



# Study on Bayesian Skew-Normal Linear Mixed Model and Its Application in Fire Insurance

Meiling Gong, Zhanli Mao \*, Di Zhang, Jianxing Ren and Songtao Zuo,  
School of Fire Protection Engineering, China People's Police University,  
Langfang, China

**Received:** 22 September 2022/**Accepted:** 20 May 2023/**Published online:** 7 June 2023

**Abstract.** Fire insurance is a crucial component of property insurance, and its rating depends on the forecast of insurance loss claim data. Fire insurance loss claim data have complicated characteristics such as skewness and heavy tail. The traditional linear mixed model is commonly difficult to accurately describe the distribution of loss. Therefore, it is crucial to establish a scientific and reasonable distribution model of fire insurance loss claim data. In this study, the random effects and random errors in the linear mixed model are firstly assumed to obey the skew-normal distribution. Then, a skew-normal linear mixed model is established using the Bayesian MCMC method based on a set of U.S. property insurance loss claims data. Comparative analysis is conducted with the linear mixed model of logarithmic transformation. Afterward, a Bayesian skew-normal linear mixed model for Chinese fire insurance loss claims data is designed. The posterior distribution of claim data parameters and related parameter estimation are employed with the R language JAGS package to obtain the predicted and simulated loss claim values. Finally, the optimization model in this study is used to determine the insurance rate. The results demonstrate that the model established by the Bayesian MCMC method can overcome data skewness, and the fitting and correlation with the sample data are better than the log-normal linear mixed model. Hence, it can be concluded that the distribution model proposed in this paper is reasonable for describing insurance claims. This study innovates a new approach for calculating the insurance premium rate and expands the application of the Bayesian method in the fire insurance field.

**Keywords:** Fire insurance loss claim, Skew-normal distribution model, Optimization, Bayesian MCMC method

## 1. Introduction

The impact of fire on the social environment and people's lives should not be underestimated, and major fire accidents occur frequently in the world, resulting in severe casualties and property losses. Property insurance should be established for compensating economic losses in fire accidents. Therefore, it is imperative to

---

\*Correspondence should be addressed to: Zhanli Mao, E-mail: [maozhl@mail.ustc.edu.cn](mailto:maozhl@mail.ustc.edu.cn)



determine insurance rates scientifically and reasonably and forecast fire insurance loss claim for insurance rate determination.

It is particularly complicated for setting insurance rates due to the heterogeneity of insurance portfolios and different insurance risks in the fire insurance field. Hence, a linear mixed model (LMM) is typically employed to handle this heterogeneity. Random effects and random errors are commonly assumed to follow a normal distribution in the typical linear mixed model framework Zhang and Davidian [1], Ghidey and Lesaffre [2], Frees and Young [3]. The typical distribution of fire insurance loss claims is characterized by asymmetry and a thick tail. Some researchers have changed the distribution to eliminate the complicated characteristics of insurance. For example, the log-phase type (Log PH) distribution was introduced to fit the heavy-tail distribution of fire losses in Ahn and Kim [4] study. Huang and Meng [5] corrected the skewness and heavy tail problem of insurance loss using a Bayesian nonparametric regression model based on Gaussian distribution. Escobar and Pflug [6] determined the worst case of insurance distribution with the distance variance of the probability model. These methods can produce more appropriate empirical results. Nevertheless, data transformation may reduce the information provided by the original data's potential generation mechanism, and only applies to a single individual and does not guarantee joint normality, presenting some limitations. Furthermore, the accuracy of a normal hypothesis is difficult to be verified, and accurate estimation of inter-subject variation cannot be guaranteed.

Consequently, the hypothesis method of skew distribution is gradually considered by many researchers. The linear mixed model could be expanded for a better understanding of asymmetric insurance loss claim data, and one plausible modification is to assume the random errors following a skewed distribution. Numerous researchers Azzalini and Capitanio [7], Capitanio and Azzalini [8]; Sahu and Dey [9]; Arellano-Valle and Bolfarine [10]; Huang et al. [11], Jung and Lee [12] have explored linear regression models with skewed distribution, systematically introduced the family of skew distribution, and established a method to obtain skew distribution by symmetric distribution.

Some researchers have applied a skew linear mixed model to medical research and the economic field Bandyopadhyay et al. [13], Huang and Chen [14], Leisen and Marin [15]. They fitted different experimental data and achieved more robust estimation results, while its application in insurance loss claims has caused little attention. In the fire insurance actuarial, the Bayesian method is mainly used for problems such as estimating loss distribution, adjusting rates, and corrections. De Simone and Piangerelli [16] applied Bayesian procedures to quantify the impact of COVID-19 epidemic cases in eight different countries and all regions of Italy on determining effective reproduction numbers. Meanwhile, a considerable number of researchers (Jara and Quintana [17]; Scollnik [18]; Bermúdez and Bermúdez [19]; Esmacili and Klüppelberg [20]; Moreno and Jiménez et al. [21]) investigated the application of the Markov Chain Monte Carlo (MCMC) Bayesian method to estimate and fit insurance losses through different regression models. The corresponding studies mentioned above demonstrate the potential of the Bayesian method in insurance actuarial.

This paper aims to investigate the use of a skew linear mixed model in the prediction of fire insurance loss claims and propose a Bayesian skew-normal linear mixed model to describe the relationship between covariates and response variables flexibly. The random errors and random effects are assumed to obey skew-normal distribution in comparison to the traditional linear mixed model for eliminating the influence of the loss distribution's skewed characteristics. The MCMC method makes Bayesian inference a viable alternative for this prediction model. Regarding parameter estimation, the Bayesian method and the MCMC method apply to the prior distribution and the posterior distribution, respectively. The parameter posterior means are Gibbs sampling using the JAGS package in R language. This approach can improve the flexibility of the Bayesian method.

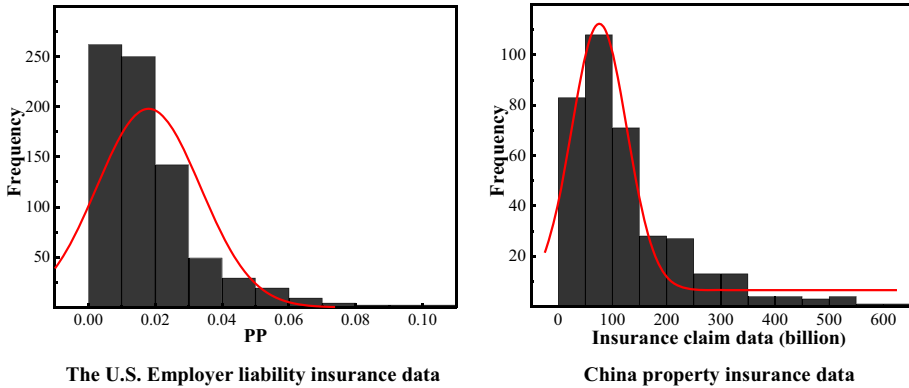
Linear mixed models (LMM) are often used to analyze repeated measurement data because they are more flexible in modeling intra-subject correlations, which are usually present in such data. The most popular continuous response LMM assumes that both the random effects and the internal errors of the subjects are normally distributed, which may be an unrealistic assumption that conceals important features of changes within and between units (or groups). In this work, a skew normal linear mixed model (SNLMM) is proposed, which relaxes the normality assumption by using multivariate skew normal distribution, including normal distribution as a special case, and is suitable for the structure of repeated measurement data and clustering data analysis. The Bayesian skew-normal linear mixed model proposed in this study outperforms the log-linear mixed model and the regression model on two fire insurance real loss datasets, expanding the method's application in the fire insurance business.

## **2. Data Sources**

The correctness of the Bayesian skew-normal linear mixed model is verified on two different property insurance datasets. The first dataset is the U.S. employer liability insurance loss claims data from Klugman [22], which is originally sourced from the National Council on Compensation Insurance and contains a large sample and small categorical covariates. The second dataset consists of a significant number of continuous covariates by introducing mainland China property insurance claims data, sourced from the China Statistics Bureau and the China Insurance Regulatory Commission (CIRC) website (<http://www.stats.gov.cn/tjsj/ndsj/>).

Employer liability insurance, classified as property insurance, is mainly used to manage claims for permanent or partial incapacity of employees. A total of 121 risk categories are included in the longitudinal dataset, and 847 data are collected in seven observation years, including occupational categories, income, and compensation. The applicability of the model is verified by exploring the data characteristics of the pure premium (PP) (PP = claim amount/income) data features.

Property insurance in mainland China is majorly employed to handle fire losses claims. The property insurance loss claims data and the gross regional product (GRP) dataset of various regions for 2011–2020 in mainland China are cited in this study. The panel data contains 340 data items for 34 regional categories over 10 observation years and are adopted in the study owing to the strong regional



**Figure 1. Frequency distribution histogram of two different datasets.**

characteristics of property insurance business size and claims. The model is applied for this dataset to investigate and estimate the association between property premium loss claims and GRP.

Figure 1 illustrates the histograms of frequency distributions of pure premiums of employer liability insurance and property insurance loss claims. As observed in Figure 1, both insurance datasets exhibit remarkable right skew and heavy tail characteristics. Thus, it is difficult for them to accurately apply to the traditional actuarial model. Fortunately, the application of the skewness distribution proposed in this study can better tackle this problem.

### 3. Methodology

#### 3.1. Bayesian Skew-Normal Linear Mixed Model

The skew-normal distribution is a further extension of the normal distribution. The Bayesian method is primarily applied to parameter estimation of the loss distribution. The traditional linear mixed model is generally insufficient to describe the insurance loss claim due to the complicated characteristics of skewness, heavy tail, and multimodality. As a result, the mature application of the Bayesian method and skewness distribution is required. Bayesian skew-normal linear mixed model (SNLMM) can better handle this type of panel data. Sahu et al. [9], Arellano-Valle and Genton [10], Arellano-Valle et al. [23], Bernardi and Petrella [26], Miljkovic and Grün [27], and other researchers have conducted correlation studies on the Bayesian skew-normal linear mixed model. In this section, the expression function studied by the previous researchers is introduced to define the Bayesian skew-normal linear mixed model (SNLMM).

*3.1.1. Skew-Normal Linear Mixed Model* Firstly, multivariate skew-normal distribution is introduced to define the proposed SNLMM. According to the multivariate skew-normal distribution studied by Arellano-Valle and Genton [10], the

density function of  $e$ -element skew-normal distribution of random variable  $X = (X_1, \dots, X_e)^T$  at point  $x$  is defined as

$$f_{SN}(x|\mu, \Sigma, \alpha) = 2\psi_e(x - \mu; \Sigma)\Psi(\alpha^T \sigma^{-1}(x - \mu)) \tag{1}$$

where  $\Sigma$  denotes the positive definite covariance matrix,  $\sigma$  represents the diagonal matrix  $\Sigma = \sigma \bar{\Sigma} \sigma$  formed by the standard deviation in  $\Sigma$ , and  $\bar{\Sigma}$  indicates the correlation matrix. The density function of  $e$ -element normal distribution  $N_e(\mu, \Sigma)$  at point  $x$  is expressed as  $\psi_e(x; \mu; \Sigma)$ , and the corresponding distribution function is expressed as  $\Psi_e(x; \mu; \Sigma)$ .

The most commonly used linear mixed model (LMM) was proposed by Laird and Ware [28] for a continuous response. Assuming that the loss data come from  $m$  different individual risks, there are  $k_i$  ( $i = 1, 2, \dots, m$ ) observations at different stages for each risk. Additionally,  $Y_i$  represents the  $k_i$  dimension vector of the continuous measurement of individual  $i$ . Then, the linear mixed model of  $Y_i$  is expressed as

$$Y_i = X_i\beta + Z_i b_i + \varepsilon_i, i = 1, 2, \dots, m \tag{2}$$

where  $X_i$  ( $k_i \times p$ ) indicates the explanatory variable matrix corresponding to the fixed effect vector  $\beta$ , and  $Z_i$  ( $k_i \times q$ ) denotes the explanatory variable matrix corresponding to the random effect vector  $b_i$ . The random effect  $b_i$  and the random error  $\varepsilon_i$  are independent of each other and obey the mean value of 0. The normal distribution of the covariance matrix is determined by  $\Sigma_b$  and  $\sigma_e^2$ :  $b_i \sim N_q(0, \Sigma_b)$  and  $\varepsilon_i \sim N(0, \sigma_e^2)$ . Covariance matrixes  $\Sigma_b$  and  $\sigma_e^2$  represent the variance between individuals and within individuals, respectively.

Generally, the parameter estimation of skew distribution is complicated. The skew-normal distribution linear mixing model proposed in this study extends the general linear mixed model, defined based on the linear model in Eq. (2). Assuming that the random effects and random errors obey the skew-normal distribution, the specific distribution is expressed as

$$b_i \overset{iid}{\sim} SN_q(0, \Sigma, \alpha_b), \varepsilon_i \overset{ind}{\sim} SN_{k_i}(0, \Sigma_i, \alpha_{ei}), i = 1, \dots, m. \tag{3}$$

The assumptions of random effects and random errors are independent of each other, contributing to the following hierarchical structure.

$$Y_i | b_i, \beta, \Sigma_i, \alpha_{ei} \overset{ind}{\sim} N_{k_i}(X_i\beta + Z_i b_i, \Sigma_i, \alpha_{ei}) \tag{4}$$

$$b_i | \Sigma, \alpha_b \overset{ind}{\sim} SN_q(0, \Sigma, \alpha_b), i = 1, \dots, m. \tag{5}$$

where  $\alpha_{ei}$  and  $\alpha_b$  are vectors, containing shape parameters  $\alpha_{ei1}, \dots, \alpha_{eini}$  and  $\alpha_{b1}, \dots, \alpha_{bq}$ , respectively. It is assumed that  $\Sigma_i = \sigma_e^2 I_{ni}$  and  $\alpha_{ei} = \alpha_e 1_{ni}$ ,  $i = 1, \dots, m$ , where  $1_g$  denotes the  $g$ -dimensional vector of ones. The standard diagonal matrix of matrix

$\Sigma_i$  is  $\sigma_i = \sigma_e I_{k_i}$ , where  $\sigma_e$  a real number. According to the characteristics of skew normal distribution, the expectation and variance of response variable  $Y_i$  are obtained as follows.

$$E[Y_i|b_i, \beta, \sigma_e^2, \delta_{ei}] = X_i\beta + Z_i b_i + \sigma_e \delta_{ei} (2/\pi)^{1/2} 1_{k_i} \tag{6}$$

$$\text{Var}[Y_i|b_i, \beta, \sigma_e^2, \delta_{ei}] = \sigma_e^2 (I_{k_i} - \delta_{ei} \delta_{ei}^T (2/\pi)). \tag{7}$$

where  $\delta_{ei} = (1 + \alpha_{ei}^T \alpha_{ei})^{-1/2} \alpha_{ei}$ ; the observation value  $y = (y_1^T, \dots, y_m^T)^T$  is conditionally distributed for a given random effect  $b = (b_1^T, \dots, b_m^T)^T$ ; other parameters are shown as

$$f(y|b, \beta, \sigma_e^2, \delta_e) = \prod_{i=1}^m 2^{k_i} \psi_{k_i}(y_i - X_i\beta - Z_i b_i; \sigma_e^2 I_{k_i}) \Psi(\alpha_e^T \sigma_e^{-1} (y_i - X_i\beta - Z_i b_i)). \tag{8}$$

The model mainly infers parameter vector  $\theta = (\beta^T, \sigma_e^T, \alpha^T)^T$ ,  $\alpha = (\alpha_e, \alpha_b^T)^T$ ,  $\alpha_b = (\alpha_{b1}, \dots, \alpha_{bq})^T$ . This model is based on the distribution of response variable  $Y_i$ . Random effects and random errors are not limited to the normal distribution and can be used to evaluate asymmetric data.

*3.1.2. Bayesian Estimation of SNLMM* The loss distribution is estimated using the Bayesian method by specifying the prior distribution of unknown parameters to determine the posterior distribution of parameters. The prior distribution of the fixed effect vector  $\beta$  is usually taken as multivariate Gaussian distribution  $N(\beta | \mu_\beta, \sigma_\beta^2)$ , with density function given by

$$p(\beta | \mu_\beta, \sigma_\beta^2) \propto \exp\left\{-\frac{1}{2}(\beta - \mu_\beta) \sigma_\beta^{-2} (\beta - \mu_\beta)\right\} \tag{9}$$

The prior covariance structure of scale parameter  $\sigma_e^2$  is the inverse Wishart distribution  $IG(\sigma_e^2 | \tau_e, T_e)$  of matrix  $T_b$ , with density function given by

$$p(\sigma_e^2 | \tau_e, T_e) = \frac{(T_e)^{\tau_e}}{\Gamma(\tau_e)} \left(\frac{1}{\sigma_e^2}\right)^{(\tau_e)+1} \exp\left\{-\frac{T_e}{\sigma_e^2}\right\} \tag{10}$$

The prior distribution of scale matrix  $\Sigma$  with random effect is  $\tau_b$  degrees of freedom and the variable covariance structure is inverse Wishart distribution  $IW(\Sigma | T_b)$  of matrix  $T_b$ . The prior distribution of skewness parameters  $\alpha_e$  and  $\alpha_b$  is a normal distribution  $N(\alpha_e | \mu_e, \gamma_e^2) I\{\alpha_e > 0\}$  and a multivariate truncated Gaussian distribution  $N_q(\alpha_b | \mu_b, \gamma_b^2) I\{\alpha_b > 0\}$ , respectively, where  $\gamma_b^2$  denotes a positive definite matrix, and  $I\{A\}$  represents the indicative value of dataset  $A$ , with density function given by

$$\pi(\alpha_e | \mu_e, \gamma_e^2) \propto \exp\left\{-\frac{1}{2\gamma_e^2}(\alpha_e - \mu_e)^2\right\} I\{\alpha_e > 0\} \tag{11}$$

and

$$\pi(\alpha_b | \mu_b, \gamma_b) \propto \exp\left\{-\frac{1}{2}(\alpha_b - \mu_b)^T \gamma_b^{-1}(\alpha_b - \mu_b)\right\} I\{\alpha_b > 0\} \tag{12}$$

Assuming that the prior distribution of the above parameters is independent of each other, according to the conditional distribution of the parameters, the posterior distribution of all relevant parameters can be obtained by

$$\begin{aligned} \pi(\theta, b, \Sigma | y) &\propto \prod_{i=1}^m SN_{n_i}(y_i | X_i \beta + Z_i b_i, \sigma_e^2 I_{n_i}, \alpha_e) \times \prod_{i=1}^m SN_q(b_i | 0, \Sigma, \alpha_b) \\ &\times N(\beta | \mu_\beta, \sigma_\beta^2) \times IG(\sigma_e^2 | \tau_e, T_e) \\ &\times IW_{\tau_b}(\Sigma | T_b) \times N(\alpha_e | \mu_e, \gamma_e^2) I\{\alpha_e > 0\} \\ &\times N_q(\alpha_b | \mu_b, \gamma_b^2) I\{\alpha_b > 0\} \end{aligned} \tag{13}$$

However, it is quite complicated a to obtain the marginal distribution of the amount of interest involved in Eq. (13) from an analytical perspective. The MCMC method is adopted in this study to perform Bayesian analysis, considering that the traditional Bayesian method is difficult to conduct on such models. In this study, the MCMC method is used to simulate the posterior distribution of each parameter, and Gibbs sampling is applied to the conditional posterior distribution of each parameter to simplify its sophisticated formula.

### 3.2. MCMC Simulation Analysis

The MCMC method involves the skew-normal distribution function, which is difficult to establish the model. The hierarchical structure studied by Arellano-Valle et al. [23] is employed in this paper to solve the complicated skew problem. The skew-normal linear mixed model is expressed in the equivalent form of simple normal distribution, making it convenient to use Gibbs sampling for posterior distribution. The complete layered structure of the model is expressed as

$$Y_i | b_i, \beta, \sigma_e^2, \alpha_e, w_{ei} \sim N_{n_i}(X_i \beta + Z_i b_i + \alpha_e w_{ei}, \sigma_e^2 I_{n_i}), i = 1, \dots, m.n \tag{14}$$

$$w_{ei} \sim N_{n_i}(0, I) I\{w_{ei} > 0\} \tag{15}$$

$$b_i | \Sigma, \alpha_b, w_{bi} \sim N_q(\alpha_b w_{bi}, \Sigma) \quad (16)$$

$$w_{bi} \sim N_q(0, I) I\{w_{bi} > 0\} \quad (17)$$

$$\beta \sim N_p(\mu_\beta, \sigma_\beta^2) \quad (18)$$

$$\sigma_e^2 \sim IG(\tau_e, T_e) \quad (19)$$

$$\Sigma \sim IW_{\tau_b}(T_b) \quad (20)$$

$$\alpha_e \sim N(\mu_e, \gamma_e^2) I\{\alpha_e > 0\} \quad (21)$$

$$\alpha_b \sim N_q(\mu_b, \gamma_b^2) I\{\alpha_b > 0\} \quad (22)$$

The corresponding programming code can be easily obtained according to the layered expressions. According to Eqs. (14)–(22), the complete conditional distribution required to realize the Gibbs sampler can be directly derived and sampled. The algorithm starts from the initial value and cycle of all variables mentioned, which generates samples from the above conditional distribution until convergence. The method mentioned by Gelman and Rubin [24] is utilized to verify the conditions, such as running several parallel chains. Regarding unbalanced datasets, existing statistical software, such as MATLAB, can be used to easily calculate the above equations. For balanced datasets, the Gibbs sampler can be implemented by OpenBUGS or JAGS software. Besides, the R language jags programming is employed to perform Gibbs sampling in this study.

#### 4. Case Study

In this section, the Bayesian skew-normal linear mixed model is established and compared with Frees et al. [3] model to verify the accuracy of SNLMM model based on the US employer liability insurance loss claim data Klugman [22]). Then, the SNLMM model is applied to China's mainland property insurance statistics [29] (<http://www.stats.gov.cn/tjsj/ndsjs/>) to explore the relationship between the relevant variables. Additionally, the premium claims are estimated to determine the premium rate based on variance analysis to determine the premium rate.



#### 4.1. U.S. Employer Liability Insurance Claims Data

Pure premium (PP) is defined as the loss of income per dollar due to permanent, partial disability. Frees selected the logarithm of pure premium (LnPP) as the response variable  $Y$ ; other explanatory variables were the observation year  $t$  and the occupation category  $w$ . Frees established a logarithmic normal linear mixed model, which used logarithmic transformation to explain heteroscedasticity. Figure 1 demonstrates that the pure premium claim data are panel data with right deviation and heavy tail. The correlation between the same set of data needs to be considered, and the logarithmic transformation exhibits data information loss and increase errors. In this study, the model is optimized by establishing a Bayesian skew-normal linear mixed model, so as to improve the reliability of data prediction.

*4.1.1. Model Analysis* Frees conducted logarithmic transformation of pure premiums under the consideration of the right-deviation characteristics of data. After logarithmic transformation, the right-deviation of the data was offset. The distribution map of logarithmic pure premium data presented symmetry. The established logarithmic linear mixed Model I was

$$\log(PP_{wt}) = \gamma_1 + d_{1w} + d_{2w}Year_t + e_{wt}$$

$$w = 1, \dots, 121; t = 1, \dots, 7. \quad (23)$$

where  $w$  denotes the occupational category and  $t$  represents the observation year. Since the pure premium does not change with the observation year, the model only contains a constant term  $\gamma_1$  as a fixed effect variable. The value of the observation year is (Year-4) /8. Additionally, the model is a general linear mixed model. The random effect  $d_w$  and the random error term  $e_{wt}$  are subject to the normal distribution with the mean value of zero, and the random effect  $d_w$  follows the binary normal distribution. Frees used t statistics and likelihood statistics to judge the adequacy of the model, which fitted the panel data model with the data to create a credibility predictor to solve the skew problem of the model.

Logarithmic transformation reduces the correlation of the same group of data. After the establishment of the linear mixed model, there are some defects in the internal structure, and the prediction data lacks certain authenticity. This study uses the proposed Bayesian skew-normal linear mixed model to analyze pure premium data. Normal distribution is a special case of skew normal distribution. Therefore, the skewed normal linear mixed model can process both skewed data and symmetric distribution data.

According to the evaluation results of the Bayesian model, DIC statistics Spiegelhalter et al. [29] and BIC criterion Yuan and Lu [30] are employed in this

paper to analyze the accuracy of the three models. The expressions of DIC and BIC are

$$DIC = 2\overline{D(\theta)} - D(\overline{\theta}) = D(\overline{\theta}) + 2k \quad (24)$$

$$BIC = \ln(n)k - 2\ln(L) \quad (25)$$

where  $D(\theta)$  is equal to the logarithmic likelihood function of negative twice,  $k$  denotes the number of effective parameters of the model, and  $L$  indicates the likelihood function. As suggested in the previous studies, smaller DIC and BIC were better in fitting and prediction effects of the model.

In this paper, the partial normal distribution is used to define the random effect  $d_w$  and the random error term  $e_{wt}$  to form a partial normal linear mixed model. This paper will compare and analyze the two models, and the response variables are pure premium logarithm. The first model is a general linear mixed model, and the random effects and random errors obey the normal distribution. The second model introduces skew normal distribution, assuming that both random errors and random effects obey skew normal distribution. Suppose that the random error term obeys the normal distribution, that is, the Frees model.

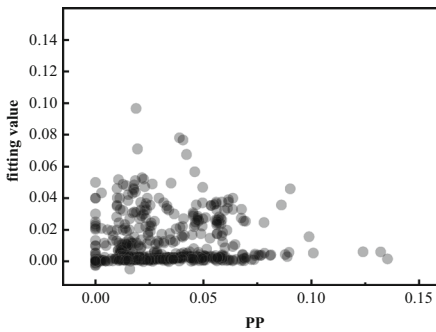
In order to facilitate the comparison of the same parameters of the two models, the same prior distribution is given. In each simulation, 100 Monte Carlo data sets are simulated from Eq. (23) according to additional specification described below, and Eq. (23) is fit to each data set via the strategy in Proposition 1 (see Appendix) using the R software jointly with the R2WinBUGS package and the following prior specifications. The prior distribution of the fixed effect is regarded as a normal distribution  $\gamma_0 \sim N(0, 10^2)$ . The prior distribution of the scale parameter of the error distribution is assumed to obey inverse gamma distribution  $\sigma_e^2 \sim IG(0.01, 0.01)$ . The prior distribution of the scale matrix of the random effect prior distribution  $d_w$  is inverse Wishart distribution  $\Sigma \sim IW_2(I_2)$ . The skewness parameter is  $\alpha_e \sim N(0, 10^2) I\{\alpha_e > 0\}$ ,  $\alpha_w \sim N(0, 10^2) I\{\alpha_w > 0\}$ . The prior distributions of parameters  $\omega_e$  and  $\omega_w$  are assumed as  $\omega_e \sim N(0, 1) I\{\omega_e > 0\}$ ,  $\omega_d \sim N(0, 1) I\{\omega_d > 0\}$ . The prior distribution considered is close to the non-informative distribution, that is, the prior distribution with large variance is considered. The parameter estimates and standard deviations of various models are shown in Table 1.

$\Sigma_{11}$  is the variance of the random effect  $d_{1w}$ ,  $\Sigma_{12}$  is the covariance of the random effects  $d_{1w}$  and  $d_{2w}$ ,  $\Sigma_{22}$  is the variance of the random effect  $d_{2w}$ . The last two lines are negative twice the log-likelihood function value and DIC information for comparative analysis of the model.

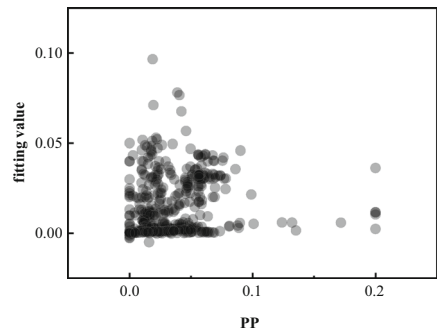
It can be seen from Table 1 that when the random error term obeys the partial normal distribution, the DIC value of the model is obviously the smallest, indicating that the model has the best fitting effect on the pure premium logarithm. However, comparing the fitting values of the models with the actual values, it can be found that the fitting values of these models for the logarithm of pure premiums are not much different. See Figures 2 and 3, this is mainly because the loga-

**Table 1**  
**The Posterior Mean of Parameters with the Logarithm of Net Premium as the Response Variable**

Parameter	Normal		Skew-normal	
	Mean	SD	Mean	SD
$\gamma_1$	-4.35	0.08	-4.59	0.16
$\sigma^2$	0.35	0.02	0.31	0.043
$\sum_{11}$	0.7	0.1	0.70	0.096
$\sum_{12}$	-0.11	0.09	-0.1	0.082
$\sum_{22}$	0.36	0.13	0.36	0.13
$\alpha_e$	-	-	0.38	0.017
$\alpha_{d1}$	-	-	0.23	0.066
$\alpha_{d2}$	-	-	0.078	0.053
-2log-likelihood	2515		1868	
DIC	2665		2162	



(a) Normal linear mixed model



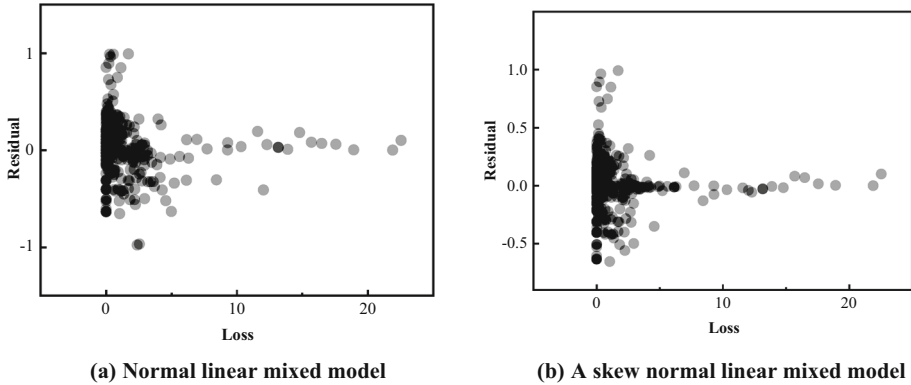
(b) A skew normal linear mixed model

**Figure 2. Scatter plot of pure premium and fitting value.**

rithm of pure premiums presents an approximate symmetrical shape, so the superiority of skewed distribution is not fully reflected.

4.1.2. *Model Optimization* Logarithmic transformation reduces the correlation of the same group of data. After the establishment of the linear mixed model, there are some defects in the internal structure, and the prediction data lacks certain authenticity. This study uses the proposed Bayesian skew-normal linear mixed model to analyze the pure premium observation data directly.

The skew-normal linear mixed model for the pure premium claim dataset is established directly in this study, and the explanatory variables mentioned above are the same as Model II. Therefore, the Bayesian skew-normal linear mixed Model III can be obtained as



**Figure 3. Residual and compensation scatter plot.**

$$PP_{tn} = \Upsilon_o + d_{on} + d_{1n}Year_n + e_m. \quad (26)$$

In this model, the random effects and random errors are assumed to follow the skew-normal distribution, and the pure premium is considered the response variable.

The skew-normal linear mixed model does not need to transform the data and can directly calculate the pure premium data on the right. In this way, the explanatory ability of the model to the data is significantly improved, and the influence of the correlation of the same set of data caused by the transformation is avoided. For the convenience of comparison, the same meaning parameters give the same prior distribution.

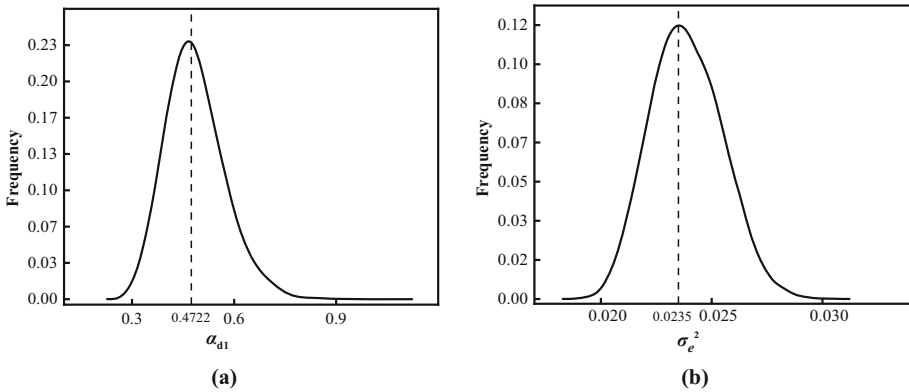
The prior distribution of parameters is obtained based on the Bayesian method, and the hierarchical expression of the model is acquired using the equivalent form of the skew-normal linear mixed model. In this paper, the hierarchical expression operation code is programmed by JAGS in R language through Gibbs sampling of MCMC method, with 30,000 iterations. The first 10,000 iterations are abandoned until the sample converges. Then, the parameter posterior distribution sample value is obtained. Meanwhile, the lag value is set as 4 to avoid related problems in the generated Markov chain. The mean value and standard deviation of the estimated values of each parameter are obtained using the sample values of the posterior distribution, as listed in Table 2.

Table 2 implies that the posterior distribution mean of the skewness parameter is close to 0.5, and the variance interval of the error parameter is relatively narrow. The dotted line in Figure 4 represents the median of the skewness parameter and error parameter. The median of skewness parameter is close to 0.5, and the median of error parameter is close to the mean, suggesting that the established skew-normal linear mixed model is more suitable for this dataset.

**4.1.3. Comparative Analysis** Since the response variables used in Tables 2 and 1 are completely different, one is the pure premium and the other is the logarithm of the pure premium, their DIC values cannot be directly compared, but the fit-

**Table 2**  
**A Posteriori Mean of Parameters with Pure Premium as Response Variable**

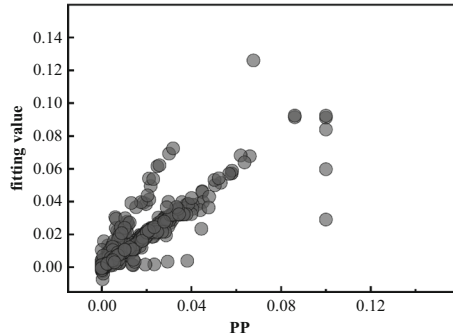
Parameter	Skew-normal			
	Mean	SD	2.5%	97.5%
$\gamma_0$	-0.0395	0.0365	-0.1257	0.1104
$\alpha_{d1}$	0.4772	0.0318	0.3421	0.5090
$\alpha_e$	0.3800	0.0547	0.0685	0.4140
$\sigma_e^2$	0.0240	0.0016	0.0114	0.0530
$\Sigma_{11}$	0.05397	0.0321	0.0440	0.1711
$\Sigma_{12}$	0.0605	0.0243	-0.0099	0.0861
-2log-likelihood	-3899			
BIC	-2.225			
DIC	-3359			



**Figure 4. The posterior distribution frequency distribution curve of parameters. (a) The posterior distribution frequency curve of the skewness parameter  $\alpha_{d1}$ . (b) The posterior distribution frequency curve of the error parameter  $\sigma_e^2$ .**

ting values and model residuals of different models can be compared and analyzed.

Figure 5 is a scatter plot between the observed values of the residuals and pure premiums of the partial slash linear mixed model and their fitted values. It can be seen that the fitted value and the observed value of the model III with the pure premium as the response variable are closer to the diagonal, and the fitting effect is better than that of the Figure 2. It can be seen that the fitted values and observed values of the partial slash linear mixed model are closer to the diagonal, and the residual fluctuation range is smaller. The residuals are concentrated near the zero value, and the fitting effect is significantly better than the normal linear



**Figure 5. Scatter plot of observations and fitting values of net premiums.**

mixed model shown in Figures 3 and 4, which further indicates that the model has a better fitting effect on the pure premium data.

Figure 6 represents the scatter plot of residuals. From the graph, it can be seen that the residual fluctuation range is small and gathers near the zero value. The fitting effect is obviously better than the model I and model II shown in Figure 3, which further shows that the model III fitting of pure premium data greatly improves the model's ability to interpret data.

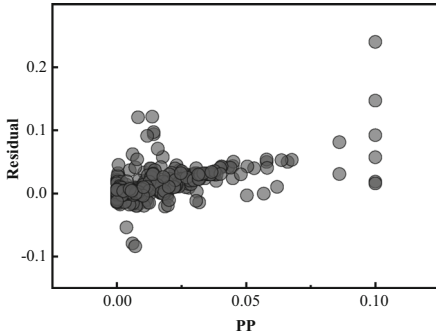
#### 4.2. Mainland China Property Insurance Statistics

Property insurance business scale and loss claims have strong regional characteristics. In this study, the relationship between regional property insurance loss claims and regional economy GRP is mainly discussed. The data of the two over the years are fitted and analyzed. The claims for property insurance losses in the next 3 years are further predicted.

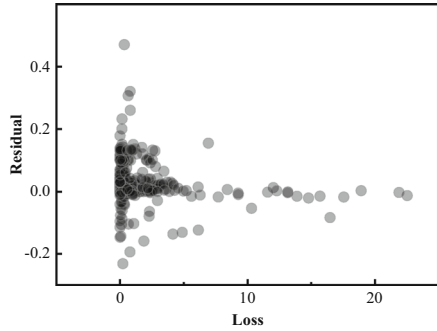
*4.2.1. Variable Assumptions* Figure 7 illustrates the relationship between total property insurance losses with GDP and economic growth over the past decade. The total insurance loss claims and GDP increase with time; the total amount of insurance loss claims grows faster when GDP grows faster; the growth rate of total insurance loss claims decreases when GDP growth slows down, implying a linear relationship between the two parameters. In recent years, China has entered a transitional period, with a large GDP base and rapid growth of per capita GDP, while the growth rate has declined.

Figure 8 exhibits the time series of claims for property insurance losses in the regions over the past decade with the observation year. Figure 8 reveals that the level of insurance loss claims increases with the increasing observation years for most regions, while there are significant differences in the level of insurance loss claims among regions with different economic levels.

According to the property insurance loss claim data in China Statistical Yearbook and China Fire Yearbook, these data variables are summarized in Table 3.

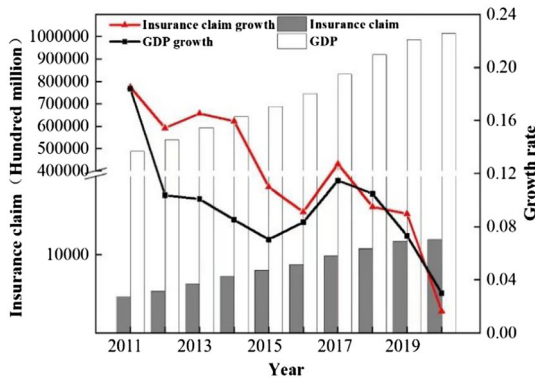


(a) Residual and net premium scatter plot

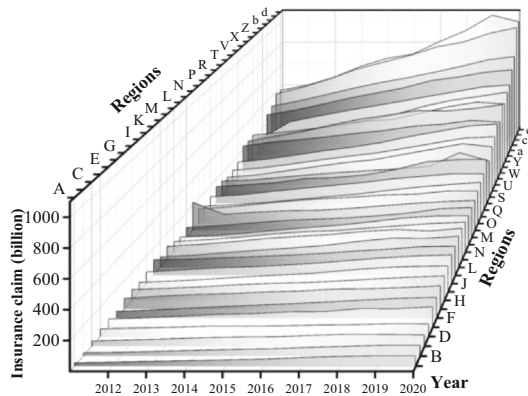


(b) Residual and claim value cost scatter plot

**Figure 6. Residual scatter plot.**



**Figure 7. Comparison and growth of premium claim with GDP.**



**Figure 8. Property insurance claims in different regions.**

As suggested in Table 3, the trend of insurance loss claims and regional economy with time is somewhat consistent, and their variances are relatively large.

A preliminary analysis indicates that it is somewhat similar to the linear trend of regional economies GRP with time, though the level of claims for property insurance losses may depend on the complicated characteristics of different regions. Moreover, the dataset demonstrates the distribution characteristics of right deviation, thick tail, and asymmetry. In this study, regional property insurance loss claim is designated as response variable  $Y$  (billion). Random variables include observation year  $t$ , regional economy GRP (billion), observation area category  $n$ , explanatory variable  $n = 34$ , and  $t = 10$  (2011, ..., 2020). Different letters instead of region names are used to protect data sources and information security. Meanwhile, no more information about the region category is provided.

**4.2.2. Model Establishment** The above analysis implies a certain linear trend between the amount of property insurance loss compensation and the observation year, as well as a correlation between the amount of property insurance loss compensation and the regional economy GRP. Furthermore, the residual error of the amount of property insurance loss compensation fitted from a single region does not exhibit a special trend. Then, the following skew-normal linear mixed model expression is obtained by

$$Y_m = \beta_o + \beta_1 GRP_m + b_{on} + b_{1n} T_n + e_m \quad (27)$$

$$n = 1, \dots, 31; t = 1, \dots, 10. \quad (28)$$

where,  $Y_m$  represents the claim for property insurance loss with time  $t$  and regional category change  $n$ , and the constant term  $\beta_0$  denotes the fixed effect variable. The random effects include the constant term  $b_{on}$ , the observation year  $b_{1n}$  and the regression coefficient  $\beta_1$  of the observation region. It mainly manages the correlation between different observation years and the observation regions. The value of the observation year is (Year-5) /10, which is convenient for calculation and statistics. Three different distribution models of random effects and random errors are adopted in this study to make a comparative analysis.

**Model 1:** The random error term  $e_n = (e_{1n}, \dots, e_{10n})^T$  and the random effect are assumed to obey the multivariate skew-normal distribution model;

**Model 2:** A model with the random error term  $e_n = (e_{1n}, \dots, e_{10n})^T$  for a multivariate normal distribution, and the random effect for the multivariate skew-normal distribution model;

**Model 3:** The response variable is the logarithm of premium claims, random effects and random errors obey the Gaussian distribution model.

The prior distributions of the same parameters of the three models are assumed to be the same to perform comparative analysis. This study is based on a large variance. The prior distribution of the fixed effect is set to independent distribution  $\gamma_0 \sim N(0, 10^2)$ , and the prior distribution of the scale parameter of the error distribution is set to  $\sigma_e^2 \sim IG(0.01, 0.01)$ . Hence, the mean value of the distribution



**Table 3**  
**Summary Statistics of Insurance Claim Data and Regional Economy GRP, by Year**

Variable	Year (t)	Mean	Median	Standard Deviation	Minimum	Maximum
Insurance claim data (billion)	1	59.51528	46.09033	47.845	178.49	2.96
	2	77.30833	60.07841	61.83	240.08	3.36
	3	102.2619	87.83932	79.055	352.56	4.02
	4	104.6056	79.86177	80.46	336.3	4.1
	5	117.0072	91.7645	87.01	403.04	5.82
	6	130.4622	99.15416	94.695	437.64	6.97
	7	140.2003	106.9386	107.28	455.63	8.72
	8	162.9508	126.851	113.46	516.78	12.6
	9	179.4919	136.3548	135.57	542.23	16.25
	10	191.9178	147.7201	147.145	642.28	17.47
	Total	1265.721	982.653	954.35	4105.03	82.27
GRP (billion)	1	16,166.90	12,391.30	13,365.89	605.83	52,673.59
	2	18,598.43	14,454.00	14,325.90	701.00	57,068.00
	3	20,345.67	16,045.00	15,613.31	802.00	62,164.00
	4	22,073.14	17,689.94	16,986.30	920.83	67,792.24
	5	23,315.09	17,831.51	18,218.96	1026.39	72,812.55
	6	25,163.55	18,499.00	20,103.05	1151.41	80,854.91
	7	27,596.12	20,818.50	22,213.03	1310.63	89,879.23
	8	29,506.69	21,984.80	23,905.15	1477.63	97,277.77
	9	31,784.95	24,757.50	25,949.29	1697.82	107,671.07
	10	32,658.55	25,115.00	26,661.81	1902.70	110,760.90
	Total	247,209.09	189,586.55	197,342.69	11,596.24	798,954.26

is 2. The prior distribution of the scale matrix of the prior distribution  $d_n$  of random effects is considered inverse Wishart distribution  $\sum \sim IW_2(I_2)$ ; the skewness parameter is assumed to  $\alpha_e \sim N(0, 10^2) I\{\alpha_e > 0\}$ ,  $\alpha_{b0} \sim N(0, 10^2) I\{\alpha_{b0} > 0\}$ , and  $\alpha_{b1} \sim N(0, 10^2) I\{\alpha_{b1} > 0\}$ . It is demonstrated that the results of the analysis are quite robust since all previous considerations are based on large variances.

4.2.3. *Data Analysis* The 40,000 iterations were conducted after the previous 10,000 iterations were abandoned, so as to analyze the data. The lag value is set as 5 to avoid the related problems in the generation chain. The JAGS programming code in Model 1 is provided in the appendix, and the posterior estimation average of model parameters are presented in Table 4.

The results of parameter estimation (posterior mean) are listed in Table 4. Following the analysis of the Table 4, BIC and DIC criteria suggest that model 1 is most suitable for data.

Figure 9 provides the Markov data chain and posterior distribution sample frequency histogram of some parameters in Model 1. As observed in Figure 9, the skewness parameter is closer to 0.5, and the error confidence interval is relatively narrow. It can be seen from the figure that there is no convergence problem in the

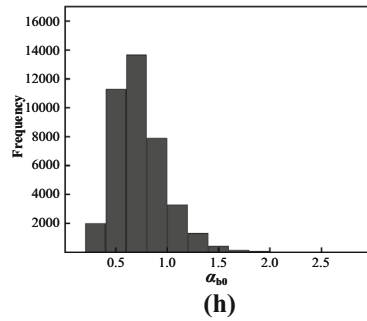
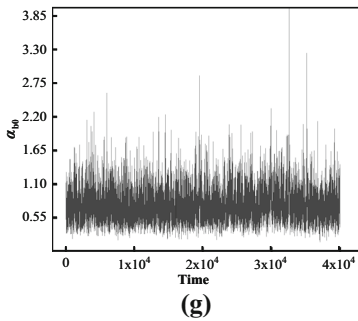
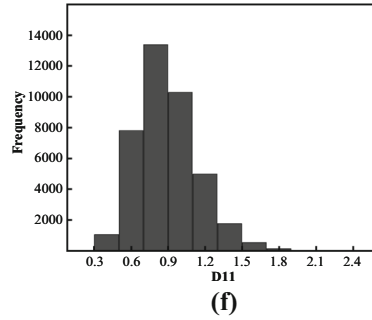
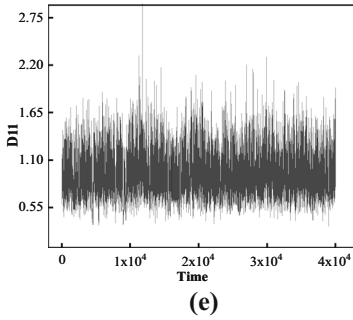
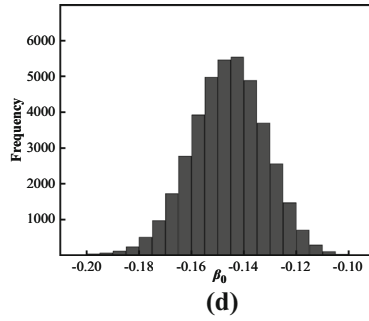
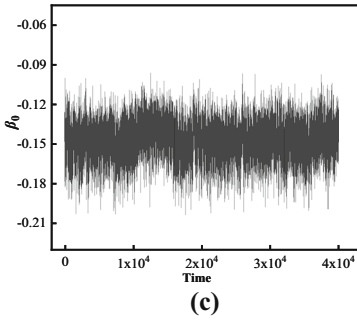
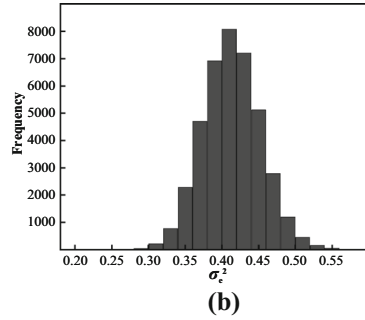
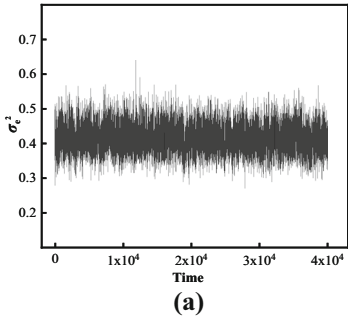
**Table 4**  
**A Posteriori Average of Three Model Parameters**

Parameters	$\alpha_{b0}$	$\alpha_{b1}$	$\beta_1$	$\alpha_c$	$\beta_0$	$\Sigma_{11}$	$\Sigma_{12}$	$\sigma_c^2$
<i>Model 1</i>								
Mean	0.6055	0.5865	0.0541	0.6052	-0.1531	1.1140	1.3680	0.4770
SD	0.4895	0.4310	0.0422	0.5311	0.3295	7.7110	2.4520	0.3890
2.5%	0.0213	0.0211	0.0001	0.0177	-0.7321	0.4822	3.0200	0.1700
97.5%	1.8610	0.4105	0.0970	2.0100	0.5553	1.6010	9.5630	0.4100
-2log-likelihood	-1004.2							
BIC	-0.843							
DIC	-275.5							
<i>Model 2</i>								
Mean	0.6080	0.6610	0.1482	-	0.1228	0.9156	8.32603	0.6500
SD	0.5053	0.5329	0.1142	-	0.2620	0.3150	1.26403	0.4990
2.5%	0.0310	0.0224	0.0021	-	-0.4841	0.4864	9.73700	0.2000
97.5%	1.9190	1.8650	0.1350	-	0.5365	1.5580	1.67403	0.4710
-2log-likelihood	-977.2							
BIC	-0.356							
DIC	-173.3							
<i>Model 3</i>								
Mean	-	-	0.2123	-	0.1386	3.6730	2.2747	1.7800
SD	-	-	0.1459	-	0.3468	7.6431	1.1849	1.1480
2.5%	-	-	0.0021	-	0.3237	0.0777	2.4103	4.6200
97.5%	-	-	0.1782	-	1.7390	1.2230	1.9238	9.6600
-2log-likelihood	1150.5							
BIC	1.348							
DIC	270.2							

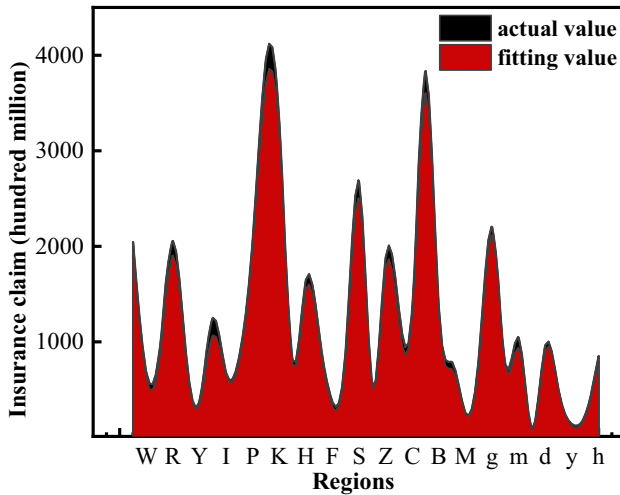
chain generated by the parameters. Figure 10 exhibits the comparison curve between the estimated value and the real value of each region. The analysis results in Figure 10 reveal that the values of the two curves are very close, and the changing trends are consistent. Hence, Model 1 has a high fitting degree.

Figure 11 demonstrates the prediction values of regional property insurance loss claims in the following 3 years. The previous analysis implied that the insurance loss claims significantly varied in different regions, and the insurance loss claims generally changed little in small regions, otherwise changed dramatically. The development trend of insurance loss claims in Figure 9 meets this result analysis.

In summary, the proposed Bayesian SNLMM model has superior performance in intra-sample fitting and sample prediction and provides a reasonable model to analyze the influence of covariates.



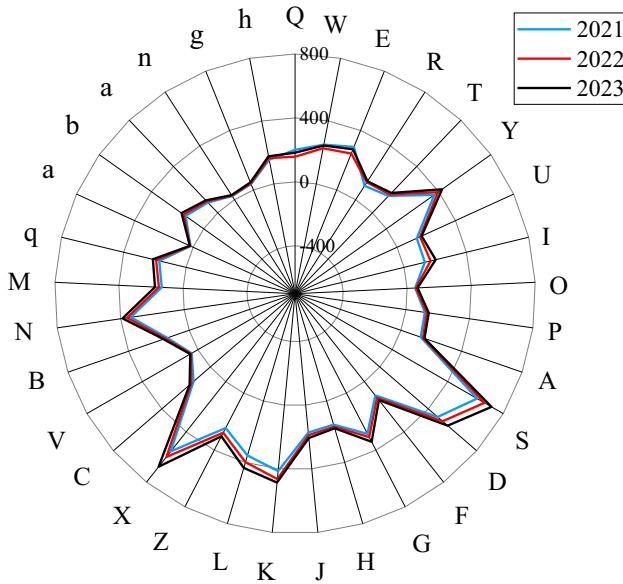
**◀Figure 9. Sample chain of simulated data (left image) and histogram of posterior distribution of parameters (right image). a is Identified as  $\sigma_e^2$  Sample chain, b is Identified as  $\sigma_e^2$  posterior distribution histogram, c is Identified as  $\beta_0$  Sample chain, d is Identified as  $\beta_0$  posterior distribution histogram, e is Identified as D1 1 Sample chain, f is Identified as D1 1 posterior distribution histogram, g is Identified as  $\alpha_{b0}$  Sample chain, h is Identified as  $\alpha_{b0}$  posterior distribution histogram.**



**Figure 10. Comparison of actual values and fitting values for each region.**

## 5. Discussions

An essential aspect of analyzing insurance loss claim data is that they tend to follow a skew distribution, because extreme events are not uncommon and exhibit a heavy tail. A Bayesian linear mixed model is proposed in this paper. Its random effects and random errors belong to skew-normal distribution. This model can identify non-normal and asymmetric behaviors that may occur in the random errors or effects. Firstly, the methods used and the Bayesian skew-normal linear mixed model are introduced. Secondly, the data fitting results of logarithmic linear mixed model and Bayesian skew-normal linear mixed model are compared based on the dataset of American Employers liability industrial insurance to empirically test the superiority of the Bayesian skew-normal linear mixed model (Pure premium is the response variable). Finally, the Bayesian SNLMM model is applied to the property insurance loss dataset in mainland China to explore the relationship between insurance claims and the regional economy GRP and predict insurance loss claims.



**Figure 11. Forecast value of property insurance claims in regions in the next 3 years.**

The results of this study suggest that both the US employer liability insurance claim dataset and the Chinese mainland property insurance loss dataset have complicated characteristics of right deviation and heavy tail. As revealed by asymmetric pure premium panel data analysis, the Bayesian skew-normal linear mixed model has a more accurate fitting effect and higher data accuracy compared with the logarithmic linear mixed model. The research based on property insurance loss claim datasets demonstrates a certain linear relationship between regional economy and insurance claims, and the trend is consistent. It is assumed that the first model simulation is more advantageous among the three different linear mixed models. Then, the insurance loss claims in different regions in the next 3 years present a trend of first decreasing and then increasing. This flexibility that other linear models cannot achieve highlights the advantages of Bayesian methods over existing actuarial models.

In previous linear mixed model research, a robust reasoning method using heavy tail distribution is proposed. For example, Pinheiro and Vidakovic [31] used maximum likelihood to describe the robust Gaussian LMM model proposed by Laird and Ware [28]. They assumed that the distribution of errors had the same degree of freedom as the distribution of random effects, and the random effects were independent. Two different stochastic processes are difficult to be controlled by the same degree of freedom parameters. The method designed in this paper can solve this problem by allocating different parameters. Modeling by modifying the known distribution is a hot topic. Many alternative structures of Arellano-Valle

et al. [23] are based on Genton [10]. In this study, the structure proposed by Arellano-Valle et al. [23] is selected mainly for simple calculation.

The fourth section analyzes the Frees data case. When the data is converted, some information may be lost. The fitting results between the models are not much different, and the superiority of the skewed distribution model is not fully reflected. When analyzing data directly, the fitting results and residual values of the SNLMM model are more accurate. Although the related density function is difficult to handle, the MCMC method can easily fit the model. This method can be implemented on computing by using simple and accessible software (such as R language, WinBUGS). Other skewed distribution variants currently available are not so easy to implement. The MCMC method and standard information standards are very effective in identifying correct normal asymmetric models, as shown in simulation studies. A relatively large advantage of this model is that it can give more flexible assumptions about the random effects and random errors of the model. However, when applying the SNLMM model to analyze the data, the prior distribution of the parameters considered is close to the non-informative distribution, that is, the prior distribution with large variance is considered, and the model is suitable for asymmetric longitudinal data.

To sum up, the same method can also be applied to other types of data, and these data have similar skewness characteristics. In other words, the Bayesian SNLMM model for objective Bayesian analysis of data, including parameter distribution and prior, is applicable to other disciplines, such as engineering and environmental science.

## 6. Conclusions

Forecasting fire insurance loss claims is a critical part of rate determination analysis. However, the insurance rates calculation is still in the exploring stage for fire insurance loss claim data, which usually has a characteristic of right skew and heavy tail and is difficult to be described by the traditional normal distribution linear mixed model.

In this paper, an asymmetric fire insurance loss claim prediction model is proposed by combining the Bayesian method with the skew-normal linear mixed model. Based on the traditional linear mixed model, it is assumed that the random effects and random errors obey the skew-normal distribution. Then, it can describe the complicated characteristics of skewness and the heavy tail of insurance loss claim data. Moreover, the Bayesian MCMC method is adopted to estimate the model parameters, and the posterior parameter mean estimation is obtained using Gibbs sampling, contributing to addressing complicated operational problems. These improvements have achieved better results compared to the normal linear mixed model in theory and practice.

Additionally, the proposed Bayesian SNLMM model has been validated on two property insurance loss claims datasets, including the latest insurance data statistics. Firstly, this model demonstrates better goodness of fitting for pure premium data, better data interpretation ability, and lower DIC compared with the log lin-

ear mixed model. Secondly, in the process of predicting the number of fire insurance claims, the accuracy of the model is high, and the influence of different data characteristics is small. Finally, the method estimates parameter samples by simple and easy-to-use software (such as R language, WinBUGS). The Bayesian skew-normal linear mixed model can describe the specific form of claim data in different categories by identifying skew patterns. This model has great flexibility and is suitable for modeling and analysis in various complicated situations.

In addition, the estimation model proposed in this paper also has shortcomings and needs further study. Firstly, the random model of different factors in MCMC simulation is explored to reduce the influence of variable weight structure and improve the accuracy. Secondly, when the model is applied to the prediction of fire insurance loss claim data, it is necessary to directly model and analyze the original longitudinal data to improve the accuracy of data prediction.

## **Acknowledgements**

The authors would like to acknowledge the financial support of Natural Science Foundation of Hebei Province (G2021507001) and China People's Police University Project (2019sycxpd006). The authors would also like to thank Intelligent Transportation Institute of Zhejiang Communication Investment Group Co., Ltd. for their great help.

## **Author Contributions**

MG: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing—original draft. ZM: Revising the manuscript critically for important intellectual content. DZ: Validation. JR: Formal analysis. SZ: data Investigation.

## **Funding**

The research is funded by Natural Science Foundation of Hebei Province (G2021507001) and China People's Police University Project (2019sycxpd006).

## **Declarations**

**Competing interest** The authors declare that there is no conflict of interests regarding the publication of this article.

## **Appendix**

To obtain the marginal distribution of  $Y_j$ , we drop the subscript  $j$ , corresponding to the  $j$  the group (or individual) to simplify notation. From Eqs. (5, 6) and the definition of skew-normal multivariate distribution in Eq. (3) with  $k = n$  and  $D =$

diag ( $\delta$ ), it follows that the marginal density of  $Y$  is obtained by computing the following integral:

$$\begin{aligned}
 f_Y(\mathbf{y}|\boldsymbol{\theta}, \lambda) &= \int_{\mathbb{R}^q} f(\mathbf{y}|\mathbf{b}, \boldsymbol{\beta}, \sigma_e^2, \delta_e) f(\mathbf{b}|\boldsymbol{\alpha}, \boldsymbol{\delta}_b) d\mathbf{b} \\
 &= \int_{\mathbb{R}^q} 2^{n+q} \phi_n(\mathbf{y}|\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{b}, (\delta_e^2 + \sigma_e^2)\mathbf{I}_n) \\
 &\quad \times \phi_q(\mathbf{b}|\mathbf{0}, \Delta_b^2 + \mathbf{D}) \Phi_q\left(\Delta_b(\Delta_b^2 + \mathbf{D})^{-1}\mathbf{b}|\mathbf{0}, (\mathbf{I}_q + \Delta_b\mathbf{D}^{-1}\Delta_b)^{-1}\right) \\
 &\quad \times \Phi_n\left(\frac{\delta_e}{\sqrt{\sigma_e^2(\delta_e^2 + \sigma_e^2)}}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{b})|\mathbf{0}, \mathbf{I}_n\right) d\mathbf{b} \tag{A-1}
 \end{aligned}$$

For which the following proposition will be used.

The values of standard information standards in the SNLMM model can be used to detect deviations from normal conditions, which will be seen in the simulation study. The specific algorithm is as follows.

Let  $Y_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \epsilon_i$ , where  $b_i \stackrel{iid}{\sim} SN_q(0, \Sigma, \alpha_b)$  and  $\epsilon_i \stackrel{ind}{\sim} SN_{ki}(0, \Sigma_i, \alpha_{ei})$  are independent. Then, the marginal distribution of  $Y_i$  is given by

$$f_{Y_i}(\mathbf{y}_i|\boldsymbol{\theta}, \lambda) = 2^{n_i+q} \phi_{n_i}(\mathbf{y}_i|\mathbf{X}_i\boldsymbol{\beta}, \Sigma_i) \Phi_{n_i+q}(\boldsymbol{\mu}_{2i} - \Gamma_i\boldsymbol{\mu}_{1i}|\mathbf{0}, \mathbf{R}_i + \Gamma_i\Lambda_i\Gamma_i^T)$$

where

$$\begin{aligned}
 \Sigma_i &= (\delta_e^2 + \sigma_e^2)\mathbf{I}_{n_i} + \mathbf{Z}_i(\Delta_b^2 + \mathbf{D})\mathbf{Z}_i^T, \quad \boldsymbol{\mu}_{1i} = \frac{\Lambda_i\mathbf{Z}_i^T(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta})}{\delta_e^2 + \sigma_e^2}, \\
 \boldsymbol{\mu}_{2i} &= \begin{pmatrix} \frac{\delta_e}{\sqrt{\sigma_e^2(\delta_e^2 + \sigma_e^2)}}(\mathbf{y}_i - \mathbf{X}_i\boldsymbol{\beta}) \\ \mathbf{0}_q \end{pmatrix}, \quad \Gamma_i = \begin{pmatrix} \frac{\delta_e}{\sqrt{\sigma_e^2(\delta_e^2 + \sigma_e^2)}}\mathbf{Z}_i \\ -\Delta_b(\Delta_b^2 + \mathbf{D})^{-1} \end{pmatrix}, \\
 \mathbf{R}_i &= \begin{pmatrix} \mathbf{I}_{n_i} & \mathbf{0} \\ \mathbf{0} & (\mathbf{I}_q + \Delta_b\mathbf{D}^{-1}\Delta_b)^{-1} \end{pmatrix}, \quad \Lambda_i = \left( (\Delta_b^2 + \mathbf{D})^{-1} + \frac{\mathbf{Z}_i^T\mathbf{Z}_i}{\delta_e^2 + \sigma_e^2} \right)^{-1}, \\
 i &= 1, \dots, m.n.
 \end{aligned}$$

Thus, denoting the log-likelihood function for  $\boldsymbol{\theta}$  and  $\lambda$  given the observed sample  $\mathbf{y} = (y_1^T, \dots, y_m^T)^T$  by  $\ell(\boldsymbol{\theta}, \lambda|\mathbf{y})$ , it can be written as



$$\ell(\theta, \lambda|y) \propto -\frac{1}{2} \sum_{i=1}^m \log|\Sigma_i| - \frac{1}{2} \sum_{i=1}^m \left\{ (y_i - X_i\beta)^T \Sigma_i^{-1} (y_i - X_i\beta) \right\} \\ + \sum_{i=1}^m \log \Phi_{n_i+q}(\mu_{2i} - \Gamma_i \mu_{1i} | 0, R_i + \Gamma_i \Lambda_i \Gamma_i^T).$$

The results are important because it allows to find values of standard information criteria which may be used to detect departures from normality, as shown in the simulation study. Besides, a closing form of this likelihood may be used to carry out classical inference with standard optimization techniques. For instance, in the R software, the *mvtnorm* package and the *optim* routine.

## References

1. Zhang D, Davidian M (2001) Linear mixed models with flexible distributions of random effects for longitudinal data. *Biometrics* 57(3):795–802. [10.1111/j.0006-341X.2001.00795x](https://doi.org/10.1111/j.0006-341X.2001.00795x)
2. Ghidry W, Lesaffre E, Eilers P (2004) Smooth random effects distribution in a linear mixed model. *Biometrics* 60(4):945–953. [10.1111/j.0006-341X.2004.00250x](https://doi.org/10.1111/j.0006-341X.2004.00250x)
3. Frees EW, Young VR, Luo Yu (2001) Case studies using panel data models. *North Am Actuarial J* 5(4):24–42. [10.1080/10920277.2001.10596010](https://doi.org/10.1080/10920277.2001.10596010)
4. Ahn S, Kim JHT, Ramaswami V (2012) A new class of models for heavy tailed distributions in finance and insurance risk. *Insurance: Math Econ* 51(1):43–52. [10.1016/j.insmatheco.2012.02.002](https://doi.org/10.1016/j.insmatheco.2012.02.002)
5. Huang Y, Meng S (2020) A Bayesian nonparametric model and its application in insurance loss prediction. *Insurance: Math Econ* 93:84–94. [10.1016/j.insmatheco.2020.04.010](https://doi.org/10.1016/j.insmatheco.2020.04.010)
6. Escobar DD, Pflug GCh (2020) The distortion principle for insurance pricing: properties, identification and robustness. *Ann Oper Res* 292(2):771–794. [10.1007/s10479-018-3119-1](https://doi.org/10.1007/s10479-018-3119-1)
7. Azzalini A, Capitanio A (1999) Statistical applications of the multivariate skew-normal distribution. *J R Stat Soc* 61(3):579–602. [10.1111/1467-9868.00194](https://doi.org/10.1111/1467-9868.00194)
8. Capitanio A, Azzalini A, Stanghellini E (2003) Graphical models for skew-normal variates. *Scand J Stat* 30(1):129–144. [10.1111/1467-9469.00322](https://doi.org/10.1111/1467-9469.00322)
9. Sahu SK, Dey DK, Branco MD (2003) A new class of multivariate skew distributions with applications to Bayesian regression models. *Can J Stat* 31(2):22
10. Arellano-Valle RB, Genton MG (2005) On fundamental skew distributions. *J Multivar Anal* 96(1):93–116. [10.1016/j.jmva.2004.10.002](https://doi.org/10.1016/j.jmva.2004.10.002)
11. Huang Y, Chen R, Dagne G, Zhu Y, Chen H (2015) Bayesian bivariate linear mixed-effects models with skew-normal/independent distributions, with application to aids clinical studies. *J Biopharm Stat* 25(3):373–396. [10.1080/10543406.2014.920660](https://doi.org/10.1080/10543406.2014.920660)
12. Jung D, Lee J (2017) New composite distributions for insurance claim sizes. *Kor J Appl Stat* 30(3):363–376. [10.5351/KJAS.2017.30.3.363](https://doi.org/10.5351/KJAS.2017.30.3.363)
13. Bandyopadhyay D, Lachos VH, Castro LM, Dey DK (2012) Skew-normal/independent linear mixed models for censored responses with applications to HIV viral loads: skew normal/independent censored linear mixed models. *Biom J* 54(3):405–425. [10.1002/bimj.201000173](https://doi.org/10.1002/bimj.201000173)

14. Huang Y, Chen J (2016) Bayesian quantile regression-based nonlinear mixed-effects joint models for time-to-event and longitudinal data with multiple features: quantile regression joint models for event time-longitudinal data. *Stat Med* 35(30):5666–5685. [10.1002/sim.7092](https://doi.org/10.1002/sim.7092)
15. Leisen F, Miguel Marin J, Villa C (2017) Objective Bayesian modelling of insurance risks with the skewed student- t distribution: F Leisen, J. Miguel Marin and C. Villa. *Appl Stochastic Models Bus Ind* . [10.1002/asmb.2227](https://doi.org/10.1002/asmb.2227)
16. De Simone A, Piangerelli M (2020) A Bayesian approach for monitoring epidemics in presence of undetected cases. *Chaos, Solitons and Fractals* 140:110167. [10.1016/j.chaos.2020.110167](https://doi.org/10.1016/j.chaos.2020.110167)
17. Jara A, Quintana F, Martín ES (2008) Linear mixed models with skew-elliptical distributions: a Bayesian approach. *Comput Stat Data Anal* 52(11):5033–5045. [10.1016/j.csda.2008.04.027](https://doi.org/10.1016/j.csda.2008.04.027)
18. Scollnik DPM (1993) A Bayesian analysis of a simultaneous equations model for insurance rate-making. *Insurance: Math Econ* 12(3):265–286. [10.1016/0167-6687\(93\)90238-K](https://doi.org/10.1016/0167-6687(93)90238-K)
19. Bermúdez Bermúdez L, Karlis D (2011) Bayesian multivariate Poisson models for insurance ratemaking. *Insurance: Math Econ* 48(2):226–236. [10.1016/j.insmatheco.2010.11.001](https://doi.org/10.1016/j.insmatheco.2010.11.001)
20. Esmaeili H, Klüppelberg C (2010) Parameter Estimation of a bivariate compound poisson process. *Insurance: Math Econ* 47(2):224–233. [10.1016/j.insmatheco.2010.04.005](https://doi.org/10.1016/j.insmatheco.2010.04.005)
21. Moreno-Jiménez JM, Salvador M, Gargallo P, Altuzarra A (2016) systemic decision making in AHP: A Bayesian approach. *Ann Oper Res* 245(1–2):261–284. [10.1007/s10479-014-1637-z](https://doi.org/10.1007/s10479-014-1637-z)
22. Klugman SA (1992) *Bayesian statistics in actuarial science: with emphasis on credibility*. Kluwer Academic Publishers, Boston
23. Arellano-Valle RB, Bolfarine H, Lachos VH (2007) Bayesian inference for skew-normal linear mixed models. *J Appl Stat* 34(6):663–682. [10.1080/02664760701236905](https://doi.org/10.1080/02664760701236905)
24. Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. *Stat Sci* 7(4):454–472. [10.1214/ss/1177011136](https://doi.org/10.1214/ss/1177011136)
25. China Statistics Bureau (INP) 2011–2020. Population figures. <http://www.stats.gov.cn/tjsj/ndsj/>. Accessed 1 March 2022
26. Bernardi M, Maruotti A, Petrella L (2012) Skew mixture models for loss distributions: a Bayesian approach. *Insurance: Math Econ* 51(3):617–623. [10.1016/j.insmatheco.2012.08.002](https://doi.org/10.1016/j.insmatheco.2012.08.002)
27. Miljkovic T, Grün B (2016) Modeling loss data using mixtures of distributions. *Insurance Math Econ* 70:387–396. [10.1016/j.insmatheco.2016.06.019](https://doi.org/10.1016/j.insmatheco.2016.06.019)
28. Laird NM, Ware JH (1982) Random-effects models for longitudinal data. *Biometrics* 38(4):963. [10.2307/2529876](https://doi.org/10.2307/2529876)
29. Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A (2002) Bayesian measures of model complexity and fit. *J R Stat Soc B* 64(4):583–639. [10.1111/1467-9868.00353](https://doi.org/10.1111/1467-9868.00353)
30. Yuan L, Ya Yan Lu (2021) Conditional robustness of propagating bound states in the continuum in structures with two-dimensional periodicity. *Phys Rev A* 103(4):043507. [10.1103/PhysRevA.103.043507](https://doi.org/10.1103/PhysRevA.103.043507)
31. Pinheiro A , Vidakovic B (1997) Estimating the square root of a density via compactly supported wavelets. *Comput Stat Data Anal* 25. [https://doi.org/10.1016/S0167-9473\(97\)00013-3](https://doi.org/10.1016/S0167-9473(97)00013-3)