward slope in different segments.

Figures 8 and 9 show that the shape parameter plots of MLE estimates the shape and scale parameters under different thresholds' level. The 95% level of confidence intervals is computed for upper and lower dashed lines of plots. According to Chen [13], we took the threshold level before which the shape parameters are stable. According to Ren and Giles [11], we took the threshold level before the plot is comparatively flat.

The gain shows an approximate upward trend in the threshold from 0.001 to 0.015. The loss shows approximate linearity with a slightly upward trend in the threshold from 0.001 to 0.010. Therefore, there is some evidence to choose thresholds from 0.001 to 0.012 for the right tail and from 0.001 to 0.009 for the left tail.

3.7. Parameter Estimation.

Given the thresholds selected in the previous section, we could estimate the shape and scale parameters of the corresponding GPA distribution. Refer to Section 2.7, equation (13) for the maximum likelihood
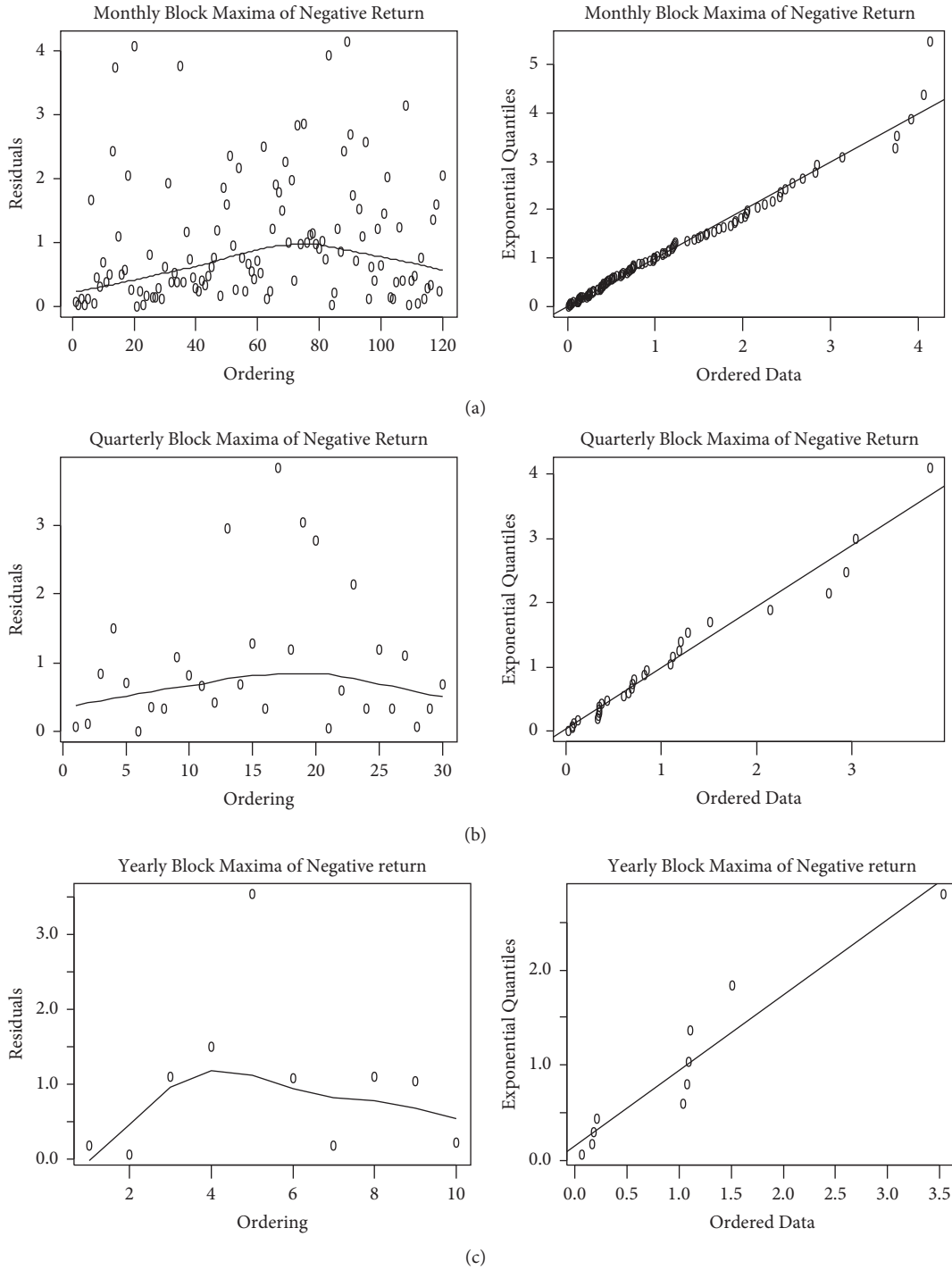
(a)

(b)

(c)

FIGURE 6: (a) Crude residual plot (left) and Q-Q plot (right) of fitted GEV distribution. (b) Crude residual plot (left) and Q-Q plot (right) of fitted GEV distribution. (c) Crude residual plot (left) and Q-Q plot (right) of fitted GEV distribution.

TABLE 4: Goodness-of-fit tests for GEV distribution.

Tests	K-S test		A-D test		Chi-square test	
Values	Statistic	p value	Statistic	p value	Statistic	p value
Month	0.0573	0.8251	2.0966	0.0814	0.4282	1
Quarter	0.2201	0.0933	0.5245	0.7205	0.1076	1
Year	0.3510	0.1319	0.2995	0.9375	0.0312	1

TABLE 5: Parameter estimation of GEV distribution based on the Maximum Likelihood method.

Parameter	Location (standard error)	Scale (standard error)	Shape (standard error)
Monthly	0.0063 (0.0003)	0.0031 (0.00002)	0.2657 (0.0981)
Quarterly	0.0101 (0.0007)	0.0039 (0.0003)	0.2618 (0.1822)
Yearly	0.0151 (0.0011)	0.0038 (0.000002)	0.5181 (0.5405)

TABLE 6: Return level based on GEV distribution for the three block frames with 95% confidence bounds.

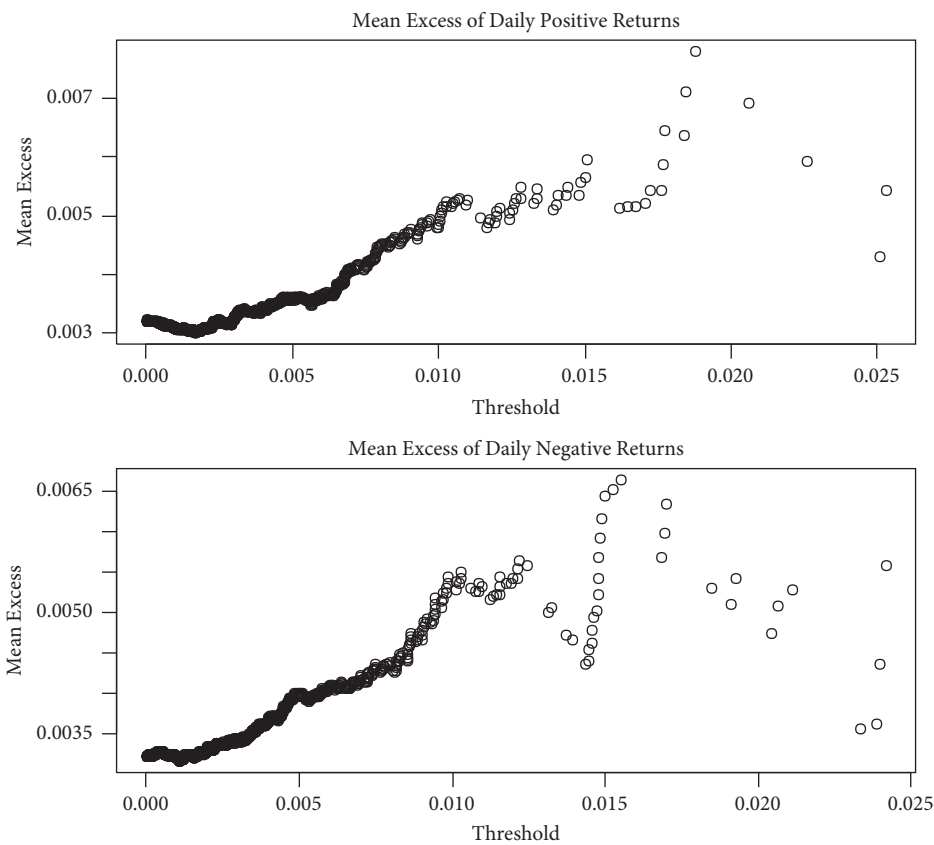| Return period	p (probability exceedances) 1/ return period	$1 - p$ (probability nonexceedances)	Monthly	Quarterly	Yearly
Once in five years	0.20	0.80	0.0293 (0.0224, 0.0449)	0.0249 (0.0194, 0.0444)	0.0237 (0.0177, 0.0568)
Once in ten years	0.10	0.90	0.0363 (0.0262, 0.0621)	0.0311 (0.0226, 0.0693)	0.0314 (0.0211, 0.0942)

FIGURE 7: Mean Excess plot for positive returns and negative returns.

function. The R software was used for modeling extreme events. Table 1 summarizes all the parameter estimates and their standard errors under different thresholds.

In Table 7, the confirmed choice of the threshold is $u = 0.008$ for positive returns and $u = 0.001$ for negative returns. For positive returns, there are 99 observations above the threshold which is 0.008. Compared to the sample size 1402, the tail only accounts for 7.06% of the total distribution. As mentioned in previous sections, EVT only analyzes the extreme events; thus, even if we started with a large sample, we contend with few observations. Moreover, compared to using weekly returns, daily returns provide a large sample size. This reduces the variance by increasing the

number of observations in the tail. It is the same situation for negative returns.

Since the maximum likelihood estimator is asymptotically normal, the associated MLE of the parameters under selected thresholds is statistically significant at the 5% significance level [13].

3.7.1. Assumptions of Poisson Distribution.

According to Cunnane [28], if the dataset contains outliers, the assumption is tested through Poisson distribution statistics with fixed confidence intervals. He also shows the use of the criterion of the Fisher Dispersion Index to check the Poisson
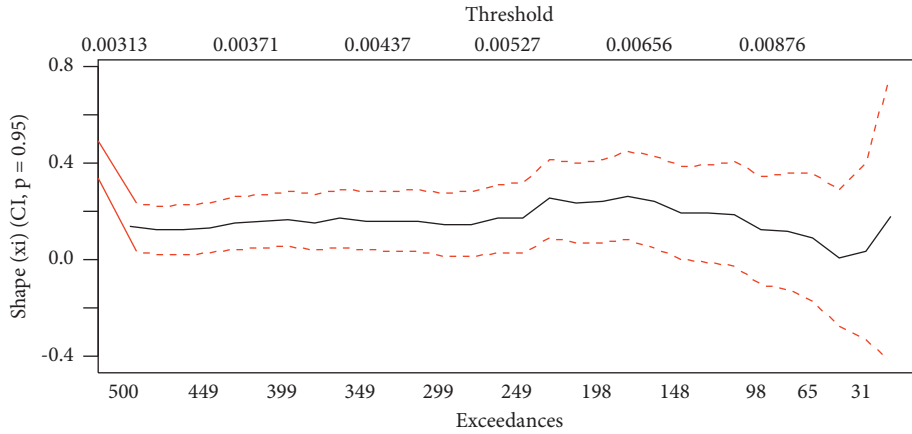
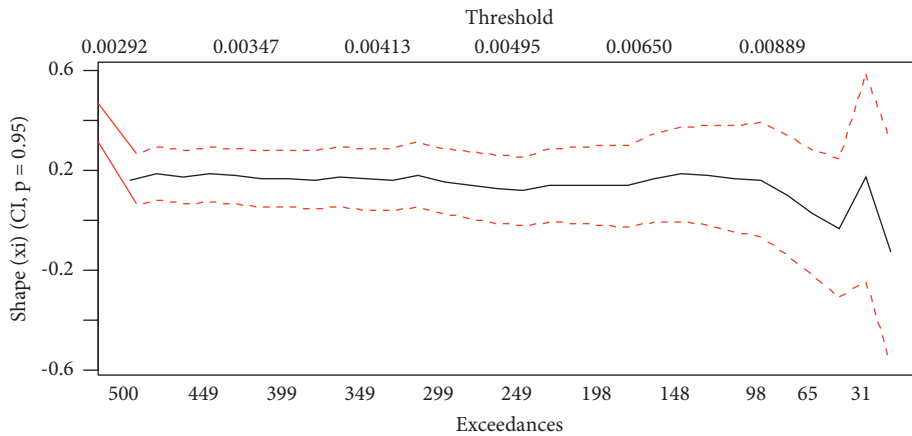Figure 8: Shape parameter plot for positive returns.

Figure 9: Shape parameter plot for negative returns.

Table 7: Maximum likelihood estimates of fitted GPA distribution under different thresholds' level.

Threshold	# of exceed	Probability less threshold	Shape parameter (standard error)	Scale parameter than (standard error)
Parameter estimate for positive returns				
0.001	1068	0.2382	0.0989 (0.02895)	0.0028 (0.00007)
0.004	375	0.7325	0.1530 (0.05438)	0.0029 (0.00015)
0.008	99	0.9294	0.1282 (0.11608)	0.0039 (0.00049)
0.012	39	0.9722	0.0978 (0.18580)	0.0045 (0.00094)
Parameter estimate for negative returns				
0.001	962	0.2623	0.1251 (0.03247)	0.0028 (0.00008)
0.003	493	0.6219	0.1762 (0.05001)	0.0029 (0.00013)
0.006	188	0.8558	0.1508 (0.08331)	0.0034 (0.00029)
0.009	81	0.9378	0.1443 (0.14659)	0.0040 (0.00061)

process assumption. The details are discussed about Poisson assumption in Section 2.8. The calculation is done in Excel by using equation (15).

In Table 8, the test statistic R is tested at 1% and 5% significance levels. For positive return, the threshold 0.008 and 0.012 of a sample of 99 and 39, respectively, and observations are considered for positive return. The test statistic R is 15.44 and 11.51 which falls in the acceptance region for both 1% and 5% significance levels which indicates that the sample obtained through the POT approach follows the Poisson process assumption. The sample of sizes 99 and 39 is

selected on the level of threshold. This sample satisfies the Poisson process assumption; therefore, further POT modeling analysis for positive return analysis is carried on the basis of this sample.

For negative return, the threshold is 0.001 and 0.003 of a sample of 962 and 493, respectively, and observations are considered for a negative return. The test statistic R is 10.14 and 17.89 which falls in the acceptance region for both 1% and 5% significance levels which indicates that the sample obtained through the POT approach follows the Poisson process assumption. A sample of sizes 962 and 493 are

TABLE 8: Output of Fisher Dispersion Index test statistics.

Threshold	Exceedances	R-statistic	C.I (1%)	C.I (5%)	Decision
Fisher Dispersion Index test for positive return					
0.001	1068	42.62	$1.74 < R < 23.59$	$2.70 < R < 19.20$	Rejected
0.004	375	30.97	$1.74 < R < 23.59$	$2.70 < R < 19.20$	Rejected
0.008	99	15.44	$1.74 < R < 23.59$	$2.70 < R < 19.20$	Accepted
0.012	39	11.51	$1.74 < R < 23.59$	$2.70 < R < 19.20$	Accepted
Fisher Dispersion Index test for negative return					
0.001	962	10.14	$1.74 < R < 23.59$	$2.70 < R < 19.20$	Accepted
0.003	493	17.89	$1.74 < R < 23.59$	$2.70 < R < 19.20$	Accepted
0.006	188	47.98	$1.74 < R < 23.59$	$2.70 < R < 19.20$	Rejected
0.009	81	59.62	$1.74 < R < 23.59$	$2.70 < R < 19.20$	Rejected

selected on the level of threshold. The samples satisfy the Poisson process assumption for a negative return. Therefore, further POT modeling for negative return analysis is carried on the basis of these two sample data.

3.8. Assumption of Generalized Pareto Distribution.

The exceedances in the POT approach are considered to identically distribute with exponential distribution, but in the presence of outliers, the exponential distribution fails to fulfill the model requirements. In recent years, GPA distribution attained high consideration in POT modeling because of its flexibility. We test the goodness-of-fit of GPA distribution above the threshold level by using K-S and A-D tests. The results of GPA distribution of goodness-of-fit for positive and negative threshold are given in Table 9.

Table 9 shows goodness-of-fit results for GPA distribution of positive and negative daily return at the threshold level; for positive return, the value of test statistic at threshold 0.008 is 0.1105 and 0.3312, respectively, and the p value are 0.1653 and 0.9128 for both K-S and A-D tests. Since $p \geq 0.050$, therefore we can say that the exceedances or the POT series follow the GPA distribution. The similar behavior is observed for the other threshold level.

It means that daily positive return of gold prices at threshold level 0.008 and 0.012 and daily negative return of gold prices at threshold level are 0.001 and 0.003 which are best fitted for GPA distribution.

EVT suggested that the excess distribution above a suitable threshold of daily returns should follow a GPA distribution. To determine how the GPA distribution fits the tails of the return distribution, we plot the empirical distribution of exceedances along with the cumulative distribution simulated by a GPA distribution and compare the results visually. Figures 10–13 are the residual diagnostic check of GPA distribution fit to daily positive and negative returns of gold prices.

Figures 10 and 11 provide the plots of excess distribution plot and tail of underlying distribution plot for gains and loss, and the plots show that the GPA distribution fitting to 99 exceedances at the threshold $u = 0.008$ with the shape parameter estimate 0.1282, and the GPA distribution fitting to 39 exceedances at the threshold $u = 0.012$ with the shape parameter estimate 0.0978. For losses, the plots give the GPA

TABLE 9: Goodness-of-fit of GPA distribution for both positive and negative returns.

Goodness-of-fit for positive return		
Threshold value	0.008	0.012
No. of exceedances	99	39
K-S test (p value)	0.1105 (0.1653)	0.1173 (0.6145)
A-D test (p value)	0.3312 (0.9128)	1.0893 (0.3132)
Goodness-of-fit for negative return		
Threshold value	0.001	0.003
No. of exceedances	962	493
K-S test (p value)	0.0325 (0.2616)	0.0448 (0.2753)
A-D test (p value)	0.2909 (0.9453)	0.4512 (0.7968)

distribution fitting is 962 exceedances at the threshold $u = 0.001$ with the shape parameter estimate of 0.1251, and the GPA distribution fitting is 493 exceedances at the threshold $u = 0.003$ with the shape parameter estimate 0.1762. For both positive and negative returns, the graph of the excess distribution function follows closely the trace of a corresponding GPA distribution, implying that the GPA distribution models the exceedances very well [11].

Figures 12 and 13 provide the plots of scatter plot of residuals and Q-Q plot for gains and loss and are the quantile plots to measure the quality of best-fitted GPA distribution. The GPA distribution is sensibly fitted to the exceedances above the threshold when the Q-Q plot shows a linear pattern. If the scatter plot of residuals does not show any pattern, it means that the peaks values are independent [14].

Overall, based on different fittings with different values of the threshold and associated parameter estimates, the GPA distribution models the tail behavior of our daily returns very well and the fits exhibit reasonable robustness to the choice of thresholds.

3.9. Risk Measures' Estimation.

The Value at Risk (VaR), Expected Shortfall (ES), and their respective confidence intervals for both positive and negative returns give different levels of the threshold and different tail quintiles. The risk measures used in this paper are VaR and ES. The VAR is a statistical technique that estimates maxima possible risk over one trade day and ES is the average amount of all losses which are greater or equal to VaR. There are two approaches, parametric and nonparametric, to compute VaR and ES. In summary statistics, we suggest that
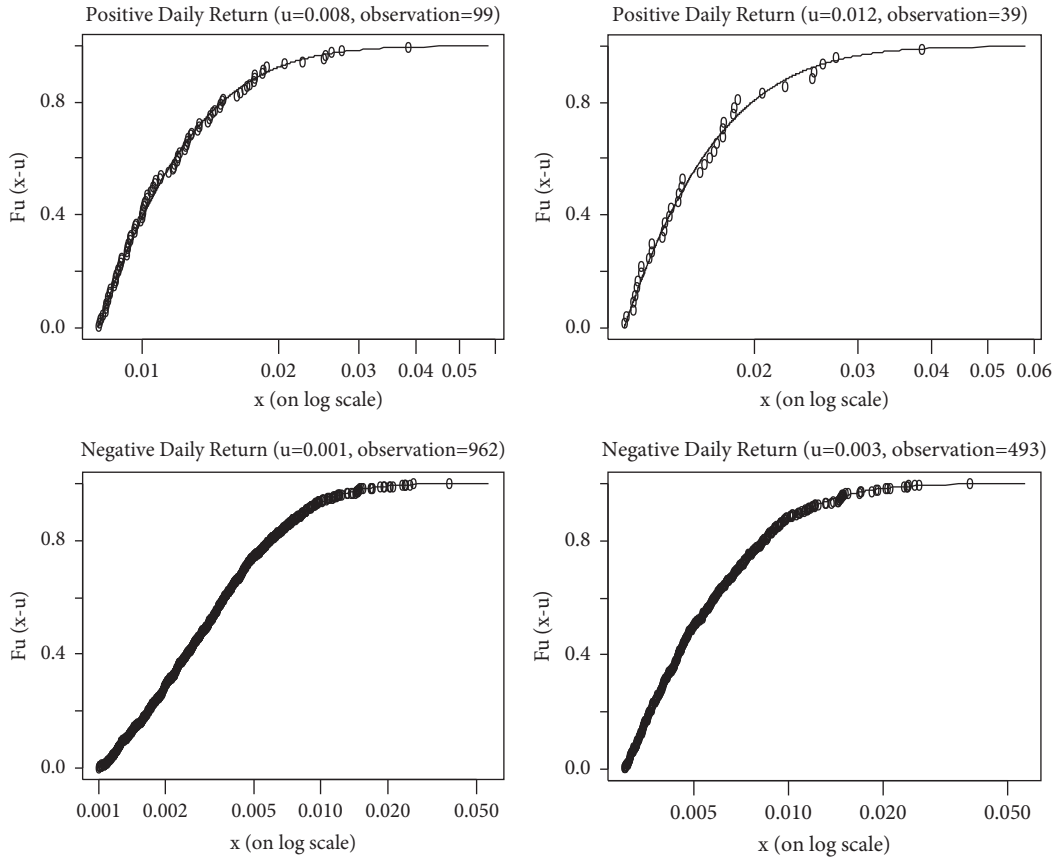
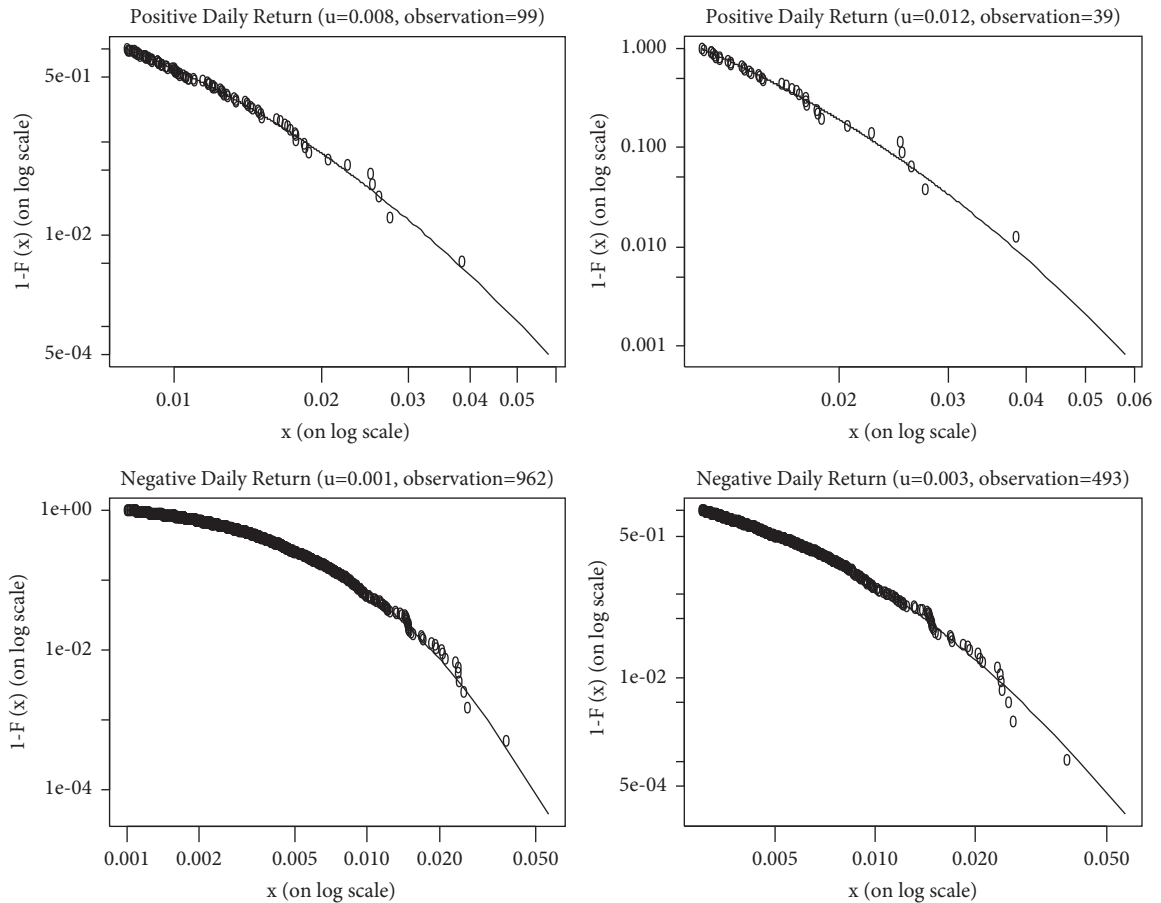FIGURE 10: Excess distribution functions' plot for positive and negative returns.

FIGURE 11: Tail of underlying distribution plot for positive and negative returns.
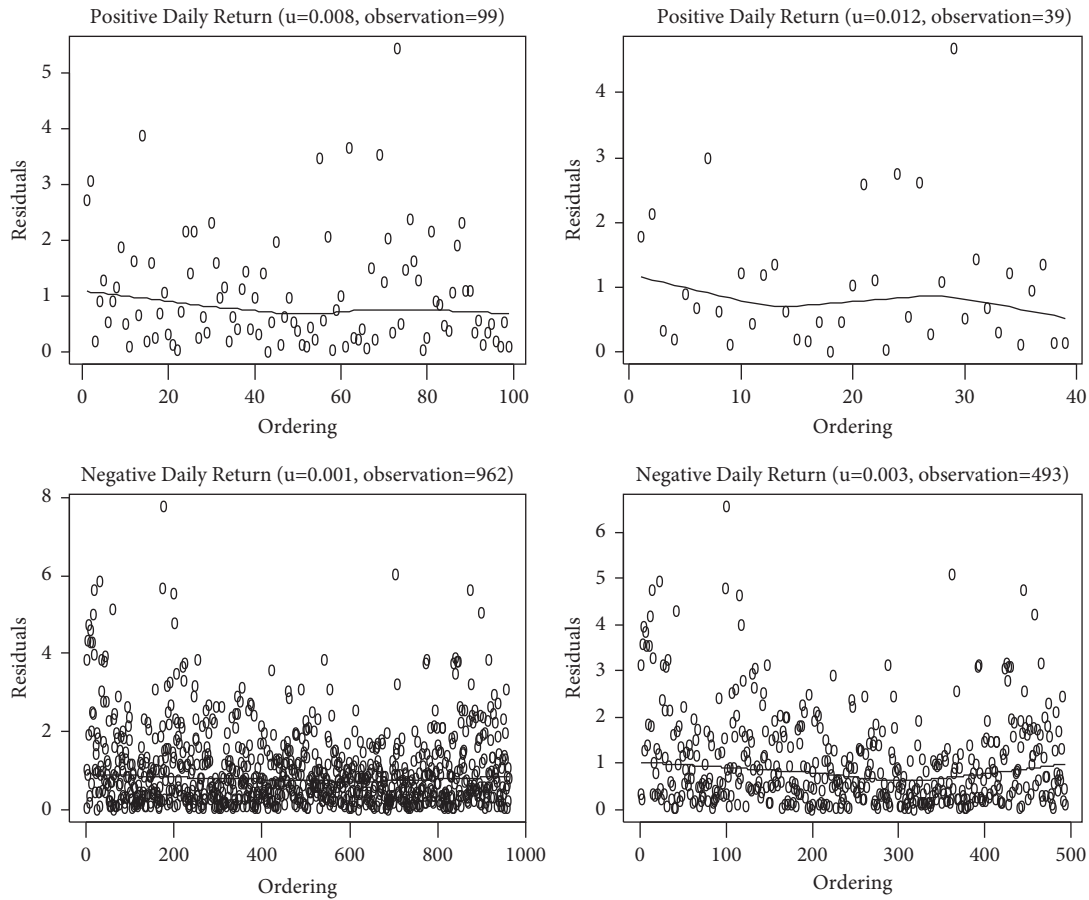
FIGURE 12: Scatter plot of residuals for positive and negative returns.

the returns' distribution has a fat tail. In fat tail distribution, the nonparametric approaches for computing VaR and ES failed. Therefore, in this paper, we used EVT for estimation of VaR and ES which is a parametric approach [14].

In this paper, Value at Risk (VaR) measures the best and worst-case scenario on the Pakistan Bullion Market value of the par-gram gold price over one trade day, given a specified degree of confidence. We first consider the cases of point estimates under the lower threshold for both tails (0.008 for the right tail and 0.001 for the left tail) with statistics shown in Table 10. The calculated VaR is 0.0325 with 1% probability (99th percentile) for the right tail. That is, given usual conditions, we expect a daily gain in the value of par-gram gold prices in the Pakistan Bullion market would not increase by more than 3.25%, and the average gain above this level will be 4.06%. In other words, the market value, with a probability of 1%, would be expected to gain by Rs.32,500 or more, and on the average basis, the gain is by Rs. 40,600 if we have an investment of 1 million Pakistani rupees in that market.

On the contrary, VaR is estimated as 0.0185 with a probability of 1% (99th percentile) for the left tail. The worst daily loss in the market value could fall below 1.85%, and the average loss below this level will be 2.42%. Put differently, if we invest 1 million Pakistani rupees in the Pakistan bullion gold market, we are 99% confident that our daily loss at

worst will not exceed Rs.18,500, and on the average basis, the loss is Rs.24,200 during one trade day.

On the contrary, VaR is estimated as 0.0185 with a probability of 1% (99th percentile) for left tail. The worst daily loss in the market value could fall below 1.85%, and the average loss below this level will be 2.42%. Put differently, if we invest 1 million Pakistani rupee in Pakistan bullion gold market, we are 99% confident that our daily loss at worst will not exceed Rs.18,500, and on the average bases, the loss is Rs.24,200 during one trade day.

Similarly, at a lower quantile of 95-level, the estimated VaR is 0.0223 for gains and 0.0112 for losses. It can be stated at, with the 95% confidence, the expected market value of par-gram gold prices would not gain by more than 2.23% for the best-case scenario or lose more than 1.12% for the worst-case scenario within the one-day duration.

The adequacy of VaR estimates' essential depends on their accuracy. The easiest way to examine this method is to construct the confidence intervals. For instance, we discuss the right tail here. An approximate 95% confidence interval constructed for the 99-level quintile VaR is (0.0257, 0.0541) under threshold is 0.008 and (0.0289, 0.0579) under threshold is 0.012; for the 95-level quantile VaR, the associated 95% confidence intervals are (0.0193, 0.0284) and (0.0232, 0.0425).
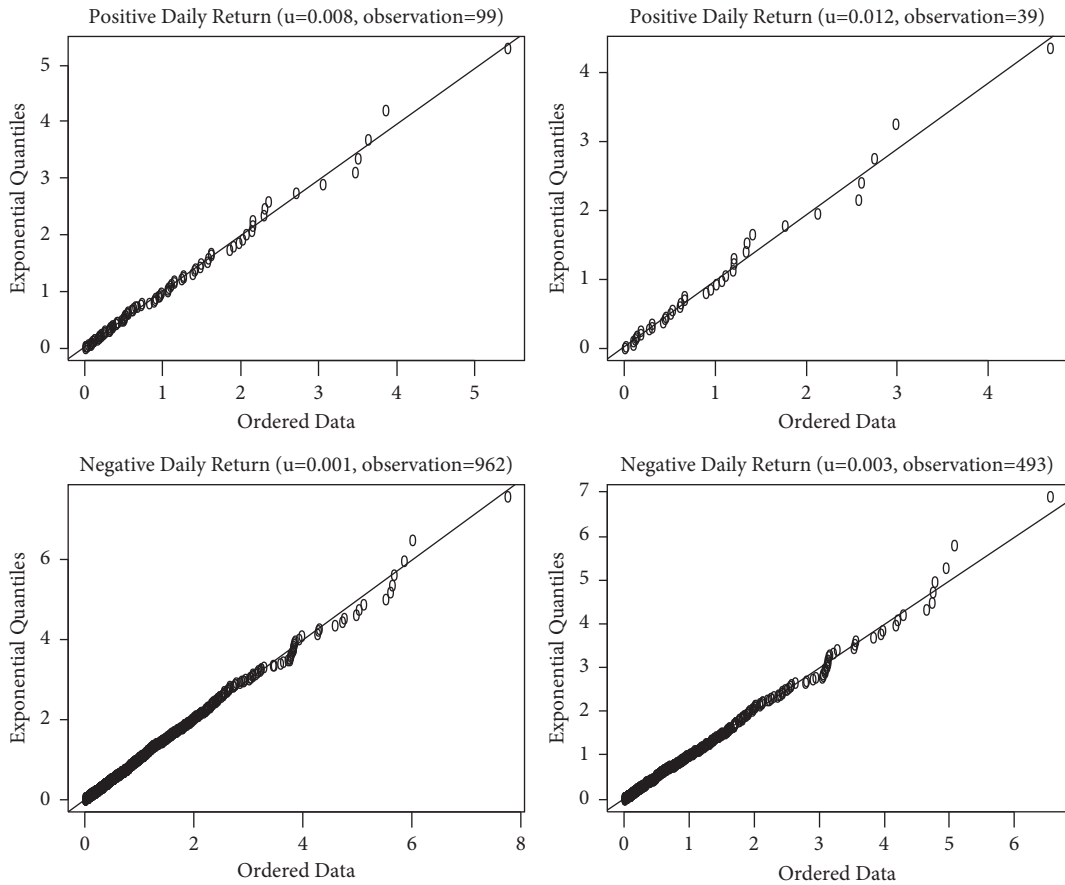
FIGURE 13: Quantile-Quantile (Q-Q) plot of residuals for positive and negative returns.

TABLE 10: Point and interval estimates of Value at Risk and Expected Shortfall of GPA distribution.

Quantile	Threshold	Estimate	CI lower	CI upper
Value at Risk for positive returns				
90-level		0.0185	0.0165	0.0217
95-level	0.008	0.0223	0.0193	0.0284
99-level		0.0325	0.0257	0.0541
90-level		0.0236	0.0204	0.0308
95-level	0.012	0.0276	0.0232	0.0425
99-level		0.0380	0.0289	0.0579
Value at Risk for negative returns				
90-level		0.0085	0.0079	0.0091
95-level	0.001	0.0112	0.0104	0.0122
99-level		0.0185	0.0165	0.0213
90-level		0.0110	0.0102	0.0120
95-level	0.003	0.0114	0.0129	0.0159
99-level		0.0231	0.0197	0.0287
Expected Shortfall for positive returns				
90-level		0.0245	0.0206	0.0357
95-level	0.008	0.0288	0.0233	0.0474
99-level		0.0406	0.0294	0.0579
90-level		0.0298	0.0244	0.0578
95-level	0.012	0.0342	0.0269	0.0579
99-level		0.0458	0.0319	0.0579

Table 10: Continued.

Quantile	Threshold	Estimate	CI lower	CI upper
Expected Shortfall for negative returns				
90-level		0.0128	0.0119	0.0142
95-level	0.001	0.0159	0.0143	0.0181
99-level		0.0242	0.0208	0.0295
90-level		0.0162	0.0144	0.0191
95-level	0.003	0.0199	0.0172	0.0246
99-level		0.0308	0.0246	0.0427

Some characteristics of the estimation can be summarized as follows: (1) under different thresholds, the estimates of VaR exhibit strong stability; (2) given the quantile level, under either the lower or higher threshold, the corresponding VaR estimate in the right tail is larger than in the left tail, but the difference is small, implying that the behavior difference in both tails is likely to be small.

4. Conclusion

The high unpredictability of gold prices in Pakistan bullion markets requires the application of effective risk management. Extreme Value Theory (EVT) is a capable method to estimate the effects of extreme events in risky markets. In this paper, we have described the EVT which can be used to model the tail-related risk measures, such as Value at Risk (VaR), Expected Shortfall (ES), and return level by applying it to the daily returns of gold prices in the Pakistan Bullion Market. We used the POT method to estimate the threshold using the mean excess plot and shape parameter plot. The MLE method was used to estimate the parameters of GEV and GPA distribution. The point and interval estimates of VaR and ES calculated under different high quantile levels show strong stability through a range of the selected thresholds, suggesting the accuracy and reliability of the estimated quantile-based risk measures.

Moreover, it is included in the paper return level forecasting in the next 5 and 10-years by analyzing the Generalized Extreme Value (GEV) distribution. It is to be noted that, with the increase in the time interval, the point estimate increases slowly and the confidence interval becomes wider. Therefore, this indicates that time is a risk factor.

Our application studies the heavy-tailed behavior in daily returns and the asymmetric characteristics in distributions, suggesting we treat positive besides negative returns separately. The results show that the EVT is a powerful technique to estimate the effects of extreme events in financial markets.

Our EVT-based VaR approach provides quantitative information for analyzing the point of potential extreme risks in bullion gold markets. Organizations and corporations may use this technique as a means of risk management. Further, for the people who want to invest in the Pakistan bullion markets, the estimates of VaR and ES provide quantitative indicators for their investment decisions.

Data Availability

We used the daily gold prices in Pakistani rupees from Bullion Rates > Gold Price History in Pakistan Rupees (PKR) for February 2007 from the following website: https://www.bullion-rates.com/gold/PKR/2007-2-history.htm.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] N. Ali, N. Zaimi, and N. M. Ali, "Statistical modelling of Malaysia trading gold price using extreme value theory approach," *Advances in Mathematics: Scientific Journal*, vol. 10, no. 1, pp. 1857–8365.

[2] K. Chaithep, S. Sriboonchitta, C. Chaiboonsri, and P. Pastpipatkul, "Value at risk analysis of gold price returns using extreme value theory," *The Empirical Econometrics and Quantitative Economics Letters*, vol. 1, no. 4, pp. 151–168, 2012.

[3] R. W. Katz, M. B. Parlange, and P. Naveau, "Statistics of extremes in hydrology," *Advances in Water Resources*, vol. 25, no. 8–12, pp. 1287–1304, 2002.

[4] A. J. McNeil, "Estimating the tails of loss severity distributions using extreme value theory," *ASTIN Bulletin*, vol. 27, no. 1, pp. 117–137, 1997.

[5] R. Gençay and F. Selçuk, "Extreme value theory and Value-at-Risk: relative performance in emerging markets," *International Journal of Forecasting*, vol. 20, no. 2, pp. 287–303, 2004.

[6] F. M. Longin, "From value at risk to stress testing: the extreme value approach," *Journal of Banking & Finance*, vol. 24, no. 7, pp. 1097–1130, 2000.

[7] E. J. Gumbel, *Statistics of Extremes*, Columbia University Press, New York, NY, USA, 1958.

[8] E. Bommier, "Peaks-over-threshold modelling of environmental data," Project Report 2014:33, Uppsala University, Uppsala, Sweden, 2014.

[9] R. A. Fisher and L. H. C. Tippett, "Limiting forms of the frequency distribution of the largest or smallest member of a sample," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 24, no. 2, pp. 180–190, 1928.

[10] M. Bhattacharyya and G. Ritolia, "Conditional VaR using EVT-towards a planned margin scheme," *International Review of Financial Analysis*, vol. 17, no. 2, pp. 382–395, 2008.

[11] F. Ren and D. E. Giles, "Extreme value analysis of daily Canadian crude oil prices," *Applied Financial Economics*, vol. 20, no. 12, pp. 941–954, 2010.

[12] J. B. Jang, "An extreme value theory approach for analyzing the extreme risk of the gold prices," *Journal of Financial Review*, vol. 6, pp. 97–109, 2007.

[13] Q. L. Chen, "The volatility of the price of gold: an application of extreme value theory," Doctoral Dissertation, University of Victoria, Victoria, Canada, 2014.

[14] K. Chinhamu, C. K. Huang, C. S. Huang, and D. Chikobvu, "Extreme risk, value-at-risk and expected shortfall in the gold market," *International Business & Economics Research Journal*, vol. 14, no. 1, pp. 107–122, 2015.

[15] A. Bücher and C. Zhou, "A horse racing between the Block Maxima method and the Peak-Over-Threshold approach," 2018, https://arxiv.org/abs/1807.00282.

[16] F. Szubzda and M. Chlebus, "Comparison of Block Maxima and Peaks Over Threshold value-at-risk models for market risk in various economic conditions," *Central European Economic Journal*, vol. 6, no. 53, pp. 70–85, 2019.

[17] I. Karahanoglu, "The VaR comparison of the fresh investment tool-BITCOIN with other conventional investment tools, gold, stock exchange (BIST100) and foreign currencies (EUR/USD VS TRL)," *Eastern Journal of European Studies*, vol. 11, no. 2, pp. 160–181, 2020.

[18] R. Wahlström, *Estimating Expected Shortfall Using an Unconditional Peaks-Over-Threshold Method under an Extreme Value Approach,* Uppsala University, Uppsala, Sweden, 2021.

[19] O. Jakata and D. Chikobvu, "Modelling extreme risk of the South African Financial Index (J580) using the generalised Pareto distribution," *Journal of Economic and Financial Studies*, vol. 12, no. 1, pp. 1–7, 2019.

[20] M. Puricelli, "Rainfall extremes modeling under shortage of data and uncertainty in the Pampean region (Argentina)," *Cuadernos de Investigación Geográfica*, vol. 44, no. 2, pp. 719–742, 2018.

[21] A. Wald and J. Wolfowitz, "An exact test for randomness in the non-parametric case based on serial correlation," *The Annals of Mathematical Statistics*, vol. 14, no. 4, pp. 378–388, 1943.

[22] J. Pickands III, "Statistical inference using extreme order statistics," *The Annals of Statistics*, vol. 3, no. 1, pp. 119–131, 1975.

[23] R. S. Tsay, *Multivariate Time Series Analysis: with R and Financial Applications*, John Wiley & Sons, Hoboken, NJ, USA, 2013.

[24] S. Coles, J. Bawa, L. Trenner, and P. Dorazio, *An Introduction to Statistical Modelling of Extreme Values (Vol. 208)*, Springer, London, UK, 2001.

[25] J. Danielsson and C. G. De Vries, "Tail index and quantile estimation with very high frequency data," *Journal of Empirical Finance*, vol. 4, no. 2-3, pp. 241–257, 1997.

[26] A. C. Davison and R. L. Smith, "Models for exceedances over high thresholds," *Journal of the Royal Statistical Society: Series B*, vol. 52, no. 3, pp. 393–425, 1990.

[27] R. L. Smith, "Maximum likelihood estimation in a class of nonregular cases," *Biometrika*, vol. 72, no. 1, pp. 67–90, 1985.

[28] C. Cunnane, "A note on the Poisson assumption in partial duration series models," *Water Resources Research*, vol. 15, no. 2, pp. 489–494, 1979.

[29] D. Rosbjerg, H. Madsen, and P. F. Rasmussen, "Prediction in partial duration series with generalized pareto-distributed exceedances," *Water Resources Research*, vol. 28, no. 11, pp. 3001–3010, 1992.

[30] G. A., Holton, History of value-at-risk: 1922–1988. Contingency Analysis Working Paper, 2002.

[31] P. Artzner, F. Delbaen, J. Eber, and D. Heath, "Thinking coherently," *Risk*, vol. 10, pp. 68–71, 1997.

[32] M. Gilli, E. Këllezi, and H. Hysi, "A data-driven optimization heuristic for downside risk minimization," *Swiss Finance Institute Research Paper, (06-2)*, Elsevier, Amsterdam, Netherlands, 2006.

[33] N. Pratiwi, C. Iswahyudi, and R. I. Safitri, "Generalized extreme value distribution for value at risk analysis on gold price," *Journal of Physics: Conference Series*, vol. 1217, no. 1, p. 012090, 2019.