The 6th International Conference on
Structural Health Monitoring of Intelligent Infrastructure
Hong Kong | 9-11 December 2013

EXTREME VALUE DISTRIBUTION MODEL OF VEHICLE LOADS BASED ON BAYESIAN PARAMETER ESTIMATION

Shunlong Li [1], Yang Xu [2], Jia Lu [3] and Hui Li [2]
[1]School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin, China.
Email: lishunlong@hit.edu.cn
[2] School of Civil Engineering, Harbin Institute of Technology, Harbin, China.
[3] School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin, China.

ABSTRACT

This paper presents the extreme value distribution model base on Bayesian parameter estimation using monitored vehicle loads data. The normal status and the dense status are considered to investigate their extreme value distribution characteristics of the vehicle loads. Analysis on the monitored vehicle loads datasets indicated that the inter-arrival times of vehicle at normal status and dense status could be described by a Poisson and Weibull process respectively. And the generalized Pareto distribution (GPD) was employed for the tail part modeling of vehicle weights. Moreover, the likelihood function of the extreme value distribution model could be constructed, and the prior distributions of the model parameters were estimated by the long-term monitored vehicle data. Then the posterior distributions over the model parameter were updated combining the prior and new sample information, with respect to prior distribution and likelihood function respectively. The application of the proposed method to the Nanjing 3rd Yangtze River Bridge demonstrates that the Bayesian extreme value estimation model predicts the extreme load more accurately than traditional methods.

KEYWORDS

Vehicle loads, extreme value distribution, Bayesian parameter estimation, Structural Health Monitoring (SHM)

INTRODUCTION

Monitored traffic action, especially the extreme values of vehicle loads, is one of the most important measures for determining fatigue deterioration (Okada et al. 1978; Moses et al. 1987) and predicting the remaining service time (Yang et al. 2004) of long-span bridges. Recently the demand for the transportation associated with prosperous economy has led to increasingly serious bridge overload in the mainland of China. The extreme value distribution models of vehicle loads in the current bridge design codes consider neither the phenomenon of increasingly serious overload nor the actual effect of load-controlling methods, and could not meet the requirement of long-term limit state analysis of bridges. Over the past decades, a considerable amount of research has been conducted related to vehicle loads modeling on bridges (Ghosn and Moses 1985; Nowak 1993; Ditlevsen and Madsen 1994; Croce and Salvatore 2001).

The extreme value distribution model can be estimated based on the monitored traffic load data, which provide the potential to insight the inherent characteristics of vehicle loads. The traffic flow was modeled as a Poisson process by (Lan et al. 2011), and the extreme value distribution function of vehicle load during the service period was extrapolated. (Tong et al. 2008) developed a multi-peak probability model to estimate the extreme value distribution of vehicle loads. (OBrien et al. 2010) proposed a semi-parametric distribution model based on in-field data from two sites. However, in comparison with the monitored weekly or monthly extreme values, the predictions generated by these methods are far from satisfactory because the estimated parameters of the extreme value distribution model could not: (1) describe the stochastic process accurately; and (2) fit the tail distribution of vehicle loads precisely.

This study estimates the extreme value distribution model of vehicle loads based on Bayesian parameter estimation, solving the above parameter identification problems. The paper is organized as follows: the extreme value distribution model incorporating the tail distribution, stochastic process and the Bayesian parameter estimation was given in Section 2; Section 3 illustrates a practical application of the proposed methods for the Nanjing 3rd Yangtze River Bridge; and finally draws some conclusions.

BAYESIAN EXTREME VALUE DISTRIBUTION ESTIMATION

Extreme Value Distribution Model

As recommended in the Unified Standard for Reliability Design of Highway Engineering Structures (GB/T 50283-1999), the vehicle weights with inter-arrival times $T_i\{T_i > 2\text{s}\}$ were considered as the normal status and the stochastic process of vehicle weights conform to a Filtered Poisson Process. The cumulative density function (CDF) of the extreme value distribution for vehicle weights at normal status could be expressed as:

$$F_{\text{MN}}(x;T) = \exp\{-\lambda_n T[1-F_N(x)]\} \tag{1}$$

where $F_D(x)$ indicates the truncated distribution of the vehicle weights; λ_n is the intensity parameter of the Poisson process; T denotes investigated reference period for extreme value distribution estimation.

Vehicle weights with time intervals $T_i\{T_i \le 2\text{s}\}$ were selected and built up a new datasets. The inter-arrival times of selected vehicle loads do not reject Weibull process hypothesis (Ministry of Transport of the PRC 1999). Then the stochastic process of vehicle loads at dense status could be modeled as Weibull process, which is considered as the promotion of Poisson process. The intensity function of Weibull process is expressed as follows

$$\lambda_d(t) = \lambda_d \beta_D t^{\beta_D - 1} (t \ge 0, \lambda_D, \beta_D > 0) \tag{2}$$

Then the extreme value distribution function of vehicle weight $F_{\text{MD}}(x;T)$ over investigated reference period $[0,T]$ can be illustrated as follows (Ministry of Transport of the PRC 1999)

$$F_{\text{MD}}(x;T) = \exp\left\{-\lambda_d T^{\beta_D}\left[1-F_{\text{D}}(x)\right]\right\} \tag{3}$$

where $F_{\text{D}}(x)$ denotes the truncated distribution of vehicle weights at the dense status.

For the truncated distribution $F_{\text{N}}(x)$ at normal status and $F_{\text{D}}(x)$ at dense statues, generalized Pareto distribution (GPD) was employed for the tail part modeling of vehicle weights. GPD has been widely used for modeling high-threshold exceedances and extreme value prediction (OBrien et al. 2010). The cumulative distribution function of vehicle weights could be could be determined by the following formulation

$$\begin{aligned} &F_{N \text{ or } D}(x) = P(X \le x) = P(X \le u) + [1 - P(X \le u)]F_u(x-u \mid \xi, \sigma) \\ &= (1-\alpha) + \alpha F_u(x-u \mid \xi, \sigma), x > u \\ &F_u(x-u \mid \xi, \sigma) = 1 - \left[1 + \frac{\xi(x-u)}{\sigma}\right]^{-\frac{1}{\xi}}, \xi \ne 0 \end{aligned} \tag{4}$$

where $F_u(x-u \mid \xi, \sigma)$ denotes the cumulative distribution function of generalized Pareto distribution function with shape parameter ξ and scale parameter σ; parameter $\alpha = 1 - P(X \le u)$ could be determined from the monitored vehicle datasets.

Substitute Eq. (4) into Eqs. (1) and (3), the extreme value distribution for normal and dense status could be expressed as

$$F_{\text{MN}}(x;T) = \exp\left\{-\alpha_n \lambda_n T\left[1 + \frac{\xi_N (x-u_N)}{\sigma_N}\right]^{-1/\xi_N}\right\} \tag{5a}$$

$$F_{\text{MD}}(x;T) = \exp\left\{-\alpha_d \lambda_d T^{\beta_D}\left[1 + \frac{\xi_D (x-u_D)}{\sigma_D}\right]^{-1/\xi_D}\right\} \tag{5b}$$

where $[\xi_N, \sigma_N, u_N]$ and $[\xi_D, \sigma_D, u_D]$ denotes the shape parameter, scale parameter and threshold at normal and dense status respectively; $\lambda_N = \alpha_n \lambda_n$ indicates the intensity parameter for vehicle weights at the tail part of normal status; the $\lambda_D = \alpha_d \lambda_d$ and β_D represents the parameter for vehicle weights at the tail part of dense status.

Bayesian Parameter Estimation

The Bayesian analysis is performed by combining the prior information $\boldsymbol{\pi}(\boldsymbol{\theta})$ and the sample information x into the posterior distribution $\boldsymbol{\pi}(\boldsymbol{\theta} \mid \mathbf{x})$ of parameter $\boldsymbol{\theta}$ given the sample observation x and expressed as

$$\text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{marginal likelihood}}$$
$$\boldsymbol{\pi}(\boldsymbol{\theta} \mid \mathbf{x}) = \frac{\mathbf{L}(\mathbf{x} \mid \boldsymbol{\theta})\boldsymbol{\pi}(\boldsymbol{\theta})}{\int_{\Theta} \mathbf{L}(\mathbf{x} \mid \boldsymbol{\theta})\boldsymbol{\pi}(\boldsymbol{\theta}) d\boldsymbol{\theta}} \tag{6}$$

$\mathbf{L}(\mathbf{x} \mid \boldsymbol{\theta})$ denotes the conditional likelihood of sample x given the parameter $\boldsymbol{\theta}$; Θ represents the parameter space of $\boldsymbol{\theta}$.

For extreme value distribution estimation at normal status, shown in Eq. (5a), the posterior distribution of model parameters $\boldsymbol{\theta} = [\xi_N, \sigma_N, \lambda_N]$ would be expressed as

$$\begin{aligned} \boldsymbol{\pi}(\xi_N, \sigma_N, \lambda_N \mid \mathbf{x}) &= \frac{\mathbf{L}(\mathbf{x} \mid \xi_N, \sigma_N, \lambda_N)\boldsymbol{\pi}(\xi_N, \sigma_N, \lambda_N)}{\iiint \mathbf{L}(\mathbf{x} \mid \xi_N, \sigma_N, \lambda_N)\boldsymbol{\pi}(\xi_N, \sigma_N, \lambda_N) d\xi_N d\sigma_N d\lambda_N} \\ &\propto \mathbf{L}(\mathbf{x} \mid \xi_N, \sigma_N, \lambda_N)\boldsymbol{\pi}(\xi_N, \sigma_N, \lambda_N) \end{aligned} \tag{7}$$

where $\mathbf{x}$ represents the observed vehicle weights; likelihood function $\mathbf{L}(\mathbf{x} \mid \xi_N, \sigma_N, \lambda_N)$ was expressed as

$$\mathbf{L}(\mathbf{x} \mid \xi_N, \sigma_N, \lambda_N) = \left(\frac{\lambda_N T}{\sigma_N}\right)^n \cdot \exp\left\{\sum_{i=1}^{n}\left[-n\left(1 + \frac{\xi_N x_i}{\sigma_N}\right)^{-\frac{1}{\xi_N}} + \log_e\left(1 + \frac{\xi_N x_i}{\sigma_N}\right)^{-\frac{1}{\xi_N}-1}\right]\right\} \tag{8}$$

In this study, the model parameters are assumed to be independently distributed, thus the prior $\boldsymbol{\pi}(\xi_N, \sigma_N, \lambda_N)$ could be determined by

$$\boldsymbol{\pi}(\xi_N, \sigma_N, \lambda_N) = \boldsymbol{\pi}(\xi_N)\boldsymbol{\pi}(\sigma_N)\boldsymbol{\pi}(\lambda_N) \tag{9}$$

A theoretical Bayesian model is readily built from the proposed GPD extreme value distribution estimation model, as shown in Eq. (7)-(9). The critical issue for practical applications is to determine the priors of the model parameters. Although the set of priors may be different for different models, there are still certain basic rules for determining an appropriate prior: (1) including as much previous experience as possible and (2) possessing certain mathematical traits for calculation efficiency and accuracy. Based on the monitored vehicle datasets, the priors for model parameters $\boldsymbol{\pi}(\xi_N)$, $\boldsymbol{\pi}(\sigma_N)$ and $\boldsymbol{\pi}(\lambda_N)$ follow normal, log-normal and log-normal distributions, respectively, the probability distribution function could be demonstrated by

$$\begin{aligned} \boldsymbol{\pi}(\xi_N) &= \frac{1}{\sqrt{2\pi\tau_{\xi_N}}} \exp\left\{-\frac{\left(\xi_N - \mu_{\xi_N}\right)^2}{2\tau_{\xi_N}^2}\right\} \\ \boldsymbol{\pi}(\sigma_N) &= \frac{1}{\sqrt{2\pi\tau_{\sigma_N}}} \exp\left\{-\frac{\left(\log\sigma_N - \mu_{\sigma_N}\right)^2}{2\tau_{\sigma_N}^2}\right\}, \\ \boldsymbol{\pi}(\lambda_N) &= \frac{1}{\sqrt{2\pi\tau_{\lambda_N}}} \exp\left\{-\frac{\left(\log\lambda_N - \mu_{\lambda_N}\right)^2}{2\tau_{\lambda_N}^2}\right\} \end{aligned} \tag{10}$$

where $[\mu_{\xi_N}, \sigma_{\xi_N}, \mu_{\sigma_N}, \sigma_{\sigma_N}, \mu_{\lambda_N}, \sigma_{\lambda_N}]$ are super-parameters that are generated from analysis of the monitored vehicle loads.

For extreme value distribution estimation at dense status, the posterior distribution is

$$\boldsymbol{\pi}(\xi_D,\sigma_D,\lambda_D,\beta_D \mid \mathbf{x}) \propto \mathbf{L}(\mathbf{x} \mid \xi_D,\sigma_D,\lambda_D,\beta_D)\boldsymbol{\pi}(\xi_D)\boldsymbol{\pi}(\sigma_D)\boldsymbol{\pi}(\lambda_D)\boldsymbol{\pi}(\beta_D) \tag{11}$$

where the likelihood function $\mathbf{L}(\mathbf{x} \mid \xi_D,\sigma_D,\lambda_D,\beta_D)$ and the priors $\boldsymbol{\pi}(\xi_D)$, $\boldsymbol{\pi}(\sigma_D)$, $\boldsymbol{\pi}(\lambda_D)$, $\boldsymbol{\pi}(\beta_D)$ could be determined the same way as at the normal status.

The posterior distributions of the model parameters were obtained by Markov Chain Monte Carlo (MCMC) method with observed vehicle weight data. It should be noted that MCMC cannot be easily performed with a large number of unknown parameters. In this study, the threshold u_N and u_D could be determined by Mean Residual Life plot method (Guess and Proschan 1985) and considered as constant during the extreme value distribution estimation procedure.

CASE STUDY: NANJING 3RD YANGTZE RIVER BRIDGE

The Nanjing 3rd Yangtze River Bridge, as shown in Figure 1, is one of the largest cable-stayed bridges in mainland China. It consists of a 648 m main span and two side spans that each has a 63 m and a 257 m section. The bridge towers consist of 4 transverse beams and are 215 m high. The tower below the bridge girder is constructed of pre-stressed concrete, and the rest of the tower is constructed of steel plates. There are a total of 4×21 pairs of stay cables.

Figure 1. Nanjing 3rd Yangtze River Bridge

During construction of the bridge, a sophisticated long-term structural health monitoring (SHM) system was designed and implemented and has been in operation since 2006. The SHM system incorporates a weight-in-motion (WIM) system, accelerometers, temperature sensors, and wind anemoscopes. The WIM system is located at the entrance to the bridge on both sides to monitor the traffic loads. The vehicle data format as well as the description of monitored vehicle parameter is presented in Table 1.

Table 1. Vehicle data format of WIM system

Data Name	Illustration	Type	Length
CardID	The card ID of each vehicle	Int	4
StatioID	The ID of toll-gate	Int	4
LaneID	The ID of vehicle lanes	SmallInt	2
PassTime	The date and time of each vehicle pass through	Datetime	8
VehicleClass	The type of vehicles: 1~6 passenger cars; 11~16 freight cars; 21~26 container cars;	Tinyint	1
AxisID	The ID of axis of each vehicle	Tinyint	1
AxisTyres	The number of tyres of each axis	Tinyint	1
AxisWeight	The weight of each axis	Int	4
AxisExceedWeight	The weight exceeded on each axis	Int	4
TotalWeight	The total weight of each vehicle	Int	4
TotalExceedWeight	The total weight exceeded on each vehicle	Int	4
Spare1		Int	4
Spare2		Int	4
Spare3		Int	4

In this study, the pass time and gross weights of vehicles collected by SHM from 2006 to 2009 were employed for the extreme value distribution estimation. Taking into account of congested traffic situations when the heavy vehicles are tightly jammed in one or more lanes, the normal status and dense status of the traffic flowing were considered. Figure 2 illustrates the histogram and empirical cumulative distribution function of vehicle weights at the normal and dense status respectively. It can be seen from Figure 2 that the truncated distribution of vehicle weights could not be fitted by log-normal distribution, which is recommended by Ministry of Transport of the PRC (1999).

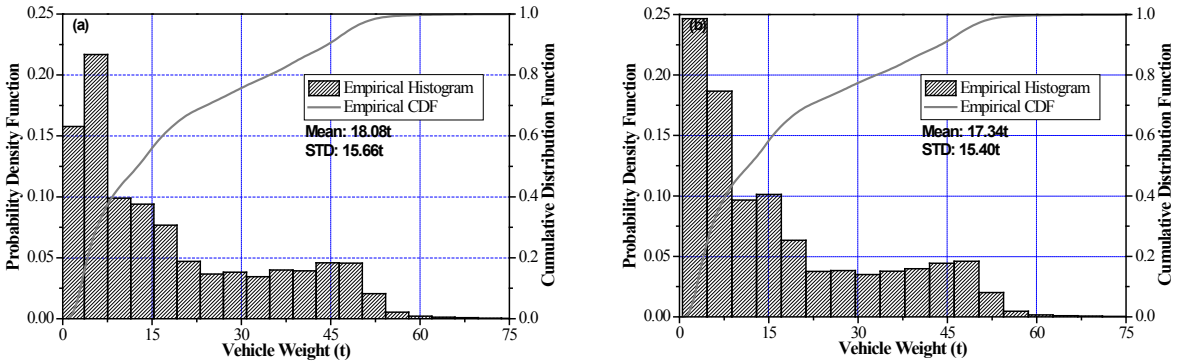

Figure 2. Histogram and empirical CDF of truncated distribution at (a) normal status and (b) dense status

Extreme Value Distribution Prediction of Vehicle Loads at Normal Status

A proper threshold u_N is critical for the extreme value distribution estimation. In this study, the mean residual life plot (Guess and Proschan 1985), as shown in Figure 3, were employed to determine the threshold of generalized Pareto distribution (GPD) in Eq. (4). It could be seen from Figure 3 that the mean excess of the observed vehicle weights over the threshold $u_N \geq 66.7\text{t}$, fluctuate around a straight line. Thus the exceedance above threshold $u_N \geq 66.7\text{t}$ conforms to the GPD. However, it should be noted that a high threshold may lead to a considerable loss of information, and therefore, the optimal threshold should be high enough to maintain GPD distribution but low enough to conserve sufficient information for the extreme value analysis. With respect to these two criteria, u_N=66.7t is selected as an appropriate threshold.

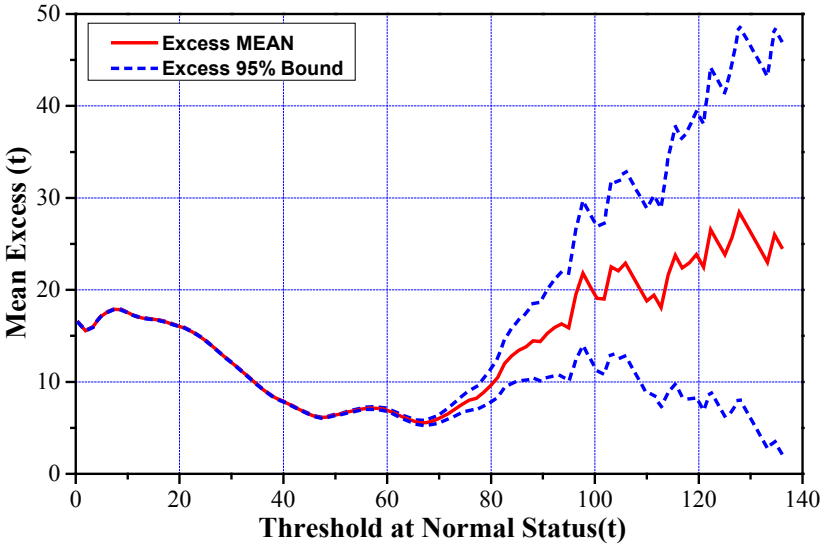

Figure 3. Mean Residual Life plot

In this study, the priors can be generated based on the monitored vehicle data, as shown in Figure 4. After using the Metropolis-Hastings Algorithm (Chib and Greenberg 1995, Gilks et al. 1996), a famous Markov Chain Monte Carlo (MCMC) method, with training data (weekly extreme values of the vehicle weights), the posterior distributions of the model parameters were obtained and are also shown in Figure 4. The estimated model

parameters $[\hat{\xi}_N, \hat{\sigma}_N, 1/\hat{\lambda}_N]$ were [0.2497, 7.8625, 58099] from the Bayesian MCMC method and [0.2122, 5.0506, 16014] from the traditional maximum likelihood estimation.

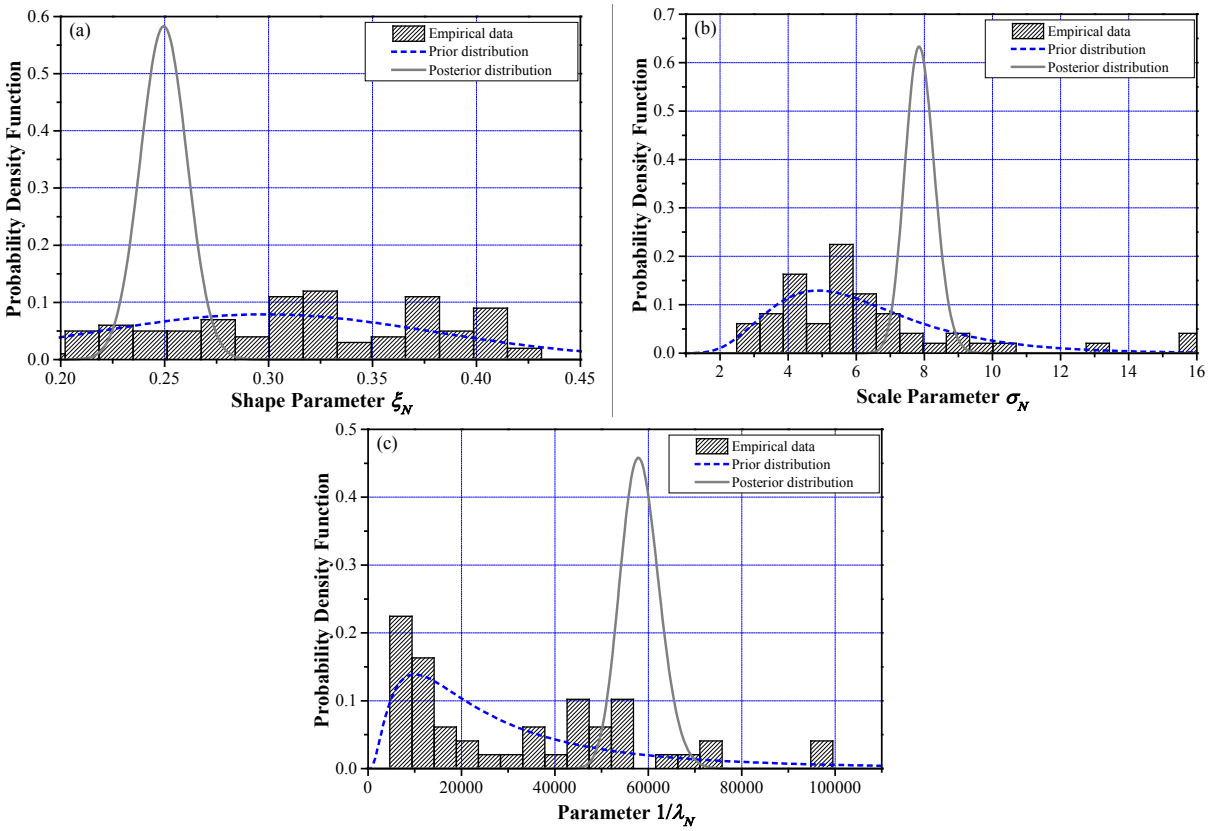

Figure 4. Prior and posterior distributions of model parameters at normal status

Once the model parameters $[\hat{\xi}_N, \hat{\sigma}_N, 1/\hat{\lambda}_N]$ were obtained, the extreme value distribution at normal status could be generated from Eq. (5a). Figure 5 shows the estimation comparison of weekly extreme value distribution using the proposed Bayesian MCMC method and the traditional maximum likelihood method. The figure shows that the Bayesian MCMC method shows much more agreement with the monitored extreme values than the traditional method.

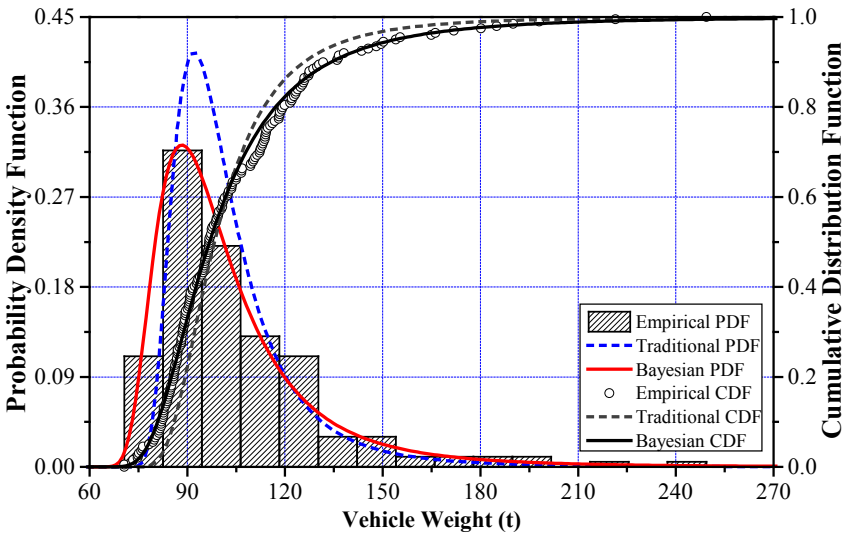

Figure 5. Comparison of the weekly extreme value distributions generated from the Bayesian model and the traditional model

The predicted extreme value distribution of vehicle weights under different reference periods calculated from Eq. (5a) are shown in Figure 6. The markers demonstrated in Figure 6 are the observed vehicle weights while the predicted distributions are shown as dashed lines. It can be seen from Figure 6 that the predicted extreme value distribution shows considerable coincidence with the observed one, which gives evidence for the correctness of the proposed method. Both the mean value and the standard deviation of extreme vehicle weights increased with the increase in reference periods.

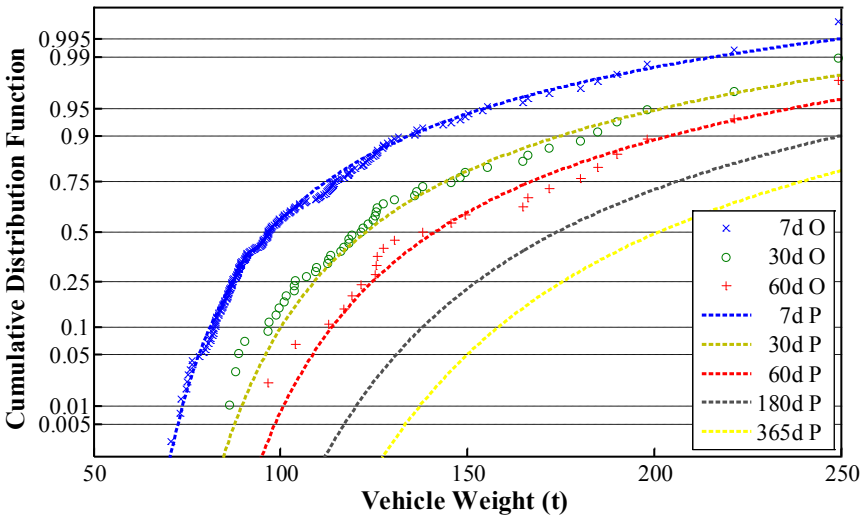

Figure 6. CDF Comparison of observed vehicle weights and the predicted distributions at normal status

Extreme Value Distribution Prediction of Vehicle Loads at Dense Status

For the vehicle weights at dense status, the threshold $u_N = 56.5$ could be obtained by mean residual life plot the same as the threshold selection procedure at normal status. Figure 7 shows the priors and the posterior distribution of model parameters at Eq. (5b) obtained by Bayesian MCMC method. The estimated model parameters $[\hat{\xi}_D, \hat{\sigma}_D, 1/\hat{\lambda}_D, \beta_D]$ were [0.2599, 5.0349, 188.2, 0.5910].

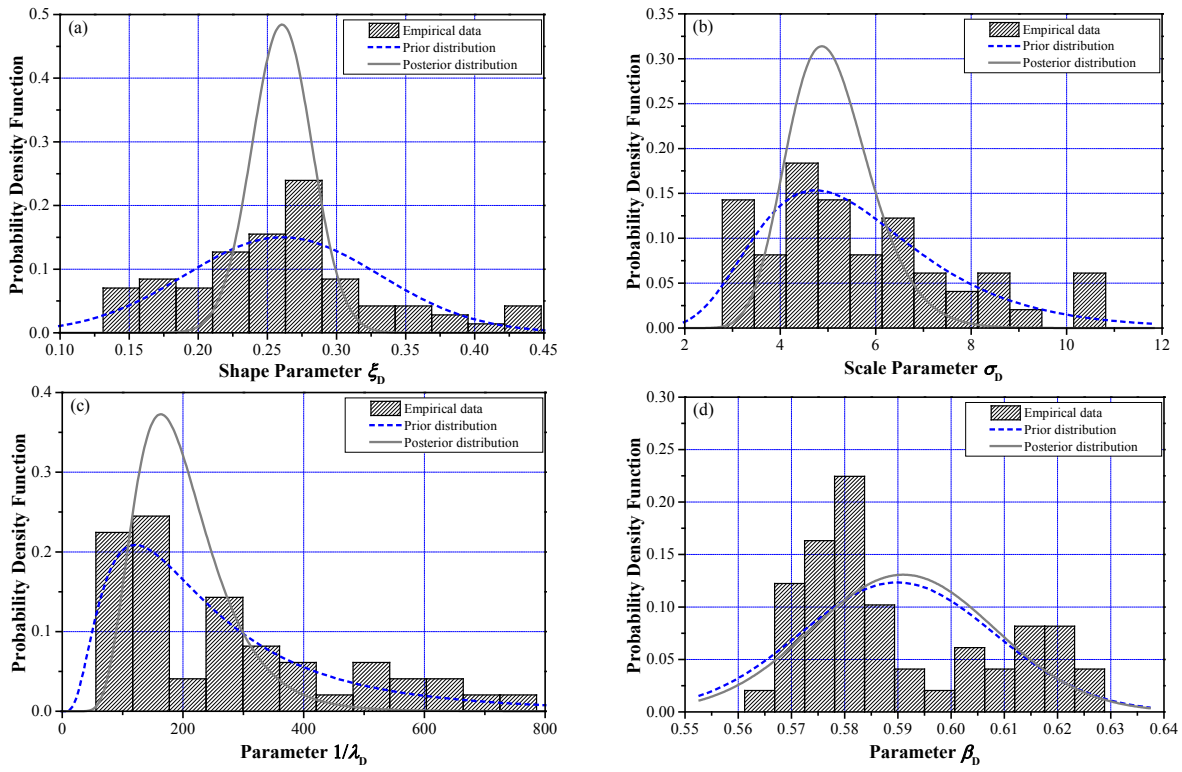

Figure 7. Prior and posterior distributions of model parameters at dense status

Figure 8 shows the CDF comparison between observed and predicted extreme value distribution at dense status. It can be seen from Figure 8 that the proposed method could generate weekly and monthly extreme value distribution in considerable agreement with the observed ones. Compared Figure 8 with Figure 6, the mean value and standard deviation of vehicle weights at normal status would be bigger than that at dense status, because the vehicle at the dense status mainly composed of small or mid-sized cars during the daytime, while the vehicles at the normal status contains the heavy trucks during the nighttime.

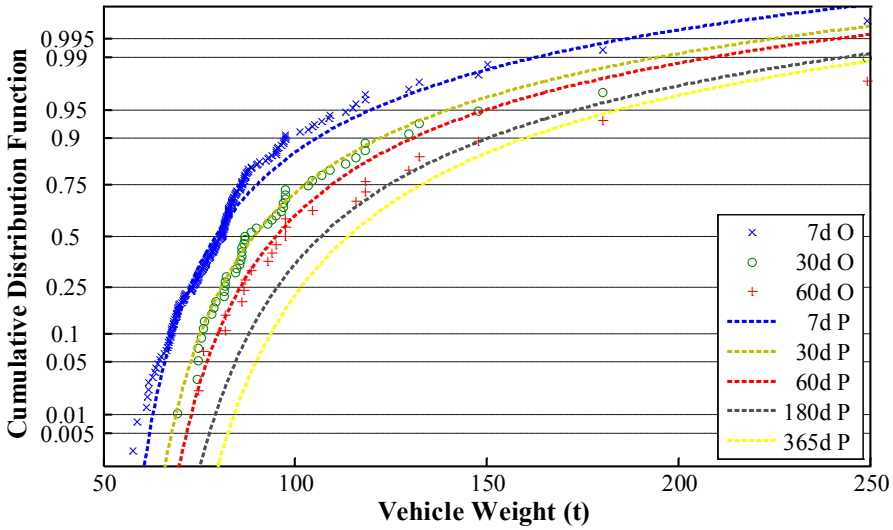

Figure 8. CDF Comparison of observed vehicle weights and the predicted distributions at dense status

CONCLUSIONS

This paper presents the Bayesian parameter estimation for extreme value distribution based on tail fitting. Analysis of collected vehicle data from the structural health monitoring system shows that the proposed Bayesian parameter estimation can lead to more satisfactory extreme value distributions. The following conclusions can be made:

(i) The inter-arrival times of vehicle weights at normal and dense status follow the exponential and Weibull distribution respectively. The application of the GPD tail distribution to the extreme value estimation was help for improvement on prediction accuracy.

(ii) The model parameters of the stochastic process and the GPD tail fitting were estimated by Bayesian MCMC method. The posterior distribution of model parameters could be achieved by incorporating the likelihood function and the prior distribution. The mean values of the Bayesian MCMC simulation samples were chosen as the model parameter and the extreme value distribution at different reference periods could be predicted.

It should be pointed that the investigated extreme value distribution models are based on the past monitoring data of traffic loads in this study. However, the parameters of Bayesian extreme value distribution can be easily updated by the new monitored datasets and the accuracy of parameter estimation could be improved with the growth of the accumulative data volume.

ACKNOWLEDGMENTS

Financial support for this study was provided by NSFC (Grant No. 51008096), MOST (Grant No. 2011BAK02B02, 2013CB036305) and the Fundamental Research Funds for the Central Universities (Grant No. HIT. NSRIF. 2013075).

REFERENCES

Chib, S. and Greenberg, E. (1995). “Understanding the metropolis-hastings algorithm”, *The American Statistician*, 49(4), 327-335.

Croce, P. and Salvatore, W. (2001). “Stochastic model for multilane traffic effects on bridges”, *Journal of Bridge Engineering*, 6(2), 136-143.

Ditlevsen, O. and Madsen, H.O. (1994). “Stochastic vehicle-queue-load model for large bridges”, *Journal of Engineering Mechanics*, 120(9), 1829-1847.

Ghosn, M. and Moses, F. (1985). “Markov renewal model for maximum bridge loading”, *Journal of Engineering Mechanics*, 111(9), 1093-1104.
Gilks, W.R., Richardson, S. and Spiegelhalter, D.J. (1996). *Markov Chain Monte Carlo in Practice*, Chapman & Hall/CRC, London.
Guess, F. and Proschan, F. (1985). “Mean residual life: theory and applications”, *AFOSR Technical Report No. 85-178*, 1-17.
Lan, C., Li, H. and Ou, J. (2011). “Traffic load modelling based on structural health monitoring data”, *Structure and Infrastructure Engineering*, 7(5), 379-386.
Ministry of Transport of the PRC (1999). *Unified Standard for Reliability Design of Highway Engineering Structures (GB/R 50283-1999)*, China Planning Press, Beijing.
Moses, F., Schilling, C. G. and Raju, K. (1987). “Fatigue evaluation procedures for steel bridges”, *NCHRP Report*, 1-299.
Nowak, A.S. (1993). “Live load model for highway bridges”, *Structural Safety*, 13(1), 53-66.
OBrien, E.J., Enright, B. and Getachew, A. (2010). “Importance of the tail in truck weight modeling for bridge assessment”, *Journal of Bridge Engineering*, 15(2), 210-213.
Okada, K., Okamura, H. and Sonoda, K. (1978). “Fatigue failure mechanism of reinforced concrete bridge deck slabs”, *Transportation research record*, 664, 136-144.
Tong, G., Aiqun, L. and Daliang, Z. (2008). “Multiple-peaked probabilistic vehicle load model for highway bridge reliability assessment”, *Journal of Southeast University (Natural Science Edition)*, 38, 763-766.
Yang, S.-I., Frangopol, D.M. and Neves, L.C. (2004). “Service life prediction of structural systems using lifetime functions with emphasis on bridges”, *Reliability Engineering & System Safety*, 86(1), 39-51.