

Article

Bayesian Predictive Analysis of Natural Disaster Losses

Min Deng , Mostafa Aminzadeh  and Min Ji *

Department of Mathematics, Towson University, Towson, MD 21252, USA; mdeng@towson.edu (M.D.); maminzadeh@towson.edu (M.A.)

* Correspondence: mji@towson.edu

Abstract: Different types of natural events hit the United States every year. The data of natural hazards from 1900 to 2016 in the US shows that there is an increasing trend in annual natural disaster losses after 1980. Climate change is recognized as one of the factors causing this trend, and predictive analysis of natural losses becomes important in loss prediction and risk prevention as this trend continues. In this paper, we convert natural disaster losses to the year 2016 dollars using yearly average Consumers Price Index (CPI), and conduct several tests to verify that the CPI adjusted amounts of loss from individual natural disasters are independent and identically distributed. Based on these test results, we use various model selection quantities to find the best model for the natural loss severity among three composite distributions, namely Exponential-Pareto, Inverse Gamma-Pareto, and Lognormal-Pareto. These composite distributions model piecewise small losses with high frequency and large losses with low frequency. Remarkably, we make the first attempt to derive analytical Bayesian estimate of the Lognormal-Pareto distribution based on the selected priors, and show that the Lognormal-Pareto distribution outperforms the other two composite distributions in modeling natural disaster losses. Important risk measures for natural disasters are thereafter derived and discussed.

Keywords: composite distributions; predictive analysis; bayesian inference; natural disaster; climate change; risk measures



Citation: Deng, Min, Mostafa Aminzadeh, and Min Ji. 2021. Bayesian Predictive Analysis of Natural Disaster Losses. *Risks* 9: 12. <https://doi.org/10.3390/risks9010012>

Received: 20 November 2020

Accepted: 18 December 2020

Published: 2 January 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Different types of natural events hit the United States (US) every year. The east coast of the US suffers hurricanes, the middle of the US sees tornadoes, the west coast of the US endures earthquake, and the south of the US bears a variety of issues such as hurricane, wind, drought, and floods¹. The data of the occurrence and damage of natural events from 1900 to present in the US, from the Emergency Events Database (EM-DAT)—International Disaster Database², showed that every year's natural disaster losses after 1980 have been increasing. Climate change is recognized as one of the contributors to this increasing natural losses. Injury, homelessness, displacement, and economic losses from natural events can have significant impact on populations and societies. Predictive analysis of future natural events is imperative to provide important information for prevention and remedy plans to reduce the human impact and economic losses from natural disasters.

Various research has attempted to model the frequencies and damages due to the natural events. Levi and Partrat (1991) analyzed hurricane losses between 1954 and 1986 in the US, and found that the amounts of losses were independent and identically distributed (i.i.d.) and independent of the frequencies of hurricanes. These assumptions are confirmed in our research, based on the EM-DAT natural disaster data from 1900 to 2016 after taking into account price inflations. The number of losses in different years is converted into 2016

¹ See <http://www.nytimes.com/interactive/2011/05/01/weekinreview/01safe.html>.

² https://www.emdat.be/emdat_db/.

dollars using the annual average Consumers Price Index (CPI)³. We found that both the numbers of natural events and the CPI adjusted total damages have been increasing over the time; however there is no visible trend in the individual CPI adjusted losses, which is consistent with the conclusion claimed by [Levi and Partrat \(1991\)](#).

To discover the distribution features of individual losses from natural disasters, we use the most recent 37 years of the CPI adjusted amounts of natural losses to investigate appropriate parametric distribution models for the amount of loss caused by individual natural event, i.e., natural loss severity. [Levi and Partrat \(1991\)](#) proposed to use lognormal distribution for the losses of natural events based on the hurricane data between 1954 and 1986 in the US. However, this distribution cannot describe the typical feature of the natural disaster losses, i.e., many small amounts of losses and occasional occurrence of huge amount of losses. Some researchers argued to use composite distributions to capture this recognized feature.

A composite distribution has a threshold for the support of the loss random variable, with the Pareto distribution usually used to model a large loss beyond the threshold amount and popular parametric continuous distributions used for a small amount of loss below the threshold. For example, [Bakar et al. \(2015\)](#) developed several Weibull distribution-based composite models for heavy tailed insurance loss data. In our paper, we compare the performance of three composite distributions, namely Exponential-Pareto (Exp-Pareto), Inverse Gamma-Pareto (IG-Pareto), and Lognormal-Pareto (LN-Pareto), and the corresponding three non-composite parametric models, based on the negative loglikelihood, the Akaike information criterion, and Bayesian information criterion. All the three model selection criteria show that the composite distributions fit the data better than the non-composite models.

We thereafter apply the three composite models to the data and select the best fitted composite model for the loss severity of natural events. For consistent model comparison, we make the first attempt to derive analytic Bayesian estimation of the LN-Pareto composite distribution, because [Aminzadeh and Deng \(2018, 2019\)](#) has analytically derived Bayesian estimators for Exp-Pareto and IG-Pareto. In addition, Bayesian method enables us to perform predictive analysis of future losses of natural disasters using the best fitted composite model.

[Cooray and Cheng \(2015\)](#) developed Bayesian estimators for the LN-Pareto composite model based on a Markov Chain Monte Carlo (MCMC) simulation algorithm. In our paper, we carefully select prior distributions for the LN-Pareto composite model and derive a closed form Bayesian estimators for the unknown parameters of the LN-Pareto distribution. The mean squared errors (MSE) of the Bayesian estimation method and the maximum likelihood (ML) estimation method show that Bayesian estimation method outperform the ML estimation method in estimating these three composite models.

The comparison results show that the LN-Pareto is the best among these three composite models for the loss severity of natural disasters. Various risk measures of natural event losses are thereafter presented based on the LN-Pareto distribution. Same risk measures based on the other two composite distributions of natural losses are also provided for comparison.

The remainder of this paper is organized as follows. Section 2 describes the natural disaster data and tests the assumption that individual loss amounts are independent and identically distributed. Section 3 introduces three composite distributions and compare these models with the corresponding non-composite models. Section 4 derives Bayesian estimators for the LN-Pareto composite distribution and demonstrates the performance of Bayesian estimation method versus the ML estimation method. Section 5 presents risk measures of the future natural disaster losses based on the LN-Pareto model and the other two composite distributions. The concluding remarks are given in Section 6.

³ The CPI data was download from Bureau of Labor Statistics <https://data.bls.gov/pdq/SurveyOutputServlet>.

2. Natural Losses in the US

2.1. The Data

The EM-DAT contains worldwide data on the occurrence and impact of natural events from 1900 to the present day. The database is compiled from various sources, including United Nation agencies, non-governmental organizations, insurance companies, research institutes and press agencies. There were a total of 258 natural events in the US from 1900 to 2016, among which 156 events happened between 1980 to 2016, accounting for 60% of the total events during this one third of the whole time period. This observation is consistent with the fact that Earth's climate is changing faster than at any point in the history as a result of human activities.

We also look into the trend of the natural loss amounts from 1900 to 2016. To eliminate the effect of price inflation, we convert the amounts of losses in each year into the year 2016 dollars. Let CPI_t be the annual average CPI in year t and y_t be the CPI adjusted amount of losses in that year, then we have

$$y_t = \text{loss amount in year } t \times \frac{CPI_{2016}}{CPI_t}, \text{ for } t = 1900, 1991, \dots, 2016.$$

After price inflation effect being adjusted, the amount of losses during 1980 to 2016 accounts for 85.3% of the total losses. Figure 1 shows the yearly number of natural events and the CPI adjusted (in 2016 dollars) total damage costs from 1900 to 2016 in the US. We can see that both the number of natural events and the CPI adjust damage costs in each year have been increasing over these years, especially after 1980.

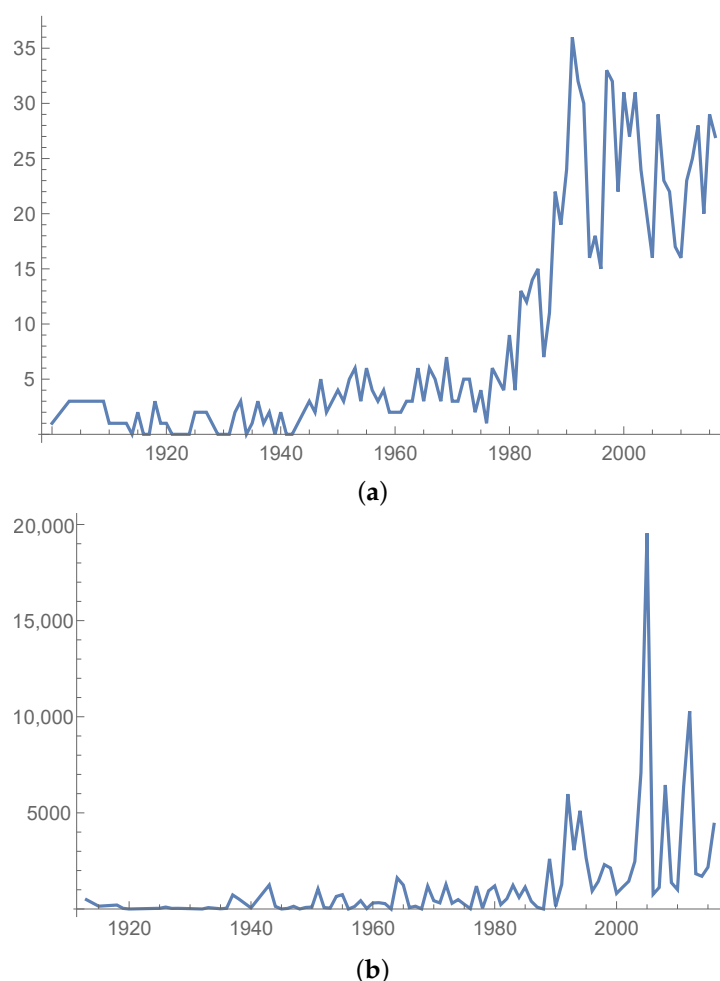


Figure 1. The number and damage costs of natural events from 1900 to 2016 in the US. (a) The number of natural events; (b) Total damage in 2016 dollars ('0000000 US\$).

To demonstrate the change in both number and damage costs of natural events caused by human activities, we break down the data to be before and after 1980. Figure 2 displays the percentage of the number of occurrence to the total number of occurrence and the percentage of the CPI adjusted damage costs to the total losses by different time periods (before and after 1980) and by different types of natural disasters. The exact numbers for this figure are listed in Tables A1 and A2 in Appendix A.

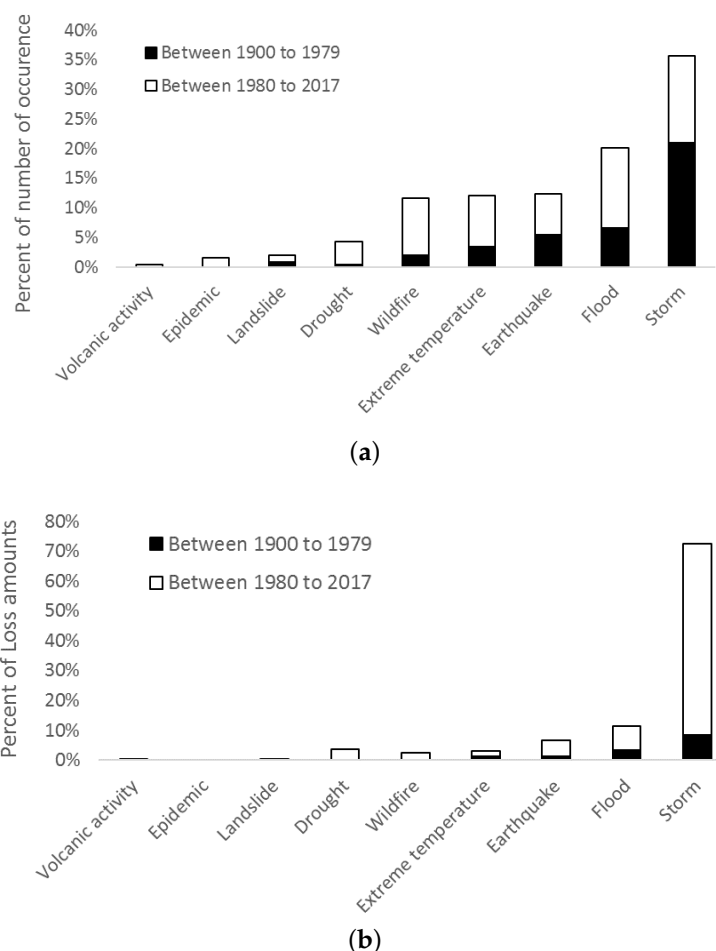


Figure 2. The percentage of the number of occurrence and damage costs of natural disasters to the total occurrence and damage losses by different types of natural events before and after 1980 in the US. (a) Percent of the number of occurrence; (b) Percent of the damage costs.

We can see that storms and floods account for most of natural events. The damage costs from natural events after 1980 account for the majority of the total costs, among which the damage costs caused by storms after 1980 account for more than half. In addition, these storm losses were caused by fewer storms after 1980 than before 1980, indicating higher average storm losses after 1980 than before 1980. However, this is not the case for other types of natural events.

Although there is an obvious increasing trend in both the numbers and the total damage costs of natural events after 1980, it is not trivial to find out the pattern in the economic losses of individual natural events. Based on the aforementioned features of the natural loss data, we use the natural losses in the most recent 37 years from 1980 to 2016. There are a total number of 462 natural events in these 37 years, and only seven events caused economic loss above 2×10^7 (in 2016 dollars). In Figure 3, we depict the scatter plot of those losses below 2×10^7 , and we can see that there is no identified trend in individual natural disaster losses.

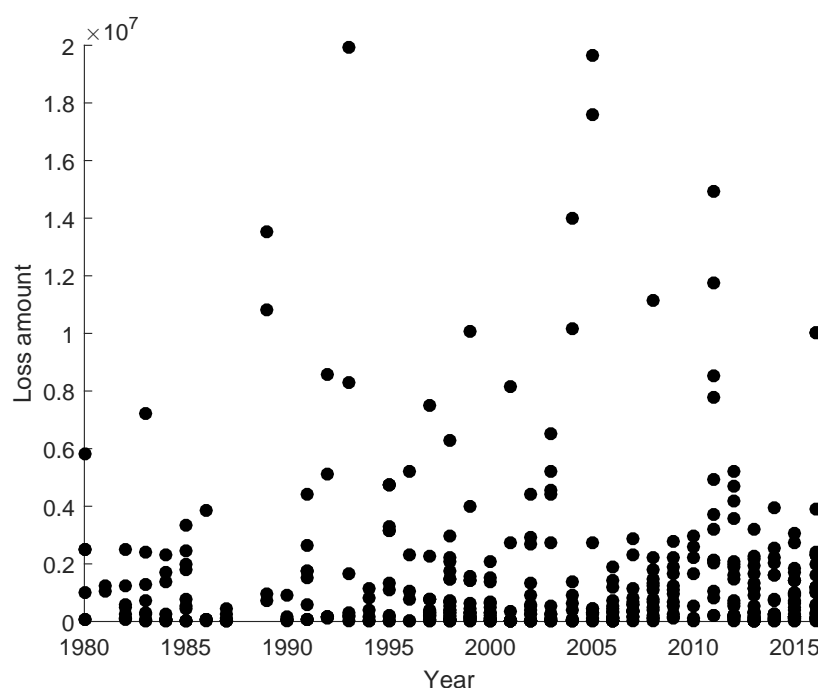


Figure 3. The scatter plot of individual natural disaster losses after 1980 in the US.

We are interested in how to appropriately model individual natural disaster losses. Motivated by this research question and the above observations, in this paper we aim to investigate the features of the severity of natural events, attempt to find appropriate loss severity models for natural disasters, and provide the predictive modeling of loss severity for risk management and insurance arrangement.

2.2. The i.i.d. Assumption for Loss Severity

Let the random variable N denote the number of occurrences per year and X_i , for $i = 1, 2, \dots, N$, be the severity random variable of the i th occurrence of natural events occurring in each year. The vector $\vec{X} = (X_1, X_2, \dots, X_N)^T$ denotes the random vector of yearly loss severity. We use the natural disaster data between 1980 to 2016, and the random variable N has a sample size of 37. Table A3 in Appendix B lists the CPI adjusted individual natural damage losses in these 37 years. It can easily be seen that random variable N takes one of the values $0, 2, 6, 8, \dots, 28$ over the 37 years. One value of N is zero, because there was no record of natural events in year 1988. In all other years, there were at least two natural events every year. Therefore, there are 36 non-zero realizations $(n_k, \vec{x}^k)_{k=1,2,\dots,36}$ of (N, \vec{X}) . The i th column in the matrix $(\vec{x}^1, \vec{x}^2, \dots, \vec{x}^{36})^T$ are the realizations of X_i , i.e., the i th occurrence of natural disasters in each of 36 years.

Let m_i be the sample size of the i th random variable X_i . From Table A3 we also see that X_1 , the 1st severity random variable, has 36 non-zero values and therefore the sample size $m_1 = 36$. The realization of X_1 is $(x_1^1, x_1^2, \dots, x_1^{36})^T$, which takes the value $(1019.45, 1056.14, \dots, 550.00)$ million in 2016 dollars. X_{28} , the 28th severity random variable with $m_{28} = 1$ and takes only one non-zero value 1822.71 millions in 2016 dollars, because there is only one year with 28 losses. The total number of the observation is $\sum_{k=1}^{28} m_k = 462$.

We want to test the general assumption that X_1, X_2, \dots, X_{28} are independent and identically distributed random variables. Since $X_{21}, X_{22}, \dots, X_{28}$ have small sample sizes, we group them together as one random variable X_{21} . Applying the nonparametric methods, the Kendall Tau test and the Spearman test (see Gibbons and Chakraborti 2003), we test the independence assumption for X_1, X_2, \dots, X_{21} as follows.

For $X_i, X_j, i \neq j$, where $i, j = 1, 2, \dots, 21$, the null hypotheses and alternative hypotheses are

$$H_0 : \rho_{X_i, X_j} = 0, \quad H_a : \rho_{X_i, X_j} \neq 0$$

Let $m = \min(m_i, m_j)$. The Kendall Tau statistic is

$$T = \frac{1}{\binom{m}{2}} \sum_{1 \leq k \leq l \leq m} A_{kl},$$

$$\text{where } A_{kl} = \begin{cases} 1 & (x_i^k - x_i^l)(x_j^k - x_j^l) > 0 \\ 0 & (x_i^k - x_i^l)(x_j^k - x_j^l) = 0 \\ -1 & (x_i^k - x_i^l)(x_j^k - x_j^l) < 0 \end{cases}$$

Let R_k and S_k be the rank of x_i^k and x_j^k among m observations respectively, for $k = 1, 2, \dots, m$. The Spearman rho test statistic is

$$R = \frac{\sum_{k=1}^m (R_k - \bar{R})(S_k - \bar{S})}{\sqrt{\sum_{k=1}^m (R_k - \bar{R})^2 \sum_{k=1}^m (S_k - \bar{S})^2}},$$

where \bar{R} and \bar{S} are the average of R_k and S_k respectively.

For X_1, X_2, \dots, X_{21} , there are total of $\binom{21}{2} = 210$ tests. We obtain 210 values of all the Kendall Tau test statistic and Spearman test statistic with the corresponding p -values. Table 1 lists the pairs of severity random variables with the null hypotheses being rejected at the significance level of 1%. All other pairs have non-significant test results and we fail to reject the null hypotheses at the significance level of 1%. Based on these test results, it is reasonable to assume that X_1, X_2, \dots are independent.

Table 1. Pairs of severity losses with significant Kendall and Spearman test results.

Pairs	Kendall Tau Statistics (p -Value)	Spearman Statistics (p -Value)
X_1, X_5	−0.373 (0.002)	−0.546 (0.001)
X_7, X_{21}	0.393 (0.0051)	0.566 (0.003)
X_9, X_{16}	−0.575 (0.009)	−0.769 (0.0034)

Next, the Kruskal-Wallis test is used to verify the identical distribution assumption. The details of this test are introduced by [Gibbons and Chakraborti \(2003\)](#). Let F_i be the cumulative distribution function of $X_i, i = 1, 2, \dots, 21$. In our Kruskal-Wallis test, the null hypotheses and alternative hypotheses are

$$H_0 : F_1 = F_2 = \dots = F_{21}, \quad H_a : F_i \neq F_j, \text{ for some } i \neq j.$$

There are a total of 462 observations. Sort all the 462 observations in an increasing order. If there are more than one observations with an identical value, then the median is assigned as the rank for these observations. Define T_i to be the sum of the ranks of all the observations of X_i , where $i = 1, 2, \dots, 21$. We have $T_1 = 8817, T_2 = 8444, \dots$, and $T_{21} = 6656.5$. It is clear that $\sum_{i=1}^{21} T_i = 462(462 + 1)/2 = 106,854$.

Under the null hypotheses, $E(T_i) = m_i(462 + 1)/2$. Therefore, $E(T_1) = 8334, E(T_2) = 88,334, \dots$, and $E(T_{21}) = 6019$. The Kruskal-Wallis Statistic is

$$KW = \frac{12}{462(462 + 1)} \sum_{i=1}^{21} \frac{1}{m_i} (T_i - E(T_i))^2 = 19.4274$$

Asymptotically, KW has Chi-Square distribution with the degree of freedom 20, with the p -value 0.494. We fail to reject H_0 at the significance level 1%. Therefore, it is reasonable to assume that X_1, X_2, \dots are identically distributed.

2.3. Non-Parametric Distribution of Loss Severity

Based on the test results, we can reasonably assume that the CPI adjusted natural disaster damage random variable is independent and identically distributed. Let X denote the CPI adjusted damage random variable with unknown distribution $F(x)$. The 462 individual damage losses are the realizations of the loss severity random variable X . As we explore an appropriate parametric distribution $F(x)$ for the natural disaster loss severity, we first look at the non-parametric distribution of X , which is also called a data dependent distribution.

Let $y_1 < y_2 < \dots < y_k$ be the k unique loss values, and s_i be the number of times the observations y_i appears in the sample. Let $r_j = \sum_{i=j}^k s_i$ be the number of observations greater than or equal to y_j . The Nelson-Aalen estimator of the cumulative hazard rate function is

$$\hat{H}(x) = \begin{cases} 0 & x < y_1 \\ \sum_{i=1}^{j-1} \frac{s_i}{r_i} & y_{j-1} \leq x < y_j, j = 2, \dots, k \\ \sum_{i=1}^k \frac{s_i}{r_i} & x \geq y_k \end{cases}$$

Therefore, $\hat{S}(x) = \exp(-\hat{H}(x))$ and $\hat{F}(x) = 1 - \hat{S}(x) = 1 - \exp(-\hat{H}(x))$. The Kolmogorov-Smirnov (K-S) confidence band of unknown distribution $F(x)$ can also be constructed based on the Nelson Aalen estimate of distribution $\hat{F}(x)$. Define the K-S statistics by $D_n = \sup_x (|\hat{F}(x) - F(x)|)$, where n is sample size. To form $100(1 - \alpha)\%$ confidence band, we select a number d such that $P(D_n \geq d) = \alpha$. Then, the lower band is $F_L(x) = \max(\hat{F}(x) - d, 0)$ and the upper band is $F_U(x) = \min(\hat{F}(x) + d, 1)$. We set $\alpha = 0.05$ and the true unknown distribution of loss lies between $F_L(x)$ and $F_U(x)$ with 95% confidence.

The plots of $\hat{F}(x)$, $F_L(x)$, and $F_U(x)$ is given in Panel (a) of Figure 4. We also display the histogram of the CPI adjusted severity of the natural events in Panel (b) of Figure 4. We can see that the histogram of the CPI adjusted severity of the natural event losses is both skewed and fat-tailed, which shows the typical feature that there are many small losses and a few very large losses.

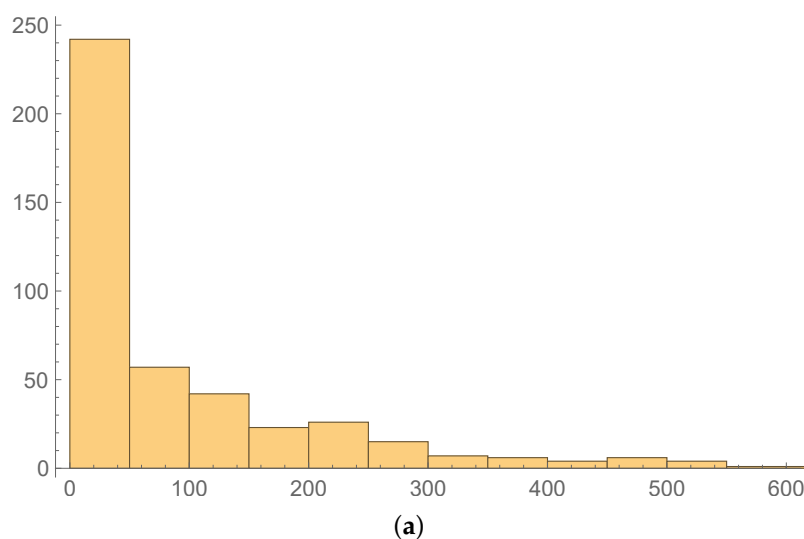


Figure 4. Cont.

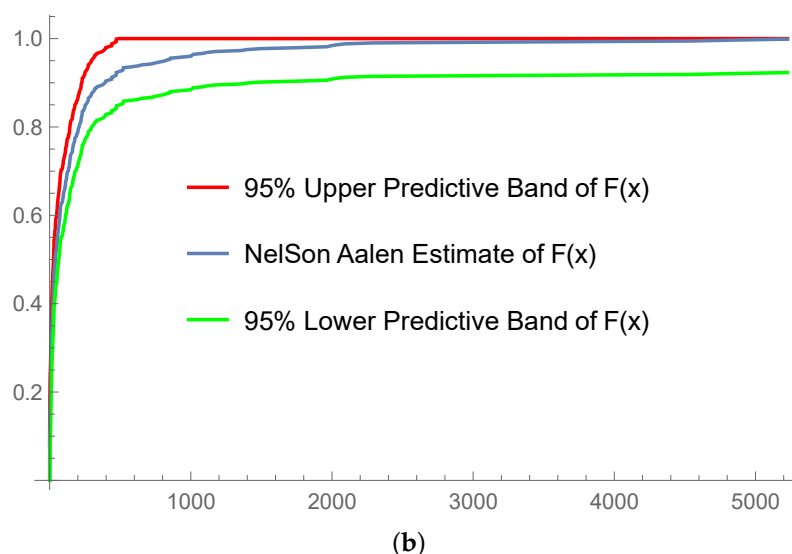


Figure 4. The histogram and the empirical distribution with 95% confidence bands of the CPI adjusted natural losses from 1980 to 2016 in the US. (a) The histogram of the CPI adjusted natural losses; (b) The empirical distribution with 95% confidence bands of the CPI adjusted natural losses.

3. Composite Models

We confirmed that the natural event severity data has the typical feature of high frequency of small losses and low frequency of large losses; while the traditional distributions such as the normal, exponential, inverse-gamma, and log-normal distributions, are not able to describe this feature. Some researchers have addressed this feature of insurance data by using composite models for insurance losses. The probability density function of a composite model consists of two distributions, namely probability density functions $f_1(x)$ and $f_2(x)$. The general composite model has the probability density function as follows.

$$f_X(x) = \begin{cases} cf_1(x) & 0 < x \leq \theta \\ cf_2(x) & \theta \leq x < \infty \end{cases},$$

where c is a normalized constant and θ is the parameter that represents the threshold of the supports for the two distributions. In order to make the composite density function smooth, it is usually assumed that the pdf $f_X(x)$ is continuous and differentiable at θ , that is $f_1(\theta) = f_2(\theta)$ and $f_1'(\theta) = f_2'(\theta)$.

3.1. Three Composite Distributions

Cooray and Ananda (2005) introduced a two-parameter continuous and differentiable composite LN-Pareto model, which is a two-parameter lognormal density up to an unknown threshold value and a two-parameter Pareto density for the rest of the model. The resulting density is similar in shape to a lognormal density with the tail behavior quite similar to a Pareto density. They applied the proposed composite model to a fire insurance data to show the importance of the proposed composite LN-Pareto distribution in describing insurance claim data.

Motivated by Cooray and Ananda (2005), Teodorescu and Vernic (2006) introduced a composite Exp-Pareto distribution, which is an exponential density up to an unknown threshold value and a two parameter Pareto density for the rest of the model. The model is reduced to a one-parameter distribution after satisfying the continuous and differentiable condition. Aminzadeh and Deng (2019) proposed an IG-Pareto composite model. Under the general assumption of smoothness and continuity, the IG-Pareto model is reduced to be a one parameter model. Let $\Phi(\cdot)$ be the cumulative distribution function (cdf) of the standard normal distribution. Table 2 summarizes the distribution density function of Exp-

Pareto, LN-Pareto, and IG-Pareto composite models, and Figure 5 plots three distribution functions with various parameter values.

Table 2. The distribution density functions of three composite models.

	$f_1(x)$	$f_2(x)$	The Composite pdf $f(x)$
Exp-Pareto	$f_1(x) = \lambda e^{-\lambda x},$ $x > 0, \lambda > 0$	$f_2(x) = \frac{\alpha \theta^\alpha}{x^{\alpha+1}},$ $x \geq \theta > 0, \alpha > 0$	$f(x) = \begin{cases} \frac{0.775}{\theta} e^{\frac{-1.35x}{\theta}}, & 0 < x \leq \theta \\ \frac{0.2\theta^{0.35}}{x^{1.35}}, & \theta \leq x < \infty \end{cases}$
IG-Pareto	$f_1(x) = \frac{\beta^\alpha x^{-\alpha-1} e^{-\beta/x}}{\Gamma(\alpha)},$ $x > 0, \alpha > 0, \beta > 0$	$f_2(x) = \frac{a\theta^a}{x^{a+1}},$ $x \geq \theta, a > 0, \theta > 0$	$f(x) = \begin{cases} \frac{c(k\theta)^\alpha x^{-\alpha-1} e^{-\frac{k\theta}{x}}}{\Gamma(\alpha)}, & 0 < x \leq \theta \\ \frac{c(\alpha-k)\theta^{\alpha-k}}{x^{\alpha-k+1}}, & \theta \leq x < \infty \end{cases}$ where $\alpha = 0.308289, k = 0.144351, c = 0.711384$
LN-Pareto	$f_1(x) = \frac{e^{-\frac{1}{2}(\frac{\ln x - \mu}{\sigma})^2}}{x\sigma\sqrt{2\pi}},$ $x > 0, \sigma > 0$	$f_2(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}},$ $x \geq \theta, \alpha > 0, \theta > 0$	$f(x) = \begin{cases} \frac{\beta\theta^\beta e^{-0.5(\frac{\beta}{k})^2 \ln^2(x/\theta)}}{(1 + \Phi(k))x^{\beta+1}}, & 0 < x \leq \theta \\ \frac{\beta\theta^\beta}{(1 + \Phi(k))x^{\beta+1}}, & \theta \leq x < \infty \end{cases}$ where $k = 0.372238898$

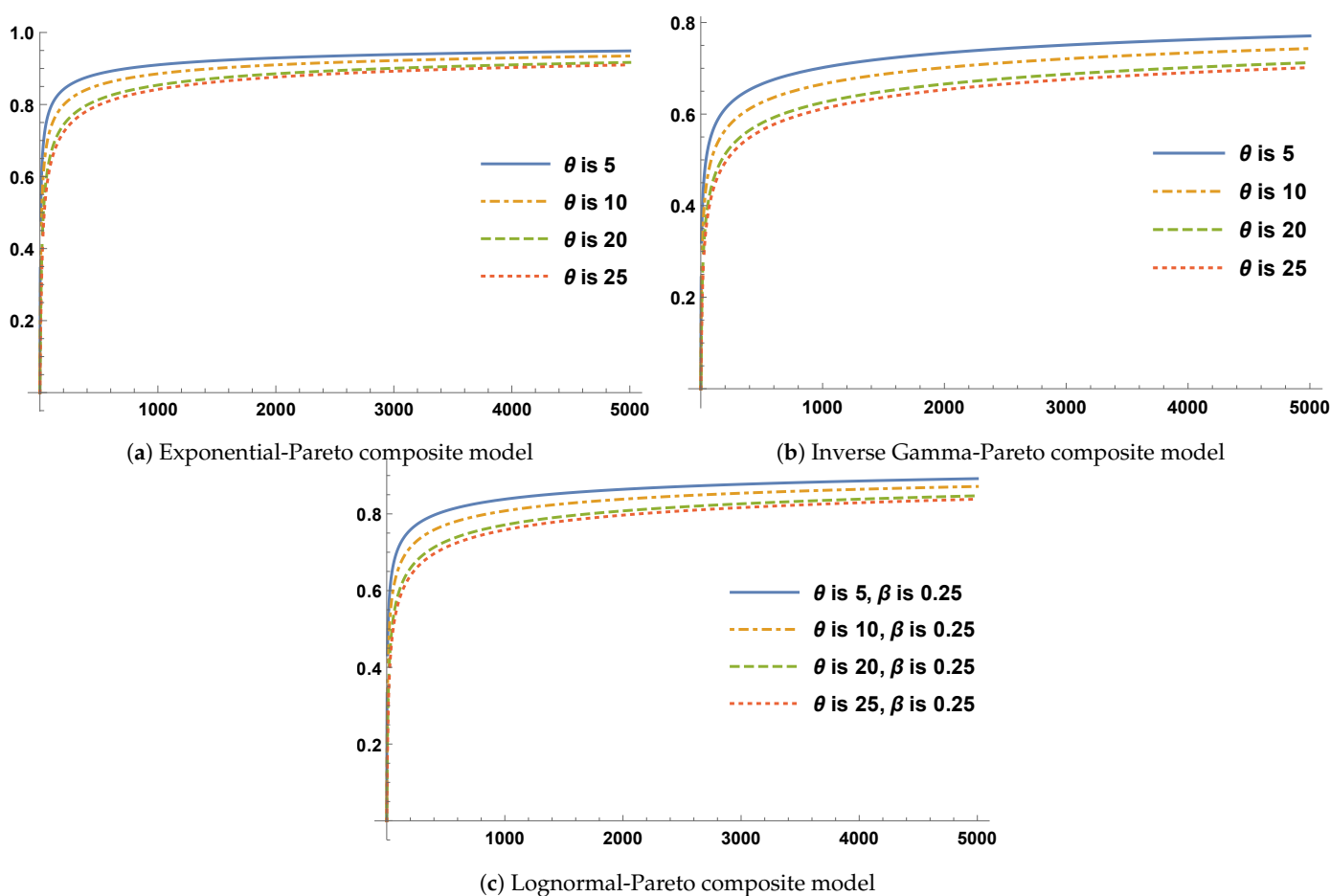


Figure 5. The cumulative distribution functions of three composite models with various parameter values.

We can see that the Nelson Aalen estimate of the unknown distribution of the natural disaster severity loss random variable is quite similar to the distributions of three composite models. This indicates that composite models might be able to describe the features of the natural disaster damage losses. In the next section, we will compare the performance of

these three composite distributions in modeling the natural disaster loss severity based on standard model comparison and selection criteria.

3.2. Model Selection for Loss Severity

The maximum likelihood estimators for the unknown parameters of the Exp-Pareto, IG-Pareto, and Lognormal-Pareto composite models have been derived by Teodorescu and Vernic (2006), Aminzadeh and Deng (2019), and Cooray and Ananda (2005) respectively. We know that the parameter θ is the unknown threshold value dividing the domain of the two distributions of a composite model, and the maximum likelihood functions changes when the value of θ changes; therefore, a grid search method has to be used to find the ML estimates. The grid search algorithm can be briefly summarized as follows.

1. Sort the sample of the natural disaster damage losses in an increasing order, i.e., $x_1 < x_2 < \dots < x_n$, where n is the sample size. Let n^* be the size of the partial sample of the first n^* losses x_1, x_2, \dots, x_{n^*} . Start from $n^* = 1$.
2. Compute the maximum likelihood estimates $\hat{\theta}$ and $\hat{\beta}$ as in Table 3 for the given n^* . If $\hat{\theta}$ is in between $x_{n^*} \leq \hat{\theta} \leq x_{n^*+1}$, we found n^* ; otherwise, increase n^* by 1.
3. Repeat Step 2 for $n^* = 2, 3, \dots$, till $x_{n^*} \leq \hat{\theta} \leq x_{n^*+1}$. The ML estimates of the parameters are found based on the correct n^* .

In our research, Mathematica software is used to code the algorithm. Table 3 lists the ML estimates of three composite distributions.

Table 3. The ML estimates of the unknown parameter in three composite models.

Composite Model	ML Estimates of Parameters
Exp-Pareto	$\hat{\theta} = \frac{1.35n^* \bar{x}_{n^*}}{1.35n^* - 0.35n}, \text{ where } \bar{x}_{n^*} = \frac{1}{n^*} \sum_{i=1}^{n^*} x_i.$
IG-Pareto	$\hat{\theta} = \frac{n^* \alpha + (\alpha - k)n^*}{kS}, \text{ where } S = \sum_{i=1}^{n^*} x_i^{-1},$ $\alpha = 0.308289, k = 0.144351.$
LN-Pareto	If $n^* = 1,$ $\hat{\theta} = x_1 \left(\prod_{i=1}^{n^*} x_i / x_1 \right)^{k^2}, \hat{\beta} = n^* \left(\sum_{i=1}^{n^*} \ln(x_i / x_1) \right)^{-1}, \text{ where } k = 0.372238898;$ otherwise, $\hat{\theta} = \exp\left(\frac{nk^2}{n^* \hat{\beta}}\right) \left(\prod_{i=1}^{n^*} x_i \right)^{\frac{1}{n^*}}, \hat{\beta} = \frac{k^2 B + \sqrt{k^4 B^2 + 4n^* n k^2 A}}{2A},$ where $A = n^* \sum_{i=1}^{n^*} (\ln x_i)^2 - \left(\sum_{i=1}^{n^*} \ln x_i \right)^2, B = n \sum_{i=1}^{n^*} \ln x_i - n^* \sum_{i=1}^n \ln x_i.$

Based on the ML estimates, we use the negative log-likelihood (NLL) value, the Akaike's Information Criterion (AIC), the Bayesian Information Criterion (BIC) goodness-of-fit measures to compare the appropriateness of these three composite models in modeling the natural disaster severity.

NLL can be used to compare models with the same number of parameters. It is equivalent to the maximum value of the likelihood function and defined as

$$NLL = -\log L(x_1, x_2, \dots, x_n | \underline{\theta}),$$

where $\underline{\theta}$ is the vector of unknown parameters. The smaller the NLL value, the larger the value of the likelihood function, and the better the fitted model.

AIC was defined by Akaike (1973) as

$$AIC = -2 \log L(x_1, x_2, \dots, x_n | \underline{\theta}) + 2q,$$

where q is the number of unknown parameters. The smaller the AIC value, the better the fitted model. The first term $-2 \log L(x_1, x_2, \dots, x_n | \theta)$ will decrease when the number of unknown parameters increases and is offset by the value of $2q$, indicating a trade-off between the goodness-of-fit and the number of parameters.

BIC was developed by Schwarz (1978) and defined as

$$BIC = -2 \log L(x_1, x_2, \dots, x_n | \theta) + q \log(n),$$

where n is sample size and q is the number of parameters. BIC penalizes a large number of parameters and a big size of sample. The smaller the BIC, the better the fitted model. Interested readers are referred to Burnham and Anderson (2002) for more details about these model selection criteria.

Based on the sample of 462 natural disaster damage losses in the US from 1980 to 2016, the values of the NLL, AIC, and BIC and the maximum likelihood estimates of the three composite models are summarized in Table 4. For comparison, we also fit three non-composite parametric models, namely the Exponential, Lognormal, and Inverse Gamma distributions, to the natural disaster data.

Table 4. The ML estimates of three composite models and three non-composite models for the severity of natural events in the US.

Model	ML Estimates of Parameters	NLL	AIC	BIC
Exp-Pareto	$\hat{\theta} = 25.561$ $n^* = 189$	2698.02	5398.05	5402.19
IG-Pareto	$\hat{\theta} = 2.86262$ $n^* = 64$	2719.35	5440.7	5444.83
LN-Pareto	$\hat{\theta} = 20.94751406$ $\hat{\beta} = 0.22044516$ $n^* = 174$	2327.74	4659.49	4667.76
Exponential $X \sim \text{Exp}(\lambda)$	$\hat{\lambda} = 5.32 \times 10^{-3}$	2881.23	5764.47	5768.61
Inverse Gamma $X \sim \text{IG}(\alpha, \beta)$	$\hat{\alpha} = 0.254669$ $\hat{\beta} = 0.527914$	2813.29	5630.58	5638.85
Lognormal $X \sim \text{LN}(\mu, \sigma)$	$\hat{\mu} = 3.50972$ $\hat{\sigma} = 2.1659785$	3058.66	6121.31	6129.58

We can see that three composite models fit the natural disaster losses better than three corresponding non-composite models. This supports the claim that a composite model can describe the distribution features of insurance data and natural disaster losses. Among three composite models, the LN-Pareto fit the data better than the other two composite models, in terms of the NLL, AIC, and BIC values. Therefore, the LN-Pareto composite distribution is the best model to conduct Bayesian predictive analysis of the natural disaster losses in the US.

4. The Bayesian Estimate

4.1. Bayesian Estimator of LN-Pareto

There is no analytical Bayesian estimator of the LN-Pareto composite distribution in the current literature. Cooray and Cheng (2015) found the Bayesian estimation for LN-Pareto based on the MCMC method. In this paper, we use conjugate priors for the two parameters of the LN-Pareto and make the first attempt to derive a closed-form Bayesian estimator without the MCMC simulation.

Recall the LN-Pareto probability density function given by Cooray and Ananda (2005), as listed in Table 2,

$$f_X(x|\theta, \beta) = \begin{cases} \frac{\beta\theta^\beta}{(1 + \Phi(k))x^{\beta+1}} e^{-0.5(\frac{\beta}{k})^2 \ln^2(x/\theta)} & 0 < x \leq \theta \\ \frac{\beta\theta^\beta}{(1 + \Phi(k))x^{\beta+1}} & \theta \leq x < \infty \end{cases}, \quad (1)$$

where $k = 0.372238898$ and $\Phi(\cdot)$ is the cdf of the standard normal distribution.

We use $\text{Gamma}(a_1, b_1)$ to be the prior distributions of β and $\text{LN}(c_1, (\frac{k}{\beta})^2)$ to be the prior distribution of $\theta|\beta$, where a_1, b_1, c_1 are hyper parameters. The prior distributions have probability density functions as follows.

$$\begin{aligned} \rho(\beta) &\propto \frac{\beta^{a_1-1} e^{-\frac{\beta}{b_1}}}{\Gamma(a_1) b_1^{a_1}}, & a_1 > 0, b_1 > 0 \\ \rho(\theta|\beta) &\propto \frac{\beta}{k} e^{-\frac{1}{2(\frac{k}{\beta})^2} (\ln(\theta) - c_1)^2}, & c_1 > 0 \end{aligned}$$

Without loss of generality, we assume that $x_1 < x_2 < \dots < x_n$ is an ordered random sample from the LN-Pareto distribution. Given n^* , the size of the partial sample of the first n^* losses x_1, x_2, \dots, x_{n^*} , such that $x_{n^*} \leq \theta \leq x_{n^*+1}$. The likelihood function can be written as

$$L(\underline{x}|\theta\beta) \propto \frac{\beta^n \theta^{n\beta}}{(\prod_{i=1}^n x_i)^{\beta+1}} e^{-0.5(\frac{\beta}{k})^2 \sum_{i=1}^{n^*} (\ln(x_i) - \ln(\theta))^2}.$$

To find posterior distributions $\pi(\beta|\underline{x})$ and $\pi(\theta|\beta, \underline{x})$, we need the joint pdf $f(\underline{x}, \theta, \beta)$ of $(\underline{x}, \theta, \beta)$, which is obtained as

$$f(\underline{x}, \theta, \beta) = L(\underline{x}|\theta, \beta) \rho(\theta|\beta) \rho(\beta) \quad (2)$$

The joint distribution function in Equation (2) can be reduced to

$$f(\underline{x}, \theta, \beta) \propto \beta^{n+a_1} e^{-\frac{\beta(1+b_1P)}{b_1}} e^{(-0.5(\frac{\beta}{k})^2 A_2)} \theta^{n\beta-1} e^{-0.5(\frac{\beta}{k})^2 (n^*+1)(\ln(\theta) - A_1)^2}, \quad (3)$$

where $P = \sum_{i=1}^n \ln x_i$, $A_1 = \frac{\sum_{i=1}^{n^*} \ln(x_i) + c_1}{n^* + 1}$, and $A_2 = \sum_{i=1}^{n^*} \ln^2(x_i) + c_1^2 - (n^* + 1)A_1^2$.

From Equation (3), the posterior probability distribution functions can be obtained as

$$\pi(\beta|\underline{x}) \propto \beta^{n+a_1-1} e^{-\frac{\beta(1+b_1P)}{b_1}} e^{(-0.5(\frac{\beta}{k})^2 A_2)}, \quad (4)$$

and

$$\pi(\theta|\beta, \underline{x}) \propto \beta \theta^{n\beta-1} e^{-0.5 \frac{(\ln(\theta) - A_1)^2}{(\frac{k}{\beta\sqrt{n^*+1}})^2}}. \quad (5)$$

It is noted that the right hand side of Equation (5) is the kernel of a lognormal distribution with parameters A_1 and $\frac{k}{\beta\sqrt{n^*+1}}$.

Our next step is to find $E[\theta|\beta, \underline{x}]$, the expectation of the posterior distribution in Equation (5), i.e., the conditional Bayes estimate of θ under the squared error loss function.

We first need to find the normalizing constant C_1 for the probability density function in Equation (5). Let $\xi = \frac{k}{\beta\sqrt{n^*+1}}$, then

$$\begin{aligned}
 & \int_0^\infty C_1 \beta \theta^{n\beta-1} e^{-0.5 \frac{(\ln(\theta)-A_1)^2}{\left(\frac{k}{\beta\sqrt{n^*+1}}\right)^2}} d\theta \\
 &= C_1 \int_0^\infty \frac{k}{\xi\sqrt{n^*+1}} \theta^{n\beta-1} e^{-0.5 \frac{(\ln(\theta)-A_1)^2}{\xi^2}} d\theta \\
 &= \frac{C_1 k}{\sqrt{n^*+1}} \sqrt{2\pi} \int_0^\infty \theta^{n\beta} \frac{1}{\theta \xi \sqrt{2\pi}} e^{-0.5 \frac{(\ln(\theta)-A_1)^2}{\xi^2}} d\theta \quad (6) \\
 &= \frac{C_1 k}{\sqrt{n^*+1}} \sqrt{2\pi} E[\theta^{n\beta}] \\
 &= \frac{C_1 k}{\sqrt{n^*+1}} \sqrt{2\pi} M_{\ln(\theta)}(n\beta) = 1,
 \end{aligned}$$

where $M_Y(t)$ denotes the moment generating function of a random variable Y . $\ln(\theta) \sim \text{Normal}(A_1, \xi^2)$ based on the probability density functions of a lognormal distribution. As a result of Equation (6) we get

$$1 = \frac{C_1 k}{\sqrt{n^*+1}} \sqrt{2\pi} e^{A_1(n\beta) + 0.5\xi^2(n\beta)^2}.$$

Therefore,

$$C_1 = \frac{\sqrt{n^*+1}}{\sqrt{2\pi}k} e^{-(A_1(n\beta) + 0.5\xi^2(n\beta)^2)}$$

and the conditional Bayes estimate of θ is

$$\hat{\theta}_{Bayes}|\beta = E[\theta|\beta, \underline{x}] = e^{(A_1 + 0.5\xi^2(2n\beta+1))} \quad (7)$$

The Bayes estimate of β can be derived based on the posterior probability distribution function of β in Equation (4). Let C_2 denote the normalizing constant for the probability distribution function of β . Let $B_1 = n + a_1$, $B_2 = \frac{b_1}{1+b_1P}$. We have

$$\begin{aligned}
 & \int_0^\infty C_2 e^{(-0.5(\frac{\beta}{k})^2 A_2)} \beta^{B_1-1} e^{-\frac{\beta}{B_2}} d\beta \\
 &= C_2 \Gamma(B_1) B_2^{B_1} \int_0^\infty C_2 e^{(-0.5(\frac{\beta}{k})^2 A_2)} \frac{\beta^{B_1-1} e^{-\frac{\beta}{B_2}}}{\Gamma(B_1) B_2^{B_1}} d\beta = 1
 \end{aligned}$$

Please note that $\frac{\beta^{B_1-1} e^{-\frac{\beta}{B_2}}}{\Gamma(B_1) B_2^{B_1}}$ is the probability density function of $\text{Gamma}(B_1, B_2)$. As a result

$$C_2 = (\Gamma(B_1) B_2^{B_1-1} E[e^{-0.5(\frac{\beta}{k})^2 A_2}])^{-1},$$

and the Bayes estimate of β is

$$\hat{\beta}_{Bayes} = E[\beta|\underline{x}] = \frac{B_1 B_2 E_1}{E_2}, \quad (8)$$

where E_1 and E_2 are the expected values $E[e^{-0.5(\frac{\beta}{k})^2 A_2}]$, when β follows $\text{Gamma}(B_1 + 1, B_2)$ and $\text{Gamma}(B_1, B_2)$ distribution respectively. Numerical integration in software Mathematica is used to compute both E_1 and E_2 . Similar to the ML estimation method, the following grid searching method is used in Bayesian estimation:

1. Sort the sample of size n in increasing order, i.e., $x_1 < x_2 < \dots < x_n$ and let n^* be the size of the partial sample of the first n^* losses x_1, x_2, \dots, x_{n^*} . Start from $n^* = 1$.
2. Compute the Bayes estimate β via (8) for the given n^* .
3. Compute the conditional Bayes estimate of θ via (7), given $\hat{\beta}_{Bayes}$ from Step 2. If $x_{n^*+1} \leq \hat{\theta}_{Bayes|\beta} \leq x_{n^*+1}$, then we found n^* . otherwise, increase n^* by 1.
4. Repeat Step 2 and 3 for $n^* = 2, 3, \dots$ till $x_{n^*+1} \leq \hat{\theta}_{Bayes|\beta} \leq x_{n^*+1}$, and we found the correct n^* .

The values of $\hat{\beta}_{Bayes}$ and $\hat{\theta}_{Bayes|\beta}$ found based on the correct n^* , represent the actual Bayes estimates of the parameters β and θ . We therefore obtain analytical Bayesian estimates of the LN-Pareto composite model without simulation.

4.2. Validation by Simulation

We employ simulation studies to validate the accuracy of the proposed Bayesian estimation method for the LN-Pareto distribution, and compare it with the ML estimation method. In each simulation, we generate a sample of n observations from the LN-Pareto composite density function given by Equation (1) for different selected values of θ and β . Then we obtain the ML estimates and Bayesian estimates of these two parameters, based on the simulated samples. We repeat the simulation for $N = 100$ times. The average value of the estimates (denoted by $\bar{\theta}$ and $\bar{\beta}$) and the mean squared errors (MSE, denoted by ϵ) are computed. MSEs show the differences between the true values and the estimated values of the parameters and intuitively indicate the performance of the estimation method.

Bayesian estimates of θ and β needs appropriate values of the hyper-parameters a_1, b_1 , and c_1 . Recall that the prior distribution for β is a Gamma distribution with hyper-parameter a_1 and b_1 . Therefore, when specifying the values of hyper-parameter a_1 and b_1 , we need to make sure the product of a_1 and b_1 , which is the expectation of the Gamma prior distribution, equal to the preset value of β . In addition, since the variance of the Gamma prior distribution is proportional to b_1^2 , we choose b_1 very small and then solve a_1 from the equation $a_1 * b_1 = \beta$. For example, when the selected true value for β is 0.5, a_1 and b_1 are chosen to be $a_1 = 100$ and $b_1 = 0.005$ such that b_1 is small and the product of a_1 and b_1 is 0.5.

Similarly, the conditional prior of $\theta|\beta$ are assumed to follow a log-normal distribution $LN(c_1, (\frac{k}{\beta})^2)$, therefore we choose c_1 to be the solution of the equation $\exp(c_1 + 0.5(\frac{k}{\beta})^2) = \theta$, based on the expectation of the log-normal distribution. Table 5 lists Bayesian estimates and the ML estimates using simulated data from the LN-Pareto composite distribution with different values of the parameters.

From Table 5, we can see that the informative Bayesian estimates outperform the ML estimates in both cases $\beta < 1$ and $\beta > 1$, for Bayesian estimates have smaller MSEs than the ML estimates in all the simulation scenarios. Please note that as sample size n increases, both ML estimation and Bayes method provide more accurate estimates for θ , in terms of smaller MSEs when n increases. However, in Equation (8) the Bayesian estimate of β is proportional to n since $B_1 = n + a_1$. As n increases, the MSE of Bayesian estimate of β increases slightly but is still much smaller than the MSE of the ML estimates. These simulation results indicate that Bayesian estimates are consistently better than the ML estimates if choosing reasonable hyper-parameters.

Table 5. Comparison of the ML and Bayesian estimates.

$\theta = 5 \quad \beta = 0.5 \quad (a_1 = 100, b_1 = 0.005, c_1 = 1.33231)$								
n	$\hat{\theta}_{ML}$	$\epsilon_{\theta_{ML}}$	$\hat{\beta}_{ML}$	$\epsilon_{\beta_{ML}}$	$\hat{\theta}_{Bayes}$	$\epsilon_{\theta_{Bayes}}$	$\hat{\beta}_{Bayes}$	$\epsilon_{\beta_{Bayes}}$
20	7.8604	4.0640	0.4747	0.1611	5.6933	2.0959	0.4641	0.0402
50	7.8942	3.8006	0.4435	0.1121	5.7941	1.2265	0.4338	0.0688
100	7.6421	3.6415	0.4333	0.1056	5.8745	1.1874	0.4058	0.0960
$\theta = 20 \quad \beta = 0.5 \quad (a_1 = 560, b_1 = 0.001, c_1 = 2.71861)$								
n	$\hat{\theta}_{ML}$	$\epsilon_{\theta_{ML}}$	$\hat{\beta}_{ML}$	$\epsilon_{\beta_{ML}}$	$\hat{\theta}_{Bayes}$	$\epsilon_{\theta_{Bayes}}$	$\hat{\beta}_{Bayes}$	$\epsilon_{\beta_{Bayes}}$
20	21.8966	12.2540	0.5878	0.2626	20.5062	5.5390	0.5341	0.0345
50	19.9977	4.9775	0.5738	0.2431	20.5262	3.3892	0.5025	0.0065
100	19.5730	3.6691	0.5467	0.1830	21.2182	3.0585	0.4629	0.0379
$\theta = 5 \quad \beta = 1.5 \quad (a_1 = 3000, b_1 = 0.0005, c_1 = 1.57865)$								
n	$\hat{\theta}_{ML}$	$\epsilon_{\theta_{ML}}$	$\hat{\beta}_{ML}$	$\epsilon_{\beta_{ML}}$	$\hat{\theta}_{Bayes}$	$\epsilon_{\theta_{Bayes}}$	$\hat{\beta}_{Bayes}$	$\epsilon_{\beta_{Bayes}}$
20	5.342	1.2443	1.510	0.3811	5.076	0.3877	1.480	0.02047
50	5.041	0.5606	1.528	0.2181	5.038	0.2526	1.451	0.0492
100	5.035	0.2344	1.516	0.1389	5.111	0.2196	1.406	0.0941
$\theta = 20 \quad \beta = 1.5 \quad (a_1 = 3000, b_1 = 0.0005, c_1 = 2.96494)$								
n	$\hat{\theta}_{ML}$	$\epsilon_{\theta_{ML}}$	$\hat{\beta}_{ML}$	$\epsilon_{\beta_{ML}}$	$\hat{\theta}_{Bayes}$	$\epsilon_{\theta_{Bayes}}$	$\hat{\beta}_{Bayes}$	$\epsilon_{\beta_{Bayes}}$
20	20.698	3.3162	1.560	0.4157	20.556	1.8798	1.460	0.04030
50	20.763	3.0164	1.468	0.2479	20.460	1.0797	1.404	0.09601
100	20.218	2.6237	1.466	0.2839	20.637	0.9563	1.324	0.1762

4.3. Bayesian Estimates of Three Composite Models

Aminzadeh and Deng (2018, 2019) have derived closed-form Bayesian estimators of the unknown parameter of Exp-Pareto and IG-Pareto distributions. The Bayesian estimate of θ in Exp-Pareto distribution is

$$\hat{\theta}_{Bayes} = \frac{b_2 + 1.35 \sum_{i=1}^{n^*} x_i}{a_2 - 0.35n + 1.35n^* - 1},$$

where a_2 and b_2 are hyper-parameters, n is the sample size and equal to 462 for our data, and n^* is the size of partial sample, as aforementioned.

The Bayesian estimate of θ in IG-Pareto distribution is

$$\hat{\theta}_{Bayes} = \frac{a_3(na + n^*k + b_3)}{(a_3k \sum_{i=1}^{n^*} \frac{1}{x_i} + 1)},$$

where a_3 and b_3 are hyper-parameters, and $k = 0.144351, a = 0.163847$, as specified in Table 2, for the IG-Pareto distribution.

Based on the Bayesian estimators given by Aminzadeh and Deng (2018, 2019) and the Bayesian estimators for the LN-Pareto distribution derived in this research, we have closed form Bayesian estimators for all three composite models. Based on the CPI adjusted natural disaster losses from 1980 to 2016 in the US, we obtain analytical Bayesian estimates of the three composite models for natural loss severity in Table 6.

Table 6. Bayesian estimates of three composite models for the severity of natural events in the US.

Model	Prior Distributions	Bayesian Estimates	NLL	AIC	BIC
Exp-Pareto	$\theta \sim \text{Inverse-Gamma}(10, 5)$	$\hat{\theta} = 23.1451$ $n^* = 183$	2697.57	5397.13	5401.27
IG-Pareto	$\theta \sim \text{Gamma}(50, 1)$	$\hat{\theta} = 4.3818$ $n^* = 77$	2699.17	5400.34	5404.48
LN-Pareto	$\theta \sim \text{LN}(1.61352, 2.857)$ $\beta \sim \text{Gamma}(20, 500, 1.1 \times 10^{-6})$	$\hat{\theta} = 19.2316$ $\hat{\beta} = 0.220173$ $n^* = 168$	2327.63	4659.27	4667.54

Bayesian estimation also shows that the LN-Pareto is the best model among all three composite distribution models for the smallest NLL, AIC, and BIC values. Comparing the NLL, AIC, and BIC values in Tables 4 and 6, we can see that Bayesian estimation method has smaller values of these three criteria when fitting three composite models to the natural losses data. Kass and Raftery (1995) claimed that difference of 10 or more is strong evidence to favor the model with a smaller BIC value. Although the advantage of Bayesian estimate over the ML estimate is marginal especially in fitting Exp-Pareto and LN-Pareto to the data, Bayesian estimation overall performs better than the ML estimation method. The advantage of Bayesian method will be significant and favorable when the sample size is small.

5. Risk Measures

In this section, we are going to investigate risk measures of the loss severity of natural events, based on the LN-Pareto composite model. Two important risk measures, Value at Risk (VaR) and Tailed Value at Risk (TVaR), are used in our research. As comparison, we also display VaR and TVaR of loss severity based on the other two composite models.

Value at Risk and Tailed Value at Risk

VaR is a point risk measurement and describes the minimum loss with the desired level of confidence. Given a level of confidence p and a cumulative distribution function $F(x)$ of a loss random variable X , VaR is defined as

$$\Pr(X \leq \text{VaR}_p(X)) = p, \quad \text{i.e.,} \quad \text{VaR}_p(X) = F^{-1}(p).$$

For example, if the VaR of the natural disaster loss severity is \$100 million at a 95% confidence level, there is a only a 5% chance that the damage from a natural event will be more than \$100 million in any natural disaster. This risk measure can be used by an insurance company to assess reinsurance need and risk management, so that the losses can be covered without putting the company at risk.

TVaR was developed as an alternative to VaR. It describes the average loss over VaR for a given confidence level p . Mathematically put,

$$\text{TVaR}_p(X) = E[X|X > \text{VaR}_p(X)] = \frac{\int_{\text{VaR}_p(X)}^{\infty} xf(x)dx}{1-p}.$$

In three composite distributions, the Pareto distribution is used to model large losses with small frequencies. However, the expectation of the Pareto distribution does not exist if the shape parameter is smaller than 1. This is true for all fitted composite distributions in our research. Therefore, we define Limited Tailed Value at Risk (LTVaR) as

$$\text{LTVaR} = E[(X \wedge b)|X > \text{VaR}_p(X)] = \frac{\int_{\text{VaR}_p(X)}^b xf(x)dy + \int_b^{\infty} bf(x)dy}{1-p},$$

where b is the maximum liability of a loss. From its definition, LTVaR is the average of the losses that are great than VaR but capped by the loss limit b , and therefore greater than the limited expectation. LTVaR is a very useful measure and can be easily implemented in insurance because this concept matches the maximum insurance benefit of an insurance policy.

Table 7 displays the derived VaR_p and $LTVaR_p$ at 85% confidence level from the three Bayesian estimated and ML estimated composite distributions. b is chosen to be 10^5 ('0000000 US\$) for $LTVaR_{0.85}$. Table 7 shows that the theoretical $VaR_{0.85}(X)$ based on three composite models are different when a different estimation method is used.

Table 7. Risk measures (in '0000000 US\$) of natural events' severity from three composite models.

	Models	$VaR_{0.85}(x)$	$LTVaR_{0.85}(x)$
Bayesian Estimation	Exp-Pareto	1073	30,723
	IG-Pareto	58,212	98,041
	LN-Pareto	11,070	75,860
ML Estimation	Exp-Pareto	1185	31,770
	IG-Pareto	38,029	94,619
	LN-Pareto	11,963	76,946

For the IG-Pareto model, the Bayesian estimation method is significantly better than the ML estimation and these two estimation methods result in dramatic difference in the theoretical $VaR_{0.85}(X)$ and $LTVaR_{0.85}(X)$. For the other two composite models, although the Bayesian estimation method has marginally lower BIC values than the ML estimation method, the theoretical $VaR_{0.85}(X)$ and $LTVaR_{0.85}(X)$ are significantly different. Therefore, we need to be cautious in the choice of model estimation methods when calibrating a composite model.

Secondly, large values of theoretical $VaR_{0.85}(X)$ and $LTVaR_{0.85}(X)$ indicates that composite distributions do take care of the fat tail problem in our real world situation and that the average loss in the worst cases will not be underestimated by a composite model. Moreover, both the ML estimation and Bayesian estimation methods confirm that the LN-Pareto fits the US natural disaster data better than the other two composite models, and our simulation validation has verified that the Bayesian estimation method for the LN-Pareto distribution is more accurate than the ML estimation; therefore, we would like to put more weight on the result from the Bayesian estimated LN-Pareto model, as highlighted by bold font in in Table 7.

6. Conclusions

In this paper, we propose using composite distributions to model natural disaster losses. A composite model piece-wisely models the typical feature of insurance losses, that is, high frequency of small amount of losses and low frequency of large amount of losses. We use the US natural disaster data from 1980 to 2016 in our research, considering the change in natural disasters' occurrence after 1980 due to climate change.

There are a total of 462 natural disasters during 1980 to 2016. After converting the amounts of losses into the year 2016 dollars, we test the assumption that natural disaster severity random variables in different years are independent and identically distributed. Our tests support the i.i.d assumption and we are able to use all the natural losses as realizations of one natural disaster severity random variable.

Based on the sample of 462 natural losses, we compare the performance of three composite distributions in modeling the natural losses in the US, namely Exp-Pareto, IG-Pareto, and LN-Pareto distributions. Based on the ML estimation method, we find that composite distributions fit the natural disaster losses better than the corresponding non-composite distributions according to the NLL, AIC, and BIC measures. In addition, we also find that the LN-Pareto model is the best one among these three composite distributions,

We make the first attempt to derive analytical Bayesian estimates for the LN-Pareto model. Simulation studies are conducted to assess the performance of the derived Bayesian estimates. The values of the MSEs from simulation show that the analytical Bayesian estimation performs better than the ML estimation method for the LN-Pareto distribution with various values for its parameters. The simulation study also reveals that Bayesian method is superior to ML in particular if the sample size is not very large.

In this research, the MCMC method is not used since we derived closed form Bayesian estimates for the LN-Pareto model. Based on the analytical Bayesian estimates of the three composite models, it is confirmed that LN-Pareto composite model is the best fit to the natural disaster losses. Bayesian estimation is proven to perform better than the ML estimation, according to the NLL, AIC, and BIC values.

Several risk measures for natural losses based on these three composite models are thereafter presented and compared. The differences in the derived risk measures from different composite distributions and different estimation methods reveal the importance of choosing an appropriate composite model for modeling natural losses and the difficulty in estimating the model.

Our research provides alternative information for insurance and risk management of natural disasters. We acknowledge the sparseness of natural disaster data. There are only 462 individual natural losses in the past 37 years and we rely on the CPI data to convert the losses to be consistent and free of the effect of price inflations over years. In the future research, we will continue to investigate the features of composite models and explore Bayesian model selection in predictive analysis of natural losses.

Author Contributions: Conceptualization, M.D., M.A. and M.J.; Methodology, M.D. and M.A.; Software, M.D. and M.A.; Validation, M.A.; Formal analysis, M.D. and M.A.; Investigations, M.D. and M.A.; Resources, M.D.; Data curation, M.D.; writing—original draft preparation, M.J.; writing—review and editing, M.D. and M.A.; visualization, M.D. and M.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Number and Loss Amounts of Natural Events from 1900 to 2016

Table A1. The number of occurrence of the natural events breakdown by time, type of natural events from 1900 to 2016.

	1900 to 1980	1980 to 2016	1900 to 2016	% of All Events
Drought	1 (9.09%)	10 (90.91%)	11	4.26%
Earthquake	14 (43.75%)	18 (56.25%)	32	12.40%
Epidemic	0 (0.00%)	4 (100.00%)	4	1.55%
Extreme temperature	9 (29.03%)	22 (70.97%)	31	12.02%
Flood	17 (32.69%)	35 (67.31%)	52	20.16%

Table A1. *Cont.*

	1900 to 1980	1980 to 2016	1900 to 2016	% of All Events
Landslide	2 (40.00%)	3 (60.00%)	5	1.94%
Storm	54 (58.70%)	38 (41.30%)	92	35.66%
Volcanic activity	0 (0.00%)	1 (100.00%)	1	0.39%
Wildfire	5 (16.67%)	25 (83.33%)	30	11.63%
Total	102 (39.53%)	156 (60.47%)	258	100.00%

Table A2. The adjusted total damage ('0000000 US\$) breakdown by time, type of natural events from 1900 to 2016.

	1900 to 1979	1980 to 2016	1900 to 2016	% of All Damages
Drought	0 (0.00%)	4371.623471 (100.00%)	4371.623471	3.62%
Earthquake	1521.333617 (19.08%)	6454.101764 (80.92%)	7975.435381	6.60%
Epidemic	0 (0.00%)	0 (0.00%)	0	0.00%
Extreme temperature	1633.100124 (43.24%)	2143.756251 (56.76%)	3776.856375	3.13%
Flood	4260.932883 (30.94%)	9511.711943 (69.06%)	13,772.64483	11.40%
Landslide	0 (0.00%)	2.027634158 (100.00%)	2.027634158	0.00%
Storm	10,143.84049 (11.57%)	77,499.01392 (88.43%)	87,642.85441	72.54%
Volcanic activity	0 (0.00%)	250.4927427 (100.00%)	250.4927427	0.21%
Wildfire	253.0904446 (8.34%)	2782.634167 (91.66%)	3035.724612	2.51%
Total	17,812.29756 (14.74%)	103,015.3619 (85.26%)	120,827.6595	100.00%

Appendix B. Damage Losses from Natural Events from 1980 to 2016

Table A3. Individual damage losses from natural event ('000000 US\$) from 1980 to 2016.

Year	# of Losses	1st Loss	2nd Loss	3th Loss	4th Loss	5th Loss	...				
1980	6	1019.45	87.38	58.25	2504.93	5825.41	2504.93				
1981	2	1056.14	1217.20								
1982	8	572.04	2487.12	84.31	99.48	248.71	104.46	497.42	1243.56		
1983	9	7229.13	2409.71	74.70	15.06	1265.10	722.91	240.97	313.26	36.15	
1984	10	1683.05	80.85	2309.98	69.30	46.20	39.27	80.85	265.65	69.30	1385.99
1985	10	2453.60	2007.49	3345.82	26.99	758.39	1784.44	446.11	22.31	516.59	0.89
1986	5	1.58	3832.23	87.59	65.70	54.75					
1987	7	450.01	8.45	242.96	33.80	122.54	21.13	10.56			
1988	0										
1989	4	13,548.78	10,839.03	735.51	967.77						
1990	6	23.32	73.45	183.63	918.16	64.27	82.63				
1991	9	59.03	2643.25	52.86	4405.41	52.86	1762.17	1497.84	1762.17	590.33	
1992	8	128.30	171.07	145.41	8553.35	5132.01	153.96	171.07	45,332.75		
1993	7	315.58	207.62	166.09	8304.74	19,931.38	1660.95	12.46			
1994	7	161.95	3.24	404.87	48,584.41	1133.64	809.74	3.40			
1995	12	196.86 3149.70	3307.18 4724.55	4724.55	1330.75	1102.39	3149.70	4724.55	157.48	15.75	15.75
1996	6	1070.78	5200.92	13.00	2294.52	764.84	30.59				
1997	17	269.17 224.31	3.74 299.07	2.99 89.72	366.37 747.69	74.77 747.69	373.84 2243.06	224.31 7476.85	299.07	747.69	149.54
1998	25	2061.41 73.62 2208.65	2945.61 0.88 736.22	1472.44 92.03 397.56	406.39 1472.44 92.03	690.57 2.94 295.22	6294.66 1774.21	220.87 544.80	147.24 663.33	88.35 92.03	6.63 295.22
1999	18	648.28 144.06	144.06 1440.62	3976.11 10.08	288.84 90.04	1440.62 288.12	216.09 0.43	132.54 1584.68	100.84 432.19	10,084.33	288.84

Table A3. Cont.

Year	# of Losses	1st Loss	2nd Loss	3th Loss	4th Loss	5th Loss	...				
2000	16	627.20 305.24	292.69 13.94	2090.65 27.88	139.38 487.82	39.72 1533.15	11.29 696.88	1393.77	231.37	69.69	125.44
2001	12	338.80 4.07	31.17 2710.41	2.44	8131.24	17.62	27.10	13.55	40.66	5.42	9.49
2002	16	533.65 400.23	267.49 933.88	6.67 2668.23	17.34 601.02	5.34 8.81	2935.05 4402.57	26.68	267.49	1334.11	26.68
2003	13	6521.93 4.43	32.61 2739.21	5217.54 260.88	138.26	4395.78	260.88	521.75	22.17	65.22	4565.35
2004	14	381.17 10,164.40	5.72 2.67	1397.61 1.27	889.39 635.28	76.23	0.22	20,328.81	79.41	13,976.06	22,869.91
2005	11	307.23 122.89	245.78	36.87	430.12	6.55	430.12	2740.48	19,662.63	17,573.48	301.08
2006	20	1428.61 119.05	714.31 18.45	1904.82 39.12	308.34 29.76	14.29 113.10	535.73 357.15	101.19 29.76	1190.51 178.58	8.33 107.15	19.05 428.58
2007	15	32.41 2315.08	578.77 347.26	810.28 347.26	150.48 694.52	2893.85 347.26	364.63	1157.54	578.77	162.06	405.14
2008	16	501.63 668.84	2.23 1114.74	780.32 1226.21	1783.58 11,147.41	113.70 122.62	1337.69 401.31	200.65	33,442.22	1449.16	2229.48
2009	12	185.71 950.91	1901.83 1678.08	111.87	268.49	2796.80	559.36	1230.59	671.23	2237.44	1118.72
2010	7	2586.57	2971.80	13.76	2201.33	110.07	1651.00	550.33			
2011	14	2133.97 4908.14	195.26 213.40	11,736.86 2133.97	14,937.82 8535.90	2027.28	7789.01	1066.99	3200.96	800.24	3734.45
2012	22	182.94 4181.42 219.52	1620.30 209.07 20,907.08	1881.64 522.68	219.52 5226.77	181.89 4704.09	627.21 104.54	2090.71 3554.20	52,267.70 1463.50	2.09 1986.17	52.27 731.75
2013	26	1648.42 25.76 206.05	1133.29 25.76 103.03	309.08 2.06 2266.58	3193.82 334.84 103.03	309.08 180.30 103.03	2163.55 309.08 1133.29	22.05 1957.50	515.13 10.30	927.24 1339.34	2.06 2.06

Table A3. Cont.

Year	# of Losses	1st Loss	2nd Loss	3th Loss	4th Loss	5th Loss	...				
2014	19	2.03	2027.63	101.38	3953.89	273.73	66.91	1622.11	709.67	172.35	101.38
		253.45	91.24	212.90	253.45	1622.11	760.36	2534.54	2230.40	20.28	
2015	28	172.14	506.31	1417.66	961.98	1012.62	162.02	1417.66	2734.06	658.20	101.26
		81.01	2.03	708.83	101.26	961.98	151.89	1417.66	2.03	1721.45	101.26
		273.41	141.77	911.35	607.57	405.05	151.89	3037.85	1822.71		
2016	25	550.00	125.00	3900.00	2000.00	2400.00	1000.00	1100.00	300.00	1000.00	150.00
		50.00	10,000.00	100.00	600.00	550.00	10,000.00	1200.00	275.00	20.00	1200.00
		100.00	2300.00	1600.00	1200.00	2300.00					

References

- Akaike, Hirotogu. 1973. Information theory and an extension of the maximum likelihood principle. In *Proceedings of the 2nd International Symposium on Information Theory, Tsahkadsor, USSR, Armenia, 2–8 September 1971*. Edited by B. N. Petrov and F. Csaki. Budapest: Akademia Kiado, pp. 267–81.
- Aminzadeh, Mostafa S., and Min Deng. 2018. Bayesian Predictive Modeling for Exponential-Pareto Composite Distribution. *Variance* 12: 59–68.
- Aminzadeh, Mostafa S., and Min Deng. 2019. Bayesian Predictive Modeling for Inverse Gamma-Pareto Composite Distribution. *Communications in Statistics-Theory and Methods* 48: 1938–54. [\[CrossRef\]](#)
- Bakar, S. A. Abu, Nor A. Hamzah, Mastoureh Maghsoudi, and Saralees Nadarajah. 2015. Modeling loss data using composite models. *Insurance: Mathematics and Economics* 61: 146–54.
- Burnham, Kenneth P., and David R. Anderson. 2002. *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach*, 2nd ed. New York: Springer.
- Cooray, Kahadawala, and Chin-I Cheng. 2015. Bayesian estimators of the lognormal-Pareto composite distribution. *Scandinavian Actuarial Journal* 6: 500–15. [\[CrossRef\]](#)
- Cooray, Kahadawala, and Malwane M. A. Ananda. 2005. Modeling actuarial data with a composite lognormal-Pareto model. *Scandinavian Actuarial Journal* 5: 321–34. [\[CrossRef\]](#)
- Gibbons, Jean Dickinson, and Subhabrata Chakraborti. 2003. *Nonparametric Statistical Inference, 4th ed., Revised and Expanded*. Statistics Textbooks and Monographs. Boca Raton: CRC Press, vol. 168.
- Kass, Robert E., and Adrian E. Raftery. 1995. Bayes Factors. *Journal of the American Statistical Association* 90: 773–95. [\[CrossRef\]](#)
- Levi, Charles, and Christian Partrat. 1991. Statistical Analysis of Natural Events in the United States. *ASTIN Bulletin: The Journal of the IAA* 21: 253–76. [\[CrossRef\]](#)
- Schwarz, Gideon. 1978. Estimating the dimension of a model. *The Annals of Statistics* 6: 461–64.
- Teodorescu, Sandra, and Raluca Vernic. 2006. A composite Exponential-Pareto distribution. *The Annals of the Ovidius, University of Constanta, Mathematics Series* 1: 99–108.