

车辆荷载作用下桥梁应变极值估计的 阈值选取*

阳霞¹, 张静¹, 任伟新¹, 袁平平²

(1. 合肥工业大学 土木与水利工程学院, 合肥 230009;

2. 中南大学 土木工程学院, 长沙 410075)

(我刊编委任伟新来稿)

摘要: 采用过阈法估计车辆荷载作用下桥梁的应变极值,合理的阈值选取十分关键.阈值选取过大,信息量少,阈值选取过小,广义 Pareto 分布模型参数估计偏差大.常用的阈值选取方法不能较好地适用于车辆荷载作用下的应变极值估计.基于太平湖大桥车辆荷载作用下 1 年的应变数据,对拟合结果较好的 3 种混合分布进行 Monte-Carlo(蒙特-卡洛)抽样,对比同一样本基于不同阈值的广义 Pareto 分布模型的极值估计结果,提出了一种经验式的阈值选取方法.与常用阈值选取方法相比,根据文中方法所得阈值估计的周应变极值分布与实测结果更为接近,估计结果更好.

关键词: 桥梁健康监测; 阈值选取方法; 应变极值; 广义 Pareto 分布

中图分类号: U441

文献标志码: A

doi: 10.21656/1000-0887.370395

引 言

车辆荷载是车行桥最主要的可变荷载,我国现行规范采用的车道荷载是根据多地实测的车辆荷载数据基于可靠度理论得到的^[1],其取值主要注重通用性,与具体桥梁的实际车辆荷载模式不一定相符.直接采用现行规范的车道荷载对桥梁进行安全评估,容易误判,从而导致因加固改造带来的资源浪费或忽略潜在的危险.尤其对于广泛存在的老旧桥梁,其设计建造时采用的标准与现行不同,若能通过实测数据估计桥梁车辆荷载及其效应的极值分布,则可较为准确地评估其在役期间的服务能力.

桥梁的实测车辆荷载效应通常呈多峰分布^[2-3],较为复杂.以往关于车辆荷载及其效应的模型研究,通常采用单个分布或 2~3 个概率模型加权的混合分布对采集的全部数据直接拟合,在高分位处容易产生较大偏差,且这些偏差在极值估计过程中被累积放大,从而大大降低极值估计的精度.高速公路管理部门针对不同车型,依据其装载乘客和货物能力的大小将车辆划分为客车、货车 2 大类,9 小类;王涛^[4]根据交通荷载调查结果,将采集到的车辆样本划分为 5 大类.因此若要比准确地拟合所有车辆荷载数据,需采用 5~9 个函数加权的混合分布,待估参数多达 20 多个,计算量庞大,且极易产生误差.

* 收稿日期: 2016-12-26; 修订日期: 2017-03-24

基金项目: 中国博士后科学基金(2015M581982);安徽省自然科学基金(1708085ME101)

作者简介: 阳霞(1986—),女,博士生(E-mail: yangx@mail.hfut.edu.cn);

张静(1984—),女,副教授,博士,硕士生导师(通讯作者. E-mail: zhangj@hfut.edu.cn).

极值理论指出^[5],极值估计通常只与概率密度函数的高尾分布相关,即车辆荷载效应的极值估计应只与重型车有关.广义 Pareto 分布 (generalized Pareto distribution, GPD) 模型作为样本右尾数据的拟合模型,广泛运用于洪水高度、风速、气温等的极值估计中,近年来在车辆荷载的极值研究中也得到了较好的应用^[6].研究表明^[7-8],基于 GPD 模型的极值估计结果与阈值的选取密切相关,对于给定的数据集,如何选择一个适当的阈值尤为关键.阈值偏大,数据量小,信息少,GPD 模型参数(形状参数)估计值的变异性增大;阈值较小,数据量大,信息多,模型参数估计值的偏差增大.目前常用的 GPD 模型阈值估计方法主要有超阈值均值图法、Hill 图解法、峰度法、均方误差最小法^[9-10].基于此类方法选取的阈值对于车辆荷载作用下的应变估计结果均不理想,不能很好地应用于工程实践.

本文基于安徽省太平湖大桥车辆荷载效应下 1 年的动应变数据,选取其右尾数据进行拟合,根据拟合效果较好的模型对应变右尾概率模型(底分布)进行假定,然后利用 Monte-Carlo 方法抽样,产生大量不同时长的子样,比较同一样本,基于不同阈值条件下 GPD 模型的极值估计结果及理论值,研究 GPD 模型的阈值选取方法,从而较好地估计出桥梁基准期内的应变极值分布,为在役桥梁的可靠性研究及安全评估提供依据,并为相同地区的桥梁设计提供参考.

1 车辆荷载作用下应变峰值分布模拟

图 1 为太平湖大桥主梁某测点剔除温度影响后,仅由车辆荷载引起的应变时程曲线.图中每根“毛刺”均代表一个车辆通过桥梁时测点产生的应变,定义每根“毛刺”的最大值为应变峰值.选取所有大于 1.8×10^{-5} 的应变峰值作为待拟合数据,分别采用 3 个混合分布(3 个正态分布的加权、一个对数正态分布与两个正态分布的加权、一个 Weibull 分布与两个正态分布的加权)进行拟合,拟合结果如图 2 所示.3 个混合分布的拟合效果较好,均通过 K-S 检验,其中以 1 个 Weibull 分布与两个正态分布加权的模型拟合效果最佳.根据拟合结果,对车辆荷载作用下应变峰值的右尾分布(底分布)进行了以下 4 种假定,每个混合分布均为 3 个简单分布的加权:

- 1) 3 个正态分布^{①②③};
- 2) 1 个对数正态分布^①+2 个正态分布^{②③};
- 3) 1 个 Weibull 分布^①+2 个正态分布^{②③};
- 4) 1 个 Weibull 分布^①+2 个正态分布^{②③}.

底分布 3) 与 4) 的概率模型相同,但选取的参数不同,具体如表 1 所示. P_1, P_2, P_3 表示各分布的权重系数, μ_N, σ_N 为正态分布的均值和标准差, μ_L, σ_L 为对数正态分布的均值与标准差, σ_W, ξ_W, μ_W 分别表示 Weibull 分布的尺度参数、形状参数和位置参数.根据太平湖大桥实测数据,假设每天通过的重车(最重的 3 种车型)为 100 辆,利用 Monte-Carlo 方法分别抽取采样时长为 1 天、1 个月、6 个月、1 年、10 年、50 年的子样本各 100 组,即各子样本容量分别为 100, 3 000, 18 000, 36 000, 360 000, 1 800 000.

表 1 底分布参数

Table 1 Parameters of the parent distributions

parent distribution	P_1	P_2	P_3	parameters of ①	parameters of ②	parameters of ③
1)	0.6	0.25	0.15	$\mu_N = 26, \sigma_N = 0.8$	$\mu_N = 30, \sigma_N = 4$	$\mu_N = 40, \sigma_N = 6$
2)	0.75	0.15	0.1	$\mu_L = 3, \sigma_L = 0.2$	$\mu_N = 28, \sigma_N = 2.5$	$\mu_N = 35, \sigma_N = 5$
3)	0.7	0.2	0.1	$\sigma_W = 5, \xi_W = 1.2, \mu_W = 32$	$\mu_N = 37, \sigma_N = 1.5$	$\mu_N = 40, \sigma_N = 4$
4)	0.75	0.1	0.15	$\sigma_W = 4.5, \xi_W = 1.3, \mu_W = 20$	$\mu_N = 20, \sigma_N = 1$	$\mu_N = 23, \sigma_N = 1.5$

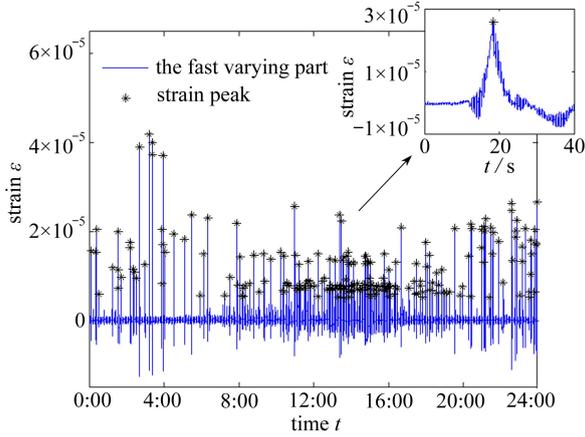


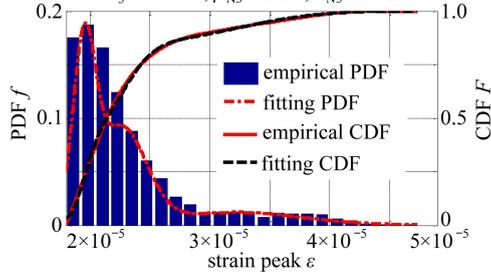
图1 车辆荷载引起的应变峰值(2015-12-27)

Fig. 1 Strain peaks induced by vehicle load (2015-12-27)

$$P_1=0.3666, \mu_{N1}=18.54, \sigma_{N1}=0.94$$

$$P_2=0.4705, \mu_{N2}=22.61, \sigma_{N2}=2.09$$

$$P_3=0.1629, \mu_{N3}=32.24, \sigma_{N3}=5.48$$



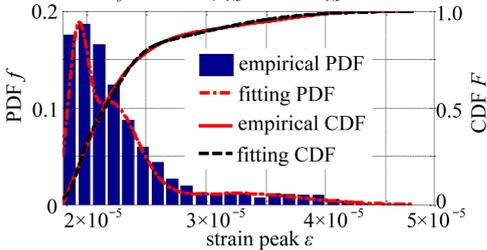
(a) 3个正态分布的加权

(a) The mixed distribution of 3 normal distributions

$$P_1=0.4055, \mu_{L1}=2.98, \sigma_{L1}=0.052$$

$$P_2=0.4388, \mu_{N2}=22.82, \sigma_{N2}=2.12$$

$$P_3=0.1558, \mu_{N3}=32.56, \sigma_{N3}=5.38$$



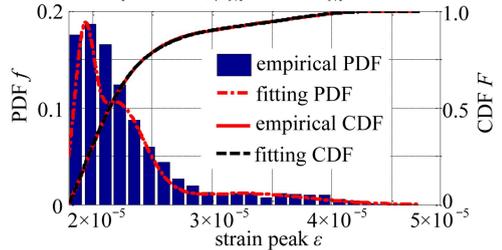
(b) 一个对数正态分布与两个正态分布的加权

(b) The mixed distribution of 1 lognormal distribution and 2 normal distributions

$$P_1=0.7592, \mu_{W1}=18, \sigma_{W1}=5.60, \zeta_{W1}=0.9797$$

$$P_2=0.0955, \mu_{N1}=19.92, \sigma_{N1}=0.72$$

$$P_3=0.1453, \mu_{N1}=21.99, \sigma_{N1}=1.34$$



(c) 一个Weibull分布与两个正态分布的加权

(c) The mixed distribution of 1 Weibull distribution and 2 normal distributions

图2 3个混合分布的拟合结果

Fig. 2 Fitting results of 3 mixed distributions

由于篇幅限制,本文就底分布3)子样的阈值估计和阈值选取作详细说明,其他3种底分布仅介绍其结论。

1.1 阈值估计

(a) 超阈值均值图法.超阈值均值图法是选取 GPD 模型阈值最常用的一种方法^[4,11].此法缺陷在于阈值的选取依赖于主观判断,易带主观误差.采用此法计算底分布 3) 各子样的超阈值均值图如图 3 所示.

(b) Hill 图法.令 $X_{(1)} > X_{(2)} > \dots > X_{(n)}$ 表示独立同分布的顺序统计量^[4],尾部指数的 Hill 统计量定义为

$$H_{k,n} = \frac{1}{k} \sum_{i=1}^k \ln \left(\frac{X(i)}{X(k)} \right), \quad k \leq n. \quad (1)$$

Hill 图定义为点 $(k, H_{k,n}^{-1})$ 构成的曲线,选取 Hill 图形中尾部指数的稳定区域起始点的横坐标 k 所对应的阈值.此法的缺陷与超阈值均值法相同,阈值的选取依赖于主观判断.底分布 3) 各子样本的 Hill 估计图如图 4 所示.

(c) 峰度法.Mcneil 和 Frey 提出了厚尾分布与正态分布相交法来确定阈值^[12],由于计算过程极为复杂难以运用于实际求解,Pierre 在此基础上提出峰度法^[13],易理解且便于运用.相较于前两种图解法,峰度法具有理论优势,且不存在主观误差.

(d) 均方误差最小法.阈值的选择直接影响模型参数估计的准确性,Caers, Maes 等^[14-15]建议采用参数估计(形状参数)的最小均方误差准则,通过试算法来选择尾部数据.此方法的计算量庞大,故一般是基于图形分析法(超阈值均值法和 Hill 图法)对阈值进行初估后,再使用此法对阈值进一步进行择选.

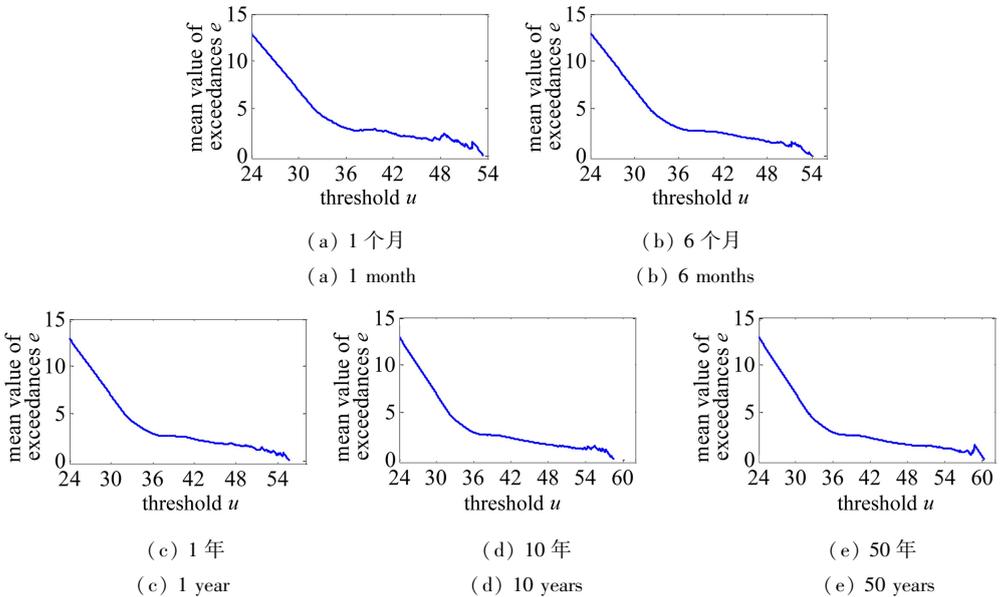


图 3 不同时长样本的超阈值均值图

Fig. 3 Mean excess function plots for sample data for different sampling periods

据以上 4 种方法估计底分布 3) 各子样的阈值如表 2 所示.从表可以看出,图形分析法估计的阈值随着样本容量增大而增大,均方误差最小法与峰度法的估计结果受样本容量影响较小,峰度法的估计结果较其他方法偏大.

其余 3 个底分布的阈值估计结果与上述规律基本一致,另外,底分布 1)、2) 采用图形分析法时,存在较大缺陷,自同一底分布抽取的不同子样本的超阈值均值图和 Hill 估计图显示出较

大的不稳定性,80%以上的子样难以判断“转折点”的位置,即阈值;不同底分布采用峰度法估计的阈值较其他方法相差较大,且时而偏大,时而偏小;均方误差最小法计算量太大,估计的阈值与图形分析法较为接近。

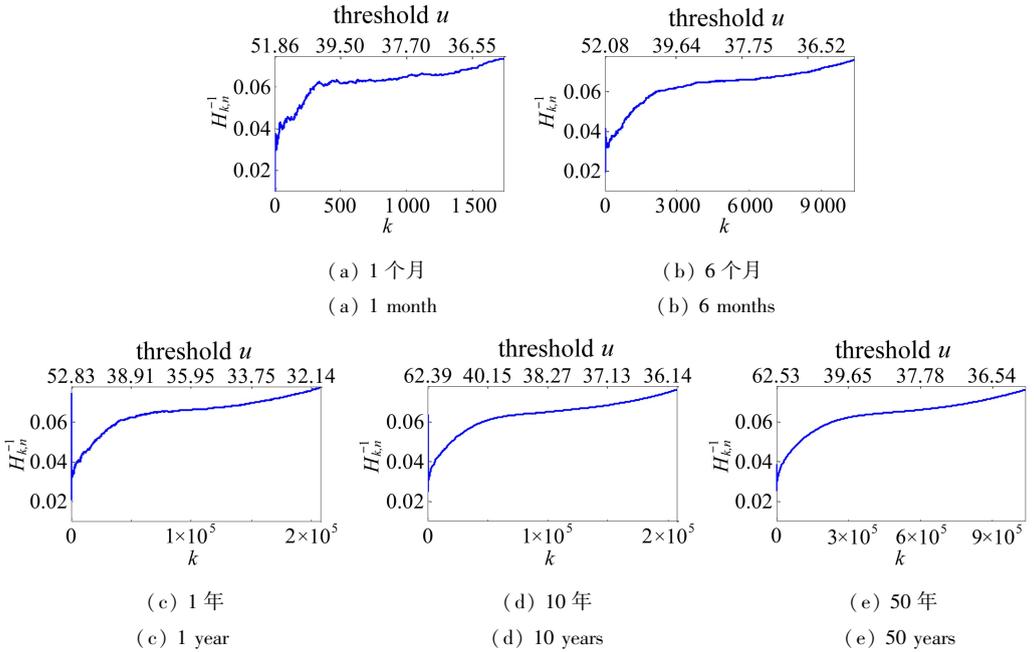


图 4 不同时长样本的 Hill 估计图

Fig. 4 Hill plots of sample data for different sampling periods

表 2 常用 4 种方法估计的阈值

Table 2 Thresholds estimated with the 4 common methods

method	sampling period t/d				
	30	180	360	3 600	18 000
mean excess function	39.6	40.5	41.0	42.4	42.9
Hill estimator	40.64	40.69	41.22	42.20	45.06
kurtosis method	45.12	45.15	45.15	45.15	45.15
minimum MSE method	42.50	42.50	43.00	44.00	44.50

表 3 超阈值样本数

Table 3 The numbers of exceedances

sampling period t/d	threshold u				
	44	46	48	50	52
1	3.49	1.55	0.58	0.20	0.05
30	99.87	39.45	13.67	4	1.04
180	593.32	239.41	80.78	24.15	6.35
360	1 187	471.58	162.30	50.14	13.12
3 600	11 827	4 748.55	1 637.62	495.10	133.26
18 000	59 132	23 651.63	8 219.55	2 481.29	672.26

1.2 阈值选取对极值估计的影响

根据上述阈值估计的分析结果,针对底分布 3) 选取不同阈值,分别采用 GPD 模型拟合各

子样本的超阈值,并估计 100 年内的极值分布.选取不同阈值时各子样本的超阈值样本数如表 3 所示.采用极大似然方法估计 GPD 模型参数^[16],并定义累积误差

$$e_a = \frac{1}{N} \sum_{i=1}^N | F(x_{i,N}) - \hat{F}(x_{i,N}) | \tag{2}$$

用于衡量不同阈值的 GPD 模型的拟合优度.式中 $F(x_{i,N})$ 为第 i 个次序统计量 $x_{i,N}$ 对应的经验累积概率, $\hat{F}(x_{i,N})$ 为根据 GPD 模型推算的 $x_{i,N}$ 对应的累积概率.不同采样时长的子样本数均为 100 组,将式(2)所得的 100 组样本的累积误差取平均,所得的平均累积误差见表 4.计算各子样本选取不同阈值时 100 年内极值的期望、方差及 95%的置信区间,结果如表 5 所示.“-”表示因数据太少,无法计算结果; μ, σ 分别是 100 组子样估计极值期望的均值与标准差;95% CI 表示 95%置信区间.基于底分布 3) 的理论模型已知,即可得其 100 年内极值期望的理论值为 62.82.

表 4 GPD 模型对不同子样的拟合优度
Table 4 Fitting goodnesses of different samples fitted by the GPD models

sampling period t/d	threshold u				
	44	46	48	50	52
30	0.026 3	0.037 7	0.066 0	0.159 3	-
180	0.018 7	0.018 2	0.025 0	0.045 3	-
360	0.017 9	0.015 4	0.018 6	0.029 7	0.062 2
3 600	0.017 7	0.013 7	0.010 8	0.011 5	0.018 4
18 000	0.017 6	0.013 6	0.009 7	0.006 5	0.009 2

表 5 极值估计结果(理论值: $E = 62.82$)
Table 5 Results of extreme value estimation (theoretical value: $E = 62.82$)

sampling period t/d		threshold u				
		44	46	48	50	52
30	(μ, σ)	(70.01, 7.77)	(68.69, 4.81)	(65.61, 5.76)	(63.34, 6.29)	-
	95% CI	(65.53, 74.63)	(61.77, 71.26)	(57.50, 73.79)	(52.29, 74.23)	-
180	(μ, σ)	(69.25, 6.50)	(66.99, 4.40)	(65.25, 3.05)	(63.43, 2.89)	-
	95% CI	(67.41, 70.79)	(64.71, 69.61)	(61.24, 68.74)	(58.91, 69.01)	-
360	(μ, σ)	(69.21, 6.42)	(66.56, 3.83)	(65.38, 2.66)	(64.25, 2.63)	(63.42, 3.10)
	95% CI	(68.11, 70.20)	(65.22, 68.07)	(62.70, 68.45)	(60.19, 68.04)	(58.20, 69.03)
3 600	(μ, σ)	(69.16, 6.34)	(66.62, 3.81)	(64.78, 2.00)	(63.82, 1.25)	(63.32, 1.14)
	95% CI	(68.77, 69.57)	(66.19, 67.13)	(63.96, 65.43)	(62.62, 65.66)	(61.68, 65.04)
18 000	(μ, σ)	(69.16, 6.34)	(66.50, 3.78)	(64.83, 2.02)	(63.60, 0.97)	(63.38, 0.84)
	95% CI	(68.98, 69.34)	(66.38, 66.82)	(64.50, 65.20)	(63.25, 64.23)	(62.31, 64.48)

从表 5 可以看出,极值的估计结果需均衡考虑偏差与方差,建议引入最小均方误差准则判断最优阈值:

$$F_{MSE} = \text{bias}^2(E - \mu) + \text{var}(E), \tag{3}$$

式中 E 表示极值期望的理论值, μ 表示估计极值期望的均值, $\text{bias}(E - \mu)$ 为极值估计的偏差, $\text{var}(E)$ 表示极值估计的方差,使两者之和即 F_{MSE} 最小的阈值为最优解.假设工程实践中只接受 $\pm 10\%$ 的误差,即估计值在 [56.54, 69.10] 范围内为可以接受.

分析表 3~5 可知,GPD 模型的极值估计结果较理论值整体偏大.基于 1 天的数据估计桥梁基准期内的极值分布是不可取的,可用的数据太少;采用 1 个月的数据也无法获得理想的极

值估计,估计结果的误差超过 10%,而阈值取 50 时,结果相对最优;采样时长为 6 个月,阈值取 48 时,极值的估计结果在可接受范围内;采样时长为 1 年,阈值取 46~52 时,估计结果均比较理想,取 48 最优;当采样时长为 10 年或 50 年,阈值取 46~52 时,估计结果均可接受,取 52 最优.当阈值取 44 时,不同时长子样的估计结果基本一致,偏差与方差都较大,95%置信区间较窄;随着阈值的增大,偏差与方差迅速减小,且减小的速度随阈值增大而放缓,95%置信区间宽度增加;继续增大阈值,偏差缓慢减小,而方差迅速增大,95%置信区间持续变宽,估计结果愈来愈不稳定.对比表 4 与表 5 可知,GPD 模型拟合效果最好时的阈值并不等于极值估计最优时的阈值.其余 3 个底分布关于阈值对估计极值的影响规律与以上大致相同.

比较表 2 与表 5 可以看出,采用超阈值均值法、Hill 图法、均方误差最小法估计的阈值均较小,但比较稳定,而峰度法估计的阈值则不稳定,特别对于底分布 1) 和 4),估计阈值与最优阈值相差很大.

通过对比分析 4 个底分布的抽样基于不同阈值的 GPD 模型估计的极值结果,针对车辆荷载作用下应变峰值,本文提出一种基于阈值-估计极值期望关系曲线图的阈值选取方法.具体步骤为:建立阈值与估计极值(100 年内)期望的关系曲线图,观察估计极值随阈值的变化,(i)若样本数量较小,即随着阈值增大,极值期望随之减小,且变化的幅度逐渐减小,但未趋于稳定(斜率绝对值始终大于 0.5),在保证超阈值样本能够在估计出 GPD 模型参数的前提下(形状参数不出现较大波动或错误),选取尽可能大的阈值;(ii)若样本数量较多,即随着阈值增大,极值期望随之减小,且逐渐趋于稳定(斜率绝对值小于 0.5),选取极值期望稳定后对应的阈值作为最优阈值,另外尽量保证超阈值样本数不小于 50.

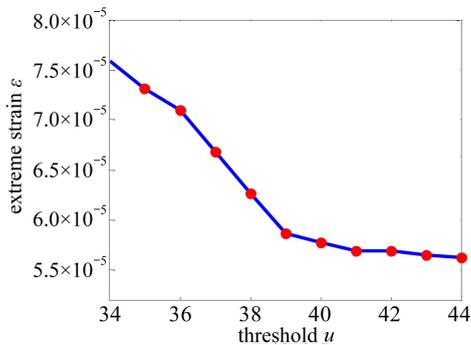


图 5 估计极值随阈值变化图

Fig. 5 The expectation of extreme strain vs. the threshold

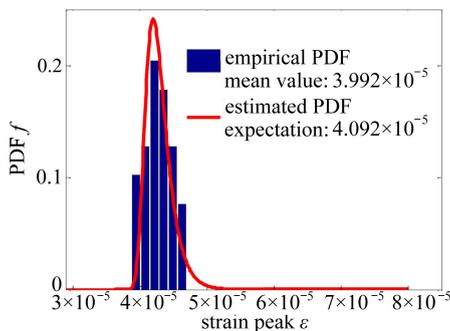


图 6 估计的周应变极值分布与实测值的频率直方图

Fig. 6 The estimated and measured weekly extreme strain distribution

2 工程实例：太平湖大桥

以安徽省黄山市太平湖大桥主梁实测 324 天的动应变数据为例,采用解析模态分解法剔除应变中的温度效应,选取每辆车通过桥梁时产生的应变峰值为研究样本,确定 GPD 模型阈值,具体步骤为:(a)选取不同阈值,采用过阈法分别估计应变峰值在 100 年内的极值期望;(b)建立阈值-应变极值期望关系曲线图,如图 5 所示;(c)从图中可知,当阈值大于 41 时,极值估计结果趋于平稳,属样本较多情形;(d)超越 41 的样本数量为 90(大于 50),满足要求,故选取 41 作为最优阈值.另外 4 种常用方法估计的阈值如表 6 所示,可见估计结果相差较大且较本文方法均偏小.为了验证本文方法选取的阈值是否合理,将实测的 324 天数据分为两部

分,前 30 天的数据用于预测周极值的分布,剩余 294 天共 42 周的数据用于验证预测结果。基于 GPD 模型的极值估计结果及实测的周极值频率直方图如图 6 所示。从图中可以看出,估计结果较好,但整体上估计的极值相较于理论值仍偏大。结合表 6,将实测周应变极值均值(3.992×10^{-5})对比图 5 可知,依本文方法所得阈值估计的周应变极值相较于 4 种常用方法更接近实测结果,估计结果更好。

表 6 实测样本的阈值估计

Table 6 Estimated thresholds of measured samples

method	mean excess function	Hill estimator	kurtosis method	minimum MSE method
estimated threshold	39.06	39.05	32.95	39.87

3 结 论

合理的阈值选取是车辆荷载作用下桥梁应变极值估计的重要基础。本文基于安徽省黄山市太平湖大桥车辆荷载作用下 1 年的应变数据,选取拟合效果较好的 3 种混合分布并采用 Monte-Carlo 方法抽样,通过比较同一样本基于不同阈值的 GPD 模型的极值估计结果与理论值,分析了 4 种常用阈值估计方法的优点及局限性,给出了一种基于经验的阈值选取方法并加以验证,得出以下结论:

(I) 1 个 Weibull 分布与两个正态分布加权的混合模型能较好地拟合车辆荷载作用下桥梁应变峰值的高尾分布。

(II) 图形分析法简单、快捷,但服从同一底分布的不同子样本的超阈值均值图与 Hill 估计图极为不稳定,部分图形无法确定阈值;采用峰度法确定的部分样本的阈值(底分布 1)、2)) 较最优阈值相差较大;均方误差最小法估计阈值计算量太大,估计结果偏小。

(III) 基于 GPD 模型估计的应变极值与阈值的选取密切相关。当样本容量较小时,在保证超阈值样本能够在估计出 GPD 模型参数的前提下,阈值越大,估计结果相对更好;若样本容量较大,阈值较小时,估计极值的偏差和方差都较大,偏差随着阈值增大而减小,且减小速率越来越慢,而方差最初随着阈值增大而减小,如若阈值持续增大,方差转而迅速增大,阈值的选取需均衡考虑偏差和方差。

(IV) 基于 GPD 模型估计的应变极值较理论值整体偏大。

(V) 太平湖大桥极值的估计结果显示,相较于常用的 4 种阈值估计方法,对于车辆荷载引起的应变峰值,据本文方法所得阈值估计的应变极值与实测值更接近,估计结果更好。方法可进一步推广应用至工程实践。

致谢 作者衷心感谢武汉理工大学道路桥梁与结构工程湖北省重点实验室 2015 年开放基金项目(大跨度斜拉桥疲劳寿命及可靠性研究:DQZDJJ201501)对本文的资助。

参考文献(References):

- [1] 公路桥涵设计通用规范: JTG/D 60—2015[S]. (General code for design of highway bridges and culverts: JTG/D 60—2015[S]. (in Chinese))
- [2] 孙守旺, 孙利民. 基于实测的公路桥梁车辆荷载统计模型[J]. 同济大学学报(自然科学版), 2012, 40(2): 198-204. (SUN Shou-wang, SUN Li-min. Statistic of vehicle loads for highway bridges[J]. *Journal of Tongji University (Natural Science)*, 2012, 40(2): 198-204. (in Chinese))

- [3] Mei G, Qin Q, Lin D J. Bimodal renewal processed models of highway vehicle loads[J]. *Reliability Engineering & System Safety*, 2004, **83**(3): 333-339.
- [4] 王涛. 高速公路桥梁交通荷载调查分析及仿真模拟[D]. 博士学位论文. 西安: 长安大学, 2010. (WANG Tao. Investigation statistics and simulation of random traffic loading of expressway bridge[D]. PhD Thesis. Xi'an: Chang'an University, 2010. (in Chinese))
- [5] 史道济. 实用极值统计方法[M]. 天津: 天津科学技术出版社, 2006: 28-32. (SHI Dao-ji. *Practical Extremum Statistical Method*[M]. Tianjin: Tianjin Science and Technology Press, 2006: 28-32. (in Chinese))
- [6] 李植淮, 李春前, 孙健康, 等. 基于 GPD 模型的车辆荷载效应极值估计[J]. 工程力学, 2012, **29**(S1): 166-171. (LI Zhi-huai, LI Chun-qian, SUN Jian-kang, et al. Estimation of extreme vehicle load effect based on GPD model[J]. *Engineering Mechanics*, 2012, **29**(S1): 166-171. (in Chinese))
- [7] 史道济, 张春英. 尾部指标估计中的阈值选择[J]. 天津理工大学学报, 2006, **22**(6): 78-82. (SHI Dao-ji, ZHANG Chun-ying. Threshold selection in tail index estimation[J]. *Journal of Tianjin University of Technology*, 2006, **22**(6): 78-82. (in Chinese))
- [8] Thompson P, Cai Y Z, Reeve D, et al. Automated threshold selection methods for extreme wave analysis[J]. *Coastal Engineering*, 2009, **56**(10): 1013-1021.
- [9] 李强. 基于 Copula 理论和 GPD 模型的金融市场风险测试研究[D]. 博士学位论文. 重庆: 重庆大学, 2012. (LI Qiang. The study of financial market risk measurement based on Copula theory and GPD model[D]. PhD Thesis. Chongqing: Chongqing University, 2012. (in Chinese))
- [10] 赵旭. 广义 Pareto 分布的统计推断[D]. 博士学位论文. 北京: 北京理工大学, 2012. (ZHAO Xu. Statistical inference of the generalized Pareto distribution[D]. PhD Thesis. Beijing: Beijing University of Technology, 2012. (in Chinese))
- [11] 段忠东, 欧进萍, 周道成. 极值风速的最优概率模型[J]. 土木工程学报, 2002, **35**(5): 11-16. (DUAN Zhong-dong, OU Jin-ping, ZHOU Dao-cheng. The optimal probabilistic distribution for extreme wind speed[J]. *China Civil Engineering Journal*, 2002, **35**(5): 11-16. (in Chinese))
- [12] Mcneil A J, Frey R. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach[J]. *Journal of Empirical Finance*, 2000, **7**(3/4): 271-300.
- [13] 花拥军, 张宗益. 基于峰度法的 POT 模型对沪深股市极端风险的度量[J]. 系统工程理论与实践, 2010, **30**(5): 786-796. (HUA Yong-jun, ZHANG Zong-yi. POT model based on kurtosis and its empirical study on extreme risk of Chinese stock markets[J]. *Systems Engineering—Theory & Practice*, 2010, **30**(5): 786-796. (in Chinese))
- [14] Caers J, Berilant J, Maes M A. Statistics for modeling heavy tailed distribution in geology—part I: methodology[J]. *Mathematical Geology*, 1999, **4**(31): 391-410.
- [15] Caers J, Maes M A. Identifying tails, bounds and end-points of random variables[J]. *Structure Safety*, 1998, **20**(1): 1-23.
- [16] 段忠东, 周道成. 极值概率分布参数估计方法的比较研究[J]. 哈尔滨工业大学学报, 2004, **36**(12): 1605-1609. (DUAN Zhong-dong, ZHOU Dao-cheng. A comparative study on parameter estimation method for extremal value distribution[J]. *Journal of Harbin Institution of Technology*, 2004, **36**(12): 1605-1609. (in Chinese))

Threshold Selection for the Extreme Value Estimation of Bridge Strain Under Vehicle Load

YANG Xia¹, ZHANG Jing¹, REN Wei-xin¹, YUAN Ping-ping²

(1. School of Civil Engineering, Hefei University of Technology, Hefei 230009, P.R.China;

2. School of Civil Engineering, Central South University, Changsha 410075, P.R.China)

(Contributed by REN Wei-xin, M. AMM Editorial Board)

Abstract: The selection of a reasonable threshold is critical to estimate the extreme strain under vehicle load on bridges with the peak-over-threshold method. Little information can be used if the threshold is too high, while the bias of parameters of the general Pareto distribution will be large if the threshold is too low. Common threshold selection methods are not suitable to be applied in estimation of the extreme strain under vehicle load. Based on 1-year strain data of the Taiping Lake Bridge, 3 types of mixed distributions for the strain peaks induced by vehicle load were chosen to generate a large number of samples with the Monte-Carlo method. The estimated extreme values of the samples based on the generalized Pareto distributions with different thresholds were compared and analyzed. Then, an empirical threshold selection method was proposed for the strain data induced by vehicle load. Finally, the Taiping Lake Bridge was chosen as the case verification. It is demonstrated that the estimated weekly extreme strain based on the threshold selected with the proposed method is more close to the measured results than those with the common methods.

Key words: bridge health monitoring; threshold selection method; extreme strain; generalized Pareto distribution

Foundation item: China Postdoctoral Science Foundation(2015M581982)

引用本文/Cite this paper:

阳霞, 张静, 任伟新, 袁平平. 车辆荷载作用下桥梁应变极值估计的阈值选取[J]. 应用数学和力学, 2017, 38(5): 503-512.

YANG Xia, ZHANG Jing, REN Wei-xin, YUAN Ping-ping. Threshold selection for the extreme value estimation of bridge strain under vehicle load[J]. *Applied Mathematics and Mechanics*, 2017, 38(5): 503-512.