2006

Statistical Computation for Extreme Bridge Traffic Load Effects

Colin C. Caprani
Dublin Institute of Technology, colin.caprani@dit.ie

E. J. O'Brien
University College Dublin

2006-01-01

Statistical computation for extreme bridge traffic load effects

Colin C. Caprani
Dublin Institute of Technology, colin.caprani@dit.ie

E.J. O'Brien
University College Dublin

Paper 139

Proceedings of the Eighth International Conference
on Computational Structures Technology,
B.H.V. Topping, G. Montero and R. Montenegro,
(Editors), Civil-Comp Press, Stirlingshire, Scotland.

Statistical Computation for Extreme Bridge Traffic Load Effects

C.C. Caprani and E.J. OBrien
School of Architecture, Landscape and Civil Engineering
University College Dublin, Ireland

Abstract

The maintenance of highway infrastructure constitutes a major expenditure in many countries. This cost can be reduced significantly by minimizing the repair or replacement of highway bridges. In the assessment of existing bridges, the strength estimate tends to be more accurate than that of traffic loading, due to the more variable nature of loading. Recent advances in the statistical analysis of highway bridge traffic loading have resulted in more accurate forecasts of the actual loading to which a bridge is subject. While these advances require extensive numerical computation, they can significantly improve the accuracy of the calculation. This paper outlines the recent advances and describes the associated computational aspects in detail.

Keywords: bridge, statistics, loading, simulation, predictive likelihood, traffic.

1 Introduction

As bridge stocks age across Europe and the world, maintenance costs represent an increasing proportion of total road infrastructure expenditure. A recent study [1] estimates the EU expenditure on the repair, rehabilitation and maintenance of bridge structures to be €4–6 bn annually. This figure only includes that of the 15 member states up to May 2004. Therefore, in the recently enlarged EU the annual spend on bridge maintenance is likely to be at least €6 bn.

The significant cost attributable to the maintenance of highway bridge infrastructure has driven much research in this area over the past number of years. The assessment of existing bridges represents an area in which significant cost savings are possible. Whilst the assessment of the capacity of an existing bridge is not yet highly accurate, the estimate of actual loading is significantly more variable.

It is therefore in the area of highway bridge traffic loading that much progress towards reducing the maintenance expenditure may be made.

In recent years statistical methods have been increasingly employed to assist in the estimation of the lifetime traffic loading to which a bridge is subject [2]–[7]. This paper describes the recent statistical methods used, and their computational aspects, to determine the lifetime traffic load effect. A new approach, termed predictive likelihood, is also described which calculates the design lifetime distribution of load effect directly from simulations of measured traffic data. This lifetime distribution of load effect is then used to determine the design value. The Eurocode for traffic loading on bridges [8] defines the (design) characteristic value to be that value which has probability of exceedance of 10% in a 100 year design life. The associated return period is 1000 years. Based upon a fitted distribution, it is usual to extrapolate to this return period, giving a single characteristic value. Instead, predictive likelihood is used to directly estimate the load effect distribution for the 100 year design life from which the 90-percentile may be taken. In general, the two approaches give differing values of design load effect. As predictive likelihood returns more information from the sample, this is considered to be a more accurate result.

In conclusion, the latest statistical models to be applied to the bridge traffic load problem are described and extended. The new approach is shown to result in an increase in information from the data, and computational methods are integral to this. Further, the numerical aspects of the problem are described and the solutions discussed. It is concluded that great improvements in the accuracy of bridge traffic loading are obtained through the use of statistical computational methods.

2 Bridge Traffic Load Simulation

2.1 Measurement and Modelling

The highway traffic data, essential to the bridge traffic load modelling process, must be obtained from suitable installations. The measurements taken must enable headway, speed and other such pertinent data to be measured. WIM technology has been developed to determine vehicle weights but also meets these requirements. This work is based on data taken from the A6 motorway near Auxerre, France. The site has 4 lanes of traffic (2 in each direction) but only the traffic recorded in the slow lanes was used and it is acknowledged that this results in conservative loading for a 2-lane bridge. In total 17 756 and 18 617 trucks were measured in the north and south slow lanes respectively, giving an average daily truck flow of 6744 trucks. This represents one week of traffic data which, it is acknowledged, is short in duration. However, the methodologies presented in this work are general and only the absolute quantitative results are affected by this short duration of measurements.

The traffic model required to simulate bridge load effects must be consistent with the measured traffic at the site it claims to represent. Yet, it is important that there is the potential for variation from the measured traffic in the model; otherwise the model would only represent multiple sets of the same traffic. By using parametric statistical distributions, the traffic model may remain sympathetic to the measurements, yet retain the capacity to differ. The recorded WIM data was analysed for the statistical distributions of the traffic characteristics of the site for each lane as follows:

- Gross Vehicle Weight: Modelled as tri-modal mixture of Normal distributions;
- Axle spacings: Modelled as uni- or bi-modal Normal distributions, as appropriate to the data;
- Axle weights for 2- and 3-axle trucks: Modelled as tri- or bi-modal Normal distributions, as appropriate to the data;
- Axle weights for 4- and 5-axle trucks: Axle weight expressed as a percentage of Gross Vehicle Weight (GVW) for the first and second axles and for the remaining tandem group. In each case, the percentage is modelled as a Normal distribution
- Composition: The measured percentage of 2-, 3-, 4- and 5-axle trucks is used;
- Speed: Modelled as a Normal distribution and considered independent of truck type and uncorrelated with GVW;
- Flow rates: For each hour of the day, the average flow rate (ignoring weekend days) was used for all the days available;
- Headway: This is modelled with a number of distributions dependent on flow, as described in [9].

Both short-term variations, such as variations from hour to hour, and long-term variations, such as an annual increase in traffic volume, exist in traffic. The major consequence of such variations for bridge loading is in the headways between trucks: increasing the number of trucks in a given time interval reduces the headways, thereby increasing the likelihood of observing two or more same-lane trucks on the bridge concurrently. In this work only variations that occur within a day are accounted for. Long term traffic growth is not allowed for and so the statistical models described are stationary. Further, it is taken that the 'economic year' is equivalent to about 50 weeks of weekday traffic. Therefore, 250 'simulation days' are taken to represent a calendar year.

2.2 Simulation of Bridge Traffic Loading

2.2.1 Generation of Traffic

Based on the models of the measured traffic, a Monte-Carlo simulation process is used to generate artificial traffic. In this way, significantly more data is made available than can be obtained through site measurements alone. The fundamental tool in this generation process is a (pseudo-)random number generator (RNG). Rubinstein [10] describes the importance of random number generation in Monte Carlo simulation and the fundamentals of computer-based RNGs.

The RNG initially used in this work was ran2() of *Numerical Recipes in C* [11]. However, as the work progressed, it was discovered that there were some inconsistencies with the results got from this algorithm: numbers very close to unity were essentially deterministic (though the machine epsilon of the computer used is 2.2×10^{-16}). In generating sequences of maxima from a parent distribution, it is essential that numbers close to unity are random also. Thus, the virtual pseudo-random number generator described by L'Ecuyer et al [12] was adopted. This generator gives excellent results, even for values very close to unity. This is due to its double precision methodology. It can have multiple separate streams, each of which is based on the multiple recursive generator MRG32k3a [13] which has a period of 2^{57}; the seeds of each stream are separated by 2^{127} steps. This is useful when multiple random deviates are required: for example, the Box-Muller transform for normal deviates requires two uniform deviates. In this case therefore, two separate streams of L'Ecuyer's RNG are ideal.

2.2.2 Load Effect Calculation and Implementation

The generated traffic streams are placed over the influence lines of interest. Both measured[14] and theoretical influence lines may be used, as well as influence lines determined from finite element analyses [2]. The use of these measured influence lines extends the applicability of the load assessment procedure from mere theoretical considerations, to considerations of the actual behaviour of the bridge under investigation. The software developed for this research requires an influence line to be specified by algebraic equations. For theoretical influence lines, the exact expressions are used whilst for measured influence lines, a number of quadratic or cubic polynomials are fit to segments of the influence line.

The main programs developed for this work are written using object oriented programming. As an example of the approach, the virtual object for the truck, a fundamental element in this work, is explained. The properties of the physical truck are programmed into the *CTruck* class (for example, number of axles, GVW etc.). *CTruck* only allows the rest of the program access to these class members through an appropriate interface, reducing logic errors in the program. In addition, *CTruck* has a number of actions it can perform, the class functions. For example, *CTruck* returns its time of arrival on the bridge, writes itself to file, or deletes itself when asked to by an external function. Critical to this research, the *CTruck* class is treated as a single piece of data and (large) arrays (C++ Standard Template Library <*vector*> class) of such objects are therefore used to contain the artificial trucks in the computer memory.

The results described in this paper are based on simulations of a 1000-day sample period of two-lane bi-directional truck traffic. The resulting load effects are determined for bridge lengths in the range 20 m to 50 m. The particular load effects considered are:

- Load Effect 1: Bending moment at the mid-span of a simply supported bridge;
- Load Effect 2: Left support shear in a simply-supported bridge;
- Load Effect 3: Bending moment at central support of a two-span continuous bridge.

To minimize computing requirements only significant crossing events were processed and are defined as multiple-truck presence events and single truck events with Gross Vehicle Weight (GVW) in excess of 40 tonnes. When a significant crossing event is identified, the comprising truck(s) are moved in 0.02 second intervals across the bridge and the maximum load effects of interest for the event identified.

3 Analysis of Extremes

3.1 Basis

An extreme value analysis is performed on the load effect data collected from the simulation process. Many authors [15]–[17] describe the basic forms of an extreme value analysis but in this work an extension to the method is used. When the extremes of interest are generated from a single statistical generating mechanism, the three Fisher and Tippett [18], [19] families can be expressed in a single form; the Generalized Extreme Value distribution (GEV) [17]:

$$G(x) = \exp\left\{-\left[1-\xi\left(\frac{x-\mu}{\sigma}\right)\right]_+^{1/\xi}\right\} \tag{1}$$

where $[h]_+ = \max(h,0)$ and the parameter vector is $\theta = (\mu,\sigma,\xi)$ – the location, scale and shape parameters of the distribution respectively. The probability density function (PDF) is:

$$g(x;\theta) = G(x;\theta)\cdot\sigma^{-1}\left\{1+\xi\left(\frac{x-\mu}{\sigma}\right)\right\}^{-1/\xi-1} \tag{2}$$

In the case of bridge traffic load effect, a loss in accuracy results when traffic loading is taken to be a single statistical generating mechanism. Bridge traffic loading is a multi-mechanism phenomenon. Currently, it seems adequate to consider the mechanisms of loading caused by different numbers of trucks concurrently present on the bridge [2], [19]. That is, the distribution of load effect caused by two-trucks on the bridge differs from that caused by three trucks on the bridge. In [2], [19] it is shown that the exact distribution of load effect may be arrived at by considering the distribution associated with each mechanism as well their relative frequency of occurrence.

Considering there to be n_t event types and n_d loading events per day, and using the law of total probability, the exact distribution of daily maximum load effect $\bar{S}$ is then given by:

$$P\left[\bar{S} \le s\right] = \left(\sum_{j=1}^{n_t} F_j(s)\cdot f_j\right)^{n_d} \tag{3}$$

In which the cumulative distribution function and frequency of the *j*th event type is $F_j(\cdot)$ and f_j respectively. In practice these quantities are difficult to ascertain and in [19] it is shown that the exact distribution may be asymptotically approximated by:

$$P\left[\bar{S} \le s\right] = G_C(s) = \prod_{j=1}^{n_t} G_j(s) \tag{4}$$

where $G_j(\cdot)$ is the GEV distribution of the *j*th event type and $G_C(\cdot)$ represents the composite distribution statistics (CDS) approach. Thus it is necessary to note each loading event according to its truck composition and to order the loading events separately, noting the maximum event of each type for each day of simulation.

3.2 Computational Aspects

Given the data for each loading event type, maximum likelihood estimation is used to estimate the parameters of the GEV distribution that best represents the observations. Maximum likelihood requires maximization of the log-likelihood function of the distribution of interest. Optimization techniques often deal with minimizing functions; hence minimization of the negative log-likelihood is usually performed. The log-likelihood function for the GEV distribution is [17]:

$$l(\mu,\sigma,\xi;x) = -n\log\sigma - \left(1-\frac{1}{\xi}\right)\sum_{i=1}^{n}\log y_i - \sum_{i=1}^{n} y_i^{1/\xi} \tag{5}$$

where $y_i = 1-\xi\left(\dfrac{x_i-\mu}{\sigma}\right) > 0$ for $i = 1,\ldots,n$.

For parameter combinations where $y_i < 0$ (which occurs when a data point x_i has fallen beyond the range of the distribution) the likelihood is zero and the log-likelihood will be numerically ill-defined. Solution of (5) is done by numerical means – there is no analytical solution. Based on the elements of the Hessian matrix of Equation (5), Prescott and Walden [21] propose a Newton-Raphson technique which is generally found to converge quickly. Hosking's algorithm based on this [22] is commonly used for GEV estimation. Good starting values for the minimization of the negative log-likelihood function of the GEV distribution are obtained from the method of probability weighted moments (PWMs) described by Hosking et al. [23]. The results of published data sets [17] are used to verify the output. Significantly, it is found, however, that there are cases in which Hosking's algorithm does not converge, or does not achieve the same minimum function value as other methods. As a result a more robust optimization method is implemented.

The Nelder-Mead (NM) optimization algorithm [24] is also known as the amoeba algorithm [11] because of its slow robust movement across the *k*-dimensional surface of a function, where *k* is the dimension of the optimization problem. The

NM algorithm is based on a simplex – a geometric shape with $k+1$ corners. Lagarias et al [25] describe, in detail, the operations of the algorithm.

In the statistical processing undertaken in this work, the PWM method is used to initiate both the Hosking and NM algorithms – processing time is not substantial in any case. The program checks to see if the Hosking algorithm has a smaller negative log-likelihood than that of the NM algorithm. If not, the results of the NM algorithm are used. While good results can be obtained with manual re-injection of the Hosking algorithm, in general this is not possible for this research – the number of individual GEV fits is substantial for each run.

3.3 Comparison of Methods

To compare the CDS distribution of Equation (4) to the single distribution approach of (1) a representative set of load effect distributions are stipulated and given in Table 1. The distributions of daily maximum load effect obtained from a 1000-day simulation were "back-calculated" using a reverse application of the stability postulate [17]. These distributions are then the parent distributions that would result in the set of observed daily maximum distributions of load effect. Further, the parameters of the distributions are normalized to reflect the underlying relationship between the mechanisms regardless of load effect and bridge length. Given this set of stipulated distributions and frequency data, Equation (3) can be used to determine the exact distribution of daily maximum load effect and consequently the exact distribution of load effect for the 100 year design life.

Parameter	1-truck	2-truck	3-truck	4-truck
ξ	0.06	0.09	0.28	0.21
σ	1.41	2.37	9.99	22.76
μ	71.93	100	67.42	21.92
$n_d f_j$	3102	2566	517	19

Table 1: Parameters of mechanisms for comparison study.

For the comparison, 1000 daily maximum data points are generated from the known distributions. Both the single and CDS methods are used to determine the distribution of daily maximum load effect and to extrapolate to determine the 100 year load effect value. This procedure is repeated 100 times to arrive at a distribution of 100 year load effect for both the single distribution and CDS approaches. GEV distributions are fit to the return level points of both the single distribution and CDS approaches and these two distributions can be compared directly to the exact distribution, as shown in Figure 1.

It can be seen that, due to the reasons outlined, the single distribution method underestimates the return level. The difference between the single and CDS distribution modes is not large. It is the skewed nature of the CDS distribution that is significant: its 90-percentile is similar to that of the exact distribution. Thus the increased fidelity offered by the CDS approach is reflected in more accurate extrapolations.

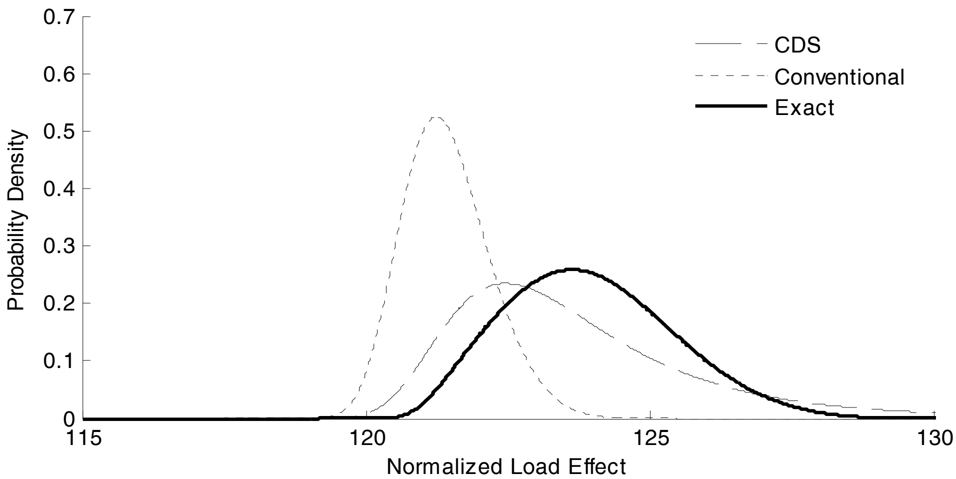


Figure 1: Distributions of 100 return level values.

4 Prediction

4.1 Basis

The predictions made from a data set are clearly variable, as may be seen from the results of the previous section in which the underlying distributions were exactly known. Therefore, in practical applications where the underlying statistical mechanisms can only be approximated, even more variability of the prediction must result. In the literature on bridge traffic load effects, few authors have considered the variability of the prediction [2]. In this work a recent statistical method termed predictive likelihood is used to estimate the distribution of the prediction [26].

In predictive likelihood, the observations are taken as the only incontrovertible known. All subsequent processing is taken to introduce variability and predictive likelihood accounts for this. Such processing uncertainty is the reason why conventional approaches are not generally repeatable – for the amounts of data normally simulated, there is considerable variation in repeated runs using the same algorithms. Predictive likelihood ranks predicted values, or predictands, based upon their likelihood given the data. It does this by calculating a distribution that maximizes the joint likelihood of observing both the data, L_y, and the predictand, L_z, jointly:

$$L_P\left(z \mid x\right) = \sup_{\theta} L_x\left(\theta; x\right) L_z\left(\theta; z\right) \tag{6}$$

Therefore, different postulated predictands will yield different values of predictive likelihood. This enables the predictands to be relatively ranked by likelihood. In **Error! Reference source not found.**, a range of possible values of predictions is shown along with the absolute data and a number of fitted distributions. The relative likelihood of each predictand is also shown. It can be seen that this process results in a distribution of predictand. Also, this process accounts for processing variability, as subsequently described.

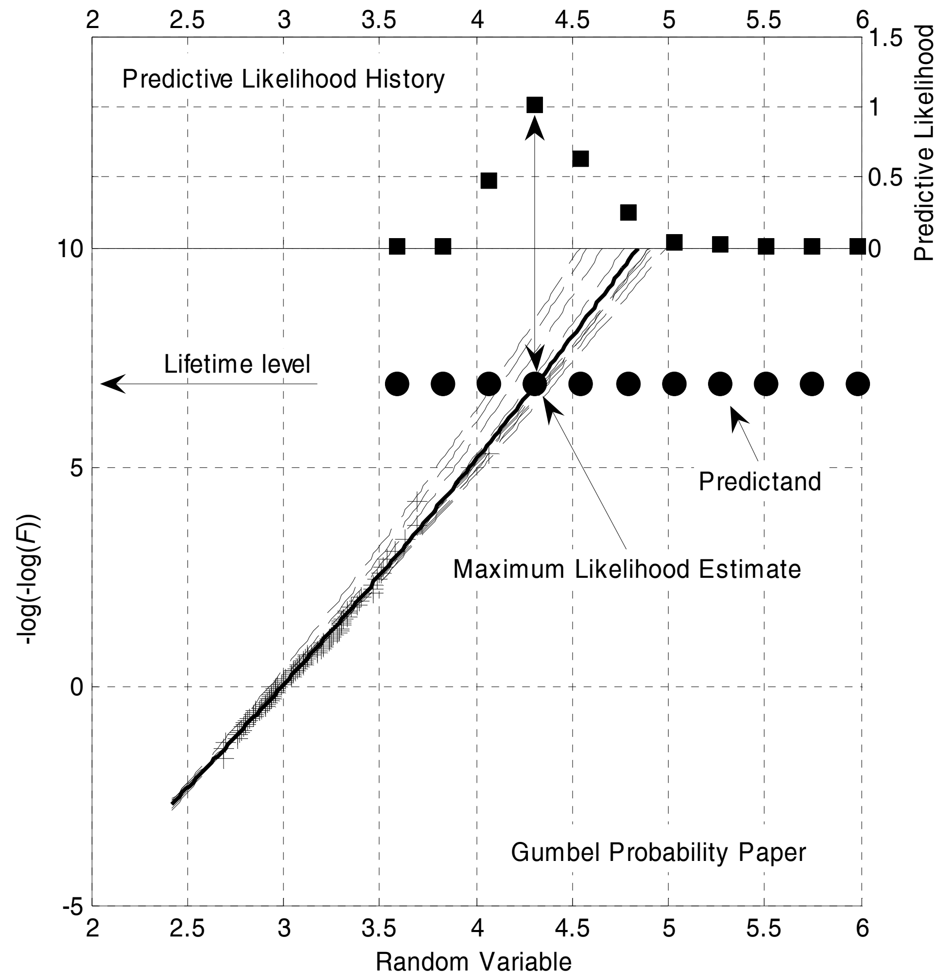


Figure 2: Sample predictive likelihood analysis.

4.2 Theory

Two modifications are required to the predictive likelihood of Equation (6) [27], [28]. The first accounts for the confidence in each parameter vector for each predictand; the second is a constant required to transform the problem into the correct domain. Allowing for these modifications, the modified profile predictive likelihood (L_{MP}) is given as:

$$L_{MP}\left(z \mid x\right)=\frac{L_P\left(z \mid x;\theta_z\right)}{\left|\frac{\partial\theta_z}{\partial\theta}\right|\sqrt{\left|\mathcal{I}\left(\theta_z\right)\right|}} \tag{7}$$

In this Equation, the square root of determinant of the Fisher information matrix, $\sqrt{\left|\mathcal{I}\left(\theta_z\right)\right|}$, (the Hessian matrix of the likelihood function) represents the confidence (information) about the parameter values. It is an inverse relationship: larger determinants represent less information and vice versa. Also, the parameter transform modification $\left|\partial\theta_z/\partial\theta\right|$ is required so that the problem is in the domain of the 'free' parameter vector, θ, which is reliant only upon the data.

The likelihood of the data for the CDS distribution is defined in this work to be the combined likelihood of each of the mechanisms of the CDS distribution:

$$\begin{aligned}\log\left[L_y\left(\theta;x\right)\right]&=l_y\left(\theta;x\right)\\&=\sum_{j=1}^{N}\left\{\sum_{i=1}^{n_j}\log\left[g_j\left(\theta_j;x_{j,i}\right)\right]\right\}\end{aligned} \tag{8}$$

where n_j is the number of data points for each event type; $x_{j,i}$ is the *i*th data point of event type *j*, and; $\theta_j=\left(\mu_j,\sigma_j,\xi_j\right)$ is the parameter vector for each $G_j\left(\cdot\right)$. The CDS distribution of a maximum of *m* sample repetitions, $G_{Z,C}\left(\cdot\right)$, is:

$$\begin{aligned}G_{Z,C}\left(z\right)&=\left[G_C\left(z\right)\right]^m\\g_{Z,C}\left(z\right)&=m\cdot g_C\left(z\right)\cdot\left[G_C\left(z\right)\right]^{m-1}\end{aligned} \tag{9}$$

in which $g_{Z,C}\left(\cdot\right)$ is its PDF. Therefore, the likelihood of the predictand, given the initial distribution is:

$$\begin{aligned}\log\left[L_z\left(\theta;z\right)\right]&=\log\left[g_{Z,C}\left(z\right)\right]\\&=\log\left\{m\cdot g_C\left(z\right)\cdot\left[G_C\left(z\right)\right]^{m-1}\right\}\end{aligned} \tag{10}$$

To determine the points of the predictive distribution, $f_{L_P}\left(z;x\right)$ (which correspond to each predictand examined), firstly the log modified predictive likelihoods are defined:

$$l_{MP}\left(z \mid x\right)=\log\left[L_{MP}\left(z \mid x\right)\right] \tag{11}$$

and the maximum value of $l_{MP}\left(z \mid x\right)$ obtained from the set of predictands is:

$$\hat{l}_{MP}\left(z \mid x\right)=\sup_{z}\left\{\log\left[L_{MP}\left(z \mid x\right)\right]\right\} \tag{12}$$

The curve of likelihood ratios is determined as:

$$f_{L_P}^{*}\left(z;x\right)=\exp\left\{l_{MP}\left(z \mid x\right)-\hat{l}_{MP}\left(z \mid x\right)\right\} \tag{13}$$

This curve is then normalized to the predictive distribution:

$$f_{L_P}\left(z;x\right)=\frac{f_{L_P}^{*}\left(z;x\right)}{\int f_{L_P}^{*}\left(z;x\right)} \tag{14}$$

Butler [29] points out that the parameter transform $\left|\partial\theta_z/\partial\theta\right|$ is constant. Therefore the normalization of the area under the curve of Equation (14) amounts to the evaluation of $\left|\partial\theta_z/\partial\theta\right|$. Therefore all terms in $L_{MP}\left(z \mid y\right)$ are known, yielding the predictive density.

4.3 Computational Aspects of Predictive Likelihood

4.3.1 Composite Distribution Predictive Likelihood Algorithm

Save for Davison [30], the statistical literature on predictive likelihood [26] does not generally consider its implementation. The algorithm used is presented here and aspects related to the numerical computations are examined.

For each value of the predictand, Equation (7) is maximized with the terms given by Equations (8) and (10). As up to four event types are involved in a typical bridge traffic loading problem, the maximization has a set of up to 12 parameters. Sequential quadratic programming optimization is used in this work to minimize the negative of the predictive likelihood function. A MATLAB toolbox is developed for this purpose as part of this work.

In each optimization, each GEV parameter vector must only operate on the data corresponding to its event-type. Therefore parameter bounds are used to enforce this requirement. The bounds used in this work are based on deviations from the ordinary maximum likelihood estimates, and are taken as:

$$\frac{\lambda_i}{\psi_i}\leq\lambda_i\leq\psi_i\lambda_i;\quad \text{where } \lambda=\left\{\mu,\sigma,\xi\right\};\ \psi=\left\{1.4,1.4,1.1\right\} \text{ and } i=1,2,3 \tag{15}$$

Whilst seemingly restrictive, the optimized parameter values are found to remain within these bounds. Also, the different bounds reflect the sensitivity of the fit to each of the parameters.

For each predictand considered an optimization is required to determine the parameter values that maximize Equation (6). Usually the range of predictands are equally and relatively closely spaced. In this work 100 predictands were considered. The start and end values of predictands are taken as 90% and 110% of the conventional extrapolation prediction. In optimization problems it is common to seed each optimization with the results of the previous one. However, in this work, even though the predictands are closely spaced, it is found necessary for each optimization to start on the initial maximum likelihood estimate, rather than the final parameter estimates of the previous predictand optimization. Though more computationally expensive, each optimization is therefore independent, and the risk of divergence of the solution is reduced. However, there remain situations in which an optimization can diverge from solution. Therefore, for each predictand, an intermediate optimization is carried out to provide initial parameter estimates – the function is constrained to return the predictand it is optimizing for:

$$z - G_C^{-1}\left(p_z;\theta\right) = 0 \tag{16}$$

where p_z is the probability level for the predictand, z. Having obtained the parameter vector that solves this constraint function, the second optimization is commenced with this parameter vector as the start point.

The solution that results for each predictand is then processed using numerical derivatives to determine the (up to 12×12-dimension) Hessian matrix of the solution – the observed Fisher information matrix. Also, the maximized value of predictive likelihood is brought forward to the analysis for the distribution of predictive likelihood.

4.3.2 Fitting the Predictive Distribution

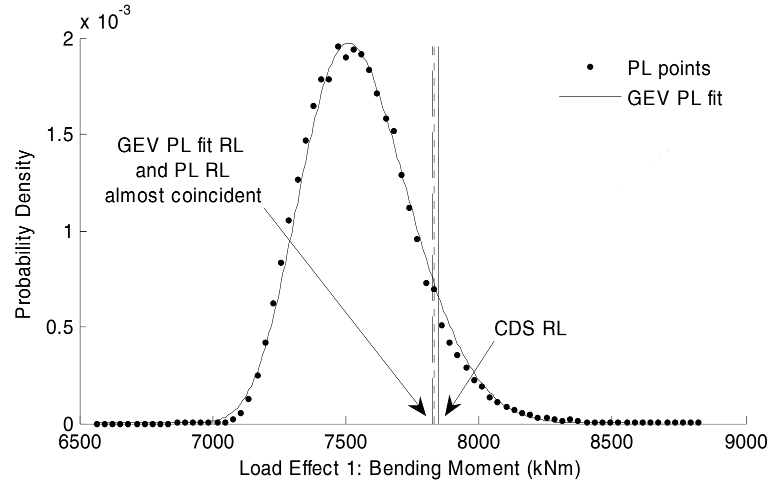
As only discrete values of L_{MP} are calculated at discrete intervals of predictand, the resulting distribution needs to be smoothed. Therefore, for this work, a GEV distribution is fitted through the discrete points that result after normalization of the area under the points to unity. A least-squares fit through these points is not appropriate as it unduly weights the larger relative likelihoods by assigning a weight of unity to all points. A number of weighting functions are possible but the one adopted in this work is to use a weight of unity for all points below the mode of the distribution, and to use a weight equal to the reciprocal of the predictand for points above the mode.

4.3.3 Effect of Data Scale and Sample Size

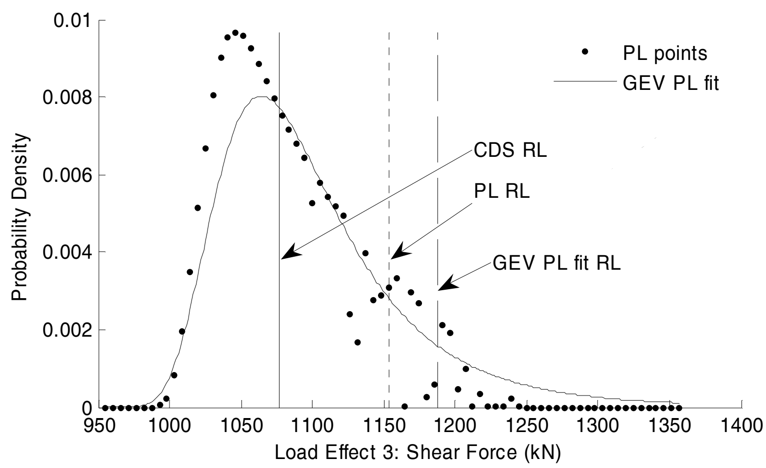
Due to the small order of numbers involved in predictive likelihood, numerical problems can arise from the state of the information matrix. An example is the numerical differentiation involved in calculating the information matrix. A useful measure of its stability with respect to numerical computations is the matrix condition number [31]. In this work, it has been found necessary to scale the input data to the predictive likelihood algorithm so that its order is less than 10. Higher order numbers exhibit severe ill-conditioning of the matrices with resultant effects on the modified predictive likelihood distribution. Inherent random variation of the data clearly affects the conditioning of the matrices. However, sample size has a more considerable effect. Significant variability remains in the determinant but the condition number is stable for sample sizes above about 150.

4.4 Application

The load effects resulting from a 1000-day simulation of traffic are analysed using predictive likelihood and the results are given in Table 2. The information matrices exhibited considerable numerical instability and consequently the modification for parameter variability is not made. This modification has been found to be generally slight in any case [2].



(a) Load Effect 1, 30 m bridge length



(b) Load Effect 3, 40 m bridge length

Figure 3: Characteristic load effect prediction (see text for details).

Two predictive distributions of 100-year lifetime-maximum load effect are presented. The GEV fits to the predictive distribution are also shown. The load effect with 10% probability of exceedance in 100 years is also indicated, both for the predictive likelihood points (PL RL) and the GEV fit to these points (GEV PL fit). Also given in each figure is the 1000-year maximum likelihood estimate of the return level (CDS RL), derived from the CDS distribution.

Some of the GEV fits to the raw predictive likelihood points are not obtained through objective means. In such cases, the upper tail is fit more closely than either the lower tail or the mode. In any case, the results have been derived from both the fits and the raw distributions and may be seen to be comparable from Table 2 – the maximum difference is about 3% for Load Effect 2, 40 m bridge length.

Given the differences between the predictive likelihood result (100-year with 10% probability of exceedance) and the conventional CDS result (1000-year return period), it is apparent that these two definitions of probability level are not equivalent when allowing for sources of variability. This has implications for the specification of acceptable probabilities and the manner in which practitioners estimate the associated design levels.

Load Effect	**Bridge Length (m)**	**Characteristic Load Effect**			**Percentage difference[a]**	
		PL[b]	**GEV PL[c]**	**CDS[d]**	**GEV**	**CDS**
1 (kNm)	20	4074	4073	4067	0.0	-0.2
	30	7830	7827	7852	0.0	0.3
	40	10814	10801	10701	-0.1	-1.0
	50	14150	14173	13893	0.2	-1.8
2 (kNm)	20	1074	1074	1067	0.0	-0.6
	30	1636	1641	1643	0.3	0.4
	40	2841	2854	2921	0.5	2.8
	50	3825	3839	3785	0.4	-1.1
3 (kN)	20	927	926	922	-0.1	-0.6
	30	969	969	963	0.0	-0.6
	40	1153	1187	1079	2.9	-6.5
	50	1235	1253	1185	1.4	-4.0

[a] Relative to numerical PL results;

[b] 90-percentile of 100-year distribution based on predictive likelihood points;

[c] 90-percentile of 100-year distribution GEV fit to predictive likelihood points;

[d] 1000-year return level based on CDS extrapolation.

Table 2: Predictive likelihood and conventional results.

5 Conclusions

In this paper the importance of the accurate assessment of bridge traffic loading is discussed. It is proposed that this is an area in which significant savings may be made by avoiding unnecessary repair and rehabilitation of existing bridges.

The bridge traffic load effect simulation process is described. In particular, measured Weigh-In-Motion data is statistically modelled to characterize the traffic at the site of measurement. Monte Carlo simulation of the modelled traffic is used to synthetically extend the amount of traffic data available. This simulated traffic is passed over bridge lengths and influence lines of interest, to determine the load effects that result. The resulting load effect data forms a population upon which a statistical analysis is carried out.

The composite distribution nature of bridge traffic loading is also described. This distribution of bridge traffic loading is a mixture of different types of loading events. For example, 1-truck presence events are very common, while 4-truck presence events are rare, but significant for loading. A statistical model that takes account of this mixture is presented and compared to theoretical examples. The computer implementation aspects of the model are also discussed.

The method of predictive likelihood is presented and applied to the bridge loading problem. An extension of predictive likelihood is presented which caters for composite distribution statistics problems. Predictive likelihood includes many sources of variability within the predictive likelihood distribution. The results of this approach are compared to a more conventional approach. The differences in lifetime load effects are considerable, yet within reason. As predictive likelihood accounts for sources of uncertainty in its estimation, it is to be preferred.

Of particular importance, the implementation of the predictive likelihood approach is described in detail. The numerical computations necessary for its implementation are described. A strategy for the implementation of predictive likelihood is described. The algorithm has several features that maximize its robustness. However, situations in which numerical problems arise are also identified and discussed.

It is also shown in this paper that predictive likelihood results differ from the more usual return period method. This will have implications for practitioners and code definitions. Also, it is shown that the predictive likelihood distribution represents a considerable increase in the information gained from a sample. Thus there is more confidence about the result in comparison with the return period approach. Overall, predictive likelihood is a valuable tool in estimating distributions of extremes of stochastic processes and its implementation is therefore presented here.

References

[1] COST 345, "Procedures Required for Assessing Highway Structures", Report of Working Group 1: "Report on the Current Stock of Highway Structures in

European Countries, the Cost of their Replacement and the Annual Cost of Maintaining, Repairing and Renewing them", available from http://cost345.zag.si/, 2004.
[2] C.C. Caprani, "Probabilistic Analysis of Highway Bridge Traffic Loading", Ph.D. Thesis, School of Architecture, Landscape, and Civil Engineering, University College Dublin, Ireland, 2005.
[3] A.S. Nowak and Y.K. Hong, "Bridge live load models", Journal of Structural Engineering, ASCE,117(9), 2757-2767, 1991.
[4] A. Bruls, P. Croce, L. Sanpaolesi and G. Sedlacek, "ENV1991 – Part 3: Traffic Loads on Bridges; Calibration of Load Models for Road Bridges", Proceedings of IABSE Colloquium, Delft, The Netherlands, IABSE-AIPC-IVBH, 439-453, 1996.
[5] S.F. Bailey and R. Bez, "'Site specific probability distribution of extreme traffic action effects", Probabilistic Engineering Mechanics, 14(1), 19-26, 1999.
[6] D.I. Cooper, "Development of short span bridge-specific assessment live loading", Safety of Bridges, Ed. P. C. Das, Thomas Telford, 64-89, 1997.
[7] C. Crespo-Minguillón and J.R. Casas, "A Comprehensive traffic load model for bridge safety checking", Structural Safety,19, 339-359, 1997.
[8] EC 1: "Basis of design and actions on structures, Part 3: Traffic loads on bridges", European Prestandard ENV 1991-3: European Committee for Standardisation, TC 250, Brussels, 1994.
[9] E.J. OBrien and C.C. Caprani, "Headway modeling for traffic load assessment of short-to-medium span bridges", The Structural Engineer, 83(16), 33-36, 2005.
[10] R. Y. Rubenstein, "Simulation and the Monte Carlo Method", John Wiley and Sons. 1981.
[11] W. H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, "Numerical Recipes in C", 2nd edn., Cambridge University Press, 1992.
[12] P. L'Ecuyer, R. Simard, E.J. Chen and W.D. Kelton, "An objected-oriented random-number package with many long streams and substreams", Operations Research, 50(6), 1073-1075, 2002.
[13] P. L'Ecuyer, "Good parameter sets for combined multiple recursive random number generators", Operations Research, 47(1), 159-164, 1999.
[14] E.J. OBrien, M.J. Quilligan and R. Karoumi, "Calculating an Influence Line from Direct Measurements", Bridge Engineering, Proceedings of the Institution of Civil Engineers, 159(BE1), 31-34, Mar. 2006.
[15] E. Castillo, "Extreme Value Theory in Engineering", New York, Academic Press, 1988.
[16] A.H.S. Ang and W.H. Tang, "Probability Concepts in Engineering Planning and Design, Vols. I and II", Wiley and Sons, 1975.
[17] S.G. Coles, "An Introduction to Statistical Modeling of Extreme Values", London, Springer-Verlag, 2001.
[18] R.A. Fisher and L.H.C. Tippett, "Limiting forms of the frequency distribution of the largest or smallest number of a sample", Proc. Cambridge Philosophical Society XXIV, Part II, 180-190, 1928.

[19] E. J. Gumbel, "Statistics of Extremes", Columbia University Press, 1958.
[20] C.C. Caprani, E.J. OBrien and G.J. Mclachlan, "Characteristic traffic load effects from a mixture of loading events on short to medium span bridges", Structural Safety, under review, 2006.
[21] P. Prescott and A.T. Walden, "Maximum likelihood estimation of the parameters of the Generalized Extreme-Value distribution", Biometrika, 67, 723-724, 1980.
[22] J.R.M. Hosking, "Algorithm AS215: maximum-likelihood estimation of the parameters of the Generalized Extreme-Value distribution", Applied Statistics, 34(3), 301-310, 1985.
[23] J.R.M. Hosking, J.R. Wallis and E.F. Wood, "Estimation of the Generalized Extreme Value distribution by the method of Probability-Weighted Moments", Technometrics, 27, 251-261, 1985.
[24] J. A. Nelder and R. Mead, "A simplex method for function minimization", The Computer Journal, 7, 308-313, 1965.
[25] J.C. Lagarias, J.A. Reeds, M.H. Wright, and P.E. Wright, "Convergence properties of the Nelder-Mead simplex method in low dimensions", SIAM Journal on Optimization, 9(1), 112-147, 1998.
[26] J.F. Bjørnstad, "Predictive likelihood: a review", Statistical Science, 5(2), 242-254, 1990.
[27] P.E. Mathiasen, "Prediction functions", Scandinavian Journal of Statistics, 6, 1-21, 1979.
[28] J.K. Lindsey, "Parametric Statistical Inference", Oxford: Oxford University Press, 1996.
[29] R.W. Butler, "Approximate predictive pivots and densities", Biometrika, 76, 489-501, 1989.
[30] A.C. Davison, "Approximate predictive likelihood", Biometrika, 73, 323-332, 1986.
[31] G. Golub and C. Van Loan, "Matrix Computations", 3rd edn., John Hopkins University Press, Baltimore, Massachusetts, 1996.