Nusantara Hasana Journal
Volume 5 No. 3 (Agustus 2025), Page: 33-40
E-ISSN : 2798-1428

APPLICATION OF PEAK OVER THRESHOLD METHOD FOR VALUE AT RISK ESTIMATION IN PROPERTY INSURANCE

Marisa

Institut of Technology Sumatera

Email: marisa@at.itera.ac.id

Abstract
Insurance companies, as providers of financial protection services, must manage risks accurately to avoid misestimating potential losses that could jeopardize financial stability. The magnitude of potential loss incurred by the insurer due to policyholder claims is commonly referred to as claim severity. A widely used risk measurement tool is Value at Risk (VaR), which estimates the maximum potential loss under the assumption of normally distributed data. However, in reality, claim amounts often exhibit extreme behavior very large values that occur with low frequency rendering conventional methods insufficient for accurate risk estimation. To address this, the present study employs Extreme Value Theory (EVT) with the Peak Over Threshold (POT) approach to model the distribution of extreme claim values. The POT method produces a Generalized Pareto Distribution (GPD), which effectively captures the heavy-tailed nature of the data. Extreme values are identified by selecting several candidate thresholds (u) using a mean excess function plot. The most appropriate threshold is then determined through the Kolmogorov–Smirnov test to ensure a good fit with the GPD. This optimal threshold is subsequently used to estimate the Value at Risk based on property insurance claim data from 2010 to 2016.
Keywords: *Property insurance; Extreme Value Theory; Peak Over Threshold; Value at Risk; Generalized Pareto Distribution*

INTRODUCTION

Insurance is defined as an agreement between the insurer and the insured, whereby the insurer, in exchange for a premium, commits to providing compensation for losses, damages, or failure to obtain expected profits resulting from uncertain events [1]. One form of insurance is general insurance, which offers protection for various assets such as buildings, machinery, inventory, and other types of property against damage, loss, or destruction. General insurance is further categorized into several branches, including property insurance, engineering insurance, and motor vehicle insurance. Property insurance includes fire insurance, home insurance packages, property all risk insurance, and earthquake insurance [2]. For instance, fire insurance covers losses resulting from fire, lightning, explosion, falling aircraft, and smoke collectively known as FLEXAS (Fire, Lightning, Explosion, impact of Aircraft, and Smoke) as stipulated in the Indonesian Standard Fire Insurance Policy (PSAKI) [3]. This type of insurance is commonly held by businesses to mitigate financial losses. In practice, claim severity, which represents the size of the loss experienced by the insured, can exhibit extreme behavior very large but infrequent. These extreme losses are best modeled using non-negative continuous probability distributions with heavy tails to accurately capture the risk of rare but high-impact events. Such events are critical in determining future premium and reserve policies [4]. One commonly used risk metric is Value at Risk (VaR), which estimates the maximum potential

Nusantara Hasana Journal
Volume 5 No. 3 (Agustus 2025), Page: 33-40
E-ISSN : 2798-1428

loss under normal market conditions. However, conventional VaR methods may underestimate extreme losses. To address this, the Peak Over Threshold (POT) method from Extreme Value Theory (EVT) is applied, which identifies extreme events as those exceeding a defined threshold and models them using the Generalized Pareto Distribution (GPD). The threshold is typically chosen using the mean excess function plot and validated with the Kolmogorov Smirnov test. This method improves the accuracy of VaR estimation, enhances risk management practices, and strengthens the insurer's financial resilience. Ultimately, it enables more reliable decision-making aligned with policyholder protection and business sustainability.

METHODS

Extreme Value Theory (EVT)

Extreme events are rare occurrences with significant and often unpredictable impacts [5]. Analyzing such events is essential for accurate risk assessment. Extreme Value Theory (EVT) is a statistical approach used to model and analyze extreme data. EVT helps quantify the risk of extreme outcomes to support better preparedness. There are two main approaches to identify extreme values, the Block Maxima (BM) method and the Peak Over Threshold (POT) method [6]. The BM approach divides data into fixed intervals and selects the maximum value from each block, resulting in the Generalized Extreme Value (GEV) distribution. However, BM may lose valuable information since it considers only one maximum value per block and is sensitive to block size selection.This study adopts the POT approach, which identifies extremes as observations exceeding a certain threshold. POT is more efficient in utilizing data, especially when the dataset is limited, and yields the Generalized Pareto Distribution (GPD) [7].

Peak Over Threshold (POT)

POT method identifies extreme values by setting a threshold that separates small losses from large ones. Losses exceeding this threshold are considered extreme values, and the excess over the threshold is modeled as a new random variable. Careful selection of the threshold value is crucial. If the threshold u is set too high, there will be too few data points to accurately estimate the model, leading to biased estimators. Conversely, a threshold that is too low will result in high variance.

Let X be a random variable representing independent and identically distributed loss amounts with distribution function F. The upper endpoint of the support of X under POT is denoted by [8]:

$$x_F = sup\{x \in R | F(x) < 1\}$$

Given a threshold $u, u \in R$ where $u < x_F$, the exceedances over the threshold (u), are defined as a new variable Y

$$Y = \begin{cases} undefined, & X \leq u \\ X - u, & u < X \leq x_F \end{cases}$$

Y represents the excess loss above the threshold u.

Mean Excess Function (MEF)

One common method for determining the threshold in loss data is by using the mean excess function plot. Let X denote the random variable representing the

Nusantara Hasana Journal
Volume 5 No. 3 (Agustus 2025), Page: 33-40
E-ISSN : 2798-1428

loss amount, and Y denote the excess loss over a threshold. The mean excess function is defined as the expected average loss that exceeds a given threshold u [9]. To estimate an appropriate threshold, the mean excess function is plotted against various values of u. The threshold is then subjectively selected from the plot, typically at the point where the distances between successive points begin to increase and the curve shows an upward concave pattern.

For empirical data, the mean excess function used to approximate the threshold is given by the following equation [10]:

$$e(u) = E(X - u|x > u) = \frac{\sum_{i=1}^{n_u}(x_i - u)}{n_u}$$

where $x_1, x_2, \dots, x_{n_u}$ represents the claim amount that exceeds the threshold value u.

In the application of EVT, to address issues arising from high threshold values in the POT method, the Pickands–Balkema–de Haan theorem can be applied. The Pickands–Balkema–de Haan theorem states that the Generalized Pareto Distribution (GPD) is the limiting distribution for values that exceed a sufficiently high threshold. In other words, as the threshold u increases, the distribution of the excess values over u [11]. Generally, a continuous random variable $Y \sim GPD\ (\gamma, \sigma)$ is said to follow a Generalized Pareto Distribution if it is characterized by a scale parameter σ and a shape parameter γ.

Value at Risk (VaR)

Risk measurement is a critical aspect of risk management. One commonly used risk metric is VaR. VaR can be employed to detect potential losses and is useful in determining a company's capital reserves. VaR is defined as the estimated maximum expected loss over a specific time period under normal market conditions, at a given confidence level α, where α ∈ (0, 1) [12]. Let X be a random variable representing the loss, then the VaR at level α is defined such that:

$$P(X \leq VaR_\alpha) = \alpha$$

This means that the probability of a loss not exceeding the VaR is α, while the probability of a loss exceeding the VaR is 1−α. Mathematically,

$$VaR_\alpha = F_X^{-1}(\alpha)$$

VaR_α is the α-quantile of the cumulative distribution function F

Value at Risk under the GPD with POT Assumption

In data modeled using EVT the risk value is calculated only for observations that exceed a specified threshold. Since the focus is on assessing the risk of large losses, a high confidence level α is chosen.

$$P(X > VaR_\alpha) = 1 - \alpha$$

To calculate the VaR, the VaR of the variable Y is required, where Y represents the excess of the claim amounts exceeding the threshold and follows a GPD

$$VaR_\alpha(Y) = F_Y^{-1}(\alpha)$$

where $F_Y(\alpha)$ is the distribution function of the GPD at confidence level α. The VaR of the random variable Y is used to determine the VaR of the random variable X, which has been modeled using EVT. The definition of the random variable Y is $Y = X - u$, therefore [13]:

$$VaR_\alpha(Y) = VaR_\alpha(X) - u$$
$$F_Y^{-1}(\alpha) = VaR_\alpha(X) - u$$
$$VaR_\alpha(X) = F_Y^{-1}(\alpha) + u$$

Nusantara Hasana Journal
Volume 5 No. 3 (Agustus 2025), Page: 33-40
E-ISSN : 2798-1428

RESULTS AND DISCUSSIONS

Descriptive Statistics

The data used in this study consists of large insurance claims from a general insurance company covering property damage or loss in Indonesia. The dataset spans the period from 2010 to 2016 and contains 11,377 observations. The analysis considers the entire dataset regardless of the year the claim occurred, and aims to estimate VaR by applying EVT. Extreme value modeling is performed using the POT approach, which assumes the GPD. To determine a suitable threshold, a MEF plot is used. Several candidate thresholds are selected based on the slope of the MEF plot and the goodness-of-fit to the GPD.

Table 1 Descriptive Statistics of Severity Claim

Statistic	Value
Sample Size	11.377
Variance($x10^{18}$)	8,651
Maximum ($x10^{11}$)	1,779
Skewness	32,854
Kurtosis	1536,4
Minimum	100.000
Median	26.640.000
Mean	342.750.000
Standard Deviation	2.941.200.000
Total Claim Amount	3.899.490.000.000

From Table 1, it can be observed that the average claim amount is Rp 342,750,000, while the median is Rp 26,640,000. The large gap between the mean and the median indicates that the data is not symmetric, suggesting a wide spread and significant variability in the dataset. The skewness value of the claim data is positive at 32.8542, indicating that the distribution is skewed to the right. Additionally, the kurtosis value is greater than 3, at 1,536.4, which suggests that the claim data follows a heavy-tailed distribution

Distribution Fit Test for Severity Claim Data

Fitting a suitable parametric distribution model to the entire dataset is essential. Since the data exhibits a heavy-tailed distribution, the Exponential, Gamma, Lognormal, Weibull, and Generalized Pareto distributions are considered to determine the best-fitting distribution. The results are as follows:

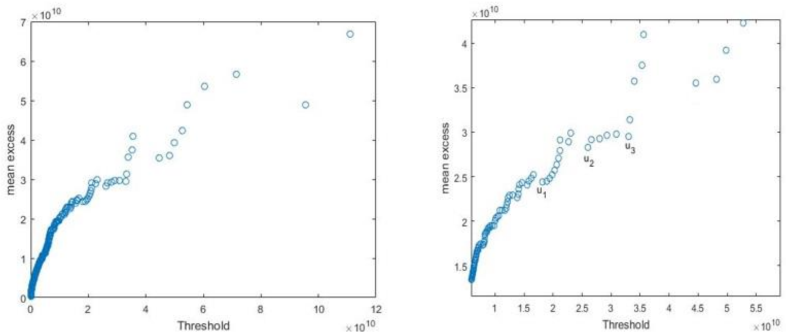
Nusantara Hasana Journal
Volume 5 No. 3 (Agustus 2025), Page: 33-40
E-ISSN : 2798-1428

Tabel 2 Parameter Estimation in Distribution Model Fitting

Distribusi	Eksponensial	Gamma	Weibul
Loglikelihood	-234.964	-224.151	-221.342
Mean (x10^8)	3,4275	3,4275	1,9221
Variansi (x10^{17})	1,0560	1,1748	2,2901
Parameter 1	$\mu = 3,4275x10^8$	α=0,2852	λ=0,464005
Parameter 2		β=1,2017x10^9	k=8,29x10^7
Distribusi	Lognormal	GPD	
Loglikelihood	-219.801	-219.753	
Mean (x10^8)	2,1512	inf	
Variansi (x10^{18})	2,3420	inf	
Parameter 1	$\mu = 17,2149$	$\gamma = 1,3482$	
Parameter 2	$\sigma = 1,9858$	$\sigma = 2,3387x10^7$	

Based on the log-likelihood values, the Generalized Pareto distribution has the highest log-likelihood, indicating that it is the best-fitting distribution for the large claim data. It is followed by the Lognormal, Weibull, Gamma, and Exponential distributions.

Mean Excess Function (MEF)



Gambar 1 Plot of Mean Excess Function (MEF)

From the MEF plot, three candidate thresholds were selected U_1, U_2, and U_3. Threshold U_1 corresponds to the 11,349th order statistic with a claim value of Rp 19,841,000,000; U_2 corresponds to the 11,359th with Rp 28,022,000,000; and U_3 corresponds to the 11,364th with Rp 34,000,000,000. Each of these thresholds produces different parameter estimates for the GPD. Therefore, to evaluate the suitability of each threshold, a goodness-of-fit test is performed using the Kolmogorov–Smirnov method.

Goodness-of-Fit Test for GPD

After identifying several potential threshold values, a goodness-of-fit test is conducted using the Kolmogorov–Smirnov test. The hypotheses are formulated as follows:

H_0 : The data follows a GPD

H_1 : The data does not follow a GPD

Nusantara Hasana Journal
Volume 5 No. 3 (Agustus 2025), Page: 33-40
E-ISSN : 2798-1428

Tabel 3 Kolmogorov-Smirnov Test Results for Candidate Thresholds

Output	**Threshold**		
	19,841	**28,022**	**34,00**
Sample Size	29	19	14
Mean ($x10^{10}$)	2,4342	2,7716	3,0612
Variance ($x10^{21}$)	1,1542	1,3840	1,5924
P-Value	0,7439	0,7598	0,5401

Based on the results of the Kolmogorov-Smirnov test on the extreme value data, using three candidate thresholds, it can be seen that the obtained P_{value} are all greater than the significance level α, 0,05. This indicates that the extreme value data under all three candidate thresholds fit well with the GPD. The p represents the probability of accepting or rejecting H_0; the larger the p, the stronger the evidence for not rejecting H_0 in the hypothesis testing. From the Kolmogorov-Smirnov test, it can be concluded that the most appropriate threshold to be used in identifying extreme values using the POT method is Rp 28,022,000,000.

To estimate the parameter values, the Maximum Likelihood Estimation (MLE) method was used with the assistance of MATLAB software, and the results obtained are as follows:

Tabel 4. GPD Parameter Estimation

Statistic	**Value**
Threshold (u)	28.022.000.000
Number of observations	11.377
Number of observations above the threshold	19
γ	0,3985
σ	1,7625

Value at Risk (VaR) Measurement

VaR can be used as a measure of risk value. The VaR calculation will be performed on the extreme value data from the years 2010 to 2016. After obtaining the estimates of the shape parameter γ and the scale parameter σ using the Maximum Likelihood Estimation (MLE) method, the VaR value can then be calculated using the following formula:

$$\text{VaR}_{\alpha}(X) = F_Y^{-1}(\alpha) + u$$

where F_Y is the cumulative distribution function (CDF) of the GPD. The VaR value is calculated at a confidence level of α, and the result is as follows:

Nusantara Hasana Journal
Volume 5 No. 3 (Agustus 2025), Page: 33-40
E-ISSN : 2798-1428

Tabel 5 Value at Risk Estimated Using the EVT Method

α	**VaR**
0,90	94.506.123.725,39
0,95	129.726.436.024,98
0,98	194.038.476.505,39
0,99	260.921.441.880,91

From the calculation of VaR using EVT, it can be observed that as the probability level α increases 0,05, the VaR also increase by nearly 1.5 times compared to the value at $VaR_{0,90}$. Similarly, for other increases in the probability level, the VaR continues to rise compared to the previous values. $VaR_{0,99}$ This means that a proportion of the loss claim data exceeds Rp 260,921,441,880.91.

CONCLUSIONS

Based on the theories and simulations presented in the previous chapters, several conclusions can be drawn from this study. Large property insurance claim data, without applying EVT, was found to follow a GPD, as indicated by goodness-of-fit tests using Exponential, Weibull, Gamma, Lognormal, and GPD distributions. This is due to the heavy-tailed nature of the data. Extreme value analysis using the POT method involved selecting several candidate thresholds based on a subjective assessment of the mean excess function plot. The Kolmogorov-Smirnov test was then used to determine the most suitable threshold, with the candidate yielding the highest p-value being selected. The chosen threshold for the 2010–2016 claim data was Rp 28,022,000,000. VaR_α was calculated to estimate potential loss, where the probability of exceeding the VaR_α value is $1 - \alpha$. The analysis showed that higher values of α led to significantly higher VaR_αvalues, indicating that fewer extreme losses are expected beyond the threshold at higher confidence levels. A comparison of VaR calculations with and without the use of EVT demonstrated that the POT approach is more effective in capturing large, infrequent claims, thereby reducing the likelihood of underestimating risk. Finally, since VaR can help identify potential risks and support the estimation of company reserves, future research is recommended to focus on reserve calculations using extreme value data.

REFERENCES

[1] Kitab Undang-Undang Hukum Dagang, Wetboek van Koophandel voor Indonesie. (2013). Jakarta: Grahamedia Press.

[2] Otoritas Jasa Keuangan Republik Indonesia. (2025). *Definisi Perusahaan Asuransi Umum*. Diakses tanggal 30 Juli 2025 dari situs resmi Otoritas Jasa Keuangan. https://www.ojk.go.id/id/kanal/iknb/pages/asuransi.aspx

[3] Sartika, D. D., Saleh, M. (2024). Perlindungan hukum terhadap nasabah dalam proses asuransi kerugian. *Jurnal Hukum*, 6 (4), 12081-12092.

[4] Buhlmann, H. (1967). Experience Rating and Credibility. *ASTIN Bulletin*, 4, 199-207.

[5] Embrechts, P., Resnick, S. I., Samorodnitsky, G. (1999). Extreme Value Theory as a Risk Management Tool. *North American Actuarial Journal*, 3 (2),

Nusantara Hasana Journal
Volume 5 No. 3 (Agustus 2025), Page: 33-40
E-ISSN : 2798-1428

30-41.

[6] Gilli, M., Kellezi, E. (2006). An Application of Extreme Value Theory for Measuring Financial Risk. *Computational Economic*, 27, 1-23.

[7] Coles, S. (2001). *An Introduction to Statistical Modelling of Extreme Values*. Bristol: Springer.

[8] Gencay, R., Faruk, S. (2004). Extreme Value Theory and Value at Risk: Relative peformance in emerging markets. *International Journal of Forecasting*, 20, 287- 303.

[9] Tse, Y. (2009). *Nonlife Actuarial Models Theory, Methods and Evaluation*. Cambridge University Press.

[10] Jong, P. D., Gillian, Z. H. (2008). Generalized Linear Models for Insurance Data. Cambridge University Press.

[11] Makarov, M. (2006). Extreme value theory and high quantile convergence. *Journal of Operational Risk*, 2, 51–57.

[12] Jockovi´c, J. 2012. Quantile Estimation for The Generalized Pareto Distribution with Application to Finance. *Yugoslav Journal of Operations Research*, 22, 297- 311.

[13] Singh, A. K, David, E. A., Roberth, J.P. (2011). *Value at Risk Estimation Using Extreme Value Theory*. Perth: Australian Mathematical Science Institute.