Statistics and Its Interface Volume 15 (2022) 29–38

Spatial Weibull regression with multivariate log gamma process and its applications to China earthquake economic loss

Hou-Cheng Yang, Yishu Xue*,
Lijiang Geng, and Guanyu Hu

Bayesian spatial modeling of heavy-tailed distributions has become increasingly popular in various areas of science in recent decades. We propose a Weibull regression model with spatial random effects for analyzing extreme economic loss. Model estimation is facilitated by a computationally efficient Bayesian sampling algorithm utilizing the multivariate Log-Gamma distribution. Simulation studies are carried out to demonstrate better empirical performances of the proposed model than the generalized linear mixed effects model. An earthquake data obtained from Yunnan Seismological Bureau, China is analyzed. Logarithm of the Pseudo-marginal likelihood values are obtained to select the optimal model, and Value-at-risk, expected shortfall, and tail-value-at-risk based on posterior predictive distribution of the optimal model are calculated under different confidence levels.

Keywords and phrases: Catastrophic risk, MCMC methods, Non-Gaussian data, Risk measure.

1. INTRODUCTION

Extreme geological disasters often cause catastrophic impact to both environmental systems and human society. For example, an earthquake can cause ground shaking, ground rupture, landslides, tsunami, etc. All such effects pose serious threats to the environment as well as humans, and cause tremendous economic losses and casualties. Study of the influential factors for, and consequences of such extreme environmental events, is of great value to both the environment and the human society.

Both the locations and outcomes of extreme events are of research interest. In reality, for example, the locations of earthquakes are not fixed but rather random. However, they tend to happen near the boundaries of tectonic plates (see Figure 2 in Xue and Hu, 2019, for an example), which indicates that there is an underlying intensity surface that govern the overall frequency of occurence. Spatial point process models assume the existence of such intensity surface and treat events as realizations from point processes, and have proven to be capable of capturing patterns in the locations of events. (Vere-Jones, 1970; Ogata, 1988; Schoenberg, 2003; Hu, Huffer and Chen, 2019). For analysis of the outcomes, regression-based models have been widely used to identify influential factors, such as the impact of earthquake magnitude on the final loss (Charpentier and Durand, 2015; Hu and Bradley, 2018; Yang, Hu and Chen, 2019). The economic losses incurred by such disastrous events, often observed to be long-tailed, have been studied using extreme value theory (EVT; Coles and Powell, 1996; Bali, 2003) in different fields such as economics and finance, insurance, environmetrics, and geology.

The bridge between spatial factors and the economic losses due to extreme events, however, have not been fully established. Li, Tang and Jiang (2016) proposed a Bayesian approach for a total of four mixture models to depict catastrophic economic losses caused by earthquakes. Covariates and spatial-dependent structures are, nevertheless, missing from the model, which can be a major disadvantage as the effect of an earthquake on the economic loss can be spatially varying. Intuitively, an earthquake that occurred in regions with high population densities is expected to cause more severe economic loss than one that struck a vast region where no human reside in. It is therefore rather crucial that spatial variation in covariate effects is allowed in the model. Although random effects have been used in previous works in conjunction with long-tailed distributions, such as in Sohn, Chang and Moon (2007) and Hernández and Giampaoli (2018), these random effects are often i.i.d., or have a certain clustered structure. The gap between, extreme distribution and spatial factors, however, remains unbridged.

In this work, we propose a hierarchical Bayesian approach for analyzing economic losses caused by extreme events. A spatial generalized linear mixed effects model (GLMM) is proposed, where spatial random effects (Banerjee, Carlin and Gelfand, 2014) are incorporated into the model to allow for information leveraging from neighbors. We choose to use the traditionally-popular Weibull distribution to model economic loss because of its heavy tail. The multivariate log-gamma distribution (MLG; Bradley et al., 2018) is used as the prior for the Weibull

*Corresponding author.

distrbution (Xu, Bradley and Sinha, 2019). As MLG enjoys conjugacy, closed forms for posterior distributions can be obtained, which facilitates efficient computation. Three risk measures based on the posterior predictive distribution are introduced. Our simulation studies show promising empirical performance of the proposed Bayesian methods as the parameter estimation is fairly accurate. The model is further illustrated with an earthquake dataset from Yunnan, China, and it identifies impact factors that influence the final incurred loss.

The rest of article are organized as follows. Section 2 gives a brief introduction and description of the motivating data. In Section 3, we develop the spatial Weibull regression model and risk measures based on the posterior predictive distribution, and examine theoretical properties of the proposed model. Furthermore, Bayesian model selection criteria Logarithm of the Pseudo-marginal likelihood (LPML) is used for model comparison in Section 4. In addition, extensive simulation studies are conducted in Section 5 to investigate empirical performance of the proposed model. In Section 6, we implement our model using an earthquake data about Yunnan Province, China from 1950 to 2014. Section 7 concludes with a brief discussion. For ease of exposition, all proofs are given in the appendices.

2. MOTIVATING DATA

Similar to Li, Tang and Jiang (2016), we analyze the direct economic losses caused by earthquakes which occurred in and close to mainland China between 1950 and 2014, collected by Yunnan Province Seismological Bureau. In this data set, earthquake magnitude is a number that characterizes the relative size of the earthquake, which is based on a measurement of the maximum motion recorded by a seismograph. The location (latitude, longitude) is recorded for each occurrence. An indicator variable denoting whether an earthquake occurred in urban or rural areas is also present. A visualization of the earthquake locations and magnitudes is given in Figure 1.

In this dataset, information of 124 earthquakes are collected, among which 77 occurred in cities and 47 occurred in rural areas. A description of the dataset is shown in Table 1. The magnitudes of the earthquakes range from 4 to 8.1, and economic loss they incurred range from 5 CNY to 10^6 CNY, making an extremely wide interval. A histogram and an exponential quantile-quantile plot of the economic losses are shown in Figure 2, from which we observe that the economic loss is heavily tailed. A scatterplot of economic losses versus magnitudes is presented in Figure 3. It is rather clear that simple linear regression cannot capture the relation between the earthquake magnitudes and the economic losses.

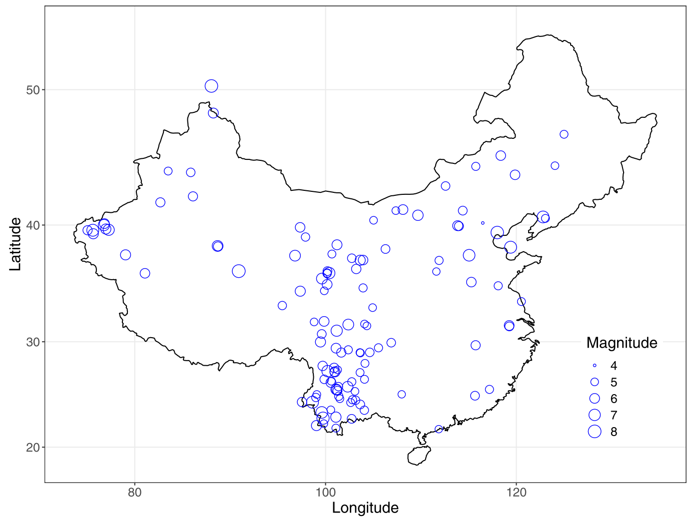

Figure 1. Visualization of earthquake locations in and close to mainland China between 1950 and 2014. Larger circle indicate more severe economic loss.

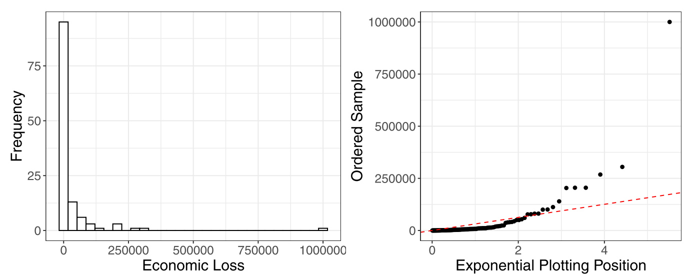

Figure 2. Histogram and exponential quantile-quantile plot of economic losses.

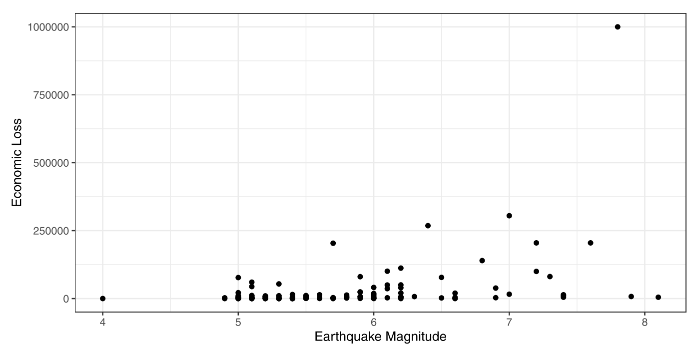

Figure 3. Scatterplot of earthquake magnitudes and economic losses.

3. METHODOLOGY

The Bayesian hierarchical model

Heavy-tailed distributions are probability distributions whose tails are not exponentially bounded, and there are even super-exponential distributions. They have been used in many areas, such as earth science, survival analysis, economics, and finance. Specifically, in economics or finance study (e.g., Rachev, 2003), the underlying risk factors are always assumed to follow heavy-tailed distributions. The

Table 1. Summary of response and covariates in the earthquake dataset

Notation	Variable Name	Type	Range/Categories	Median/Count
Z	Economic Loss	Numerical	[5, 10^6]	5000
X_1	Earthquake Magnitude	Numerical	[4, 8.1]	5.5
X_2	Urban Indicator	Binary	$\{0, 1\}$	$\{47(X_2 = 0), 77(X_2 = 1)\}$

Weibull distribution is one of the most important heavy-tailed distributions that is used as a probabilistic model of the amount of loss associated with actuarial and financial risk management (Gebizlioglu, Şenoğlu and Kantar, 2011). With this feature, the Weibull distribution is a reasonable choice for us to model the economic losses caused by earthquakes.

The Weibull probability density function is

$$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} \exp\left(-\left(\frac{x}{\lambda}\right)^{k}\right), \quad x > 0,\ k > 0,\ \lambda > 0, \tag{1}$$

where k and λ are the shape and scale parameters, respectively. In order to obtain an efficient conjugate form under Bayesian setting, we rewrite the probability density function alternatively as

$$f(x) = kb \cdot x^{k-1} \exp\left(-x^k b\right), \quad x > 0,\ b = \lambda^{-k} > 0,\ k > 0. \tag{2}$$

For the rest of this paper, we denote the distribution expressed by (2) as Weibull(k, b).

To bring in spatial information, latent Gaussian models (Gelfand and Schliep, 2016) have been conventionally used. Introducing spatial random effect into the Weibull model, for locations $\boldsymbol{s} = (s_1, \ldots, s_n)$, denote the losses as $Z(s_1), \ldots, Z(s_n)$, we have

$$Z_{s_i} \sim \text{Weibull}(k, \mu_i),$$

where the vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ satisfy:

$$\log(\boldsymbol{\mu}) = \boldsymbol{X}(s)\boldsymbol{\beta} + \boldsymbol{W}, \tag{3}$$

where $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)^\top$, $\boldsymbol{W}$ is an n-dimensional vector of spatial random effects, $\boldsymbol{X}(\boldsymbol{s})$ is an $n \times p$ covariate matrix, and $s_i \in \mathcal{R}^2$. Furthermore, it is often assumed that

$$\boldsymbol{W} \mid \phi, \sigma_w \sim \text{MVN}(\boldsymbol{0}_n, \sigma_w^2 \boldsymbol{H}(\phi)), \tag{4}$$

where $\boldsymbol{W} = (w(s_1), \ldots, w(s_n))^\top$, $\boldsymbol{H}(\phi)$ is the $n \times n$ spatial correlation matrix with ϕ being the range parameter, and MVN denotes the multivariate normal distribution. In our paper, we use an exponential covariogram to define $\boldsymbol{H}(\phi)$, i.e.,

$$\boldsymbol{H}(\phi) = \exp(-\text{dist}/\phi), \tag{5}$$

where "dist" denotes an $n \times n$ matrix whose (i, j)th entry is the distance between locations s_i and s_j.

A Bayesian approach involves specifying prior distributions for unknown parameters. Under the setting described above, the joint distribution of the data, processes, and parameters is written as the product of the following distributions:

$$\begin{aligned}
&\textbf{Data Model}: Z(s_i) \mid \boldsymbol{W}, \boldsymbol{\beta}, \sigma^2, \phi \overset{\text{ind}}{\sim} \text{Weibull}(k, \mu(s_i)) \\
&\text{for } i = 1, \ldots, n; \forall k \in (0, +\infty); \\
&\textbf{Process Model}: \boldsymbol{W} \mid \phi, \sigma_w \sim \text{MVN}(\boldsymbol{0}_n, \sigma_w^2 \boldsymbol{H}(\phi)) \\
&\textbf{Parameter Model 1}: \boldsymbol{\beta} \sim \text{MVN}(\boldsymbol{0}_p, \sigma^2 \boldsymbol{I}_p) \\
&\textbf{Parameter Model 2}: \log(\sigma) \sim \text{N}(0, 1) \\
&\textbf{Parameter Model 3}: \log(\sigma_w) \sim \text{N}(0, 1) \\
&\textbf{Parameter Model 4}: \phi \sim \text{DU}(a_1, b_1)
\end{aligned}, \tag{6}$$

where DU is shorthand for discrete uniform distribution. Lognormal priors are given to σ and σ_w for better convergence results, which has been later verified in numerical experiments. For ϕ, DU is assumed following Banerjee, Carlin and Gelfand (2014), which argued that compared to the variance components, σ and σ_w, we are less interested in the spatial range, and an informative prior such as discretized uniform over a finite set of points is recommended. The formulation in (6) requires tuning of Metropolis–Hasting steps (Chib and Greenberg, 1995) within the Gibbs sampler, as Gaussian process does not maintain conjugacy for non-Gaussian data. To improve the computational efficiency for Weibull regression, a conjugate prior is desired. Bradley et al. (2018) proposed a multivariate log-gamma (MLG) distribution as the conjugate prior for Poisson spatial regression model, and established connection between multivariate log-gamma distribution and multivariate normal distribution. The following construction demonstrates that MLG is also an ideal prior choice for Weibull regression model because of their conjugacy. Similar to Bradley et al. (2018), we define the n-dimensional random vector $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_n)^\top$, which consists of n mutually independent log-gamma random variables with shape and scale parameters organized into n-dimensional vectors $\boldsymbol{\alpha} \equiv (\alpha_1, \ldots, \alpha_n)^\top$ and $\boldsymbol{\kappa} \equiv (\kappa_1, \ldots, \kappa_n)^\top$, respectively. Then the n-dimensional random vector $\boldsymbol{q}$ is defined as

$$\boldsymbol{q} = \boldsymbol{\mu} + \boldsymbol{V}\boldsymbol{\gamma}, \tag{7}$$

where $\boldsymbol{V} \in \mathcal{R}^n \times \mathcal{R}^n$ and $\boldsymbol{\mu} \in \mathcal{R}^n$. Bradley et al. (2018) called $\boldsymbol{q}$ the multivariate log-gamma random vector. The probability density function of the random vector $\boldsymbol{q}$ can be defined as:

$$f(\boldsymbol{q} \mid \boldsymbol{\mu}, \boldsymbol{V}, \boldsymbol{\alpha}, \boldsymbol{\kappa}) = \frac{1}{\det(\boldsymbol{V})} \left(\prod_{i=1}^{m} \frac{1}{\Gamma(\alpha_i)\kappa_i^{\alpha_i}} \right) \cdot \exp\left[\boldsymbol{\alpha}^\top \boldsymbol{V}^{-1}(\boldsymbol{q}-\boldsymbol{\mu}) - \boldsymbol{\kappa}^{(-1)^\top} \exp\{\boldsymbol{V}^{-1}(\boldsymbol{q}-\boldsymbol{\mu})\} \right], \tag{8}$$

where "det" represents the determinant function. We use $\mathrm{MLG}(\boldsymbol{\mu}, \boldsymbol{V}, \boldsymbol{\alpha}, \boldsymbol{\kappa})$ as a shorthand for the probability density function in (8). From Bradley et al. (2018), we know that the latent multivariate log-gamma process is a saturated process of the latent Gaussian process. If $\boldsymbol{q}$ follows a multivariate log-gamma distribution $\mathrm{MLG}(\mathbf{0}, \alpha^{1/2}\boldsymbol{V}, \alpha\mathbf{1}, 1/\alpha\mathbf{1})$, as $\alpha \to \infty$, $\boldsymbol{\beta}$ will converge in distribution to a multivariate normal distribution vector with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{V}\boldsymbol{V}^\top$. In practice, choosing $\alpha = 10\,000$ is sufficient for this normal approximation.

This property of the MLG distribution makes it a favorable choice in our scenario. Substituting the Gaussian distribution in (6) with MLG, we have the following hierarchical model:

(9)

Data Model: $Z(s_i) \mid \boldsymbol{W}, \boldsymbol{\beta}, \sigma^2, \phi, \overset{\text{ind}}{\sim} \mathrm{Weibull}(k, \mu(s_i))$ for $i = 1, \dots, n,\ \forall k \in (0, +\infty)$

Process Model: $\boldsymbol{W} \mid \phi, \sigma_w \sim \mathrm{MLG}(\mathbf{0}_n, \boldsymbol{\Sigma}_W^{1/2}, \alpha_W \mathbf{1}_n, \kappa_W \mathbf{1}_n)$

Parameter Model 1: $\boldsymbol{\beta} \sim \mathrm{MLG}(\mathbf{0}_p, \boldsymbol{\Sigma}_\beta^{1/2}, \alpha_\beta \mathbf{1}_p, \kappa_\beta \mathbf{1}_p)$

Parameter Model 2: $\log(\sigma) \sim \mathrm{N}(0, 1)$

Parameter Model 3: $\log(\sigma_w) \sim \mathrm{N}(0, 1)$

Parameter Model 4: $\phi \sim \mathrm{DU}(a_1, b_1)$

where k denotes the shape parameter of the Weibull distribution, $\mu(\boldsymbol{s}_i) = \boldsymbol{X}(s_i)\boldsymbol{\beta} + \boldsymbol{W}$, $\boldsymbol{\Sigma}_W = \sigma_w^2 \boldsymbol{H}(\phi)$, $\boldsymbol{\Sigma}_\beta = \sigma^2 \boldsymbol{I}_p$, $\alpha_W > 0, \alpha_\beta > 0, \kappa_W > 0$, and $\kappa_\beta > 0$. Based on the results of Bradley et al. (2018), the full conditionals of $\boldsymbol{\beta}$ and $\boldsymbol{W}$ will be the conditional MLG distribution (cMLG). As there is no analytic forms for the posterior distributions of σ_w and ϕ, in this work we use Metropolis–Hasting algorithm (Chib and Greenberg, 1995) to obtain posterior samples. Slice sampling (Neal et al., 2003) might be an alternative approach, but we do not discuss it here, and refer interested readers to the original text. Full conditional distributions are presented in Appendix A.

Long-tailed property justification

We present theoretical justification for usage of MLG as a conjugate prior for Weibull distribution. The Weibull distribution is "long-tailed" with shape parameter greater than 0 but less than 1, which is an important subclass of heavy-tailed distributions (Asmussen, 2003). A random variable X is said to have a long right tail if for all $k > 0$,

$$\lim_{x\to\infty} P[X > x + k \mid X > x] = 1. \tag{10}$$

For the Weibull distribution, the long tail probability can be written as

$$LP = P[Z > z + \delta \mid Z > z] = \frac{\exp\big(-b(z+\delta)^k\big)}{\exp(-bz^k)}, \tag{11}$$

where $0 < k < 1$ is the shape parameter of the Weibull distribution. Under the spatial setting, $b(s) = \exp\{\boldsymbol{X}^\top(s)\boldsymbol{\beta} + w(s)\}$ for any $s \in D$. thus the log of the long-tail probability can be expressed as

$$\log(LP) = \exp\{\boldsymbol{X}(s)^\top \boldsymbol{\beta} + w(s)\}(z^k - (z+\delta)^k). \tag{12}$$

For the latent Gaussian model of $w(s)$, it is rather straightforward to find the expected value of $\log(LP)$ using the moment generating function for the normal distribution. The expected log long-tail probability under Gaussian model is given as

$$\mathrm{E_G}\{\log(LP) \mid \sigma_\beta^2, \sigma_w^2\} = \exp\left\{ \frac{1}{2} \left(\sum_{i=1}^{p} X_i^2(s,t)\sigma_\beta^2 + \sigma_w^2 \right) \right\} \cdot (z^k - (z+\delta)^k). \tag{13}$$

For the multivariate log-gamma model, the expected log long-tail probability is

$$\mathrm{E_{MLG}}\left\{\log(LP) \mid \sigma_\beta^2, \sigma_w^2, \alpha_w, \alpha_\beta, \kappa_\beta, \kappa_w\right\} = \frac{\left(\frac{\kappa_\beta^{p\alpha_\beta}}{\Gamma(\alpha_\beta)^p}\right)\left(\frac{\kappa_w^{\alpha_w}}{\Gamma(\alpha_w)}\right)}{\left(\frac{\kappa_\beta^{p\alpha_\beta + \sum_{i=1}^p X_i(s,t)\sigma_\beta}}{\prod_{i=1}^p \Gamma(\alpha_\beta + X_i(s,t)\sigma_\beta)}\right)\left(\frac{\kappa_w^{\alpha_w+1}}{\Gamma(\alpha_w+1)}\right)} (z^k - (z+\delta)^k). \tag{14}$$

Proofs for Equations (13) and (14) are provided in Appendix B of Hu and Bradley (2018). A relationship between Equations (13) and (14) is provided in Proposition 1 below.

Proposition 1. *Assume that $\boldsymbol{\beta}$ and $\boldsymbol{W}$ follow the MLG distribution as defined in Equation (8). Then, we have the following,*

$$E_{MLG}\{\lim_{\alpha\to\infty} \log(LP) \mid \sigma_\beta^2 = \alpha\sigma_1^2, \sigma_w^2 = \alpha\sigma_2^2, \alpha_w = \alpha, \alpha_\beta = \alpha, \kappa_\beta = \alpha, \kappa_w = \alpha\} = E_G\left\{\log(LP) \mid \sigma_\beta^2 = \sigma_1^2, \sigma_w^2 = \sigma_2^2\right\},$$

where E_{MLG} is the expected value with respect to the multivariate log-gamma distribution, and E_G is the expected value with respect to the Gaussian distribution, $\sigma_1^2 > 0$, and $\sigma_2^2 > 0$.

Proof. Pass the limit through the expectation, and apply Proposition 1 from Bradley et al. (2018). □

Based on results in Proposition 1, the long-tail probability of Gaussian process model implied by the Weibull distribution is a special case of the long-tail probabilities implied by (9). Choosing relatively smaller value of α will model the long-tail property better than Gaussian process, since the Gaussian process is known to have a very thin tail, but the Log Gamma distribution is a heavy tail distribution.

Risk measures based on the posterior predictive distribution

While in conventional model fitting, researchers are concerned with the final point prediction which often occurs at the mean, with catastrophes or extreme environmental events such as earthquakes or hurricanes, risk measures different from the mean are preferred, which are often of high importance to policy sellers in the insurance industry. The most popular among them include value-at-risk (VaR; Duffie and Pan, 1997), expected shortfall (ES; Acerbi and Tasche, 2002) and tail-value-at-risk (TVaR; Bargès, Cossette and Marceau, 2009). VaR is a popular risk measure because of its simplicity and easiness to be understood (Dowd and Blake, 2006). Given α between 0 and 1, VaR is defined as the $100(1-\alpha)$th percentile of the density function of loss, and denoted as q_α. Being an alternative measure, ES is defined to be the negative of the expectation of the tail beyond the VaR. Noticing that the VaR is only a numerical value and does not describe the loss pattern in the tail beyond itself, Bargès, Cossette and Marceau (2009) proposed the more general TVaR, which is defined as

$$\text{TVaR}_\alpha = \frac{1}{1-\alpha}\int_\alpha^1 \text{VaR}_u \mathrm{d}u, \tag{15}$$

and captures the entire tail beyond the specified percentile rather than one point, making it a measure for the average risk. For a new catastrophic event, we would like to make predictions on the risk. In Bayesian analysis, the posterior predictive distribution is the distribution of possible unobserved values conditional on the observed values (Gelman et al., 2013). Let $p(\boldsymbol{\theta}|\boldsymbol{Z})$ denote the posterior distribution of all parameters $\boldsymbol{\theta}$ given data $\boldsymbol{Z}$, then the posterior predictive distribution of new data $\boldsymbol{Z}^*$ is given as

$$p(\boldsymbol{Z}^* \mid \boldsymbol{Z}) = \int p(\boldsymbol{Z}^* \mid \boldsymbol{\theta}, \boldsymbol{Z}) p(\boldsymbol{\theta} \mid \boldsymbol{Z}) \mathrm{d}\boldsymbol{\theta}. \tag{16}$$

The VaR, ES, and TVaR based on the posterior predictive distribution of $\boldsymbol{Z}^*$ are defined as:

$$\text{VaR}_\alpha(\boldsymbol{Z}^*) = q_\alpha(\boldsymbol{Z}^*), \tag{17}$$

$$\text{ES}_\alpha(\boldsymbol{Z}^*) = E[\boldsymbol{Z}^* \mid \boldsymbol{Z}^* \geq \text{VaR}_\alpha(\boldsymbol{Z}^*)], \tag{18}$$

$$\text{TVaR}_\alpha(\boldsymbol{Z}^*) = \frac{1}{1-\alpha}\int_\alpha^1 \text{VaR}_u(\boldsymbol{Z}^*)\mathrm{d}u. \tag{19}$$

4. BAYESIAN MODEL SELECTION CRITERION

As seen in the model construction (9), tuning of parameter k for Weibull distribution is needed. To select the most suitable model parameter, we adapt the Bayesian model assessment criterion, logarithm of the Pseudo-marginal likelihood (LPML; Ibrahim, Chen and Sinha, 2013) to our Weibull regression with spatial random effects scenario. The LPML is defined as

$$\text{LPML} = \sum_{i=1}^n \log(\text{CPO}_i), \tag{20}$$

where CPO_i is the conditional predictive ordinate (CPO) for the i-th subject. CPO is calculated based on the leave-one-out-cross-validation, which estimates the probability of observing data y_i in the future after having already observed data $y_1, \cdots, y_{i-1}, y_{i+1}, \cdots, y_n$. The CPO for the i-th subject is defined as

$$\text{CPO}_i = f(y_i \mid \boldsymbol{y}_{-i}) \equiv \int f(y_i \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta} \mid \boldsymbol{y}_{-i})\mathrm{d}\boldsymbol{\theta}, \tag{21}$$

where $\boldsymbol{y}_{-i} = \{y_1, \cdots, y_{i-1}, y_{i+1} \cdots, y_n\}$, and

$$\pi(\boldsymbol{\theta} \mid \boldsymbol{y}_{-i}) = \frac{\prod_{j\neq i} f(y_j \mid \boldsymbol{\theta})\pi(\boldsymbol{\theta})}{c(\boldsymbol{y}_{-i})}, \tag{22}$$

where $c(\boldsymbol{y}_{-i})$ is the normalizing constant. The CPO_i in Equation (21) can be expressed as

$$\text{CPO}_i = \frac{1}{\int \frac{1}{f(y_i|\boldsymbol{\theta})}\pi(\boldsymbol{\theta} \mid \boldsymbol{y}_{-i})\mathrm{d}\boldsymbol{\theta}}. \tag{23}$$

Therefore, a Monte Carlo estimate of CPO_i in Equation (23) is given by

$$\widehat{\text{CPO}}_i^{-1} = \frac{1}{B}\sum_{b=1}^B \frac{1}{f(y_i \mid \boldsymbol{\theta}_b)}, \tag{24}$$

where $\boldsymbol{\theta}_b$ is the b-th MCMC sample of $\boldsymbol{\theta}$ from $\pi(\boldsymbol{\theta} \mid \boldsymbol{y})$. For model (9), we have following estimation for CPO:

$$\widehat{\text{CPO}}_i^{-1} = \frac{1}{B}\sum_{i=1}^B \frac{1}{f(Z(s_i) \mid \boldsymbol{\beta}_b, \boldsymbol{X}(s_i), \widehat{w}(s_i))}, \tag{25}$$

where $\{\boldsymbol{\beta}_b, b = 1, \cdots, B\}$ denotes a Gibbs sample of $\boldsymbol{\beta}$ from $\pi(\boldsymbol{\beta} \mid \text{Data})$, and $\widehat{w}(\boldsymbol{s_i})$ is posterior mean of spatial random effects on location $\boldsymbol{s_i}$. Finally, the logarithm of the Pseudo marginal likelihood (LPML) is defined as

$$\text{LPML} = \sum_{i=1}^n \log(\widehat{\text{CPO}_i}). \tag{26}$$

In the context of model selection, we select the best model which has the largest LPML value under CPO.

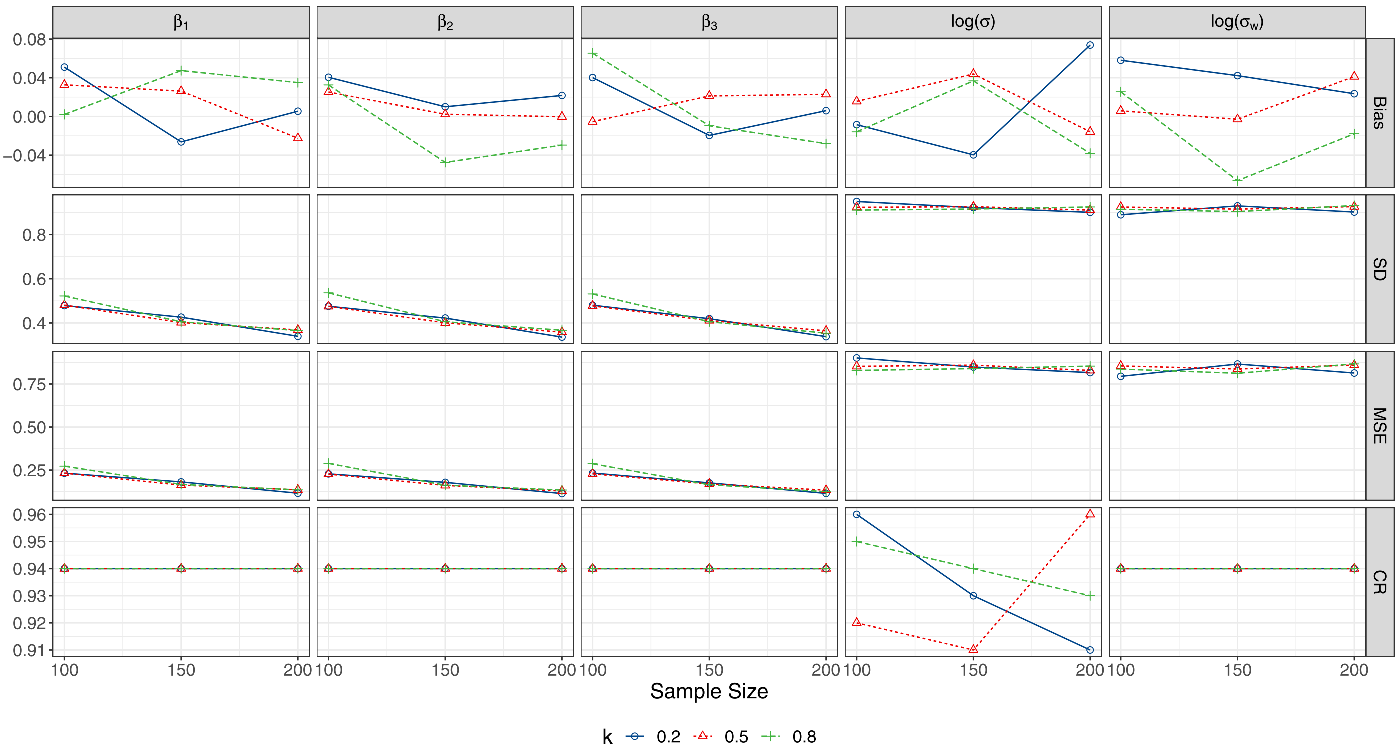

Figure 4. Visualization of simulation performances.

5. SIMULATION STUDY

For simplicity, we base our simulation study on the spatial domain $D = [0,3] \times [0,3]$, and the locations s_i for $i = 1, \cdots, n$ are generated in an uniform manner over D. The proposed model (9) is fitted. In the first three simulation designs, we set $\boldsymbol{\beta} = (-1,-1,-1)^\top$, and the $X(s_i)$'s are independently generated from the standard uniform distribution $\mathrm{U}(0,1)$. The shape parameter for Weibull distribution is set to be $k \in \{0.2, 0.5, 0.8\}$. The vector of spatial random effects $\boldsymbol{W}$ is generated from $\mathrm{MLG}(\mathbf{0}, \sigma_w \boldsymbol{H}^{1/2}(\phi), \alpha_w, \kappa_w)$ where $\phi = 5$, $\sigma_w = 1$, $\alpha_w = 1$, $\kappa_w = 1$, and $\boldsymbol{H}$ is the matrix of Euclidean distances between pairs in the n generated locations. A DU(1,10) prior is given to ϕ following Chapter 6 of Banerjee, Carlin and Gelfand (2014). Statistical inference is performed using MCMC with chain length of 5000, and we drop the first 2000 iterations as burn-in. We calculate the bias, standard deviation (SD), mean squared error (MSE), and coverage rate (CR) of the 95% highest posterior density (HPD; Chen and Shao, 1999) interval for each parameter based on posterior mean and posterior quantiles. These four evaluation metrics are visualized in Figure 4.

A few observations can be made from Figure 4. For each value of k, even with sample size of 100, the model parameters are estimated with very small average bias. The average bias, with only few exceptions, decrease in terms of absolute value with increase in the sample size. The SD and MSE for the three covariate effect estimators decrease monotonically with respect to increase in sample size for each value of k. Estimation performance for $\log(\sigma)$ and $\log(\sigma_w)$ also improve with sample size, but the influence of sample size is not as pronounced as for the three β's. The average CR remains around its 0.95 nominal value.

To demonstrate the performance of our proposed model under a different scenario where the underlying spatial random effects are generated from the multivariate normal distribution, in the last simulation scenario for each k, we generate $\boldsymbol{W}$ from multivariate normal distribution with mean zero and covariance matrix $\boldsymbol{\Sigma}_W = \sigma_w^2 \boldsymbol{H}(\phi)$. We assume true $\phi = 5$ and $\sigma_w = 1$. The covariate $X(s_i)$'s are independently generated from the standard uniform distribution $\mathrm{U}(0,1)$. We again fit the model in (9) over 100 replicates. In each replicate, we run 5000 iterations and drop the first 2000 iterations as burn-in. The estimation performance results are shown in Table 2. The biases of parameter estimates increased as a consequence of model misspecification. In all three scenarios, the parameter estimates are biased to the positive direction, which indicates the effect of covariates are underestimated as the MLG is long-tailed, and its long-tail captured more information than needed. The SD, however, is smaller when compared to the corresponding scenarios when $\boldsymbol{W}$ is generated from MLG. The CR still remain close to its nominal level. In addition, the estimates for $\log(\sigma)$ and $\log(\sigma_w)$ remain very precise even if the model is misspecified.

Table 2. Performance measures when spatial random effect generate from multivariate normal distribution for $n = 200$

Setting	parameter	Bias	SD	MSE	CR
$k = 0.2$	β_1	0.206	0.306	0.138	0.94
	β_2	0.091	0.304	0.100	0.94
	β_3	0.107	0.306	0.105	0.94
	$\log(\sigma)$	−0.011	0.939	0.881	0.97
	$\log(\sigma_w)$	−0.010	0.927	0.859	0.94
$k = 0.5$	β_1	0.172	0.305	0.123	0.94
	β_2	0.199	0.305	0.133	0.94
	β_3	0.170	0.304	0.121	0.94
	$\log(\sigma)$	0.026	0.911	0.830	0.94
	$\log(\sigma_w)$	−0.019	0.914	0.835	0.94
$k = 0.8$	β_1	0.389	0.303	0.238	0.94
	β_2	0.411	0.298	0.258	0.94
	β_3	0.389	0.300	0.241	0.94
	$\log(\sigma)$	−0.001	0.935	0.874	0.95
	$\log(\sigma_w)$	0.024	0.906	0.821	0.94

To verify the ability of LPML in choosing an appropriate shape parameter, we generated 100 additional datasets with true shape parameter $k = 0.8$. To demonstrate that when the shape parameter is far away from the truth, the LPML will assume small values, we set the candidates to be $k \in \{0.2, 0.4, 0.6, 0.8\}$. In 89 out of the 100 replicates, LPML correctly selected the true parameter. This indicates that LPML can select the appropriate parameter with high accuracy, and therefore we apply it to choice of candidate models in our real data application.

6. A REAL DATA EXAMPLE

We analyze the earthquake data, which includes 124 earthquakes that occurred at different locations and corresponding direct economic losses in Yunnan Province caused by these earthquakes. In this dataset we have three covariates: a continuous variable magnitude characterizing the relative size of an earthquake, a binary variable county indicating whether the earthquake occurs in city or rural area, and another binary variable indicating whether the earthquake location is in Yunnan or not. The model in (9) is considered, with $\alpha_\beta = 10\,000$ and $\kappa_\beta = 0.0001$, which are the values that lead to an MLG that approximates a multivariate normal distribution for $\boldsymbol{\beta}$. We fixed $\alpha_w = 1$ and $\kappa_w = 1$ which means that the spatial random effect are not approximate to multivariate normal distribution, so it can catch the heavy tail dependence structure. The full conditionals in Appendix A are used to run a Gibbs sampler. The number of iterations of the Gibbs sampler is 25 000, and we drop the first 20 000 iterations as burn-in.

We fit Weibull regression models with nine different shape parameter values varying from 0.025 to 0.0975 with 0.025 increments, and present the corresponding LPML values of these models in Figure 5. The model with $k = 0.7$ turned out to have the largest LPML, and is selected as the best model. Further comparison of the three models that have the largest LPML values is performed, and the results are presented in Appendix B. We also compare model (6) to model (9) and both have two scenario: with and without spatial random effect. The result is shown in Table 3. We find that with a spatial random effect and a MLG prior (same as model (9)) has a best result among others. The posterior mode for ϕ is 1, indicating moderate spatial dependency. Table 4 shows the posterior estimation result of the selected model. We see that the 95% HPD of β_1 and β_2 does not contain zero, which means that both covariates magnitude and county are significant. The HPD interval for β_3 contains 0, which indicates that Yunnan province may still suffer from huge economic loss even due to an earthquake that happened somewhere else. Since the posterior estimations of both β_1 and β_2 are negative, we conclude that (i) earthquakes with higher magnitudes cause larger economic losses; (ii) earthquakes occurring in city area cause larger economic losses.

The three risk measurements VaR, ES and TVaR are computed based on posterior predictive distribution of the selected model under three different confidence levels. The measures are shown in Table 5. They together can help in-

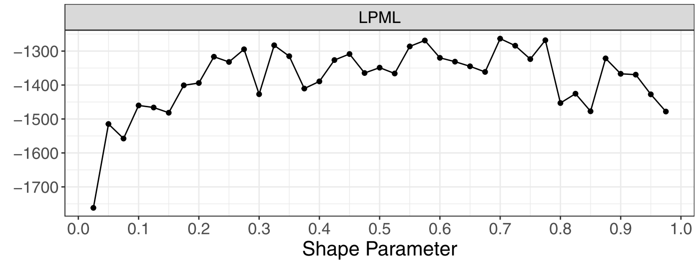

Figure 5. Plot of LPML values for candidate models with different Weibull shape parameters.

Table 3. LPML values corresponding to the best k for the spatial and non-spatial models

Prior	Model	LPML
MLG	Spatial Model	−1263
	Non-Spatial Model	−1275
MVN	Spatial Model	−7423
	Non-Spatial Model	−8542

Table 4. Posterior parameter estimates based on the best model ($k = 0.7$)

	Posterior Mean	Standard Error	HPD Interval
β_1	−0.889	0.047	(−0.977, −0.792)
β_2	−1.345	0.313	(−1.928, −0.710)
β_3	−0.508	0.353	(−1.224, 0.165)
$\log(\sigma_w)$	−0.582	1.026	(−2.539, 1.343)

Table 5. *Risk Measure calculated based on the final model selected by LPML (Unit: CNY)*

	VaR	ES	TVaR
90%	418,444	4,655,642	4,350,369
95%	1,833,555	8,016,257	7,192,600
99%	12,402,260	16,740,000	14,591,333

surance companies or the government make further decisions about catastrophic insurance coverages, reinsurance levels, or catastrophic reserves. $\mathrm{VaR}_{90\%}(\boldsymbol{Z}) = 418{,}444$ implies that the insurance coverage should be 418,444 CNY to cover 90% loss resulted from earthquakes. $\mathrm{ES}_{90\%}(\boldsymbol{Z}) = 4{,}655{,}642$ means that the expected excess of loss beyond 418,444 CNY is 4,655,642 CNY, which can be used to calculate the reinsurance premium when considering an excess of loss reinsurance. $\mathrm{TVaR}_{90\%}(\boldsymbol{Z}) = 4{,}350{,}369$ indicates that the mean loss above 418,444 is 4,350,369 CNY.

7. DISCUSSION

In this article, we propose an efficient Bayesian spatial model to analyze extreme losses caused by catastrophes. Our main methodological contribution is to use multivariate log-gamma process model for both regression coefficients and spatial random effects within a hierarchical spatial regression model. Multivariate log-gamma process models have the computational advantage of being conjugate with the Weibull likelihood, and therefore allow by-pass of tuning for Metropolis–Hastings algorithms. Additionally, our simulation results indicate that multivariate log-gamma process models have good estimation performance even when the data are generated from Gaussian process model. The results in this article can be applied to analyze the losses caused by many different perils (hurricane, tornado, earthquake, wildfire, etc.), and thus the methodology is of independent interest.

Three topics beyond the scope of this paper are worth further investigation. In this work we considered four covariates. When the number of covariates becomes large, introducing Bayesian variable selection to the MLG model becomes a necessary procedure. In this work, we considered using MLG as the prior for Weibull distribution, which belongs to the family of generalized extreme value (GEV) distributions. Using the MLG as prior for other members of GEV family, such as Gumbel and Fréchet distributions, are also of research interest. Finally, in our model formulation, the spatial random effect is dependent only on distances between pairs of locations. There could be other factors that control the spatial random effect, such as similarities in infrastructure, and incorporating random neighborhood structures similar to Gao and Bradley (2019) in spatial extreme value modeling is devoted for future research.

APPENDIX A. FULL CONDITIONALS DISTRIBUTIONS FOR WEIBULL DATA WITH LATENT MULTIVARIATE LOG-GAMMA PROCESS MODELS

From the hierarchical model in (9), the full conditional distribution for $\boldsymbol{\beta}$ satisfies

$$\begin{aligned} &f(\boldsymbol{\beta} \mid \cdot) \\ &\propto f(\boldsymbol{\beta}) \prod f(Z \mid \cdot) \\ &\propto \exp\Bigg[\sum_i (\boldsymbol{X}(s_i)^\top \boldsymbol{\beta} + w(s_i)) \\ &\qquad - \sum_i \big(Z_i^k \exp\big((\boldsymbol{X}(s_i)^\top \boldsymbol{\beta} + w(s_i))\big)\big)\Bigg] \\ &\quad \times \exp\left\{\alpha_\beta \mathbf{1}_p^\top \Sigma_{\boldsymbol{\beta}}^{-1/2} \boldsymbol{\beta} - \kappa_\beta \mathbf{1}_p^\top \exp\left(\Sigma_{\boldsymbol{\beta}}^{-1/2} \boldsymbol{\beta}\right)\right\}. \end{aligned} \tag{27}$$

Rearranging the terms we have

$$f(\boldsymbol{\beta} \mid \cdot) \propto \exp\left\{\boldsymbol{\alpha}_\beta^\top \boldsymbol{H}_\beta \boldsymbol{\beta} - \boldsymbol{\kappa}_\beta^\top \exp(\boldsymbol{H}_\beta \boldsymbol{\beta})\right\}, \tag{28}$$

which implies that $f(\boldsymbol{\beta} \mid \cdot)$ is equal to $\mathrm{cMLG}(\boldsymbol{H}_\beta, \boldsymbol{\alpha}_\beta, \boldsymbol{\kappa}_\beta)$, where "cMLG" is the conditional multivariate log gamma distribution from Bradley et al. (2018).

Similarly, the full conditional distribution for $\boldsymbol{W}$ satisfies

$$\begin{aligned} &f(\boldsymbol{W} \mid \cdot) \\ &\propto f(\boldsymbol{W}) \prod f(Z \mid \cdot) \\ &\propto \exp\Bigg[\sum_i (\boldsymbol{X}(s_i)^\top \boldsymbol{\beta} + w(s_i)) \\ &\qquad - \sum_i \big(Z_i^k \exp\big((\boldsymbol{X}(s_i)^\top \boldsymbol{\beta} + w(s_i))\big)\big)\Bigg] \\ &\quad \times \exp\left\{\alpha_W \mathbf{1}_n^\top \Sigma_W^{-1/2} \boldsymbol{W} - \kappa_W \mathbf{1}_n^\top \exp\left(\Sigma_W^{-1/2} \boldsymbol{W}\right)\right\}. \end{aligned} \tag{29}$$

Rearranging the terms we have

$$f(\boldsymbol{W} \mid \cdot) \propto \exp\left\{\boldsymbol{\alpha}_W^\top \boldsymbol{H}_W \boldsymbol{W} - \boldsymbol{\kappa}_W^\top \exp(\boldsymbol{H}_W \boldsymbol{W})\right\}, \tag{30}$$

which implies that $f(\boldsymbol{W} \mid \cdot)$ is equivalent to $\mathrm{cMLG}(\boldsymbol{H}_W, \boldsymbol{\alpha}_W, \boldsymbol{\kappa}_W)$. Thus we obtain the following full-conditional distributions to be used within a Gibbs sampler:

$$\begin{aligned} \boldsymbol{\beta} &\sim \mathrm{cMLG}(\boldsymbol{H}_\beta, \boldsymbol{\alpha}_\beta, \boldsymbol{\kappa}_\beta) \\ \boldsymbol{W} &\sim \mathrm{cMLG}(\boldsymbol{H}_W, \boldsymbol{\alpha}_W, \boldsymbol{\kappa}_W) \\ \log(\sigma) &\propto \mathrm{MLG}(\mathbf{0}, \boldsymbol{\Sigma}_\beta^{1/2}, \alpha_\beta \mathbf{1}_p, \kappa_\beta \mathbf{1}_p) \times \mathrm{N}(0,1) \\ \log(\sigma_w) &\propto \mathrm{MLG}(\mathbf{0}, \boldsymbol{\Sigma}_W^{1/2}, \alpha_w \mathbf{1}_n, \kappa_w \mathbf{1}_n) \times \mathrm{N}(0,1) \\ \phi &\propto \mathrm{MLG}(\mathbf{0}_n, \boldsymbol{\Sigma}_W^{1/2}, \alpha_w \mathbf{1}_n, \kappa_w \mathbf{1}_n) \times \mathrm{DU}(a_1, b_1) \end{aligned} \tag{31}$$

Table 6. Parameters of the full conditional distribution

Parameter	Form
$\boldsymbol{H}_\beta$	$\begin{bmatrix} \boldsymbol{X} \\ \Sigma_{\boldsymbol{\beta}}^{-1/2} \end{bmatrix}$
$\boldsymbol{\alpha}_\beta$	$\begin{bmatrix} \mathbf{1}_{n\times 1} \\ \alpha_\beta \mathbf{1}_{p\times 1} \end{bmatrix}$
$\boldsymbol{\kappa}_\beta$	$\begin{bmatrix} \mathbf{Z}^{\top k} \mathbf{1}_{n\times 1} \exp(\mathbf{W}^\top)) \\ \kappa_\beta \mathbf{1}_p^\top \end{bmatrix}$
$\boldsymbol{H}_W$	$\begin{bmatrix} I_n \\ \boldsymbol{\Sigma}_W^{-1/2} \end{bmatrix}$
$\boldsymbol{\alpha}_W$	$\begin{bmatrix} \mathbf{1}_{n\times 1} \\ \alpha_W \mathbf{1}_{n\times 1} \end{bmatrix}$
$\boldsymbol{\kappa}_W$	$\begin{bmatrix} \mathbf{Z}^{\top k} \mathbf{1}_{n\times 1} \exp\big((\boldsymbol{X}\boldsymbol{\beta})^\top\big) \\ \kappa_W \mathbf{1}_n^\top \end{bmatrix}$

A motivating feature of this conjugate structure is that it is relatively straightforward to simulate from a cMLG. For $\log(\sigma)$, $\log(\sigma_w)$ and ϕ, we consider using a Metropolis-Hasting algorithm or slice sampling procedure. The parameters of the conditional multivariate log-gamma distribution are summarized in Table 6.

APPENDIX B. COMPARISON OF TOP THREE CANDIDATE MODELS FOR REAL DATA

We compare the root mean squared percentage error regression loss (RMSPE) of the three models that correspond to the largest LPML values. Their respective shape parameters are 0.700, 0.775, and 0.575. Define the RMSPE ratio for a bandwidth k' as

$$\text{RATIO}_{k'} = \frac{\text{RMSPE}_{k=0.7}}{\text{RMSPE}_{k=k'}}.$$

The values of $\text{RATIO}_{0.775}$ and $\text{RATIO}_{0.575}$ are calculated, respectively, to be 0.838 and 0.451, from the RMSPE perspective verifying that $k = 0.7$ yields the best candidate model.

Table 7. Posterior parameter estimates based on the proposed model with $k \in \{0.775, 0.575\}$)

k		Posterior Mean	SE	HPD Interval
0.775	β_1	−0.958	0.054	(−1.065 −0.851)
	β_2	−1.516	0.327	(−2.156, −0.875)
	β_3	−0.524	0.385	(−1.278, 0.231)
	$\log(\sigma_w)$	−0.693	0.967	(−2.588, 1.202)
0.575	β_1	−0.699	0.066	(−0.828, −0.569)
	β_2	−1.109	0.429	(−1.951, −0.268)
	β_3	−0.445	0.456	(−1.341, 0.449)
	$\log(\sigma_w)$	−0.202	0.908	(−1.983, 1.579)

The parameter estimates based on models with $k \in \{0.775, 0.575\}$ are show in Table 7. Compared to results presented in Table 4, there is no change in significance of any parameter based on HPD intervals.

Received 6 August 2020

REFERENCES

Acerbi, C. and Tasche, D. (2002). Expected Shortfall: A Natural Coherent Alternative to Value at Risk. *Economic Notes* **31** 379–388.

Asmussen, S. (2003). Steady-State Properties of of GI/G/1. *Applied Probability and Queues* 266–301. MR1978607

Bali, T. G. (2003). An Extreme Value Approach to Estimating Volatility and Value at Risk. *The Journal of Business* **76** 83–108.

Banerjee, S., Carlin, B. P. and Gelfand, A. E. (2014). *Hierarchical Modeling and Analysis for Spatial Data*. Chapman and Hall/CRC. MR3362184

Bargès, M., Cossette, H. and Marceau, E. (2009). TVaR-based Capital Allocation with Copulas. *Insurance: Mathematics and Economics* **45** 348–361. MR2591311

Bradley, J. R., Holan, S. H., Wikle, C. K. et al. (2018). Computationally Efficient Multivariate Spatio-Temporal Models for High-Dimensional Count-Valued Data (with Discussion). *Bayesian Analysis* **13** 253–310. MR3773410

Charpentier, A. and Durand, M. (2015). Modeling Earthquake Dynamics. *Journal of Seismology* **19** 721–739.

Chen, M.-H. and Shao, Q.-M. (1999). Monte Carlo Estimation of Bayesian Credible and HPD Intervals. *Journal of Computational and Graphical Statistics* **8** 69–92. MR1705909

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis–Hastings Algorithm. *The American Statistician* **49** 327–335.

Coles, S. G. and Powell, E. A. (1996). Bayesian Methods in Extreme Value Modelling: A Review and New Developments. *International Statistical Review/Revue Internationale de Statistique* **64** 119–136. MR3341082

Dowd, K. and Blake, D. (2006). After VaR: the Theory, Estimation, and Insurance Applications of Quantile-based Risk Measures. *Journal of Risk and Insurance* **73** 193–229.

Duffie, D. and Pan, J. (1997). An Overview of Value at Risk. *Journal of Derivatives* **4** 7–49.

Gao, H. and Bradley, J. R. (2019). Bayesian Analysis of Areal Data with Unknown Adjacencies Using the Stochastic Edge Mixed Effects Model. *Spatial Statistics* **31** 100357. MR3955693

Gebizlioglu, O. L., Şenoğlu, B. and Kantar, Y. M. (2011). Comparison of Certain Value-at-risk Estimation Methods for the Two-parameter Weibull Loss Distribution. *Journal of Computational and Applied Mathematics* **235** 3304–3314. MR2784045

Gelfand, A. E. and Schliep, E. M. (2016). Spatial Statistics and Gaussian Processes: A beautiful Marriage. *Spatial Statistics* **18** 86–104. MR3573271

Gelman, A., Stern, H. S., Carlin, J. B., Dunson, D. B., Vehtari, A. and Rubin, D. B. (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC. MR3235677

Hernández, F. and Giampaoli, V. (2018). The Impact of Misspecified Random Effect Distribution in a Weibull Regression Mixed Model. *Stats* **1** 48–76.

Hu, G. and Bradley, J. (2018). A Bayesian Spatial–temporal Model with Latent Multivariate Log-gamma Random Effects with Application to Earthquake Magnitudes. *Stat* **7** e179. MR3796093

Hu, G., Huffer, F. and Chen, M.-H. (2019). New Development of Bayesian Variable Selection Criteria for Spatial Point Process with Applications Technical Report No. 18-05, University of Connecticut, Department of Statistics.

Ibrahim, J. G., Chen, M.-H. and Sinha, D. (2013). *Bayesian Survival Analysis*. Springer Science & Business Media. MR1876598

Li, Y., Tang, N. and Jiang, X. (2016). Bayesian Approaches for Analyzing Earthquake Catastrophic Risk. *Insurance: Mathematics and Economics* **68** 110–119. MR3492696

Neal, R. M. et al. (2003). Slice Sampling. *The Annals of Statistics* **31** 705–767. MR1994729

Ogata, Y. (1988). Statistical Models for Earthquake Occurrences and Residual Analysis for Point Processes. *Journal of the American Statistical Association* **83** 9–27.

Rachev, S. T. (2003). *Handbook of Heavy Tailed Distributions in Finance: Handbooks in Finance* **1**. Elsevier.

Schoenberg, F. P. (2003). Multidimensional Residual Analysis of Point Process Models for Earthquake Occurrences. *Journal of the American Statistical Association* **98** 789–795. MR2055487

Sohn, S. Y., Chang, I. S. and Moon, T. H. (2007). Random effects Weibull regression model for occupational lifetime. *European Journal of Operational Research* **179** 124–131. MR2281805

Vere-Jones, D. (1970). Stochastic Models for Earthquake Occurrence. *Journal of the Royal Statistical Society. Series B (Methodological)* **32** 1–62. MR0272087

Xu, Z., Bradley, J. R. and Sinha, D. (2019). Latent Multivariate Log-Gamma Models for High-Dimensional Multi-Type Responses with Application to Daily Fine Particulate Matter and Mortality Counts. *arXiv preprint arXiv:1909.02528*.

Xue, Y. and Hu, G. (2019). Online Updating of Information Based Model Selection in the Big Data Setting. *Communications in Statistics – Simulation and Computation*. Forthcoming.

Yang, H.-C., Hu, G. and Chen, M.-H. (2019). Bayesian Variable Selection for Pareto Regression Models with Latent Multivariate Log Gamma Process with Applications to Earthquake Magnitudes. *Geosciences* **9** 169.

Hou-Cheng Yang
Department of Statistics
Florida State University
Tallahassee, FL 32306
USA
E-mail address: hy15e@my.fsu.edu

Yishu Xue
Department of Statistics
University of Connecticut
Storrs, CT 06269
USA
E-mail address: yishu.xue@uconn.edu

Lijiang Geng
Department of Statistics
University of Connecticut
Storrs, CT 06269
USA
E-mail address: lijiang.geng@uconn.edu

Guanyu Hu
Department of Statistics
University of Missouri Columbia
Columbia, MO 65201
USA
E-mail address: guanyu.hu@missouri.edu