



HAL
open science

A mixture peaks over threshold approach for predicting extreme bridge traffic load effects

Xiao Yi Zhou, Franziska Schmidt, François Toutlemonde, Bernard Jacob

► **To cite this version:**

Xiao Yi Zhou, Franziska Schmidt, François Toutlemonde, Bernard Jacob. A mixture peaks over threshold approach for predicting extreme bridge traffic load effects. *Probabilistic Engineering Mechanics*, 2016, 43, pp.121-131. 10.1016/j.probengmech.2015.12.004 . hal-01498789v3

HAL Id: hal-01498789

<https://hal.science/hal-01498789v3>

Submitted on 22 Jun 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Mixture Peaks over Threshold Approach for Predicting Extreme Bridge Traffic Load Effects

Xiao-Yi Zhou ¹, Franziska Schmidt², François Toutlemonde ³ and Bernard Jacob ⁴

1 **ABSTRACT**

2 Traditionally, bridge traffic load effects are considered as identically and indepen-
3 dently distributed random variables. However, load effects resulting from different
4 loading events in terms of simultaneously involved vehicles/trucks do not have the
5 same statistical distributions. To consider this, a novel method has been developed for
6 predicting characteristic value and maximum value distribution of traffic load effects
7 on bridges. The proposed method is based on the conventional peaks-over-threshold
8 method, which uses the generalized Pareto distribution. The principle is to (1) separate
9 the traffic load effects by types of loading event, (2) model the upper tail of the load ef-
10 fect for each type with generalized Pareto distribution, and (3) integrate them together
11 according to their respective weights in the total population. Numerical studies have
12 been conducted to demonstrate the feasibility of the proposed method in predicting
13 characteristic value or quantile and extreme value distribution for bridge traffic load

¹Formerly, Ph.D. student, Materials and Structures Department, IFSTTAR (French Institute of Science and Technology for Transport, Development and Networks); Université Paris-Est, France; Currently, Research associate, Ph.D., School of civil engineering and geosciences, Newcastle University, United Kingdom.

²Research engineer, Ph.D., Materials and Structures Department, IFSTTAR; Université Paris-Est, France.

³Chief Scientist, Materials and Structures Department, IFSTTAR; Université Paris-Est, France.

⁴Senior engineer and Scientific delegate, IFSTTAR; Université Paris-Est, France.

14 effects. Results show that the proposed approach is efficient to conduct extreme value
15 analysis for data having mixture probability distribution function.

16 **Keywords:** Traffic load effects; Peaks-over-threshold; Mixture peaks-over-threshold;
17 Bridge; Extreme value; Generalized Pareto distribution

18 **INTRODUCTION**

19 Assessing the condition of existing bridges is of increasing concern in bridge
20 management as more and more bridges step into their ageing stage worldwide,
21 and a deteriorated bridge raises a risk to safety and welfare loss for the users.
22 Although extensive efforts have been devoted to elaborate load-carrying capacity
23 models, the role of traffic loading in existing bridge structures has increasingly
24 received attention in recent years as potential benefits have been revealed in
25 terms of optimally allocating the limited maintenance and management budgets
26 (COST 345 2002; Frangopol et al. 2008; Fu and You 2009; Li et al. 2012). In
27 addition, the growth of traffic has been reported in recent years worldwide: for
28 instance in Europe the road freight transport has increased by 35% between 1995
29 and 2010. This has led the regulators introducing truck weight limit regulations
30 and allowing the introduction of higher and longer vehicles in some member
31 states, such as Scandinavia. These changes may have aggressive impacts on
32 bridge structures in terms of maximum load and load effect, fatigue damage,
33 probability of failure and etc. (Desrosiers and Grillo 1973; Ghosn and Moses
34 2000; Righiniotis 2006; Gindy and Nassif 2007; Tong et al. 2008; Fu et al. 2011;
35 Zhou et al. 2014; O'Brien et al. 2014). The topic of multi-hazard analysis
36 combines traffic loading with seismic or wind loading (Cai and Chen 2004; Zhu
37 and Frangopol 2012; Ghosh et al. 2013). Therefore, an accurate prediction of the
38 extreme traffic load effects on bridges is desired, especially for evaluating existing
39 bridge structures.

40 Indeed the estimation of a high quantile or tail distribution is not an easy task,

41 making inference about the extremal behaviour, in a domain where the samples
42 only contain a very small amount of data. Moreover, extrapolation beyond the
43 range of the data is necessary to know something about areas where there are
44 no observation at all (Leadbetter et al. 1983; Coles 2001; de Haan and Ferreira
45 2006). This issue belongs to extreme value statistics, which has been extensively
46 developed in the last 60 years, although it can be tracked back to the early 20th
47 century. Extreme value predictive techniques have been used in many disciplines
48 including structural engineering, and extensive research has been conducted in
49 recent decades on bridge traffic load effects. The methods in the literature on
50 extreme traffic load effects on bridges can be broadly classified into two major
51 categories: (1) tail distribution methods, and (2) periodic maxima methods.

52 The primary objective of the first category of methods is to find the underlying
53 distribution of bridge traffic load effects, then the maximum value distribution
54 can be easily computed by raising the distribution to a certain power (Coles
55 2001). Using Normal distribution (Nowak 1993; Sivakumar et al. 2011), Gumbel
56 distribution (Cooper 1997; Fu and You 2009), Weibull distribution (O'Brien et al.
57 1995) to bridge traffic load effects belongs to this category of method. In addition,
58 the Rice formula based level-crossing method adopted in (Cremona 2001) can also
59 be classified into the first category as the mathematical assumption implies that
60 the traffic load effect is normally distributed.

61 The second category of methods aim at fitting a series of local maxima, taken
62 from successive independent samples of observations over a given time period, to
63 a standard extreme value distribution. Then the extreme characteristic values
64 (or values with a given return period) for expected probabilities of exceedance
65 can be computed. Fitting daily or yearly maxima to Weibull distribution (Bailey
66 and Bez 1999), or to Gumbel distribution (Fu and You 2009) and to generalized
67 extreme value distribution (Messervey et al. 2010; Park and Sohn 2006; Enright

68 et al. 2013) belongs to this category. A comprehensive review and quantitative
69 comparison of the prediction methods of extreme traffic load effects on bridges
70 can be found in (O'Brien et al. 2015).

71 It has been widely accepted in the extreme value statistics research commu-
72 nity that the generalized Pareto distribution (GPD) based peaks-over-threshold
73 approach (POT) is as effective as generalized extreme value distribution (GEV)
74 based block-maxima method (BM) to estimate extreme value. However, the use
75 of POT approach has seldom been reported in bridge traffic load effects, although
76 the POT approach has significant advantages. Many papers in other disciplines
77 have proved that it may provide more accurate estimates than the BM method
78 in modelling extreme values (Madsen et al. 1997; O'Brien et al. 2015). Moreover
79 its mathematical form leads to very simple formulation.

80 Most of the previous works assume that bridge traffic load effects are iden-
81 tically and independently distributed (iid), which is a main condition to apply
82 the extreme value theory (Coles 2001). However, it has been shown that bridge
83 traffic load effects are induced by different types of loading events, depending on
84 the number of trucks being simultaneously on the bridge deck. Thus the periodic
85 maximum (usually daily maximum) used in the estimation may not come from
86 the same type of distribution, which does not comply with the iid assumption
87 (Harman and Davenport 1979). Desrosiers and Grillo (1973) stated that the mul-
88 tiple presence of trucks depends significantly on the bridge length, truck speed
89 and traffic volume based on field data collected from several highway locations
90 (Connecticut Route 5, I-91 at the Depot Hill Road, and I-91 at the Connecticut
91 Route 68). These findings have been confirmed in (Gindy and Nassif 2007) with
92 recent traffic data collected from 25 WIM sites in New Jersey between 1993 and
93 2003. Moreover, Messervey et al. (2010) states that the periodic maxima usually
94 do not come from a single distribution as the number of events varies day by day.

95 It is possible to select an optimal periodic length (Messervey et al. 2010), but it
96 may waste data because of the reduced number of extremes used from these data.
97 Another solution by (Caprani et al. 2008) named composite statistic distribution
98 method accounts for the variation of loading distribution based on block maxima
99 method and models extreme load effects from the same type of loading event.

100 In order to address the non-identically distributed traffic load effects, a novel
101 extreme value analysis method has been proposed. The proposed method is
102 based on the conventional peaks-over-threshold method (CPOT), which relies on
103 the generalized Pareto distribution. The principle is to classify the traffic load
104 effects by types of loading event. Then the CPOT is used to derive the upper
105 tail of load effect distribution for each loading event category with generalized
106 Pareto distribution. Finally the upper tail distribution is the weighed average of
107 the upper tail distributions by loading event.

108 In the following sections, the mathematical background and the details of
109 derivation of the novel method are presented. Numerical studies, including a
110 theoretical example and a real traffic load effect example, are conducted to illus-
111 trate the capacity of the proposed method, and its performance is assessed by
112 comparing with the conventional methods and the recently developed composite
113 statistic distribution method (Caprani et al. 2008).

114 **METHODOLOGY**

115 **The generalized Pareto distribution and Peaks-over-Thresholds ap-** 116 **proach**

117 Let X_1, \dots, X_n be a sequence of independently and identically distributed
118 random variables with distribution function F . When the value taken by X_i
119 exceeds some high threshold u , this value can be treated as an extreme event.
120 The behavior of those extremes can be described by the conditional distribution

121 function of the excesses, $x = X - u$, over the threshold u :

$$F_u(x) = Pr \{X - u \leq x | X > u\} = \frac{F(x + u) - F(u)}{1 - F(u)}, \quad (1)$$

122 for $0 \leq x < x_0 - u$.

123 The Balkema-de Haan-Pickands theorem (Balkema and de Haan 1974; Pickands
124 III 1975) states that, for a certain class of distributions, the generalized Pareto
125 distribution (GPD) is the limiting distribution for the distribution of the excesses,
126 as the threshold tends to the right endpoint. The distribution function of GPD
127 is usually expressed as:

$$H(x; \xi, \sigma) = \begin{cases} 1 - [1 + \xi(\frac{x-u}{\sigma})]^{-1/\xi} & \xi \neq 0, \\ 1 - \exp(-\frac{x-u}{\sigma}) & \xi = 0, \end{cases} \quad (2)$$

128 where u is the threshold value, $\sigma > 0$, and the support is $x \geq 0$ when $\xi \geq 0$ and
129 $0 \leq x - \sigma/\xi$. The GPD comprises three known distribution types, depending
130 on the value of parameter ξ . When $\xi > 0$, the function is equivalent to a re-
131 parametrized version of the usual Pareto distribution; if $\xi < 0$, the distribution
132 is called a type II Pareto distribution; $\xi = 0$ gives the exponential distribution.

133 According to Eqs. (1) and (2), the distribution function $F(x)$ can thus be
134 expressed as:

$$F(x) = (1 - \varsigma_u) + \varsigma_u H(x; \xi, \sigma, u), \quad (3)$$

135 where $\varsigma_u = Pr \{X > u | X \geq 0\} = 1 - F(u)$ represents the survival function, while
136 $F_u(x)$ is the cumulative distribution function (CDF) of $x > u$ only.

137 The quantile x_m that is exceeded on average once every m observations is the

138 solution of:

$$x_m = \begin{cases} u + \frac{\sigma}{\xi} \left[(m\zeta_u)^\xi - 1 \right] & \xi \neq 0 \\ u + \sigma \log(m\zeta) & \xi = 0 \end{cases} \quad (4)$$

139 provided m is sufficiently large to ensure that $x_m > u$.

140 **Derivation of the mixture Peaks-over-Thresholds method**

141 Now, let X_1, \dots, X_n be a sequence of independently but non-identically dis-
 142 tributed random variables with distribution function F , which is a mixture dis-
 143 tribution consisting of m components, expressed as:

$$F(x) = \sum_{j=1}^m F_j(x) \cdot \varphi_j, \quad (5)$$

144 where the j -th component (distribution function of the j -th sub-population) F_j
 145 belongs to the domain of maximum attraction, and φ_j is the weight of X belonging
 146 to the j -th sub-population, with $\sum_{j=1}^m \varphi_j = 1$. Straightforwardly, the survivor
 147 function is expressed as:

$$\bar{F}(x) = 1 - F(x) = \sum_{j=1}^m [1 - F_j(x)] \varphi_j. \quad (6)$$

148 Assume that for a given threshold u_j the exceedances of j -th component
 149 could be reliably described by a generalized Pareto distribution, from Eq.(3) the
 150 survivor function of the j -th component can be formulated:

$$1 - F_j(x) \equiv [1 - H_j(x - u_j)][1 - F_j(u_j)] \quad (7)$$

151 Substituting Eq.(7) into Eq.(6), the survivor function of the mixture distribution

152 can be expressed as:

$$\bar{F}(x) = \sum_{j=1}^{n_t} [1 - H_j(x - u_j)][1 - F_j(u_j)]\varphi_j. \quad (8)$$

153 Therefore, the tail of the mixture distribution can be represented by:

$$F(x) = 1 - \bar{F}(x) = 1 - \sum_{j=1}^{n_t} [1 - H_j(x - u_j)][1 - F_j(u_j)]\varphi_j. \quad (9)$$

154 As shown in Eq.(9), the quantile for this mixture distribution can not be
155 obtained directly. Hence, iteration is needed to find optimal estimate \hat{x}_m that
156 satisfies the following equation:

$$[1 - F(\hat{x}_m)] - \frac{1}{m} \leq \epsilon. \quad (10)$$

157 with ϵ as a given small value.

158 **Approach for threshold selection in the use of Mixture Peaks-Over-** 159 **Thresholds method**

160 In the application of the MPOT method, an essential step is to select an ap-
161 propriate threshold u_j for each component of the mixture models of load effects
162 to which the asymptotic GPD is approximated. The threshold selection requires
163 consideration of the trade-off between bias and variance: a too high threshold
164 reduces the number of exceedances and thus increases the estimated variance,
165 whereas a low threshold can reduce the estimated variance but increase the bias
166 (Scarrott and MacDonald 2012). Graphical diagnosis approaches, e.g. the mean
167 residual life plot, are commonly used for such a selection, but they require the
168 practitioner to have substantial expertise and can be rather subjective. More-
169 over, they may be time-consuming if there are many thresholds to be selected.

170 Hence, graphical diagnosis approaches are not fully suitable for our problem. Au-
 171 tomatic threshold selection approach with appropriate measure is preferable to
 172 avoid subjective judgement and to apply the proposed approach efficiently as sev-
 173 eral thresholds are needed to be selected in the MPOT method. Several types of
 174 automatic threshold selection rules exist. The simplest ones are the fix number
 175 rules such as the upper 10% rule, the square root rule $k = \sqrt{n}$ or its modification
 176 $k = n^{2/3} / \log [\log(n)]$, but they are usually lacking of theoretical background.

177 Therefore, we adopt the automatic method based on goodness-of-fit test statis-
 178 tics. The Anderson-Darling (AD) and Cramer - von Mises (CM) test proposed
 179 by Choulakian and Stephens (2001) to examine the goodness-of-fit for GPD have
 180 been adopted:

$$\begin{aligned}
 W_n^2 &= \frac{1}{12n} + \sum_{i=1}^n \left(z_i - \frac{1 - 1/2}{n} \right)^2 \quad \text{for CM test,} \\
 A_n^2 &= -n - \frac{1}{n} \sum_{i=1}^n (2i - 1) \{ \ln z_i + \ln (1 - z_{n+1-i}) \} \quad \text{for AD test.} \quad (11)
 \end{aligned}$$

181 It is worth mentioning that the collection of optimal thresholds for individual com-
 182 ponents may not be the optimal threshold combination for the mixture model.
 183 An additional procedure is needed to find an optimal combination of the indi-
 184 vidual thresholds. Again, a goodness-of-fit test is used to make the decision.
 185 However, only a non-parametric test is reasonable to be used due to the feature
 186 of mixture model: the Kolmogorov-Smirnov (KS) test has been chosen in this
 187 work. In statistics, the KS test is a non-parametric test and qualifies a distance
 188 between the empirical distribution function of the sample and the cumulative
 189 distribution function of the reference distribution. In addition, the generalized
 190 Pareto distribution has an important property that will be used to find the opti-
 191 mal combination of thresholds. If excesses of a sample over the optimal threshold,

192 u_0 , can be reasonably modelled by a GPD with shape parameter ξ and σ_0 , then
 193 the excesses over thresholds larger than the optimum will follow GPDs with same
 194 shape parameter ξ but different scale parameter, σ_u that linearly depends on the
 195 threshold value $\sigma_u = \sigma_0 + \xi(u - u_0)$. Therefore, the solution is to find a set of
 196 u_1, \dots, u_m that satisfy:

$$D_n = \sup \{F_n(x) - F(x; u_1, \dots, u_m)\}. \quad (12)$$

197 **The estimation of GPD parameters**

198 Estimating the distribution parameters of GPD is another decisive point that
 199 influences the performance of the MPOT method. Various estimators have been
 200 proposed to estimate the parameters of GPD. The applicability of a certain
 201 method depends on the features of the considered data. A comprehensive re-
 202 view and qualitative comparison of different parameter estimation methods has
 203 been provided by (de Zea Bermudez and Kotz 2010), and a quantitative study has
 204 been conducted in (Zhou 2013) to evaluate the performance of various parameter
 205 estimation methods when applying peaks-over-threshold method on traffic load
 206 effect data. The method of moment (MM), the power weighted moment method
 207 (PWM) and the maximum likelihood method (ML) are commonly used in the
 208 literature. It has been widely accepted that the maximum distribution of bridge
 209 traffic load effects belongs to an upper bounded Weibull distribution which has
 210 a shape parameter $\xi < 0$. Hence, the MM, PWM and ML methods are suit-
 211 able to traffic load effects. In addition, the minimum density power divergence
 212 (MDPD) method is used in this work due to its excellent performance in the case
 213 of contaminated data (Juarez and Schucany 2004).

214 **TRAFFIC DATA AND BRIDGE TRAFFIC LOAD EFFECTS**

215 **Description of Weigh-in-Motion traffic data**

216 Traffic data from the A9 motorway near Saint-Jean-de-Védas (SJDV), in
217 southeastern France, was used in this study. Weights and dimensions of trucks,
218 which travelled in the slow and fast lanes in one direction of the 6-lane motorway,
219 were recorded by using a piezo-ceramic Weigh-in-Motion (WIM) system from
220 January 2010 to May 2010. A total number of 581,011 trucks representing traffic
221 of 86 days were drawn from the original data by excluding unreasonable record-
222 ings, weekends and system inactivity days. The traffic composition displayed in
223 Fig.1a shows that the 5-axle truck is the dominant type of truck on this site rep-
224 resenting 76.4% in traffic volume. The histogram of gross vehicle weight (GVW)
225 is presented in Fig.1b. To see the contribution from each type of truck, a stacked
226 plot is given. It can be seen that the 5-axle truck governs the leading mode of
227 the GVW histogram.

228 **Traffic loading Monte Carlo simulation**

229 If traffic data can be recorded by WIM for a sufficiently long period of time,
230 such as a year, then the load effects induced by the measured traffic can be
231 directly used to estimate the extreme load effect. Long term data, however, are
232 not always available, due to the limitation of storage for huge amount of data for
233 continuous recording, the problem of the equipment, the limitation of budget for
234 conducting long term measuring, etc. Using limited data to predict extreme value
235 distribution is thus a common situation in practice. The estimate of characteristic
236 value may have large variance if extrapolation is based on limited data. Hybrid
237 method that integrates extreme value analysis approaches with traffic simulation
238 techniques is a practical solution. Using microscopic traffic simulation techniques
239 to generate long-term traffic loads or load effects has been demonstrated as an
240 efficient and accurate approach to study bridge traffic load effect in recent years

241 (O'Connor and O'Brien 2005; Chen and Wu 2011; Enright and O'Brien 2012).

242 In the present study, a simulation program is developed to generate virtual
243 traffic and to calculate traffic load effects on bridges. The basic principle is to
244 generate traffic flow with the same features as those extracted from measured
245 traffic data, such as aforementioned 86 days' WIM data. This is realized in
246 following steps:

- 247 1. Calculating traffic composition: In this study, vehicles are categorized into
248 classes according to their silhouettes as illustrated in Fig.2.
- 249 2. Establishing statistical models for characteristics of each class of vehicle,
250 including gross vehicle weight (GVW), distribution of GVW to individ-
251 ual axle or axle group, vehicle speed, vehicle configuration in terms of
252 axle spacing and vehicle length, and lateral position of the vehicle in the
253 lane. The best fit is selected among the normal, bi- and tri-modal normal
254 distribution.
- 255 3. Establishing vehicle moving model: Time headway distribution model,
256 which describes the time distance between the rear axle of the front truck
257 and the front axle of the following truck, is fundamental to traffic flow
258 modelling in traffic simulation. A refined hourly truck flow rate depended
259 headway model proposed by O'Brien and Caprani (2005) is adopted in
260 the present study. Headways of less than 4 seconds are modelled using
261 quadratic curves for different flow rates, and a negative exponential distri-
262 bution is used for larger headways.
- 263 4. Simulating traffic flow: Assume the simulation program started at time t ,
264 a group of n_t vehicles is generated by using headway model in step (3);
265 each vehicle of these n_t vehicles is randomly assigned a vehicle class with
266 the traffic composition information that is a uniform distributed random

267 variable ranging from 0 to 1, and the vehicle characteristics are generated
268 according to the assigned class.

269 5. Calculating load effects: Once the traffic data is generated, it is passed to
270 the load effect calculation subroutine. The calculation is activated when
271 a vehicle arrives on the bridge, then this vehicle is assumed as leading
272 vehicle and passes the bridge in a time step Δt .

273 At each step, the program searches and counts the number of vehicles, N , on
274 the bridge. The load effect, $LE(t_n)$, at time, t_n , induced by these N vehicles can
275 be obtained by using:

$$LE(t_n) = \sum_{j=1}^N \sum_{k=1}^{n_j} \phi S_i(x_j^k, y_j^k) P_j^k, \quad (13)$$

276 where:

277 N : number of vehicles on the bridge,

278 n_j : number of axles of the j th vehicle,

279 ϕ : dynamic amplification factor,

280 S_i : influence surface for load effect of interest produced by a unit load of i -th
281 type of tyre,

282 x_j^k : longitudinal position of the k -th axle of the j -th vehicle

$$x_j^k = v_j \cdot (t_n - t_j^0) - d_j^k,$$

283 v_j : speed of j th vehicle,

284 t_j^0 : arrival time of the first axle of the j th vehicle, when passing over the position

285 $x = 0$,

286 d_j^k : distance between steering axle and the k th axle of the j th vehicle,

287 y_j^k : transversal position of the k -th axle of the j -th vehicle,

288 P_j^k : load of the k -th axle of the j th vehicle.

289 **Classifying load effects by loading event**

290 Recording traffic load effects and loading events simultaneously, Fig.3 shows
291 that several single truck loading events have induced a larger load effect than those
292 induced by 2-truck loading events. In order to use all possible relatively large load
293 effects efficiently, the full time history of effects induced by traffic passing over
294 the bridge is retained first, then the local extremes and corresponding types of
295 loading events (comprising the number of trucks) are identified. Fig. 4 illustrates
296 such a process, the time history of the traffic load effect is drawn in blue line and
297 the local extremes are marked with red stars:

- 298 1. The process starts with a single loading event when the first truck arrives
299 the bridge.
- 300 2. Then another truck (2nd truck) arrives on the bridge generating a 2-truck
301 loading event.
- 302 3. The first arrived truck leaves the bridge and the loading becomes a single
303 truck event again.
- 304 4. Then a new truck (3rd truck) enters the bridge and the loading becomes
305 a 2-truck event again,
- 306 5. The 2nd arrived truck exits the bridge (single loading event),
- 307 6. Then a new truck (4th truck) arrives so that a new 2-truck loading event
308 is generated,
- 309 7. Finally the 3rd truck exits the bridge and the loading event is a single
310 truck loading event again.

311 In this process, a total of four trucks has arrived on the bridge and produced
312 4 extreme single truck loading events and 3 extreme two-truck loading events.
313 The local extreme for each loading event is identified and marked in Fig.4. Using
314 this procedure, local extremes for various types of loading event are identified.
315 Fig.5 shows histograms of traffic load effects induced by simulated traffic for
316 illustration purpose, and it can be seen that local extremes induced by different
317 types of loading events are not identically distributed. The classical extreme value
318 theory can thus not be directly applied to these mixed data as it requires data of
319 independent and identical distribution.

320 Previous studies have demonstrated that three types of load effects are critical
321 for short to median length bridges: (I1) bending moment at mid-span and (I2)
322 shear force at end-support of a simply supported bridge, and (I3) hogging moment
323 at middle support of a two-span continuous bridge. In this study, these three
324 types of load effects are studied with span lengths of 20m, 30m, 40m and 50
325 m. Considering the time consumption, 1500-day's traffic data were generated by
326 the developed traffic simulation program using statistical inputs extracted from
327 SJDV traffic data. For the three types of load effects, six categories of loading
328 events have been identified from the simulation. These six categories of truck
329 arrangements are 1-truck, 2-truck, 3-truck, 4-truck, 5-truck, and 6-truck loading
330 events. It should be noted that the 1-truck case includes situations from only
331 one axle of the truck to the whole truck being on the bridge. Similarly, 2-truck
332 loading events include all possible combinations of two trucks, from both trucks
333 having only one axle on the bridge to both trucks having all axles on the bridge
334 simultaneously. This is also the case for all loading types.

335 Two sets of loading event composition are listed in Table 1 for the three types
336 of load effects, with four types of bridge lengths. The first group is for load
337 effects over 90th percentile, and the second group is for load effects above 95th

338 percentile. Fig. 6 shows that the governing type of loading event changes with
339 increased bridge length. For a bridge length of 20 m, 2-truck and 3-truck loading
340 events govern the upper tail. For a bridge length of 30 m, it can be seen from
341 Fig. 6 that the governing event is 3-truck loading event. For bridge lengths of 40
342 and 50 m, 3-truck events are still the governing but some 4- and 5-truck events
343 occur at the upper end of the simulation period. In addition, the composition of
344 loading events are different between the data over 90th percentile and those over
345 95th percentile. In general, it demonstrates the importance to classify the load
346 effects by loading events in predicting extreme value distribution or characteristic
347 value.

348 **EVALUATING THE PERFORMANCE OF THE MPOT APPROACH**

349 To show how the MPOT method works for realistic bridge traffic load effects,
350 two numerical studies have been conducted and are reported in this section. The
351 first example is to examine the performance of the MPOT method for a set of data
352 generated from a mixture normal distribution, and the second example is to eval-
353 uate the MPOT method for bridge traffic load effects generated by Monte Carlo
354 traffic microsimulation. In both examples, a comparison of the relative accuracy
355 of the present MPOT and of the conventional peaks-over-threshold (CPOT) is
356 performed.

357 **Theoretical example**

358 The normal distribution is widely used in bridge engineering: for example
359 gross vehicle weights are usually modelled by normal distribution or mixture
360 normal distribution. In the first example, the performance of MPOT method is
361 evaluated by using a random event having a parent distribution of mixture normal
362 distribution with two components, $F(X < x) = \varphi_1 \Phi(\frac{x-\mu_1}{\sigma_1}) + \varphi_2 \Phi(\frac{x-\mu_2}{\sigma_2})$. The
363 core distribution is $N(420, 30)$ with the relative frequency of occurrence $\varphi_1 = 0.9$,

364 and the "contaminating" distribution is $N(380, 45)$ with the relative frequency
365 of $\varphi_2 = 0.1$. Assuming a thousand events of this type occurring every day,
366 three thousand days' events are simulated with a total of $n = (3000 \times 1000) =$
367 $3,000,000$ -elements sample. In the simulation process, values from the $N(420, 30)$
368 are denoted as event one, while those from the $N(380, 45)$ are denoted as event
369 two. These 3,000,000 sample are thus classified into two groups.

370 To approximate the upper tail of the distribution of the simulated sample,
371 the two aforementioned CPOT and MPOT methods are applied. For the CPOT
372 method, an optimal GPD is needed to be found, while for the MPOT method two
373 optimal GPDs with one for each subgroup of events are required. The goodness-
374 of-fit based threshold selection approach is used first to select the optimal thresh-
375 old, then the GPD parameters for the exceedances are estimated by using the
376 four previously mentioned estimators. Following this procedure, the threshold
377 and GPD parameter estimates for the CPOT are obtained and tabulated in Ta-
378 ble 2, and the corresponding results for the MPOT method are listed in Table
379 3.

380 Using these estimates, the upper tail distribution can be obtained from Eq.(3)
381 for CPOT and Eq.(9) for MPOT. They are shown in a log-scale plot in Fig.7
382 along with the empirical distribution function of the sample. It can be seen that
383 both CPOT and MPOT methods capture the main part of the distribution very
384 well, but the discrepancy between empirical distribution and fitted distribution
385 becomes larger when getting close to the upper tail. The CDF obtained from
386 MPOT captures the upper tail with significantly less bias than that from the
387 CPOT. Indeed, the MPOT follows the trend of the data, while the CPOT strongly
388 deviates. By using the estimates of GPD, the quantile or characteristic values for
389 a certain return period can be calculated from Eq.(4) for CPOT or Eq.(10) for
390 MPOT. Fig.8 compares the characteristic values for a return period of 100-year

391 calculated with CPOT and MPOT methods with the real one (which is known
392 because the underlying distribution is known). It indicates that both approaches
393 have good performance on quantile estimation, with maximum error less than
394 2%. The return levels estimated with conventional method are even much closer
395 to the true value.

396 For reliability analysis, the maximum value distribution of load effects is
397 required. After obtaining the upper tail distribution, it is straightforward to
398 calculate the maximum value distribution function using $F^n(x)$. The CDFs of
399 maximum value distribution with CPOT and MPOT methods are displayed in a
400 Gumbel plot in Fig.9, where the true distribution is given as well. It can be seen
401 that the MPOT based maximum value distribution matches the true distribution
402 well, while the CPOT based maximum value distribution differs from the true
403 distribution, particularly at the upper tail.

404 Although the CPOT can provide a relatively accurate estimate of character-
405 istic value, especially for low return period, as the advanced MPOT method, it
406 can not predict the upper tail of the distribution in sufficient accuracy as sig-
407 nificant deviation is found in maximum value distribution when comparing with
408 the bench mark. It is of particular importance to estimate the maximum value
409 distribution for reliability-based structural assessment. It therefore illustrates the
410 importance to consider the inherent distribution for load effects.

411 **Simulated traffic load effect example**

412 The previous simple example showed that the MPOT method has better per-
413 formance than the CPOT method when the data are not identically distributed.
414 Now we will evaluate its performance for bridge traffic load effects, which are gen-
415 erated by the previously mentioned microscopic Monte Carlo traffic simulation
416 program. Table 1 has shown that the upper tail of distribution for bridge traffic

417 load effects consists of contributions from different loading events, and Fig.5 has
418 displayed that load effects from different loading events have various distribution
419 features in terms of distribution type or parameters. To show how these features
420 influence the distribution function estimation or high quantile prediction and to
421 demonstrate the advantage of the proposed method, a comparative study between
422 the CPOT method and the MPOT method is performed.

423 To exclude the influence of the threshold selection, we firstly conducted the
424 comparison with fixed thresholds at 90th, 92nd, 94th, 96th, and 98th percentiles.
425 Again, the estimators of MM, PWM, ML, and MDPD are used to estimate the
426 distribution parameters for the involved GPDs. We used the graphical method to
427 evaluate the performance of MPOT and CPOT methods. For instance, Fig. 10
428 shows the comparison between CPOT method and MPOT method for bending
429 moment at the mid-span of a simply-supported bridge with span of 40m I1 load
430 effect, and the distribution parameters are estimated by ML method. The graphs
431 on the left in Fig.10 illustrate the empirical survival function (black dots) fitted
432 function with CPOT estimates (red solid lines) and with mixture POT estimates
433 (green dash lines) for various thresholds, while the graphs on the right side show
434 corresponding these results in a logarithm scale plot. It can be seen that the
435 MPOT method approximates the excesses over threshold with good accuracy,
436 while the CPOT method approximates the majority of the data well but has
437 poor approximation for the high tail. It is commonly accepted that the high tail
438 is extremely important in the extreme value analysis such as quantile estimation.
439 A quantitative method has been adopted to compare the performance of the two
440 methods. The results of root-mean-square-error reported in Table 4 confirm that
441 the MPOT method improves the modelling as a majority of the values for MPOT
442 are smaller than those for CPOT. Therefore, the MPOT has better performance
443 than the CPOT method in capturing the upper tail of the distribution.

444 This preliminary study has demonstrated that the MPOT method has the
445 potential to provide more accurate prediction than the CPOT method. When
446 studying bridge traffic load effects, the prediction of characteristic values for long
447 return periods, such as the 1000-year characteristic value for traffic load model
448 in Eurocode (CEN 2003), is a critical issue. Here we will illustrate the difference
449 between CPOT and MPOT on this characteristic value prediction. Except for
450 these two GPD based methods, the comparison also includes the GEV distribu-
451 tion based BM method. In the preliminary study, fixed thresholds are used to
452 compare the performance of CPOT and MPOT methods under consistent con-
453 ditions. But it should be noted that a fixed threshold may not be optimal to
454 approximate the upper tail distribution. Thus, in the following study, threshold
455 for each GPD is selected by using the goodness-of-fit statistics based automatic
456 method for both CPOT and MPOT methods. With an illustration purpose, the
457 selected optimal threshold and corresponding distribution parameters for each
458 component of the mixture distribution are listed in Table 5 for the I1 load ef-
459 fect with bridge length of 40m. The tail distribution consists of load effects
460 from 2-truck, 3-truck and 4-truck loading events, thus three sets of threshold and
461 parameters have to be estimated. It can be seen from the results that each com-
462 ponent has different tail distribution. For instance, distribution for load effects
463 resulting from 2-truck loading events has a Pareto distribution with shape pa-
464 rameter $\xi > 0$, while those from 3-truck loading events and 4-truck loading events
465 have type II Pareto distribution with $\xi < 0$. Similar procedures are applied to
466 other load effect cases, then the optimal threshold and corresponding distribution
467 parameters are obtained. For the BM method, daily maxima are identified from
468 the simulated load effects, then GEV distribution is fitted to each set of daily
469 maxima.

470 For the load effects I1, I2 and I3 with span lengths of 20m, 30m, 40m and 50m,

471 the 100-year and 1000-year return period characteristic values are calculated by
472 the BM, CPOT and MPOT approaches. Results from the BM and CPOT meth-
473 ods are given in Table 6 for characteristic values for 100-year return period and
474 in Table 7 for characteristic values for 1000-year return period in terms of relative
475 difference with respect to the corresponding results from the MPOT method. The
476 differences between conventional and mixture estimates are smaller for 100-year
477 return level than for 1000-year return level. For example, the difference between
478 the convention method and the proposed method for 100-year return level of load
479 effect I1 with span of 30 m shown in Table 6 is around -6.31% for MM case,
480 while the difference for 1000-year return level in Table 7 is around 13.5% . It
481 confirms the common impression that the extrapolation to remote future is not
482 stable. As expected, the difference between conventional method and mixture
483 method is smaller for load effects for shorter spans, either the BM or the POT.
484 For instance, the difference is -8.49% for BM for 100-year return level of load
485 effect I1 at length of 20 m in Table 6, but it increases to about 17% at span length
486 of 50m. The composition of loading events becoming more complex when span
487 length increases, and more types of loading events thus become the governing
488 loading events. Among the three types of load effects, the performances of the
489 methods are different. The differences are larger for load effects of I3 than for the
490 other two. As stated in Harman and Davenport (1979), the load effect of I3 is
491 more sensitive to the multiple presence of trucks. This shows that the differences
492 for return level of type I3 load effect between conventional method and mixture
493 method becomes larger with the increase of span length.

494 To further demonstrate the accuracy of the proposed MPOT method, a com-
495 parison study between the present MPOT method and the composite distribution
496 statistic (CDS) approach proposed by (Caprani et al. 2008), which fit GEV dis-
497 tribution to block maxima for load effects resulting from same loading event, has

498 been performed to predict the characteristic values for 100-year return period
499 and 1000-year period. The relative differences between these two approaches are
500 given in Table 8. The two methods seem to provide consistent results. In general,
501 the differences are less than 10%, it can be concluded that the two loading event
502 depended methods have similar performance. However, it is also clear that some
503 of the differences are significant, especially for longer span lengths.

504 It is clear from Fig.10 that the CPOT method is strongly governed by the
505 relative frequency extremes, it thus results in the upper tail with less observed
506 extremes poorly fitted. While the proposed MPOT method considering the con-
507 tribution by type of loading event that results in a well captured tail. Quantitative
508 comparison in terms of characteristic value for 100- and 1000-year return period
509 further demonstrates the difference between the two methods. Due to the lack of
510 sufficient long-term measured traffic data, although it is impossible to provide an
511 directly comparison between predict method and measurement, the comparison
512 between the MPOT method and the CDS method provides confidence that the
513 MPOT method can provide sufficiently accurate prediction.

514 **CONCLUSIONS**

515 Special caution should be taken when estimating the high quantile or finding
516 the extreme value distribution for bridge traffic load effects. A novel method is
517 proposed in the present paper to study extreme value distribution of bridge traffic
518 load effects and properly predict the characteristic values for long return periods.
519 The proposed method is based on the generalized Pareto distribution as the clas-
520 sic Peaks-over-Threshold method. But conversely to the GPD which is seldom
521 fitted to load effects resulting from the same loading event defined by number of
522 simultaneously involved trucks/vehicles, since bridge traffic load effects generally
523 result from different loading events, the proposed method accounts for various

524 numbers of simultaneous trucks/vehicles on the bridges. Thus, the upper tail of
525 the load effect distribution can be approximated by a mixture generalized distri-
526 bution, and the method is thus named Mixture Peaks-over-Threshold Approach.
527 Numerical studies have been conducted to demonstrate the capability of the pro-
528 posed method in predicting characteristic values and extreme value distribution
529 for bridge traffic load effects. In a theoretical example with known distribution,
530 comparison between conventional extreme value estimation methods and the pro-
531 posed method shows that the proposed MPOT method has better performance
532 to capture the upper tail of the parent distribution and the maximum value dis-
533 tribution. In the traffic load effects example, the differences can be seen between
534 the conventional methods and the proposed method for predicting characteristic
535 values. Consistent results have been obtained from the proposed MPOT method
536 and the composite statistic distribution method. It is believed that the proposed
537 MPOT method provides more accurate and reasonable prediction as it considers
538 the non-identically distributed nature of load effects.

539 **ACKNOWLEDGMENT**

540 The authors gratefully acknowledge the financial support provided for this
541 study by the Marie Curie Initial Training Network TEAM (Training in European
542 Asset Management) project that has been funded by the European Community's
543 7th Framework Programme.

544 **REFERENCES**

- 545 Bailey, S. F. and Bez, R. (1999). "Site specific probability distribution of extreme
546 traffic action effects." *Probabilistic Engineering Mechanics*, 14(12), 19–26.
- 547 Balkema, A. A. and de Haan, L. (1974). "Residual life time at great age." *The*
548 *Annals of Probability*, 792–804.

- 549 Cai, C. S. and Chen, S. R. (2004). “Framework of vehicle–bridge–wind dynamic
550 analysis.” *Journal of Wind Engineering and Industrial Aerodynamics*, 92(78),
551 579–607.
- 552 Caprani, C. C., O’Brien, E. J., and McLachlan, G. J. (2008). “Characteristic
553 traffic load effects from a mixture of loading events on short to medium span
554 bridges.” *Structural Safety*, 30(5), 394–404.
- 555 CEN (2003). *Eurocode 1: Actions on Structures - Part 2: Traffic Loads on*
556 *Bridges*. European Committee for Standardization, Brussels, Belgium.
- 557 Chen, S. R. and Wu, J. (2011). “Modeling stochastic live load for long-span
558 bridge based on microscopic traffic flow simulation.” *Computers & Structures*,
559 89(910), 813–824.
- 560 Choulakian, V. and Stephens, M. A. (2001). “Goodness-of-fit tests for the gener-
561 alized pareto distribution.” *Technometrics*, 43(4), 478–484.
- 562 Coles, S. (2001). *An introduction to statistical modeling of extreme values*.
563 Springer Series in Statistics. Springer-Verlag London Limited, London.
- 564 Cooper, D. I. (1997). “Development of short span bridge-specific assessment live
565 loading.” *The safety of bridges*, P. C. Das, ed. Thomas Telford, 64–89.
- 566 COST 345 (2002). *Procedures required for assessing highway structures: final*
567 *report*. European Commission Directorate General Transport and Energy.
- 568 Cremona, C. (2001). “Optimal extrapolation of traffic load effects.” *Structural*
569 *Safety*, 23(1), 31–46.
- 570 de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*.
571 Springer series in operations research and financial engineering. Springer.
- 572 de Zea Bermudez, P. and Kotz, S. (2010). “Parameter estimation of the gener-
573 alized Pareto distribution - Part I & II.” *Journal of Statistical Planning and*
574 *Inference*, 140, 1353–1388.
- 575 Desrosiers, R. and Grillo, R. (1973). *Estimating the Frequency of Multiple Truck*

576 *Loadings on Bridges: Final Report*. University of Connecticut.

577 Enright, B., Carey, C., and Caprani, C. (2013). “Microsimulation evaluation of
578 Eurocode load model for American long-span bridges.” *Journal of Bridge En-*
579 *gineering*, 18(12), 1252–1260.

580 Enright, B. and O’Brien, E. J. (2012). “Monte Carlo simulation of extreme traf-
581 fic loading on short and medium span bridges.” *Structure and Infrastructure*
582 *Engineering*, 1–16.

583 Frangopol, D., Strauss, A., and Kim, S. (2008). “Bridge reliability assessment
584 based on monitoring.” *Journal of Bridge Engineering*, 13(3), 258–270.

585 Fu, G., Liu, L., and Bowman, M. (2011). “Multiple presence factor for truck load
586 on highway bridges.” *Journal of Bridge Engineering*, 18(3), 240–249.

587 Fu, G. and You, J. (2009). “Truck loads and bridge capacity evaluation in China.”
588 *Journal of Bridge Engineering*, 14(5), 327–335.

589 Ghosh, J., Caprani, C., and Padgett, J. (2013). “Influence of traffic loading on the
590 seismic reliability assessment of highway bridge structures.” *Journal of Bridge*
591 *Engineering*, 19(3), 04013009.

592 Ghosn, M. and Moses, F. (2000). “Effect of changing truck weight regulations on
593 U.S. bridge network.” *Journal of Bridge Engineering*, 5(4), 304–310.

594 Gindy, M. and Nassif, H. H. (2007). “Multiple presence statistics for bridge live
595 load based on weigh-in-motion data.” *Transportation Research Record: Journal*
596 *of the Transportation Research Board*, 125–135.

597 Harman, D. J. and Davenport, A. G. (1979). “A statistical approach to traffic
598 loading on highway bridges.” *Canadian Journal of Civil Engineering*, 6(4),
599 494–513.

600 Juarez, S. F. and Schucany, W. R. (2004). “Robust and efficient estimation for
601 the generalized pareto distribution.” *EXTREMES*, 7, 237–251.

602 Leadbetter, R. M., Lindgren, G., and Rootzen, H. (1983). *Extremes and related*

603 *properties of random sequences and processes*. Springer-Verlag, New York Hei-
604 delberg Berlin.

605 Li, S., Zhu, S., Xu, Y.-L., Chen, Z.-W., and Li, H. (2012). “Long-term condition
606 assessment of suspenders under traffic loads based on structural monitoring
607 system: Application to the Tsing Ma Bridge.” *Structural Control and Health*
608 *Monitoring*, 19(1), 82–101.

609 Madsen, H., Rasmussen, P. F., and Rosbjerg, D. (1997). “Comparison of annual
610 maximum series and partial duration series methods for modeling extreme hy-
611 drologic events: 1. at-site modeling.” *Water Resources Research*, 33(4), 747–
612 757.

613 Messervey, T. B., Frangopol, D. M., and Casciati, S. (2010). “Application of the
614 statistics of extremes to the reliability assessment and performance prediction
615 of monitored highway bridges.” *Structure and Infrastructure Engineering*, 7(1-
616 2), 87–99.

617 Nowak, A. S. (1993). “Live load model for highway bridges.” *Structural Safety*,
618 13, 53–66.

619 O’Brien, E. and Caprani, C. C. (2005). “Headway modelling for traffic load as-
620 sessment of short to medium span bridges.” *The Structural Engineer*, 86(16),
621 33–36.

622 O’Brien, E., Schmidt, F., Hajializadeh, D., Zhou, X. Y., Enright, B., Caprani,
623 C. C., Wilson, S., and Sheils, E. (2015). “A review of probabilistic methods of
624 assessment of load effects in bridges.” *Structural Safety*, 53, 44–56.

625 O’Brien, E. J., Bordallo-Ruiz, A., and Enright, B. (2014). “Lifetime maximum
626 load effects on short-span bridges subject to growing traffic volumes.” *Structural*
627 *Safety*, 50, 113–122.

628 O’Brien, E. J., Sloan, D. T., Bulter, K. M., and Kirkpatrick, J. (1995). “Traffic
629 load ‘fingerprinting’ of bridges for assessment purposes.” *The Structural Engi-*

- 630 *neer*, 73(19), 320–324.
- 631 O’Connor, A. and O’Brien, E. J. (2005). “Traffic load modelling and factors
632 influencing the accuracy of predicted extremes.” *Canadian Journal of Civil*
633 *Engineering*, 32(1), 270–278.
- 634 Park, H. W. and Sohn, H. (2006). “Parameter estimation of the generalized ex-
635 treme value distribution for structural health monitoring.” *Probabilistic Engi-*
636 *neering Mechanics*, 21(4), 366–376.
- 637 Pickands III, J. (1975). “Statistical inference using extreme order statistics.” *The*
638 *Annals of Statistics*, 3(1), 119–131.
- 639 Righiniotis, T. D. (2006). “Effects of increasing traffic loads on the fatigue re-
640 liability of a typical welded bridge detail.” *International Journal of Fatigue*,
641 28(8), 873–880.
- 642 Scarrott, C. and MacDonald, A. (2012). “A review of extreme value threshold
643 estimation and uncertainty quantification.” *REVSTAT - Statistical Journal*,
644 10(1), 33–60.
- 645 Sivakumar, B., Ghosn, M., Moses, F., and TranSystems Corporation (2011).
646 “Protocols for collecting and using traffic data in bridge design.” *National Co-*
647 *operative Highway Research Program (NCHRP) Report 683*, Lichtenstein Con-
648 sulting Engineers, Inc., Washington, D. C.
- 649 Tong, G., Aiqun, L., and Jianhui, L. (2008). “Fatigue life prediction of welded
650 joints in orthotropic steel decks considering temperature effect and increasing
651 traffic flow.” *Structural Health Monitoring*, 7(3), 189–202.
- 652 Zhou, X., Schmidt, F., Toutlemonde, F., and Jacob, B. (2014). “Applying Weigh-
653 in-Motion traffic data to reliability based assessment of bridge structures.”
654 *Safety, Reliability, Risk and Life-Cycle Performance of Structures and Infras-*
655 *tructures*. CRC Press, 3831–3838.
- 656 Zhou, X.-Y. (2013). “Statistical analysis of traffic loads and their effects on

657 bridges using Weigh-in-Motion data collected in France,” Ph.D. thesis, Uni-
658 versité Paris-Est.

659 Zhu, B. and Frangopol, D. (2012). “Risk-based approach for optimum mainte-
660 nance of bridges under traffic and earthquake loads.” *Journal of Structural*
661 *Engineering*, 139(3), 422–434.

662 **List of Tables**

663	1	Probabilities for six categories of loading events for data above 90 th	
664		and 95 th percentile	30
665	2	Parameter estimates for the CPOT method by various estimators	31
666	3	Parameter estimates for the MPOT method by various estimators	31
667	4	Root mean square error at various thresholds	31
668	5	Optimal threshold selection for I1 (bending moment at mid-span	
669		of simply supported bridge) with bridge length of 40m	32
670	6	Percentage difference of 100-year return level between conventional	
671		and mixture method (%)	33
672	7	Difference in 1000-year return level between conventional and mix-	
673		ture model (%)	33
674	8	Difference (mixture POT vs. mixture GEV)	34

TABLE 1: Probabilities for six categories of loading events for data above 90th and 95th percentile

Data	Type of loading event	20 m			30 m			40 m			50 m		
		I1	I2	I3	I1	I2	I3	I1	I2	I3	I1	I2	I3
$X_{0.90}$	1-truck	0.047	0.079	0.068	0.050	0.067	0.056	0.047	0.051	-	0.155	0.043	-
	2-truck	61.37	59.49	61.77	36.95	45.51	9.00	11.73	11.16	0.29	2.20	2.20	0.19
	3-truck	37.60	38.76	36.36	59.87	49.28	81.46	79.06	70.91	50.18	75.77	63.08	37.07
	4-truck	0.99	1.67	1.80	3.10	4.94	9.29	8.79	16.55	45.96	20.13	30.11	54.96
	5-truck	-	-	-	0.037	0.194	0.198	0.37	1.31	3.57	1.71	4.42	7.53
	6-truck	-	-	-	-	-	-	-	-	-	0.044	0.142	0.240
$X_{0.95}$	1-truck	0.094	0.079	0.136	0.100	0.073	0.113	0.094	0.081	-	-	0.085	-
	2-truck	39.87	42.38	41.97	13.72	26.73	3.87	4.25	4.07	0.57	1.38	1.59	0.29
	3-truck	58.38	54.98	55.00	81.11	65.50	83.99	84.04	73.84	28.86	72.35	57.96	18.02
	4-truck	1.67	2.55	2.89	5.03	7.37	11.74	11.18	20.04	65.14	23.92	34.24	69.99
	5-truck	-	-	-	0.050	0.328	0.282	0.44	1.94	5.43	2.26	5.87	11.31
	6-truck	-	-	-	-	-	-	-	-	-	0.089	0.255	0.384

TABLE 2: Parameter estimates for the CPOT method by various estimators

Estimator	Shape	Scale	Location	No. exceedances	KS, p-value
MM	-0.0767	10.21	510.52	1321	0.8823
PWM	-0.0930	10.37	510.52	1321	0.9735
ML	-0.0583	10.03	510.52	1321	0.6936
MDPD	-0.0760	10.20	510.52	1321	0.8726

TABLE 3: Parameter estimates for the MPOT method by various estimators

Item	Parameter	Estimator			
		MM	PWM	ML	MDPD
Comp. 1	Shape, ξ	-0.173	-0.105	-0.177	-0.177
	Scale, σ	9.9	10.0	10.0	10.0
	Location, μ	515.2	508.0	515.2	515.2
	No. exceed.	707	1500	707	707
	KS p-value	0.908	0.909	0.922	0.922
Comp. 2	Shape, ξ	-0.056	-0.058	-0.053	-0.057
	Scale, σ	15.9	16.0	15.9	16.0
	Location, μ	479.1	479.1	479.1	479.1
	No. exceed.	1371	1371	1371	1371
	KS p-value	0.926	0.903	0.945	0.918
Mixture	KS p-value	0.964	0.866	0.979	0.974

TABLE 4: Root mean square error at various thresholds

Threshold	No.	Method	MM	PWM	ML	MDPD
$X_{0.90}$	6403	CPOT	0.0091	0.0083	0.0035	0.0066
		MPOT	0.004	0.0059	0.0032	0.0062
$X_{0.92}$	5122	CPOT	0.0079	0.0079	0.0034	0.0063
		MPOT	0.0033	0.0054	0.0032	0.0065
$X_{0.94}$	3842	CPOT	0.0099	0.0083	0.0064	0.0099
		MPOT	0.0061	0.0079	0.0042	0.0071
$X_{0.96}$	2561	CPOT	0.0095	0.0084	0.0048	0.0083
		MPOT	0.0051	0.0071	0.0039	0.0069
$X_{0.98}$	1281	CPOT	0.0086	0.0086	0.0041	0.0061
		MPOT	0.0035	0.0059	0.0033	0.0068

TABLE 5: Optimal threshold selection for I1 (bending moment at mid-span of simply supported bridge) with bridge length of 40m

Statistic	Estimator	2-truck			3-truck			4-truck			KS p-value
		Shape	Scale	Threshold	Shape	Scale	Threshold	Shape	Scale	Threshold	
AD	MM	0.0628	262.5	6540	-0.2185	830.4	6540	-0.1874	1114.0	6540	0.09
	PWM	0.0952	253.5	6540	-0.2056	821.6	6540	-0.1918	1118.2	6540	0.10
	ML	0.0725	259.9	6540	-0.2771	812.7	6864	-0.1887	1115.5	6540	0.74
	MDPD	0.0884	256.4	6540	-0.2728	808.9	6864	-0.1873	1113.9	6540	0.55
CM	MM	0.0536	269.8	6571	-0.2567	798.0	6864	-0.1874	1114.0	6540	0.66
	PWM	0.0952	253.5	6540	-0.2483	792.7	6864	-0.1918	1118.2	6540	0.48
	ML	0.0725	259.9	6540	-0.2813	803.3	6921	-0.1887	1115.5	6540	0.86
	MDPD	0.0884	256.4	6540	-0.2728	808.9	6864	-0.1873	1113.9	6540	0.55

TABLE 6: Percentage difference of 100-year return level between conventional and mixture method (%)

Load effect	Length	BM/GEV	POT/GPD			
			MM	PWM	ML	MDPD
I1	20	-8.49	0.11	0.43	0.19	0.17
	30	-9.56	-6.31	-10.40	-8.18	-9.66
	40	-14.63	-8.27	-7.90	-1.82	-7.19
	50	-16.98	15.78	-2.71	20.32	21.18
I2	20	5.12	-0.47	1.54	0.20	0.36
	30	-20.60	-3.02	-0.32	-6.33	-3.66
	40	-9.51	-3.02	-16.38	-16.02	-21.40
	50	-11.22	0.08	-2.73	1.20	1.04
I3	20	-29.92	-4.55	-7.11	-1.63	-4.30
	30	-15.22	-5.89	-9.69	-4.11	-5.92
	40	-8.28	5.76	20.09	16.59	23.89
	50	-17.85	8.44	24.03	9.14	14.40

TABLE 7: Difference in 1000-year return level between conventional and mixture model (%)

Load effect	Length	BM/GEV	POT/GPD			
			MM	PWM	ML	MDPD
I1	20	-10.62	0.24	0.64	0.30	0.28
	30	-16.20	-13.50	-22.30	-22.38	-24.88
	40	-29.67	-9.78	-11.11	-1.00	-9.44
	50	-36.45	34.53	-1.39	44.65	46.16
I2	20	8.65	-0.80	1.93	0.05	0.28
	30	-25.62	-8.36	-8.90	-11.17	-8.81
	40	-11.39	-4.48	-36.71	-36.18	-42.90
	50	-13.91	1.26	-2.68	2.69	2.58
I3	20	-41.27	-8.28	-12.52	-3.83	-7.92
	30	-17.82	-7.10	-13.42	-6.72	-10.99
	40	-10.21	9.40	34.22	28.00	40.81
	50	-17.65	15.34	40.60	16.50	24.94

TABLE 8: Difference (mixture POT vs. mixture GEV)

Load effect	Length	100-year				1000-year			
		MM	PWM	ML	MDPD	MM	PWM	ML	MDPD
I1	20	0.43	-0.65	-0.57	0.89	0.48	-0.74	-0.65	1.02
	30	2.14	0.13	1.22	9.21	8.13	8.84	11.73	17.17
	40	-0.14	-0.05	0.25	0.46	1.80	0.14	1.32	8.82
	50	-1.78	0.50	-0.10	-1.38	-2.44	0.70	-0.14	-1.85
I2	20	0.13	-1.91	-1.38	1.74	0.08	-2.45	-1.79	2.14
	30	-3.81	3.98	0.60	-3.03	-0.92	3.74	0.41	-3.95
	40	16.27	15.95	23.47	32.22	51.41	50.41	67.44	87.83
	50	2.79	-0.85	-0.99	-1.23	3.93	-1.07	-1.35	-1.53
I3	20	5.27	-4.12	-0.21	15.40	8.39	-6.18	-0.32	25.15
	30	3.34	-1.77	-0.06	2.40	6.06	-0.29	4.23	13.06
	40	-0.07	-0.72	0.00	5.57	-0.12	-1.28	-0.03	10.09
	50	-5.15	1.49	0.36	-2.09	-7.19	2.18	0.51	-2.80

675 **List of Figures**

676 1 Characteristics of measured traffic data 36

677 2 Classification of vehicles/trucks 36

678 3 Time history of load effects 37

679 4 Time history and local extreme 37

680 5 Histogram of load effects due to various types of loading events . . 38

681 6 Probabilities for six types of loading events (left) over 90th percentile 39

682 7 Gumbel scaled cumulative distribution probability plot. 39

683 8 Comparison of estimates of the characteristic values obtained from

684 between CPOT and MPOT. 40

685 9 Extreme value distribution from CPOT and MPOT with true dis-

686 tribution. 40

687 10 Diagnosis plot for threshold excess model fitted to load effect. . . 41

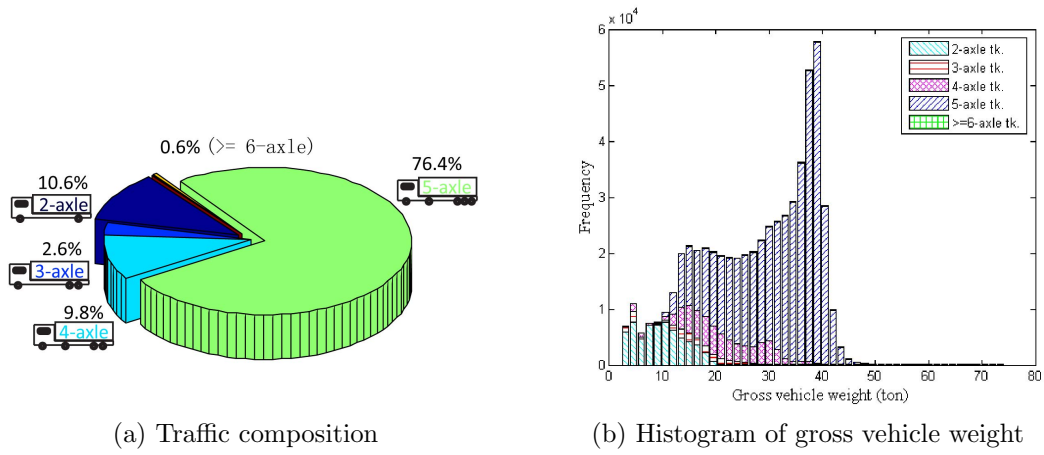


FIG. 1: Characteristics of measured traffic data



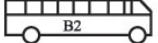


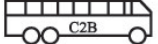

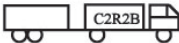
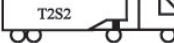

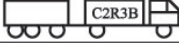
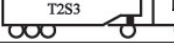

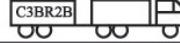
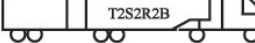
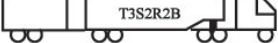

2-axle	 	
3-axle	 	 
4-axle	 	
5-axle	 	 
6-axle		
7-axle		
8-axle		

FIG. 2: Classification of vehicles/trucks

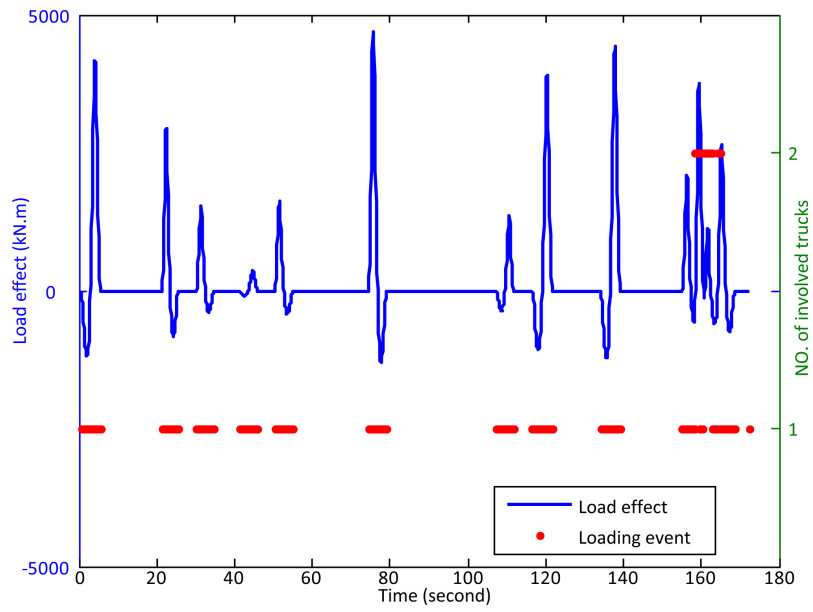


FIG. 3: Time history of load effects

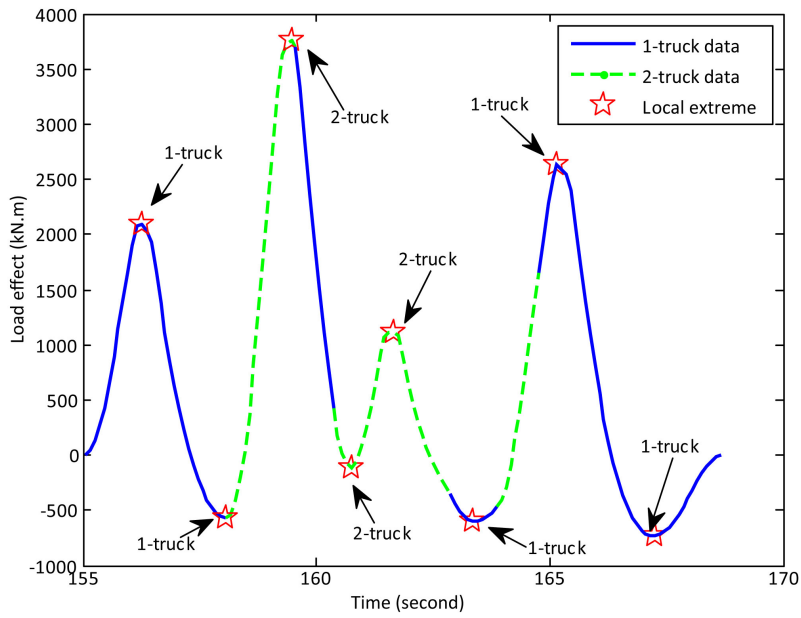


FIG. 4: Time history and local extreme

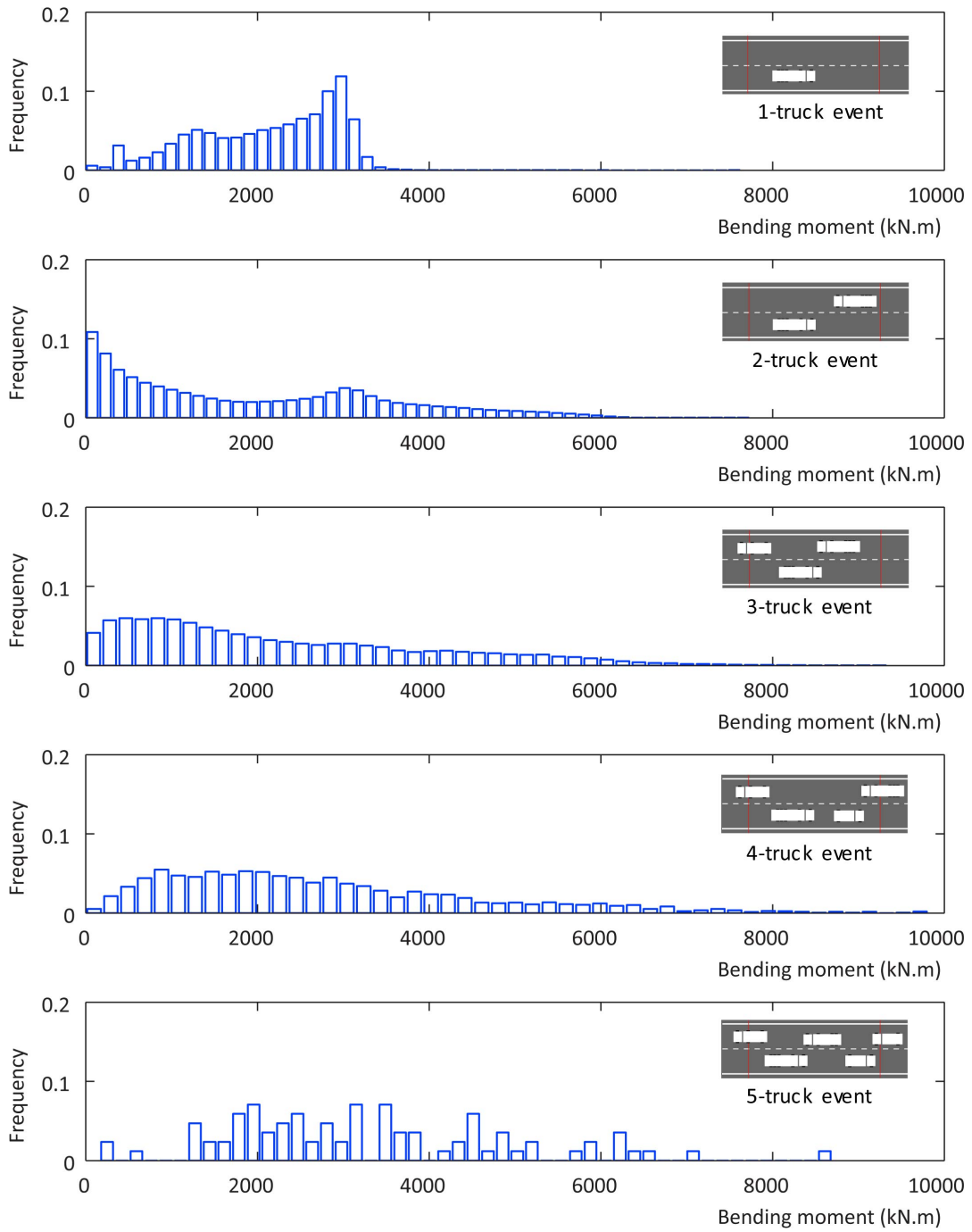


FIG. 5: Histogram of load effects due to various types of loading events

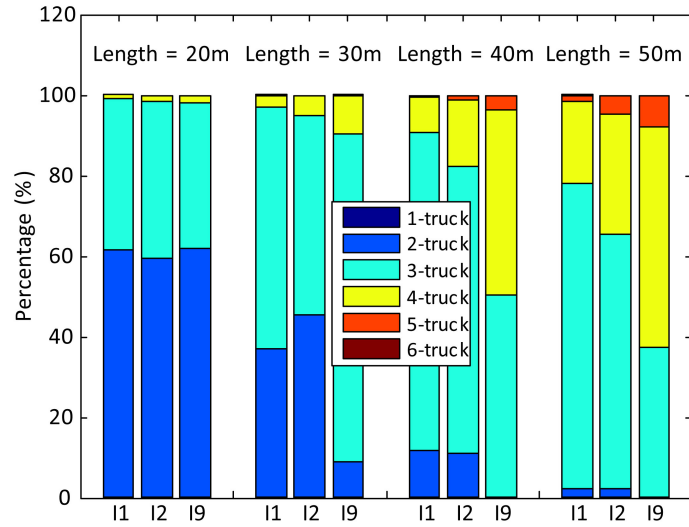


FIG. 6: Probabilities for six types of loading events (left) over 90th percentile

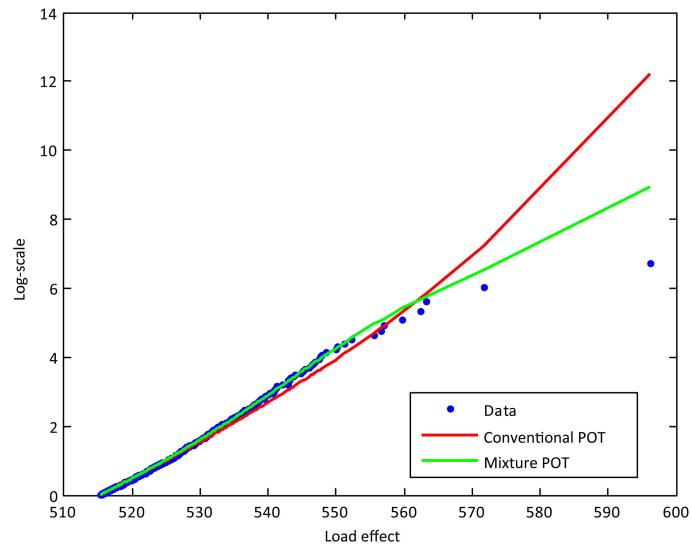


FIG. 7: Gumbel scaled cumulative distribution probability plot.

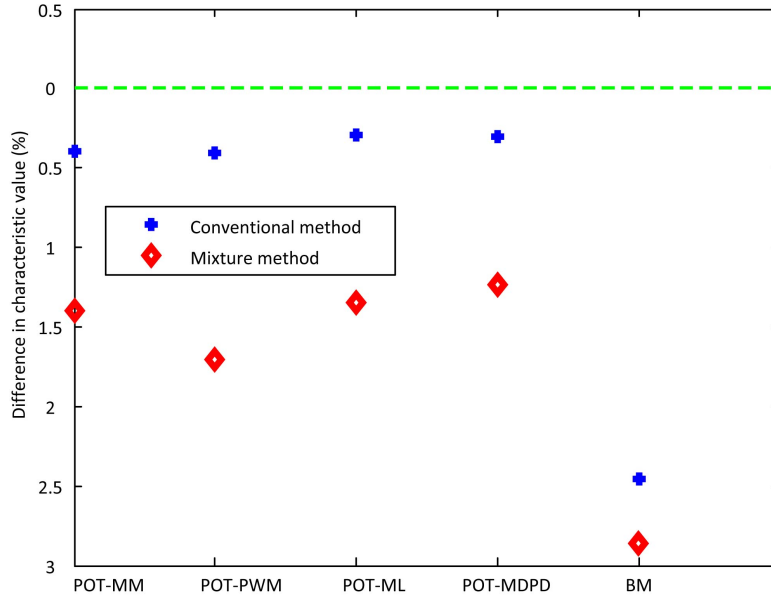


FIG. 8: Comparison of estimates of the characteristic values obtained from between CPOT and MPOT.

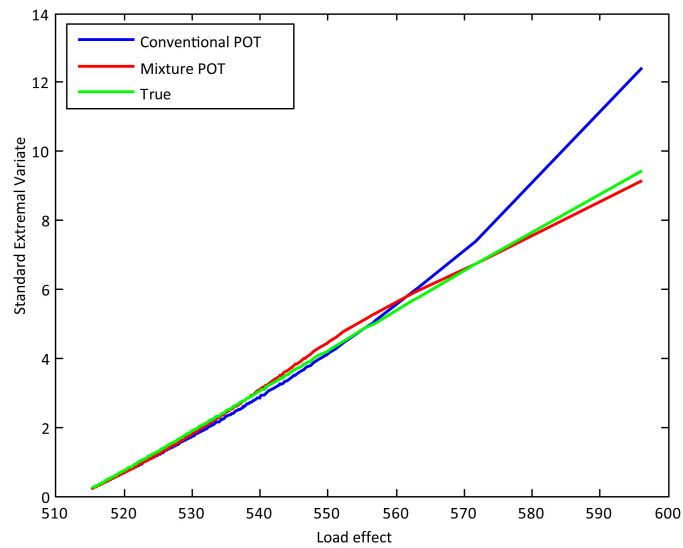


FIG. 9: Extreme value distribution from CPOT and MPOT with true distribution.

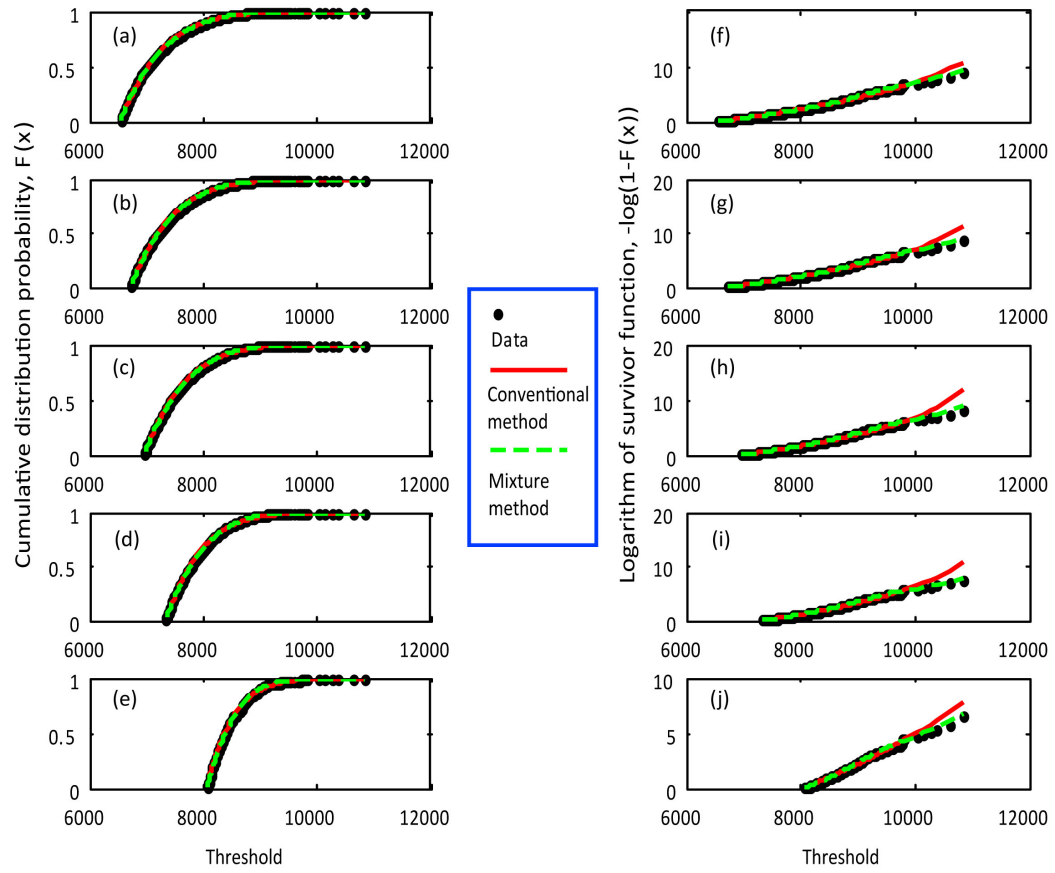


FIG. 10: Diagnosis plot for threshold excess model fitted to load effect.