Int. J. Appl. Math. Comput. Sci., 2023, Vol. 33, No. 4, 637–650
DOI: 10.34768/amcs-2023-0045

CLAIM MODELING AND INSURANCE PREMIUM PRICING UNDER A BONUS–MALUS SYSTEM IN MOTOR INSURANCE

WEENAKORN IEOSANURAK [a], BANPHATREE KHOMKHAM [b], ADISAK MOUMEESRI [c,*]

[a]Department of Mathematics, Faculty of Science
Khon Kaen University
Khon Kaen, 40002, Thailand
e-mail: weenie@kku.ac.th

[b]Department of Computer Science, Faculty of Science
Srinakharinwirot University
Bangkok, 10110, Thailand
e-mail: banphatree@g.swu.ac.th

[c]Department of Statistics, Faculty of Science
Silpakorn University
Nakhon Pathom, 73000, Thailand
e-mail: moumeesri_a@silpakorn.edu

Accurately modeling claims data and determining appropriate insurance premiums are vital responsibilities for non-life insurance firms. This article presents novel models for claims that offer improved precision in fitting claim data, both in terms of claim frequency and severity. Specifically, we suggest the Poisson-GaL distribution for claim frequency and the exponential-GaL distribution for claim severity. The traditional method of assigning automobile premiums based on a bonus-malus system relies solely on the number of claims made. However, this may lead to unfair outcomes when an insured individual with a minor severity claim is charged the same premium as someone with a severe claim. The second aim of this article is to propose a new model for calculating bonus-malus premiums. Our proposed model takes into account both the number and size of claims, which follow the Poisson-GaL distribution and the exponential-GaL distribution, respectively. To calculate the premiums, we employ the Bayesian approach. Real-world data are used in practical examples to illustrate how the proposed model can be implemented. The results of our analysis indicate that the proposed premium model effectively resolves the issue of overcharging. Moreover, the proposed model produces premiums that are more tailored to policyholders' claim histories, benefiting both the policyholders and the insurance companies. This advantage can contribute to the growth of the insurance industry and provide a competitive edge in the insurance market.

Keywords: bonus-malus system, claim severity, exponential-GaL distribution, motor insurance, number of claims, Poisson-GaL distribution.

1. Introduction

The bonus-malus system (BMS), also known as a no-claim discount or experience rating, is a widely adopted mechanism used in the field of car insurance. It is designed to reward policyholders for good driving behavior and penalize those who frequently make claims or exhibit risky driving habits. The BMS operates on the principle of adjusting insurance premiums based on an individual's claims history, thereby encouraging safer driving practices and providing financial incentives for maintaining a claim-free record.

The BMS has become a widely accepted and integral part of car insurance in many countries. Its implementation has contributed to promoting safer driving habits and reducing the number of claims. Additionally, it has incentivized policyholders to be more cautious on the road, leading to a decrease in accidents and ultimately resulting in lower costs for both insurers and insured

*Corresponding author

individuals.

The Bayesian method is applied in the BMS of car insurance to estimate parameters related to driving behavior and risk. It involves using prior distributions to incorporate prior beliefs about the parameters, likelihood functions to quantify the agreement between data and parameter values, and Bayes' theorem to obtain posterior distributions. These posterior distributions represent updated probabilities of parameters after incorporating observed data. The Bayesian method allows for personalized risk assessment, fair premium adjustments, and encourages safer driving behavior. It provides a framework to incorporate prior knowledge, update beliefs with observed data, and make informed decisions in the BMS of car insurance. The principles of Bayesian premium calculation have been extensively utilized in the literature for many years. A comprehensive exploration of these principles can be found in the book of Bühlmann (1970).

The Poisson distribution models the occurrence of rare events, such as accidents, over a fixed time period. Policyholders' claim frequencies are assumed to follow a Poisson distribution, where the parameter represents the average number of claims. In the BMS of motor insurance, claim frequency data are not usually modeled using a Poisson distribution. The Poisson distribution assumes independent and constant rate events, which may not apply to claim frequencies. Instead, more complex distributions like the negative binomial or mixed Poisson distributions are commonly used. These distributions account for overdispersion, where the variance exceeds the mean, commonly observed in claim frequency data. Several papers, such as the works of Tremblay (1992), Lemaire (1995), Walhin and Paris (1999), Bulbul and Baykal (2016), and Tzougas *et al.* (2019a), have proposed the use of mixed Poisson distributions with different prior distributions to model claim frequency. These studies have also presented models for calculating bonus-malus premiums that focus solely on the claim frequency aspect. They have contributed to the advancement of methods considering various prior distributions and mixed Poisson models in the computation of bonus-malus premiums based on claim frequency data. Furthermore, Willmot (1986) as well as Karlis and Xekalaki (2005) provided a comprehensive overview of the existing research and literature related to Poisson mixtures. This included a discussion of various methodologies, techniques, and applications involving mixed Poisson distributions.

Upon analyzing the research conducted by Moumeesri *et al*. (2020), it was discovered that previous assessments failed to differentiate between the payment of premiums for minor and major losses when policyholders file claims. In response to this gap, several researchers, such as Frangos and Vrontos (2001), Mert and Saykan (2005), Ni *et al*. (2014), Emad and Ali (2016), Tzougas *et al.* (2019b), Moumeesri *et al*. (2020), Pongsart *et al*. (2021), and Moumeesri and Pongsart (2022), have proposed the premium model based on the BMS. These researchers have explored diverse distributions for the frequency and severity components in order to develop an optimal Bayesian BMS framework.

Numerous studies have explored contemporary advancements in insurance mathematics, with a specific focus on diverse modeling techniques, including non-homogeneous Poisson processes, fuzzy methodologies, and heavy-tailed claim value distributions. These approaches play a pivotal role in risk assessment and management within the insurance industry. Tank and Tuncel (2015) explored survival probabilities based on non-homogeneous claim occurrences and discussed surplus distribution characteristics in a compound binomial model. Matsui and Rolski (2016) introduced a mixed Poisson cluster model that improves prediction accuracy by accounting for clustering behavior through non-homogeneous Poisson processes. Romaniuk (2017) adopted a fuzzy approach to analyze catastrophe bonds, demonstrating how uncertainties in the insurer's shares can impact risk estimation. Furthermore, there have been several studies on fuzzy approaches, including those conducted by Nowak and Romaniuk (2013), Grzegorzewski *et al*. (2020) as well as Grzegorzewski and Romaniuk (2022). Cojocaru (2017) contributed to ruin probability analysis by proposing a multivariate model using non-homogeneous periodic Poisson processes, with consideration for dependent heavy-tailed claims. These studies collectively emphasized the significance of innovative modeling techniques to capture the complexities of insurance risks and enhanced decision-making processes within the industry. Further integration of these methods promises to advance insurance mathematics and risk management strategies.

When considering the similarities and differences between the BMS and modeling approaches, it becomes evident that both modeling approaches and the BMS share the common goal of risk assessment and management. They employ predictive modeling and address uncertainty, with modeling approaches catering to diverse insurance domains through intricate mathematical analyses. In contrast, the BMS is specialized for promoting safe driving behaviors in motor insurance, relying on predefined rules and adhering to regulatory and market constraints. While modeling approaches offer broader applications, the BMS provides a targeted approach within the motor insurance sector, demonstrating a balance between complexity and practicality.

In a study conducted by Nedjar and Zeghdoudi (2016), a novel probability distribution called the Gamma Lindley distribution (GaL) was introduced. This distribution aims to offer enhanced flexibility when analyzing lifetime data. It is commonly assumed that mixing the Poisson distribution with the GaL distribution yields improved accuracy when modeling claim frequency data. Conversely, when it comes to modeling claim severity data, mixing the exponential distribution with the GaL distribution is deemed more appropriate.

In this study, we present the models for claim frequency and severity, along with a new model that utilizes Bayesian BMS to determine the suitable premium. To assess the claim frequency distribution, we mix the Poisson distribution with the GaL distribution. This mixing offers a better fit for the claim data set compared with using traditional distributions (such as the exponential and Lindley distributions) due to its thicker tail. Regarding the claim severity distribution, previous studies have shown that the exponential-Lindley distribution does not adequately fit the claim data. Therefore, we propose using the exponential-GaL distribution instead.

To derive the posterior structure functions for the claim frequency and claim severity distributions, we employ the Bayesian method. The mean values of these functions are then utilized to determine appropriate premiums for policyholders. Insurance companies can employ our proposed model to establish premiums that align with the actual driving behavior of policyholders. Additionally, this model has the potential to be extended to other types of insurance, such as agricultural crop insurance, to improve the performance of the insurance sector, generate economic benefits, increase government income, and promote national development.

The paper's structure is outlined as follows. Section 2 provides a detailed description of the methodology, which is divided into two subsections: the claim frequency distribution and the claim severity distribution. This section covers mixing distributions, the Bayesian method, and premium prediction. In Section 3, a practical application of the proposed bonus-malus model is demonstrated using real insurance claim data. This section is further divided into two parts. The first part discusses the bonus-malus premiums obtained from the claim frequency data, while the second part focuses on the combined data of claim frequency and claim severity components. Additionally, this section discusses parameter estimation and the goodness of fit test. Finally, Section 4 provides a summary of the findings.

2. Methodology

We make the assumption that the claim frequency and claim severity for each policyholder are treated as independent variables. In order to analyze these aspects individually, we segregate the claim frequency and claim severity distributions into distinct subsections.

2.1. Poisson-GaL model for frequency distribution. In automobile insurance, policyholders face varying levels of risk related to accident occurrence. This risk is quantified by a risk parameter, which plays a crucial role in determining insurance premiums and coverage. The risk parameter is treated as a random variable, influenced by individual characteristics and circumstances. By analyzing this parameter, insurance companies can assess the probability and potential severity of future claims, enabling them to offer appropriate coverage and fair premiums. To model claim frequency data accurately, we propose mixing the Poisson distribution with the GaL distribution (Nedjar and Zeghdoudi, 2016), which provides a better fit and thicker tails compared with the Poisson distribution alone.

2.1.1. Mixing distribution. Consider the scenario where the number of claims of each policyholder, represented as k, is assumed to follow a Poisson distribution with a parameter ϑ. The probability mass function (pmf) for this distribution can be expressed as

$$f(k|\vartheta) = \frac{e^{-\vartheta}\vartheta^k}{k!}, \quad k = 0, 1, 2, \ldots, \vartheta > 0.$$

The expected value, denoted as $E[K|\vartheta]$, of a Poisson random variable K is $E[K|\vartheta] = \vartheta$.

Each policyholder is assigned a constant that represents the expected inherent risk of their insurance coverage. This constant corresponds to the mean number of claims for each insured and is denoted as ϑ. We assume that the parameter ϑ follows a GaL distribution, characterized by the parameters θ and β. Consequently, the probability density function (pdf) of ϑ can be expressed in the following form:

$$\pi(\vartheta) = \frac{\theta^2\big[(\beta + \beta\theta - \theta)\vartheta + 1\big]e^{-\theta\vartheta}}{\beta(1+\theta)}, \quad \vartheta, \theta, \beta > 0.$$

The expected value of the random variable GaL is $E[\vartheta] = \frac{2\beta(1+\theta)-\theta}{\theta\beta(1+\theta)}$. The mixed Poisson distribution with a GaL distribution or unconditional distribution for Poisson-GaL is derived in the following manner:

$$\begin{aligned} f(k) &= \int_0^\infty f(k|\vartheta)\,\pi(\vartheta)\,\mathrm{d}\vartheta \\ &= \int_0^\infty \frac{e^{-\vartheta}\vartheta^k}{k!} \cdot \frac{\theta^2\big[(\beta + \beta\theta - \theta)\vartheta + 1\big]e^{-\theta\vartheta}}{\beta(1+\theta)}\,\mathrm{d}\vartheta \\ &= \frac{\theta^2\big[(\beta + \beta\theta - \theta)(k+1) + \theta + 1\big]}{\beta(\theta+1)^{k+3}}, \end{aligned} \tag{1}$$

where $\theta > 0, \beta > 0$ and $k = 0, 1, 2, \ldots$.

The pmf plots of the Poisson-GaL distribution, which illustrate the unconditional distribution of claim frequency

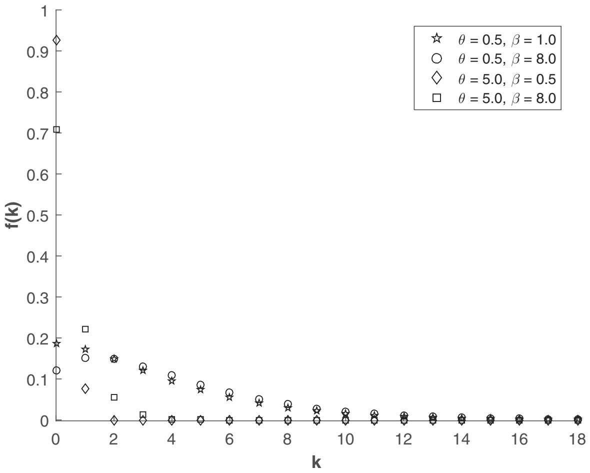

Fig. 1. Plots pmf of the Poisson-GaL distribution with various values of θ and β.

for all policyholders in the portfolio, are depicted in Fig. 1.

The mixed Poisson with GaL distribution, denoted as the Poisson-GaL distribution, can effectively handle zero-inflated data. The distribution's flexibility allows it to model situations where there is an excess of zeros compared with what would be expected in a standard Poisson distribution. This is particularly useful for scenarios where certain events are more likely to result in zero occurrences.

The Poisson-GaL distribution can accommodate zero-inflated effects by incorporating both a Poisson component (to model non-zero occurrences) and a GaL component (to account for the excess zeros). This combination allows for a more accurate representation of the data, making it a suitable choice for data sets with zero-inflated characteristics.

2.1.2. Bayesian methodology. The calculation of bonus-malus premiums has been extensively researched, leading to the adoption of various methods. Among these approaches, the Bayesian methodology emerges as one of the most commonly employed techniques. The primary goal of utilizing the Bayesian methodology in bonus-malus premium calculations is to establish the posterior distribution function. This approach proves effective when policyholder-specific data are available, such as historical claim records or policyholder profiles. By incorporating Bayesian techniques, insurers can gain a more comprehensive understanding of individual risk profiles and make more accurate assessments when determining appropriate bonus-malus premiums. This methodology holds particular value in insurance settings where personalized data plays a critical role in evaluating and classifying policyholders.

Consider a sample denoted as $k = (k_1, k_2, \dots, k_t)$, where t represents the specified time. Let N be the total number of claims made by a policyholder during the t-year period, given by $N = \sum_{i=1}^{t} k_i$, where k_i represents the number of claims made by the policyholder in year i ($i = 1, 2, \dots, t$). The likelihood function is defined as follows:

$$L(\vartheta; k_1, k_2, \dots, k_t) = \prod_{i=1}^{t} \frac{e^{-\vartheta}\vartheta^{k_i}}{k_i!} \propto e^{-\vartheta t}\vartheta^{N}.$$

The prior distribution is represented as

$$\pi(\vartheta) \propto \big[(\beta + \beta\theta - \theta)\vartheta + 1\big]e^{-\theta\vartheta}.$$

Applying Bayes' theorem, the posterior distribution function can be derived for an individual policyholder or a group of policyholders with claim history denoted as $k_1, k_2, \dots, k_t$. The posterior distribution function is directly proportional to the product of the prior distribution and the likelihood function. Therefore,

$$\pi^*(\vartheta | k_1, k_2, \dots, k_t) \propto e^{-(t+\theta)\vartheta}\big[(\beta + \beta\theta - \theta)\vartheta + 1\big]\vartheta^{N}.$$

Given

$$\int_0^\infty \pi^*(\vartheta | k_1, k_2, \dots, k_t)\,\mathrm{d}\vartheta \propto \int_0^\infty e^{-(t+\theta)\vartheta}\big[(\beta + \beta\theta - \theta)\vartheta + 1\big]\vartheta^{N}\,\mathrm{d}\vartheta,$$

we can deduce that

$$\int_0^\infty \pi^*(\vartheta | k_1, k_2, \dots, k_t)\,\mathrm{d}\vartheta \propto \int_0^\infty A e^{-(t+\theta)\vartheta}\big[(\beta + \beta\theta - \theta)\vartheta + 1\big]\vartheta^{N}\,\mathrm{d}\vartheta = 1$$

where A is a constant. Consequently,

$$A = \frac{(t+\theta)^{N+2}}{\Gamma(N+1)\big[(\beta + \beta\theta - \theta)(N+1) + t + \theta\big]}.$$

Hence, the posterior distribution function for the frequency component can be represented as follows:

$$\pi^*(\vartheta | k_1, k_2, \dots, k_t) = \frac{(t+\theta)^{N+2}e^{-(t+\theta)\vartheta}\big[(\beta + \beta\theta - \theta)\vartheta + 1\big]\vartheta^{N}}{\Gamma(N+1)\big[(\beta + \beta\theta - \theta)(N+1) + t + \theta\big]}. \quad (2)$$

2.1.3. Premium prediction. The net premium principle is indeed an important concept in the insurance industry. It involves determining the premium amount for an insurance policy based on the expected value of losses.

The net premium represents the expected number or mean of the number of claims that policyholders are likely to make. This article focuses on applying the net premium principle, which serves as a fundamental concept stating that premiums should align with the expected value of losses. In this context, the net premium represents the anticipated average or mean number of claims arising from each policyholder.

The expected number of claims for a policyholder with a claim history $k_1, k_2, \ldots, k_t$, or the mean of the posterior distribution function from Eqn. (2) for the Poisson-GaL distribution, can be calculated as follows:

$$\begin{aligned}\hat{\vartheta}_{t+1} &= E[\vartheta | k_1, k_2, \ldots, k_t] \\ &= \frac{(N+1)\big[(\beta+\beta\theta-\theta)(N+2)+t+\theta\big]}{(t+\theta)\big[(\beta+\beta\theta-\theta)(N+1)+t+\theta\big]}. \end{aligned} \quad (3)$$

Suppose that the initial premium or base premium at time $t = 0$ is set at 100. In that case, at time $t+1$, the premium rate can be calculated based on the number of claims and can be represented in the following manner:

$$\begin{aligned}\text{Premium}_{t+1} &= \frac{\theta\beta(1+\theta)(N+1)}{(t+\theta)(2\beta+2\beta\theta-\theta)} \\ &\quad \cdot \frac{\big[(\beta+\beta\theta-\theta)(N+2)+t+\theta\big]}{\big[(\beta+\beta\theta-\theta)(N+1)+t+\theta\big]} 100, \end{aligned} \quad (4)$$

where $\theta > 0, \beta > 0$ and $N = \sum_{i=1}^{t} k_i$. The premium rate is influenced by the parameters of the Poisson-GaL distribution (θ and β), the observation period covering t years for the policyholder, and the total count of claims (N).

2.2. Exponential-GaL model for severity distribution. In insurance portfolios, it is common to utilize heavy-tailed distributions for modeling claim severity. By combining these claim severity distributions with the prior distribution, the resulting mixed claim severity distribution exhibits thicker tails, thereby offering an improved fit to the actual claim severity data.

2.2.1. Mixing distribution. Within the existing body of literature, the mixed exponential distribution has gained significant usage for modeling severity distributions, as evidenced by Frangos and Vrontos (2001), Mert and Saykan (2005), and Tzougas *et al.* (2019b). In this particular article, we introduce a novel approach that involves utilizing a mixed exponential distribution alongside a GaL prior distribution for modeling severity distribution, as outlined below:

Consider a random variable X representing the claim size for each insured individual. Suppose that the claim size, denoted by x, follows an exponential distribution. The pdf for the exponential distribution with parameter λ is defined as

$$f(x|\lambda) = \lambda e^{-\lambda x}, \quad x > 0, \quad \lambda > 0.$$

The expected value of the random variable X is $E[X|\lambda] = 1/\lambda$.

When examining insurance policyholders, it is important to acknowledge that the rate parameter λ, which characterizes certain distributions, can differ among policyholders. To establish a prior distribution for λ, one proposed method is to employ the GaL distribution with parameters α and δ. The pdf of the GaL distribution is defined as

$$\pi(\lambda) = \frac{\alpha^2\big[(\delta+\delta\alpha-\alpha)\lambda+1\big]e^{-\alpha\lambda}}{\delta(1+\alpha)}, \qquad \lambda, \alpha, \delta > 0.$$

The expected value of the random variable Λ can be calculated as $E[\Lambda] = \frac{2\delta(1+\alpha)-\alpha}{\alpha\delta(1+\alpha)}$. The subsequent steps can be followed to derive the unconditional distribution or mixed exponential distribution with GaL distribution of the claim size x:

$$\begin{aligned} f(x) &= \int_0^\infty f(x|\lambda)\,\pi(\lambda)\,\mathrm{d}\lambda \\ &= \int_0^\infty \lambda e^{-\lambda x} \cdot \frac{\alpha^2\big[(\delta+\delta\alpha-\alpha)\lambda+1\big]e^{-\alpha\lambda}}{\delta(1+\alpha)}\,\mathrm{d}\lambda \\ &= \frac{\alpha^2(2\delta+2\delta\alpha-\alpha+x)}{\delta(1+\alpha)(x+\alpha)^3}, \end{aligned} \quad (5)$$

where $x > 0, \alpha > 0$ and $\delta > 0$. The cumulative distribution function (cdf) of the exponential-GaL distribution can be represented by

$$F(x) = \frac{\alpha^2}{\delta(\alpha+1)} \cdot \left(\frac{\delta\alpha+\delta}{\alpha^2} - \frac{x+\delta\alpha+\delta}{x^2+2\alpha x+\alpha^2}\right).$$

The pdf plots of the exponential-GaL distribution, representing the unconditional distribution of claim size x, are depicted in Fig. 2. These parameter values determine the characteristics of the distribution and can be estimated based on the specific application or data being analyzed.

2.2.2. Bayesian methodology. The total number of claims made by a policyholder during the t-year period can be represented as $N = \sum_{i=1}^{t} k_i$. Let x_k denote the size of claim k, where k takes values from 1 to N. Thus, the claim size history of the policyholder during the t-year period can be depicted as a vector $x = (x_1, x_2, \ldots, x_N)$. The total claim size incurred during the t-year period is denoted as $S = \sum_{k=1}^{N} x_k$. The likelihood function, which represents the probability distribution of the observed claim sizes, is given by

$$L(\lambda; x_1, x_2, \ldots, x_N) = \prod_{k=1}^{N} \lambda e^{-\lambda x_k} = \lambda^N e^{-\lambda S}.$$

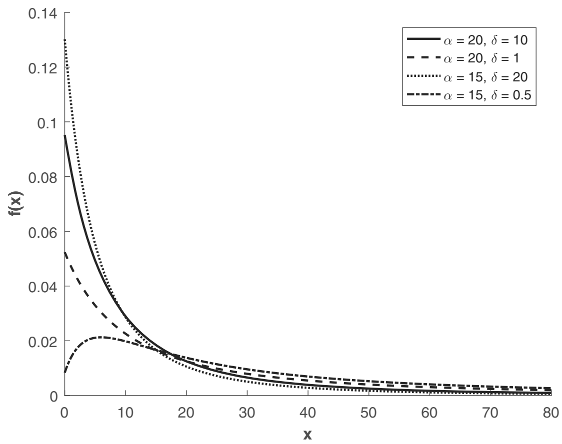

Fig. 2. Exponential-GaL distribution density plots with various values of α and δ.

The prior distribution is

$$\pi(\lambda) \propto [(\delta + \delta\alpha - \alpha)\lambda + 1]e^{-\alpha\lambda}.$$

Applying Bayes' theorem, the posterior distribution function is proportional to the product of the likelihood function and the prior distribution, given by:

$$\pi^*(\lambda|x_1, x_2, \ldots, x_N) \propto e^{-(S+\alpha)\lambda}\big[(\delta + \delta\alpha - \alpha)\lambda + 1\big]\lambda^N.$$

Considering the integral

$$\int_0^\infty \pi^*(\lambda|x_1, x_2, \ldots, x_N)\,\mathrm{d}\lambda \propto \int_0^\infty e^{-(S+\alpha)\lambda}\big[(\delta + \delta\alpha - \alpha)\lambda + 1\big]\lambda^N\,\mathrm{d}\lambda,$$

then we have

$$\int_0^\infty \pi^*(\lambda|x_1, x_2, \ldots, x_N)\,\mathrm{d}\lambda = \int_0^\infty Be^{-(S+\alpha)\lambda}\big[(\delta + \delta\alpha - \alpha)\lambda + 1\big]\lambda^N\,\mathrm{d}\lambda = 1,$$

where B is a constant. Solving for B, we find

$$B = \frac{(S+\alpha)^{N+2}}{\Gamma(N+1)\big[(\delta + \delta\alpha - \alpha)(N+1) + S + \alpha\big]}.$$

Hence, the posterior distribution function for the severity distribution can be expressed as

$$\pi^*(\lambda|x_1, x_2, \ldots, x_N) = \frac{(S+\alpha)^{N+2}e^{-(S+\alpha)\lambda}\big[(\delta + \delta\alpha - \alpha)\lambda + 1\big]\lambda^N}{\Gamma(N+1)\big[(\delta + \delta\alpha - \alpha)(N+1) + S + \alpha\big]}. \quad (6)$$

2.2.3. Premium prediction. To calculate premiums for the claim severity distribution, we employ the net premium principle, which mirrors the methodology used for the claim frequency distribution. The expected value in Eqn. (6) of the exponential-GaL distribution is determined through the following computation:

$$\begin{aligned}\hat{\lambda}_{t+1} &= E[\lambda|x_1, x_2, \ldots, x_N] \\ &= \frac{(N+1)\big[(\delta + \delta\alpha - \alpha)(N+2) + S + \alpha\big]}{(S+\alpha)\big[(\delta + \delta\alpha - \alpha)(N+1) + S + \alpha\big]}.\end{aligned}$$

The posterior distribution represents the estimated distribution of the risk parameter based on observed data. Technically, it serves as a conditional probability distribution of the risk parameter λ given the observed data x. Our primary objective is to calculate the expected value of the observed data (loss) given the risk parameter. In other words, it is the expected value of a future event (loss) given the preceding event (risk parameter), but we lack precise knowledge of the preceding event or risk parameter. To address this uncertainty, we employ a Bayesian method to identify the risk parameter, which corresponds to the mean of the posterior distribution function.

Recall that the expected value of the exponential random variable is $1/\lambda$, denoted as $E[X|\lambda] = 1/\lambda$. Since we lack the exact value of λ, we employ the Bayesian method, which is based on observed data, to determine the risk parameter (λ). The mean of the posterior distribution function from Eqn. (6) for the exponential-GaL distribution can be derived as $\hat{\lambda} = E[\lambda|x_1, x_2, \ldots, x_N]$. From $E[X|\lambda] = 1/\lambda$, the expression for $E[x_k|\lambda] = 1/\hat{\lambda}$ can be derived as

$$E[x_k|\lambda] = \frac{(S+\alpha)\big[(\delta + \delta\alpha - \alpha)(N+1) + S + \alpha\big]}{(N+1)\big[(\delta + \delta\alpha - \alpha)(N+2) + S + \alpha\big]}. \quad (7)$$

To maintain equity among policyholders in the portfolio, it is crucial for the premium paid by each policyholder to reflect the combined influence of the number and size of their claims. As the specific losses (x_k) arising from individual claims vary, policyholders with an equal number of claims will be subject to varying premiums. The Bayesian bonus-malus premium is designed to address this disparity and is determined by multiplying the Bayesian premium, which considers the frequency component in Eqn. (3), with the severity component outlined in Eqn. (7). This premium can be mathematically expressed as follows:

$$\begin{aligned}&\text{Premium}_{t+1}\\&=\frac{(N+1)\big[(\beta+\beta\theta-\theta)(N+2)+t+\theta\big]}{(t+\theta)\big[(\beta+\beta\theta-\theta)(N+1)+t+\theta\big]}\\&\cdot\frac{(S+\alpha)\big[(\delta+\delta\alpha-\alpha)(N+1)+S+\alpha\big]}{(N+1)\big[(\delta+\delta\alpha-\alpha)(N+2)+S+\alpha\big]},\end{aligned}\tag{8}$$

where $\theta > 0, \beta > 0, \alpha > 0, \delta > 0, N = \sum_{i=1}^{t} k_i$, and $S = \sum_{k=1}^{N} x_k$. By utilizing this formula, policyholders are charged premiums that are in proportion to the number and size of their claims, while additionally considering the overall risk level of the portfolio.

Equation (8) reveals that the premium depends on various factors, which encompass the parameters of the Poisson-GaL distribution (θ and β), the parameters of the exponential-GaL distribution (α and δ), the observation period spanning t years for the policyholder, along with the total number of claims (N) and the total amount (S) of the filed claims.

In order to determine the bonus-malus premiums for a policyholder using the proposed model, certain information is required, including the policyholder's claim count, policy age, and total claim amounts. These data points are typically available within the portfolio.

The initial premium, known as the Bayesian bonus-malus premium at time $t = 0$, represents the base premium that is initially charged to new policyholders upon their enrollment in the insurance policy. The formulation for the base premium is as follows:

$$\text{Premium}_1 = \frac{2\beta(1+\theta)-\theta}{\theta\beta(1+\theta)} \cdot \frac{\alpha\delta(1+\alpha)}{2\delta(1+\alpha)-\alpha}. \tag{9}$$

3. Numerical application

3.1. Claim data. In this paper, the calculation of model premiums involved the utilization of a specific data set. The data set used was comprised of one-year automobile insurance policies that were taken out either in 2004 or 2005. For access to the data set used in this study, interested individuals can visit the website of the Faculty of Business and Economics at Macquarie University in Sydney, Australia (De Jong and Heller, 2008).

The data set employed in this analysis encompassed a total of 67,856 policies within the portfolio. Among these policies, 4,624 policyholders had made at least one claim. Table 1 presents a breakdown of the claim frequency, showing that there were 4,333 policyholders with one claim, 271 with two claims, 18 with three claims, and 2 with four claims.

To provide insights into the claim severity within the data set, Fig. 3 presents a histogram representation. The severity data are displayed on a logarithmic scale for better visualization.

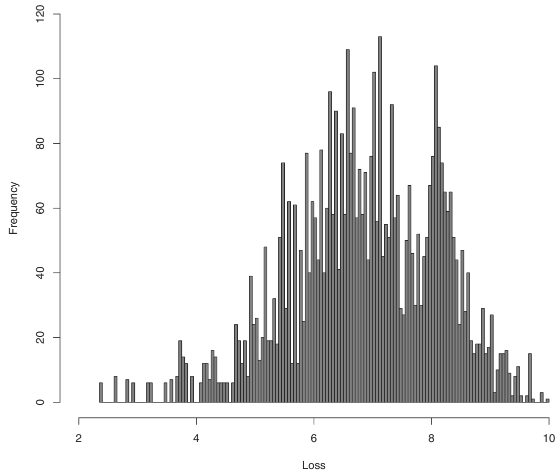

Fig. 3. Histogram displaying the claim severity data on a logarithmic scale.

3.2. Parameter estimation and model fitting. For the purpose of parameter estimation and goodness of fit testing, we analyze the claim frequency distribution and claim severity distribution separately.

3.2.1. Claim frequency distribution. The method of maximum likelihood estimation (MLE) is a commonly employed statistical technique for estimating model parameters. Its fundamental concept revolves around identifying parameter values that optimize the likelihood function, which represents the probability density of the observed data. The underlying principle of maximum likelihood is to estimate parameter values that maximize the likelihood of the observed data. The MLE methodology for the Poisson-GaL distribution is outlined as follows:

Consider a random sample of size n, denoted as $k_1, k_2, \dots, k_n$, drawn from a Poisson-GaL distribution with the pmf given by Eqn. (1). To determine the most likely values of the parameters θ and β, we need to maximize the likelihood function L, defined as

$$L(\theta, \beta; k_i) = \prod_{i=1}^{n} \frac{\theta^2\big[(\beta+\beta\theta-\theta)(k_i+1)+\theta+1\big]}{\beta(\theta+1)^{k_i+3}}.$$

Next, we derive the log-likelihood function as follows:

$$\begin{aligned}&\ln L(\theta, \beta; k_i)\\&\quad= 2n\ln\theta - n\ln\beta - \ln(\theta+1)\sum_{i=1}^{n}(k_i+3)\\&\quad+\sum_{i=1}^{n}\ln\big[(\beta+\beta\theta-\theta)(k_i+1)+\theta+1\big].\end{aligned}$$

To obtain the estimators $\hat{\theta}$ and $\hat{\beta}$ for the parameters θ and β, we need to solve the following equations:

$$\frac{\partial}{\partial\theta}\ln L(\theta,\beta;k_i)=0 \quad \text{and} \quad \frac{\partial}{\partial\beta}\ln L(\theta,\beta;k_i)=0.$$

Due to the absence of a closed-form solution for estimating the parameters θ and β, the Newton–Raphson method is employed as a numerical iteration technique to solve for the solution. This method facilitates the estimation process by iteratively refining the parameter estimates until convergence is achieved.

In order to evaluate the suitability of the Poisson-GaL distribution for the frequency component, we conducted a Chi-Square goodness of fit test. This statistical test is utilized to assess whether the observed sample distribution significantly deviates from the expected probability distribution. To perform the test, the sample data are divided into intervals, and the number of data points falling within each interval is compared with the expected number of points in each interval. The Chi-Square goodness of fit test calculates the level of agreement between the observed and expected frequencies using the formula

$$\chi^2=\sum_{i=1}^{n}\frac{(O_i-E_i)^2}{E_i}.$$

Here, the value of the Chi-Square goodness of fit test, denoted as χ^2, is obtained by comparing the observed frequency count, O_i, with the expected frequency count, E_i, for each category or level of the variable i.

Moreover, we use the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) as statistical criteria for model selection, which are widely employed in statistics and data analysis to assist researchers and practitioners in choosing the most appropriate model from a set of candidate models. This selection is based on the models' fit to the data and the number of parameters utilized.

The AIC served as the criterion for selecting the appropriate model. The model with the lowest AIC value is deemed the best fit among all other models. The AIC estimation equation is expressed as follows:

$$\text{AIC}=-2\ln L+2m,$$

where L denotes the likelihood function value of the model, and m represents the number of estimated parameters in the model (Akaike, 1974). The BIC (Stone, 1979) is a device for selecting a model based on its balance between how well it fits the data and how complex it is. A model with a lower BIC value is considered to have a better fit. This is the equation utilized to approximate the BIC of a model:

$$\text{BIC}=-2\ln L+2m\ln N,$$

where N is the number of data.

In this paper, we utilize the Poisson-GaL distribution to analyze the distribution of claim frequencies. The maximum likelihood estimator is employed to estimate the parameters of the Poisson-GaL distribution, resulting in estimated values of $\hat{\theta}$ =18.5607 and $\hat{\beta}$ =1.4608. To assess the goodness of fit, we employ the Chi-Square test, which yields a test statistic value of χ^2 = 1.2140. Furthermore, we compare the observed claim frequency values with several other distributions, including the traditional Poisson distribution (PD), Poisson-exponential distribution (PED), Poisson–Lindley distribution (PLD) (Sankaran, 1970), and Poisson–Lindley (2P) distribution (PLD(2P)) (Shanker and Mishra, 2014). The findings indicate that the Poisson-GaL distribution (PGD) provides a superior fit to the claim frequency data compared with the other distributions. When we take into account the AIC and BIC values, which account for both goodness of fit and model complexity, it becomes evident that the PGD does not perform as well as anticipated. Additionally, we compare the expected claim frequencies, which are presented in Table 1.

3.2.2. Claim severity distribution. The following methodology outlines the process for the MLE of the exponential-GaL distribution.

Consider a vector $X=(x_1,x_2,\dots,x_n)^T$, representing a set of identically and independently distributed observations from the exponential-GaL distribution with the pdf described in Eqn. (5). To determine the parameter values α and δ that yield the best fit to the observed data in X, we aim to maximize the likelihood function L, defined as

$$L(\alpha,\delta;x_i)=\prod_{i=1}^{n}\frac{\alpha^2(2\delta+2\delta\alpha-\alpha+x_i)}{\delta(1+\alpha)(x_i+\alpha)^3}.$$

Subsequently, the log-likelihood function can be expressed as

$$\begin{aligned}\ln L(\alpha,\delta;x_i)&=2n\ln\alpha-n\ln\delta-n\ln(1+\alpha)\\&\quad+\sum_{i=1}^{n}\ln(2\delta+2\delta\alpha-\alpha+x_i)\\&\quad-3\sum_{i=1}^{n}\ln(x_i+\alpha).\end{aligned}$$

The estimators $\hat{\alpha}$ and $\hat{\delta}$ for the parameters α and δ, respectively, are obtained by equating the partial derivatives of the log-likelihood function with respect to α and δ to zero:

$$\frac{\partial}{\partial\alpha}\ln L(\alpha,\delta;x_i)=0 \quad \text{and} \quad \frac{\partial}{\partial\delta}\ln L(\alpha,\delta;x_i)=0.$$

Table 1. Observed frequency and expected frequency for estimated parameter values.

Number of claim	Observed frequency	Expected frequency PD	PED	PLD	PLD (2P)	PGD
0	63,232	63,094.32	63,253.84	63,252.68	63,233.26	63,233.47
1	4333	4590.55	4290.03	4292.03	4327.62	4327.24
2	271	167.00	290.96	290.30	277.04	277.18
3	18	4.05	19.73	19.58	17.01	17.03
4	2	0.07	1.34	1.32	1.01	1.02
5+	0	0.00	0.09	0.09	0.06	0.06
Total	67,856	67,856	67,856	67,856	67,856	67,856
Estimated parameters		$\hat{\theta}=0.0728$	$\hat{\theta}=13.7444$	$\hat{\theta}=14.6238$	$\hat{\alpha}=0.0987$	$\hat{\theta}=18.5607$
					$\hat{\beta}=18.5921$	$\hat{\beta}=1.4608$
	χ^2	204.0396	2.3840	2.2571	1.2171	1.2140
	AIC	36,205.00	36,102.89	36,102.75	36,103.49	36,103.49
	BIC	36,225.25	36,123.14	36,123.01	36,143.99	36,143.99

Table 2. Claim severity distributions: pdf and cdf.

Claim severity distributions	pdf		cdf
ED	$f(x;\lambda)=\lambda e^{-\lambda x}$,	$x>0,\lambda>0$	$F(x;\lambda)=1-e^{-\lambda x}$
ELD	$f(x;\delta)=\frac{\delta^2}{\delta+1}\cdot\frac{x+\delta+2}{(x+\delta)^3}$,	$x>0,\delta>0$	$F(x;\delta)=\frac{(\delta+1)x^2+(\delta+2)\delta x}{(\delta+1)(x+\delta)^2}$
ELD(2P)	$f(x;\alpha,\delta)=\frac{\delta^2}{\alpha\delta+1}\cdot\frac{\alpha x+\alpha\delta+2}{(x+\delta)^3}$,	$x>0,\alpha>0,\delta>0$	$F(x;\alpha,\delta)=\frac{(\alpha\delta+1)x^2+(\alpha\delta+2)\delta x}{(\alpha\delta+1)(x+\delta)^2}$
EGD	$f(x;\alpha,\delta)=\frac{\alpha^2(2\delta+2\delta\alpha-\alpha+x)}{\delta(1+\alpha)(x+\alpha)^3}$,	$x>0,\alpha>0,\delta>0$	$F(x;\alpha,\delta)=\frac{\alpha^2}{\delta(\alpha+1)}\cdot\left(\frac{\delta\alpha+\delta}{\alpha^2}-\frac{x+\delta\alpha+\delta}{x^2+2\alpha x+\alpha^2}\right)$

Since the estimation of parameters α and δ does not have closed-form solutions, a numerical iteration technique is utilized to determine the estimates.

In this study, we utilize the Kolmogorov–Smirnov test (K-S test), the AIC and BIC to evaluate the performance and suitability of statistical models for claim severity distributions.

The K-S test is utilized as a goodness of fit test for claim severity distribution to assess whether a given distribution is suitable for a particular data set. The K-S test statistic is defined as

$$D=\max|F^n(x)-F(x)|,$$

where $F(x)$ represents the theoretical cumulative distribution of claim severity distributions and $F^n(x)$ is defined as

$$F^n(x)=\frac{1}{n}\left[\text{number of observations}\leq x\right],$$

with n being the sample size.

To compare the fit of different models, we have derived the formulas for the mixed distribution that results from mixing the exponential distribution with various prior distributions, including the Lindley (Lindley, 1958), Lindley (2P) (Shanker and Mishra, 2013), and GaL distributions (Nedjar and Zeghdoudi, 2016). The pdf and cdf of the resulting mixed exponential distribution are presented in Table 2.

Table 3 presents a comparative analysis of the estimated parameters, K-S statistic, AIC, and BIC for different claim severity distributions, including exponential (ED), exponential-Lindley (ELD), exponential-Lindley (2P) (ELD(2P)), and exponential-GaL (EGD). The K-S statistic, AIC, and BIC values of the EGD indicate the lowest among the other models. This implies that the proposed EGD model is the most suitable fit for the claim severity data compared to the other models.

The probability-probability (P-P) plots of the ED, ELD, ELD(2P), and EGD distributions are illustrated in Fig. 4. The results suggest that the exponential-GaL distribution provides the best fit for the claim severity data.

3.3. Insurance premium pricing. In this section, we calculate the bonus-malus premiums using two different approaches. Firstly, we focus solely on the claim frequency component. Secondly, we take into account both the claim frequency component and the claim severity component.

3.3.1. Claim frequency component. The Bayesian bonus-malus premiums, which solely rely on the frequency component, are determined, and calculated using Eqn. (4). The results of these calculations are

Table 3. Comparison of different claim severity distributions, namely, ED, ELD, ELD (2P), and EGD, considering the estimated parameters, K-S statistic, AIC, and BIC.

	Claim severity distributions			
	ED	ELD	ELD (2P)	EGD
Estimated parameters	$\hat{\lambda}$ =0.000533	$\hat{\delta}$ =977.053376	$\hat{\alpha}$ =16.875	$\hat{\alpha}$ =2404.941
			$\hat{\delta}$ =976.250	$\hat{\delta}$ =11079.447
K-S statistic	0.103389	0.087968	0.087986	0.043401
AIC	78,956.80	79,416.58	79,418.58	78,830.29
BIC	78,971.68	79,431.46	79,448.33	78,860.04

Table 4. Bonus-malus premiums based on the PGD for the frequency component.

t	Number of claims				
	0	1	2	3	4
0	100.00				
1	94.06	176.07	253.35	328.27	401.85
2	88.74	166.68	240.23	311.56	381.62
3	83.96	158.19	228.35	296.43	363.29
4	79.65	150.49	217.55	282.65	346.60
5	75.73	143.47	207.70	270.07	331.34
6	72.17	137.06	198.67	258.53	317.34
7	68.91	131.16	190.36	247.91	304.45

Table 5. Bonus-malus premiums based on the PED for the frequency component.

t	Number of claims				
	0	1	2	3	4
0	100.00				
1	93.22	186.44	279.65	372.87	466.09
2	87.30	174.59	261.89	349.19	436.49
3	82.08	164.17	246.25	328.33	410.42
4	77.46	154.92	232.37	309.83	387.29
5	73.33	146.65	219.98	293.30	366.63
6	69.61	139.22	208.83	278.45	348.06
7	66.26	132.51	198.77	265.02	331.28

presented in Table 4.

Based on the findings presented in Table 4, policyholders who do not file any claims in the first year are rewarded with a bonus equal to 5.94% of the base premium. Conversely, individuals who experience a single claim in the first year have to pay a malus corresponding to 76.07% of the base premium. Premium amounts vary based on the occurrence or absence of claims, with reductions observed after claim-free years and increases following claim occurrences.

In order to facilitate a comparative analysis, we also calculate the Bayesian bonus-malus premiums using the traditional Poisson-exponential model. The outcomes of this computation are provided in Table 5.

Based on the findings provided in Table 5, policyholders who do not file any claims in the first year receive a bonus equivalent to 6.78% of the base premium. Conversely, policyholders who have one claim during the first year are required to pay a malus corresponding to 86.44% of the base premium.

Based on our observations, it is evident that the Bayesian bonus-malus premiums calculated using the traditional PED model are stricter for policyholders categorized as bad drivers, in comparison with the proposed PGD model. Conversely, premiums derived from the PED model are more forgiving for policyholders classified as good drivers, in contrast to the proposed PGD model.

The proposed PGD model exhibits a reduced degree of penalization when compared with the traditional PED model. Therefore, it underscores the model's capacity to address the problem of overcharges.

3.3.2. Claim frequency and severity components. The proposed model is employed to illustrate the premiums, as depicted in Eqn. (8). Tables 6–8 presents various scenarios where the claim amount of the insured party varies. Specifically, we examine situations where the policyholder's total claim amount is $S = 1500$, 4500, and 10000, along with the corresponding results presented in Tables 6–8, respectively.

Tables 6–8 illustrate the premiums required over a span of seven years for different claim numbers. The premiums presented in these tables are calculated using Eqn. (8). These premiums are computed using the PGD model for the frequency component and the EGD model for the severity component. The initial base premium is set at 87.49 (calculated using (9)) and gradually decreases with each consecutive claim-free year.

For instance, if a policyholder submits a claim of 1500 in the first year, the corresponding premium indicated in Table 6 is 166.76. If an additional claim of 3000 is filed in the second year, resulting in a total claim amount of 4500 over the two-year period, the premium increases to 301.72, as shown in Table 7. However, if no claims are made in the third year, the premium decreases to 286.81, as displayed in Table 7, considering the cumulative claim amount of 4500 over three years.

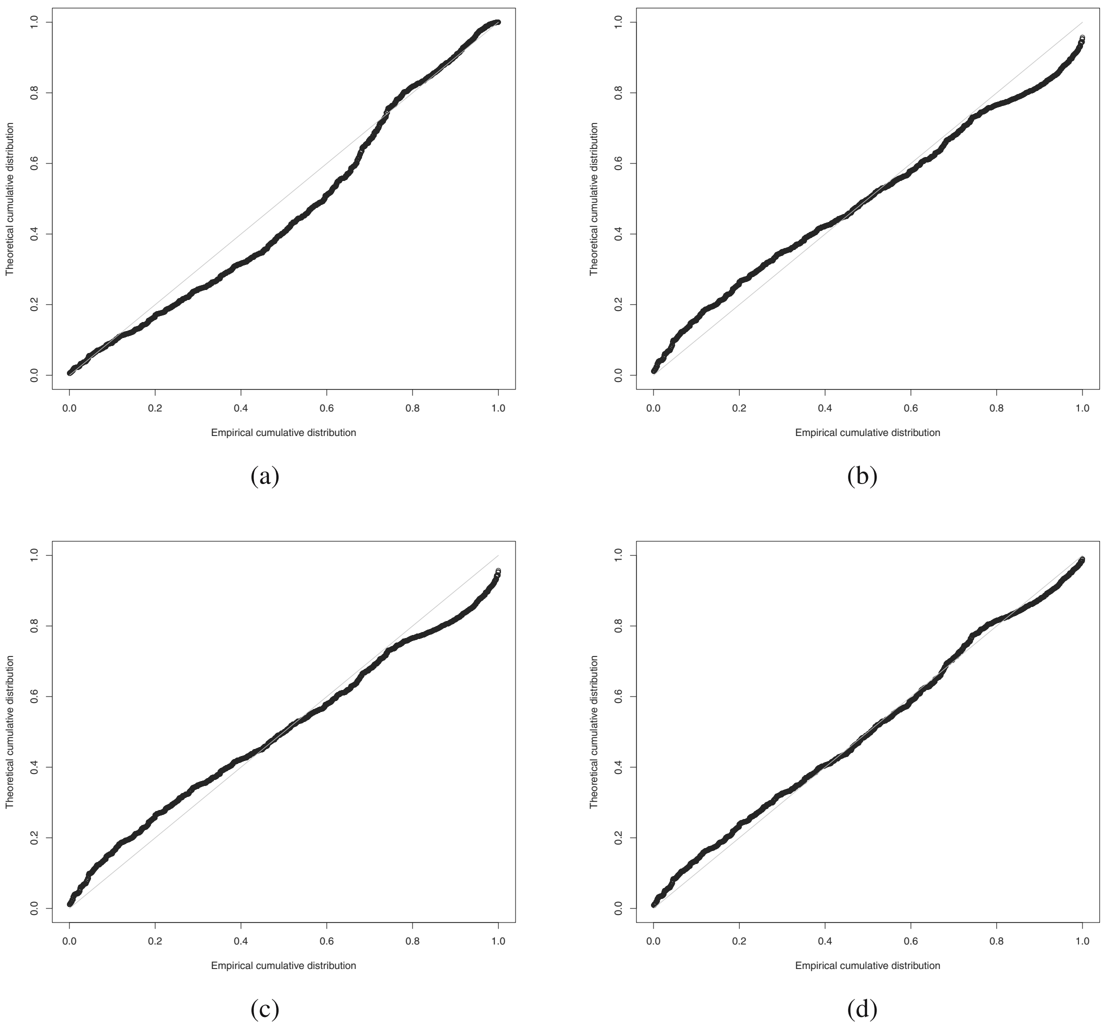

Fig. 4. P-P plots of ED (a), ELD (b), ELD(2P) (c), and EGD (d).

In the case of a 5500 claim in the fourth year, an additional surcharge is applied, raising the premium to 510.24. This premium accounts for three accidents and a total claim amount of 10,000 over four years, as demonstrated in Table 8. It is worth noting that as the number of claims increases, the premium escalates proportionately for the same cumulative claim amount.

Under the same situation, if a policyholder submits two claims with a total claim amount of 1500 in the first year, the corresponding premium indicated in Table 6 is 179.95. However, if no claims are made in the second year, the premium decreases to 170.63, as displayed in Table 6, considering the cumulative claim amount of 1500 over two years.

In the case of a 3000 claim in the third year, an additional surcharge is applied, raising the premium to 297.85. This premium accounts for three accidents and a total claim amount of 4,500 over three years, as demonstrated in Table 7.

In the case of a 5500 claim in the fourth year, an additional surcharge is applied, raising the premium to 521.39. This premium accounts for four accidents and a total claim amount of 10,000 over four years, as demonstrated in Table 8.

For better comprehension, Tables 9 and 10 provide a summary of the bonus-malus premiums for the situations in the first and second cases, respectively, considering varying total claim amounts derived from the premiums in Tables 6–8.

Table 11 showcases the premiums (malus) that policyholders are required to pay in the first year when a claim occurs, utilizing the PGD for claim frequency and the EGD for claim severity. The table includes data encompassing accidents with frequencies ranging from 1 to 4 and corresponding total claim amounts varying from AUD 300 to AUD 50,000. The graphical representation

Table 6. Bonus-malus premiums using the PGD for frequency and EGD for severity, with a total claim amount of $S = 1500$.

t	Number of claims (N)				
	0	1	2	3	4
0	87.49				
1	82.29	**166.76**	179.95	186.53	466.09
2	77.64	157.86	170.63	177.04	436.49
3	73.46	149.82	162.20	168.44	410.42
4	69.69	142.53	154.53	160.61	387.29
5	66.26	135.88	147.53	153.46	366.63
6	63.14	129.80	141.11	146.91	348.06
7	60.29	124.22	135.22	140.87	331.28

Table 7. Bonus-malus premiums using the PGD for frequency and EGD for severity, with a total claim amount of $S = 4500$.

t	Number of claims (N)				
	0	1	2	3	4
0	87.49				
1	82.29	294.87	318.21	329.84	336.48
2	77.64	279.14	**301.72**	313.05	319.54
3	73.46	264.92	**286.81**	297.85	304.19
4	69.69	252.03	273.25	284.01	290.22
5	66.26	240.28	260.87	271.37	277.44
6	63.14	229.53	249.53	259.77	265.72
7	60.29	219.66	239.10	249.10	254.92

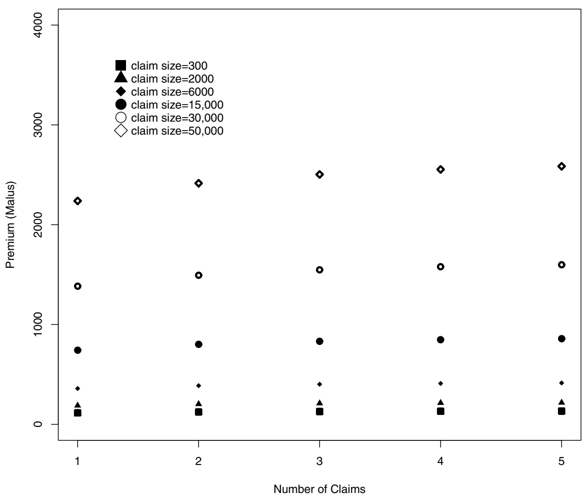

Fig. 5. Premiums based on PGD and EGD with a various number of claims and different claim sizes in the first year.

of these results can be observed in Fig. 5.

As the number of claims increases, there is a corresponding upward trend in the premiums. Similarly, an increase in the claim amount also leads to higher premiums. The variation in premiums can be observed by considering both the total claim amount and the number of claims filed.

4. Conclusions

A new and innovative model is introduced to determine bonus-malus premiums, which takes into consideration both claim frequency and claim severity aspects. The model utilizes two mixing distributions: the GaL distribution combined with the Poisson distribution for claim frequency, and the GaL distribution combined with the mixed exponential distribution for claim severity. The Bayesian methodology is employed to calculate the premiums.

In order to demonstrate the effectiveness of the model, actual automobile insurance data are used as an illustrative example. The results obtained from fitting the claim frequency and severity surpass those achieved by traditional models. Additionally, traditional models often imposed severe penalties on bad drivers when claims occurred, potentially causing issues for insurers if policyholders switched companies the following year due to high premiums. The findings indicate that the proposed model effectively addresses the problem of overcharging.

The proposed model offers substantial benefits to insurance practitioners, impacting insurers, policyholders, and the industry overall. It enhances risk assessment and pricing accuracy, supports profitability and financial stability, fosters customer loyalty, and positively influences driving behavior. Additionally, it provides a competitive edge and generates data-driven insights, aligns with regulatory standards, and showcases a commitment to road safety.

Acknowledgment

The authors would like to extend sincere appreciation to the editor and the anonymous referees for their valuable and constructive feedback.

References

Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control* **19**(6): 716–723.

Bulbul, S. and Baykal, K. (2016). Optimal bonus malus system design in motor third party liability insurance in Turkey: Negative binomial model, *International Journal of Economics and Finance* **8**(8): 205–211.

Bühlmann, H. (1970). *Mathematical Methods in Risk Theory*, Springer, New York.

Cojocaru, I. (2017). Ruin probabilities in multivariate risk models with periodic common shock, *Scandinavian Actuarial Journal* **2017**(2): 159–174.

Table 8. Bonus-malus premiums using the PGD for frequency and EGD for severity, with a total claim amount of S =10,000.

t	Number of claims (N)				
	0	1	2	3	4
0	87.49				
1	82.29	529.77	571.68	592.57	604.49
2	77.64	501.49	542.07	562.42	574.07
3	73.46	475.96	515.27	535.10	546.49
4	69.69	452.79	490.91	**510.24**	521.39
5	66.26	431.68	468.67	487.52	498.44
6	63.14	412.37	448.29	466.69	477.38
7	60.29	394.64	429.55	447.51	457.98

Table 9. Summary of the premium for the first case situation with varying total claim amounts.

t	Number of claims (N)				Total claim amount (S)
	0	1	2	3	
0	87.49				
1		166.76			1500
2			301.72		4500
3			286.81		4500
4				510.24	10000

Table 10. Summary of the premium for the second case's situation with varying total claim amounts.

t	Number of claims (N)				Total claim amount (S)
	0	1	2	3	
0	87.49				
1		179.95			1500
2		170.63			4500
3			297.85		4500
4				521.39	10000

Table 11. Analysis of premiums for the different number of claims and varying sizes of claims within the first year.

Total claim amount (S)	Number of claims (N) 1	2	3	4
300	115.51	124.65	129.21	131.81
2000	188.11	202.99	210.42	214.65
6000	358.93	387.33	401.49	409.57
15,000	743.32	802.11	831.43	848.15
30,000	1384.06	1493.46	1548.02	1579.13
50,000	2238.57	2415.36	2503.53	2553.82

De Jong, P. and Heller, G. (2008). *Generalized Linear Models for Insurance Data*, Cambridge University Press, New York.

Emad, A. and Ali, I. (2016). Bayesian approach for bonus-malus systems with Gamma distributed claim severities in vehicles insurance, *British Journal of Economics, Management, and Trade* **14**(1): 1–9.

Frangos, N. and Vrontos, S. (2001). Design of optimal bonus-malus systems with a frequency and a severity component on an individual basis in automobile insurance, *Astin Bulletin* **31**(1): 1–22.

Grzegorzewski, P., Hryniewicz, O. and Romaniuk, M. (2020). Flexible resampling for fuzzy data, *International Journal of Applied Mathematics and Computer Science* **30**(2): 281–297, DOI: 10.34768/amcs-2020-0022.

Grzegorzewski, P. and Romaniuk, M. (2022). Bootstrap methods for epistemic fuzzy data, *International Journal of Applied Mathematics and Computer Science* **32**(2): 285–297, DOI: 10.34768/amcs-2022-0021.

Karlis, D. and Xekalaki, E. (2005). Mixed Poisson distributions, *International Statistical Review* **73**(1): 35–58.

Lemaire, J. (1995). Bonus-malus systems in automobile insurance, *Insurance: Mathematics and Economics* **3**(16): 227.

Lindley, D. (1958). Fiducial distributions and Bayes' theorem, *Journal of the Royal Statistical Society B* **20**(1): 102–107.

Matsui, M. and Rolski, T. (2016). Prediction in a mixed Poisson cluster model, *Stochastic Models* **32**(3): 460–480.

Mert, M. and Saykan, Y. (2005). On a bonus-malus system where the claim frequency distribution is geometric and the claim severity distribution is Pareto, *Hacettepe Journal of Mathematics and Statistics* **34**(1): 75–81.

Moumeesri, A., Klongdee, W. and Pongsart, T. (2020). Bayesian bonus-malus premium with Poisson–Lindley distributed claim frequency and lognormal-Gamma distributed claim severity in automobile insurance, *WSEAS Transactions on Mathematics* **19**(46): 443–451.

Moumeesri, A. and Pongsart, T. (2022). Bonus-malus premiums based on claim frequency and the size of claims, *Risks* **10**(9): 181.

Nedjar, S. and Zeghdoudi, H. (2016). On Gamma Lindley distribution: Properties and simulations, *Journal of Computational and Applied Mathematics* **298**: 167–174.

Ni, W., Constantinescu, C. and Pantelous, A. (2014). Bonus-malus systems with Weibull distributed claim severities, *Annals of Actuarial Science* **8**(2): 217–233.

Nowak, P. and Romaniuk, M. (2013). A fuzzy approach to option pricing in a Levy process setting, *International Journal of Applied Mathematics and Computer Science* **23**(3): 613–622, DOI: 10.2478/amcs-2013-0046.

Pongsart, T., Moumeesri, A., Mayureesawan, T. and Phaphan, W. (2021). Computing Bayesian bonus-malus premium distinguishing among different multiple types of claims, *Lobachevskii Journal of Mathematics* **42**(13): 3208–3217.

Romaniuk, M. (2017). Analysis of the insurance portfolio with an embedded catastrophe bond in a case of uncertain parameter of the insurer's share, *in* A. Grzech *et al.* (Eds), *Information Systems Architecture and Technology: Proceedings of 37th International Conference on Information Systems Architecture and Technology, ISAT 2016. Part IV: Advances in Intelligent Systems and Computing*, Springer, Berlin/Heidelberg, pp. 33–43.

Sankaran, M. (1970). The discrete Poisson–Lindley distribution, *Biometrics* **26**(1): 145–149.

Shanker, R. and Mishra, A. (2013). A two-parameter Lindley distribution, *Statistics in Transition—New Series* **14**(1): 45–56.

Shanker, R. and Mishra, A. (2014). A two-parameter Poisson–Lindley distribution, *International Journal of Statistics and Systems* **9**(1): 79–85.

Stone, M. (1979). Comments on model selection criteria of Akaike and Schwartz, *Journal of the Royal Statistical Society B* **41**(2): 276–278.

Tank, F. and Tuncel, A. (2015). Some results on the extreme distributions of surplus process with nonhomogeneous claim occurrences, *Hacettepe Journal of Mathematics and Statistics* **44**(2): 475–484.

Tremblay, L. (1992). Using the Poisson inverse Gaussian in bonus-malus systems, *Astin Bulletin* **22**(1): 97–106.

Tzougas, G., Hoon, W. and Lim, J. (2019a). The negative binomial-inverse Gaussian regression model with an application to insurance ratemaking, *European Actuarial Journal* **9**(1): 323–344.

Tzougas, G., Yik, W. and Mustaqeem, M. (2019b). Insurance ratemaking using the exponential-lognormal regression model, *Annals of Actuarial Science* **14**(1): 42–71.

Walhin, J. and Paris, J. (1999). Using mixed Poisson processes in connection with bonus-malus systems, *ASTIN Bulletin* **29**(1): 81–99.

Willmot, G. (1986). Mixed compound Poisson distributions, *ASTIN Bulletin* **16**(1): 59–79.

Weenakorn Ieosanurak received her MSc and PhD degrees in applied mathematics from Khon Kaen University in 2015 and 2018, respectively. In 2020 she became a lecturer and in 2022 an assistant professor at the same university. Her research interests encompass the fields of image processing and probability theory, with a particular focus on parameter estimation.

Banphatree Khomkham holds a Bachelor's degree in mathematics from Khon Kaen University and a Master's degree in applied mathematics and computational science from the Department of Mathematics and Computer Science at Chulalongkorn University, Bangkok, Thailand. She also holds a PhD degree in computer science and information technology from the same department at Chulalongkorn University. Currently, she serves as a lecturer in the Department of Computer Science within the Faculty of Science at Srinakharinwirot University. Her primary research interests encompass image processing, pattern recognition, and machine learning.

Adisak Moumeesri received his BSc degree (Hons.) in mathematics in 2014, his MSc degree in applied mathematics in 2017, and his PhD degree in applied mathematics in 2021, all from Khon Kaen University, Thailand. His research interests include theory of distributions, probability theory, risk and insurance, parameter estimations, and credibility in insurance.

Received: 10 June 2023
Revised: 23 August 2023
Re-revised: 5 September 2023
Accepted: 11 September 2023