Available online at www.ijournalse.org

Emerging Science Journal

(ISSN: 2610-9182)

Vol. 9, No. 1, February, 2025

Estimating Ruin Probability in an Insurance Risk Model Using the Wang-PH Transform Through Claim Simulation

Weenakorn Ieosanurak [1], Adisak Moumeesri [2*]

[1] Department of Mathematics, Faculty of Science, Khon Kaen University, Khon Kaen 40002, Thailand.

[2] Department of Statistics, Faculty of Science, Silpakorn University, Nakhon Pathom 73000, Thailand.

Abstract

The accurate estimation of ruin probability is a fundamental challenge in non-life insurance, impacting financial stability, risk management strategies, and operational decisions. This study aims to propose an approach for estimating ruin probability using claim simulation enhanced by the Wang-PH transform to fit various loss distributions, including Gamma, Weibull, Lognormal, Log-logistic, Inverse Weibull, and Inverse Gaussian, to actual claim data. Methods involve the transformation of loss distributions via the Wang-PH transform and rigorous evaluation to select the optimal distribution model that best reflects actual claim characteristics. This model serves as the foundation for estimating finite-time ruin probability through claim simulation, employing the acceptance-rejection technique to generate random samples. Additionally, a regression-based methodology estimates the minimum capital reserve required to safeguard against financial risk. Findings indicate the proposed method's computational efficiency, making it a valuable tool for insurers and risk analysts in assessing and mitigating financial risks in the non-life insurance sector. The novelty of this study lies in the integration of the Wang-PH transform with empirical data fitting and simulation techniques, applied to estimating ruin probability and determining capital reserves.

Keywords:

Claim Simulation;
Loss Distribution;
Minimum Capital Reserve;
Non-Life Insurance;
Ruin Probability; Wang Transform.

Article History:

Received:	04	June	2024
Revised:	25	November	2024
Accepted:	11	December	2024
Published:	01	February	2025

1- Introduction

The challenge of maintaining an adequate capital reserve is currently a prominent issue faced by non-life insurance companies in Thailand. The complexities associated with capital calculations have become a central point of concern. These companies need to allocate a sufficient initial capital amount to ensure that the potential financial loss from risks does not exceed an acceptable threshold. In the contemporary landscape of non-life insurance, ensuring a sufficient and well-managed capital fund is paramount. It serves as a financial buffer against unexpected claim events and market fluctuations, safeguarding the insurer's ability to meet its obligations to policyholders and stakeholders alike. The determination of an appropriate capital fund involves a comprehensive assessment of various risk factors, market conditions, regulatory requirements, and the company's own risk appetite.

Actuaries, being experts in assessing and managing risk, play a pivotal role in addressing the capital reserve problem. Their analytical expertise allows them to navigate the complex landscape of risk evaluation, enabling them to quantify the potential financial impact of various scenarios on an insurer's capital reserves. By accurately estimating the capital required to support the insurer's operations and fulfill obligations, actuaries contribute to enhancing the company's financial stability and resilience. At its core, loss distributions represent the patterns and magnitudes of financial losses incurred by an insurance company due to claim events. These distributions capture the variability and uncertainty inherent in the insurance business, serving as a foundation for quantifying potential financial risks. The choice of an

* **CONTACT**: moumeesri_a@silpakorn.edu

DOI: http://dx.doi.org/10.28991/ESJ-2025-09-01-011

appropriate loss distribution model is fundamental as it directly affects the accuracy of risk assessments. Klugman et al. [1] discussed loss distribution and its modeling, emphasizing its importance as a method for estimating risk measures. McNeil et al. [2] and Dhaene et al. [3] described risk measures and their classification. One of the risk measures we consider for insurance pricing in this study is the premium calculation principle.

Numerous authors have discussed the premium calculation principle for both financial and insurance risks, with Wang being particularly notable for establishing the premium calculation principle. Noteworthy works include those by Wang [4-7]. Wang [4] proposed proportional hazards (PH) transform principle for calculating premiums. The calculation is in the form: $H(X) = \int_0^\infty [S_X(t)]^c \, dt$ for some $0 < c < 1$, where $S_X(t) = \Pr(X > t) = 1 - F_X(t)$. Wang [8] presented a pricing method based on the transformation as follows: $F^*(x) = \Phi[\Phi^{-1}(F(x)) + \theta]$. Here, Φ represents the standard normal cumulative distribution function (cdf), $F(x)$ is the cdf of the distribution of interest, and θ is a constant. This transformation, known as the Wang transform, is widely recognized among financial engineers and risk managers. Given the challenge of fitting loss distributions to actual claims data, these transformations provide essential tools for pricing insurance premiums based on transformed functions.

Recent studies have advanced premium calculation methods by addressing asymmetric risks and sector-specific variables. Calderín-Ojeda et al. [9] developed a premium calculation framework based on Conditional Tail Expectation (CTE) and asymmetric loss functions, enhancing traditional risk measures. Psarrakos [10] introduced a novel approach using the mode of unimodal weighted distributions as a premium principle, offering an alternative to traditional methods. These developments suggest that integrating such methodologies could significantly improve risk assessments and capital reserve adequacy in the non-life insurance market.

Ruin probability, a critical metric in insurance risk assessment, quantifies the likelihood that an insurance company's capital reserve will be exhausted due to large and unexpected claim events. Accurate estimation of ruin probability hinges on a comprehensive understanding of loss distributions and the adequacy of the capital reserve. By integrating these components, insurers can gauge the vulnerability of their financial position and make informed decisions to manage risk effectively. The development of risk assessment in the field of insurance has been a significant endeavor. This evolution began in the early 1900s when a Swedish actuary named Filip Lundberg was credited with establishing the foundation for risk theory. Among Lundberg's contributions is the creation of a basic model for non-life insurance. The central focus of any insurance firm is the timing and size of claims, both of which impact the company's capital. Insurers are obligated to safeguard their stakeholders from unexpected losses and risks. The insurer's stability is gauged by its capital reserves, the severity of claims, and its premium income. Lundberg's model makes three assumptions.

- Time instances of claim occurrences are denoted as T_i, where $i = 0, 1, 2, ...$ The sequence of these times, $T_0 \leq T_1 \leq T_2 \leq \cdots$, is referred to as the claim occurrence process.
- Claim sizes at time T_i are represented as Y_i. These claim sizes are independent and identically distributed, forming the claim size process.
- The claim occurrence and claim size processes are independent of each other.

In addition to the three aforementioned assumptions, the number of claims within a time interval $[0, t]$ was defined by the equation: $N(t) = \max\{i \geq 1: T_i \leq t\}$, with $N(t)$ termed as the claim count process.

Furthermore, insurance companies examine the total claim size, denoted as $S(t) = \sum_{i=1}^{N(t)} Y_i$, $t \geq 0$, for a specified time interval $[0, t]$.

In the 1930s, the renowned Swedish statistician Harald Cramer [11] significantly advanced collective risk theory by integrating the total claim size process, $S(t)$, with claim arrival times, T_i, generated by a Poisson process. When the claim number process is defined as a homogeneous Poisson process, the resulting model that merges claim sizes and claim arrival times is referred to as the Cramer-Lundberg model. In the Cramer-Lundberg model, let $p(t)$ represent the premium income in the time interval $[0, t]$. It is assumed that $p(\cdot)$ is a deterministic linear function, meaning $p(t) = ct$, for $t \geq 0$, where $c > 0$ is a constant known as the premium rate. Given the total claim amount $S(t)$ for $t \geq 0$, $U(t) = u + p(t) - S(t) = u + ct - \sum_{i=1}^{N(t)} Y_i$. The $\{U(t), t \geq 0\}$ process is referred to as the risk process, or surplus process, of the model, where $u > 0$ represents the capital reserve.

In the classical risk process, the claim severity Y_n occurs at time T_n such that $0 \leq T_1 \leq T_2 \leq \cdots$. In this process, the probability of insolvency (ruin) only arises at claim time T_n, where $n \in \mathbb{N}$. The discrete-time risk process is defined as follows: $U_n(u) = u + cT_n - \sum_{k=1}^{n} Y_k$, where $U_0(u) = u \geq 0$ is the initial capital reserve, $c > 0$ is the premium rate for one unit of time, computed by the expected value principle as: $c = (1 + \theta)E[Y]$, where Y is a random variable of claim severity and $\theta \geq 0$ represents the safety loading of the insurer. Chan & Zhang [12] extended this concept by considering the discrete-time risk process under the assumption $T_n = n$, where the risk process is defined as: $U_n(u) = u + cn - \sum_{k=1}^{n} Y_k$, where $n \in \mathbb{N}$. Here, $U_0(u) = u \geq 0$ represents the initial capital, $c > 0$ is the premium rate for one unit of time, and the process $\{Y_n: n \in \mathbb{N}\}$ consists of a sequence of independent and identically distributed claim random variables at the claim arrival time $T_n = n$.

Furthermore, the probability of ruin over a unit of time can be defined as: $\Phi_n(u) = \Pr(U_k(u) < 0$ for some $k = 1, 2, \dots, n \mid U_0(u) = u$. From this equation, if $U_n(u) < 0$, it indicates that the insurance company doesn't have sufficient surplus to cover the obligations specified in the insurance contracts. This situation might lead to financial instability, potentially resulting in the suspension or regulation of business operations by the regulatory authority. The concept of surplus is crucial for insurance companies. Maintaining an appropriate surplus ensures financial stability and the ability to meet insurance claims promptly. However, an excessively large surplus might lead to reduced probabilities of ruin. Conversely, a lower surplus might restrict the company's ability to invest in other profitable ventures, such as stocks or real estate.

The analysis of these models aims to compute the probability of ruin, which represents the likelihood of financial distress for the insurance company. This is a complex mathematical problem and requires advanced techniques to evaluate. Sattayatham et al. [13] further extended this analysis by introducing definitions for the minimum initial capital that corresponds to acceptable levels of risk. Several studies [14-19] explore ruin probability in the discrete-time surplus process. Recent studies have deepened our understanding of ruin probability by exploring various complex scenarios. Yıldırım Külekci et al. [20] examined the impact of dependent extreme losses using a GARCH-EVT-Copula model, revealing the importance of considering loss dependence in risk assessments. Liu et al. [21] addressed delayed claims in renewal risk models, providing uniform asymptotic formulas for tail probabilities. Denisov et al. [22] explored the effects of level-dependent premium rates, showing how near-critical premium rates can produce heavy-tailed ruin probabilities. Li et al. [23] developed asymptotic formulas for finite-time ruin probabilities in multi-line risk models with stochastic returns, while Xu et al. [24] applied similar techniques to non-stationary claim-number processes. These studies collectively enhance the tools available for insurers to evaluate and mitigate financial risks effectively.

Despite the extensive research, the literature still lacks studies on the application of advanced statistical transforms, such as the Wang-PH transform, in modeling loss distributions for accurate ruin probability estimation. Additionally, there is a need for more empirical studies that validate these theoretical models using real-world data from various insurance sectors, particularly in emerging markets like Thailand. Another gap is the limited exploration of the interplay between different risk factors, such as claim frequency and severity, in determining ruin probabilities.

To address these gaps, our study focuses on the discrete-time surplus process of a company operating within a limited timeframe. The company aims to reserve sufficient capital to keep the probability of ruin below a specified risk level. We analyze the interplay between ruin probability, safety loading, and capital reserve. To transform the loss distribution data, we employ the Wang-PH transform, followed by fitting the resulting distributions to actual claim data. After thorough evaluation, we identify the optimal distribution that best represents the characteristics of the actual claim data. This model serves as the basis for estimating the finite-time ruin probability through claim simulation. We utilize the acceptance-rejection technique to generate random samples for simulation. Additionally, we apply a regression-based methodology to estimate the minimum capital reserve needed to mitigate financial risk.

The structure of the article is as follows: Section 2 begins with an examination of right-skewed loss distributions such as Gamma, Weibull, Lognormal, Log-logistic, Inverse Weibull, and Inverse Gaussian, known for their heavy tails and fitted to actual claim data sets. Section 3 utilizes the Wang-PH transform to enhance traditional loss distributions, addressing the diverse nature of claims distributions. Section 4 covers parameter estimation, which is conducted using maximum likelihood estimation (MLE) and the local minimum Kolmogorov-Smirnov estimator (LMKSE) with the random neighborhood search (RNS) technique, implemented via Scilab programming. Section 5 assesses the goodness of fit (GOF) test to evaluate the compatibility between theoretical probability distributions and observed data. Section 6 discusses the calculation of the net premium based on the expected value of losses, using relevant loss distributions, in Thai Baht (THB). This is followed by a simulation of ruin probability through claim simulation and analysis of the results in Section 8. Finally, the conclusions in Section 9 summarize the findings and their implications for the non-life insurance sector.

2- Loss Distributions

The loss distributions examined in this study display a right-skewness or skewed right distribution. Claim severity loss distributions are often characterized by heavy tails or fat tails. Gamma, Weibull, Lognormal, Log-logistic, Inverse Weibull, and Inverse Gaussian distributions are of particular interest as they are used to fit the actual claim data sets.

2-1- Gamma Distribution

Assume that X follows a Gamma distribution with shape parameter α and scale parameter β, denoted as $X \sim \text{Gam}(\alpha, \beta)$. The probability density function (pdf) and cumulative distribution function (cdf) of $X \sim \text{Gam}(\alpha, \beta)$ are given by Equations 1 and 2, respectively.

$$f(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}}, \qquad \alpha > 0, \beta > 0, x > 0 \tag{1}$$

$$F(x; \alpha, \beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \int_0^x t^{\alpha-1} e^{-\frac{t}{\beta}} dt, \quad \alpha > 0, \beta > 0, x > 0 \tag{2}$$

The expected value of $X \sim \text{Gam}(\alpha, \beta)$ is $E[X] = \alpha\beta$.

2-2- Weibull Distribution

Assume that X follows a Weibull distribution with scale parameter λ and shape parameter η, denoted as $X \sim \text{Weib}(\lambda, \eta)$. The pdf and cdf of $X \sim \text{Weib}(\lambda, \eta)$ are given by Equations 3 and 4, respectively.

$$f(x; \lambda, \eta) = \frac{\eta}{\lambda}\left(\frac{x}{\lambda}\right)^{\eta-1} e^{-\left(\frac{x}{\lambda}\right)^{\eta}}, \qquad \lambda > 0, \eta > 0, x \geq 0 \tag{3}$$

$$F(x; \lambda, \eta) = 1 - e^{-\left(\frac{x}{\lambda}\right)^{\eta}}, \qquad \lambda > 0, \eta > 0, x \geq 0 \tag{4}$$

The expected value of $X \sim \text{Weib}(\lambda, \eta)$ is $E[X] = \lambda \Gamma\left(\frac{\eta+1}{\eta}\right)$, where $\Gamma(\cdot)$ is the gamma function.

2-3- Lognormal Distribution

Let's assume that X follows a Lognormal distribution with location parameter μ and scale parameter σ, denoted as $X \sim \text{LN}(\mu, \sigma)$. The pdf and cdf of $X \sim \text{LN}(\mu, \sigma)$ are given by Equations 5 and 6, respectively.

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{\frac{-(\ln x - \mu)^2}{2\sigma^2}}, \qquad \mu \in \mathbb{R}, \sigma > 0, x > 0 \tag{5}$$

$$F(x; \mu, \sigma) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right), \qquad \mu \in \mathbb{R}, \sigma > 0, x > 0 \tag{6}$$

where $\Phi(\cdot)$ is the cdf of the standard normal distribution. The expected value of $X \sim \text{LN}(\mu, \sigma)$ is $E[X] = e^{\mu + \frac{\sigma^2}{2}}$.

2-4- Log-logistic Distribution

Let's assume that X follows a Log-logistic distribution with location parameter μ and scale parameter σ, denoted as $X \sim \text{Loglo}(\mu, \sigma)$. The pdf and cdf of $X \sim \text{Loglo}(\mu, \sigma)$ are given by Equations 7 and 8, respectively.

$$f(x; \mu, \sigma) = \frac{e^{\left(\frac{\ln x - \mu}{\sigma}\right)}}{\sigma x\left(1 + e^{\left(\frac{\ln x - \mu}{\sigma}\right)}\right)^2}, \qquad \mu > 0, \sigma > 0, x > 0 \tag{7}$$

$$F(x; \mu, \sigma) = \int_0^x \frac{e^{\left(\frac{\ln t - \mu}{\sigma}\right)}}{\sigma t\left(1 + e^{\left(\frac{\ln t - \mu}{\sigma}\right)}\right)^2} dt, \qquad \mu > 0, \sigma > 0, x > 0 \tag{8}$$

The expected value of $X \sim \text{Loglo}(\mu, \sigma)$ is $E[X] = \frac{e^{\mu}(\pi\sigma)}{\sin(\pi\sigma)}$, if $0 < \sigma < 1$, else undefined.

2-5- Inverse Weibull Distribution

Let's assume that X follows an inverse Weibull distribution with shape parameter α and scale parameter β, denoted as $X \sim \text{IW}(\alpha, \beta)$. The pdf and cdf of $X \sim \text{IW}(\alpha, \beta)$ are given by Equations 9 and 10, respectively.

$$f(x; \alpha, \beta) = \frac{\alpha}{x}\left(\frac{\beta}{x}\right)^{\alpha} e^{-\left(\frac{\beta}{x}\right)^{\alpha}}, \qquad \alpha > 0, \beta > 0, x > 0 \tag{9}$$

$$F(x; \alpha, \beta) = 1 - e^{-\left(\frac{\beta}{x}\right)^{\alpha}}, \qquad \alpha > 0, \beta > 0, x > 0 \tag{10}$$

The expected value of $X \sim \text{IW}(\alpha, \beta)$ is $E[X] = \beta\Gamma\left(1 - \frac{1}{\alpha}\right)$, for $\alpha > 1, \alpha \leq 1$, undefined, where $\Gamma(\cdot)$ is the gamma function.

2-6- Inverse Gaussian Distribution

Let's assume that X follows an Inverse Gaussian distribution with mean parameter μ and shape parameter λ, denoted as $X \sim \text{IG}(\mu, \lambda)$. The pdf and cdf of $X \sim \text{IG}(\mu, \lambda)$ are given by Equations 11 and 12, respectively.

$$f(x; \mu, \lambda) = \sqrt{\frac{\lambda}{2\pi x^3}}\, exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right], \qquad \mu > 0, \lambda > 0, x > 0 \tag{11}$$

$$F(x; \mu, \lambda) = \Phi\left(\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} - 1\right)\right) + exp\left(\frac{2\lambda}{\mu}\right)\Phi\left(-\sqrt{\frac{\lambda}{x}}\left(\frac{x}{\mu} + 1\right)\right), \qquad \mu > 0, \lambda > 0, x > 0 \tag{12}$$

where $\Phi(\cdot)$ is the standard normal distribution cdf. The expected value of $X \sim \text{IG}(\mu, \lambda)$ is $E[X] = \mu$.

3- Wang-PH Transform

Traditional models or loss distributions may not adequately fit claims data, leading to the utilization of mixture models to represent the diverse nature of claims distributions. One approach to modeling such mixture data includes transforming conventional loss distributions into new models using specific transformation functions. In this study, the Wang-PH transform is applied to several right-skewed loss distributions known for their heavy tails, including the Gamma, Weibull, Lognormal, Log-logistic, Inverse Weibull, and Inverse Gaussian distributions. The loss distributions selected for this study were chosen for their ability to capture the right-skewness and heavy tails characteristic of actual claim data. These distributions are commonly used in actuarial science for their flexibility in modeling claim severity, particularly in scenarios involving significant variation and the possibility of extreme losses.

In this analysis, individual claim policies are considered, with the claim amount X_i representing the payment for the i^{th} policy. The following assumptions are specified:

- ***Assumption 1:*** (Policy independence): For n distinct policies, the response variable X_i for the i^{th} policy is assumed to be independent of $X_1, X_2, \dots, X_n$.
- ***Assumption 2:*** Severity losses are categorized as non-catastrophic losses.
- ***Assumption 3:*** The loss distributions exhibit a right-skewed pattern.

The transformed models in this study involve the conversion from a loss distribution or non-transformed model $F(x)$ to a transformed model $F^*(x)$.

When considering any random variable X with its survival function denoted as $S_X(x)$, the Equation 13:

$$S_Y(x) = [S_X(x)]^c, \text{ where } c > 0 \tag{13}$$

defines a different random variable Y with its survival function represented as $S_Y(x)$. This transformation from X to Y is known as the proportional hazards (PH) transform. In this context, we can express it as:

$$S^*(x) = [S(x)]^c \tag{14}$$

$$1 - F^*(x) = [1 - F(x)]^c \tag{15}$$

$$F^*(x) = 1 - [1 - F(x)]^c \tag{16}$$

here, $F(x)$ represents the cdf of the loss distribution, and c is a positive constant. It's clear that $F^*(x)$ also qualifies as the cdf.

Let Φ represent the cdf of the standard normal distribution, defined as $\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}s^2} ds$. Additionally, consider θ as a real-valued parameter. The Wang transform converts a cdf, denoted as $F(x)$, into a new function, $F^*(x)$, using the formula:

$$F^*(x) = \Phi[\Phi^{-1}(F(x)) + \theta] \tag{17}$$

The proposed transformation method used in this study is the Wang-PH transform, which can be expressed as follows:

$$F^*(x) = \Phi\left\{\Phi^{-1}\left[1 - \left(1 - F(x)\right)^c\right] + \theta\right\} \tag{18}$$

here, Φ represents the cdf of the standard normal distribution, $F(x)$ is the cdf of the original loss distributions, and θ and c are constants.

The objective is to derive the pdf of the Wang-PH transform by differentiating its cdf. This involves applying the chain rule and the general form of the Leibniz integral rule. Consequently, the pdf of the Wang-PH transform can be expressed as follows:

$$f^*(x) = cf(x)[1 - F(x)]^{c-1} \exp\left\{-\theta\Phi^{-1}\left[1 - \left(1 - F(x)\right)^c\right] - \frac{\theta^2}{2}\right\} \tag{19}$$

In the Equation 19, $F(x)$ denotes the cdf of the original (non-transformed) models, $f(x)$ represents the pdf of the original models, and $\theta, c \in \mathbb{R}$ represent the parameters associated with the Wang transform and the PH transform, respectively. The Wang-PH transform is a statistical tool used to adjust loss distributions in risk models, making them more aligned with the actual risk profile. It is particularly valuable in insurance risk modeling, where accurately capturing the tail behavior of loss distributions is crucial for estimating ruin probabilities and determining appropriate capital reserves. The likelihood function of the Wang-PH transform can be defined as:

$$L = \prod_{i=1}^{n} f^*(x_i) \tag{20}$$

To simplify calculations, the log-likelihood function is obtained as:

$$\ln(L) = \ln \prod_{i=1}^{n} f^*(x_i) = \sum_{i=1}^{n} \ln \left[c f(x_i)[1 - F(x_i)]^{c-1} \exp\left(-\theta \Phi^{-1}(1 - [1 - F(x_i)]^c) - \frac{\theta^2}{2}\right)\right] = \\ n \ln c - \frac{\theta^2 n}{2} + \sum_{i=1}^{n} \ln f(x_i) + (c-1) \sum_{i=1}^{n} \ln\big(1 - F(x_i)\big) - \theta \sum_{i=1}^{n} \Phi^{-1}(1 - [1 - F(x_i)]^c) \tag{21}$$

4- Parameter Estimation

The maximum likelihood estimation (MLE) is utilized to estimate the parameters of the loss distributions. For the transformed models, the parameters are estimated using the local minimum Kolmogorov-Smirnov estimator (LMKSE), with the implementation of the random neighborhood search (RNS) technique. The parameter estimation process is carried out using Scilab programming in this study.

Consider a vector $X = (\boldsymbol{x_1}, \boldsymbol{x_2}, \dots, \boldsymbol{x_n})'$ as an independent observation, where each x_i represents the amount paid for the i^{th} contract. The loss distributions are fitted to the dataset using MLE. To find the most likely value of the parameter vector Θ of the loss distributions that best explains the outcomes in X, the likelihood function L is maximized. The likelihood function is defined as:

$$L(\Theta) = \prod_{i=1}^{n} f(x_i; \Theta) \tag{22}$$

where $f(x_i; \Theta)$ represents the pdf of the loss distributions.

To simplify calculations, the log-likelihood function is used, given by:

$$\ln L(\Theta) = \sum_{i=1}^{n} \ln[f(x_i; \Theta)] \tag{23}$$

To estimate the parameters (Θ), the maximum of the log-likelihood function is found by taking the partial derivative with respect to Θ_j and setting it equal to zero:

$$\frac{\partial}{\partial \Theta_j} \ln L(\Theta) = 0 \tag{24}$$

Solving the Equation 24 provides the estimated parameters $\widehat{\Theta}$ through the partial derivative method described above.

The randomized neighborhood search (RNS) technique is a numerical optimization method designed for objective functions that may be discontinuous and non-differentiable. RNS iteratively moves from an initial solution to a better solution while adhering to specified constraints. It is particularly well-suited for optimizing functions that are difficult to differentiate, making it useful for many global optimization problems. RNS ensures optimality by efficiently finding good solutions. In this study, we will employ the RNS technique to estimate parameters using statistical tests. The local minimum Kolmogorov-Smirnov estimator (LMKSE) is a method used for parameter estimation by minimizing the statistical value D of the Kolmogorov-Smirnov (K-S) test. The K-S test statistic is defined as:

$$D = \max |F^n(x) - F(x)| \tag{25}$$

where $F(x)$ represents the theoretical cdf of the non-transformed models, and $F^n(x)$ is defined as:

$$F^n(x) = \frac{1}{n} [\text{number of observations} \le x] \tag{26}$$

with n representing the sample size.

The procedure for minimizing the statistical value in parameter estimation with LMKSE, using the RNS technique, is described as follows:

- Step 1: Set the parameter range for c as $[0,5]$ and for θ as $[-0.5, 0.5]$, and determine the parameters of the loss distributions using the MLE method, denoted as Θ.
- Step 2: Begin with the initial iteration $i = 0$ and let the initial parameter vector (c, θ, Θ) be denoted as Z_0. Calculate the statistical value, δ_0, according to Equation 25.
- Step 3: Generate a new parameter vector Z_i^* by randomly perturbing the current parameter vector Z_i. This can be achieved by selecting a uniform random variable r from the interval $[0,1]$, and updating Z_i^* as $Z_i^* = Z_i + r\, Z_i$, where $r \in [-0.5, 0.5]$.

- Step 4: Compute the statistical value δ_i using the parameter vector Z_i^*.
- Step 5: Compare the statistical values δ_{i-1} for the $(i-1)^{\text{th}}$ iteration and δ_i for the ith iteration. If $\delta_i < \delta_{i-1}$, update Z_i as $Z_i^* = Z_i$ and proceed to Step 6. Otherwise, return to Step 3.
- Step 6: If $|\delta_i - \delta_{i-1}| \leq 10^{-5}$, consider the process complete. Otherwise, return to Step 3.

This iterative process continues until the difference in statistical values falls below a predefined threshold, indicating convergence and completion of the parameter estimation process. This parameter estimation method is implemented using Scilab programming.

The parameter estimation process for the different loss distributions is a multi-step approach that begins with initial estimation using MLE, followed by an optimization phase using the Wang-PH transform. This involves minimizing the K-S statistic through the LMKSE method, supported by RNS to ensure a robust fit to the actual claim data. The entire process is computationally intensive and implemented in Scilab, resulting in parameters that optimize the fit between the transformed models and the empirical data, thereby enhancing the accuracy of the subsequent risk assessments.

5- Goodness of Fit Test

The goodness of fit (GOF) test assesses the compatibility between a theoretical probability distribution function and a given set of observations, measuring how well the distribution fits the random sample. The K-S test statistic D quantifies the discrepancy between the empirical cumulative distribution function $F^n(x)$ and the cumulative distribution function $F(x)$ of the non-transformed and transformed models. The D-value in the K-S test is calculated as the maximum absolute difference between the cumulative distributions:

$$D = \max|F^n(x) - F(x)| \tag{27}$$

The K-S test statistic measures the largest vertical distance between the empirical distribution and the theoretical distribution. A smaller value of D indicates a better fit between the models and the observed data. The Akaike Information Criterion (AIC) is utilized as a criterion for selecting the most suitable model. Among all the models considered, the one with the lowest AIC value is deemed the best fit. The AIC is calculated using the following estimation equation:

$$\text{AIC} = -2\ln(L) + 2m \tag{28}$$

where $\ln(L)$ represents the natural logarithm of the likelihood function value of the model, and m represents the number of estimated parameters in the model [25].

6- Insurance Premium Calculation

The net premium principle holds significant importance within the insurance industry as it aids in determining appropriate premium amounts for insurance policies. This principle relies on evaluating the expected value of losses to establish the net premium. The net premium signifies the anticipated number or average of claims policyholders are expected to file. The current study primarily focuses on the application of the net premium principle, which serves as a foundational concept stating that premiums should correspond to the expected value of losses. Specifically, this study calculates the insurance premium, denoted as $H[X]$, in Thai Baht (THB) by considering the relevant loss distributions.

The net premium principle involves calculating the expected value or mean of claim amounts without considering any risk factors or risk loading. This expected value is denoted as $E[X]$. As a result, the net premium for the loss variable X, represented by $H[X]$, is defined as follows: $H[X] = E[X]$. The expected value of the loss variable X can be calculated using the integral expressions:

$$E[X] = \int_0^\infty xf(x)dx \text{ or } E[X] = \int_0^\infty [1 - F(x)]dx \tag{29}$$

here, $f(x)$ represents the pdf of the loss distributions and $F(x)$ represents the cdf of the loss distributions. These integral formulas provide a way to calculate the expected value of the loss X based on its probability distribution.

The expected value, $E[X]$, for non-transformed models can be determined based on the respective distributions, where the parameters are estimated using the MLE method.

Gamma: $$E[X] = \hat{\alpha}\hat{\beta} \tag{30}$$

Weibull: $$E[X] = \hat{\lambda}\Gamma\left(\frac{\hat{\eta}+1}{\hat{\eta}}\right) \tag{31}$$

Lognormal: $$E[X] = e^{\hat{\mu}+\frac{\hat{\sigma}^2}{2}} \tag{32}$$

Log-logistic: $$E[X] = \frac{e^{\hat{\mu}}(\pi\hat{\sigma})}{\sin(\pi\hat{\sigma})} \tag{33}$$

Inverse Weibull: $$E[X] = \hat{\beta}\Gamma\left(1 - \frac{1}{\hat{\alpha}}\right) \tag{34}$$

Inverse Gaussian: $E[X] = \hat{\mu}$ (35)

The net premium of loss X in the transformed models, denoted as $H[X]$, is defined using the following formula:

$$H[X] = E^*[X] = \int_0^\infty [1 - F^*(x)]dx \tag{36}$$

where $E^*[X]$ is the expected value obtained via the Wang-PH transform. In the Equation 36, $F^*(x)$ represents the Wang-PH transform given by:

$$F^*(x) = \Phi\{\Phi^{-1}[1 - (1 - F(x))^c] + \theta\} \tag{37}$$

here, Φ denotes the cdf of the standard normal distribution, $F(x)$ represents the cdf of the original loss distributions, and θ and c are constants used in the Wang-PH transform.

7- Simulation

Ruin probability is a critical concept in risk theory and insurance, representing the likelihood of a financial entity facing insolvency or bankruptcy. In the context of insurance, it assesses the probability that the cumulative losses incurred by the insurer exceed the available surplus, leading to financial ruin.

The focus of this study is on the discrete-time risk process or surplus process, which operates under the condition that insolvency, or ruin probability, can exclusively happen at claim arrival times denoted as $T_n = n$, with n belonging to the set $\{1,2,3, \dots\}$. To account for this, a value of $Z_n = 1$ is assigned when n is an element of the set $\{1,2,3, \dots\}$. As a result, the surplus process can be adjusted using the following formulation:

$$U_n(u) = u + \rho \sum_{k=1}^{n} Z_k - \sum_{k=1}^{n} X_k == u + \rho n - \sum_{k=1}^{n} X_k\,,\ n \in \mathbb{N}, = U_{n-1}(u) + \rho - X_n, \quad n \in \mathbb{N}. \tag{38}$$

Here, $U_0(u) = u \geq 0$ represents the initial capital reserve, and $\rho > 0$ denotes the premium rate for a single unit of time. The sequence $\{X_n : n \in \mathbb{N}\}$ refers to a series of independent and identically distributed (i.i.d.) claim random variables at the claim arrival $T_n = n$, where $n \in \mathbb{N}$. The premium rate ρ is calculated according to the expected value principle, i.e.,

$$\rho = (1 + \vartheta)H[X] \tag{39}$$

In Equation 39, ϑ takes on values of 0, 0.05, 0.10, 0.15, and 0.20, representing the safety loading of an insurer. $H[X]$ represents the expected value of claim severity, calculated using Equation 36. Mathematically, ruin probability is often denoted by the symbol $\varphi_n(u)$, where u represents the initial surplus or capital of the insurer. The function $\varphi_n(u)$ quantifies the probability that the surplus falls below zero as a result of accumulating insurance claims. Consequently, the ruin probability at one of the times $k = 1,2, \dots, n$ is defined by:

$$\varphi_n(u) = \Pr(U_k(u) < 0 \text{ for some } \ k = 1,2, \dots, n \mid U_0(u) = u) \tag{40}$$

The research conducted by Sattayatham et al. [13] introduced the definition of the minimum initial capital as follows:

Consider the surplus process $\{U_n(u), n \in \mathbb{N}\}$, which is influenced by the claim process $\{X_n, n \in \mathbb{N}\}$, and let $\rho > 0$ be a premium rate. Given $\alpha \in (0,1)$ and $n \in \mathbb{N}$, an initial capital $u \geq 0$, is deemed an ***acceptable capital reserve*** corresponding to $(\alpha, n, \rho, \{X_n, n \in \mathbb{N}\})$ if $\varphi_n(u) \leq \alpha$. Specifically, if there exists:

$$u^* = \min\{u | \varphi_n(u) \leq \alpha\} \tag{41}$$

It is referred to as the ***minimum initial capital*** corresponding to $(\alpha, n, \rho, \{X_n, n \in \mathbb{N}\})$, denoted as:

$$u^* \coloneqq \text{MIC}\left((\alpha, n, \rho, \{X_n, n \in \mathbb{N}\})\right) \tag{42}$$

The acceptance-rejection technique is a traditional sampling technique used to generate samples from a distribution that is challenging or impossible to simulate using inverse transformation, as described by Gamerman & Lopes [26]. The acceptance-rejection technique is utilized in this study to generate a random variable X that conforms to the desired continuous density $f(x)$. The procedure involves generating a random number Y from a distribution $g(y)$ and accepting this value with a probability proportional to the ratio $\frac{f(Y)}{g(Y)}$. This allows us to effectively sample from the desired distribution and overcome the limitations of direct simulation.

If we let m be a constant satisfying the condition $m \geq \frac{f(y)}{g(y)}$ for all y, The desired variates can be generated by following the outlined procedure below. The constant m is necessary to ensure that the height of $g(y)$ can be adjusted if needed to surpass $f(y)$. Points are generated from $m \cdot g(y)$, and those falling within the curve of $f(y)$ are accepted as samples from the desired density, while those outsides are rejected. The procedure is outlined as follows:

- Step 1: Choose a density $g(y)$ that is easy to sample from.
- Step 2: Determine a constant m such that $m = \sup_y \frac{f(y)}{g(y)}$.
- Step 3: Generate a random number Y from the density $g(y)$.
- Step 4: Generate a uniform random number U.
- Step 5: If $U \leq \frac{f(Y)}{m \cdot g(Y)}$, accept $X := Y$ as a sample from the desired density. If not, return to step 3 and repeat the process until the desired number of samples is obtained.

The acceptance-rejection technique faced challenges, particularly in ensuring efficiency when the scaling constant m was large, leading to low acceptance rates and slower simulations. This was mitigated by carefully selecting a proposal distribution $g(y)$ that closely matched the target distribution, minimizing m and improving efficiency. Additionally, validating the generated samples was crucial for accurate simulation results, achieved through goodness-of-fit tests and P-P plots that compared the empirical samples to the theoretical distribution. These strategies ensured the generation of valid samples, contributing to the reliability of the simulation of ruin probabilities in the study.

The flowchart in Figure 1 represents a comprehensive method for simulating ruin probability over a set number of iterations, using a statistical approach with the Wang-PH transform to generate samples and evaluate risk over time.

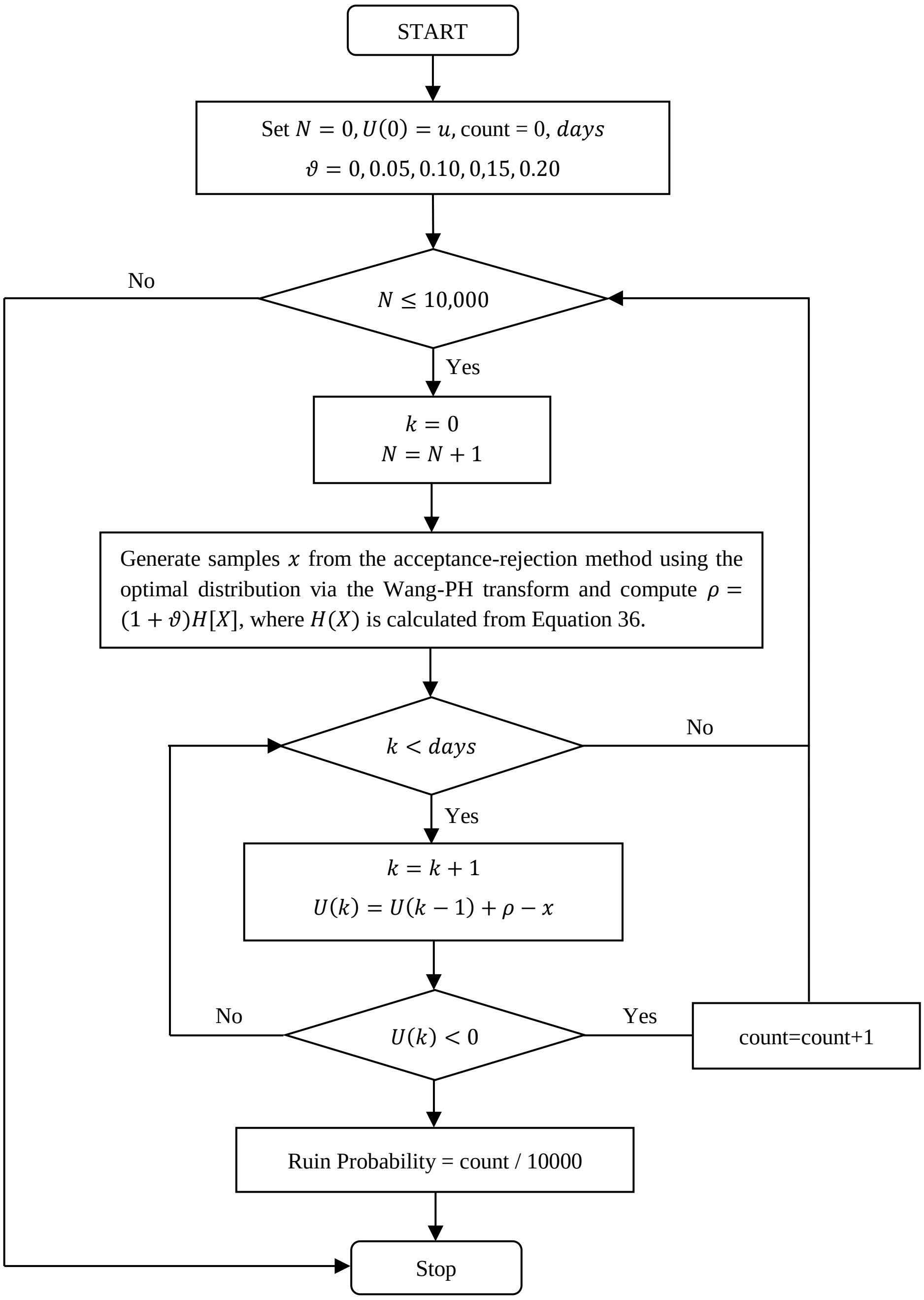

Figure 1. The flowchart of algorithm of simulation

8- Results and Discussion

The application of both non-transformed and transformed models is conducted on actual motor insurance claims. The dataset used for analysis is provided by a non-life insurance company in Thailand and consists of policy claims for the year 2009. The specific type of coverage in the dataset is 2+ under a voluntary plan. A total of 1,296 policy claims were observed for analysis, and Figure 2 displays the histogram of claim severity data on a log scale.

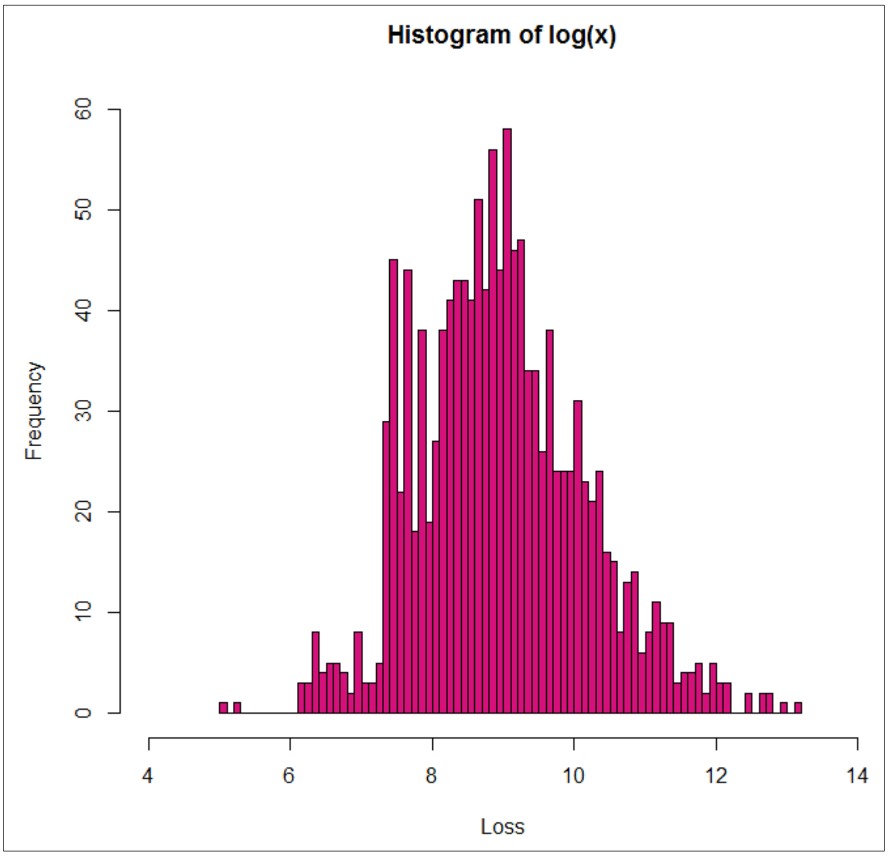

Figure 2. Histogram (log scale) of claim data

Table 1 displays the estimated parameters, statistical values, and premiums for the non-transformed models and the Wang-PH transform in relation to each loss distribution. With a significance level of 0.05, the critical value for the K-S test is 0.037778. Upon examination, it is evident that all D-values exceed the critical value for all loss distributions. Hence, it can be concluded that the non-transformed models do not provide a suitable fit for the actual data set.

Considering the inability of the loss distributions or non-transformed models to be fitted to any actual data sets, the alternative approach involves applying the transformed models of loss distributions, specifically the Wang-PH transform. By varying the values of θ, c, and re-parameterization, the new estimated parameters result in a reduction of the D-value for each model. Among the transformed models, the Lognormal distribution stands out as the most favorable fit. It consistently exhibits the lowest values in all statistical tests, followed by the Inverse Gaussian, Log-logistic, Inverse Weibull, Weibull, and Gamma distributions, respectively. Notably, all loss distributions based on the transformed models provide a better fit to the data set compared to the non-transformed models.

Figure 3 illustrates a P-P plot, showcasing the comparison of loss distributions. The plot provides a visual representation of how well the empirical cumulative distribution aligns with the theoretical cumulative distribution. Figure 4 presents a P-P plot specifically for loss distributions obtained through the Wang-PH transform. It is evident that all loss distributions based on the transformed models provide a better fit to the dataset compared to the non-transformed models. In Figure 5, a P-P plot focuses on the Lognormal distribution when processed through the Wang-PH transform. This figure is presented separately to enhance clarity, allowing for a more detailed examination of the graph. Notably, it demonstrates the Lognormal distribution as the best fit to the actual claim data.

Table 1. Summary of model fitting and insurance premiums (THB)

Loss distributions	Items	Non-transform	Wang-PH transform	Loss distributions	Items	Non-transform	Wang-PH transform
Gamma	α	0.7528	1.1256	Log-logistic	μ	8.9244	8.5702
	β	23,053.21	17,409.36		σ	0.6654	0.5397
	θ	-	0.1610		θ	-	0.1371
	c	-	1.4408		c	-	0.5843
	D-value	0.1453	0.0668		D-value	0.0380	0.0258
	p-value	<0.01	0.0058		p-value	<0.01	0.7639
	AIC	27,821.09	28,549.12		AIC	27,354.32	27,344.31
	Premium	17,353.90	12,085.91		Premium	18,092.53	36,402.81
Weibull	λ	14,398.56	12,911.34	Inverse Weibull	α	0.9057	0.9827
	η	0.7840	1.0640		β	4429.28	6010.16
	θ	-	0.0550		θ	-	0.0803
	c	-	1.0218		c	-	0.7226
	D-value	0.1126	0.0646		D-value	0.0554	0.0334
	p-value	<0.01	0.0086		p-value	<0.01	0.4439
	AIC	27,704.45	28,035.66		AIC	27,434.59	27,270.16
	Premium	16,557.37	11,774.07		Premium	-	-
Lognormal	μ	8.9667	7.8413	Inverse Gaussian	μ	17,353.90	14,881.19
	σ	1.1787	0.6850		λ	5,523.65	6,929.40
	θ	-	0.0289		θ	-	0.1119
	c	-	0.2451		c	-	0.9056
	D-value	0.0463	**0.0216**		D-value	0.0436	0.0237
	p-value	<0.01	**0.9228**		p-value	<0.01	0.8523
	AIC	27,349.89	**27,344.81**		AIC	27,377.19	27,414.66
	Premium	15,700.31	14,732.85		Premium	17,353.90	15,057.84

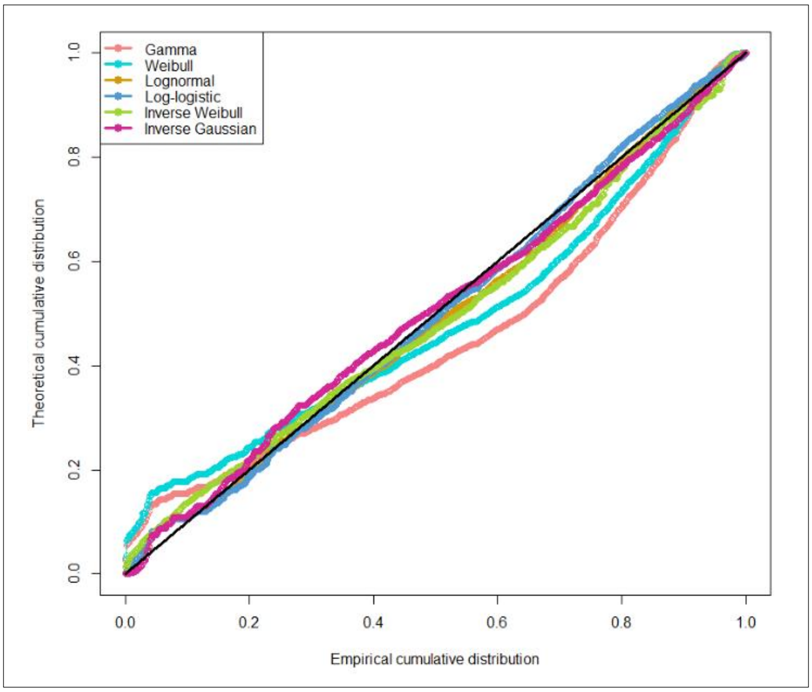

Figure 3. P-P plot of loss distributions

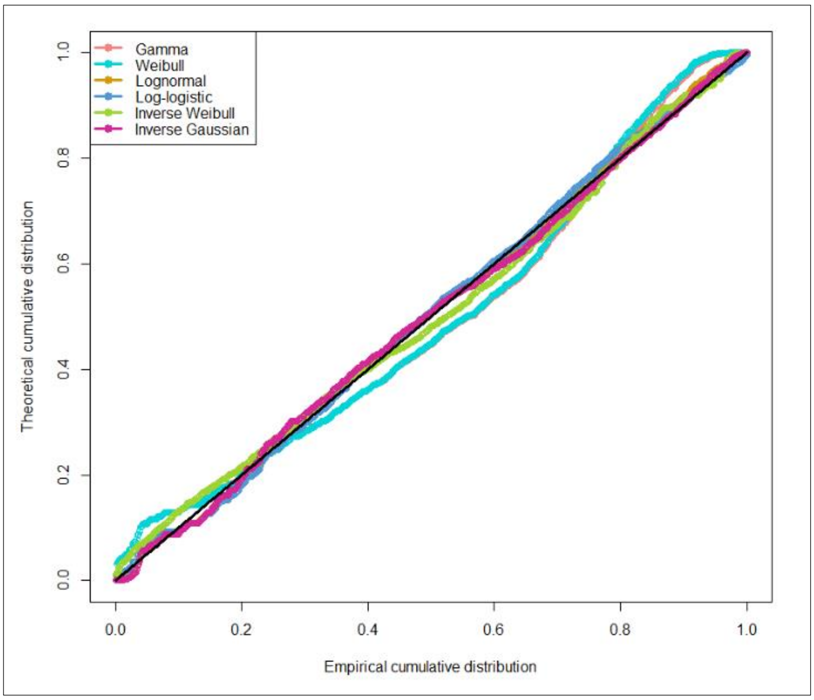

Figure 4. **P-P plot of loss distributions via Wang-PH transform**

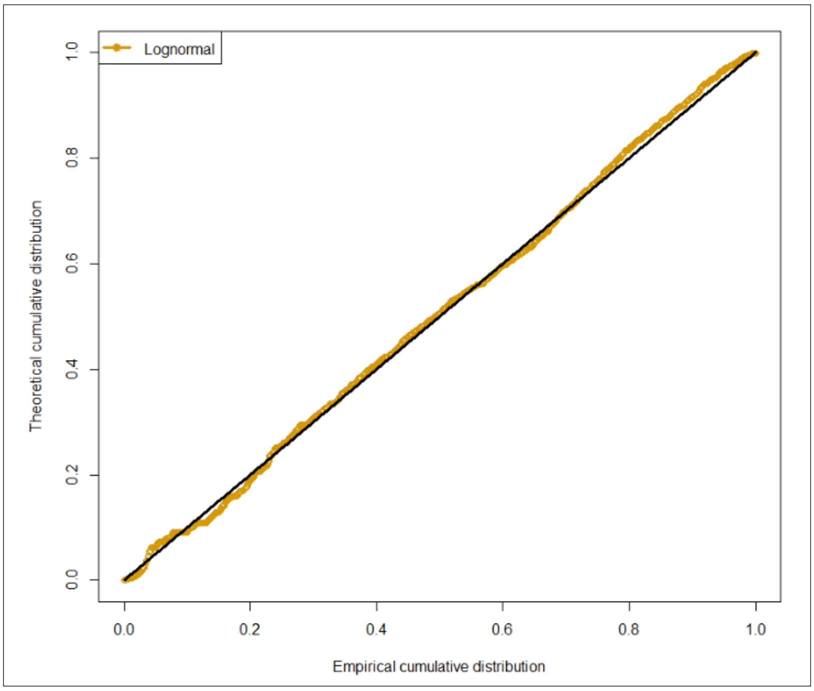

Figure 5. **P-P plot of Lognormal via Wang-PH transform**

The application of the Wang-PH transform resulted in a significantly improved model fit across all examined distributions. This improvement is particularly evident in the Lognormal distribution, where the D-value decreased substantially, indicating a much better alignment between the transformed model and the actual claim data. This finding suggests that the Wang-PH transform is highly effective in capturing the nuances of the loss distribution, especially in cases where the data exhibit heavy tails or significant skewness. The Wang-PH transform's ability to adjust the tail behavior of loss distributions is critical in insurance modeling. In practice, claim data often display tail behaviors that standard distributions cannot adequately capture, leading to underestimation or overestimation of risk. The transformation modifies the original distribution to better reflect real-world data, thus providing more accurate risk assessments and premium calculations.

However, when claim data shows low variability or is more symmetric, traditional models like Gamma or Weibull distributions are sufficient without requiring transformation. The Wang-PH transform does not notably enhance accuracy in these cases, making traditional models more efficient and user-friendly. While Wang-PH transformed models excel in handling extreme tail risks and skewed distributions common in insurance claims, they add computational complexity. Traditional models, though less precise in capturing tail risks, are faster and more efficient, fitting scenarios with symmetric or low-variability data.

The Wang-PH transform appears to reduce the premium compared to the non-transformed approach for the Gamma, Weibull, Lognormal and Inverse Gaussian distributions. However, it is worth noting that for the Log-logistic distribution, the Wang-PH transform increases the premium compared to the non-transformed approach. This result suggests that the transform may not be appropriate for this distribution in the context of premium calculations. The impact of the Wang-PH transform on premium calculations varies across different loss distributions. In some cases, it leads to lower premiums (Gamma, Weibull, Lognormal, and Inverse Gaussian), indicating a potential increase in risk tolerance. This reduction can be interpreted as a reflection of a more accurate assessment of risk, where the transformation accounts for the actual distribution of claim severity. In practical terms, this allows insurers to price their policies more competitively without compromising on the adequacy of their risk reserves. Conversely, for the Log-logistic distribution, it leads to higher premiums, suggesting a more conservative approach to risk assessment. Actuaries and risk analysts should carefully assess the appropriateness of the chosen transformation method in the context of the specific insurance application and the underlying data characteristics.

Considering a dataset for claim size from a non-life insurance company in Thailand in 2009, Table 1 reveals that the Lognormal distribution, transformed by the Wang-PH transform, demonstrates the best fit for the claim severity data. Therefore, the assumption is made that X_n follows the Lognormal distribution transformed by the Wang-PH transform for all n. We present the capital reserves u as 0, 150000, 300000, 450000, 600000, 750000, 900000 and 1000000 Thai Baht (THB). Let's consider different values for the safety loading, ϑ, of a company, specifically 0, 0.05, 0.10, 0.15, and 0.20. We will also explore three different durations: one year ($n = 1$), three years ($n = 3$), and five years ($n = 5$). In this study, the discrete-time risk process is studied by the equation:

$$U_n(u) = U_{n-1}(u) + \rho - X_n, \qquad n \in \mathbb{N}. \tag{43}$$

here, $U_0(u) = u \geq 0$ represents the initial capital reserve. The variable X_n follows the Lognormal distribution transformed by the Wang-PH transform for all n. Additionally, $\rho > 0$ represents the premium rate for one day, calculated as:

$$\rho = (1 + \vartheta)H[X] \tag{44}$$

where ϑ is the safety loading and $H[X]$ denotes the expected value of claim severity calculated by Equation 36 over n days. The focus is on the surplus process, assuming that the possibility of insolvency (ruin) only arises at claim arrival $T_n = n$, where $n \in \mathbb{N}$. Therefore, the ruin probability at any given time k, where $k = 1,2,3,\dots,n$, can be calculated as:

$$\varphi_n(u) = Pr(U_k(u) < 0\ for\ some\ k = 1,2,\dots,n | U_0 = u) \tag{45}$$

For obtaining simulation results of the surplus process, 10000 paths are employed.

Tables 2 to 4 present the results of approximating the finite-time ruin probability for n = 1 year, 3 years, and 5 years, as depicted in Figures 6 to 8, respectively. It is observed that as the safety loading (ϑ) increases, the finite-time ruin probability generally decreases. This is intuitive because higher safety loading implies more capital reserves are set aside to cover potential losses, reducing the risk of financial ruin. As the initial capital reserves (u) increase, the finite-time ruin probability also decreases. More substantial capital reserves act as a buffer against unexpected losses, reducing the likelihood of financial ruin. Increasing safety loading makes the firm safer, but it also ties up more capital that could potentially be used for investments or other purposes. These calculations are critical for risk assessment. They help organizations determine the adequacy of their capital reserves and the level of risk they are willing to tolerate. Depending on their risk appetite and regulatory requirements, organizations can use this data to make informed decisions about capital allocation, risk mitigation, and financial stability. As the time horizon (n) increases, the ruin probabilities generally increase. This suggests that over longer time periods, there is a higher risk of the institution experiencing financial ruin if capital reserves remain unchanged.

Table 2. Approximating the finite-time ruin probability of the surplus process via simulation for n = 1 year

Safety loading (ϑ)	Initial capital reserves (u)							
	0	150,000	300,000	450,000	600,000	750,000	900,000	1,000,000
0.00	0.8947	0.5226	0.2933	0.1532	0.0755	0.0347	0.0165	0.0084
0.05	0.8324	0.3835	0.1698	0.0754	0.0306	0.0125	0.0044	0.0023
0.10	0.7657	0.2749	0.0982	0.0340	0.0116	0.0038	0.0012	0.0006
0.15	0.7059	0.2018	0.0580	0.0158	0.0042	0.0014	0.0006	0.0005
0.20	0.6555	0.1509	0.0364	0.0077	0.0021	0.0006	0.0003	0.0001

Table 3. Approximating the finite-time ruin probability of the surplus process via simulation for n = 3 years

Safety loading (ϑ)	Initial capital reserves (u)							
	0	150,000	300,000	450,000	600,000	750,000	900,000	1,000,000
0.00	0.9246	0.6254	0.4332	0.2984	0.2042	0.1450	0.1042	0.0854
0.05	0.8552	0.4284	0.2394	0.1396	0.0880	0.0626	0.0494	0.0438
0.10	0.7880	0.3058	0.1384	0.0752	0.0472	0.0350	0.0278	0.0234
0.15	0.7280	0.2218	0.0892	0.0424	0.0248	0.0150	0.0118	0.0098
0.20	0.6720	0.1676	0.0552	0.0234	0.0098	0.0056	0.0036	0.0024

Table 4. Approximating the finite-time ruin probability of the surplus process via simulation for n = 5 years

Safety loading (ϑ)	Initial capital reserves (u)							
	0	150,000	300,000	450,000	600,000	750,000	900,000	1,000,000
0.00	0.9493	0.7282	0.5675	0.4407	0.3375	0.2562	0.1937	0.1576
0.05	0.8751	0.4960	0.2930	0.1739	0.1035	0.0624	0.0378	0.0278
0.10	0.8042	0.3352	0.1496	0.0664	0.0307	0.0141	0.0069	0.0045
0.15	0.7419	0.2361	0.0808	0.0289	0.0097	0.0036	0.0018	0.0008
0.20	0.6870	0.1752	0.0470	0.0133	0.0037	0.0014	0.0004	0.0002

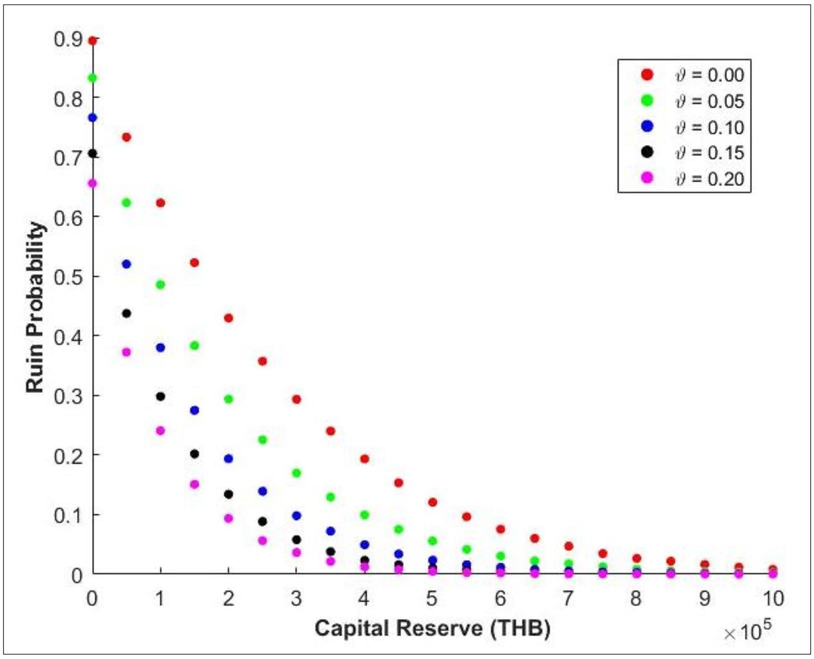

Figure 6. The relation between ruin probabilities and initial capital reserves for n = 1 year

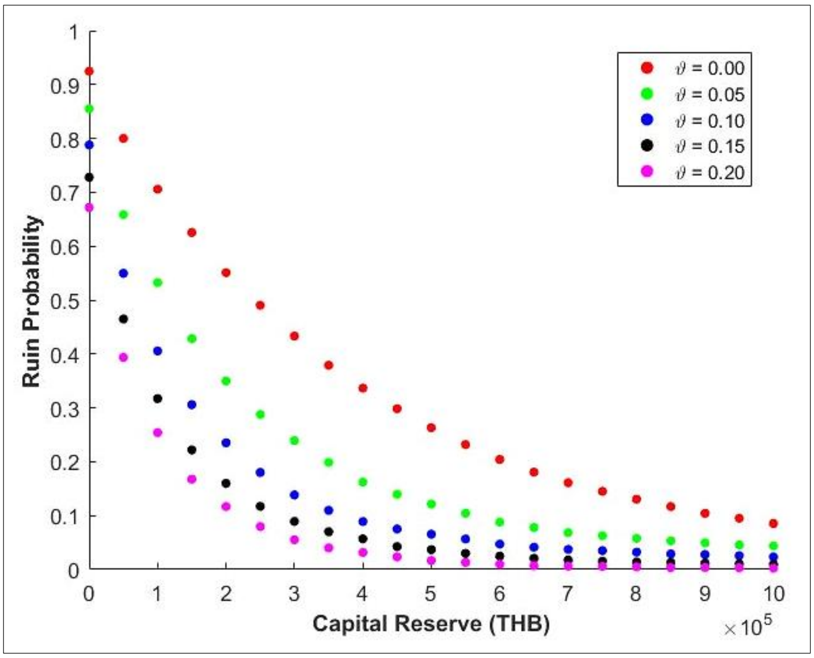

Figure 7. The relation between ruin probabilities and initial capital reserves for n = 3 years

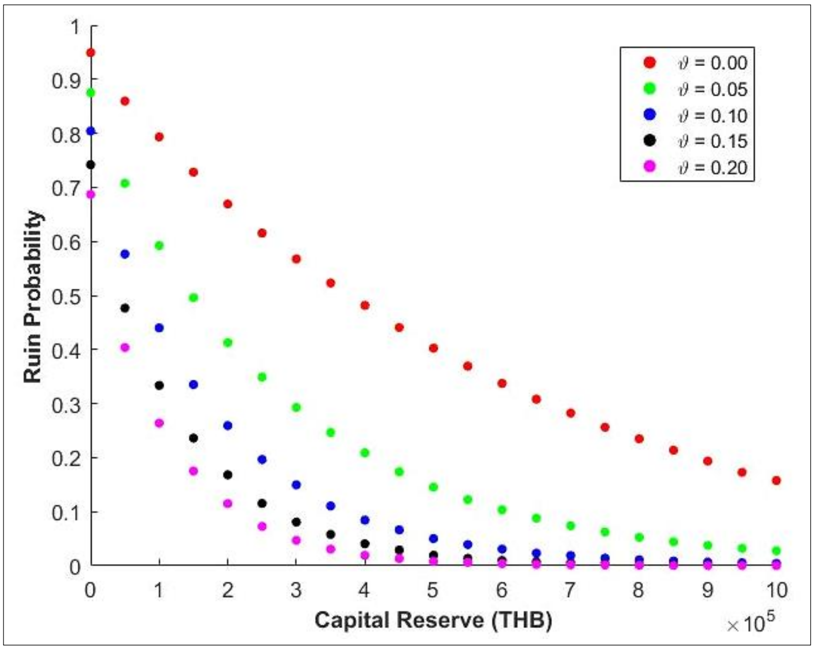

Figure 8. The relation between ruin probabilities and initial capital reserves for n = 5 years

In practical terms, these results provide valuable insights into how safety loading and initial capital reserves impact the financial risk profile of an organization. They can be used to inform risk management strategies and financial planning decisions to ensure the organization remains solvent and resilient in the face of unexpected events.

Figures 6 to 8 reveal a relationship between the ruin probability and the capital reserve, which can be described by an exponential function:

$$\varphi_n(u) = \gamma e^{\delta u} \tag{46}$$

In this study, we employ data transformation to estimate the parameters and establish the relationship between φ and u in the exponential equation:

$$\varphi = \gamma e^{\delta u} \tag{47}$$

Taking the natural logarithm of both sides, the equation becomes:

$$\ln \varphi = \ln \gamma + \delta u \tag{48}$$

Let's define $z = \ln \varphi$, $a_0 = \ln \gamma$, and $a_1 = \delta$. Therefore,

$$z = a_0 + a_1 u \tag{49}$$

The parameters (a_0, a_1) are estimated using linear regression analysis in Equation 49 and the least squares method. The parameter estimations are computed using the following formulas:

$$a_1 = \frac{n\sum_{i=1}^{n} u_i z_i - \sum_{i=1}^{n} u_i \sum_{i=1}^{n} z_i}{n\sum_{i=1}^{n} u_i^2 - \left(\sum_{i=1}^{n} u_i\right)^2} \text{ and } a_0 = \bar{z} - a_1\bar{u} \tag{50}$$

Once a_0 and a_1 are determined, the original constants of the model $\varphi = \gamma e^{\delta u}$ can be obtained, where $\delta = a_1$ and $\gamma = e^{a_0}$. Thus,

$$\delta = \frac{n\sum_{i=1}^{n} u_i z_i - \sum_{i=1}^{n} u_i \sum_{i=1}^{n} z_i}{n\sum_{i=1}^{n} u_i^2 - \left(\sum_{i=1}^{n} u_i\right)^2} \text{ and } \gamma = e^{\bar{z} - a_1\bar{u}} \tag{51}$$

where u_i represents the capital reserve values, and φ_i corresponds to the ruin probabilities obtained from simulation for each capital reserve value u_i (i = 1, 2, 3, ..., 21). Since u_i starts at 0 and increases by 150,000 up to 1000000, we have $n = 21$. Note that the values shown in the Tables 2 to 4 are the results for selected values of u. Tables 5 to 7 show the R-squared values and parameters of exponential regression for the respective 1-year, 3-year, and 5-year periods, respectively (also, Figures 9 to 11).

Table 5. **Exponential regression parameters and R-squared for a 1-year simulation**

Safety loading (ϑ)	Parameters		R^2
	γ	δ	
0.00	1.0847703	- 0.0000046	0.9571569
0.05	0.9531300	- 0.0000059	0.9771554
0.10	0.8028419	- 0.0000071	0.9958134
0.15	0.5919356	- 0.0000078	0.9759606
0.20	0.5034191	- 0.0000088	0.9493077

Table 6. **Exponential regression parameters and R-squared for a 3-year simulation**

Safety loading (ϑ)	Parameters		R^2
	γ	δ	
0.00	0.8878975	-0.0000024	0.9985196
0.05	0.6332186	-0.0000030	0.9307327
0.10	0.4637410	-0.0000034	0.8261559
0.15	0.3865585	-0.0000042	0.7814896
0.20	0.3526579	-0.0000055	0.7811460

Table 7. **Exponential regression parameters and R-squared for a 5-year simulation**

Safety loading (ϑ)	Parameters		R^2
	γ	δ	
0.00	0.9608174	- 0.0000018	0.9989916
0.05	0.8303421	- 0.0000034	0.9981847
0.10	0.7191637	- 0.0000052	0.9915185
0.15	0.6173419	- 0.0000067	0.9767126
0.20	0.5519538	- 0.0000081	0.9664280

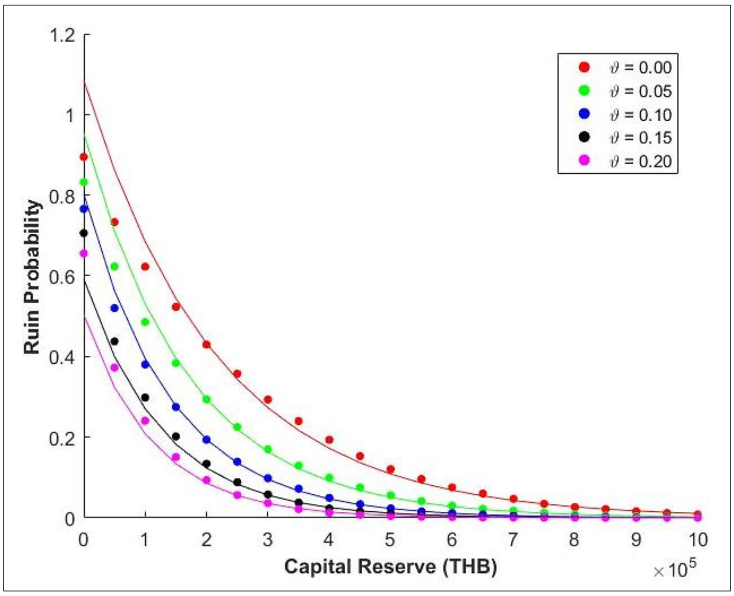

Figure 9. Ruin probabilities and capital reserves relation via exponential regression lines for n = 1 year

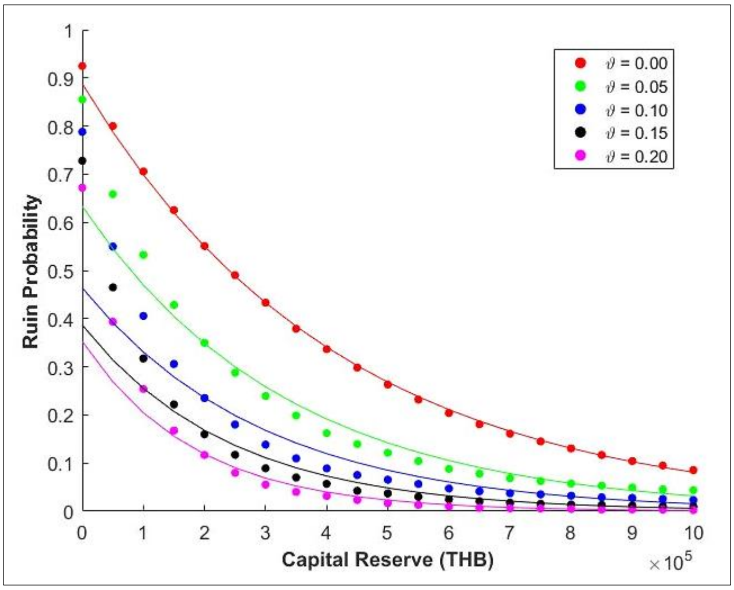

Figure 10. Ruin probabilities and capital reserves relation via exponential regression lines for n = 3 years

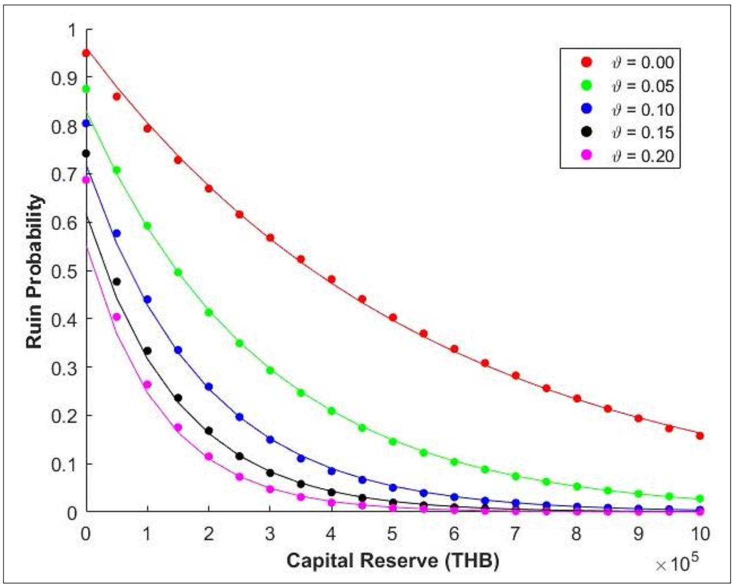

Figure 11. Ruin probabilities and capital reserves relation via exponential regression lines for n = 5 years

The results show that the R-squared values are relatively high for all levels of safety loading. This indicates that the exponential regression model is a good fit for the data and provides valuable insights into the relationship between safety loading, capital reserve, and finite-time ruin probability. These regression parameters can be used to model and predict the finite-time ruin probability for various levels of safety loading within the studied range. This information is valuable for decision-makers in risk management, as it helps them understand how different levels of safety loading affect financial stability and risk exposure

The maximum acceptable risk, denoted by α, is established. Consequently, the ruin probability under the regulation must not exceed α. In other words:

$$\varphi(u) \leq \alpha \tag{52}$$

Additionally, the capital reserve must satisfy the following inequality:

$$\gamma e^{\delta u} \leq \alpha \tag{53}$$

Therefore, we can derive the condition for the capital reserve as:

$$u \geq \frac{1}{\delta} \ln\left(\frac{\alpha}{\gamma}\right) \tag{54}$$

For a non-dangerous portfolio or when the premium rate is sufficiently high, the capital reserve u might be negative. In such cases, having a capital reserve is not necessary. Consequently, the minimum capital reserve (MCR) is determined as:

$$\mathrm{MCR} = \max\left\{0, \frac{\ln\alpha - \ln\gamma}{\delta}\right\} \tag{55}$$

The MCR is estimated using an exponential regression model, which models the relationship between ruin probability $\varphi(u)$ and capital reserve u as $\varphi(u) = \gamma e^{\delta u}$. Parameters γ and δ are estimated via linear regression on the log-transformed ruin probability data. The minimum capital reserve u^* is then derived by solving $\phi(u^*) = \alpha$, leading to $u^* = \frac{1}{\delta}\ln\left(\frac{\alpha}{\gamma}\right)$. This approach assumes constant risk parameters, independence of claim arrivals and severities, and homogeneity of claims. The model's accuracy is sensitive to the stability of these assumptions, particularly the constancy of γ and δ, the independence of claims, and the log-linear relationship between capital and ruin probability. In practice, it is essential to validate these assumptions regularly, especially in dynamic risk environments, to ensure that the estimated capital reserves remain reliable and accurate.

Tables 8 to 10 provide the minimum capital reserve for different acceptable risks α across various time horizons: one year, three years, and five years. Meanwhile, Figures 12 to 14 illustrate the relationship between safety loading and the minimum capital reserve for ruin probabilities (0.01, 0.05, and 0.10), representing 1-year, 3-year, and 5-year durations. These findings highlight the importance of insurance companies maintaining a minimum capital reserve to ensure that the ruin probability does not exceed the acceptable risks α=0.01, 0.05, and 0.10.

Table 8. The link between minimum capital reserve (MCR), safety loading, and acceptable risk over a 1-year period

$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$	
ϑ	MCR	ϑ	MCR	ϑ	MCR
0.00	1,018,162.11	0.00	668,507.72	0.00	517,919.78
0.05	775,963.71	0.05	501,919.41	0.05	383,894.96
0.10	617,385.93	0.10	390,814.76	0.10	293,235.87
0.15	520,128.06	0.15	314,993.97	0.15	226,647.52
0.20	443,360.30	0.20	261,275.49	0.20	182,855.83

Table 9. The link between minimum capital reserve (MCR), safety loading, and acceptable risk over a 3-year period

$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$	
ϑ	MCR	ϑ	MCR	ϑ	MCR
0.00	1,873,025.41	0.00	1,201,082.51	0.00	911,692.46
0.05	1,387,107.37	0.05	848,934.98	0.05	617,156.74
0.10	1,131,690.67	0.10	656,968.56	0.10	452,516.88
0.15	878,050.04	0.15	491,378.69	0.15	324,848.40
0.20	653,412.19	0.20	358,253.07	0.20	231,134.95

Table 10. The link between minimum capital reserve (MCR), safety loading, and acceptable risk over a 5-year period

$\alpha = 0.01$		$\alpha = 0.05$		$\alpha = 0.10$	
ϑ	MCR	ϑ	MCR	ϑ	MCR
0.00	2,577,678.02	0.00	1,668,930.67	0.00	1,277,554.49
0.05	1,283,999.32	0.05	816,382.43	0.05	614,990.80
0.10	823,037.68	0.10	513,219.68	0.10	379,788.33
0.15	617,197.50	0.15	376,261.27	0.15	272,495.68
0.20	496,739.87	0.20	297,414.01	0.20	211,569.03

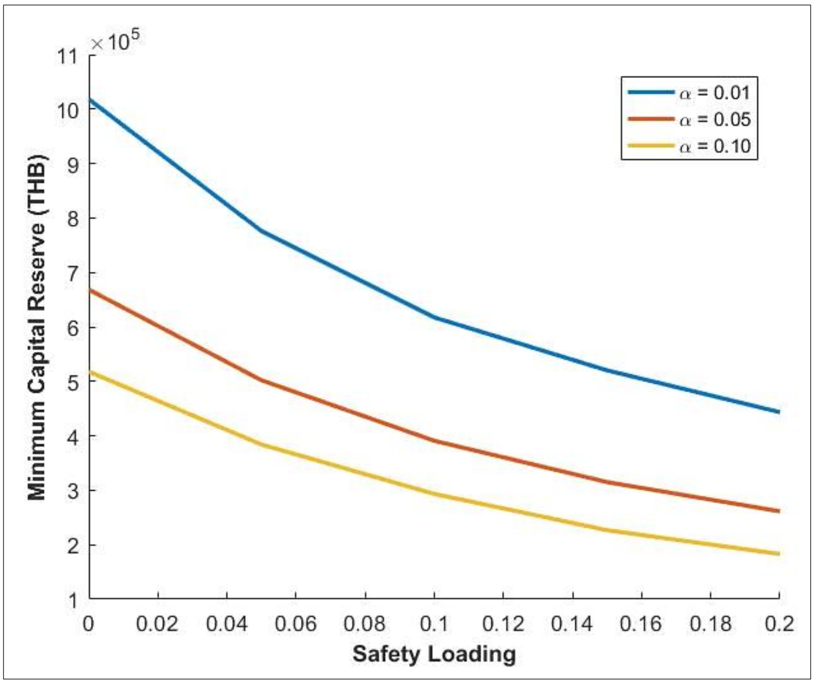

Figure 12. The MCR with the link between safety loading and ruin probabilities (α = 0.01, 0.05, 0.10) in 1 year

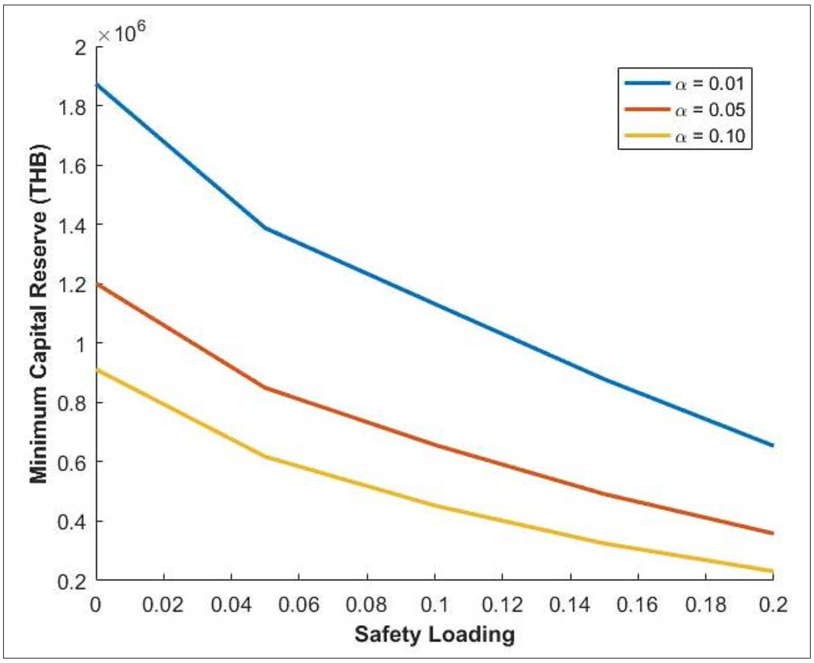

Figure 13. The MCR with the link between safety loading and ruin probabilities (α = 0.01, 0.05, 0.10) in 3 years

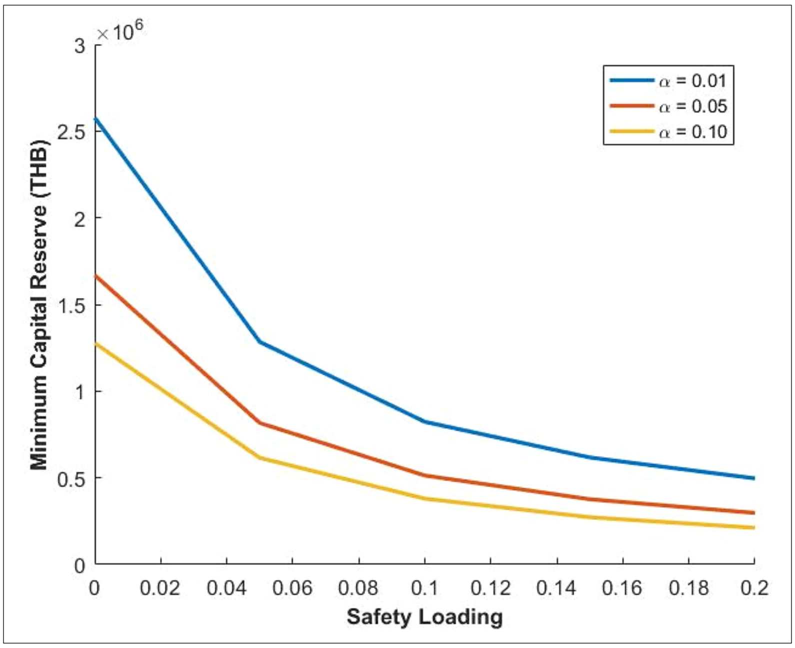

Figure 14. The MCR with the link between safety loading and ruin probabilities (α = 0.01, 0.05, 0.10) in 5 years

As safety loading (ϑ) increases, the minimum capital reserve (MCR) tends to decrease. This suggests that higher safety loading allows for a lower level of capital reserves while maintaining the desired level of confidence in avoiding financial ruin.

Different acceptable risk levels (α) have a significant impact on the required minimum capital reserves. As α becomes more stringent (e.g., α = 0.01), the MCR increases, indicating that a higher level of capital is necessary to maintain a very low risk of financial ruin. Conversely, at a higher α (e.g., α = 0.10), the MCR is lower, allowing for a more moderate risk tolerance.

These findings have significant implications for risk management decisions. Organizations can use this information to determine the minimum capital reserves needed to meet their risk tolerance and regulatory requirements, aiding in optimizing capital allocation to strike a balance between risk mitigation and capital efficiency. The analysis of the minimum capital reserve required to maintain ruin probability within acceptable limits (α levels) provides practical guidelines for insurers. The results show that higher safety loading allows for lower MCR, which means that insurers can maintain solvency with less capital if they adjust their premium rates appropriately. This finding is significant for insurers aiming to optimize their capital usage. It suggests that by carefully calibrating safety loading, insurers can free up capital for other investments or reduce the overall cost of maintaining solvency, while still adhering to regulatory and risk management requirements.

Insurers can leverage insights from this study to enhance strategic capital management. By understanding the link between safety loading and MCR, they can better balance risk and return, possibly increasing safety loading to free up capital for growth opportunities. This is particularly relevant in emerging markets like Thailand, where claim data may be more variable. Applying the Wang-PH transform in such contexts can lead to more accurate pricing and reserve setting, essential for maintaining solvency in volatile markets. The study underscores the role of advanced statistical techniques in bolstering financial stability and competitiveness.

9- Conclusion

This study demonstrates the limitations of non-transformed models in fitting an actual claim dataset, reinforcing the importance of using transformed models, particularly the Wang-PH transform, in insurance risk modeling. Among the transformations applied, the Lognormal distribution consistently provided the best fit, followed by the Inverse Gaussian, Log-logistic, Inverse Weibull, Weibull, and Gamma distributions. These findings indicate that transformed models offer a more accurate representation of claim data, which is critical for precise risk assessment and decision-making in the insurance industry.

The study also reveals valuable insights into the relationship between safety loading, minimum capital reserves, and finite-time ruin probabilities. As safety loading and initial capital reserves increase, the probability of ruin decreases, thereby enhancing the financial stability of insurance companies. The high R-squared values observed in the exponential regression model confirm its effectiveness in predicting ruin probabilities across varying levels of safety loading. This relationship between safety loading and capital reserves is instrumental in optimizing risk management strategies and capital allocation. Furthermore, the results highlight the critical role of maintaining an adequate minimum capital reserve, showing how adjustments in safety loading can impact reserve requirements under different risk tolerances and time horizons. These insights enable organizations to make informed decisions, ensuring compliance with regulatory requirements and promoting financial stability while efficiently utilizing their capital resources.

10- Declarations

10-1- Author Contributions

Conceptualization, A.M. and W.I.; methodology, A.M.; software, W.I.; validation, A.M. and W.I.; formal analysis, A.M.; investigation, W.I.; resources, A.M.; writing—original draft preparation, A.M. and W.I.; writing—review and editing, A.M. and W.I.; visualization, W.I.; project administration, A.M.; funding acquisition, A.M. All authors have read and agreed to the published version of the manuscript.

10-2- Data Availability Statement

The data presented in this study are available on request from the corresponding author.

10-3- Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

10-4- Institutional Review Board Statement

Not applicable.

10-5- Informed Consent Statement

Not applicable.

10-6- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

11- References

[1] Klugman, S. A., Panjer, H. H., & Willmot, G. E. (2013). Loss Models. Wiley Series in Probability and Statistics, John Wiley & Sons, Hoboken, United States. doi:10.1002/9781118787106.

[2] McNeil, A. J., Frey, R., & Embrechts, P. (2015). Quantitative risk management: concepts, techniques and tools-revised edition. Princeton University Press, Princeton, United States.

[3] Dhaene, J., Vanduffel, S., Goovaerts, M. J., Kaas, R., Tang, Q., & Vyncke, D. (2006). Risk measures and comonotonicity: A review. Stochastic Models, 22(4), 573–606. doi:10.1080/15326340600878016.

[4] Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. Insurance: Mathematics and Economics, 17(1), 43–54. doi:10.1016/0167-6687(95)00010-P.

[5] Wang, S. (1996). Premium Calculation by Transforming the Layer Premium Density. ASTIN Bulletin, 26(1), 71–92. doi:10.2143/ast.26.1.563234.

[6] Wang, S. S., Young, V. R., & Panjer, H. H. (1997). Axiomatic characterization of insurance prices. Insurance: Mathematics and Economics, 21(2), 173–183. doi:10.1016/S0167-6687(97)00031-0.

[7] Young, V. R. (1999). Discussion of Christofides' Conjecture Regarding Wang's Premium Principle. ASTIN Bulletin, 29(2), 191–195. doi:10.2143/ast.29.2.504610.

[8] Wang, S. S. (2000). A Class of Distortion Operators for Pricing Financial and Insurance Risks. The Journal of Risk and Insurance, 67(1), 15. doi:10.2307/253675.

[9] Calderín-Ojeda, E., Gómez-Déniz, E., & Vázquez-Polo, F. J. (2023). Conditional Tail Expectation and Premium Calculation under Asymmetric Loss. Axioms, 12(5), 496. doi:10.3390/axioms12050496.

[10] Psarrakos, G. (2024). Calculating premium principles from the mode of a unimodal weighted distribution. ASTIN Bulletin, 54(3), 791–803. doi:10.1017/asb.2024.18.

[11] Cramér, H. (1959). On the Mathematical Theory of Risk. Centraltryckeriet, Stockholm, Sweden.

[12] Chan, W. S., & Zhang, L. (2006). Direct Derivation of Finite-Time Ruin Probabilities in the Discrete Risk Model with Exponential or Geometric Claims. North American Actuarial Journal, 10(4), 269–279. doi:10.1080/10920277.2006.10597426.

[13] Sattayatham, P., Sangaroon, K., & Klongdee, W. (2013). Ruin probability-based initial capital of the discrete-time surplus process. Variance, 7(1), 74-81.

[14] Pavlova, K. P., & Willmot, G. E. (2004). The discrete stationary renewal risk model and the Gerber–Shiu discounted penalty function. Insurance: Mathematics and Economics, 35(2), 267–277. doi:10.1016/j.insmatheco.2004.04.006.

[15] Li, S. (2005). On a class of discrete time renewal risk models. Scandinavian Actuarial Journal, 2005(4), 241–260. doi:10.1080/03461230510009745.

[16] Li, S. (2005). Distributions of the surplus before ruin, the deficit at ruin and the claim causing ruin in a class of discrete time risk models. Scandinavian Actuarial Journal, 2005(4), 271–284. doi:10.1080/03461230510009808.

[17] Klongdee, W., Sattayatham, P., & Boonta, S. (2013). On approximating the ruin probability and the minimum initial capital of the finite-time risk process by separated claim technique of motor insurance. Far East Journal of Mathematical Sciences (FJMS), (IV), 597-604.

[18] Jing, H., Peng, J., Jiang, Z., & Bao, Q. (2022). Asymptotic estimates for finite-time ruin probability in a discrete-time risk model with dependence structures and CMC simulations. Communications in Statistics - Theory and Methods, 51(11), 3761–3786. doi:10.1080/03610926.2020.1801740.

[19] Xu, C., Wang, K., & Wu, X. (2024). The finite-time ruin probability of a risk model with stochastic return and subexponential claim sizes*. Communications in Statistics - Theory and Methods, 53(6), 2194–2204. doi:10.1080/03610926.2022.2122840.

[20] Yıldırım Külekci, B., Korn, R., & Selcuk-Kestel, A. S. (2024). Ruin probability for heavy-tailed and dependent losses under reinsurance strategies. Mathematics and Computers in Simulation, 226, 118–138. doi:10.1016/j.matcom.2024.06.018.

[21] Liu, X., Gao, Q., & Kan, L. (2024). Uniform Asymptotics for a Renewal Risk Model with a Random Number of Delayed Claims. Journal of Industrial and Management Optimization, 20(11), 3415–3431. doi:10.3934/jimo.2024059.

[22] Denisov, D., Gotthardt, N., Korshunov, D., & Wachtel, V. (2024). Probabilistic approach to risk processes with level-dependent premium rate. Insurance: Mathematics and Economics, 118, 142–156. doi:10.1016/j.insmatheco.2024.06.002.

[23] Li, M., Chen, Z., Cheng, D., & Zhou, J. (2024). Uniform asymptotics for finite-time ruin probabilities of a bidimensional compound risk model with stochastic returns. Statistics & Probability Letters, 207, 110013. doi:10.1016/j.spl.2023.110013.

[24] Xu, C., Wang, K., & Peng, J. (2024). Uniform Asymptotics for Finite-Time Ruin Probabilities of Risk Models with Non-Stationary Arrivals and Strongly Subexponential Claim Sizes. Wuhan University Journal of Natural Sciences, 29(1), 21–28. doi:10.1051/wujns/2024291021.

[25] Akaike, H. (1974). A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control, 19(6), 716–723. doi:10.1109/TAC.1974.1100705.

[26] Gamerman, D., & Lopes, H. F. (2006). Markov Chain Monte Carlo. Chapman and Hall/CRC, New York, United States. doi:10.1201/9781482296426.