Original Research Article

EXTREME VALUE THEORY MODELING OF GEOCHEMICAL ANOMALIES: BLOCK MAXIMA APPROACH

Abstracts

Geochemical anomalies are important in mineral exploration because they are connected to the development of deposits. The extreme value index (EVI) is a critical parameter in extreme value analysis since it identifies the tail heaviness of the underlying distribution and is the key parameter necessary for the estimate of subsequent extreme occurrences. In this paper, the block maxima (BM) approach of extreme value theory (EVT) was utilized to model geochemical anomalies of Au concentration using the generalized extreme value (GEV) distribution. The maximum likelihood estimation approach was used to estimate the model parameters. The EVI (ξ) estimates from the MLE were larger than -0.5, indicating that the estimators had approached asymptotic normality and that ML estimators are feasible. The Fréchet family of GEV distributions provided an excellent match for the dataset, based on the positive shape parameter estimate. Statistical inference is accomplished by evaluating multiple return levels that correspond to the return periods, with the findings suggesting that the possibility of a higher Au concentration return level is highly unlikely.

Keywords: Extreme value theory, generalized extreme value distribution, geochemical anomalies, part per billion.

1 INTRODUCTION

Ghana's most important solid mineral resources are gold, diamonds, manganese, and bauxite. For over two decades, gold has been the most important mineral produced in the country, accounting for about 90 percent of all mining profits [1]. Minerals are essential for every country's economic growth. Since minerals have such a large impact on Ghana's economy, mineral shortages can be avoided if the mineral industry appropriately predicts mineral deposits.

Mineral resource forecast sometimes has proven to be a difficult task due to it volatility. Despite having natural wealth, Ghana is vulnerable since it does not control mineral generation but dependent on foreign cash. In order to cope with new challenges, it is critical to analyze and construct a model that can accurately forecast mineral earnings. The major goal of accurate mineral production data forecasting is to make the most effective and optimal use of natural resources in economic growth. Since datasets are nonlinear, nonstationary, and have time-varying characteristics, mineral resource prediction is a critical issue. Many researchers are attempting to predict nonlinear datasets with complicated time-varying properties with high accuracy.

When looking at a data set, the goal is generally to figure out what the typical observations from the underlying distribution are. This is commonly accomplished by calculating a measure of central tendency, such as the mean, before creating a confidence interval. Normal distributions and time series methods make up the majority of common statistical techniques. These are all based on symmetric distributions, which fail to account for the tail behavior of fat (heavily) tailed and asymmetric distributions [2];[3]. In geology and mining, long-tailed or heavy-tailed distributions have been identified. High sample values in contrast to the center of the distribution are common in grade distributions for precious metals such as gold [4] and size distributions for oil and gas resources [5].

In the domain of mineral exploration and mineral resource assessment, delineating geo-anomalies to aid generation of resource estimate is a standard procedure. A geo-anomaly is distinguished from its surroundings by major changes in composition, texture, structure, and origin [6]. Although no single theory for describing the mechanisms of the formation of geochemical anomalies in deeply weathered and transported terrains with regolith or other land cover has been accepted [7], several mechanisms assume the upward migration of ions to the surface have been proposed, including electrochemistry, diffusion, groundwater pumping, convection, capillary rise, and vegetation [8][9]. Yongqing and Pengda [10] established the geological anomaly (maximum observations) as an extreme value based on a mathematical model (as discussed by [11]). According to Chen *et al* [12], extreme value theory (EVT) is a statistical science that investigates the limiting distribution of the lowest and greatest value and assesses the risk of extreme occurrences. Since most EVT applications are motivated by the need to predict the likelihood of large observations, such as geological anomalies, this thesis will focus on the right tail of the underlying distribution, which entails using the block maxima approach to look at extremely large observations (geochemical anomalies).

2.0 LITERATURE REVIEW

The three-parameter lognormal density is a well-known and often used density distribution for skew data. This method has been employed in a number of case studies [13], and the theory underlying it is well established ([14],[15]). However, the geological complexity of many mineral deposits has driven the creation of bigger distribution models. These models are based on worries about genetic deposition.

Based on large sampling databases, the compound and mixed lognormal distributions have been suggested to characterize the grade distribution of gold and the size distribution of diamonds ([16],[17],[18]). These rules are designed to correct deviations from two- or three-parameter lognormal distributions. Sichel developed a mixed lognormal distribution to describe diamond size distributions. Geological constraints were used to translate the mixing requirements of lognormal size distributions. The distribution's center can be described as more or less lognormal, with Pareto tails. It has been used in a variety of disciplines (geology). Due to the broad tail of geological data, the bulk of their distributions belong to the pareto type family, which includes lognormal. J. Caers et al [19], stressed the practical use of extreme value theory in a case study of size distributions for alluvial diamonds, which is used to identify a lognormal from a mixed log normal distribution. According to their results, the Mixed Lognormal Distribution does not provide a straightforward and practical visual strategy for estimating a distribution's tail and, as a result, discriminating between multiple feasible distribution models. J. Caers et al [19], utilized extreme value theory to illustrate the divergence from the lognormal model induced by sorting and the formation of a large mixed lognormal distribution. [20] employed GPD to simulate a geological anomaly. Using their proposed approach, anomalies in the Jiguanzui Cu-Au mining zone were detected. According to their findings, the anomalous area of Cu and Au corresponds to the spectrum of ore bodies identified in actual engineering investigation. As a consequence, they concluded that the EVT geological anomaly model can identify anomalies and has a high indicating function in mineral exploration. The block maxima approach (GEVD) in EVT is utilized in this work to model geochemical anomalies of Au.

3.METHODOLOGY

3.1 Model Formulation (Distribution of Maxima)

Extreme Value Theory (EVT) is the study of probabilistic extremes, with a particular emphasis on asymptotic behavior as sample size approaches infinity [21]. Let $X_1, X_2, .., X_n$ be a sequence of independent random variables with the same distribution, F. The model focuses on the statistical behavior of $M_n = max(X_1, X_2, .., X_n)$,where $X_i{}^s$is typically indicate values of a process monitored on a regular time scale and n is the number of observations in a day, which then corresponds to the daily maximum.

In theory, the distribution may be calculated precisely for all values of n:

$$P(M_n \leq x) = P(\max(X_1, X_2, .., X_n) \leq x)$$

=F(x)…F(x)= $[F(x)]^n$ **(1)**

In practice, the challenge derives from the fact that the distribution function F is unknown. This leads to an asymptotic method that needs establishing what feasible limit distributions are available forM_n as $n \rightarrow \infty$. This difficulty is overcome in the Central Limit Theorem (CLT) by permitting linear scaling, so that

$\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ → N (0,1),**(2)** where $\mu_n = \mu \ and \ \sigma_n$are linear re-scalings that avoid degenerate limits.

The same method is used to find the limits of the distribution of geochemical anomalies (M_n) , instead looking for limiting distributions of $\frac{M_n - b_n}{a_n}$ $b_n \ and \ a_n$ where are sequences of normalizing coefficients such that $F^n\left(\frac{M_n - b_n}{a_n}\right)$ leads to a non-degenerate distribution as $n \rightarrow \infty$. We specifically want seek { $a_n > 0$ } and { b_n } such that:

$F^n\left(\frac{M_n - b_n}{a_n}\right) \rightarrow G(z)$, (3) where G(z) is independent of n.

3.11 Extremal Types Theorem

If there are sequences of constants { $a_n > 0$ } and { b_n } such that, as n→ ∞

P$\left(\frac{M_n - b_n}{a_n} \leq z\right) \rightarrow G(z)$, where G is a non-generate distribution function, G belongs to one of the following families [20]:

i. If Ɛ = 0, $Gumbel\ type$

$\wedge(x) = e^{-e^{-x}}$, -∞ $< x < \infty$ **(4)**

The Gumbel domain is $(-\infty, +\infty)$

ii. If $\mathcal{E} < 0$, $Weibull\ type$

$$\psi_\alpha(x) = \begin{cases} e^{-(-x)^{-\alpha}}, & if\ x<0 \\ 1 & if\ x \geq 0 \end{cases} \quad \textbf{(5)}$$

The Weibull domain is $(-\infty, \mu - \frac{\sigma}{\mathcal{E}})$

iii. $if\ \mathcal{E} < 0$, *Fréchet type*

$$\phi_\alpha(x) = \begin{cases} 0 \quad , & if\ x < 0 \\ e^{-x^{-\alpha}}, & if\ x \leq 0 \end{cases} \quad \textbf{(6)}$$

The Fréchet domain is $(\mu - \frac{\sigma}{\mathcal{E}}, \infty)$

According to fisher Tippet theorem [22], the three family of distribution can combine to a single distribution called generalized Extreme value (GEV)

The distribution GEV is given:

$$G(x)=\exp[-(1+\mathcal{E}\,(\frac{X-\mu}{\sigma})]^{\frac{-1}{\mathcal{E}}}, \text{where if } 1+\mathcal{E}\left(\frac{X-\mu}{\sigma}\right) > 0 \quad \textbf{(7)}$$

GEV is made up of three parameters:

1. $\mathcal{E}$ shape parameter (Extreme Value Index), $-\infty < \mathcal{E} < \infty$
2. σ - Scale parameter (Dispersion of the M_n), $\sigma > 0$
3. μ - location parameter (mean of the M_n), $-\infty < \mu < \infty$

3.2 Maximum Likelihood (ML) Method for GEVD

To estimate the parameters of the GEV distribution, you can use Maximum Likelihood (ML), Probability Weighted Moments (PWM), or the L-moments technique. The ML methodology has an advantage over other parameter estimation methods in that it can adapt to changes in model structure. That is, even if the estimating equations of a model vary, the essential process remains mostly same. Furthermore, the ML has a set of "off-the-shelf" huge sample inference properties [23].

Let's use a GEVD distribution on x_1,...,x_n independent random variables (block maxima). The log likelihood function is as follows when $\xi \neq 0$:

$$\ell(x_i,\mu,\sigma,\xi) = -n\log\sigma - \left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n}\log\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right] - \sum_{i=1}^{n}\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}} \quad (8)$$

where i = 1,...,n and 1+ξ $(\frac{x_{i-\mu}}{\sigma})$ > 0,[22]. If this requirement is not met, L (x_1, x_2 ..., x_n ; μ, σ, ξ) = 0 and ℓ (x_1, x_2...,x_n ;μ,σ, ξ) =−∞. In a specific case when ξ = 0 the log likelihood for equation (8) is as follows:

$$\ell(x_i,\mu,\sigma) = -n\log\sigma - \sum_{l=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right) - \sum_{j=1}^{n}\exp\left\{-\left(\frac{x_i-\mu}{\sigma}\right)\right\} \text{[23] (9).}$$

The MLE technique is the most often used parametric inference tool in most statistical models, including EVT. MLE, on the other hand, has certain limitations when it comes to determining the form parameter of the GEVD [24].

The MLE approach computes point estimates of form parameters while taking into account certain characteristics, and there are some circumstances when the shape is not estimable using the MLE. When $\xi > -0.5$, MLEs meet the standard requirements.

3.3 Exceedance Probability and Return level

After the estimations of μ, σ and ξ and the GEV distribution have been obtained, extreme quantiles, exceedance probability, and return levels and their corresponding periods are all equally important extreme occurrences in any extreme value investigation. An important aim of extreme value modeling is to estimate an extreme quantile that corresponds to a certain return period, q, or the biggest value recorded in q years [25]. The extreme quantile of the GEV may be determined by inverting the GEV distribution function.

$$q_{x,p} = \begin{cases} \bar{\mu} - \frac{\bar{\sigma}}{\bar{\xi}}\left[1-\{-log-p)\}^{\bar{\xi}}\right], \bar{\xi} \neq 0 \\ \bar{\mu} - \bar{\sigma}\log[-log(1-p)], \quad \bar{\xi} = 0 \end{cases} \quad (10)$$

where H $(q_{x,p}) = 1 - p \; and \; 0 \leq p \; \leq 1$

The parameters μ, σ, ξ are the corresponding ML estimations of $\hat{\mu}$, $\hat{\sigma} and \; \hat{\xi}$ respectively. In layman's terms, $q_{x,p}$ is the return level associated with the 1/p return period.

3.4 Model Verification for the Generalized Extreme Value Distribution

According to [23], judging the correctness of an extrapolation from a GEV model is difficult, albeit some judgment may be made based on observable data. The probability plot, quantile plot, and density plot are useful graphical tools for evaluating a GEV model's goodness-of-fit [22]. To compare the empirical and fitted distributions, a probability plot and a quantile plot are utilized.

If we consider $x_{(1)} \leq \; x_{(2)} \leq \cdots . \leq x_{(k)}$, the empirical distribution function evaluated at $x_{(i)}$ is defined as $\tilde{F}(x_{(i)} = \frac{i}{k+1}$ [23].

The model-based estimations are derived as follows:

$$\hat{F}(x_{(i)}) = exp\{-[1 + \; \hat{\gamma}(\frac{x_i - \hat{\mu}}{\hat{\sigma}})]^{\frac{-1}{\hat{\gamma}}}\} \quad (11)$$

If the GEV model is applicable, $\tilde{F}\left(x_{(i)}\right) \approx \; \hat{F} \; (x_{(i)}; i = 1,2, \dots , k$ [23]. As a result, a probability plot will consist of the points [23]

$$\{(\hat{F}(x_{(i)}), \hat{F} \; (x_{(i)}); i = \; 1,2, \dots , k\}$$

and should result in a straight one-to-one line of points; any deviation from linearity shows that the GEV model is failing. Furthermore, the quantile plot is made up of points. [23]

$\{\hat{F}^{-1}$ (i/k+1), $x_{(i)}); i = \; 1,2, \dots , k\}$

As a result of (11) we have [23].

$$\hat{G}^{-1}(\frac{1}{k+1}) = \hat{\mu} - \frac{\hat{\sigma}}{\hat{\gamma}} \; [1 - \{-log(\frac{1}{k+1})^{-\hat{\gamma}}]$$

The quantile plot should likewise show a straight one-to-one line of points. Any deviation from linearity in the graphic indicates the model's failure.

4. RESULTS AND DISCUSSION

4.1 Source of Data

The research's data came from 15 exploration lines acquired from Small-Scale Mining Company in Wassa-Amenfi in Western Region of Ghana. A total of 2750 soil samples were collected on grid lines, through soil auger drilling procedure to an average depth of 3 m. The Ghana Geological Survey Department used graphite furnace-atomic absorption spectrometry (GF-AAS) [26,27] to determine the gold content in parts per billion in these samples.

4.2 Summary Statistics

This section's descriptive analysis highlights the most essential aspects of the data collected. The table 1 below provides a short overview of all the factors under investigation.

Table 1: A short overview of all the factors under investigation.

Mean	Maximum	Minimum	Std.	Skewness	Kurtosis
52.07	990.0	0.1	78.19	6.25	46.45

In general, the index's maximum and lowest Au contents are considerably distinct. The standard deviation is also high, indicating that Au content varies greatly. Positive skewness is also apparent, showing that the right tail is especially severe, indicating non-symmetric yields for the Au concentration.

4.3 Stationarity Test

The augmented Dickey-Fuller (ADF) and sequence correlation analysis are the most common approaches employed in stable test zones to analyze the self-correlation of geological data. The ADF may be used to retrieve the sample data's test results (Table 3.2). The null hypothesis that the data is not stationary should be rejected since Au's p-value of 0.067 is within the significant level 5 percent rejection zone. As a result, the Au sequences do not have unit roots; instead, they are stationary sequences that confirm the stationary tendency of the time series.

Table 2: The ADF test of Au Concentration

Null hypothesis: Data is not Stationary

Element	DF- value	P-Value
Au	-13.32	0.067

4.4 Block Maxima Approach (GEVD)

Because the samples are stationary, the maximum Au concentration is modeled using a stationary model. Although stationarity does not ensure independence, we may assume the data are independent and use the traditional EVT to characterize the stationary sequence as an independent sequence because the blocking technique decreases data reliance. The block maxima method is currently used to compute our daily gold (Au) concentration in part per billion (ppb). The correct selection of the periods that define the blocks is critical to the success of this method. The recommended periods are block sizes of 30 observations due to the nature of the gold concentration and also to ensure that there is enough data for the extremal type theorem to hold. The research period's data (i.e., 2750 observations) is divided into 30 non-overlapping sub-samples, with the highest observed value picked. As a consequence, there are 91 observations in all, each of which provides the daily Au concentration.

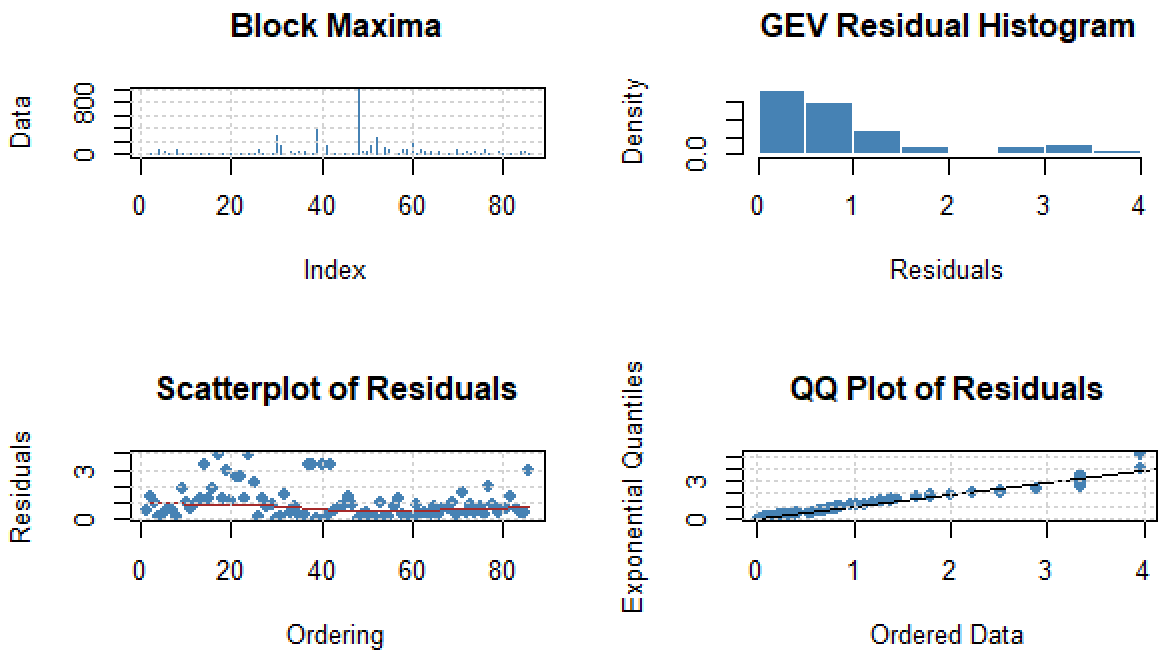

Figure 1: The daily maximum Au conc. of GEV fit.

A sample plot of block maxima on the left and a histogram depicting the GEV of residuals on the right are the two top plots in Figure 1. (right). At the bottom is a scatter plot of the residuals that shows the time of block maxima. The purpose is to search the data for a possible temporal trend. To aid in judging this, a simple fitted curve (using the fExtremes package in R) is superimposed, and there is evidence of a systematic trend. The solid line is the smoothing of scattered residuals obtained by a spline method. In the bottom right corner, the residuals are presented in a QQ-plot. The QQ plot of the fitted model, which is based on the GEV fitted to all 91 block maxima, stays close to the straight line.

4.5 Parameter Estimation (MLE)

The dataset is fitted to the GEV model using the maximum likelihood estimator (MLE). The positive shape parameter (ξ) estimates in table 3 indicate that the underlying distribution corresponds to the Fréchet domain of attraction. Furthermore, the Wald confidence interval for the form parameter excludes 0, indicating that the Au concentration distribution is beyond the Gumbel region of attraction.

Table 3: Parameter Estimates for the GEV Model Using MLE

Parameter	Estimates	Standard Error	Confidence Interval
Scale (σ)	7.1810	0.14276	6.9901 - 7.4608
Location (μ)	6.2696	0.11300	6.0481 - 6.4910
Shape (ξ)	0.9913	0.02088	0.9503 – 1.0321

According to [28], the profile log-likelihood provides greater accuracy for estimating confidence intervals, as shown in Figures 2a, 2b, and 2c.

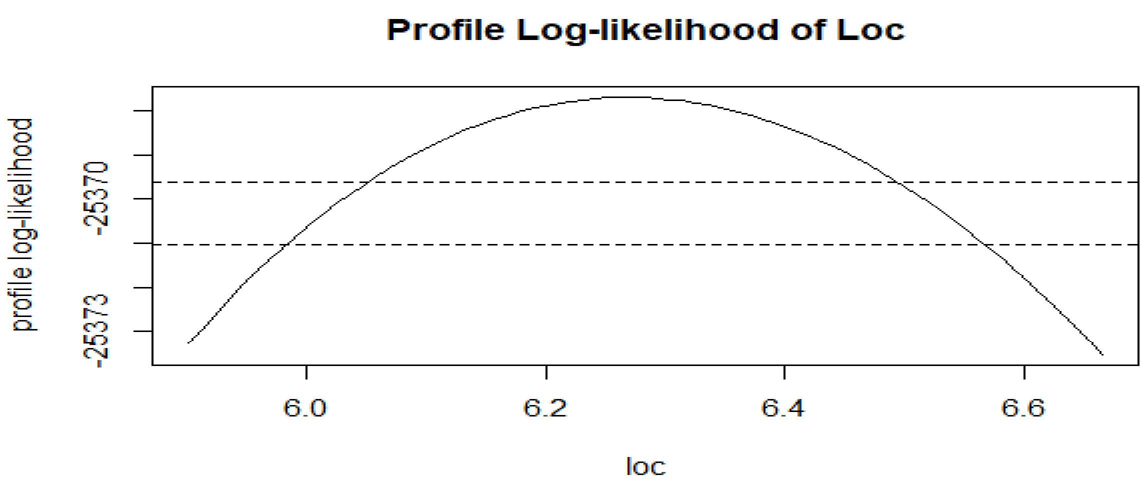

Figure 2a: Profile Likelihood for Location parameter

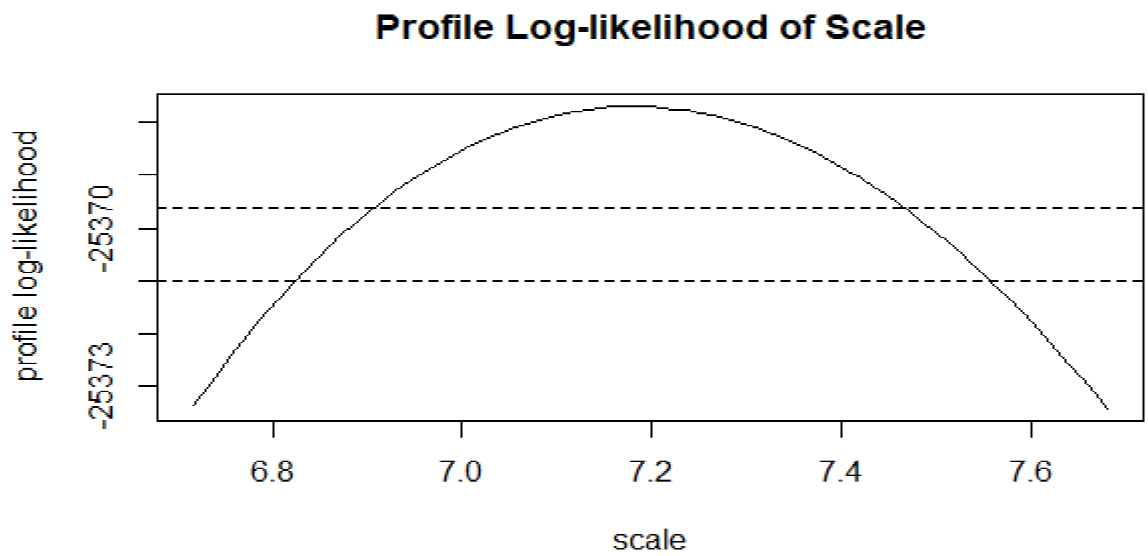

Figure 2b: Profile Likelihood for Scale parameter

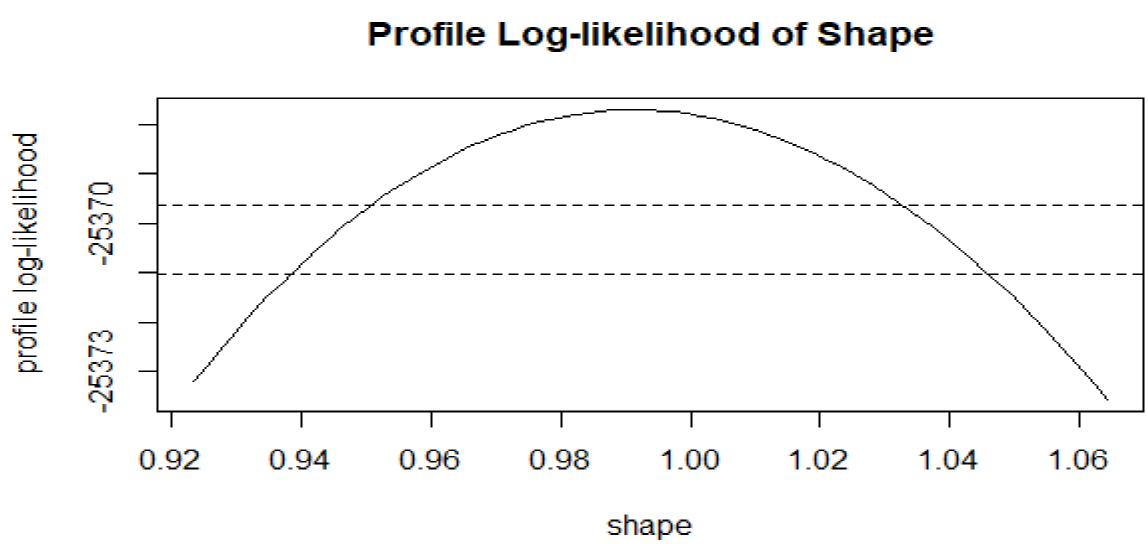

Figure 2c: Profile Likelihood for Shape parameter

4.6 Model Diagnostics

The fitted GEV models for the monthly maximum gold concentration are further evaluated using diagnostic plots. A QQ-plot (a), a PP-plot (b), and a histogram overlay of the fitted GEV's density curves are shown in the pictures (c). The graphs are shown in Figure 3. For the QQ-plot and the PP-plot, a good fit should result in a straight one-to-one line of points. In most cases, the QQ plot is preferred over the probability plot. Because the QQ-plot and the PP-plot are both linear, the model must be right. As a consequence, the GEV model fitted the data well. Based on the histogram, the density appears to be in line with the data points. As a result, we conclude that the diagnostic plots support the fitted model.

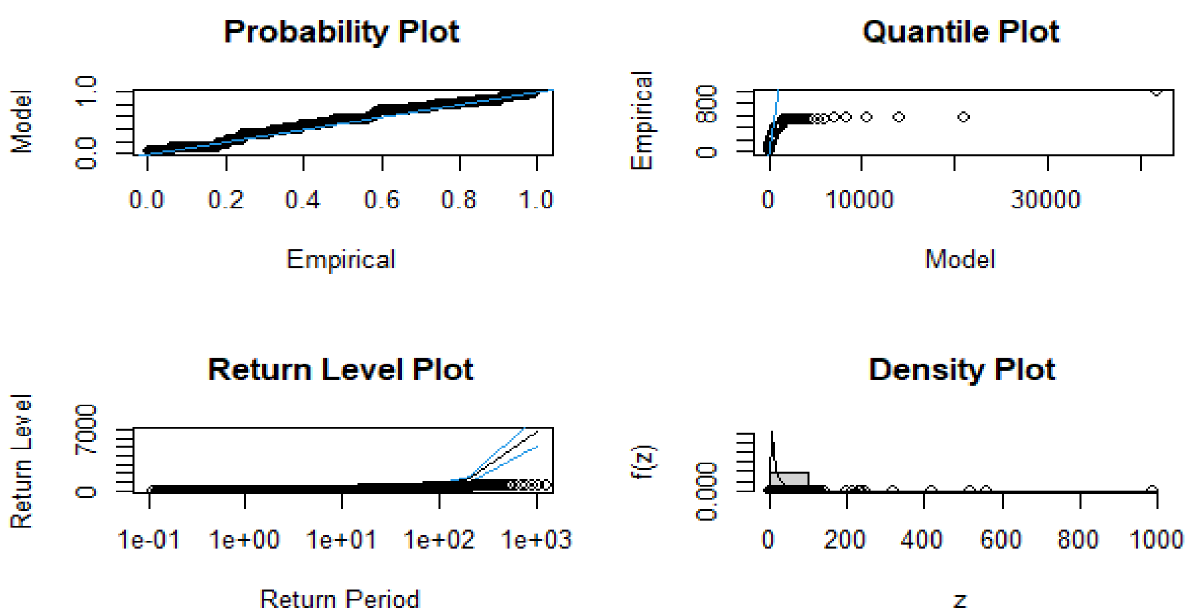

Figure 3. Diagnostic Plots for GEV

The diagnostic graphic and positive estimate derived from the shape parameter clearly show that the GEV model belongs to the Fréchet domain of attraction; we then estimate the distribution's lower endpoint. The GEV in the Fréchet domain distribution, according to [23], has an unlimited upper bound. The GEV model is defined as follows:

$$G(x)=\exp[-(1+0.99\,(\frac{X-6.27}{7.18})]^{\frac{-1}{0.99}}$$

Since GEVD has an infinite upper bound and a finite lower bound in the Fréchet domain. The term $\mu - \sigma/\xi$ represents the lower limit. According to the research, the bottom bound for the left tail is -0.974, which is substantially lower than the lowest gold concentration ever recorded for the time period in question.

4.7 Return level and Exceedance Probability

The return level of the distribution's tails and their chance of exceeding it were calculated in table 4.

Table 4: **Estimate of Return level and Exceedance Probability**

Return Period	Return level	Exceedance Probability
2 years	9.4432	0.11
5 years	31.067	0.03
10 years	66.433	0.06
20 years	136.63	0.007
50 years	345.58	0.003
100years	691.413	0.001

According to our findings in table 4, larger gold concentrations are associated with a lower right-tail exceedance probability, meaning that the possibility of a return level with a higher value is extremely low.

5. CONCLUSION

In the mining industry, modeling the frequency of occurrence of Au concentration is critical for assessing the effects of maximum Au concentration on ore deposit production. The GEVD is used to model the maximum daily Au concentration in the western part of Ghana. The diagnostic tools, the P-P and Q-Q plots, which are shown in Figures 1 and 3, respectively, are used to establish this. The Fréchet family is a suitable distribution for modeling maximum daily Au concentration in Ghana due to the positive value and asymptotic behavior of the shape parameter of GEVD. The validity of this claim is established by calculating the confidence interval for the shape parameters, which is determined to be within positive ranges. The outcome of this results was confirmed by Qin et al [20]. In their study, Generalized Pareto Distribution (GPD) approach of EVT was applied to geological anomaly and concluded that geological anomalies of Au concentration are heavy tail distribution and belong to Fréchet domain of attraction. Statistical inference was performed by examining numerous return levels matching to the return periods, with the results indicating that the likelihood of a return level with a greater value is extremely unlikely.

REFERENCES

[1] Aryee, B. N. (2001). Ghana's mining sector: its contribution to the national economy. *Resources Policy*, *27*(2), 61-75.

[2] Gençay, R., Selçuk, F., & Whitcher, B. (2005). Multiscale systematic risk. *Journal of International Money and Finance*, *24*(1), 55-70.

[3] Wentzel, D. C., & Mare, E. (2007). Extreme value theory—An application to the South African equity market. *Investment Analysts Journal*, *36*(66), 73-77.

[4] Sichel, HS*, Kleingeld, WJ** & Assibey-Bonsu, W. (1992). A comparative study of three frequency-distribution models for use in ore evaluation. *Journal of the Southern African Institute of Mining and Metallurgy*, *92*(4), 91-99.

[5] Seyedghasemipour, S. Javad, and B. B. Bhattacharyya. "The loghyperbolic: An alternative to the lognormal for modeling oil field size distribution." *Mathematical geology* 22.5 (1990): 557-571.

[6] Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A. and Gocayne, J.D., 2001. The sequence of the human genome. *science*, *291*(5507), pp.1304-1351.

[7] Mishra, A. K., & Singh, V. P. (2010). Changes in extreme precipitation in Texas. *Journal of Geophysical Research: Atmospheres*, *115*(D14).

[8] Anand, R.R. and Robertson, I.D., 2012. The role of mineralogy and geochemistry in forming anomalies on interfaces and in areas of deep basin cover: implications for exploration. *Geochemistry: Exploration, Environment, Analysis*, *12*(1), pp.45-66.

[9] Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr—Temporal segmentation algorithms. *Remote Sensing of Environment*, *114*(12), 2897-2910.

[10] Yongqing, C., & Pengda, Z. (1998). Zonation in primary halos and geochemical prospecting pattern for the Guilaizhuang gold deposit, eastern China. *Nonrenewable Resources*, *7*(1), 37-44.

[11] Bhole, M. V., Manson, A. L., Seneviratne, S. L., & Misbah, S. A. (2012). IgE-mediated allergy to local anaesthetics: separating fact from perception: a UK perspective. *British journal of anaesthesia*, *108*(6), 903-911.

[12] Chen, M., Suzuki, A., Thakkar, S., Yu, K., Hu, C., & Tong, W. (2016). DILIrank: the largest reference drug list ranked by the risk for developing drug-induced liver injury in humans. *Drug Discov Today*, *21*(4), 648-653.

[13] Aitchison, J., & Brown, J. A. C. (1969). THE LOGNORMAL DISTRIBUTION, WITH SPECIAL REFERENCE TO ITS USES IN ECONOMICS,

[14] Johnson, N. L., & Kotz, S. (1975). A vector multivariate hazard rate. *Journal of Multivariate Analysis*, *5*(1), 53-66.

[15] Sichel, H. S. (1987). Some advances in lognormal theory. In *Proceedings qfthe Twentieth International Symposium on Application qf Computers and Mathematics in the Mineral Industries* (pp. 3-8).

[16] Sichel, H. S. (1973). Statistical valuation of diamondiferous deposits. *Journal of the Southern African Institute of Mining and Metallurgy*, *73*(7), 235-243.2

[17] Sichel, HS*, Kleingeld, WJ** & Assibey-Bonsu, W. (1992). A comparative study of three frequency-distribution models for use in ore evaluation. *Journal of the Southern African Institute of Mining and Metallurgy*, *92*(4), 91-99. 2

[18] Caers, Jef, Jan Beirlant, and Marc A. Maes. "Statistics for modeling heavy tailed distributions in geology: Part I. Methodology." *Mathematical geology* 31, no. 4 (1999): 391-410. [2

[19] Caers, J., Vynckier, P., Beirlant, J., & Rombouts, L. (1996). Extreme value analysis of diamond-size distributions. *Mathematical geology*, *28*(1), 25-43.

[20] Qin, F., Liu, B., & Guo, K. (2016). Using EVT for Geological Anomaly Design and Its Application in Identifying Anomalies in Mining Areas. *Mathematical Problems in Engineering*, *2016*.

[21] Diebold, F. X., Schuermann, T., & Stroughair, J. D. (1998). Pitfalls and opportunities in the use of extreme value theory in risk management. In *Decision technologies for computational finance* (pp. 3-12). Springer, Boston, MA.

[22] Stephenson, A., & Tawn, J. (2004). Bayesian inference for extremes: accounting for the three extremal types. *Extremes*, *7*(4), 291-307.

[23] Moritz, M. A. (1997). Analyzing extreme disturbance events: fire in Los Padres National Forest. *Ecological Applications*, *7*(4), 1252-1262.

[24] Coles, S., Bawa, J., Trenner, L., & Dorazio, P. (2001). *An introduction to statistical modeling of extreme values* (Vol. 208, p. 208). London: Springer.

[25] Sigauke, C., Verster, A., & Chikobvu, D. (2013). Extreme daily increases in peak electricity demand: Tail-quantile estimation. *Energy Policy*, *53*, 90-96.

[26] Jonathan, P., & Ewans, K. (2013). Statistical modelling of extreme ocean environments for marine design: a review. *Ocean Engineering*, *62*, 91-109.

[27] Xie, X.J.; Wang, X.Q.; Zhang, Q.; Zhou, G.H.; Cheng, H.X.; Liu, D.; Cheng, Z.Z.; Xu, S.F. Multi-scale geochemical mapping in China. Geochem. Explor. Environ. Anal. 2008, 8, 333–341. [CrossRef]

[28] Xie, X.J.; Mu, X.Z.; Ren, T.X. Geochemical mapping in China. J. Geochem. Explor. 1997, 60, 99–113