UNP JOURNAL OF STATISTICS AND DATA SCIENCE
Volume 4 Nomor 3 Tahun 2026 Halaman 306-319.
ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Modeling Tail Risk of Employment Injury Claims using the Extreme Value Theory Framework

Sri Dewi Anugrawati*, **Nurwahidah, Nur Aeni, Risnawati Ibnas**
Department of Mathematics, Universitas Islam Negeri Alauddin Makassar, Gowa, Indonesia
*Corresponding author: sridewi.anugrawati@uin-alauddin.ac.id

Submitted : 02 April 2026
Revised : 28 Agustus 2026
Accepted : 30 Agustus 2026

ABSTRACT

This study investigates the implementation of Extreme Value Theory (EVT) in modeling tail risk associated with employment injury insurance claims, addressing the critical need for robust risk estimation in the presence of rare, high-severity losses. The dataset consists of 1,177 historical claims from BPJS Ketenagakerjaan Makassar recorded between 2016 and 2023. Preliminary diagnostic procedures, including hill plots and mean excess function analysis, reveal heavy-tailed behaviour with a tail index greater than one, implying the possibility of infinite variance—a characteristic often underestimated by traditional actuarial models. Due to the limited number of extreme exceedances, the block maxima framework was adopted, and a Generalized Extreme Value (GEV) distribution was fitted to weekly maxima using L-moment estimation to ensure parameter convergence. The resulting GEV parameters (location = 50.27, scale = 67.19, shape = 0.443) confirm a Fréchet-type distribution, consistent with a heavy-tailed risk profile. Comparative analysis shows that the GEV model yields significantly more conservative risk measures than the empirical approach. At the 99.5% confidence level, the GEV-VaR (1484.09 million IDR) is over three times higher than the empirical estimate (450.56 million IDR), while the GEV-ES reaches 2659.08 million IDR. Additionally, return level analysis indicates a one-year return level of 768.51 million IDR. These findings demonstrate that empirical methods substantially underestimate extreme loss potential, potentially threatening insurer solvency. Overall, this study provides a statistically rigorous framework for capital reserve and reinsurance optimization, offering a more conservative and theoretically justified approach to managing catastrophic occupational risks.

Keywords: *Extreme Value Theory, Block Maxima, Value at Risk, Expected Shortfall, Employment Injury Insurance, Heavy-tailed Distribution.*

I. INTRODUCTION

The practice of worker insurance in Indonesia has been developed over many years. The government committed to ensuring that all workers hold insurance policy to cover risks encountered in the workplace. Consequently, insurance companies are required to pay out employee policies in accordance with Indonesian government regulations. Injury insurance is a form of worker coverage that indemnifies insured risks occurring at the workplace or illnesses caused by the work environment. Once a claim is approved, the insurer must pay a specified amount based on the policy contract. Managing these agreements is crucial for insurance companies, as they must maintain adequate liquid reserves to fulfill contractual obligations by accurately forecasting future claim payouts. Predicting total losses entails not only determining the required capital reserves to meet obligations, but also understanding the risk exposure associated with each claim payout. The capacity to absorb these losses is vital to ensuring an insurer's solvency and overall financial health. Therefore, insurance companies must periodically estimate their Value at Risk (VaR) for benefit claims.

Risk measurement is essential for assessing corporate insolvency risks within a specified confidence interval, as it quantifies potential capital losses over a given time horizon (Alexander, 2010). Various methods have been developed to evaluate risk in the insurance sector, including Monte Carlo simulation under Weibull, Burr XII, Birnbaum-Saunders, and Pareto distributions (Abbasi & Guillen, 2013; Wang et al., 2005) as well as semiparametric models in life insurance designed to optimize Conditional Value at Risk (CVaR) (Staino et al., 2023).

However, the risk profile of employment injury insurance claims deviates significantly from other lines, such as general health or personal accident insurance. While health insurance typically deals with high-frequency but relatively

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

low-severity claims (thin-tailed), worker compensation claims are characterized by unique "fat-tailed" behavior. In occupational safety, rare but catastrophic events such as major industrial accidents can result in claim amounts that far exceed the average. Traditional actuarial models frequently fail to capture these extreme tails, leading to a severe underestimation of risk (Embrechts et al., 1997; McNeil, 1997).

While VaR estimations using Extreme Value Theory (EVT) have been applied to securities (Williams et al., 2018), the stock markets (Kratz, 2019; Najamuddin et al., 2024; Singh et al., 2011), energy markets (Marimoutou et al., 2009), fire insurance (Wainaina & Waititu, 2014), and motor vehicle insurance (Adesina et al., 2016), a noticeable research gap remains regarding the extreme tail risk of employee injury claims. This study fills this critical gap by developing a comprehensive risk estimation framework using EVT. The novelty of this research lies in its specific application to occupational injury data, which exhibits severe asymmetry and heavy-tailedness, and the comparative analysis between the Block Maxima, Peak Over Treshold (POT), and empirical approaches. Furthermore, this study follows a structured analytical framework: first, identifying tail behavior through diagnostic plots; second, estimating parameters using the L-moments framework to ensure estimation stability; and third, calculating both VaR and Expected Shortfall (ES) to establish a more robust safety buffer for insurers compared to traditional actuarial methods.

II. METHODOLOGY

A. Data and Pre-processing

The empirical analysis utilizes claim severity data from workers accident insurance provided by the Social Security Agency for Health (BPJS Ketenagakerjaan) in Makassar City, Indonesia. The dataset comprises 1,177 historical claim records spanning from 2016 to 2023. To ensure data integrity, a pre-processing stage was conducted, including the removal of incomplete records and the synchronization of claim dates into weekly intervals. The data exhibits high leptokurtosis and positive skewness, which justifies the departure from Gaussian-based models toward Extreme Value Theory (EVT).

B. The EVT Framework: Selection between GEV and GPD

The extreme value theory model is a concept for modeling rare events that is commonly applied in finance, insurance, medicine, climate change, and many other fields. Rare events refer to events that occur with a small probability (Dey & Yan, 2016). This probability forms an extreme value distribution which describes the boundary behaviour of the maximum or minimum values of n independent and identically distributed random variables, with a certain normalization constant. For instance, if $X_1, X_2, \cdots, X_n$ represent n independent and identically distributed random variables, then the distribution of the largest order statistic $X_{n,n}$, when it has a non-degenerate bounded distribution, will approach one of three types of extreme value distributions as $n \to \infty$: (1) Gumbel Distribution, (2) Frechet distribution, (3) Weibull distribution (Ahsanullah, 2016). The three types of distribution can be combined into single distribution function called generalized extreme value (GEV) with the form

$$G(z) = exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} \tag{1}$$

which defined on the set $\left\{z: 1+\xi\frac{(z-\mu)}{z} > 0\right\}$ with parameters $-\infty < \mu < \infty$, $\sigma > 0$, and $-\infty < \xi < \infty (Coles, 2001)$. Another distribution that used in extreme values modelling is Generalized Pareto Distribution (GPD). It is an asymmetric distribution to model exceeding high threshold in many field (W. Chen et al., 2024; McNeil & Saladin, 1997; Song & Song, 2012) with cumulative distribution function,

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

$$G(z)=\begin{cases}1-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}, \xi \neq 0\\ 1-exp\left[\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}, \xi = 0\end{cases} \quad (2)$$

where $\xi, \mu \in \mathbb{R}$, and σ are the shape, location, and scale parameters respectively. If $\mu = 0$, the GPD will reduce to the two-parameters GPD (ξ, σ). To identify the most robust model for occupational risk, we use:

1. Block Maxima (GEV): Data are partitioned into weekly block and the maximum claim within each block is extracted. These maxima are fitted to the Generalized Extreme Value (GEV) distribution. The choice of block size in the block maxima approach involves a fundamental trade-off between bias and variance. Larger block sizes such as monthly or annual blocks ensure better convergence to the asymptotic GEV distribution reducing bias but severely reduce the number of observations available for parameter estimation. Parameters are estimated using the L-Moment method to ensure convergence, especially when dealing with the heavy-tailed nature of industrial accident data where Maximum Likelihood Estimation (MLE) may fail.
2. Peaks-over-Threshold (GPD): This approach focuses on all observations exceeding a high threshold (μ). The threshold selection is determined through diagnostic tools, specifically the Mean Excess Plot and Hill Plot, to ensure the excesses follow a Generalized Pareto Distribution (GPD).

The assumption of stationarity of the weekly maximum claim series spanning from 2016 to 2023 will be evaluated using the formal Augmented Dickey-Fuller (ADF) test to ensure the statistical validity of the Generalized Extreme Value (GEV) framework with the hypothesis $H_0 : \gamma = 0$ posits that the weekly maxima series contains a unit root (is non-stationary), while the alternative hypothesis $H_0 : \gamma < 0$ implies that the process is stationary(Dickey & Fuller, 1979; Gujarati, 2021).

Moreover, a formal goodness-of-fit assessment is required to ensure that the fitted GEV or GPD distribution accurately represents the stochastic behavior of the data. Standard distance-based tests, such as the Kolmogorov-Smirnov (KS) and Anderson-Darling (AD) tests, conventionally assume that the distribution parameters are fixed and known a priori. However, when the location ($\hat{\mu}$) , scale ($\hat{\sigma}$), and shape ($\hat{\xi}$) parameters are estimated directly from the empirical sample using L-moments, these standard tests become overly conservative, leading to severe Type I error inflation and systematic over-rejection of the null hypothesis. This study implements a Parametric Bootstrap Goodness of Fit procedure based on the Anderson-Darling statistic (A_n) for the GEV distribution. The computational framework is structured as follows (Coles, 2001; Laio, 2004):

1. The empirical GEV parameters are estimated from the original N weekly maximum observations, and the baseline test statistic ($A_{n,obs}$) is calculated.
2. Synthetic datasets of size N are generated by sampling parametrically from a GEV.
3. For each of the $B = 1{,}000$ bootstrap replications, the GEV parameters are re-estimated via L-moments from the synthetic data, and the bootstrap test statistic ($A_{n,b}$) is computed against its own estimated parameters.
4. The bootstrapped p-value is determined as the proportion of bootstrap statistics that exceed or equal the observed statistic

$$p_{boot} = \frac{1}{B}\sum_{b=1}^{B} I\left(A_{n,b} \geq A_{n,obs}\right) \quad (3)$$

where $I(.)$ denotes the indicator function. A $p_{boot} > 0.05$ indicates that the null hypothesis cannot be rejected, formally validating the GEV model adequacy.

C. Risk Metrics: VaR and Expected Shortfall (ES)

Risk is a measure of the uncertainty of future losses for any given date horizon. It is related to the variability of the future value of a position, due to market changes or uncertain events (Artzner et al., 1999). To quantify the potential financial loss, we estimate two primary risk metrics at the 95%, 99%, and 99.5% confidence levels:

1. Value at Risk (VaR)

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

UNP JOURNAL OF STATISTICS AND DATA SCIENCE
Volume 4 Nomor 3 Tahun 2026 Halaman 306-319.
ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Represents the maximum potential loss over a given horizon that will not be exceeded with a certain probability. A well-known quantitative risk measurement is Value at Risk (VaR), which is the total loss due to market risk that will be experienced by an asset or portfolio within a certain time interval and with a level of certainty chosen by the investor (Choudhry & Tanna, 2007). VaR_p defined as a measure of loss where there is a small probability $1 - p$ of exceeding the pth quantile of the loss distribution (Abbasi & Guillen, 2013; Wang et al., 2005). Furthermore, VaR at level α at time t is denoted by VaR_t^{σ} that represents α-quantile of the underlying return distributin at that time and can be defined as (3) (J. M. Chen, 2018; Lazar et al., 2024; Nolde & Ziegel, 2017; Ziegel, 2016):

$$VaR_t^{\alpha} = \inf\left\{Y_t \mid F_Y\left(Y_t \mid \Omega_{t-1} \geq \alpha\right)\right\} \tag{4}$$

where $F_Y(.\,|\Omega_{t-1})$ is the cumulative distribution function of asset return Y_t over a given time horizon subject to the available information Ω_{t-1}. At the level of α, VaR can be simply expressed in the form of inverse cumulative distribution function (Duffie & Pan, 1997; Lazar et al., 2024):

$$VaR_t^{\alpha} = F_Y^{-1}\left(\alpha \mid \Omega_{t-1} \geq \alpha\right) \tag{5}$$

Value at Risk (VaR) is generally calculated for confidence intervals ranging from 95% to 99.5%. Specifically, for values $0.95 \leq q < 1$, VaR at level q corresponds to the q-th quantile of the distribution F in equation (3) and (4) expressed as (W. Chen et al., 2024; Fernandez, 2003):

$$VaR_p = u + \frac{\hat{\sigma}}{\hat{\xi}'}\left[\left(\frac{n}{n - n_u}(1 - p)\right)^{-\xi}\right] \tag{6}$$

where n is sample size of observations and n_u is the number of observations that smaller than threshold

2. Expected Shortfall (ES)

While Value at Risk (VaR) specifies a maximum threshold at a given confidence level α, it fails to capture the underlying structure or heaviness of the tail beyond that point. Because regulators and corporate risk managers require insights into both the likelihood and the financial severity of extreme losses, alternative tail risk metrics such as Expected Shortfall (ES) are widely adopted. For a random variable X following a cumulative distribution function $F(x)$, the Expected Shortfall at confidence level α, denoted as $ES_\alpha(X)$, represents the expected magnitude of loss given that the Value at Risk threshold has been exceeded (Cruz et al., 2015).

$$ES_\alpha(X) = \frac{1}{1-\alpha}\int_{\alpha}^{1} VaR_p(X)\,dp \tag{7}$$

These estimates are then compared against Parametric VaR (Normal Distribution) to demonstrate the degree of underestimation in non-EVT models.

D. Return Level Analysis

To provide a time-based risk interpretation for insurance stakeholders, we derive the Return Level (z_p). For a return period of $\frac{1}{p}$ weeks, the return level is the claim amount expected to be exceeded once every $\frac{1}{p}$ weeks on average. This is calculated using the quantile function of the fitted GEV distribution (Coles, 2001).

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

UNP JOURNAL OF STATISTICS AND DATA SCIENCE
Volume 4 Nomor 3 Tahun 2026 Halaman 306-319.
ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

III. RESULT AND DISCUSSION

The number claims are described in the **Figure 1**

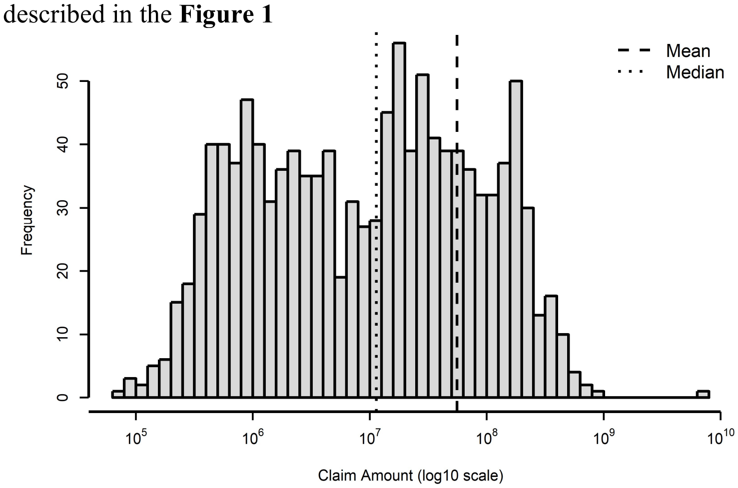


Figure 1. Frequency and Severity of Employment Injury Insurance Claim
Source: compiled by the authors

Table 1 presents descriptive statistics of insurance claims, including minimum, quartile, median, mean, and maximum values, which help identify potential extreme values in occupational accident insurance claim data.

Table 1. Descriptive statistics of amount claim (million IDR)

Minimum Value	1st Quartile	Median	Mean	3rd Quartile	Maximum Value
6.853e+04	1.340e+06	1.138e+07	5.577e+07	5.208e+07	7.000e+09

This analysis reveals that the maximum claim amount reached IDR 7,000 million, suggesting the presence of extreme claims that warrant attention for improved risk management. To ensure data accuracy, this maximum claim was checked against historical BPJS Ketenagakerjaan records. The maximum claim of IDR 7,000 million stands out as a potential outlier because it far exceeds the average claim size, beyond three standard deviations from the mean. Standard statistical tests often flag such extreme values as outliers. However, an audit of BPJS Ketenagakerjaan records confirmed that this figure represents a genuine major workplace accident with multiple casualties. In traditional statistics, extreme values are often removed to maintain normality. In Extreme Value Theory (EVT), however, this 7,000 million IDR claim is not a bad data point, but a true heavy-tail driver.

Additional analytical results indicate a skewness value of 26.13 and kurtosis (excess) of 805.3601. This signifies that the loss distribution is highly asymmetric, featuring a long tail on the right side. While most claims are relatively small, there are a few exceptionally large claims that significantly skew the mean to the right. The bell curve distribution is classified as extremely leptokurtic, characterized by a very pronounced peak and notably fat tails as illustrated in **Figure 2**.

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

 ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

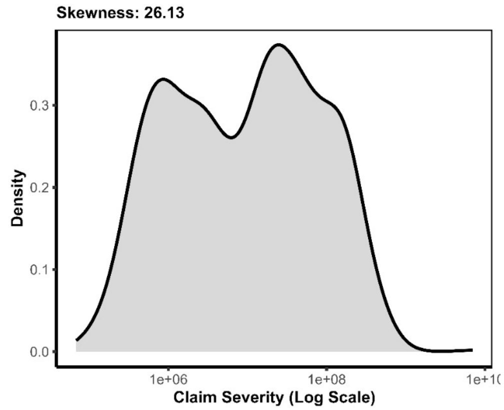


Figure 2. Probability Density Function of Severity Claim
Source: compiled by the authors

The distribution of claim severity exhibits a pronounced right-skew, as evidenced by an exceptionally high skewness value of 26.13. This significant deviation from symmetry indicates that the distribution is heavily tailed, with a long right tail extending towards very large claim amounts. The use of a logarithmic scale on the x-axis for claim severity visually confirms this property, compressing a wide range of values and revealing that while the probability density is concentrated at lower severities, there exists a non-negligible mass of probability in the extreme upper tail. Such a shape is characteristic of loss distributions in insurance and operational risk contexts, where many frequent, low-severity events coexist with rare but catastrophic losses.

This heavy-tailed nature has critical implications for risk assessment and capital modelling. Standard thin-tailed distributions (normal or exponential distribution) would severely underestimate the likelihood of extreme losses. Consequently, the analysis of these tail events necessitates the application of Extreme Value Theory (EVT). Techniques such as modelling exceedances over a high threshold with the Generalized Pareto Distribution (GPD) or fitting block maxima with the Generalized Extreme Value (GEV) distribution become essential. These methods provide a robust statistical framework for quantifying tail risk, estimating metrics like Value-at-Risk (VaR) and Expected Shortfall (ES), and ultimately informing more adequate pricing, reserving, and capital allocation decisions to address potential solvency-threatening events.

Threshold Selection

A hill plot will be used to determine the threshold for estimating the tail index α of a heavy-tailed distribution. This plot, as shown in **Figure 3**, illustrates the value of the Hill estimator against the statistical sequence (k), which is the largest number of observations used as the threshold.

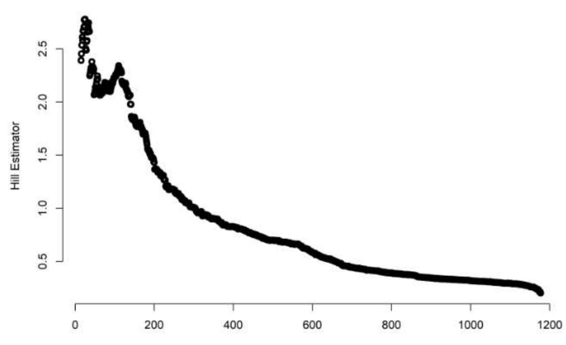


Figure 3. Threshold Selection Based on Hill Plot
Source: compiled by the authors

The Hill plot exhibits substantial fluctuations for small values of k , which is expected when only a few extreme observations are used. As k increases, the estimator becomes more stable, and a relatively flat region appears

 DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

approximately in the range $k = 200$ to $k = 400$. This region can be considered a suitable candidate for selecting the threshold in the estimation of the tail index. Beyond $k > 500$, the estimator gradually decreases, indicating the influence of non-extreme observations that bias the tail index downward. The overall magnitude of the Hill estimator (approximately between 1.2 and 1.5 in the stable region) suggests that the claim amount distribution exhibits heavy-tailed behavior with potentially infinite variance, which is consistent with the typical characteristics of insurance loss data.

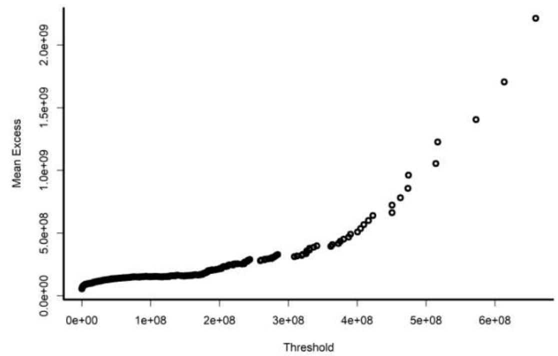


Figure 4. Threshold Selection Based on Mean Excess Plot
Source: compiled by the authors

Further, the mean excess plot displays a distinctly upward-sloping trend, confirming the heavy-tailed nature of the claim severity distribution. Following the Pickands–Balkema–de Haan theorem (Coles, 2001), a suitable threshold u for the Peaks-over-Threshold (POT) approach is chosen where the mean excess function becomes approximately linear. From the plot, a threshold in the range of $u = 0.8$ to $u = 1.2$ billion appears appropriate, as the mean excess stabilizes into a consistent linear pattern beyond this point while retaining sufficient excesses for reliable estimation. Below this range, the mean excess is too volatile due to non-tail observations and above it, data become too sparse. This threshold will be used to fit the Generalized Pareto Distribution (GPD) to the exceedances for extreme quantile estimation.

The complementary insights from the Hill plot and the Mean Excess Plot consistently indicate a heavy-tailed distribution with a substantial tail risk. The Hill estimator stabilizes around $\hat{\xi} \approx 1.45$ for a range of upper order statistics confirming a tail index greater than 1, which implies infinite variance and a high propensity for extreme losses. This is directly corroborated by the Mean Excess Plot, which exhibits a clear, upward-sloping linear trend rather than a horizontal pattern. Such a trend is the characteristic signature of a heavy-tailed distribution under the Peak-over-Threshold framework, as predicted by the Pickands–Balkema–de Haan theorem. The consistent upward slope indicates that the conditional expectation of exceedances grows with the threshold, reinforcing the need for a high threshold to isolate the tail behavior for a valid Generalized Pareto Distribution (GPD) fit. Thus, both diagnostic tools converge on the same conclusion: the claim severity distribution possesses a heavy right tail, mandating the use of Extreme Value Theory for accurate risk assessment of catastrophic losses.

Fitting Data with Extreme Value Method

Analysing the data using extreme value theory means matching the data with the GEV or GPD distribution.

Table 2. The Number and Percentage of Exceedances Data

Threshold (u) (in billion)	Number of Exceedances	Percentage of Data
800	2	0.170
900	1	0.085
1000	1	0.085
1100	1	0.085
1200	1	0.085

Source: compiled by the authors

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Table 2 presents the number of exceedances for various thresholds, revealing a critical limitation for Peaks-over-Threshold (POT) analysis. While diagnostic plots suggested a threshold range of 0.8–1.2 billion, the data show that only one observation exceeds 900 million, accounting for merely 0.085% of the dataset. This extreme sparsity in the upper tail violates the fundamental requirement of POT methodology, which typically needs at least 50–100 exceedances for reliable parameter estimation of the Generalized Pareto Distribution (GPD).

Consequently, the initial threshold recommendation based on graphical diagnostics must be revised considering this data constraint. Due to insufficient exceedances. Therefore, we pivot to alternative extreme value methodologies such as the Generalized Extreme Value distribution applied to block maxima, which requires fewer extreme observations. This pragmatic adaptation acknowledges the real-world data constraint while still addressing the research objective of assessing extreme loss potential.

Given our dataset contains 1,177 claims spanning an 8-year period, adopting an annual block size would yield only 8 data points, which is statistically insufficient to achieve convergence in the L-moment estimation framework. Similarly, monthly blocks would provide only 96 observations, which still limits the precision of the shape parameter (ξ) estimation. Therefore, a weekly block size (345 blocks across 8 years) was selected to maintain a sufficient sample size for robust parameter estimation, minimizing the variance of the estimators. To justify this selection, the data was pre-screened to ensure that weekly maxima consistently represented the upper tail of the overall claim distribution, successfully balancing the empirical trade-off between model bias and parameter variance in a finite sample context.

The Augmented Dickey-Fuller (ADF) test applied to the weekly maximum claims yielded a test statistic of -6.5597 with a p-value of 0.01. Since the p-value is well below the 5% significance level, the null hypothesis of a unit root is rejected, confirming that the extreme claim series is stationary and stable for GEV parameter estimation.

The results of GEV parameter estimation with weekly blocks on claim amount data are shown in the following **Table 3**

Table 3. GEV Parameter Estimates

Parameter	Value
Location	50.27
Scale	67.19
Shape	0.4433

Source: compiled by the authors

Given convergence challenges with Maximum Likelihood Estimation, the Generalized Extreme Value (GEV) distribution was fitted using the method of L-moments (Hosking, 1990) which provides robust parameter estimates for moderate sample sizes. Weekly block maxima were derived from the claim severity data by extracting the maximum claim amount for each ISO calendar week. The L-moment method yielded the following GEV parameter estimates: location μ = 50.27 million, scale σ = 67.19 million, and shape $\xi = 0.443$. The positive shape parameter indicates a Fréchet-type distribution with heavy tails, confirming the heavy-tailed nature observed in earlier diagnostic plots. This value implies that while the variance of weekly maxima is finite (since $\xi < 0.5$), higher moments may not exist. To validate the robustness of the weekly block selection and evaluate potential model dependency on temporal aggregation, a block size sensitivity analysis was conducted across weekly, bi-weekly, and monthly maxima that can be seen in **Table 4**

Table 4. Sensitivity Analysis of GEV Parameter Estimates and VaR (99.5%) Across Different Block Maxima Choices

Block Choice	Sample Size (N)	Location (μ)	Scale (σ)	Shape (ξ)
Weekly	345	50.27	67.19	0.4433
Bi-Weekly	227	69.46	85.55	0.4261
Monthly	96	149.71	121.15	0.4822

Source: compiled by the authors
Note: Parameters are estimated using the L-moments method in Million IDR

As presented in **Table 4**, the sensitivity analysis confirms that the heavy-tailed property ($\xi > 0$) remains highly consistent across all time scales, with the shape parameter (ξ) remaining stable within the range of 0.4433 to 0.4822. This validates that the heavy-tailed Fréchet classification is an inherent characteristic of the claim process rather than

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

a result of the chosen block size. However, selecting the weekly block size (N = 345) remains optimal because it maintains an adequate sample size for precise parameter estimation, effectively avoiding the parameter variance inflation and severe data reduction observed in monthly aggregation (N = 96).

The distribution of weekly maxima thus exhibits a propensity for extreme values, though not as severe as the infinite-variance case suggested by the initial Hill estimator applied to the full dataset.

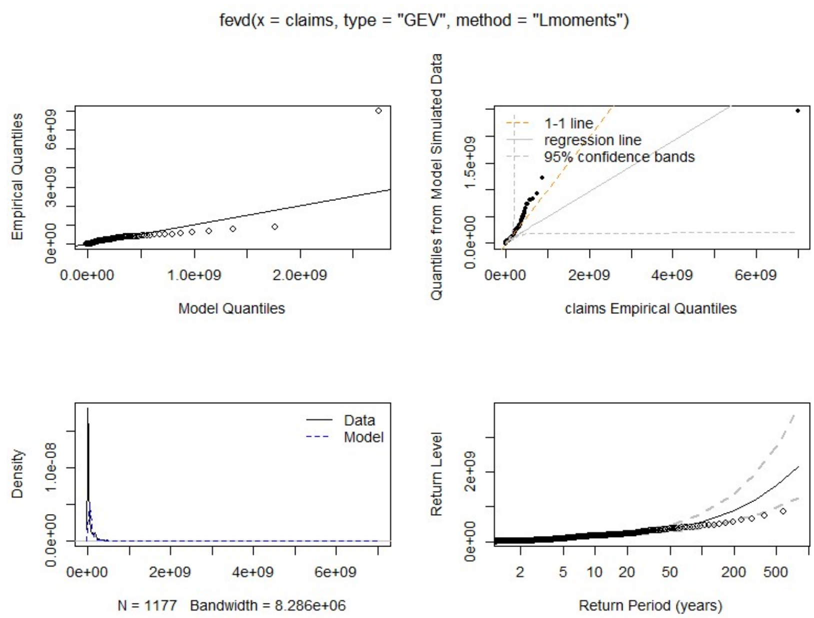


Figure 5. GEV Diagnostic Plots for Weekly Maximum Claims
Source: compiled by the authors

The probability plot and quantile-quantile plot show excellent agreement between the empirical data and the fitted model. The return level plot confirms that extreme claims are within the 95% confidence intervals, and the density plot illustrates the distribution fit.

A Parametric Bootstrap Goodness-of-Fit procedure with $B = 1,000$ replications was implemented. In each iteration, synthetic samples of size $N = 345$ were simulated from the fitted GEV model, and parameters were re-estimated using L-moments to construct a robust empirical null distribution of the Anderson-Darling test statistic. The resulting Bootstrapped Anderson-Darling test yielded a p-value of 0.948. Since this value is substantially larger than the standard significance level ($\alpha = 0.05$), fail to reject the null hypothesis. This formal test provides definitive, quantifiable proof that the Generalized Extreme Value (GEV) distribution provides an exceptionally robust fit for the weekly maximum employment injury claims, matching the excellent visual alignment previously observed in the Q-Q and Return Level plots.

Value at Risk (VaR) and Expected Shortfall (ES) Estimation

VaR estimates were computed using both EVT and a parametric approach based on normal distribution assumptions. The GEV EVT-based VaR and Expected shortfall calculated at various confidence levels while the Parametric VaR derived using the mean and standard deviation of the entire data set, assumed normal distributed. The results of the VaR calculation using the two methods were shown in **Table 5.**

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Table 5. Table of Value-at-Risk (VaR) and Expected Shortfall (ES) Estimates with 95% Bootstrap Confidence Intervals (in Million IDR)

CI	VaR GEV(95% CI)	GEV ES (95% CI)	Emp.VaR	Emp.ES	Ratio (GEV/Emp): Var/ES	N Exc
90%	309.70 (326.54 – 2015.23)	643.67 (984.30 – 347389.73)	73.36	223.31	4.22 / 2.88	236
95%	464.18 (558.72 – 5907.92)	912.95 (1552.99 – 691243.05)	164.42	331.19	2.82 / 2.76	118
99%	1063.38 (1642.76 – 65994.78)	1949.06 (4257.12 – 3382035.33)	362.79	755.53	2.93 / 2.58	24
99.5%	1484.09 (2589.11 – 182734.67)	2659.08 (6497.36 – 6655874.73)	450.56	1113.14	3.29 / 2.39	12

Source: compiled by the authors

Note: Values in parentheses represent the 95% bootstrap confidence intervals based on 1,000 iterations

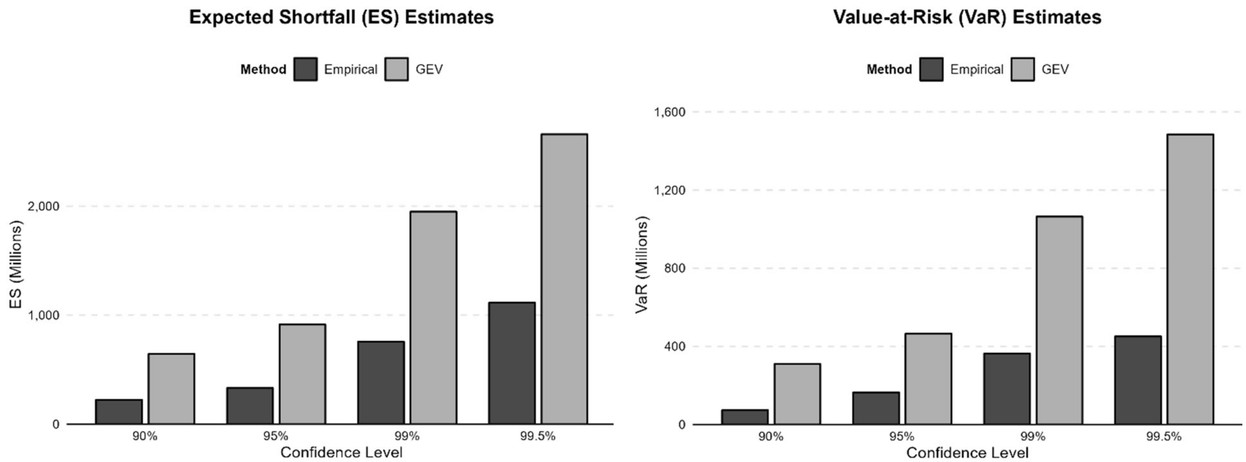
Table 5 presents the findings of the Value at Risk (VaR) and Expected Shortfall (ES) computations, elucidating a pronounced divergence between the non-parametric empirical methodology and the parametric Extreme Value Theory (EVT) framework. The results across all evaluated confidence levels show that the GEV-based VaR estimates are consistently higher than their empirical counterparts. At the 90% and 95% confidence levels, the ratios between GEV and empirical VaR are 4.22 and 2.82, respectively, while at the more extreme levels of 99% and 99.5%, the ratios remain elevated at 2.93 and 3.29. This pattern indicates that the GEV model effectively captures a heavier tail in the claim severity distribution compared to the empirical estimates derived directly from historical data.

As for Expected Shortfall (ES), the GEV model also gives more conservative and reliable risk estimates than the empirical method at all levels. At the 90% and 95% confidence levels, the GEV-ES estimates are 2.88 and 2.76 times higher than the empirical ES, respectively. This conservative gap is maintained at the upper tail extremes, with ratios of 2.58 at the 99% level and 2.39 at the 99.5% level. This systematic divergence is theoretically driven by the positive shape parameter which defines a heavy-tailed Fréchet distribution. While the empirical ES relies strictly on the limited number of historical exceedances within the tail (only 24 observations at 99% and 12 observations at 99.5%), making it highly vulnerable to sample volatility, the GEV framework extrapolates the asymptotic behavior of the tail. Consequently, the updated GEV-ES provides a more reliable safety margin, allowing the insurance portfolio to withstand extreme shocks beyond what is recorded in historical data.

Furthermore, to evaluate parameter uncertainty, 95% bootstrap confidence intervals (1,000 iterations) were calculated for both VaR and ES estimates, as presented in Table 4 also. The analysis shows that the confidence intervals widen significantly at higher tail extremes, particularly at the 99% and 99.5% confidence levels. This widening is expected and mathematically justified in heavy-tailed distributions because the upper tail relies on a small number of extreme exceedances, the estimation of extreme quantiles carries higher sampling variability. Rather than a weakness, this broad interval reflects the true uncertainty of catastrophic risks and reinforces why empirical methods which fail to capture this parameter variability are insufficient for long-term solvency planning.

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

UNP JOURNAL OF STATISTICS AND DATA SCIENCE
Volume 4 Nomor 3 Tahun 2026 Halaman 306-319.
ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X



(a) The comparison of emperical ES and GEV ES

(b) The comparison of emperical VaR and GEV Var

Figure 6. The comparison of Value at Risk (VaR) and Expected Shortfall (ES)
Source: compiled by the authors

The contrasting behaviour of VaR and ES ratios with VaR ratios consistently above 1.0 and ES ratios declining below 1.0 beyond the 95% confidence level highlights the complementary nature of these risk measures. For regulatory capital purposes, the GEV-based VaR more recommended which estimate of 1,484.09 million IDR at the 99.5% confidence level, as it provides a prudential buffer against model uncertainty. However, for internal risk management and reinsurance structuring, the empirical ES estimates, despite their wide confidence intervals, offer valuable insights into potential loss severity conditional on exceedance events. Further, the graph in **Figure 6** shows the comparison of VaR and ES between empirical and extreme value distribution

Figure 6b illustrates a significant discrepancy between the VaR estimated using the normal distribution and the Extreme Value Theory (EVT) approach. At the 99.5% confidence level, the EVT-based VaR is approximately three times higher than the Normal VaR. This confirms that the employee insurance claim data in Makassar possesses a heavy-tail characteristic, where extreme losses occur more frequently than a Gaussian model would predict. The Expected Shortfall (ES) results in **Figure 6a** further emphasize the financial risk exposure. While VaR only provides a threshold, ES estimates the average loss when that threshold is breached. The substantial gap between Normal ES and EVT ES (reaching over 1,600 million IDR at 99.5%) indicates that if a catastrophic accident occurs, the insurance company relying on standard models would face a massive capital shortfall, threatening its solvency.

Overall, these findings imply that the GEV model produces more conservative and stable VaR estimates, especially at higher confidence levels, reflecting the heavy-tailed nature of the claim severity distribution. Meanwhile, empirical ES estimates become unreliable at extreme quantiles due to the limited number of tail observations and substantial sampling uncertainty. The GEV model provides a more robust representation of tail risk, particularly when the available extreme data points are sparse. Therefore, the GEV model appears more appropriate for modelling and assessing extreme claim risks, offering a more stable framework for tail risk estimation compared to empirical methods, which are highly sensitive to sample size and variability in the extreme tail.

These results also affirm: (1) the insufficiency of empirical distributions in accurately representing insurance tail risk, (2) the necessity for a thorough reassessment of traditional capital allocation methodologies, and (3) the critical importance of adopting EVT for compliance with regulatory mandates and for the enhancement of risk management practices. The outcomes align with the research conducted by (McNeil & Frey, 2000), which highlights the superiority of EVT in the modelling of extreme insurance losses.

Return Level

Return level estimation transforms abstract statistical parameters into actionable business intelligence by quantifying the relationship between loss magnitude and event frequency. This enables insurers to make evidence-based decisions across underwriting, pricing, capital management, and risk transfer – ultimately enhancing financial

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

resilience while optimizing the trade-off between risk and return in an increasingly volatile risk landscape. Based on the injury claim amount data of BPJS Ketenagakerjaan Makassar, we estimate the return level of claim amount in several period as shown in **Table 6**.

Table 6. Return Level Estimates for Insurance Claim Severities

Return Period	Return Periods (Weeks)	Return Period (Years)	Return Level (in million IDR)
1 month	4	0.08	162.01
3 months	13	0.25	362.95
6 months	26	0.5	535.62
1 year	52	1	768.51
2 years	104	1.99	1,083.94
5 years	260	4.98	1,680.14

Source: compiled by the authors

The return level analysis, derived from the Generalized Extreme Value (GEV) distribution fitted to weekly block maxima, provides concrete estimates of extreme claim severities across various time horizons. As presented in **Table 6**, the estimated return levels increase substantially with return periods, reflecting the heavy-tailed nature of the claim distribution. The annual return level (52-week period) is estimated at 768.51 million IDR, indicating that the maximum claim expected to occur once per year exceeds three-quarters of a billion rupiah. More extreme events show significantly higher magnitudes: the 5-year return level reaches 1,680.14 million IDR, representing catastrophic losses that would be expected approximately once every five years

The progression of return levels demonstrates the practical implications of the estimated shape parameter. The ratio between 5-year and annual return levels is approximately 2.19, suggesting that claims at the 5-year horizon are more than double the annual extreme events. This non-linear growth pattern is characteristic of heavy-tailed distributions and underscores the importance of long-term risk planning. Notably, even relatively frequent events such as the monthly return level of 162.01 million IDR, represent substantial losses that require careful financial provisioning.

Moreover, the estimated Value-at-Risk (VaR) measures, presented alongside their corresponding return periods in **Table 7**, reveal the temporal frequency of extreme loss events.

Table 7. Value-at-Risk Estimates and Corresponding Return Periods

Confidence Levels	Return Periods (Weeks)	Return Period (Years)	Return Level (in million IDR)
90%	10	0.2	309.7
95%	20	0.4	464.18
99%	100	1.9	1063.38
99.5%	200	3.8	1488.09

Source: compiled by the authors

The analysis of Value-at-Risk estimates in relation to their return periods provides a comprehensive view of the portfolio's risk profile across different time horizons. The progression from frequent 10-week events (90% VaR: 309.70 million IDR) to rare 200-week catastrophes (99.5% VaR: 1484.09 million IDR) informs a tiered risk management strategy that balances capital efficiency with solvency protection. The 99.5% VaR has a return period of about 3.8 years, showing that massive claim shocks can happen in the medium term and require formal risk transfer. On the other hand, the more frequent 90% and 95% VaR events show that BPJS Ketenagakerjaan must maintain high operational resilience to handle daily claim volatility.

IV. CONCLUSION

This research demonstrates the critical importance of Extreme Value Theory (EVT) for accurately assessing and managing tail risk in insurance portfolios where catastrophic losses are rare but financially significant. Our analysis of 1,177 claims reveals a severely heavy-tailed distribution, a characteristic that conventional empirical methods and thin-tailed distributions fail to capture. The transition from the Peaks-over-Threshold (POT) to the Block Maxima method

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

due to data sparsity highlights a common practical challenge in insurance risk modelling, underscoring the robustness of the Generalized Extreme Value (GEV) distribution in handling sparse extreme data.

The comparison of risk measures presents a stark contrast: the GEV-based 99.5% Value-at-Risk (1484.09 million IDR) is more than three times higher than the empirical estimate (450.56 million IDR). This divergence has profound implications for financial solvency; relying solely on historical empirical quantiles would lead to a severe underestimation of capital requirements. Furthermore, the translation of these estimates into return periods such as the 5-year return level of 1680.14 million IDR provides a concrete benchmark for tiered risk management. Frequent, moderate losses should be managed through operational capital, while rare, severe losses necessitate risk transfer mechanisms like reinsurance.

In conclusion, while empirical approaches reflect observed history, EVT provides a forward-looking, probabilistic framework essential for navigating the unknowns of insurance risk. Despite the study's limitation in disaggregating claims by specific industrial sectors, it establishes a rigorous baseline for occupational risk assessment. Future work should enhance this analysis by incorporating covariates to account for non-stationarity, applying Bayesian methods to refine parameter estimates, and exploring industry-specific tail behaviors to further optimize insurance pricing and solvency strategies.

ACKNOWLEDGMENT

We extend our sincere appreciation to our students: Sri Ayu Anugrah, Rukayati, and Nurhikmah, for their valuable contributions in preparing the claim data used in this paper. We also express our profound gratitude to BPJS Ketenagakerjaan Makassar for granting access to the work accident claim dataset, without the data, this research would not have been possible.

REFERENCES

Abbasi, B., & Guillen, M. (2013). Bootstrap control charts in monitoring value at risk in insurance. *Expert Systems with Applications*, *40*(15), 6125–6135. https://doi.org/10.1016/j.eswa.2013.05.028

Adesina, Olumide. S., Adeleke, I., & Oladeji, T. . F. (2016). Using Extreme Value Theory to Model Insurance Risk of Nigeria’s Motor Industrial Class of Business. *The Journal of Risk Management and Insurance*, *20*(1), 40–51. https://jrmi.au.edu/index.php/jrmi/article/view/150/137

Ahsanullah, M. (2016). *Extreme Value Distributions* (Vol. 8). Atlantis Press. https://doi.org/10.2991/978-94-6239-222-9

Alexander, C. (2010). *Market risk analysis. 4: Value-at-risk models* (Repr. with corr). Wiley.

Artzner, P., Delbaen, F., Eber, J., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance*, *9*(3), 203–228. https://doi.org/10.1111/1467-9965.00068

Chen, J. M. (2018). On Exactitude in Financial Regulation: Value-at-Risk, Expected Shortfall, and Expectiles. *Risks*, *6*(2), 61. https://doi.org/10.3390/risks6020061

Chen, W., Zhao, X., Zhou, M., Chen, H., Ji, Q., & Cheng, W. (2024). Statistical Inference and Application of Asymmetrical Generalized Pareto Distribution Based on Peaks-Over-Threshold Model. *Symmetry*, *16*(3), 365. https://doi.org/10.3390/sym16030365

Choudhry, M., & Tanna, K. (2007). *An Introduction to Value-at-Risk* (4th ed). John Wiley & Sons, Ltd.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer London. https://doi.org/10.1007/978-1-4471-3675-0

Cruz, M. G., Peters, G. W., & Shevchenko, P. V. (2015). *Fundamental Aspects of Operational Risk and Insurance Analytics: A Handbook of Operational Risk* (1st edn). Wiley. https://doi.org/10.1002/9781118573013

Dey, D. K., & Yan, J. (Eds). (2016). *Extreme Value Modeling and Risk Analysis: Methods and Applications* (0 edn). Chapman and Hall/CRC. https://doi.org/10.1201/b19721

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, *74*(366), 427. https://doi.org/10.2307/2286348

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Duffie, D., & Pan, J. (1997). An Overview of Value at Risk. *The Journal of Derivatives*, *4*(3), 7–49. https://doi.org/10.3905/jod.1997.407971

Embrechts, P., Klüppelberg, C., & Mikosch, T. (1997). *Modelling Extremal Events*. Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-33483-2

Fernandez, V. (2003). *Extreme Value Theory: Value at Risk and Returns Dependence Around the World*.

Gujarati, D. N. (2021). *Basic econometrics* (Sixth edition, special Indian edition). McGraw Hill Education (India) Private Limited.

Hosking, J. R. M. (1990). L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *52*(1), 105–124. https://doi.org/10.1111/j.2517-6161.1990.tb01775.x

Kratz, M. (2019). Introduction to Extreme Value Theory: Applications to Risk Analysis and Management. In J. De Gier, C. E. Praeger, & T. Tao (Eds), *2017 MATRIX Annals* (Vol. 2, pp. 591–636). Springer International Publishing. https://doi.org/10.1007/978-3-030-04161-8_51

Laio, F. (2004). Cramer–von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters. *Water Resources Research*, *40*(9), 2004WR003204. https://doi.org/10.1029/2004WR003204

Lazar, E., Pan, J., & Wang, S. (2024). On the estimation of Value-at-Risk and Expected Shortfall at extreme levels. *Journal of Commodity Markets*, *34*, 100391. https://doi.org/10.1016/j.jcomm.2024.100391

Marimoutou, V., Raggad, B., & Trabelsi, A. (2009). Extreme Value Theory and Value at Risk: Application to oil market. *Energy Economics*, *31*(4), 519–530. https://doi.org/10.1016/j.eneco.2009.02.005

McNeil, A. J. (1997). Estimating the Tails of Loss Severity Distributions Using Extreme Value Theory. *ASTIN Bulletin*, *27*(1), 117–137. https://doi.org/https://doi.org/10.2143/AST.27.1.563210

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, *7*(3–4), 271–300. https://doi.org/10.1016/S0927-5398(00)00012-8

McNeil, A. J., & Saladin, T. (1997). The Peaks over Thresholds Method for Estimating High Quantiles of Loss Distributions. *Proceedings of 28th International ASTIN Colloquium*.

Najamuddin, F. F., Herdiani, E. T., & Jaya, A. K. (2024). Value at risk Estimation Using Extreme Value Theory Approach in Indonesia Stock Exchange. *BAREKENG: Jurnal Ilmu Matematika Dan Terapan*, *18*(2), 0695–0706. https://doi.org/10.30598/barekengvol18iss2pp0695-0706

Nolde, N., & Ziegel, J. F. (2017). Elicitability and backtesting: Perspectives for banking regulation. *The Annals of Applied Statistics*, *11*(4). https://doi.org/10.1214/17-AOAS1041

Singh, A. K., Allen, D. E., & Powell, R. J. (2011, December 12). Value at Risk Estimation using Extreme Value Theory. *Chan, F., Marinova, D. and Anderssen, R.S. (Eds) MODSIM2011, 19th International Congress on Modelling and Simulation.* 19th International Congress on Modelling and Simulation. https://doi.org/10.36334/modsim.2011.D6.singh

Song, J., & Song, S. (2012). A quantile estimation for massive data with generalized Pareto distribution. *Computational Statistics & Data Analysis*, *56*(1), 143–150. https://doi.org/10.1016/j.csda.2011.06.030

Staino, A., Russo, E., Costabile, M., & Leccadito, A. (2023). Minimum capital requirement and portfolio allocation for non-life insurance: A semiparametric model with Conditional Value-at-Risk (CVaR) constraint. *Computational Management Science*, *20*(1), 12. https://doi.org/10.1007/s10287-023-00439-1

Wainaina, H., & Waititu, A. (2014). Modeling Insurance Returns with Extreme Value Theory (A Case Study for Kenya's Fire Industrial Insurance Class of Business). *Mathematical Theory and Modeling*, *4*(2), 48–64. https://www.iiste.org/Journals/index.php/MTM/article/view/10515/10705

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480

ISSN(*Print*) : 3025-5511
ISSN(*Online*): 2985-475X

Wang, C.-P., Shyu, D., & Huang, H.-H. (2005). Optimal Insurance Design Under a Value-at-Risk Framework. *The Geneva Risk and Insurance Review*, *30*(2), 161–179. https://doi.org/10.1007/s10713-005-4677-0

Williams, R., Van Heerden, J. D., & Conradie, W. J. (2018). Value at Risk and Extreme Value Theory: Application To The Johannesburg Securities Exchange. *Studies in Economics and Econometrics*, *42*(1), 87–114. https://doi.org/10.1080/10800379.2018.12097328

Ziegel, J. F. (2016). Coherence and Elicitability. *Mathematical Finance*, *26*(4), 901–918. https://doi.org/10.1111/mafi.12080

DOI: https://doi.org/10.24036/ujsds/vol4-iss3/480