



**HAL**  
open science

## A Review of Probabilistic Methods of Assessment of Load Effects in Bridges

Ej Obrien, Franziska Schmidt, Donya Hajializadeh, Xiao Yi Zhou, Bernard Enright, Colin Caprani, S Wilson, Emma Sheils

► **To cite this version:**

Ej Obrien, Franziska Schmidt, Donya Hajializadeh, Xiao Yi Zhou, Bernard Enright, et al.. A Review of Probabilistic Methods of Assessment of Load Effects in Bridges. *Structural Safety*, 2015, 53, pp 44-66. 10.1016/j.strusafe.2015.01.002 . hal-01213059

**HAL Id: hal-01213059**

**<https://hal.science/hal-01213059>**

Submitted on 7 Oct 2015

**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 Review of Probabilistic Methods of Assessment of Load Effects in Bridges**

O'Brien, E.J.<sup>1</sup>, Schmidt, F.<sup>2</sup>, Hajializadeh, D.<sup>1</sup>, Zhou, X.-Y.<sup>2</sup>, Enright, B.<sup>3</sup>, Caprani, C.C.<sup>3</sup>, Wilson, S.<sup>1</sup>, Sheils, E.<sup>1</sup>

<sup>1</sup> University College of Dublin

<sup>2</sup> Ifsttar – Université Paris Est

<sup>3</sup> Trinity College Dublin

### **Abstract**

This paper reviews a range of methods used to infer characteristic values and probabilities of failure from a sample of measured or simulated load effect data. The popular methods of Peaks-Over-Threshold and Generalized Extreme Value (GEV) are considered but also other methods including the Box-Cox approach, fitting to a Normal distribution and the Rice formula. These five methods are fitted to the tail of the daily maximum data. Bayesian Updating and Predictive Likelihood are also considered, but are fitted to the entire data set. In general, the five tail fitting methods are reasonably accurate at inferring characteristic annual maximum values from 1000 days of data and the other methods less so. All methods are considerably less accurate at inferring probabilities of failure than characteristic values.

**Keywords:** Review, Bridge, Load, Traffic, Assessment, POT, Peaks-Over-Threshold, Extreme Value, GEV, Box-Cox, Rice, Predictive Likelihood.

### **1. Introduction**

A necessary part of bridge management is the assessment of the safety of bridge structures. In its simplest form, a bridge is safe when its capacity to resist load exceeds the load applied. More precisely, a bridge can be considered safe when there is an acceptably low probability that load exceeds capacity. A great deal of work has been carried out on methods of evaluating the load-carrying capacity of bridges and the associated uncertainties. Load-carrying capacity can be reduced by different forms of deterioration, depending on factors such as the structural material, the quality of workmanship during construction, the age of the structure, the environment and the loading history. To carry out a more accurate assessment of the load-carrying capacity, non-destructive and/or destructive tests can be carried out to get more detailed site specific information on these deterioration mechanisms to reduce uncertainty and associated conservatism (Al-Harthy et al., 2011, Frangopol & Liu, 2007, Richard et al., 2012, Rücker et al., 2005, Suo & Stewart, 2009). These inspection results can be incorporated into time-dependent reliability-based assessments to give up-to-date structure-specific deterioration rates. These in turn can be used to accurately predict the capacity of the structure and to schedule maintenance and repairs (Melchers, 1999, Orcesi & Cremona, 2009, Orcesi & Cremona, 2010, Sheils et al., 2010).

Traffic loading on bridges, perhaps the greatest source of uncertainty, has received less attention and is the focus of this paper. Historical developments in the field of traffic loading are reviewed. A wide

range of statistical/probabilistic approaches have been applied to the problem with no clear ‘winner’ emerging.

Two Extreme Value examples are used here as benchmark tests, against which a range of approaches are compared. The first example is the problem of finding the maximum of numerous normally distributed random variables, a problem for which the exact theoretical solution is known. The many methods of estimating characteristic maxima are shown to give large errors, to varying degrees.

The second example is based on a carefully calibrated traffic load simulation model. The simulation is run for 5 000 years so that, while the exact solution is unknown, it can be estimated very well and there is a high degree of confidence in the lifetime maximum results. As for the first example, several methods of prediction, using modest quantities of data, are tested. Again, the errors in the predictions are found to be significant and to vary considerably, depending on the approach adopted.

## **2. Review of Literature**

Load effects (LE’s) – bending moments, shear forces, etc. – result from traffic passing over a bridge. The process varies in time with many periods of zero LE when there is no traffic on the bridge and peaks corresponding to heavy vehicle crossings or more complex vehicle meeting or overtaking scenarios. The majority of the local peaks in LE are due to cars which are relatively light and there have been many efforts to simplify the problem by excluding consideration of these data. The methods of statistical inference used in the literature to predict the extremes of traffic LE’s are quite diverse.

### *Tail Fitting*

In the context of this problem, many approaches fit a distribution to the tail of the Cumulative Distribution Function (CDF). This can be justified by the fact that the distribution is often made up of a mixture of load effect types – for example, LE’s due to 2-axle trucks and those due to heavy low-loader vehicles. For bridge traffic loading, the heavier vehicles tend to dominate, with the lighter ones making very little contribution to the probability of exceedance at the extremes. The tail can be chosen by engineering judgement when the cumulative distribution is seen to change at a particular probability level. Alternatively, some authors have fitted to the top  $2\sqrt{n}$  of a distribution of  $n$  data, based on theoretical considerations (Castillo, 1988). Others have fitted to the top 30% of data (Enright, 2010) based on sensitivity analyses.

Two of the tail fitting approaches are particularly popular – Peaks-Over-Threshold (POT) and Block Maximum. POT considers the extent by which the peaks of LE exceed a specified threshold. The POT LE’s are fitted to a probability distribution such as the Generalized Pareto distribution. In the Block Maximum approach, only the maximum LE’s in given blocks of time (days, years, etc.), are considered. This has the advantage of time referencing the data which is necessary when calculating lifetime maximum probabilities of exceedance. Block maximum LE’s can be fitted to one of a range of distribution types such as Generalised Extreme Value (GEV) (incorporating Gumbel, Weibull and Fréchet), or Normal. Fitting block maximum values to GEV and Normal distributions will be considered here.

The Block Maximum approach has the disadvantage that only one LE in each block of time is considered, even if several very large LE’s are recorded. The POT approach addresses this issue but

the selection of the threshold, below which LE's are discarded, is subjective. The Box-Cox approach is more general and aims to address the disadvantages of both POT and GEV. The Rice formula is also investigated as it was used for the extrapolations in the background study supporting the development of the Eurocode for traffic loading on bridges. However, while the Rice formula is a fitting to tail data, it is applied to a histogram of 'upcrossings' past a threshold, not to a CDF, and assumes a normally distributed process.

### *Full Distribution Fitting*

Bayesian Updating is another approach that can be applied to bridge traffic loading. A probability distribution is assumed for the block maximum LE's and is updated using available LE data. While only tail data could be used, in this work, the Bayesian approach is used to update the entire distribution, not just the tail. Predictive Likelihood also seeks to develop a probability distribution for all LE's but uses a frequentist likelihood approach, assigning likelihoods on the basis of the quality of the fit to the measured data.

## **2.1 Peaks Over Threshold (POT)**

Block Maximum approaches use only the maximum LE in each block of time. There is therefore a risk that some important data is discarded: if two unrelated extreme loading events occur in the same block of time, only one of the resulting LE's is retained. In such a case, the POT approach would retain both LE's as valid data.

To find characteristic maximum values of LE, data above the threshold must be fitted to a probability distribution. Coles (2001) provides a brief outline proof that the Generalized Pareto (GP) distribution approximates the CDF of such POT data well. Crespo-Minguillón & Casas (1997) use the GP distribution to model the excesses of weekly maximum traffic LE's over a threshold. James (2003) applies the POT method to analyse load effects on railway bridges. Gindy & Nassif (2006) analyse load effects caused by combined data from over 33 Weigh-in-Motion sites over an 11-year measurement period, and compare extreme values as predicted by both GP and GEV distributions.

A significant drawback of the POT approach is the issue of selecting the threshold. There are many different kinds of loading scenario on a typical bridge. For example, there are usually many single-vehicle crossings of standard 5-axle trucks. The probability distribution of LE's due to such an event type may be quite different from that due to large cranes or that due to 2-truck meeting events (Caprani et al., 2007). If the threshold is too low, there may be an excessive mixing of extreme event types with other less critical types which can result in convergence to an incorrect characteristic LE. On the other hand, if the threshold is too high, there will be too few peaks above the threshold, leading to high variance and unreliable results.

The basic principle in selecting a threshold is to adopt as low a threshold as possible, while maintaining a consistent trend. Two methods are available (Coles, 2001): one is an exploratory technique carried out prior to model estimation; the other is an assessment of the stability of parameter estimates, based on the fitting of models across a range of different thresholds. Crespo-Minguillón & Casas (1997) apply the latter method and select the optimal threshold based on the overall minimum least-squares value.

Having selected the threshold, the next step is to estimate the parameters of the GP (or other) distribution. Bermudez & Kotz (2010) consider several methods of estimating these parameters including the method of moments, the probability weighted method, the maximum likelihood method, and Bayesian updating. Crespo-Minguillón & Casas (1997) adopt a methodology that is based on the minimization of the weighted sum of squared errors. James (2003) and Gindy & Nassif (2006) use maximum likelihood estimation.

## 2.2 Block Maximum – Extreme Value Distributions

Extreme value theory is based around the extreme value theorem, proved by Gnedenko (1943) and based on initial work by Fisher & Tippett (1928) and Gumbel (1935). For a sequence of independent random variables  $X_1, X_2, \dots$ , with distribution function  $F(x) = \text{Prob}(X \leq x)$ , the distribution of  $\max(X_1, \dots, X_n)$  is  $F(x)^n$ . As  $n$  gets large, this degenerates to 0 if  $F(x) < 1$ , as is usual. The Fisher-Tippett theorem shows that a non-degenerate distribution can be found using a linear function of  $x$ , say  $a_n + b_n x$ . Then, there is a non-trivial limit to  $F(a_n + b_n x)$  and this limit must be in the form of the Generalised Extreme Value distribution (GEV), also known as the Fisher-Tippett distribution (Jenkinson, 1955, Von Mises, 1936):

$$F_{\text{GEV}}(x) = \begin{cases} \exp\left(-\left(1 + \xi\left(\frac{x - \mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}\right), & \text{if } \xi \neq 0 \\ \exp\left(-\exp\left(-\frac{x - \mu}{\sigma}\right)\right) & , \text{if } \xi = 0 \end{cases} \quad \text{Equation 1}$$

defined in terms of parameters  $\mu$ ,  $\sigma$  and  $\xi$  where  $\mu \in \mathbb{R}$  is the location parameter,  $\sigma > 0$  the scale parameter and  $\xi \in \mathbb{R}$  the shape parameter, such that  $1 + \xi(x - \mu)/\sigma > 0$ . Hence, for an appropriately large  $n$ , the exact distribution,  $F(x)^n$ , converges asymptotically to  $F_{\text{GEV}}(x)$ . For the Normal distribution, the theorem holds and it is well known that its limiting distribution is the Gumbel, the  $\xi = 0$  case of the GEV. However, convergence is slow (Cramér, 1946).

Each block maximum LE is the maximum of many traffic loading scenarios. As convergence may be slow, Caprani (2005) and OBrien et al. (2010) have fitted block maximum LE data with a ‘Normal to the power of  $n$ ’, i.e., a Normal distribution raised to some power,  $n$ , whose value is found by fitting to the data. This has merit for smaller data samples. Ghosn et al. (2003) determine the distribution of lifetime maximum LE by raising the parent distribution of LE to an appropriate power. In this way they determine the mean and coefficient of variation of the maximum LE. Caprani (2005) describes a probabilistic convolution method to obtain bending moments for single truck loading events and obtains the distribution of lifetime maximum LE by raising the parent distribution to an appropriate power. Other authors attempt to calculate the exact distribution of extreme load effect, based on a fit to the parent distribution (Bailey, 1996, Bailey & Bez, 1994, Cooper, 1995, Getachew, 2005, Ghosn & Moses, 1985, Nowak & Hong, 1991, Nowak et al., 1993). This is done by raising the initial distribution to an appropriate power.

Most researchers fit block maximum LE data to one of the extreme value distributions described by the GEV equation: Gumbel, Fréchet or Weibull (also known as Types I, II or III). The three types of distribution have distinct forms of behaviour, corresponding to the different forms of the tail in the

original distribution function (Fisher & Tippett, 1928). Coles (2001) establishes the conditions under which the Gumbel, Fréchet and Weibull distributions are the limiting forms for various parent distributions (Gumbel, 1958).

In early applications of Extreme Value theory, it was usual to adopt one of the three distributions, and then to estimate the relevant parameters. There are two weaknesses with this: first, a technique is required to choose which of the three distributions is most appropriate for the data at hand; second, once such a decision is made, subsequent inferences presume this choice to be correct, and do not allow for the uncertainty such a selection involves, even though this uncertainty may be substantial (Coles, 2001). Nevertheless, many studies (Caprani & O'Brien, 2006, Caprani et al., 2008, Kanda & Ellingwood, 1991, O'Connor & O'Brien, 2005) indicate that LE data is either Weibull or Gumbel and, given that Gumbel is a special case of Weibull (with shape parameter,  $\xi = 0$ ), an assumption that LE is always of the form of Equation 1, with  $\xi \leq 0$ , seems reasonable.

Grave (2001) uses a weighted least-squares approach to fit Weibull distributions to critical LE's. O'Connor (2001) fits Gumbel and Weibull distributions to a population of 'extreme' LE's. O'Brien et al. (2003) plot hourly maximum strain values on Gumbel probability paper. A least-squares, straight-line fit is made to the upper  $2\sqrt{n}$  data points in a similar manner to Grave (2001) and O'Connor (2001). González et al. (2003) also use the Gumbel and Weibull distributions to extrapolate bridge load effect. Getachew (2005) fits the Generalized Extreme Value distribution to the LE's from simulated 2-truck meeting events representing two weeks of traffic. Bailey (1996) describes the use of plots of the mean and standard deviation of load effects, to estimate the appropriate extreme value distribution. Bailey (1996), Bailey & Bez (1994) and Bailey & Bez (1999) describe a qualitative analysis of 500 simulated upper tails of mean maximum load effects plotted against the number of events that contribute. They determine that the Weibull distribution is most appropriate to model these tails and use maximum likelihood estimation. (Cooper, 1997) presents a traffic model of about 81 000 measured truck events, and uses it to determine the distribution of LE's due to a 'single event'. He raises this distribution to powers to determine the distribution of LE for 1, 4, 16, 256 and 1024 such events. A Gumbel distribution is then fitted to this 1024-event distribution and used to extrapolate to a 2400-year return period. Cooper (1997) converts histograms of two-week traffic LE's into CDF's, which he then raises to a power equal to the number of trucks per day, to give the distribution of daily block maxima.

Moyo et al. (2002) plot daily maximum strain values on Gumbel probability paper and use a least-squares fit to determine the parameters of the distribution. Buckland et al. (1980) use a Gumbel distribution to fit the 3-monthly maximum LE's and extrapolate to find characteristic values. Getachew (2005) uses the GEV distribution to model the parent distribution of load effect, but not as an asymptotic approximation to the distribution of extreme values. Sivakumar et al. (2011) adopt the Gumbel distribution to project the statistics of the maximum LE's for different return periods.

### **2.3 Box-Cox Approach**

Researchers commonly debate the merits of the POT method relative to the Block Maximum approach. The Box-Cox transform (Box & Cox, 1964) is used by Bali (2003) to introduce a more general extreme value distribution that encompasses the Generalised Pareto and Generalised Extreme Value distributions (Caprani & O'Brien, 2009, Rocco, 2010). This transformation offers the possibility of improving the rate of convergence to the limiting extreme value form, since different distributions

converge at different rates. This approach restricts the methodology to cases where the extreme data are strictly positive (Wadsworth et al., 2010) but still encompasses a wide variety of practical problems including traffic loading on bridges. The use of the Box-Cox transformation in extreme value analysis was considered before in an entirely different context by Eastoe & Tawn (2009).

The Box-Cox-GEV extreme value distributions are given by Bali (2003) as:

$$H(x) = \left(\frac{1}{\lambda}\right) \left( \left[ \exp \left\{ -[h(x)]_+^{1/\xi} \right\} \right]^\lambda - 1 \right) + 1 \quad \text{Equation 2}$$

in which

$$h(x) = 1 - \xi \left( \frac{x - \mu}{\sigma} \right) \quad \text{Equation 3}$$

The parameters of this distribution are those of the GEV ( $\mu$ ,  $\sigma$ ,  $\xi$ ) plus a ‘model parameter’,  $\lambda$ . As  $\lambda \rightarrow 1$ , Box-Cox converges to the GEV distribution. Conversely, as  $\lambda \rightarrow 0$ , by L’Hôpital’s Rule, it converges to the GP distribution. To apply this model, a high threshold is set on the parent distribution (Caprani & OBrien, 2009, Rocco, 2010). Bali (2003) uses a threshold of two standard deviations about the sample mean. Caprani & OBrien’s (2009) thresholds are taken in steps of 0.5 standard deviations in the range from  $-2.5$  to  $+2.5$  standard deviations about the sample mean. Tötterman (2010) suggests that the additional parameter should increase the accuracy for Box-Cox, compared with GEV and GP.

Bali & Theodossiou (2008) evaluate the performance of three extreme value distributions including the GP, GEV and Box-Cox. The empirical results show that the asymptotic distribution of the maximal and minimal returns fits the Box-Cox-GEV distribution in this case. A likelihood ratio test between the GEV and Box-Cox results in a rejection of the former (Bali & Theodossiou, 2008, Caprani & OBrien, 2009).

## 2.4 Block Maximum – Normal Distribution

Block maximum data is often fitted with extreme value distributions as each data point represents the maximum of a number of parent values. However, block maximum data is also sometimes fitted to a Normal distribution. Nowak (1999) uses a form of Normal (Gaussian) probability paper, i.e., he fits the block maximum data to a Normal distribution and extrapolates to find the characteristic maximum. In an earlier study, Nowak (1993) uses 2.4 hours as the block size and fits the maximum-per-block data to a Normal distribution. This distribution is then raised to an appropriate power to obtain the 75-year maximum LE distribution.

To calibrate the traffic load model for the AASHTO load and resistance factor design (LRFD) approach, Nowak and others use Normal probability paper to extrapolate the maximum LE’s for time periods from 1 day to 75 years, based on a set of 9250 heavy vehicles representing about two weeks of heavy traffic measured on a highway in Ontario (Kulicki et al., 2007, Moses, 2001, Nowak, 1994, Nowak, 1995, Nowak, 1999, Nowak & Hong, 1991, Nowak et al., 1993, Sivakumar et al., 2011). The expected values of the lifetime maximum LE’s are found by fitting a straight line to the tails of the data on Normal probability paper.

Kulicki et al. (2007) identify the fact that block maximum LE's due to measured trucks are not Normal but fits tail data to the Normal distribution. In the background studies for Eurocode 1, Flint & Jacob (1996) fit half-normal curves to the ends of the histograms of LE. They adopt a least-squares best fit method to estimate the distribution parameters. Multimodal (bimodal or trimodal) Gumbel and Normal distributions are also used.

## 2.5 Rice Formula

The Rice formula, introduced by Rice (1945) and described more recently by Leadbetter et al. (1983), can be used to find a parametric fit to statistical data. Ditlevsen (1994) suggests that a load effect created by the traffic on a long span bridge can be modelled as a Gaussian random process. Under that hypothesis, the mean rate  $v(x; \sigma, m, \dot{\sigma})$  of up-crossings for a threshold level,  $x > 0$  during a reference period  $T_{ref}$ , can be expressed by the Rice formula:

$$v(x) = \frac{\dot{\sigma}}{2\pi\sigma} \exp\left[-\frac{(x-m)^2}{2\sigma^2}\right] \quad \text{Equation 4}$$

where,  $x$  is the threshold value of LE,  $m$  is the mean value,  $\sigma$  is the standard deviation and  $\dot{\sigma}$  is the first derivative of  $\sigma$  with respect to time.

The CDF can be found from the definition of return period which is the mean period between two occurrences, or the value with an expectation of being crossed one time during the return period,  $R$  (Cremona, 2001):

$$F(x) = \exp\left[-Rv_0 \times \exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right)\right] \quad \text{Equation 5}$$

where,  $v_0$  is  $\sigma/2\pi\sigma$ .

Cremona (2001) suggests the Kolmogorov test (DeGroot, 1986) to select the optimal number of class intervals and starting points. Getachew (2003) adopts Cremona's approach for the analysis of traffic LE's on bridges induced by measured and Monte Carlo simulated vehicle data. O'Connor & OBrien (2005) compare the predicted extremes of simply supported moment for a range of span lengths by the Rice formula, Gumbel and Weibull Extreme Value distributions: they find about 10% difference between Rice and the others. Finally, Jacob (1991) uses Rice's formula to predict characteristic LE's for the cases of free and congested traffic in background studies for the development of the Eurocode.

## 2.6 Fitting Distributions to Extreme Data & Bayesian Inference

The concept of Bayesian Updating stems from Bayes' Theorem and is a major pillar of modern statistics. Bayesian Updating involves the adoption of an initial (prior) probability distribution, perhaps based on past experience, and updating it on the basis of measured data to give a posterior distribution (Basu, 1964, Bhattacharya, 1967, Holla, 1966).



Sinha & Sloan (1988) use Bayesian Inference to find the full 3-parameter Weibull distribution from measured data. They propose the use of Bayes Linear Estimate to approximate the posterior expectations and formulate the corresponding calculations for the Weibull parameters. Smith & Naylor (1987) work with the 3-parameter Weibull distribution, comparing Maximum Likelihood with Bayesian estimators, using specially adapted versions of numerical quadrature to perform the posterior calculations. Although the priors they work with are arbitrary, they are chosen to reflect a range of potential scientific hypotheses. They report that the Bayesian inferential framework as a whole proves more satisfactory for their data analysis than the corresponding likelihood-based analysis. The issue of prior elicitation is pursued by Singpurewalla & Song (1988), who restrict attention to the 2-parameter Weibull model. The predictive density function (Aitchison & Dunsmore, 1980) is defined as:

$$f(y|x) = \int f(y|\theta)f(\theta|x)d\theta \quad \text{Equation 6}$$

where  $x$  represents historical data,  $y$  a future observation,  $\theta$  the vector of parameters describing the distribution,  $f(y|\theta)$  the likelihood and  $f(\theta|x)$  the posterior distribution of  $\theta$  given  $x$ . Thus, the predictive distribution averages the distribution across the uncertainty in  $\theta$  as measured by the posterior distribution. Lingappaiah (1984) develops bounds for the predictive probabilities of extreme order statistics under a sequential sampling scheme, when sampling is carried out from either an exponential or Pareto population. From a practical viewpoint, the most important issues arising from the Bayesian literature are the elicitation and formulation of genuine prior information in extreme value problems, and the consequent impact such a specification has on subsequent inferences. Coles & Tawn (1996) consider a case study in which expert knowledge is sought and formulated into prior information as the basis for Bayesian analysis of extreme rainfall.

## 2.7 Predictive Likelihood

The relatively new theory of frequentist Predictive Likelihood can be used to estimate the variability of the predicted value, or predictand. Fisher (1959) is the first clear reference to the use of likelihood as a basis for prediction in a frequentist setting. A value of the predictand ( $z$ ) is postulated and the maximized joint likelihood of the observed data ( $y$ ) and the predictand is determined, based on a probability distribution with given parameters. The graph of the likelihoods thus obtained for a range of values of the predictand, yields a predictive distribution. Such a predictive likelihood is known as the profile predictive likelihood. Denoting a normed likelihood by  $\bar{L}(\theta; x)$ , this is given by:

$$L_P(z|y) = \sup_{\theta} \bar{L}_y(\theta; y) \bar{L}_z(\theta; z) \quad \text{Equation 7}$$

This formulation states that the likelihood of the predictand,  $z$ , given the data,  $y$ , is proportional to the likelihood of both the data ( $L_y$ ) and the predictand ( $L_z$ ) for a maximized parameter vector,  $\theta$  (Caprani & OBrien, 2010).

Mathiasen (1979) appears to be the first to study Fisher's Predictive Likelihood and notes some of its problems. Foremost in this work is the problem that it does not take into account the parameter variability for each of the maximizations of the joint likelihood function required (Bjornstad, 1990,

Lindsey, 1996). Lejeune & Faulkenberry (1982) propose a similar predictive likelihood, but include a normalizing function.

Predictive Likelihood is a general concept and in the literature many versions have been proposed. Cooley and Parke have a number of papers dealing with the prediction issue (Cooley & Parke, 1987, Cooley & Parke, 1990, Cooley et al., 1989). However, their method relies on the assumption that the parameters are normally distributed. Leonard (1982) suggests a similar approach while Davison & Hinkley (1997) use a different form of Predictive Likelihood.

Caprani & OBrien (2010) use the Predictive Likelihood method proposed by Butler (1986), based on that of Fisher (1959) and Mathiasen (1979) and also considered by Bjornstad (1990). Lindsey (1996) describes the reasoning behind its development. This Predictive Likelihood is the Fisherian approach, modified so that the variability of the parameter vector resulting from each maximisation is taken into account.

### 3. Simple Extreme Value Problem

To assess the safety of a bridge, a limited quantity of data is generally used to infer a probability of failure, a characteristic maximum or a statistical distribution of maximum load effects. Probability of failure is clearly the most definitive measure of bridge safety. However, it is strongly influenced by resistance which varies greatly from one example to the next. In order to retain the focus on load effect, the resistance distribution is here assumed to be a mirrored version of the exact LE distribution, shifted sufficiently to the right to give an annual probability of failure of  $10^{-6}$  – see Figure 1.

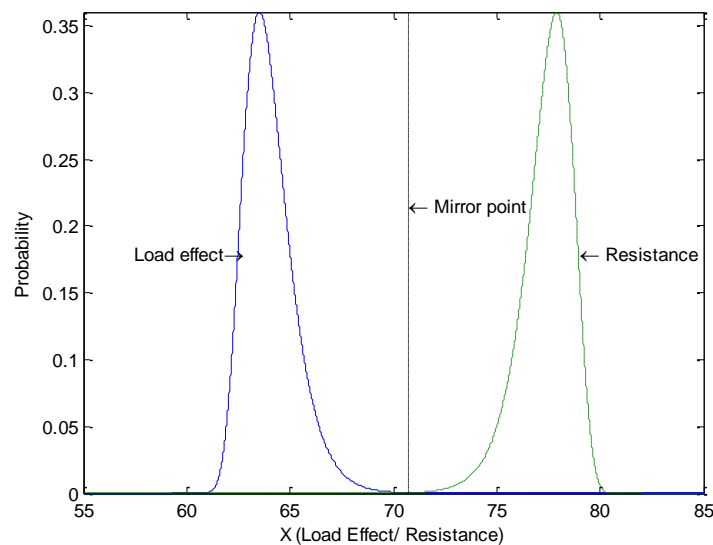


Figure 1- Load Effect and Mirror Resistance

A simple example is used here to compare the alternative methods of extrapolation. A Normally distributed random variable (such as vehicle weight in tonnes) is first considered:

$$Z \sim N(40, 5)$$

Equation 8

Three thousand values of  $Z$  are considered in a given block, say per day, with maximum:

$$X = \max(Z_i) \quad i = 1, 2, \dots, 3000 \quad \text{Equation 9}$$

Typically, a finite number of days of data is available and extreme value distributions are inferred from a dataset of daily maximum values. Hence, a finite number of daily maxima ( $X$  values) may be used to infer, for example, annual maximum distributions. In all cases, the days are considered to be working days and a year is taken to consist of 250 such days.

The exact solution to this problem is readily calculated. The annual maximum can be expressed as:

$$F_Y(z) = [F_X(z)]^{250} = [F_Z(z)]^n \quad \text{Equation 10}$$

where  $n$  is the number of values in a year, equal to  $(250 \times 3000 =) 750\,000$ .

### 3.1 Methods of Inference

Three alternative quantities of daily maximum data are considered: 200, 500 and 1000 working days. A wide range of statistical extrapolation methods are tested in each case to estimate the distribution for annual maximum LE:

- Peaks Over Threshold (POT) data, fitted to the Generalized Pareto distribution;
- Generalized Extreme Value (GEV) fit to tail of daily maximum data;
- Box-Cox fit to tail of daily maximum data;
- Normal distribution fit to tail of daily maximum data;
- Fit of Upcrossing frequency data tail to Rice formula;
- Bayesian fit to all daily maximum data;
- Predictive Likelihood (PL) fit to all daily maximum data.

In each case, the probability distribution of LE is inferred and the theorem of total probability is used with the exact resistance distribution to determine the probability of failure (defined as LE exceeding resistance).

Figure 2 uses Gumbel probability paper to illustrate the first four methods of tail fitting to the CDF's: POT, GEV, Box-Cox and Normal. For all four cases, a least squares fit is found for the top 30% of values from 1000 daily maximum LE's. The exact distribution is shown for comparison. All distributions give good fits, with the Normal being more 'bounded' than the others in this example, i.e., tending more towards an asymptote at extremely low probabilities.

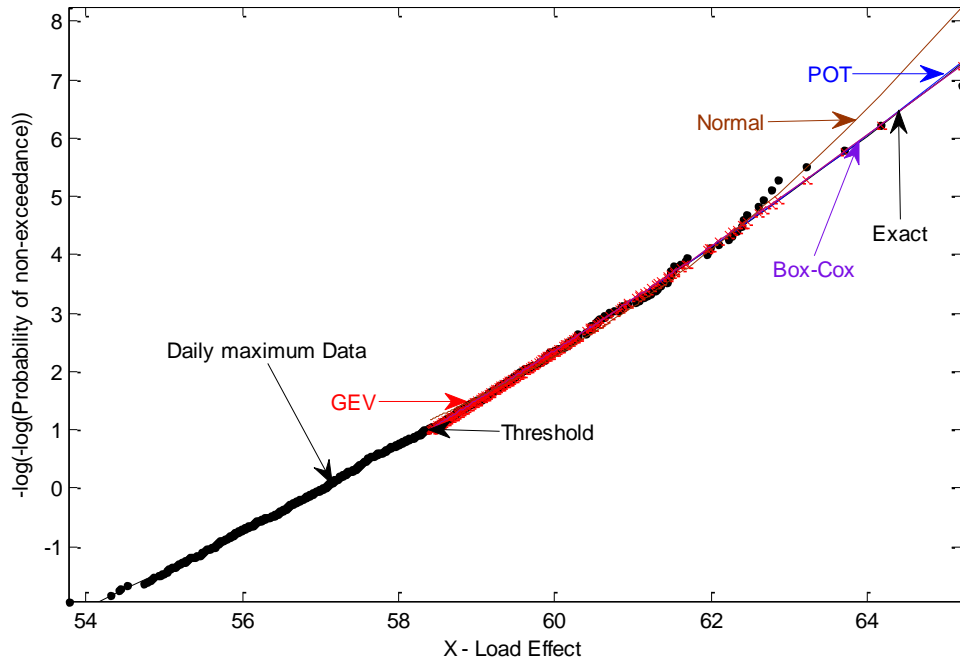


Figure 2 – Best fit distributions using four tail-fitting methods inferred from 1000 days of data

The Rice formula fit is illustrated in Figure 3 which gives the histogram of upcrossings above each threshold, for the same 1000 daily maxima. While Cremona (2001) has considered a variable quantity of data, the top 30% is used here to provide a direct comparison with the other tail fitting methods.

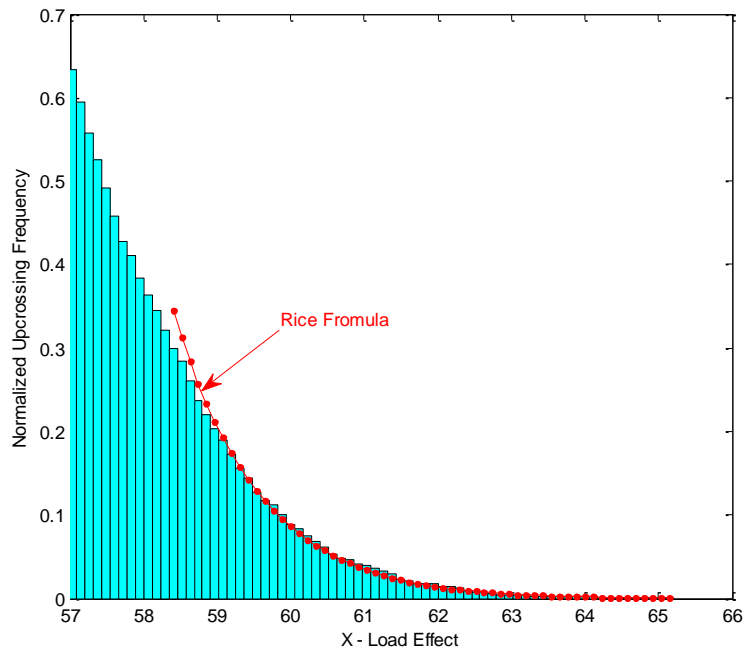


Figure 3 – Rice Formula Fit to Tail for 1000 days of data

Bayesian Updating is the sixth method considered. In this case, unlike the tail fitting methods, all 1000 daily maximum LE's are used. The method is therefore a Bayesian approach applied to block maximum data. The data is assumed to be GEV except that, in this case, a family of GEV

distributions is considered. The GEV parameter values are initially assumed to be equally probable within specified ranges (uniform prior distributions). The daily maximum data is then used to update their probabilities.

The final method applied to this problem, Predictive Likelihood, is also based on the entire dataset of 1000 block maximum values and an assumed GEV distribution. The method is based on the concept of calculating the joint likelihood of a range of possible values at a given level of probability (predictands), given the value of that predictand and the available daily maxima. For example, Figure 4 shows the joint fit to Point A, given the daily maximum data and the joint fit to Point B, given that same set of daily maxima. The likelihood of actually observing Point A is less than that of Point B, given the measurements available. In this way, the joint likelihoods of a wide range of possible predictands are calculated and used to infer a probability distribution for a given time period, such as a year.

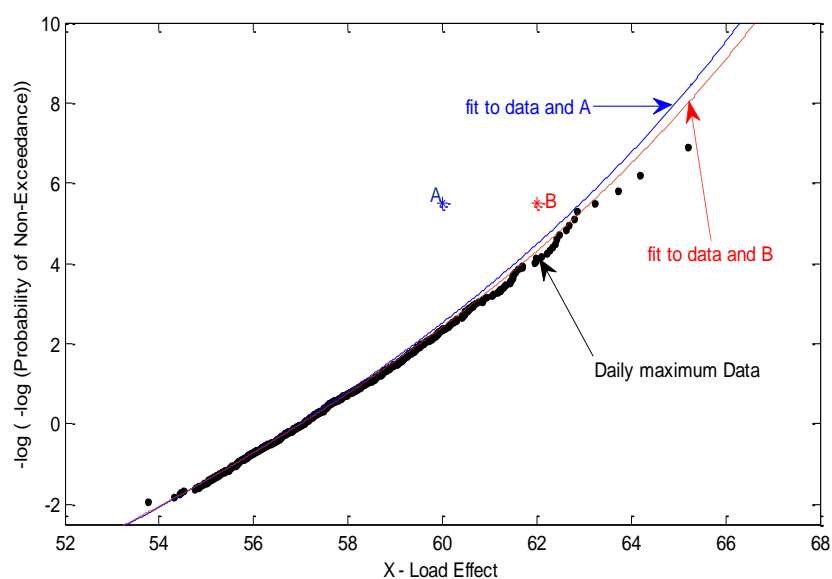


Figure 4- Predictive Likelihood

### 3.2 Inference of Annual Maximum Results from Daily Maximum Data

For the first four tail fitting methods – POT, GEV, Box-Cox and Normal – the parameters of the daily maximum distributions are inferred from the best fits to the top 30% of the daily maximum data, i.e., the block size is one day. Allowing for public holidays and weekends, 250 days are assumed per year. The annual maximum distribution can then be found by raising the CDF for daily maximum to the power of 250.

The Rice formula approach is also a tail fitting method but, in this case, the CDF for annual maximum is found directly from Equation 5. Bayesian Updating and Predictive Likelihood both infer the annual maximum distribution directly as described above.

Figure 5 illustrates the annual maximum CDF's inferred from all 7 approaches, together with the corresponding exact distribution. For this example, most of the tail fitting methods and Predictive Likelihood are more bounded than the exact solution, while Bayesian Updating is less so. The

horizontal line corresponds to a return period of 75 years and it can be seen that all methods except Bayesian Updating, Box-Cox and GEV are slightly non-conservative.

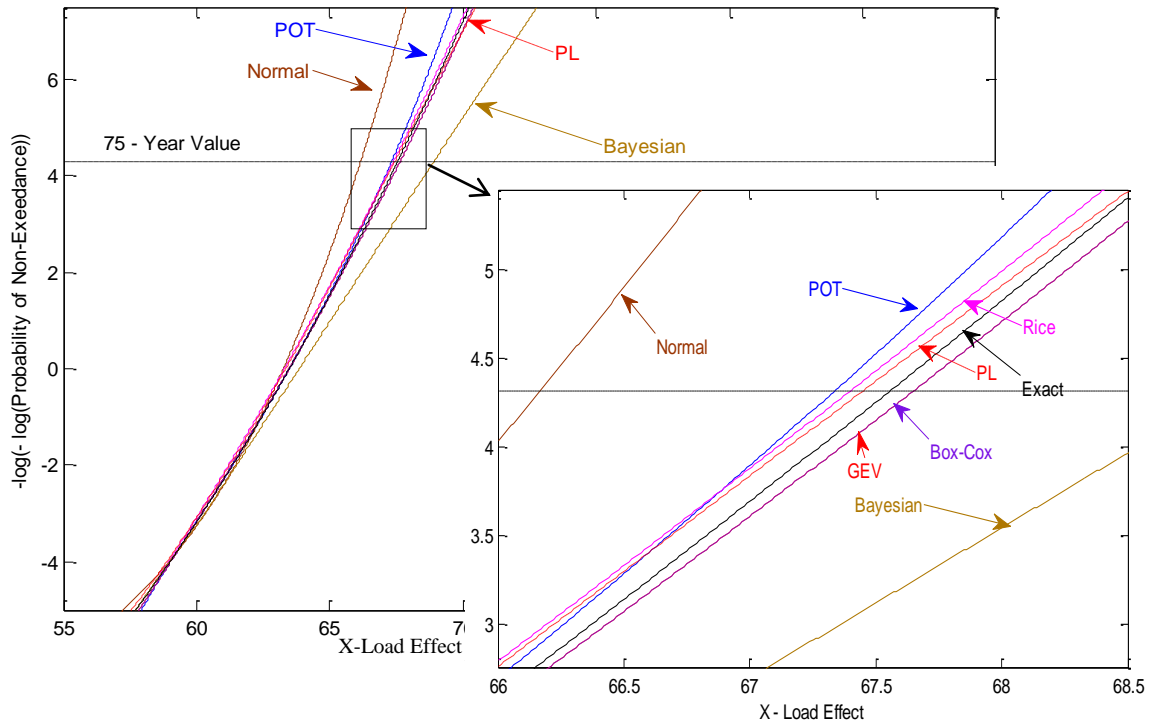


Figure 5 – Inferred Annual Maximum CDF's

Characteristic values are calculated for a 75-year return period. The process is repeated for three different quantities of daily maximum data: 200, 500 and 1000 days. For each of the three quantities, the characteristic values are calculated 20 times so that a measure of the variability in the results can be found. Figure 6 shows the mean of the 20 runs in each case,  $\pm$  one standard deviation.

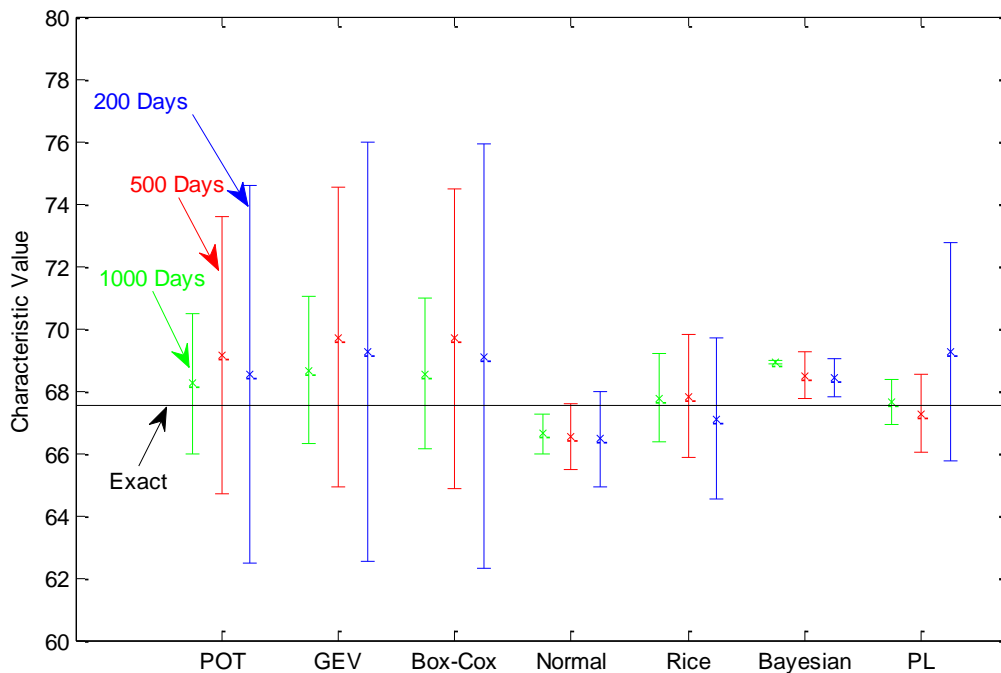


Figure 6 – Inferred 75-year Characteristic Values (mean  $\pm$  one standard deviation)

For 1000 days of data, the results are moderately accurate in most cases, generally falling in the 66 to 71 range. For POT, GEV and Box-Cox, the exact value falls within the error bars and the mean error is less than 1 from that value. Errors in individual results are less good, being as high as 6.1 in one case for GEV. There is no significant difference between these three methods. The Rice formula is relatively good. Results from tail fitting to a Normal distribution do not include the exact value in the error bars. However, the mean error is only 0.94 from the exact value and all the results are reasonably close. Predictive Likelihood is good – the mean is very close to the exact value and the error bars are small. For Bayesian Updating, the error bars are very small – results are highly repeatable – but it is consistently a little conservative for these 20 examples.

Not surprisingly, results are considerably less accurate when fewer days of data are available for inference. With 500 days of data, Normal includes the exact result within its error bars. Bayesian Updating looks better than before with the error bars coming close to the exact solution for both 500 and 200 days of data. Rice is again better than POT, GEV and Box-Cox with a mean very close to the exact and reasonably small error bars. For 200 days of data, PL looks less good than before, with the error bars becoming greater than Normal and Rice.

In order to compare inferred probabilities of failure, the exact annual maximum probability density function is mirrored to give a resistance distribution that implies a failure probability of  $10^{-6}$ . This resistance distribution is then used with each of the inferred distributions to determine the apparent probability of LE exceeding resistance. The calculated probabilities are illustrated in Figure 7.

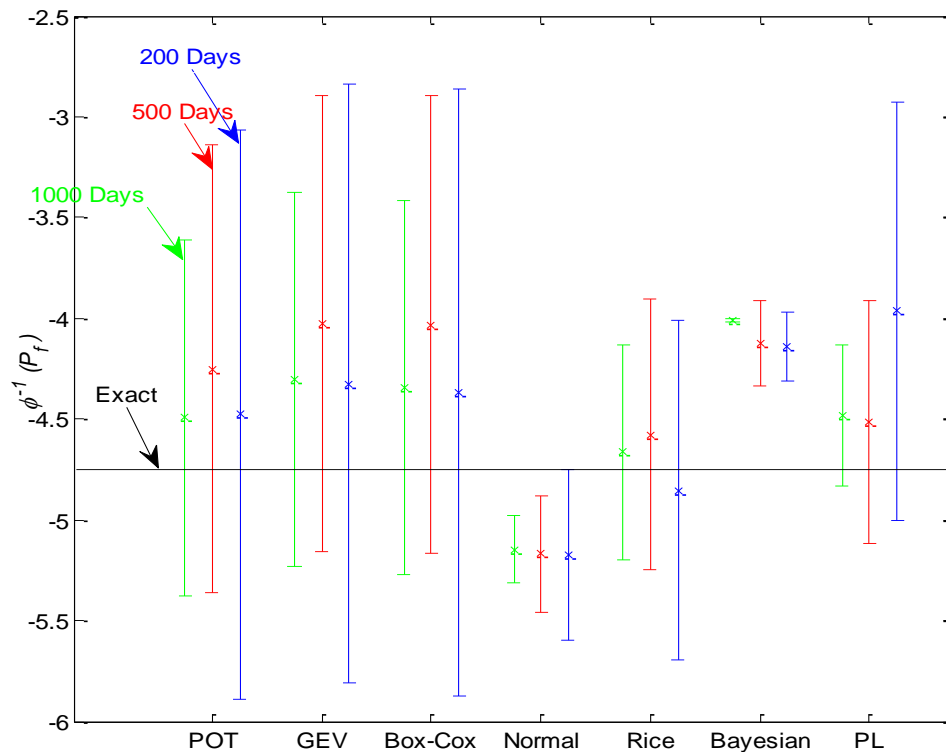


Figure 7 – Mean  $\pm$  One Standard Deviation of Inferred Probabilities of Failure,  $P_f$  plotted to a Normal scale ( $\phi$  = cumulative distribution function for Normal distribution)

Even when plotted on a Normal distribution scale, the probabilities for this example are quite inaccurate. This exercise is analogous to an extrapolation from 200 - 1000 days of data to 1 million years (i.e., annual probability of failure of  $10^{-6}$ ). While the variability in the results is hardly surprising, it has significant implications for any Reliability Theory calculation.

As before, for inference using POT, GEV and Box-Cox, the exact value falls within the error bars. On a scale, the mean error from 1000 days of data is less than about 0.5 from the exact value. Errors in individual results are considerably worse, being as high as 2.1 in the case of one outlier for GEV. The Rice formula is again relatively good, perhaps benefiting from not having an inferred daily maximum distribution raised to the power of 250.

Predictive Likelihood is relatively good and, while results from the Normal distribution do not include the exact value in the error bars, all results are reasonably close to the exact. Bayesian Updating is similar to the results for characteristic value. The error bars are again small and the mean is not near the exact value.

#### 4. Traffic Load Effect Problem

As part of the European 7<sup>th</sup> Framework *ARCHES* project [1], extensive WIM measurements were collected at five European sites: in the Netherlands, Slovakia, the Czech Republic, Slovenia and



Poland. The *ARCHES* site in Slovakia is used as the basis for the simulation model presented here. Measurements were collected at this site for 750 000 trucks over 19 months in 2005 and 2006. The traffic is bidirectional, with average daily truck traffic (ADTT) of 1100 in each direction. Very heavy trucks were recorded at all sites, with a maximum gross vehicle weight (GVW) of 117 t being recorded in Slovakia.

A detailed description of the methodology adopted is given by Enright & OBrien (2012), and is summarised here. For Monte Carlo simulation, it is necessary to use a set of statistical distributions based on observed data for each of the random variables being modelled. For gross vehicle weight and vehicle class (defined here simply by the number of axles), a semi-parametric approach is used as described by OBrien et al. (2010). This involves using a bivariate empirical frequency distribution in the regions where there are sufficient data points. Above a certain GVW threshold value, the tail of a bivariate Normal distribution is fitted to the observed frequencies which allows vehicles to be simulated that may be heavier than, and have more axles than, any measured vehicle. Results for lifetime maximum loading vary to some degree based on decisions made about extrapolation of GVW, and about axle configurations for these extremely heavy vehicles, and these decisions are, of necessity, based on relatively sparse observed data.

Bridge load effects for the spans considered here (Table 1) are very sensitive to wheelbase and axle layout. Within each vehicle class, empirical distributions are used for the maximum axle spacing for each GVW range. Axle spacings other than the maximum are less critical and trimodal Normal distributions are used to select representative values. The proportion of the GVW carried by each individual axle is also simulated in this work using bimodal Normal distributions fitted to the observed data for each axle in each vehicle class. The correlation matrix is calculated for the proportions of the load carried by adjacent and non-adjacent axles for each vehicle class, and this matrix is used in the simulation using the technique described by Iman & Conover (1982).

Traffic flows measured at the site are reproduced in the simulation by fitting Weibull distributions to the daily truck traffic volumes in each direction, and by using hourly flow variations based on the average weekday traffic patterns in each direction. A year's traffic is assumed to consist of 250 weekdays, with the very much lighter weekend and holiday traffic being ignored. This is similar to the approach used by Caprani et al. (2008) and Cooper (1995). For same-lane multi-truck bridge loading events, it is important to accurately model the gaps between trucks, and the method used here is based on that presented by OBrien & Caprani (2005). The observed gap distributions up to 4 seconds are modelled using quadratic curves for different flow rates, and a negative exponential distribution is used for larger gaps.

The modelled traffic is bidirectional, with one lane in each direction, and independent streams of traffic are generated for each direction. In simulation, many millions of loading events are analysed, and for efficiency of computation, it is necessary to use a reasonably simple model for transverse load distribution on two-lane bridges. For bending moment the maximum LE is assumed to occur at the centre of the bridge, with equal contribution laterally from each lane. In the case of shear force at the supports of a simply supported bridge, the maximum occurs when each truck is close to the support, and the lateral distribution is very much less than for mid-span bending moment. In this case a reduction factor of 0.45 is applied to the axle weights in the second lane. This factor is based on finite element analyses performed for different types of bridge (OBrien & Enright, 2012). The load effects and bridge lengths examined in the simulation runs are summarized in Table 1.

Table 1 – Load effects and bridge lengths

	Load Effect	Bridge Lengths (m)
LE1	Mid-span bending moment, simply supported bridge	15, 35
LE2	Shear force at start/end of a simply supported bridge	15, 35
LE3	Central support hogging moment, 2-span continuous bridge	35

Two series of simulation runs are performed – one to represent possible measurements over 1000 days, repeated 20 times, and another to represent the benchmark (‘exact’) results, consisting of 5000 years of traffic. For the benchmark run, the outputs consist of annual maximum LE’s, and these can be used to calculate the characteristic values and annual maximum distributions to a high degree of accuracy.

Sample results are plotted on Gumbel probability paper in Figure 8 for the 5000-year simulation run. Two load effects are shown – shear force (LE2) on a simply supported 15 m bridge, and hogging moment (LE3) over the central support of a two-span bridge of total length 35 m. Due to the randomness inherent in the process, there is some variability in the results, particularly in the upper tail region (top 1% of data approximately). Weibull fits to the upper 30% tail are used to smooth this variability (as shown in figure), and these are used to calculate the characteristic values. This long-run simulation process is considered to be highly accurate, subject to the assumptions inherent in the model and is used as the benchmark against which the accuracy of all other methods is measured.

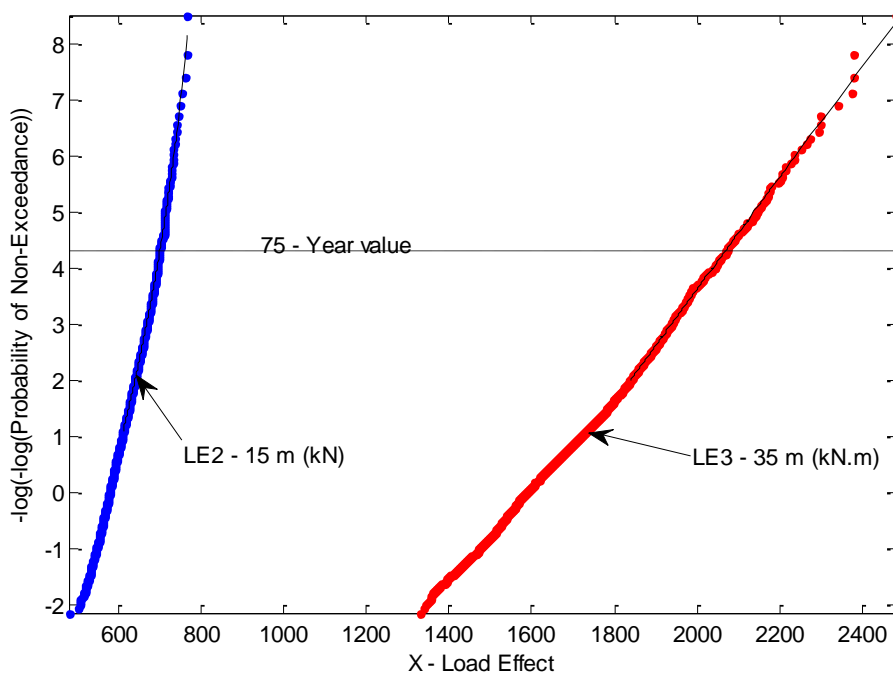


Figure 8 – Annual maximum LE’s from 5000 years of simulation

#### 4.1 Results of Inference Based on Long-run Simulation Data

The assumed measurements consisting of 1000 simulated daily maxima are used as the basis for extrapolation using each method to estimate the ‘true’ results calculated from the long-run simulation. For the five tail fitting methods, the distributions are fitted to the top 30% of data. For some load effects and spans, the distribution of the data is multi-modal (see Figure 9), i.e., there is a change in slope – around 400 kN in this case – implying data from a different parent distribution. In the case illustrated, there is a change around this point from (i) daily maxima arising from regular trucks and cranes to (ii) maxima arising from extremely heavy and rare low-loader vehicles.

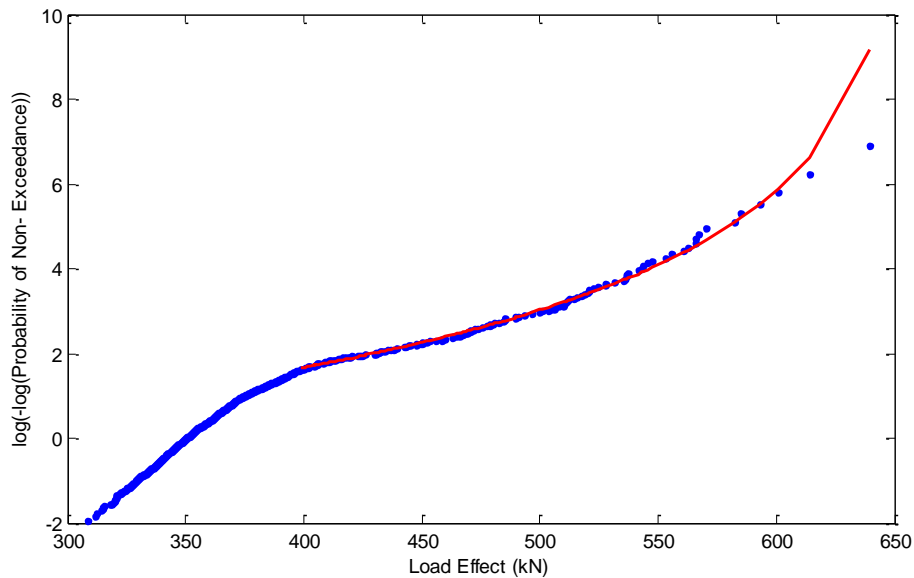


Figure 9 – Daily Maximum LE2 - 15 m span Data and Inferred Distributions

The 75-year characteristic maximum LE’s are inferred from the assumed measurements. This process is carried out for the 5 load effects and repeated 20 times to determine the variability in results. The results are illustrated in Figure 10 which shows, in each case (i) the median value, (ii) the 25% to 75% range (boxed), (iii) the 0.7% to 99.3% range (median  $\pm$  2.7 standard for normally distributed data) (dashed lines) and (iv) individual outliers beyond that range.

Figure 10 shows that the first three tail fitting methods are reasonably good, with modest range and median value close to the benchmark result from the 5000 year run. As for the simple example, fitting to a Normal distribution gives a lesser range of results which, in this case, are all reasonably close to the benchmark. The Rice method is generally better than all the others.

Predictive Likelihood gives poor results for these traffic loading problems. Characteristic values are sometimes under-estimated and other times over-estimated, with no clear trend. Sensitivity studies of these results show that there is significant influence of the Fréchet (unlimited) tail of the GEV distribution that is used to jointly maximize the likelihood of observing the data and the predictand. If the fit is limited so that Fréchet tails are not permitted (admitting only Weibull or Gumbel tails), as could be argued from the physical bounds of the traffic loading phenomenon, then the fits improve.

Bayesian Updating is surprisingly poor for this class of problem. As for the simpler example, the 20 results are consistent but deviate significantly from the benchmark results. Several variations were

tested in attempts to find a Bayesian approach that is consistently good. The GEV distribution was fitted to the top 30% of data, as an alternative to fitting it to all the data. Different numbers of parameters of the GEV distribution were updated: two ( $\sigma$  and  $\xi$ ) and three ( $\mu$ ,  $\sigma$  and  $\xi$ ). Different prior distributions were assumed for these parameters – Normal and Uniform. For the latter, different ranges were tested for the parameter values. None of these variations produced consistently better results for the five LE's and spans. The results shown are based on use of all the data; updating just two parameters ( $\sigma$  and  $\xi$ ) with a Uniform prior distribution and a limit on the range of  $\xi$  to be non-positive.

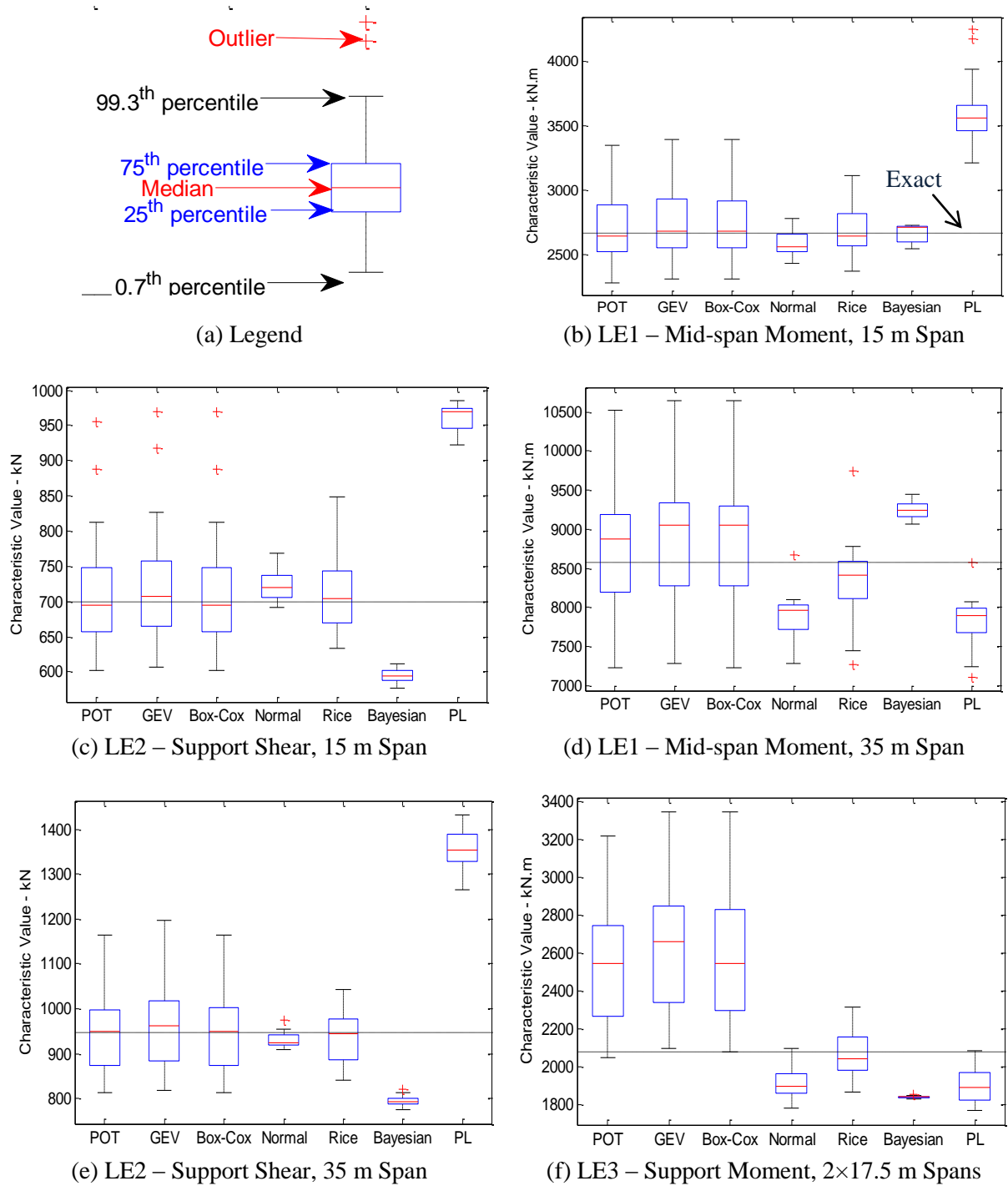
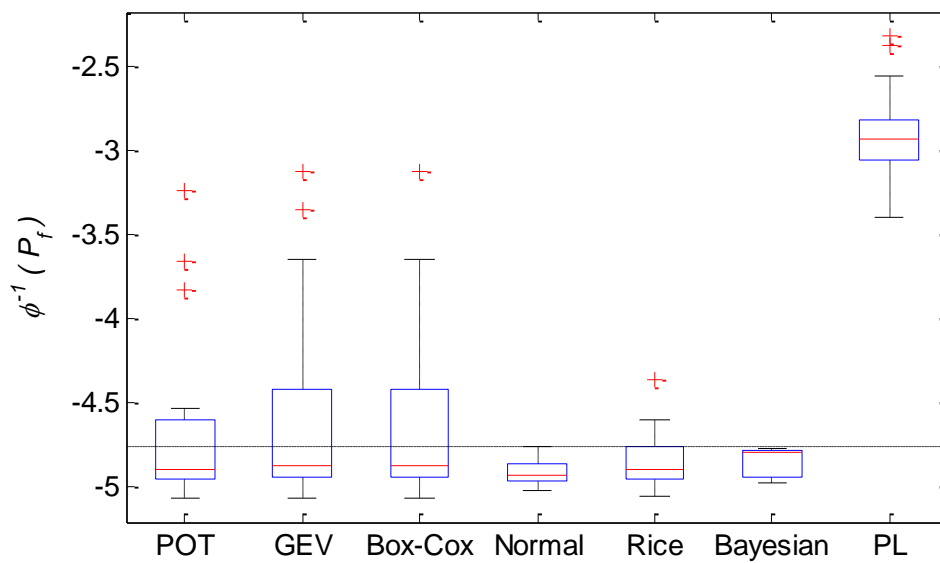


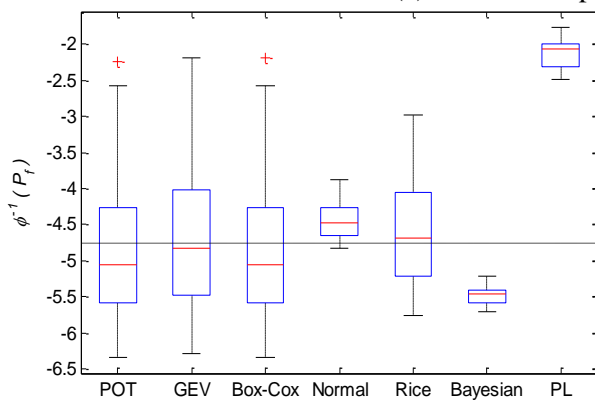
Figure 10 – Range of Inferred 75-year Characteristic Values from 1000 Days of Data

Annual probabilities of failure are also inferred for the five combinations of load effect and span. As before, the probability of failure for the benchmark example is set at  $10^{-6}$  in each case and the resistance distribution is taken to be a mirrored version of the benchmark LE distribution.

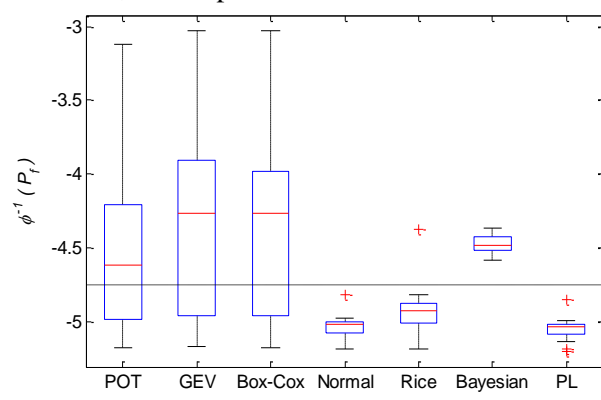
The results are illustrated in Figure 11. As for the simple example, the errors in the probabilities, even when plotted on a Normal scale, are much higher than for characteristic values. Most of the tail fitting methods – POT, GEV, Box-Cox and Rice formula – give relatively good results, with the Rice formula generally beating the others. As before, when fitting to a Normal distribution, the benchmark result is sometimes outside the 25%-75% range, but not by a great deal. As for the characteristic values, Bayesian Updating and Predictive Likelihood are less accurate than the other methods.



(a) LE1- Mid-span Moment, 15 m Span



(b) LE2 – Support Shear, 15 m Span



(c) LE1 – Mid-span Moment, 35 m Span

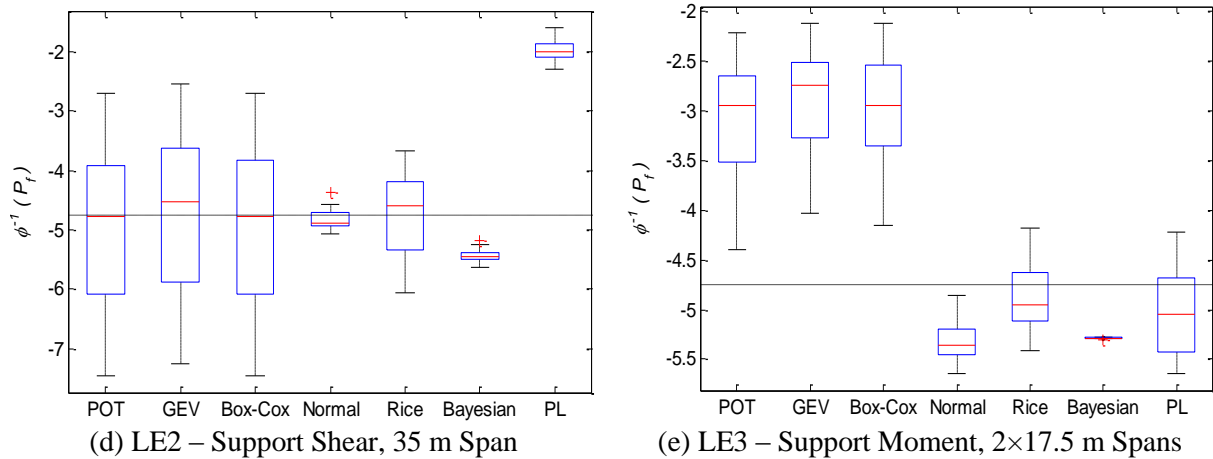


Figure 11 – Probabilities of Failure Inferred from Sets of 1000 Daily Maximum LE's

## 5. Conclusions

In this paper, seven methods of statistical inference are critically reviewed. Each method is also tested using two examples. The first example is derived from a Normal distribution and the exact solution is known. A total of 3000 normally distributed values (e.g., vehicle weights) are considered per day and the daily maxima are used to infer the characteristic maximum and the probability of failure in a year. In the second example, a sophisticated algorithm is used to generate a train of vehicles with weights and axle configurations consistent with measured Weigh-in-Motion data. Five different combinations of load effect and span are considered and, in each case, characteristic values and probabilities of failure are again calculated. In these cases, the exact solutions are not known but the simulation is run for 5000 years to obtain accurate benchmark references against which inferences based on 1000 days of data can be compared.

Of the seven methods considered, five are tail-fitting approaches, i.e., a distribution is fitted to the tail of the data. Peaks-Over-Threshold (POT) is popular in some sectors but is not time-referenced and selecting the threshold is a subjective process. Fitting the tail of block-maximum data to a Generalized Extreme Value (GEV) is perhaps the most popular used for bridge traffic loading, with a typical block size of a day. Box-Cox could be considered to be a hybrid between POT and GEV. These three methods are generally good for inferring the characteristic values, both for the simple and the more complex examples. There is no theoretical justification for fitting block maximum data to the tail of a Normal distribution but it is sometimes done. It is found here to give reasonably accurate results, with a small standard deviation. Finally, the Rice formula is an indirect approach as it is the upcrossing frequencies that are fitted to the formula, rather than the data itself. Nevertheless, it performs well in these tests, generally better than POT and GEV.

Bayesian Updating is used here to fit the block maximum data to a family of GEV distributions. The parameters of the GEV are allowed to vary, their associated probabilities being updated as the data is considered. The results from the Bayesian approach are found to be generally poor and several variations in the approach did not produce consistently good results. Finally, Predictive Likelihood is considered, a method where the likelihood of each inferred characteristic value is considered, given the available data. This is also found to give poor results.

All seven methods are used to infer the annual probabilities of failure as well as the characteristic values. To avoid the need for any assumption on the distributions for resistance, the benchmark load effect distribution is mirrored and this mirrored version is used in the calculation of probability of failure.

The inferred failure probabilities are considerably less accurate than the inferred characteristic values, perhaps not surprising given that such a small failure probability was being considered ( $10^{-6}$  in a year). As for characteristic values, the tail fitting methods are better than the others but none of the methods gives an accurate inference with 1000 days of data.

## Acknowledgment

This work is supported under the *TEAM* project. *TEAM* is a Marie Curie Initial Training Network and is funded by the European Commission 7<sup>th</sup> Framework Programme (PITN-GA-209-238648).

## REFERENCES

- Aitchison, J. & Dunsmore, I. R. Statistical prediction analysis. Great Britain at University Press, Cambridge: Syndicate of the Cambridge University Press, 1980
- Al-Harthy, A. S., Stewart, M. G. & Mullard, J. Concrete cover cracking caused by steel reinforcement corrosion. Magazine of Concrete Research, 2011, 63(9): 655-667.
- Bailey, S. F. Basic principles and load models for the structural safety evaluation of existing road bridges. Doctorate of Philosophy, Polytechnique Fédérale de Lausanne, 1996
- Bailey, S. F. & Bez, R. A parametric study of traffic load effects in medium span bridges. Proceedings of the Fourth International Conference on Short and Medium Span Bridge Engineering, , Halifax, Canada, Citeseer, 1994.
- Bailey, S. F. & Bez, R. Site specific probability distribution of extreme traffic action effects. Probabilistic engineering mechanics, 1999, 14(1): 19-26.
- Bali, T. G. The generalized extreme value distribution. Economics letters, 2003, 79(3): 423-427.
- Bali, T. G. & Theodossiou, P. Risk measurement performance of alternative distribution functions. Journal of Risk and Insurance, 2008, 75(2): 411-437.
- Basu, A. Estimates of reliability for some distributions useful in life testing. Technometrics, 1964, 6(9): 215-219.
- Bermudez, P. D. Z. & Kotz, S. Parameter estimation of the generalized pareto distribution-part ii. Journal of Statistical Planning and Inference, 2010, 140(6): 1374-1388.
- Bhattacharya, S. K. Bayesian approach to life testing and reliability estimation. Journal of the American Statistical Association, 1967, 62(317): 48-62.
- Bjornstad, J. F. Predictive likelihood: A review. Statistical Science, 1990, 5(2): 242-254.
- Box, G. E. P. & Cox, D. R. An analysis of transformations. Journal of the Royal Statistical Society. Series B (Methodological), 1964, 26(2): 211-252.
- Buckland, P. G., McBryde, J. P., Zidek, J. V. & Navin, F. P. D. Proposed vehicle loading of long-span bridges. Journal of the Structural Division, 1980, 106(4): 915-932.
- Butler, R. W. Predictive likelihood inference with applications. Journal of the Royal Statistical Society. Series B (Methodological), 1986, 48(1): 1-38.
- Caprani, C., McNally, C., O'Brien, E. & Bordallo-Ruiz, A. The structural reliability of bridges subject to time-dependent deterioration. Proceedings of the 11th. Conference on Civil, Structural and Environmental Engineering Computing, Stirling, Scotland, Civil-Comp Press, 2007.
- Caprani, C. C. Probabilistic analysis of highway bridge traffic loading. Doctorate of Philosophy, University College Dublin, 2005

- Caprani, C. C. & O'Brien, E. J. Statistical computation for extreme bridge traffic load effects. Proceedings of the 8th. International Conference on Computational Structures Technology, Stirling, Scotland, Civil-Comp Press, 2006.
- Caprani, C. C. & O'Brien, E. J. Estimating extreme highway bridge traffic load effects. International Conference on Structural Safety and Reliability, ICOSSAR '09. Osaka, Japan. 2009.
- Caprani, C. C. & O'Brien, E. J. The use of predictive likelihood to estimate the distribution of extreme bridge traffic load effect. Structural safety, 2010, 32(2): 138-144.
- Caprani, C. C., O'Brien, E. J. & McLachlan, G. J. Characteristic traffic load effects from a mixture of loading events on short to medium span bridges. Structural safety, 2008, 30(5): 394-404.
- Castillo, E. Extreme value theory in engineering (statistical modeling and decision science). Boston: Academic Press, 1988
- Coles, S. An introduction to statistical modeling of extreme values. Great Britain: Springer Verlag, 2001
- Coles, S. & Tawn, J. Modelling extremes: A bayesian approach. Applied Statistics, 1996, 45: 463–478.
- Cooley, T. F. & Parke, W. R. Likelihood and other approaches to prediction in dynamic models. Journal of econometrics, 1987, 35(1): 119-142.
- Cooley, T. F. & Parke, W. R. Asymptotic likelihood-based prediction functions. Econometrica: Journal of the Econometric Society, 1990, 58(5): 1215-1234.
- Cooley, T. F., Parke, W. R. & Chib, S. Predictive efficiency for simple non-linear models. Journal of econometrics, 1989, 40(1): 33-44.
- Cooper, D. The determination of highway bridge design loading in the united kingdom from traffic measurements. First European Conference on Weigh-In-Motion of Road Vehicles, Zurich, Switzerland, 1995.
- Cooper, D. Development of short span bridge-specific assessment live loading. London: Safety of Bridges, Thomas Telford, 1997
- Cramér, H. Mathematical methods of statistics. Princeton: Princeton University Press, 1946
- Cremona, C. Optimal extrapolation of traffic load effects. journal of Structural Saftey, 2001, 23(16): 31-46.
- Crespo-Minguillón, C. & Casas, J. R. A comprehensive traffic load model for bridge safety checking. Structural safety, 1997, 19(4): 339-359.
- Davison, A. C. & Hinkley, D. V. Bootstrap methods and their application. New York, USA: Cambridge University Press, 1997
- DeGroot, M. H. Probability and statistics. Boston, MA, USA, Addison-Wesley Publishing. 1986: 816.
- Ditlevsen, O. Traffic loads on large bridges modeled as white-noise fields. Journal of engineering mechanics, 1994, 120(23): 681-694.
- Eastoe, E. F. & Tawn, J. A. Modelling non stationary extremes with application to surface level ozone. Journal of the Royal Statistical Society: Series C (Applied Statistics), 2009, 58(1): 25-45.
- Enright, B. Simulation of traffic loading on highway bridges. Doctorate of Philosophy, University College Dublin, 2010
- Enright, B. & O'Brien, E. J. Monte carlo simulation of extreme traffic loading on short and medium span bridges. Structure and Infrastructure Engineering, 2012: 1-16.
- Fisher, R. A. & Tippett, L. H. C. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Mathematical Proceedings of the Cambridge Philosophical Society, 1928, 24(02): 180-190.
- Fisher, S. R. A. Statistical methods and scientific inference. Indiana University, Indiana, USA: Oliver and Boyd, 1959
- Flint, A. R. & Jacob, B. Extreme traffic loads on road bridges and target values of their effects for code calibration. Proceedings of IABSE Colloquium, IABSE-AIPC-IVBH, Delft, The Netherlands, 1996.
- Frangopol, D. M. & Liu, M. Maintenance and management of civil infrastructure based on condition, safety, optimization, and life-cycle cost\*. Structure and Infrastructure Engineering, 2007, 3(1): 29-41.



- Getachew, A. Traffic load effects on bridges. Doctorate of philosophy, Royal Institute of Technology, 2003
- Getachew, A. a. O., E.J. Simplified site specific models for determination of characteristic traffic load effects for bridges. 4th International Conference on Weigh-In-Motion – ICWIM4. E. J. O. e. al. National Taiwan University. 2005: 341-350.
- Ghosn, M. & Moses, F. Markov renewal model for maximum bridge loading. Journal of engineering mechanics, 1985, 111(9): 1093-1104.
- Ghosn, M., Moses, F. & Wang, J. Design of highway bridges for extreme events Transportation Research Board National Research, 2003
- Gindy, M. & Nassif, H. H. Comparison of traffic load models based on simulation and measured data. Joint International Conference on Coputing and Decision Making in Civil and Building Engineering, Montreal, Canada, 2006.
- Gnedenko, B. V. Sur la distribution limite du terme maximum d'une serie aleatoire. The Annals of Mathematics, 1943, 44(3): 423-453.
- González, A., O'Connor, A. J. & O'Brien, E. J. An assessment of the influence of dynamic interaction modelling on predicted characteristics load effects in bridges. Proceedings of the 3rd International Conference on Current and Future Trends in Bridge Design, Construction and Maintenance, Shanghai, China, Thomas Telford, 2003.
- Grave, S. Modelling of site-specific traffic loading on short to medium span bridges. Doctorate of Philosophy, Trinity College Dublin, 2001
- Gumbel, E. Les valeurs extrêmes des distributions statistiques. Ann. Inst. H. Poincaré, 1935, 5(2): 115-158.
- Gumbel, E. J. Statistics of extremes. Mineola, New York, USA: Columbia University Press, 1958
- Holla, M. Bayesian estimates of the reliability function. Australian Journal of Statistics, 1966, 8(1): 32-35.
- Iman, R. L. & Conover, W. A distribution-free approach to inducing rank correlation among input variables. Communications in Statistics-Simulation and Computation, 1982, 11(3): 311-334.
- Jacob, B. Methods for the prediction of extreme vehicular loads and load effects on bridges. Report of Subgroup 8, Eurocode 1. Traffic Loads on Bridges Paris, France, LCPC. 1991, 8.
- James, G. Analysis of traffic load effects on railway bridges. Doctorate of Philosophy, Royal Institute of Technology, 2003
- Jenkinson, A. F. The frequency distribution of the annual maximum (or minimum) values of meteorological elements. Quarterly Journal of the Royal Meteorological Society, 1955, 81(348): 158-171.
- Kanda, J. & Ellingwood, B. Formulation of load factors based on optimum reliability. Structural safety, 1991, 9(3): 197-210.
- Kulicki, J. M., Prucz, Z., Clancy, C. M., Mertz, D. R. & Nowak, A. S. Updating the calibration report for aashto lfrd code. Final Rep. for National Cooperative Highway Research Program (NCHRP). Washington DC, USA, AASHTO. 2007, 186: 20-27.
- Leadbetter, M. R., Lindgren, G. & Rootzen, H. Extremes and related properties of random sequences and processes. The University of California, California, USA: Springer-Verlag, 1983
- Lejeune, M. & Faulkenberry, G. D. A simple predictive density function. Journal of the American Statistical Association, 1982, 77(379): 654-657.
- Leonard, T. Comment on "a simple predictive density function". Journal of the American Statistical Association, 1982, 77(379): 657-658.
- Lindsey, J. K. Parametric statistical inference. Oxford University Press, USA: Oxford Science Publications, 1996
- Lingappaiah, G. Bayesian prediction regions for the extreme order statistics. Biometrical journal, 1984, 26(1): 49-56.
- Mathiasen, P. E. Prediction functions. Scandinavian Journal of Statistics, 1979, 6(1): 1-21.
- Melchers, R. E. Structural reliability: Analysis and prediction. New York, USA: John Wiley & Sons New York, 1999
- Moses, F. Calibration of load factors for lfr bridge evaluation. Report of National Cooperative Highway Research Program (NCHRP)

- Washington, DC, USA, National Academy Press. 2001, 454.
- Moyo, P., Brownjohn, J. & Omenzetter, P. Highway bridge live loading assessment and load carrying estimation using a health monitoring system. Proceedings of the first European workshop on structural health monitoring, Cachan (Paris), France, 2002.
- Nowak, A. S. Live load model for highway bridges. *Structural safety*, 1993, 13(1-2): 53-66.
- Nowak, A. S. Load model for bridge design code. *Canadian Journal of Civil Engineering*, 1994, 21(1): 36-49.
- Nowak, A. S. Reply: Load model for bridge design code. *Canadian Journal of Civil Engineering*, 1995, 22(2): 293-293.
- Nowak, A. S. Calibration of lrfd bridge design code. Report of the National Cooperative Highway Research Program (NCHRP). Washington, DC, USA, Transportation Research Board, National Research Council; National Academy Press. 1999, 368.
- Nowak, A. S. & Hong, Y. K. Bridge live-load models. *Journal of Structural Engineering*, 1991, 117(9): 2757-2767.
- Nowak, A. S., Nassif, H. & DeFrain, L. Effect of truck loads on bridges. *Journal of transportation engineering*, 1993, 119: 853-867.
- O'Connor, A. J. Probabilistic traffic load modelling for highway bridges. Doctorate of Philosophy, Trinity College Dublin, 2001
- O'Connor, A. J. & OBrien, E. J. Traffic load modelling and factors influencing the accuracy of predicted extremes. *Canadian Journal of Civil Engineering*, 2005, 32(1): 270-278.
- OBrien, E. J., Caprani, C., Žnidari, A. & Quilligan, M. Site-specific probabilistic bridge load assessment. Proceedings of the 3rd International Conference on Current and Future Trends in Bridge Design, Construction and Maintenance, Shanghai, China., Thomas Telford, 2003.
- OBrien, E. J. & Caprani, C. C. Headway modelling for traffic load assessment of short to medium span bridges. *Structural Engineering*, 2005, 83(16): 33-36.
- OBrien, E. J. & Enright, B. Using weigh-in-motion data to determine aggressiveness of traffic for bridge loading. *Journal of Bridge Engineering*, 2012, 1: 258.
- OBrien, E. J., Enright, B. & Getachew, A. Importance of the tail in truck weight modeling for bridge assessment. *Journal of Bridge Engineering*, 2010, 15: 210.
- Orcesi, A. D. & Cremona, C. F. Optimization of management strategies applied to the national reinforced concrete bridge stock in france. *Structure and Infrastructure Engineering*, 2009, 5(5): 355-366.
- Orcesi, A. D. & Cremona, C. F. A bridge network maintenance framework for pareto optimization of stakeholders/users costs. *Reliability Engineering & System Safety*, 2010, 95(11): 1230-1243.
- Rice, S. The mathematical analysis of random noise. *Bell Syst Tech journal*, 1945, 24: 24-156.
- Richard, B., Adelaide, L., Cremona, C. & Orcesi, A. A methodology for robust updating of nonlinear structural models. *Engineering structures*, 2012, 41: 356-372.
- Rocco, M. Extreme value theory for finance: A survey. Statistics and Informatics Working Paper No. 3. University of Bergamo, Italy, Department of Mathematics. 2010.
- Rücker, W., Hille, F. & Rohrman, R. Guideline for the assessment of existing structures. Proceedings of the 5th International Conference on Bridge Management, Surrey, UK, Thomas Telford, London, 2005.
- Sheils, E., O'Connor, A. J., Schoefs, F. & Breysse, D. Investigation of the effect of the quality of inspection techniques on the optimal inspection interval for structures. *Structure and Infrastructure Engineering*, 2010, 8(6): 557-568.
- Singpurewalla, N. & Song, M. S. Reliability analysis using weibull lifetime data and expert opinion. *Reliability*, IEEE Transactions on, 1988, 37(3): 340-347.
- Sinha, S. & Sloan, J. Bayes estimation of the parameters and reliability function of the 3-parameter weibull distribution. *Reliability*, IEEE Transactions on, 1988, 37(4): 364-369.
- Sivakumar, B., Ghosn, M. & Moses, F. Protocols for collecting and using traffic data in bridge design. Report of National Cooperative Highway Research Program. Washington, DC. USA, Transportation Research Board. 2011, 683.
- Smith, R. L. & Naylor, J. A comparison of maximum likelihood and bayesian estimators for the three-parameter weibull distribution. *Applied Statistics*, 1987, 36(3): 358-369.

- Suo, Q. & Stewart, M. G. Corrosion cracking prediction updating of deteriorating rc structures using inspection information. *Reliability Engineering & System Safety*, 2009, 94(8): 1340-1348.
- Tötterman, P. Applying extreme value theory and tail risk measures to reduce portfolio losses. Masters Hanken School of Economics, 2010
- Von Mises, R. La distribution de la plus grande de n valeurs. *Rev. math. Union interbalcanique*, 1936, 1(1).
- Wadsworth, J., Tawn, J. & Jonathan, P. Accounting for choice of measurement scale in extreme value modeling. *The Annals of Applied Statistics*, 2010, 4(3): 1558-1578.

[1] [http://arches.fehrl.org/index.php?m=7&id\\_directory=1611](http://arches.fehrl.org/index.php?m=7&id_directory=1611)