Engineering Structures 59 (2014) 386–392

Contents lists available at ScienceDirect

Engineering Structures

journal homepage: www.elsevier.com/locate/engstruct

A novel Bayesian extreme value distribution model of vehicle loads incorporating de-correlated tail fitting: Theory and application to the Nanjing 3rd Yangtze River Bridge

Yiming Gu [a,b], Shunlong Li [a], Hui Li [a,*], Zhiming Guo [c]

[a] School of Transportation Science and Engineering, Harbin Institute of Technology, 150090 Harbin, China
[b] School of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, United States
[c] Administration Agency of Nanjing, No. 3 Yangtze River Bridge, Nanjing 210000, China

ARTICLE INFO

Article history:
Received 1 January 2012
Revised 16 October 2013
Accepted 17 October 2013
Available online 6 December 2013

Keywords:
Vehicle loads
Extreme value statistics
De-correlated tail fitting
Bayesian updating
Structure health monitoring

ABSTRACT

Vehicle loads play an important role in fatigue deterioration and overload collapses of bridges. In this paper, a novel de-correlated tail-based extreme value (EV) distribution model of vehicle load is proposed. The monitored data show that occurrences of vehicle loads are correlated. Additionally, it is more reasonable to employ the tail region of a distribution when estimating extreme loads. Moreover, a Bayesian form of this new model is constructed, and an extension of this model, the Confidence Index (CI), is defined and may be promising for applications. The monitored vehicle weight on the Nanjing 3rd Yangtze River Bridge is used to demonstrate that the proposed tail-based de-correlated EV model predicts the extreme load more accurately than traditional methods and that the Bayesian approach can further increase the precision of this estimate. Finally, the calculated CI of the complete prediction process offers a comprehensive guideline for the estimate precision.

1. Introduction

Being an essential component of a transportation system, long-span bridges have a significant impact on the national economy and on social stability. Authorities worldwide have recognized the importance of structural health monitoring (SHM) for bridge safety to ensure ordinary operation during the structure's service lifetime. SHM systems, in conjunction with proper software, can identity possible defects in the bridge structure, forecast the potential risk of collapse, and predict the remaining service time of the bridge [5,9].

Monitored traffic action, especially the extreme values (EVs) of vehicle loads, is one of the most important measures for determining fatigue deterioration [21,15,17,14] and predicting the remaining service time [27,7,8] of long-span bridges. The demand for the transportation of goods that is associated with a prosperous economy has led to increasingly serious bridge overload. In fact, recent collapses of transportation bridges as a result of overloading from trucks have attracted much attention. Load restrictions and fines for overloading might help to control the traffic loads on bridges. However, current bridge design codes consider neither the phenomenon of increasingly serious overload nor the actual effect of load-controlling methods on the EV distribution of traffic loads.

Over the past decades, a considerable amount of research has been conducted related to vehicle loads modeling on bridge [18,20,28]. With the help of structural health monitoring techniques, the EV distribution can be estimated based on the monitored traffic load data [10,13], 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. [9], and the extreme value distribution function of vehicle load during the service period was extrapolated. Tong et al. [24] developed a multi-peak probability model to estimate the extreme value distribution of vehicle loads. OBrien et al. [19] proposed a semi-parametric distribution model based on in-field data from two sites. However, in comparison with the monitored data, the predictions generated by these traditional methods are far from satisfactory because the traditional methods involve several flawed assumptions, such as the following: (1) vehicle loads are independent of one another; and (2) the fitting of the entire distribution range is precise. A study of monitored traffic loads shows that the extreme values of the traffic loads are actually correlated and that the tail region of the distribution range, which is much more important for the EV estimate, is far from accurate.

In Sections 2 and 3, a general method for solving the above problems is presented. Specifically, in Section 2, the distribution of the tail, rather than the distribution of the entire range, is

* Corresponding author.
E-mail address: lihui@hit.edu.cn (H. Li).

0141-0296/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.engstruct.2013.10.029

employed during the process of estimating the EV. Additionally, the correlation of vehicle loads is excluded using a Peak Over Threshold (POT) method. In Section 3, a novel Bayesian approach to this de-correlated tail-based model is presented, and two possible applications, Bayesian updating (BU) and a Probabilistic Stochastic Model (PSM), are consecutively proposed. Section 4 illustrates a practical application of the methods proposed in Sections 2 and 3 for the Nanjing 3rd Yangtze River Bridge. An analysis of the collected data shows that excluding the correlations, employing a tail distribution model and incorporating the Bayesian model can lead to a more satisfactory estimate of EV distributions.

2. EV distribution estimation based on de-correlated tail

As recommended in the Unified Standard for Reliability Design of Highway Engineering Structures (GB/T 50283-1999), the stochastic process of traffic loading conforms to a Filtered Poisson Process (FPP). The cumulative density function (CDF) of the EV distribution for a FPP is:

$$F_{eV}(x) = \exp\{-\lambda T(1 - F(x))\}, \tag{1}$$

where $F(x)$ is the section distribution of the traffic loads, λ is the intensity parameter of the stochastic process, and T is the return period required for estimating EVs.

To estimate the EV distribution of the traffic loads, the section distribution of the traffic loads should first be obtained. Many previous studies have noted that the section distribution of the entire range does not reflect the tail region, which is mainly responsible for precise EV estimation. Thus, the tail distribution is incorporated into the EV distribution estimate in this study.

2.1. GPD fit of tail distribution

The Generalized Pareto Distribution (GPD), which was proven (see [1] to be the limit distribution of scaled excess over high threshold values under certain conditions, is often used to fit tail datasets. The cumulative density function of the GPD is given as follows:

$$F_{\gamma,\sigma}(x) = 1 - \left[1 + \frac{\gamma(x-u)}{\sigma}\right]^{-\frac{1}{\gamma}}, \tag{2}$$

where σ and γ represent the scale and shape parameters, respectively, and u is the chosen threshold. Substituting Eq. (2) into Eq. (1), the cumulative density function of the EV distribution can be obtained as follows:

$$F_{\gamma,\sigma,\lambda}(x) = \exp\left\{-\lambda T\left[1 + \frac{\gamma(x-u)}{\sigma}\right]^{-1/\gamma}\right\}. \tag{3}$$

2.2. Select threshold

A proper threshold is critical for estimating the EV distribution. To choose a proper threshold, the Mean Excess Function (MEF) of the GPD [22] is employed and is defined as:

$$e(u) = E(x - u|x > u) = \frac{\sigma + \gamma u}{1 - \gamma}. \tag{4}$$

Eq. (4) shows that $e(u)$ is a linear function of the threshold, u. From the definition of the MEF, it's sample estimator is:

$$e_n(u) = \frac{1}{N_u}\sum_i (x_i - u), \tag{5}$$

where N_u is the number of the samples excess threshold u, and x_i is the sample values that excess threshold u. If the excess above a threshold, u_0, conforms to the GPD, the MEF of the samples over the threshold, u_0, will fluctuate around a straight line. The point set $\{u, e_n(u)\}$ is considered to be the mean residual life plot (MSLP) [4]. Therefore, the appropriate choice of the threshold, u_0, should satisfy the condition that the $u > u_0$ region of the MSLP approaches a straight line.

However, with increasing threshold, the Standard Deviation (STD) of sample estimator of MEF is $e_n(u)$ is also increasing:

$$\mathrm{STD}[e_n(u)] = \sqrt{\frac{\sigma^2}{N_u(1-r)^2(1-2r)}}.$$

In order to capture the increase of linearity as well as the loss of credibility (increasing variance), Standardized Residual (SR) is defined as:

$$\mathrm{SR}(u) = e_n(u) - e_{fit}(u) + STD[e_n(u)], \tag{6}$$

where $e_{fit}(u)$ is the linear fit of the MRL above u and evaluated at u.

2.3. Exclude data correlation

Generally, the basic assumption that the loads are independent of one another is used for the EV estimate. However, based on an analysis of a large amount of monitored data, it is found that this assumption is not correct in some cases. For example, the same heavy truck may pass over a bridge and then return on that same bridge. Heavy vehicles may pass simultaneously as a motorcade, and motorcades of heavy trucks have been frequently monitored by the SHM system. Therefore, it becomes necessary to exclude the correlation of vehicles.

In this study, only the consecutive time correlation of vehicles, such as a motorcade, is considered. The Peak Over Threshold (POT) method [16] is well developed and widely employed in extreme value statistics. The algorithm of the POT method is:

(1) select clusters of exceedance over the threshold according to a certain criterion;
(2) select the maximum of each cluster, assuming that the maximum values of each cluster are independent;
(3) fit these maximum values to GPD.

The clustering method in POT determines the de-correlation functionality of that POT algorithm. For example, if the data are clustered in 1 min of time, the temporal correlation within 1 min is mostly eliminated [11], and these 1 min cluster maximums can be treated as independently and identically distributed if the data has an auto-correlation with a lag less than 1 min.

Other methods that account for data correlation in the EV estimate, including longer-span temporal and spacial correlations, will be studied in the future.

3. Estimate Bayesian EV

The occurrence of traffic loads may change during the operation of a bridge. Therefore, the traffic load model should be updated accordingly to reflect the real status of traffic loads based on the monitored data. In this study, a Bayesian analysis [3,23,25,26] is employed to update the traffic load model using the monitored data. The Bayesian analysis is performed by combining the prior information, $\pi(\boldsymbol{\theta})$, and the sample information, $\mathbf{x}$, into the posterior distribution of $\boldsymbol{\theta}$ given the sample observation, $\mathbf{x}$. Noting that $\boldsymbol{\theta}$ and $\mathbf{x}$ have joint density:

$$h(\mathbf{x}, \boldsymbol{\theta}) = L(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}), \tag{7}$$

where $\pi(\boldsymbol{\theta})$ is the prior distribution of the model parameters, $\boldsymbol{\theta}$, and $L(\bullet)$ is the likelihood function, and thus, $L(\mathbf{x}|\boldsymbol{\theta})$ is the conditional likelihood of sample $\mathbf{x}$ given the parameter $\boldsymbol{\theta}$.