BAREKENG: Journal of Mathematics and Its Applications
March 2026 Volume 20 Issue 1 Page 0347–0366
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol20iss1pp0347-0366

# FLOOD REINSURANCE PREMIUM PRICING BASED ON THE STANDARD DEVIATION PRINCIPLE WITH POT-BASED THRESHOLDS FOR MORTALITY AND PROPERTY DAMAGE RISKS

**Vanessa Anggriawan**[1], **Ferry Jaya Permana**[2], **Benny Yong**[3*]

[1,2,3]*Center for Mathematics and Society, Faculty of Science, Universitas Katolik Parahyangan*
*Jln. Ciumbuleuit No 94, Jawa Barat, 40141, Indonesia*

*Corresponding author's e-mail: *benny_y@unpar.ac.id*

***Article Info***

***Article History:***

*Received: 23rd March 2025*
*Revised: 28th May 2025*
*Accepted: 22nd July 2025*
*Available online: 24th November 2025*

***Keywords:***

*Generalized pareto distribution;*
*Lognormal distribution;*
*Peaks over threshold;*
*Poisson distribution;*
*Reinsurance premium;*
*Standard deviation principle.*

***ABSTRACT***

*Disasters that occur in Indonesia lead to financial loss. One approach to mitigating the financial impact is through the utilization of natural disaster insurance. Although natural disasters occur with a relatively small frequency, the associated losses are substantial. Insurance companies need to carefully consider the characteristics of natural disaster data, as these events can lead to significant claims and potentially result in the bankruptcy of insurance companies. Insurance companies can reduce the risk of bankruptcy by transferring some risk to reinsurance companies. In this paper, the disaster reinsurance premium is determined by considering both the mortality and economic risks using the peaks over threshold (POT) model under the standard deviation principle. The Poisson, generalized Pareto, and lognormal distributions are used to determine the premium, with parameters estimated using the maximum likelihood method. A simulation analysis is conducted using synthetic data generated with RStudio software, which includes the frequency of floods per year over 20 years, as well as the number of deaths and the number of houses damaged in each flood event. The threshold is determined using the percentage method, where 10% of the data is considered extreme values. The POT model is applied to various retention cases. The simulation results show that the risk of the number of damaged houses has a greater impact on the premium amount that the insurance company must pay to the reinsurance company than the risk of the number of deaths. Additionally, cases with retention values below the threshold result in the highest reinsurance premiums, while cases with retention values above the threshold result in the lowest reinsurance premiums. This paper also shows that the reinsurance premium changes almost linearly with the increase in the extreme value percentage. This study is among the first to apply the peaks over threshold model in combination with multiple distributions for reinsurance premium estimation in the Indonesian context. The findings provide new insights into the sensitivity of reinsurance premiums to damage thresholds and retention levels, offering a practical tool for insurers in disaster-prone regions.*



***How to cite this article:***

V. Anggriawan, F. J. Permana and B. Yong., "FLOOD REINSURANCE PREMIUM PRICING BASED ON THE STANDARD DEVIATION PRINCIPLE WITH POT-BASED THRESHOLDS FOR MORTALITY AND PROPERTY DAMAGE RISKS", *BAREKENG: J. Math. & App.*, vol. 20, iss. 1, pp. 0347-0366, Mar, 2026.


Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

**Research Article · Open Access**

## 1. INTRODUCTION

Disasters occur relatively infrequently but result in significant losses. There have been many papers conducted by researchers that present mathematical modeling related to disasters [1]-[4]. This poses a risk of losses and even bankruptcy for insurance companies. Therefore, insurance companies consider setting a maximum payment limit to policyholders and transferring disaster risk to other institutions, such as reinsurance companies.

Reinsurance is the process by which an insurance company (ceding company) transfers part or all of the insured risk to a reinsurance company (reinsurer). In the context of natural disaster reinsurance, the insurance company transfers the risk of losses due to natural disasters to the reinsurance company. This means that the insurance company is obliged to pay premiums to the reinsurance company. In determining the reserves and pricing of a reinsurance contract, the reinsurance company requires data on the number and size of claims submitted by policyholders each year due to natural disasters. This data is used to estimate potential risks and determine appropriate premiums.

In research [5], the determination of disaster reinsurance premiums was investigated by considering one type of risk, namely the risk of the number of deaths, using the peaks over threshold (POT) model based on the standard deviation principle. The determination of disaster reinsurance premiums using the POT model with the standard deviation principle was proposed in research [6]. According to this study, the use of the standard deviation principle in premium determination is beneficial as it considers additional costs, such as operational and service costs, which are usually borne by the premium payer. Furthermore, the formula for determining disaster reinsurance premiums can be explicitly written when using the POT model [7]-[10], where the premium determination with the POT model uses a threshold value to determine extreme values. The POT model is applied to determine the premium amount for various retention cases, namely cases where the threshold is equal to the retention limit, the threshold is smaller than the retention limit, and the threshold is larger than the retention limit. Research [6] also mentioned that claims from policyholders usually do not come from a single type of risk but from several types of risks, such as claims for death, illness, injury, or property damage. Research [11] has demonstrated the determination of premiums by considering two types of risks, namely the risk of the number of deaths and the number of damaged houses. However, the case discussed in that study was only one of the nine possible retention case combinations.

This research addresses the gap in disaster reinsurance premium pricing models that typically consider only a single risk factor. Previous studies have explored models involving either mortality or property damage, but few have examined both simultaneously across multiple retention-threshold scenarios. This paper contributes a comprehensive pricing framework that integrates two key risk factors—number of deaths and number of damaged houses—using the Peaks Over Threshold (POT) model with the standard deviation principle. The parameters of the generalized Pareto distribution are estimated via maximum likelihood, and extreme value theory (EVT) is applied to identify losses exceeding a defined threshold. This study introduces a general premium formula that accommodates all nine possible combinations of insurer retention and reinsurer threshold for dual risks. Simulation is performed using synthetic flood disaster data over 20 years generated in RStudio. Key findings reveal that the risk of house damage contributes more significantly to the reinsurance premium than the risk of deaths, and that lower insurer retention leads to higher premium costs. Moreover, the relationship between the percentage of extreme values and the premium amount is observed to be approximately linear.

## 2. RESEARCH METHODS

The premium determination in this study uses the standard deviation principle. The general formula for determining premiums for nine different cases will be derived by determining the expectation, second moment, and probability of each risk involved. After that, a simulation of premium determination for nine different cases will be conducted with the help of RStudio software. The data used is obtained by generating the frequency of flood disasters, as well as the number of deaths and the number of damaged houses due to these flood disasters. Then, the maximum likelihood estimator of the Poisson distribution for the flood disaster frequency data, as well as the generalized Pareto and lognormal distributions for the number of deaths and the number of damaged houses in previous flood disasters, will be determined. The Poisson distribution is used because it models the frequency of independent events occurring over a fixed period, which fits the annual count of flood disasters. The generalized Pareto distribution is chosen due to its effectiveness in modeling extreme values, making it suitable for representing the number of deaths that exceed a certain threshold. The lognormal distribution is used for the number of damaged houses, as such loss data are typically positively skewed and continuous, which lognormal models can represent well. Furthermore, the Kolmogorov-Smirnov test will be conducted to assess goodness-of-fit and determine thresholds using the percentage method for the number of deaths and the number of damaged houses.

### 2.1 Poisson Distribution

Suppose the random variable $X$ follows a Poisson distribution with parameter $\lambda$, which represents the number of events in a given time interval. The probability mass function (pmf) of the Poisson-distributed random variable $X$, denoted as $X \sim Poisson(\lambda)$ [12], is given by

$$p_X(x) = \frac{\lambda^x e^{-\lambda}}{x!},$$

where $\lambda > 0$ and $x \in A$ where $A = \{x | x = 0,1,2, \dots\}$. The mean of $X$ is

$$\mathrm{E}(X) = \lambda$$

and the variance of $X$ is

$$\mathrm{Var}(X) = \lambda.$$

As shown in Hogg et al. [12], the maximum likelihood estimator of $\lambda$ is

$$\hat{\lambda} = \bar{x}.$$

### 2.2 Generalized Pareto Distribution

Suppose the continuous random variable $\boldsymbol{X}$ follows a generalized Pareto distribution with two parameters, scale $(\boldsymbol{\sigma})$ and shape $(\boldsymbol{\xi})$, denoted as $\boldsymbol{X} \sim \boldsymbol{GP}(\boldsymbol{\sigma}, \boldsymbol{\xi})$. The probability density function (pdf) of $\boldsymbol{X}$ [13] is given by

$$\boldsymbol{f_X(x)} = \begin{cases} \frac{1}{\sigma}\left(1 + \xi\frac{x}{\sigma}\right)^{-\frac{1}{\xi}-1}, & \text{if } x \geq 0,\ \sigma > 0,\ \xi \neq 0; \\ \frac{1}{\sigma}e^{-\frac{x}{\sigma}}, & \text{if } x \geq 0,\ \sigma > 0,\ \xi = 0. \end{cases},$$

The cumulative distribution function (cdf) of $\boldsymbol{X}$ is

$$\boldsymbol{F_X(x)} = \begin{cases} 1 - \left(1 + \xi\frac{x}{\sigma}\right)^{-\frac{1}{\xi}}, & \text{if } x \geq 0,\ \sigma > 0,\ \xi \neq 0; \\ 1 - e^{-\frac{x}{\sigma}}, & \text{if } x \geq 0,\ \sigma > 0,\ \xi = 0 \end{cases},$$

the mean of $\boldsymbol{X}$ is

$$\mathrm{E}(\boldsymbol{X}) = \frac{\boldsymbol{\sigma}}{1 - \boldsymbol{\xi}}, \text{ for } 0 \leq \xi < 1,$$

and the variance of $\boldsymbol{X}$ is

$$\mathrm{Var}(\boldsymbol{X}) = \frac{\boldsymbol{\sigma}^2}{(1 - 2\boldsymbol{\xi})(1 - \boldsymbol{\xi})^2}, \text{ for } 0 \leq \xi < \frac{1}{2}.$$

### 2.3 Lognormal Distribution

Suppose the continuous random variable $X$ follows a lognormal distribution with two parameters, $\mu$ and $\sigma$, denoted as $X \sim Lognormal(\mu, \sigma)$. The probability density function (pdf) of $X$ [13] is given by

$$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right),$$

where $x > 0$, $\mu \in R$, and $\sigma > 0$. The mean of $X$ is

$$\mathrm{E}(X) = \int_0^\infty \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu)^2}{2\sigma^2}\right) dx$$

and the variance of $X$ is

$$\mathrm{Var}(X) = \left(e^{\sigma^2} - 1\right)e^{2\mu+\sigma^2}.$$

As shown in Klugman et al. [13], the maximum likelihood estimators of $\mu$ and $\sigma^2$ are

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} \ln x_i,$$

$$\widehat{\sigma^2} = \frac{\sum_{i=1}^{n}(\ln x_i - \hat{\mu})^2}{n}.$$

### 2.4 Peaks over Threshold

Peaks over threshold (POT) builds a model based on a specified threshold. The POT model can only model data that exceeds the threshold. Therefore, if the losses borne by the insurance company exceed the threshold, the losses will be modeled using POT. In its application, the POT model utilizes all available data [14]. Let the random variables $\boldsymbol{X_1}, \boldsymbol{X_2}, \boldsymbol{X_3}, \dots, \boldsymbol{X_n}$ represent the data indicating the number of deaths. Let $\boldsymbol{u}$ be the threshold value, so there is a random variable $\boldsymbol{Y} = \boldsymbol{X} - \boldsymbol{u}$ representing the excess loss. The random variable $\boldsymbol{Y}$ represents the payment amount that follows a generalized Pareto distribution. The cumulative distribution function (cdf) of $\boldsymbol{Y}$ is

$$\boldsymbol{F(y) = P(X - u \le y \mid X > u) = \frac{F(y+u) - F(u)}{1 - F(u)}}, \qquad \textbf{(1)}$$

for $\boldsymbol{u} > \boldsymbol{0}$ and $\boldsymbol{y} > \boldsymbol{0}$. By letting $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{u}$, (1) can be written as

$$\boldsymbol{F(x-u) = \frac{F(x) - F(u)}{1 - F(u)}},$$

which yields

$$\boldsymbol{F(x) = F(x-u)\big(1 - F(u)\big) + F(u)}$$

$$= \begin{cases} \boldsymbol{\big(1 - F(u)\big)\left(1 - \left(1 + \xi\frac{x-u}{\sigma}\right)\right)^{-\frac{1}{\xi}} + F(u)}, & \text{if } \boldsymbol{x \ge 0,\ \sigma > 0,\ \xi \ne 0}; \\ \boldsymbol{\big(1 - F(u)\big)\left(1 - e^{-\frac{x-u}{\sigma}}\right)^{-\frac{1}{\xi}} + F(u)}, & \text{if } \boldsymbol{x \ge 0,\ \sigma > 0,\ \xi = 0}. \end{cases}$$

### 2.5 Standard Deviation Principle

The standard deviation principle equation is

$$V = \mathrm{E}(Z) + \rho\sqrt{\mathrm{Var}(Z)},$$

where $Z$ represents the total amount of claims that must be paid by the reinsurance company to the insurance company due to the occurring disaster, $\mathrm{E}(Z)$ represents the pure premium that must be paid, and $\rho\sqrt{\mathrm{Var}(Z)}$ represents the loading factor required for the claims process. The value of $\rho$ typically ranges from $0.1 \le \rho \le 0.5$ [6].

### 2.6 Kolmogorov-Smirnov Test

Suppose $X_1, X_2, \dots, X_n$ is a random sample of size n with an unknown distribution function denoted as $F_X^*(x)$. Then, let $\widehat{F_X}(x)$ be the empirical distribution function of the data, and $F_X(x)$ be a model distribution that is hypothesized to fit the data distribution. The hypotheses to be tested [13] are:

$$H_0: F_X^*(x) = F_X(x);$$

$$H_1: F_X^*(x) \ne F_X(x).$$

If $H_0$ is accepted, it means the data comes from the distribution $F_X(x)$. Conversely, if $H_0$ is rejected, it means the data does not come from the distribution $F_X(x)$. Define a test statistic $D$ as the maximum difference between $\widehat{F_X}(x)$ and $F_X(x)$ which is

$$D = \max_x \left| F_X(x) - \widehat{F_X}(x) \right|.$$

Define a significance level denoted by α, which is the probability of rejecting $H_0$ when $H_0$ is true, and the p-value is the minimum α value for which $H_0$ is rejected. The approximate p-value of $D$ can be obtained from Kolmogorov-Smirnov critical values tables, or the exact value can be obtained using RStudio software. If the p-value $> \alpha$, then $H_0$ is accepted, whereas if the p-value $\leq \alpha$, then $H_0$ is rejected.

## 3. RESULTS AND DISCUSSION

### 3.1 Premium Determination Model

Let the random variable $N(t)$ denote the number of natural disaster events at time $t$ years. The random variable $N(t)$ follows a Poisson distribution with an event rate per unit time of λ, denoted as $N(t) \sim Poisson(\lambda)$. Let the random variable $X_i$ denote the number of people who died in the $i$-th disaster and the random variable $W_i$ denote the number of houses damaged in the $i$-th disaster. The value $d$ is the retention of the insurance company for the risk of the number of deaths, the value $r$ is the retention of the insurance company for the risk of the number of houses damaged, the value $c_1$ is the claim coefficient per one death, the value $c_2$ is the claim coefficient per one house damaged, and $N(t)$ denotes the number of natural disaster events at time $t$ years. The total claim in determining the reinsurance premium considering two risks is denoted as $Z$. The random variable $Z$ is defined as

$$Z = \sum_{i=1}^{N(t)} \left(c_1(X_i - d) + c_2(W_i - r)\right)_+$$

with $\left(c_1(X_i - d) + c_2(W_i - r)\right)_+ = max\{c_1(X_i - d) + c_2(W_i - r), 0\}$. We assume that the random variables $X_i$ and $W_i$ are independent. The expectation of Z is determined as follows:

$$\begin{aligned}
\mathrm{E}(Z) = \mathrm{E}\left(\sum_{i=1}^{N(t)} \left(c_1(X_i - d) + c_2(W_i - r)\right)_+\right) &= \mathrm{E}\left(\mathrm{E}\left(\textstyle\sum_{i=1}^{N(t)} \left(c_1(X_i - d) + c_2(W_i - r)\right)_+ \middle| N(t)\right)\right) \\
&= \mathrm{E}\big(N(t)\big)\mathrm{E}\left(\left(c_1(X_i - d) + c_2(W_i - r)\right)_+\right) \\
&= \lambda\Big(\mathrm{E}\left(\left(c_1(X_i - d) + c_2(W_i - r)\right)_+ \middle| X_i > d, W_i > r\right) P(X_i > d)P(W_i > r) \\
&+ \mathrm{E}\left(\left(c_1(X_i - d) + c_2(W_i - r)\right)_+ \middle| X_i > d, W_i > r\right) P(X_i < d)P(W_i < r)\Big) \\
&= \lambda\mathrm{E}\left(\left(c_1(X_i - d) + c_2(W_i - r)\right)_+ \middle| X_i > d, W_i > r\right) P(X_i > d)P(W_i > r) \\
&= \lambda\big(c_1 E((X_i - d)_+ | X_i > d)P(X_i > d) + c_2 E((W_i - r)_+ | W_i > r)P(W_i > r)\big).
\end{aligned}$$

Furthermore, the variance of Z is determined as follows:

$$\mathrm{Var}(Z) = \mathrm{Var}\left(\sum_{i=1}^{N(t)} \left(c_1(X_i - d) + c_2(W_i - r)\right)_+\right)$$

$$\begin{aligned}
&= \mathrm{E}\left(\mathrm{Var}\left(\textstyle\sum_{i=1}^{N(t)} \left(c_1(X_i - d) + c_2(W_i - r)\right)_+ \middle| N(t)\right)\right) \\
&\qquad + \mathrm{Var}\left(\mathrm{E}\left(\textstyle\sum_{i=1}^{N(t)} \left(c_1(X_i - d) + c_2(W_i - r)\right)_+ \middle| N(t)\right)\right) \\
&= \mathrm{E}\big(N(t)\big)\mathrm{Var}\left(\left(c_1(X_i - d) + c_2(W_i - r)\right)_+\right) + \mathrm{Var}\big(N(t)\big)\mathrm{E}^2\left(\left(c_1(X_i - d) + c_2(W_i - r)\right)_+\right)
\end{aligned}$$

$$= \lambda\left(\mathrm{Var}\left(\left(c_1(X_i-d)+c_2(W_i-r)\right)_+\right)+\mathrm{E}^2\left(\left(c_1(X_i-d)+c_2(W_i-r)\right)_+\right)\right)$$

$$= \lambda\left(E\left(\left(\left(c_1(X_i-d)+c_2(W_i-r)\right)_+\right)^2\right)-\mathrm{E}^2\left(\left(c_1(X_i-d)+c_2(W_i-r)\right)_+\right)\right.$$
$$\left.+\,\mathrm{E}^2\left(\left(c_1(X_i-d)+c_2(W_i-r)\right)_+\right)\right)$$

$$= \lambda\mathrm{E}\left(\left(\left(c_1(X_i-d)+c_2(W_i-r)\right)_+\right)^2\right)$$

$$= \lambda\mathrm{E}\left(\left(\left(c_1(X_i-d)+c_2(W_i-r)\right)_+\right)^2\middle|X_i>d,W_i>r\right)P(X_i>d)P(W_i>r)$$
$$+\,\lambda\mathrm{E}\left(\left(\left(c_1(X_i-d)+c_2(W_i-r)\right)_+\right)^2\middle|X_i>d,W_i>r\right)P(X_i<d)P(W_i<r)$$

$$= \lambda\mathrm{E}\left(\left(\left(c_1(X_i-d)+c_2(W_i-r)\right)_+\right)^2\middle|X_i>d,W_i>r\right)P(X_i>d)P(W_i>r)$$

$$= \lambda\left(c_1^2\mathrm{E}(((X_i-d)_+)^2|X_i>d)P(X_i>d)+c_2^2\mathrm{E}(((W_i-r)_+)^2|W_i>r)P(W_i>r)\right).$$

Therefore, the premium determination model is

$$V = \lambda(c_1E((X_i-d)_+|X_i>d)P(X_i>d)+c_2E((W_i-r)_+|W_i>r)P(W_i>r)$$
$$+\rho\sqrt{\lambda\left(c_1^2\mathrm{E}(((X_i-d)_+)^2|X_i>d)P(X_i>d)+c_2^2\mathrm{E}(((W_i-r)_+)^2|W_i>r)P(W_i>r)\right)}. \quad (2)$$

In the reinsurance premium determination model considering two risks, there are two threshold values: $u_X$ as the threshold value for the risk of the number of deaths and $u_W$ as the threshold value for the risk of the number of houses damaged. However, the relationship between $u_X$ and $d$, as well as $u_W$ and $r$, cannot be determined in practice. There are three cases that can occur in the relationship between $u_X$ and $d$, namely:

1. Case 1a: $u_X = d$;
2. Case 2a: $u_X < d$;
3. Case 3a: $u_X > d$.

The formulas for determining the reinsurance premium for these cases are described below.

1. Case 1a: $u_X = d$
   This case occurs when the threshold of the reinsurance company and the retention of the insurance company are the same for the risk of the number of deaths. Note that

$$\mathrm{E}((X_i-d)_+|X_i>d) = \int_d^\infty (x-d)\frac{f(x)}{P(X_i>d)}dx = \int_0^\infty t\frac{f(t+d)}{1-P(X_i<d)}dt$$
$$= \int_0^\infty t\frac{f(t+u_X)}{1-P(X_i<u_X)}dt = \int_0^\infty t\, f_{u_X}(t)dt = \frac{\sigma_X}{1-\xi_X}$$

$$\mathrm{E}(((X_i-d)_+)^2|X_i>d) = \int_d^\infty (x-d)^2\frac{f(x)}{P(X_i>d)}dx = \int_0^\infty t^2\frac{f(t+d)}{1-P(X_i<d)}dt$$
$$= \int_0^\infty t^2\frac{f(t+u_X)}{1-P(X_i<u_X)}dt = \int_0^\infty t^2 f_{u_X}(t)\,dt = \frac{2\sigma_X^2}{(1-\xi_X)(1-2\xi_X)}.$$

   Let $n$ denote the sample size. Let $n_{u_X}$ denote the sample size that exceeds $d$. Thus, it can be written that $P(X_i>d)=\frac{n_{u_X}}{n}$.

2. Case 2a: $u_X < d$
   This case occurs when the threshold of the reinsurance company is lower than the retention of the insurance company for the risk of the number of deaths. Based on the theorem in [15]-[16], when the value of $d>0$ is very large, the excess distribution of the generalized Pareto distribution can be approximated by its own distribution, so $F_d(t)=G_{\xi_X,\sigma_X+\xi_X(d-u_X)}(t)$. Note that

$$\mathrm{E}((X_i-d)_+|X_i>d) = \int_d^\infty (x-d)\frac{f(x)}{P(X_i>d)}dx = \int_0^\infty t\frac{f(t+d)}{1-P(X_i<d)}dt$$

$$= \int_0^\infty t\, f_d(t)\, dt = \frac{\sigma_X + \xi_X(d - u_X)}{1 - \xi_X},$$

$$\mathrm{E}(((X_i - d)_+)^2 | X_i > d) = \int_d^\infty (x-d)^2 \frac{f(x)}{P(X_i > d)}\, dx = \int_d^\infty t^2 \frac{f(t+d)}{1 - P(X_i < d)}\, dt$$

$$= \int_d^\infty t^2 \frac{f(t+u_X)}{1 - P(X_i < u_X)}\, dt = \int_0^\infty t^2 f_{u_X}(t)\, dt = \frac{2\left(\sigma_X + \xi_X(d-u_X)\right)^2}{(1-\xi_X)(1-2\xi_X)}.$$

In this case, the probability of the sample size exceeding the threshold can be written as follows

$$P(X_i > d) = P(X_i > d | X_i > u_X) P(X_i > u_X) = \left(1 - P(X_i < d | X_i > u_X)\right) P(X_i > u_X)$$
$$= \left(1 - \frac{F(d) - F(u_X)}{1 - F(u_X)}\right)\left(1 - F(u_X)\right).$$

3. Case 3a: $u_X > d$

This case occurs when the threshold of the reinsurance company is higher than the retention of the insurance company for the risk of the number of deaths. When the threshold is higher than the retention, the POT model cannot detect the data. Therefore, the loss amount must be estimated with another distribution. According to [17]-[22] the distribution that can be used to approximate disaster loss data well is the lognormal distribution. Therefore, the lognormal distribution is used to estimate the loss amount. Note that

$$\mathrm{E}(((X_i - d)_+)^2 | X_i > d) = \int_d^\infty (x-d)^2 \frac{f(x)}{P(X_i > d)}\, dx$$

$$= \frac{1}{1 - F(d)}\left(\int_d^{u_X} (x-d)^2 f(x)\, dx + \int_{u_X}^\infty \left((x - u_X) + (x - d)\right)^2 f(x)\, dx\right).$$

Let $J_1 = \int_d^{u_X} (x-d)^2 f(x)\, dx$ and $J_2 = \int_{u_X}^\infty \left((x-u_X) + (x-d)\right)^2 f(x)\, dx$. Note that $J_1 = \int_d^{u_X} (x-d)^2 f(x)\, dx = \int_d^{u_X} x^2 f(x)\, dx - 2d \int_d^{u_X} x\, f(x)\, dx + d^2 \int_d^{u_X} f(x)\, dx$. Then

$$\int_d^{u_X} x^2 f(x)\, dx = \int_d^{u_X} x^2 \frac{1}{x\sigma_X\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu_X)^2}{2\sigma_X^2}\right) dx$$

$$= \frac{1}{\sigma_X\sqrt{2\pi}} \int_{\frac{\ln d - \mu_X}{\sigma_X}}^{\frac{\ln u_X - \mu_X}{\sigma_X}} \exp\left(-\frac{1}{2}y^2\right) \sigma_X \exp(2\mu_X + 2\sigma_X y)\, dy$$

$$= \exp(2\mu_X + 2\sigma_X^2) \int_{\frac{\ln d - \mu_X}{\sigma_X}}^{\frac{\ln u_X - \mu_X}{\sigma_X}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y - 2\sigma_X)^2\right) dy$$

$$= \exp(2\mu_X + 2\sigma_X^2)\left(\phi\left(\frac{y - 2\sigma_X}{1}\right)\Big|_{\frac{\ln d - \mu_X}{\sigma_X}}^{\frac{\ln u_X - \mu_X}{\sigma_X}}\right)$$

$$= \exp(2\mu_X + 2\sigma_X^2)\left(\phi\left(\frac{\ln u_X - \mu_X - 2\sigma^2}{\sigma_X}\right) - \phi\left(\frac{\ln d - \mu_X - 2\sigma^2}{\sigma_X}\right)\right).$$

Using a similar approach, it is obtained that

$$\int_d^{u_X} x f(x)\, dx = \int_d^{u_X} x \frac{1}{x\sigma_X\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu_X)^2}{2\sigma_X^2}\right) dx$$
$$= \exp\left(\mu_X + \frac{1}{2}\sigma_X^2\right)\left(\phi\left(\frac{\ln u_X - \mu_X - \sigma^2}{\sigma_X}\right) - \phi\left(\frac{\ln d - \mu_X - \sigma^2}{\sigma_X}\right)\right),$$

$$\int_d^{u_X} f(x)\, dx = \int_d^{u_X} \frac{1}{x\sigma_X\sqrt{2\pi}} \exp\left(-\frac{(\ln x - \mu_X)^2}{2\sigma_X^2}\right) dx = \phi\left(\frac{\ln u_X - \mu_X}{\sigma_X}\right) - \phi\left(\frac{\ln d - \mu_X}{\sigma_X}\right).$$

Therefore,

$$J_1 = \exp(2\mu_X + 2\sigma_X^2)\left(\phi\left(\frac{\ln u_X - \mu_X - 2\sigma^2}{\sigma_X}\right) - \phi\left(\frac{\ln d - \mu_X - 2\sigma^2}{\sigma_X}\right)\right) - 2d\exp\left(\mu_X + \frac{1}{2}\sigma_X^2\right)\left(\phi\left(\frac{\ln u_X - \mu_X - \sigma^2}{\sigma_X}\right) - \phi\left(\frac{\ln d - \mu_X - \sigma^2}{\sigma_X}\right)\right) + d^2\left(\phi\left(\frac{\ln u_X - \mu_X}{\sigma_X}\right) - \phi\left(\frac{\ln d - \mu_X}{\sigma_X}\right)\right).$$

Also note that

$$\begin{aligned} J_2 &= \int_{u_X}^{\infty}\big((x-u_X)+(u_X-d)\big)^2 f(x)\,dx \\ &= \int_{u_X}^{\infty}(x-u_X)^2 f(x)\,dx + 2(u_X-d)\int_{u_X}^{\infty}(x-u_X)f(x)\,dx + (u_X-d)^2\int_{u_X}^{\infty}f(x)\,dx \\ &= \big(1-F(u_X)\big)\int_0^{\infty} z^2\frac{f(t+u_X)}{1-F(u_X)}\,dt + 2(u_X-d)\big(1-F(u_X)\big)\int_0^{\infty} t\frac{f(t+u_X)}{1-F(u_X)}\,dt + (u_X-d)^2\big(1-F(u_X)\big) \\ &= \big(1-F(u_X)\big)\left(\frac{2\sigma_X^2}{(1-\xi_X)(1-2\xi_X)} + 2(u_X-d)\frac{\sigma_X}{1-\xi_X} + (u_X-d)^2\right). \end{aligned}$$

Therefore,

$$\begin{aligned} &\mathrm{E}(((X_i-d)_+)^2|X_i>d) \\ &= \frac{1}{1-P(X_i<d)}\Bigg(\exp(2\mu_X+2\sigma_X^2)\left(\phi\left(\frac{\ln u_X-\mu_X-2\sigma^2}{\sigma_X}\right) - \phi\left(\frac{\ln d-\mu_X-2\sigma^2}{\sigma_X}\right)\right) \\ &- 2\exp\left(\mu_X+\frac{1}{2}\sigma_X^2\right)\left(\phi\left(\frac{\ln u_X-\mu_X-\sigma^2}{\sigma_X}\right) - \phi\left(\frac{\ln d-\mu_X-\sigma^2}{\sigma_X}\right)\right) \\ &+ d^2\left(\phi\left(\frac{\ln u_X-\mu_X}{\sigma_X}\right) - \phi\left(\frac{\ln d-\mu_X}{\sigma_X}\right)\right) \\ &+ \big(1-F(u_X)\big)\left(\frac{2\sigma_X^2}{(1-\xi_X)(1-2\xi_X)} + 2(u_X-d)\frac{\sigma_X}{1-\xi_X} + (u_X-d)^2\right)\Bigg). \end{aligned}$$

Using a similar method, it is obtained that

$$\begin{aligned} \mathrm{E}((X_i-d)_+|X_i>d) &= \frac{J_1+J_2}{1-P(X_i<d)} \\ &= \frac{1}{1-P(X_i<d)}\left(\int_d^{u_X}(x-d)f(x)\,dx + \int_{u_X}^{\infty}\big((x-u_X)+(u_X-d)\big)f(x)\,dx\right) \end{aligned}$$

$$= \frac{1}{1 - P(X_i < d)} \left( \exp\left(\mu_X + \frac{1}{2}\sigma_X^2\right)\left(\phi\left(\frac{\ln u_X - \mu_X - \sigma_X^2}{\sigma_X}\right) - \phi\left(\frac{\ln d - \mu_X - \sigma_X^2}{\sigma_X}\right)\right) - d\left(\phi\left(\frac{\ln u_X - \mu_X}{\sigma_X}\right) - \phi\left(\frac{\ln d - \mu_X}{\sigma_X}\right)\right) + \left(1 - F(u_X)\right)\left(\frac{\sigma_X}{1 - \xi_X} + (u_X - d)\right)\right).$$

The probability for the sample size exceeding the retention $d$ can be written as $P(X_i > d) = 1 - F(d)$, where $F(d)$ follows a lognormal distribution with parameters μ and σ. The probability for the sample size exceeding the threshold $u_X$ can be written as $P(X_i > u_X) = 1 - F(u_X) = \frac{n_{u_X}}{n}$.

Using a similar approach to the previous three cases, the expectation calculations for these cases are obtained.

1. Case 1b: $u_W = r$

   This case occurs when the threshold of the reinsurance company and the retention of the insurance company are the same for the risk of the number of houses damaged. Note that

$$\mathrm{E}((W_i - r)_+ | W_i > r) = \frac{\sigma_W}{1 - \xi_W},$$

$$\mathrm{E}((\mathrm{W_i} - \mathrm{r})_+)^2 | \mathrm{W_i} > \mathrm{r}) = \frac{2\sigma_{\mathrm{W}}^2}{(1 - \xi_{\mathrm{W}})(1 - 2\xi_{\mathrm{W}})},$$

$$P(W_i > r) = \frac{n_{u_W}}{n}.$$

2. Case 2b: $u_W < r$

   This case occurs when the threshold of the reinsurance company is lower than the retention of the insurance company for the risk of the number of houses damaged. Note that

$$\mathrm{E}((W_i - r)_+ | W_i > r) = \mathrm{E}(W) = \frac{\sigma_W + \xi_W(r - u_W)}{1 - \xi_W},$$

$$\mathrm{E}((W_i - r)_+)^2 | W_i > r) = \frac{2\left(\sigma_W + \xi_W(r - u_W)\right)^2}{(1 - \xi_W)(1 - 2\xi_W)},$$

$$P(W_i > r) = \left(1 - \frac{F(r) - F(u_W)}{1 - F(u_W)}\right)\left(1 - F(u_W)\right).$$

3. Case 3b: $u_W > r$

   This case occurs when the threshold of the reinsurance company is higher than the retention of the insurance company for the risk of the number of houses damaged. Note that

$$\begin{aligned}
&\mathrm{E}(((W_i - r)_+)^2 | W_i > r) \\
&= \frac{1}{1 - P(W_i < r)} \Bigg( \exp(2\mu_W + 2\sigma_W^2) \left( \phi\left( \frac{\ln u_W - \mu_W - 2\sigma_W^2}{\sigma_W} \right) - \phi\left( \frac{\ln r - \mu_W - 2\sigma_W^2}{\sigma_W} \right) \right) \\
&- 2r \exp\left( \mu_W + \frac{1}{2}\sigma_W^2 \right) \left( \phi\left( \frac{\ln u_W - \mu_W - \sigma_W^2}{\sigma_W} \right) - \phi\left( \frac{\ln r - \mu_W - \sigma_W^2}{\sigma_W} \right) \right) \\
&+ r^2 \left( \phi\left( \frac{\ln u_W - \mu_W}{\sigma_W} \right) - \phi\left( \frac{\ln r - \mu_W}{\sigma_W} \right) \right) \\
&+ \left(1 - F(u_W)\right) \left( \frac{2\sigma_W^2}{(1 - \xi_W)(1 - 2\xi_W)} + 2(u_W - r)\frac{\sigma_W}{1 - \xi_W} + (u_W - r)^2 \right) \Bigg),
\end{aligned}$$

$$\begin{aligned}
&\mathrm{E}((W_i - r)_+ | W_i > r) \\
&= \frac{1}{1 - P(W_i < r)} \Bigg( \exp\left( \mu_W + \frac{1}{2}\sigma_W^2 \right) \left( \phi\left( \frac{\ln u_W - \mu_W - \sigma_W^2}{\sigma_W} \right) - \phi\left( \frac{\ln r - \mu_W - \sigma_W^2}{\sigma_W} \right) \right) - r\left( \phi\left( \frac{\ln u_W - \mu_W}{\sigma_W} \right) - \phi\left( \frac{\ln r - \mu_W}{\sigma_W} \right) \right) \\
&+ \left(1 - F(u_W)\right) \left( \frac{\sigma_W}{1 - \xi_W} + (u_W - r) \right) \Bigg).
\end{aligned}$$

Additionally, the probability for the sample size exceeding the retention $r$ can be written as $P(W_i > r) = 1 - F(r)$, where $F(r)$ follows a lognormal distribution with parameters $\mu_W$ and $\sigma_W$. The probability for the sample size exceeding the threshold $u_W$ can be written as $P(W_i > u_W) = 1 - F(u_W) = \frac{n_{u_W}}{n}$.

Combination of these two retentions results in nine types of cases, namely:

1. Case 1: $u_X = d$ and $u_W = r$;
2. Case 2: $u_X = d$ and $u_W < r$;
3. Case 3: $u_X = d$ and $u_W > r$;
4. Case 4: $u_X < d$ and $u_W = r$;
5. Case 5: $u_X < d$ and $u_W < r$;
6. Case 6: $u_X < d$ and $u_W > r$;
7. Case 7: $u_X > d$ and $u_W = r$;
8. Case 8: $u_X > d$ and $u_W < r$;
9. Case 9: $u_X > d$ and $u_W > r$.

By substituting each expectation, second moment, and probability obtained according to the type of case into Eq. (2), a premium determination model for nine different cases can be obtained.

### 3.2 Generating Data

The frequency data of flood disasters, the number of deaths, and the number of houses damaged due to floods in Indonesia cannot be obtained completely. Therefore, synthetic data is used for the simulation. This synthetic data is generated using RStudio software. The flood disaster frequency data over 20 years ($\boldsymbol{n = 20}$) is assumed to follow a Poisson distribution with parameter $\boldsymbol{\lambda}$, the number of deaths is assumed to follow a generalized Pareto distribution with parameters $\boldsymbol{\sigma_X}$ and $\boldsymbol{\xi_X}$, and the number of houses damaged is assumed to follow a generalized Pareto distribution with parameters $\boldsymbol{\sigma_W}$ and $\boldsymbol{\xi_W}$.

The flood disaster frequency data over 20 years is generated with $\boldsymbol{\lambda = 250}$. This data is assumed to represent the number of flood disasters occurring throughout Indonesia. The chosen value of $\boldsymbol{\lambda = 250}$ is obtained by calculating the average from the flood disaster frequency data obtained from the National Disaster Management Agency (BNPB) [21]. The generated data is shown in Fig. 1.

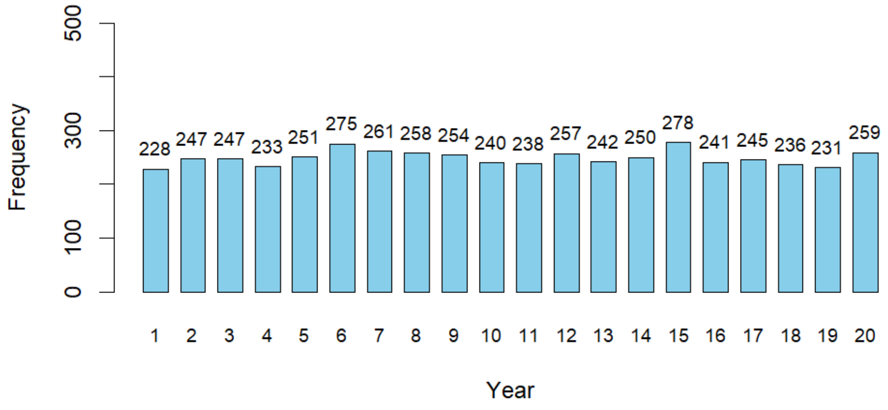
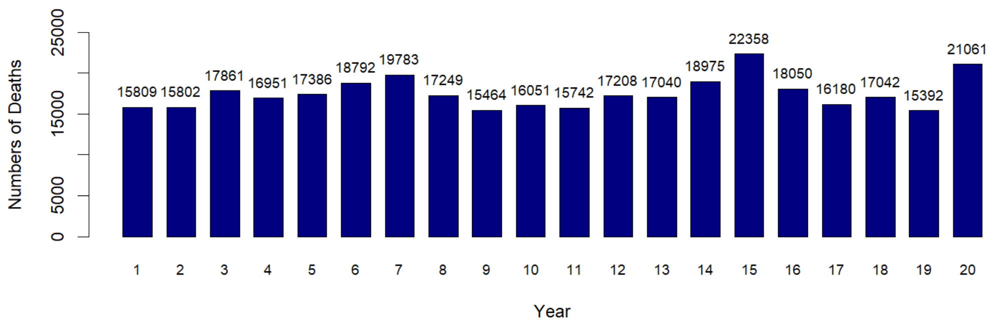
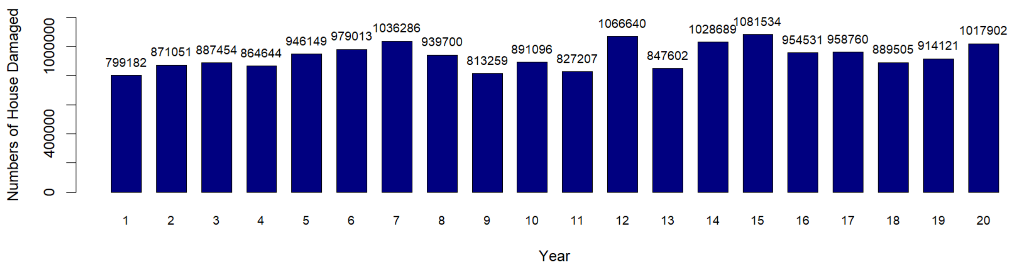



**Figure 1**. Bar Chart of Flood Disaster Frequency Over 20 Years

The data on the number of deaths is generated, assumed to follow a generalized Pareto distribution with parameters $\boldsymbol{\sigma_X}$ and $\boldsymbol{\xi_X}$, and the data on the number of houses damaged is generated, assumed to follow a generalized Pareto distribution with parameters $\boldsymbol{\sigma_W}$ and $\boldsymbol{\xi_W}$. The data on the number of deaths is generated with $\boldsymbol{\sigma_X = 53.704}$ and $\boldsymbol{\xi_X = 0.208}$, while the data on the number of houses damaged is generated with $\boldsymbol{\sigma_W = 3375.367}$ and $\boldsymbol{\xi_W = 0.076}$. These parameter values are chosen following the parameter values obtained in [11]. After the data is generated, the data is randomly divided according to the flood disaster frequency per year. Fig. 2 (a) shows the number of deaths per year due to floods, and Fig. 2 (b) shows the number of houses damaged per year due to floods.




(a)




**(b)**

**Figure 2**. Bar Chart Showing the Annual Impact of Floods Over a 20-year Period
(a) Number of Deaths, (b) Number of Houses Damaged

### 3.3 Simulation Data Processing

The frequency of disaster events is modeled using a Poisson distribution with an event rate per unit time of $\lambda$. Therefore, the Poisson parameter is estimated from the flood disaster frequency data that has been generated. The parameter $\lambda$ is estimated using the maximum likelihood method with the help of RStudio software. The parameter estimation result is $\hat{\lambda} = 249$.

Goodness-of-fit test is conducted between the empirical flood disaster frequency data and the model, which is the Poisson distribution, using the Kolmogorov-Smirnov test. The α value used is 5%. With RStudio, a p-value of 0.9807 is obtained. It indicates that there is a fit between the empirical data and the model. Also, Fig. 3 shows that there is a fit between the quantiles of the flood disaster frequency data and the Poisson distribution, indicating that the flood disaster frequency data follows a Poisson distribution.

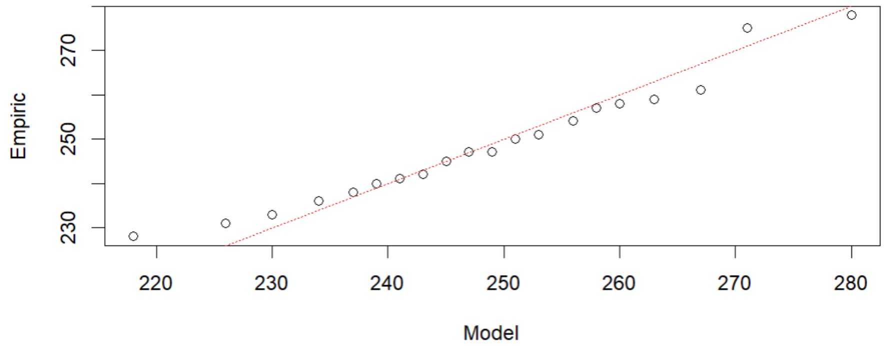



**Figure 3**. QQ-plot of Empirical Data Against the Poisson Distribution Model

For premium determination in cases where the reinsurance company's threshold is equal to the insurance company's retention and cases where the reinsurance company's threshold is lower than the insurance company's retention, the generalized Pareto distribution approach is used. For premium determination in cases where the reinsurance company's threshold is higher than the insurance company's retention, the lognormal distribution approach is used. This applies to both types of risks, namely the risk of the number of deaths and the number of houses damaged. Using RStudio software, the parameter estimation results are obtained using the maximum likelihood method, as shown in Tables 1 and 2.

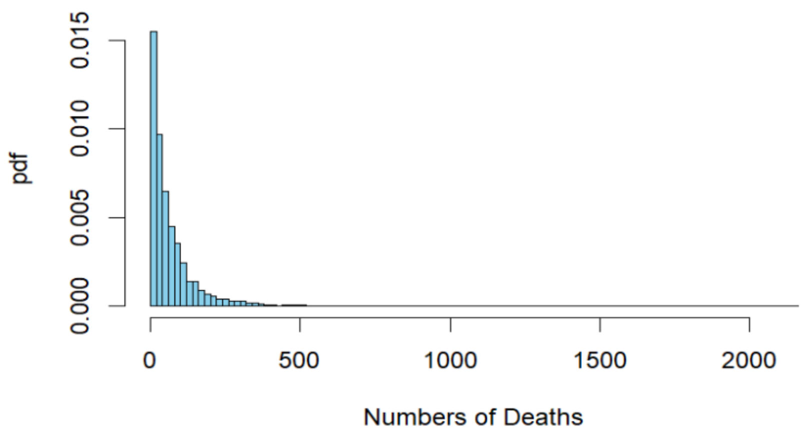
**Table 1.** Parameter Estimation Results for Generalized Pareto and Lognormal Distributions for the Number of Deaths

| Distribution | Generalized Pareto | Lognormal |
|---|---|---|
| Parameter | $\widehat{\sigma_X} = 54.07$<br>$\widehat{\xi_X} = 0.24$ | $\widehat{\mu_X} = 3.57$<br>$\widehat{\sigma_X} = 1.29$ |

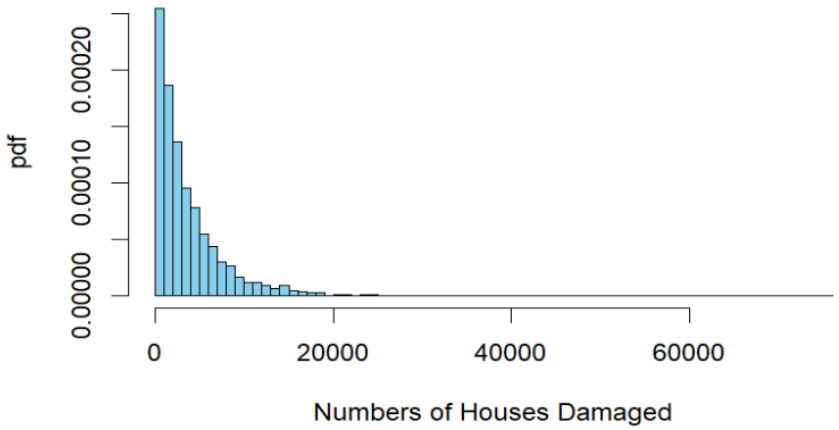
**Table 2**. Parameter Estimation Results for Generalized Pareto and Lognormal Distributions for the Number of Houses Damaged

| Distribution | Generalized Pareto | Lognormal |
|---|---|---|
| Parameter | $\widehat{\sigma_W} = 3{,}334.05$<br>$\widehat{\xi_W} = 0.11$ | $\widehat{\mu_W} = 7.59$<br>$\widehat{\sigma_W} = 1.33$ |

The data on the number of deaths and the number of houses damaged due to floods fits the pdf of the generalized Pareto distribution. This fit can be shown by the histograms in Fig. 4.




(a)

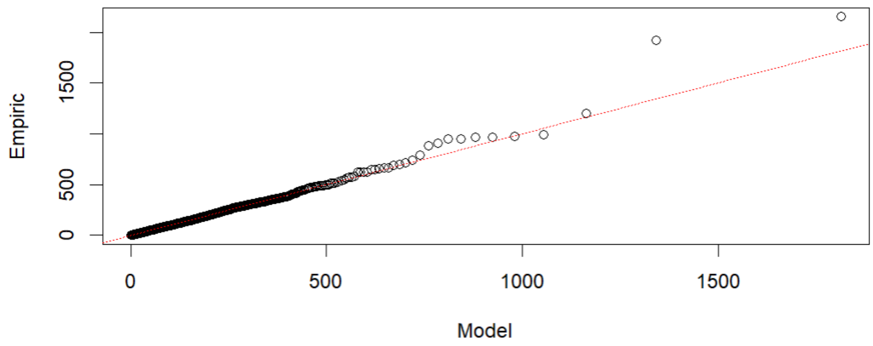



(b)

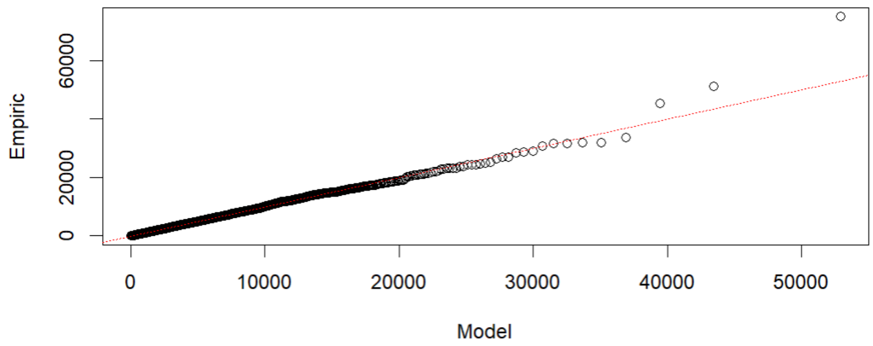
**Figure 4**. Histogram Illustrating the Distribution of Flood Impacts
(a) Number of Deaths, (b) Number of Houses Damaged

The data on the number of deaths and the number of houses damaged are heavy-tailed and skewed to the right. This is supported by the kurtosis and skewness of each data set. The data on the number of deaths has a kurtosis of 79 and a skewness of 6, while the data on the number of houses damaged has a kurtosis of 29 and a skewness of 3. Data with heavy tails is characteristic of data that follows a generalized Pareto distribution.

Goodness-of-fit test is conducted between the data on the number of deaths following a generalized Pareto distribution and the model from the parameter estimation results using the Kolmogorov-Smirnov test. Using RStudio software, a p-value of 0.2185 is obtained at $\alpha = 0.05$. This indicates that there is a fit between the empirical data and the model. Applying a similar method to the data on the number of houses damaged, a p-value of 0.9858 is obtained at $\alpha = 0.05$. This indicates that there is a fit between the empirical data and the model. The QQ-plot images are shown in Fig. 5. It can be concluded that the data on the number of deaths and the data on the number of houses damaged follow a generalized Pareto distribution.




(a)




(b)

**Figure 5**. QQ-plots for Flood Impacts
(a) QQ-plot for the Number of Deaths, (b) QQ-plot for the Number of Houses Damaged

### 3.4 Threshold Selection

The kurtosis of the data on the number of deaths is 79 and the data on the number of houses damaged is 29. This indicates that the data has extreme values. Then, the percentage method [14] is used. According to [22], the appropriate value of $\boldsymbol{m}$ for determining the threshold is 10%. Therefore, it can be written that $\boldsymbol{n_{u_X}} = \boldsymbol{n_{u_W}} = \mathbf{10\%} \times \mathbf{4,971} = \mathbf{497.1} \approx \mathbf{497}$, where $\boldsymbol{n_{u_X}}$ denotes the sample size exceeding $\boldsymbol{d}$ and $\boldsymbol{n_{u_W}}$ denotes the sample size exceeding $\boldsymbol{r}$. This means that there are 497 observations classified as extreme values out of the $\mathbf{4,971}$ observations. Then, the threshold of the data can be determined as follows $\boldsymbol{u_X} = \boldsymbol{u_W} = \mathbf{497} + \mathbf{1} = \mathbf{498}$. The data in the 498-th position, when sorted from the largest to the smallest, is the threshold of the data for both the number of deaths and the number of houses damaged. For the data on the number of deaths, the threshold is $\boldsymbol{u_X} = \mathbf{162}$, while for the data on the number of houses damaged, the threshold is $\boldsymbol{u_W} = \mathbf{8,684}$.

### 3.5 Determination of Premium Amount for Each Case

The assumptions used in determining the premium amount are as follows:

1. $\rho = 0.3$;
2. $c_1$ = Rp10,000,000 per death;
3. $c_2$ = Rp5,000,000 per house damaged;
4. When $u_X > d$, $d$ is assumed to be 62. When $u_X < d$, $d$ is assumed to be 262;
5. When $u_W > r$, $r$ is assumed to be 7,684. When $u_W < r$, $r$ is assumed to be 9,684.

Table 3 shows the pure premium, additional premium, and reinsurance premium for each case described.

**Table 3**. Comparison of Pure Premium, Additional Premium, and Reinsurance Premium for the Each Case

| Case | Pure Premium | Additional Premium | Reinsurance Premium |
|---|---|---|---|
| Case 1: $u_X = d$ and $u_W = r$ | Rp484,105,800,000 | Rp42,397,148,354 | Rp526,502,948,354 |
| Case 2: $u_X = d$ and $u_W < r$ | Rp403,137,972,000 | Rp39,178,484,221 | Rp442,316,456,221 |
| Case 3: $u_X = d$ and $u_W > r$ | Rp580,598,280,000 | Rp39,694,566,530 | Rp620,292,846,530 |
| Case 4: $u_X < d$ and $u_W = r$ | Rp476,622,852,000 | Rp42,390,653,102 | Rp519,013,505,102 |
| Case 5: $u_X < d$ and $u_W < r$ | Rp395,655,024,000 | Rp39,171,455,267 | Rp434,826,479,267 |
| Case 6: $u_X < d$ and $u_W > r$ | Rp573,115,332,000 | Rp39,687,628,978 | Rp612,802,960,978 |
| Case 7: $u_X > d$ and $u_W = r$ | Rp494,720,670,000 | Rp42,406,309,838 | Rp537,126,979,838 |
| Case 8: $u_X > d$ and $u_W < r$ | Rp413,752,842,000 | Rp39,188,398,174 | Rp452,941,240,174 |
| Case 9: $u_X > d$ and $u_W > r$ | Rp591,213,150,000 | Rp39,704,351,620 | Rp630,917,501,620 |

From Table 3, the pure premium amounts for the nine different cases are obtained. Cases 1, 2, and 3 share the commonality of having the reinsurance company's threshold equal to the insurance company's retention $d$ for the risk of the number of deaths, which is 162. The highest pure premium among these three cases occurs in Case 3, where the insurance company has a retention $r$ for the risk of the number of houses damaged that is smaller than the reinsurance company's threshold ($u_W > r$). This is followed by Case 1, where the retention $r$ for the risk of the number of houses damaged and the reinsurance company's threshold are equal ($u_W = r$). The third is Case 2, where the insurance company has a retention $r$ for the risk of the number of houses damaged that is larger than the reinsurance company's threshold ($u_W < r$).

Similarly, for Cases 4, 5, and 6, involving retention $u_X < d$, the highest pure premium occurs in Case 6 with retention $u_W > r$, followed by Case 4 with retention $u_W = r$, and Case 5 with retention $u_W < r$. For Cases 7, 8, and 9, involving retention $u_X > d$, the highest pure premium occurs in Case 9 with retention $u_W > r$, followed by Case 7 with retention $u_W = r$, and Case 8 with retention $u_W < r$. It can be seen that the pure premium involving retention $u_W > r$, where the insurance company's retention $r$ for the risk of the number of houses damaged is smaller than the reinsurance company's threshold, has a higher premium compared to cases with retention $u_W = r$, where the insurance company's retention $r$ for the risk of the number of houses damaged and the reinsurance company's threshold are equal, and $u_W < r$, where the insurance company's retention $r$ for the risk of the number of houses damaged is larger than the reinsurance company's threshold.

Additionally, Cases 1, 4, and 7 share the commonality of having the reinsurance company's threshold equal to the insurance company's retention $r$ for the risk of the number of houses damaged, which is 8,684.

The highest pure premium among these three cases occurs in Case 7, where the insurance company has a retention $d$ for the risk of the number of deaths that is smaller than the reinsurance company's threshold ($u_X > d$). This is followed by Case 1, where the retention $d$ for the risk of the number of deaths and the reinsurance company's threshold are equal ($u_X = d$). The third is Case 4, where the insurance company has a retention $d$ for the risk of the number of deaths that is larger than the reinsurance company's threshold ($u_X < d$).

The same pattern occurs for Cases 2, 5, and 8. For cases involving retention $u_W < r$, the highest pure premium occurs in Case 8 with retention $u_X > d$, followed by Case 2 with retention $u_X = d$, and Case 5 with retention $u_X < d$. For cases involving retention $u_W > r$, namely Cases 3, 6, and 9, the highest pure premium occurs in Case 9 with retention $u_X > d$, followed by Case 3 with retention $u_X = d$, and Case 6 with retention $u_X < d$. It can be seen that the pure premium involving retention $u_X > d$, where the insurance company's retention $d$ for the risk of the number of deaths is smaller than the reinsurance company's threshold, has a higher premium compared to cases with retention $u_X = d$, where the insurance company's retention $d$ for the risk of the number of deaths and the reinsurance company's threshold are equal, and $u_X < d$, where the insurance company's retention $d$ for the risk of the number of deaths is larger than the reinsurance company's threshold.

When the premiums are ranked, the pure premium and reinsurance premium follow the same order. However, the additional premium does not follow the same order. This can occur because the determination of the additional premium is based on the variance value, which affects the randomness of the data. Additionally, it is known that the sum of the pure premium and the additional premium results in the reinsurance premium. From Table 3, it can be seen that the additional premium is relatively small compared to the pure premium, so the additional premium does not significantly impact the order of the insurance premiums, even though it has a different order.

Therefore, it can be concluded that for each risk, cases involving retention values less than the threshold ($u_W > r$ and $u_X > d$) result in the highest reinsurance premiums, while cases involving retention values greater than the threshold ($u_W < r$ and $u_X < d$) result in the lowest reinsurance premiums. This occurs because when the insurance company has a retention less than the threshold, it means the insurance company's ability to bear the loss is small, so the reinsurance company has to bear a larger risk of loss. The larger the risk of loss that the reinsurance company has to bear, the higher the premium that the insured, i.e., the insurance company, has to pay. Conversely, when the insurance company has a retention greater than the threshold, it means the insurance company's ability to bear the loss is large, so the reinsurance company has to bear a smaller risk of loss. The smaller the risk of loss that the reinsurance company has to bear, the lower the premium that the insured has to pay.

Based on Table 3, it can also be seen that the highest reinsurance premiums occur in Cases 9, 3, and 6, where retention $u_W > r$ is always involved. For the fourth to sixth highest premiums, they occur in Cases 7, 1, and 4, with retention $u_W = r$ always involved. For the seventh to ninth highest premiums, they occur in Cases 8, 2, and 5, with retention $u_W < r$ always involved. Therefore, it is concluded that the risk of the number of houses damaged has a greater impact compared to the risk of the number of deaths on the reinsurance premium amount.

### 3.6 The Impact of Changes in the Percentage of Extreme Values on Reinsurance Premiums

Table 4 shows the reinsurance premiums for Case 1, where $u_X = d$ and $u_W = r$, when the percentage of extreme values varies from 5% to 15%.

**Table 4**. Reinsurance Premiums for Case 1 at Different Extreme Value Percentages

| Extreme Value Percentage | $n_{u_X} = n_{u_W}$ | $u_X = u_W$ | $u_X = d$ | $u_W = r$ | Reinsurance Premium | Premium Increase | Percentage Increase |
|---|---|---|---|---|---|---|---|
| 5% | 249 | 250 | 241 | 12.048 | Rp272,032,211,104 | | |
| 6% | 298 | 299 | 217 | 11.334 | Rp323,304,169,900 | Rp51,271,958,796 | 18.8% |
| 7% | 348 | 349 | 200 | 10.445 | Rp374,346,059,266 | Rp51,041,889,366 | 15.8% |
| 8% | 398 | 399 | 184 | 9.617 | Rp425,205,802,308 | Rp50,859,743,042 | 13.6% |
| 9% | 447 | 448 | 173 | 9.090 | Rp475,916,686,528 | Rp50,710,884,220 | 11.9% |
| 10% | 497 | 498 | 162 | 8.684 | Rp526,502,948,354 | Rp50,586,261,826 | 10.6% |
| 11% | 547 | 548 | 153 | 8.323 | Rp576,982,884,330 | Rp50,479,935,976 | 9.6% |
| 12% | 597 | 598 | 145 | 7.987 | Rp627,370,709,054 | Rp50,387,824,724 | 8.7% |

| Extreme Value Percentage | $n_{u_X} = n_{u_W}$ | $u_X = u_W$ | $u_X = d$ | $u_W = r$ | Reinsurance Premium | Premium Increase | Percentage Increase |
|---|---|---|---|---|---|---|---|
| 13% | 646 | 647 | 140 | 7.624 | Rp677,677,726,647 | Rp50,307,017,593 | 8.0% |
| 14% | 696 | 697 | 133 | 7.263 | Rp727,913,102,447 | Rp50,235,375,800 | 7.4% |
| 15% | 746 | 747 | 126 | 6.966 | Rp778,084,390,008 | Rp50,171,287,561 | 6.9% |

From Table 4, it can be seen that the higher the percentage of extreme values selected, the higher the reinsurance premium for Case 1. Conversely, the lower the percentage of extreme values selected, the lower the reinsurance premium for Case 1. This is logical because a lower percentage of extreme values indicates a lower probability of extreme events, resulting in lower risk borne by the reinsurance company. Consequently, the premium paid is not too high. Conversely, a higher percentage of extreme values indicates a higher probability of extreme events, resulting in higher risk borne by the reinsurance company. Therefore, the premium paid by the insurance company is higher.

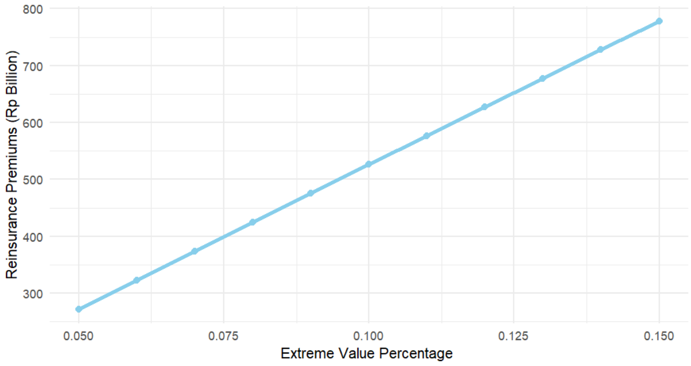



**Figure 6**. Movement of Reinsurance Premiums for Case 1 with the Movement of Extreme Value Percentages

Fig. 6 shows that the reinsurance premium for Case 1 appears to move linearly with the movement of the extreme value percentage. However, the movement of the reinsurance premium for Case 1 with the movement of the extreme value percentage is only nearly linear. Table 4 shows that the increase in reinsurance premiums is not the same but is in the same range, which is around Rp50,000,000,000. Table 4 also shows that the higher the percentage of extreme values, the smaller the percentage increase in reinsurance premiums.

### 3.7 The Impact of Changes in Risk Retention of the Number of Deaths on Reinsurance Premiums

The threshold used in this analysis is $u_X = 162$ and $u_W = 8.684$, obtained by assuming 10% of the observations are classified as extreme values, as described in subsection 3.4. The cases discussed are cases involving retention $u_W = r = 8.684$, namely cases 1, 4, and 7. The retention value $d$ used in this analysis is 162 for Case 1, while the retention value used in Case 4 is $162(1+10\%) = 178$, $162(1+10\%)^2 = 196$, and $162(1+10\%)^3 = 216$. In Case 7, the retention value used is $162(1-10\%) = 146$, $162(1-10\%)^2 = 131$, and $162(1-10\%)^3 = 118$. The reinsurance premiums with different retention $d$ values are shown in Table 5.

**Table 5**. Reinsurance Premiums with Various Retention $d$ Values

| $i$ | Case Type | $d$ | Reinsurance Premium | Premium Increase | Percentage Increase |
|---|---|---|---|---|---|
| 1 | Case 4 | 216 | Rp 521,963,122,092 | | |
| 2 | Case 4 | 196 | Rp 523,057,926,573 | Rp 1,094,804,481 | 0.21% |
| 3 | Case 4 | 178 | Rp 525,486,401,164 | Rp 2,428,474,591 | 0.46% |
| 4 | Case 1 | 162 | Rp 526,502,948,354 | Rp 1,016,547,190 | 0.19% |
| 5 | Case 7 | 146 | Rp 531,233,404,737 | Rp 4,730,456,383 | 0.90% |
| 6 | Case 7 | 131 | Rp 535,758,216,854 | Rp 4,524,812,117 | 0.85% |
| 7 | Case 7 | 118 | Rp 538,268,535,043 | Rp 2,510,318,189 | 0.47% |

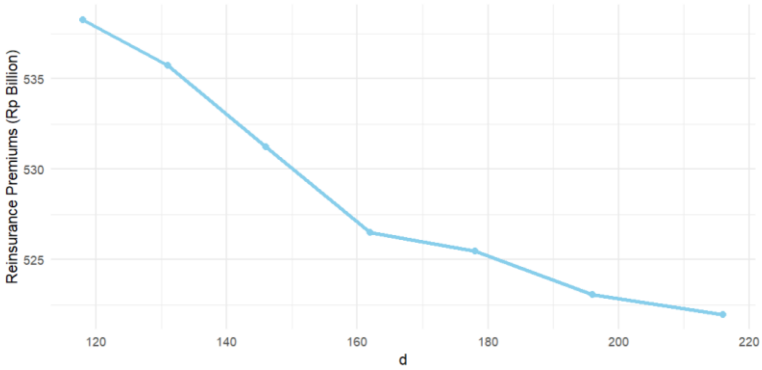



**Figure 7**. Movement of Retention Value $d$ Against Reinsurance Premiums

From Table 5 and Figure 7, it can be concluded that the higher the retention $d$, the lower the reinsurance premium that must be paid. Conversely, the lower the retention $d$, the higher the reinsurance premium that must be paid. Table 5 shows no pattern of percentage increase in premiums due to the increase in retention value.

### 3.8 The Impact of Changes in Risk Retention of the Number of Houses Damaged on Reinsurance Premiums

The threshold used in this analysis is $u_X = 162$ and $u_W = 8.684$, obtained by assuming 10% of the observations are classified as extreme values, as described in subsection **3.4**. The cases discussed are those involving retention $u_X = d = 162$, namely cases 1, 2, and 3. The retention value $r$ used in this analysis is 8.684 for case 1, while the retention value used in case 2 is $8.684(1 + 10\%) = 9.554$, $8.684(1 + 10\%)^2 = 10.511$, and $8.684(1 + 10\%)^3 = 11.586$. In case 3, the retention value used is $8.684(1 - 10\%) = 7.819$, $8.684(1 - 10\%)^2 = 7.035$, and $8.684(1 - 10\%)^3 = 6.327$. The reinsurance premiums with different retention $r$ values are shown in Table 6.

**Table 6**. Reinsurance Premiums with Various Retention $r$ Values

| i | Case Type | $r$ | Reinsurance Premium | Premium Increase | Percentage Increase |
|---|---|---|---|---|---|
| 1 | Case 2 | 11.586 | Rp 306,107,516,607 | | |
| 2 | Case 2 | 10.511 | Rp 401,487,131,994 | Rp 95,379,615,387 | 31.16% |
| 3 | Case 2 | 9.554 | Rp 440,553,563,959 | Rp 39,066,431,965 | 9.73% |
| 4 | Case 1 | 8.684 | Rp 526,502,948,354 | Rp 85,949,384,395 | 19.51% |
| 5 | Case 3 | 7.819 | Rp 596,104,992,754 | Rp 69,602,044,400 | 13.22% |
| 6 | Case 3 | 7.035 | Rp 719,015,002,129 | Rp 122,910,009,375 | 20.62% |
| 7 | Case 3 | 6.327 | Rp 795,213,182,872 | Rp 76,198,180,743 | 10.60% |

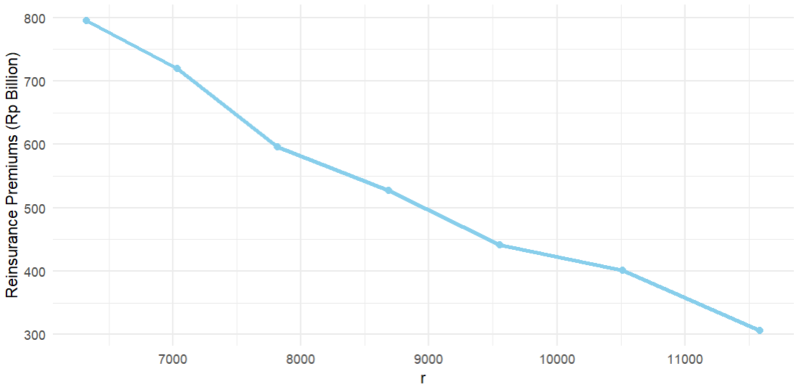



**Figure 8**. Movement of Retention Value $r$ Against Reinsurance Premiums

From Table 6 and Fig. 8, it can be concluded that the higher the retention $r$, the lower the reinsurance premium that must be paid. Conversely, the lower the retention $r$, the higher the reinsurance premium that must be paid. Table 6 shows no pattern of percentage increase in premiums due to the increase in retention value.

## 4. CONCLUSION

Based on the author's analysis in the previous sections, the following conclusions were obtained:

1. The risk of the number of damaged houses has a greater impact than the risk of death on the amount of premium that the insurance company must pay to the reinsurance company. Additionally, for each risk, cases involving retention values less than the threshold result in the highest reinsurance premiums, while cases involving retention values greater than the threshold result in the lowest reinsurance premiums.
2. In cases where the threshold is equal to the retention for each risk, a low percentage of extreme values results in lower reinsurance premiums, while a high percentage of extreme values results in higher reinsurance premiums. Changes in the percentage of extreme values result in an almost linear change, where each 1% increase in the percentage of extreme values results in an increasein reinsurance premiums of approximately Rp50,000,000,000.
3. In cases involving the threshold equal to the retention for the risk of the number of damaged houses, the higher the retention of the risk of death, the lower the reinsurance premium that must be paid. Conversely, the lower the retention of the risk of death, the higher the reinsurance premium that must be paid.
4. In cases involving the threshold equal to the retention for the risk of death, the higher the retention of the risk of the number of damaged houses, the lower the reinsurance premium that must be paid. Conversely, the lower the retention of the risk of the number of damaged houses, the higher the reinsurance premium that must be paid.

## Author Contributions

Vanessa Anggriawan: Data Curation, Formal Analysis, Investigation, Methodology, Resources, Software, Visualization, Draft Preparation, Writing-Original Draft. Ferry Jaya Permana: Conceptualization, Formal Analysis, Funding Acquisition, Project Administration, Supervision, Validation, Writing-Original Draft. Benny Yong: Conceptualization, Formal Analysis, Funding Acquisition, Project Administration, Supervision, Validation, Draft Preparation, Writing-Original Draft, Writing-Review and Editing. All authors discussed the results and contributed to the final manuscript.

## Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

## Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions, which have greatly improved the quality of this manuscript.

## Declarations

The authors declare no conflict of interest.

## REFERENCES

[1] I. Hasbiyati, M. Abdullah, R. Salambue, and A. Ahriyati, "NORTHWEST CORNER METHOD FOR NATURAL DISASTER NOTIFICATION," *Barekeng: Jurnal Ilmu Matematika dan Terapan*, vol. 16, no. 1, pp. 271-280, Mar. 2022. doi: https://doi.org/10.30598/barekengvol16iss1pp269-278

[2] Z. Qonita, S. Karima, A. Rusdiansyah, and R. Riyandari, "NUMERICAL MODELING OF THE 1998 PAPUA NEW GUINEA TSUNAMI USING THE COMCOT," *Barekeng: Jurnal Ilmu Matematika dan Terapan*, vol. 18, no. 1, pp. 349-360, Mar. 2024. doi: https://doi.org/10.30598/barekengvol18iss1pp0349-0360

[3] Y. Hikmah, "LOGISTIC MODELING TO PREDICT THE INTEREST OF THE INDONESIAN PEOPLE FOR BUYING FLOOD IMPACTED INSURANCE PRODUCTS," *Barekeng: Jurnal Ilmu Matematika dan Terapan*, vol. 17, no. 1, pp. 323-330, Apr. 2023. doi: https://doi.org/10.30598/barekengvol17iss1pp0323-0330

[4] D. L. A. Primananda and M. Muhajir, "MODELLING EARTHQUAKE DISASTER DAMAGE DUE DEPTH OF EPICENTER AND MAGNITUDE USING SPATIAL REGRESSION," *Barekeng: Jurnal Ilmu Matematika dan Terapan*, vol. 17, no. 3, pp. 1221-1234, Sep. 2023. doi: https://doi.org/10.30598/barekengvol17iss3pp1221-1234
[5] A.S. Jalandri, S. Nurrohmah, and I. Fithriani, "PRICING CATASTROPHE REINSURANCE RISK PREMIUM USING PEAKS OVER THRESHOLD (POT) MODEL," *Physics*, vol. 1, no. 1, pp. 1742-6596, 2021. doi: https://doi.org/10.1088/1742-6596/1725/1/012086
[6] W. Chao, "PRICING CATASTROPHE REINSURANCE UNDER THE STANDARD DEVIATION PREMIUM PRINCIPLE," *AIMS Math*, vol. 7, pp. 4472-4484, 2022. doi: https://doi.org/10.3934/math.2022249
[7] P. Strickler, "R¨UCKVERSICHERUNG DES KUMULRISIKOS IN DER LEBENSVERSICHERUNG," *XVI International Congress of Actuaries in Brussels*, 1, pp. 666–679, 1960.
[8] H. Q. Xiao and S. W. Meng, "EVT AND ITS APPLICATION TO PRICING OF CATASTROPHE REINSURANCE (IN CHINESE)," *J. Appl. Stat. Manage.*, vol. 32, pp. 240–246, 2013. doi: http://dx.doi.org/10.13860/j.cnki.sltj.2013.02.003
[9] M. Leppisaari, "MODELING CATASTROPHIC DEATHS USING EVT WITH A MICROSIMULATION APPROACH TO REINSURANCE PRICING," *Scandinavian Actuarial Journal*, vol. 2016, pp. 113–145, 2014. doi: https://doi.org/10.1080/03461238.2014.910833
[10] E. Ekheden, and O. Hössjer, "PRICING CATASTROPHE RISK IN LIFE (RE)INSURANCE," *Scandinavian Actuarial Journal*, vol. 2014, pp. 352–367, 2014. doi: https://doi.org/10.1080/03461238.2012.695747
[11] H. A. Surya, H. Napitupulu, and Sukono, "DOUBLE RISK CATASTROPHE REINSURANCE PREMIUM BASED ON HOUSES DAMAGED AND DEATHS," *Mathematics*, vol. 11, pp. 810, 2023. doi: https://doi.org/10.3390/math11040810
[12] R.V. Hogg, J.W. McKean, and A.T. Craig, *INTRODUCTION TO MATHEMATICAL STATISTICS*, 8th edition, Boston:Pearson, 2019.
[13] S.A. Klugman, H. H. Panjer, and G.E. Willmot, *LOSS MODELS: FROM DATA TO DECISIONS*, 5th edition, Hoboken:Wiley, 2019.
[14] J. Beirlant, Y. Goegebeur, J. Teugels, and J. Segers, *STATISTICS OF EXTREMES: THEORY AND APPLICATIONS*, John Wiley & Sons, Ltd, 2005. doi: https://doi.org/10.1002/0470012382
[15] A. A. Balkema, and L. De Haan, "RESIDUAL LIFE TIME AT GREAT AGE. ANNALS OF PROBABILITY," *The Annals of Probability*, vol. 2, pp. 792–804. doi: https://doi.org/10.1214/aop/1176996548
[16] J. Pickands III, "STATISTICAL INFERENCE USING EXTREME ORDER STATISTICS," *The Annals of Probability*, vol. 3, pp. 119–131, 1975. doi: https://doi.org/10.1214/aos/1176343003
[17] J. Liu, and Y. Li, "EMPIRICAL STUDY ON EARTHQUAKE LOSSES DISTRIBUTION AND CAT BOND PRICING IN CHINA," *Finance Trade Res.*, vol. 20, pp. 82–88.
[18] V. Pisarenko and M Rodkin, *HEAVY-TAILED DISTRIBUTIONS IN DISASTER ANALYSIS*. Springer, 2010. doi: https://doi.org/10.1007/978-90-481-9171-0
[19] K. Burnecki, A. Misiorek, and R. Weron, *STATISTICAL TOOLS FOR FINANCE AND INSURANCE*, Springer Verlag, 2005. doi: https://doi.org/10.1007/3-540-27395-6_13
[20] P. Maneerat, S-A. Niwitpong, and S. Niwitpong, "SIMULTANEOUS CONFIDENCE INTERVALS FOR ALL PAIRWISE COMPARISONS OF THE MEANS OF DELTA-LOGNORMAL DISTRIBUTIONS WITH APPLICATION TO RAINFALL DATA," *PLOS ONE*, vol. 16, no. 7, pp. 1-25, 2021. doi: https://doi.org/10.1371/journal.pone.0253935
[21] Badan Nasional Penanggulangan Bencana (BNPB), *DIBI*. [Online]. available: https://dibi.bnpb.go.id/ [Accessed: 30 January 2025].
[22] P. Embrechts, C. Klüppelberg, and T. Mikosch, *MODELLING EXTREMAL EVENTS*. Cambridge University Press, 2014. doi: https://doi.org/10.2143/AST.28.2.519071