The Use of Predictive Likelihood to Estimate the Distribution of Extreme Bridge Traffic Load Effect

Colin C. Caprani* and Eugene J. OBrien+

* Department of Civil and Structural Engineering, Dublin Institute of Technology, Ireland

+ School of Architecture, Landscape and Civil Engineering, University College Dublin, Ireland

* Corresponding author.

Email: colin.caprani@dit.ie; Tel.: +353-402-2993; Fax: +353-1-402-3720

ABSTRACT

To assess the safety of an existing bridge, the loads to which it may be subject in its lifetime are required. Statistical analysis is used to extrapolate a sample of load effect values from the simulation period to the required design period. Complex statistical methods are often used and the end result is usually a single value of characteristic load effect. Such a deterministic result is at odds with the underlying stochastic nature of the problem. In this paper, predictive likelihood is shown to be a method by which the distribution of the lifetime extreme load effect may be determined. An estimate of the distributions of lifetime maximum load effect facilitates the reliability approach to bridge assessment. Results are presented for some cases of bridge loading, compared to a return period approach and significant differences identified. The implications for the assessment of existing bridges are discussed.

1

KEYWORDS

Bridge, Statistics, Load, Predictive Likelihood, Probabilistic, Extreme Value, Traffic,

Monte Carlo, Simulation

1. INTRODUCTION

To assess the safety of structures, it is necessary to have estimates for the load or load effect to which it is subject. Statistical approaches are commonly adopted as the tools through which loads with an acceptably small probability of exceedance are determined. The assessment of existing bridges is a particular case when such analyses are very useful. In general, it is particularly expensive to repair or replace deteriorated bridges due to the cost of the new structure, disruption to traffic and the cost of resulting delays. Large savings may be made by proving that many bridges are safe without intervention and statistical analysis of bridge loading facilitates this.

The estimation of loading usually involves the following steps. At the site of interest, a Weigh-In-Motion (WIM) system is used to measure the traffic characteristics of the site. After sufficient traffic data has been measured, the statistical distributions for the traffic characteristics are determined. These distributions then form the basis for Monte Carlo simulation of vehicle traffic at the site – such simulations may be carried out for a much longer period than is practical to record. The use of Monte Carlo simulation facilitates the occurrence of unobserved values of traffic characteristics whilst remaining faithful to the original observations. However, some authors proceed directly from measured traffic to estimate load effect (see for example, [1] and [2]). The generated vehicle traffic is then processed using influence lines to give a population of load effects. Finally, an extreme value form of analysis is performed on these results and used to estimate the load effect with an acceptably small probability of exceedance. For example, the Eurocode for bridge loading [3] requires that the design load effect be calculated for a 1000-year return period. This is commonly interpreted as the load effect

that has an approximate probability of exceedance of 10% in a 100-year design life. If the design life is taken as 100 years, the objective is to best estimate the distribution of 100-year (lifetime) load effect, so that (among other things) its 90th-percentile (approximately the 1000-year level) can be established.

The characteristic value for some long return period is a function of the parameters of the distribution chosen to represent the population of load effect. These parameters are estimated from a sample of load effect and are therefore subject to variability. Indeed, even if data could be obtained for the full return period, the inherent variability of traffic loading means that, in general, another such set of data would result in another characteristic value. Of course, there must be some particular value of load effect which has, for example, a 10% probability of exceedance in 100 years, but such a value needs to be derived from a distribution which takes into account many sources of variability. Various methods exist in the statistical literature for calculating such distributions – the delta method [4] and bootstrapping [5] being two. However predictive likelihood has advantages over these as it balances possible predictions against the observed data and can account for more sources of variability.

There has been a focus in many previous studies (including the background work for the Eurocode [3]) on the calculation of characteristic load effect for a specified return period (e.g., 1000 years). This paper demonstrates that, with the application of predictive likelihood [6] to the same data, an estimate of the distribution of maximum-in-lifetime (e.g., 100 years) load effect can be determined. This provides considerably richer information to complement the single value of characteristic load effect. It has particular potential for use in reliability theory [7].

2. LOAD EFFECT PREDICTION

Conventional Extrapolation Procedure

Usually, for a particular site, Monte Carlo simulation of statistically modelled traffic is carried out for a bridge(s) and load effect(s) of interest. The resulting load effect data is then used as the basis of an extreme value theory analysis [8]. Whilst the Peaks-Over-Threshold approach is applicable, the block maxima approach [4] is used in this study. A suitable period in which a sufficient number of loading events occur is required for convergence to an extreme value distribution and in this work a period of one day is used as it meets stationarity requirements [9].

To model the daily-maxima, the Generalized Extreme Value (GEV) distribution [4], [10]–[11] is used. This can be shown to incorporate the three asymptotic distributions of [12]. Its cumulative distribution function is:

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

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

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

Maximum likelihood estimation is based on maximisation of the likelihood function:

5

$$L(\theta; y) = \prod_{i=1}^{n} f(y_i; \theta) \tag{3}$$

in which $f(y; \theta)$ is the probability density function of the chosen model and y_i is an observed data point. Usually the log-likelihood function is used, which for the GEV distribution, is [4]:

$$\log[L_y(\theta; y)] = l_y(\theta; y)$$
$$= -n \log \sigma - \left(1 - \frac{1}{\xi}\right) \sum_{i=1}^{n} \log y_i - \sum_{i=1}^{n} y_i^{1/\xi} \tag{4}$$

In equation (4), the subscripts y indicate reliance on the data, of which y_i is an individual observation. The major benefit of using this distribution is that, through inference on ξ, the data determines the correct tail behaviour, avoiding the need to make a subjective *a priori* judgement on which Fisher-Tippett [12] extreme value form to adopt [4].

The sample daily maxima load effect values are used to make inference on the distribution of the daily maxima load effect. Maximum likelihood estimation is used here as it yields information of further use and is a best asymptotically normal estimator [13]. Considering there to be 250 working days per year, the 1000-year return period (as used in the Eurocode [1]) corresponds to a probability of 1:250 000. An example of such extrapolation is shown in Figure 1 on Gumbel probability scale [14].

Bridge Load Prediction

Authors have used many different methods to predict the lifetime bridge load effect from measured or simulated load effect data. In many studies by Nowak and others [15]–[18] straight lines are superimposed on the tails of the distributions and extrapolated to determine the characteristic load effect values. In other studies by Nowak [2], [19], curved lines on normal probability paper are used for the extrapolation. Based on measured traffic samples [20]–[21] consider and compare several methods of extrapolation of the basic histogram of load effect. Grave et al [22] use a weighted least-squares approach to fit Weibull distributions to load effect values. This process is repeated to give an estimate of the distribution of characteristic values. These authors use the upper $2\sqrt{n}$ data points as recommended by Castillo [11] for data that may not be convergent to an extreme value population. Bailey and Bez [23]–[24] determine that the Weibull distribution is most appropriate to model the tails of the load effect distributions and used maximum likelihood estimation. Cooper [1], [25], uses measured truck loading events to determine the distribution of load effect. The author raises this distribution to a power to establish the distribution of the maximum load effect from 4.5 days of traffic. This is fit with a Gumbel distribution which is used to extrapolate to a 2400 year return period. Crespo-Minguillón and Casas [26] adopt a Peaks-Over-Threshold approach and use the Generalized Pareto Distribution to model the exceedances of weekly maximum traffic effects over a certain threshold. An optimal threshold is selected based on the overall minimum least-squares value and it is the distribution that corresponds to this threshold that is used as the basis for extrapolation.

A wide range of methods are used in the literature. While parameter uncertainty is considered by some (for example, [27]), the variability of the characteristic load effect is not generally assessed.

3. PREDICTIVE LIKELIHOOD

Fisherian Predictive Likelihood

Given a set of observations, parametric statistical inference requires the selection of a statistical model and estimation of the parameters of that model, that best represent the data. For a given model, there are many possible parameter vectors, θ, representing many possible distributions. Using the maximum likelihood estimator, the most likely distribution, $\hat{\theta}$, given the data, y, is that which maximizes the likelihood function. From this parameter vector, the maximum likelihood estimate of the characteristic value, z (the predictand), is identified for a given probability level. The predictand itself is a random variable since it is any possible realization of the characteristic value. Therefore it has the domain of the set of all real positive numbers. Predictive likelihood finds the most likely distribution, given both the data and a postulated possible realization of the predictand value. It does this by maximising the likelihood functions of the data, L_y, and the predictand, L_z, jointly:

$$L_P\left(z \mid y\right) = \sup_{\theta} L_y\left(\theta; y\right) L_z\left(\theta; z\right) \tag{5}$$

Equation (5) represents the Fisherian predictive likelihood [28] (also termed profile predictive likelihood), L_P. Predictive likelihood can therefore be viewed as a method which 'ranks' possible realizations of the predictand according to how likely they are to occur given the data. This 'ranking' results in the predictive distribution.

An example is illustrated in Figure 2. A random data sample from a GEV distribution

with parameter vector $\theta = (300, 20, 0.1)$ is fit using maximum likelihood estimation.

This is shown in Figure 2(a) as the solid black line. This stipulated distribution of the

data is shown in Figure 2(b) along with the distribution of the lifetime-maximum load

effect. Denoting the PDF of the data by $g(\cdot)$, the likelihood function for the data vector,

y is:

$$L_y(\theta; y) = \prod_{i=1}^{n} g(y_i; \theta) \tag{6}$$

For a postulated value of z, and denoting the PDF of the predictand by $g_z(\cdot)$, the

likelihood function is:

$$L_z(\theta; z) = g_z(z; \theta) \tag{7}$$

as there is only a single value, z. Similarly to maximum likelihood estimation, it is

easier to use the log-likelihoods – maximization of this function is equivalent to

maximization of the likelihood function itself. Therefore, equation (5), in conjunction

with equations (6) and (7), is written as:

$$\begin{aligned} l_P(z \mid x) &= \log\left[L_P(z \mid y)\right] \\ &= \sup_{\theta}\left\{\log\left[L_y(\theta; y)\right] + \log\left[L_z(\theta; z)\right]\right\} \\ &= \sup_{\theta}\left\{\sum_{i=1}^{n}\log\left[g(y_i; \theta)\right] + \log\left[g_z(z; \theta)\right]\right\} \end{aligned} \tag{8}$$

10

For a given predictand (at a probability of 0.99 in the example), the joint likelihood of both the data and predictand is maximized. By repeating the process for a range of alternative possible predictands, a range of possible distributions are found. The predictive likelihood values, obtained from equation (8), for ten values of predictand are shown in Figure 2(a). For each of the predictive likelihood maximizations, the ten GEV fits to the data are also shown in that figure. It is to be noted that these distributions are not 'forced' to go through the predictand as the distribution results from 'balancing' both the data and the predictand.

The value of this approach is that additional information is available: for each predictand, the maximized predictive likelihood value is available from equation (8). Their relative values provide the predictive distribution. The calculation of this predictive distribution – $f_{L_P}(z; y)$ – is explained later (equations (18)-(21)) but an example is shown in Figure 2(a). It can be seen from Figure 2(a) that the most likely value of the predictand from the predictive likelihood distribution (its mode) coincides, as expected, with the maximum likelihood estimate of the predictand.

Modified Predictive Likelihood

Mathiasen [29] notes some problems with Fisherian predictive likelihood. Of particular relevance to this work is that each function maximization does not account for the variability of the derived parameter vector, θ.

Many forms of predictive likelihood have been proposed in the literature to overcome the problems associated with the Fisherian formulation. In this work, the predictive likelihood method proposed by Butler [30], based on that of Fisher [28] and Mathiasen

[29] and also considered by Bjørnstad [6], is used. Lindsey [31] describes the reasoning behind its development. This predictive likelihood is the Fisherian, modified so that the variability of the parameter vector resulting from each maximization is taken into account.

Likelihood is a multi-dimensional function of the parameters of the parent distribution (components of θ). Edwards [32] describes the meaning of the likelihood (multi-dimensional) surface at the point of maximum likelihood. In particular, the determinant of the Fisher information matrix, $\left| \mathcal{I}(\theta) \right|$, (the Hessian matrix of the likelihood function) may be seen as the volume under the (multi-dimensional) surface of the likelihood function at the maximum likelihood estimate (MLE). A narrow likelihood function indicates more confidence (information) about the parameter values. Therefore larger volumes (determinants) represent less information and vice versa. Mathematically, it is the absolute value of the determinant that is taken as volumes cannot be negative. Further, it is the square root of the determinant that is used as a measure of the variability of the parameter vector in the region of the estimate as described by Edwards [32]. Hence, this metric, $\sqrt{\left| \mathcal{I}(\theta_z) \right|}$, forms a measure of credibility for a given parameter vector. To allow for the effect of parameter variability on Fisherian predictive likelihood, this metric is used to weight each value of Fisherian predictive likelihood obtained (for the parameter vector of each predictand considered). In this way, the variability of the parameters is allowed for, in a relative sense. As the resulting function is normalized to a distribution, this relative measure is adequate.

One further modification is required to the Fisherian predictive likelihood. The parameter vector determined for each value of the predictand is dependent on both the

data and the predictand (denoted θ_z). A modification, termed the parameter transform modification is required so that the problem is in the domain of the 'free' parameter vector, θ, which is reliant only upon the data. Thomasian [33] provides further information on parameter transformations. That which is relevant here is $|\partial\theta_z / \partial\theta|$.

Allowing for these modifications to the Fisherian predictive likelihood, the modified profile predictive likelihood (L_{MP}) is given as:

$$L_{MP}(z\,|\,y) = \frac{L_P(z\,|\,y;\theta_z)}{\left|\dfrac{\partial\theta_z}{\partial\theta}\right|\sqrt{\left|\mathcal{I}(\theta_z)\right|}} \tag{9}$$

Butler [30] points out that the parameter transform $|\partial\theta_z / \partial\theta|$ is constant. Therefore normalization of the area under $L_P(z\,|\,y;\theta_z)\big/\sqrt{\left|\mathcal{I}(\theta_z)\right|}$ amounts to evaluation of $|\partial\theta_z / \partial\theta|$ and hence $L_{MP}(z\,|\,y)$ yields the predictive density of the predictand, $f_{L_P}(z;y)$.

Bridge Traffic Load Effect Formulation

Bridge traffic loading does not generally provide an homogenous (or independent and identically distributed) population of data. Caprani et al [34] have shown that bridge load effects are caused by a mixture of different types of loading event such as 1-truck and 2-truck loading events. These different loading event types each have a different distribution of load effect. The distribution of daily maximum load effect is derived from the theorem of total probability to be:

$$P\left[\overline{Y} \le y\right] = \left(\sum_{j=1}^{N} F_j(y) \cdot f_j\right)^{n_d} \tag{10}$$

where $\overline{Y}$ is the daily maximum load effect; y is the level of interest; $F_j(\cdot)$ is the parent distribution of load effect for event-type j which has relative frequency of occurrence f_j, and; n_d is the number of events that occurs in a day. The maximum number of event types is N. Hence, $f_j n_d$ is the expected number of events of type j that occur each day. Caprani et al [34] show that equation (10) asymptotically approaches a composite distribution of daily maximum load effect. For the N different types of loading event, the composite distribution, $G_C(\cdot)$, of daily maximum load effect is given by:

$$G_C(y) = \prod_{j=1}^{N} G_j(y) \tag{11}$$

where $G_j(\cdot)$ is any extreme value distribution fit to the daily maximum load effect data caused by loading event type j. Considering that, in this work, the GEV distribution (equation (1)) is used to model the distribution of daily maximum load effect for each of the event types, equation (11) becomes:

$$G_C(y) = \exp\left\{-\sum_{j=1}^{N}\left[1 - \xi_j\left(\frac{y - \mu_j}{\sigma_j}\right)\right]^{1/\xi_j}\right\} \tag{12}$$

14

This approach is termed Composite Distribution Statistics (CDS). It is possible to use CDS to account for the different types of vehicle that comprise the traffic stream, but this is not done in this work. For the purposes of calculating likelihoods, the composite probability density function, $g_C(y)$, is evaluated numerically in this work.

Caprani [9] shows that the parent distribution of load effect, for a wide range of bridge lengths and load effects, is well described by the GEV distribution. Therefore, by the stability postulate, each loading event-type only needs to occur at least once in a day in order that its distribution of daily maximum load effect can be described by the GEV distribution. Thus, to determine the distribution of lifetime load effect from equation (11), we assume that each loading event type occurs at least once (and once is sufficient, as just explained) in each day. When this is the case, the distribution of maximum load effect in *m* days, for the individual loading event-type *j* is:

$$G_{Z,j}(y) = \left[G_j(y) \right]^m \tag{13}$$

Thus the joint probability of occurrence of the level of interest *y*, when all loading event types are considered, is:

$$G_{Z,C}(y) = \left[G_1(y) \right]^m \ldots \left[G_j(y) \right]^m \ldots \left[G_N(y) \right]^m = \prod_{j=1}^{N} \left[G_j(y) \right]^m = \left[G_C(y) \right]^m \tag{14}$$

Thus the distribution of a maximum of *m* sample repetitions is given by equation (11) raised to the power *m*, as expected. The probability density function of equation (14) is thus:

$$g_{Z,C}(z) = m \cdot g_C(z) \cdot \left[G_C(z)\right]^{m-1} \tag{15}$$

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

$$\log\left[L_y(\theta; y)\right] = l_y(\theta; y)$$
$$= \sum_{j=1}^{N}\left\{-\sum_{i=1}^{n_j}\left[1 - \xi_j\left(\frac{y_{j,i} - \mu_j}{\sigma_j}\right)\right]^{1/\xi_j}\right\} \tag{16}$$

where n_j is the number of data points for each event type; $y_{j,i}$ is the ith data point of event type j, and; $\theta_j = \left(\mu_j, \sigma_j, \xi_j\right)$ is the parameter vector for each $G_j(\cdot)$. Based on equation (15), the likelihood of the predictand, given the initial distribution is:

$$\log\left[L_z(\theta; z)\right] = \log\left[g_{Z,C}(z)\right]$$
$$= \log\left\{m \cdot g_C(z) \cdot \left[G_C(z)\right]^{m-1}\right\} \tag{17}$$

Thus the distributions required for use in the predictive likelihood approach have been defined with consideration to the underlying stochastic process.

Establishing the Predictive Distribution

Curves of log-predictive likelihood are used to determine the predictive distribution, $f_{L_P}(z; y)$. Firstly, the log-predictive likelihoods are defined as:

$$l_{MP}(z \mid y) = \log\left[L_{MP}(z \mid y)\right] \qquad (18)$$

and its maximum value is defined as:

$$\hat{l}_{MP}(z \mid y) = \sup_{z}\left\{\log\left[L_{MP}(z \mid y)\right]\right\} \qquad (19)$$

Then, the curve of likelihood ratios is determined as:

$$f_{L_P}^{*}(z; y) = \exp\left\{l_{MP}(z \mid y) - \hat{l}_{MP}(z \mid y)\right\} \qquad (20)$$

This curve is then normalized to the predictive distribution:

$$f_{L_P}(z; y) = \frac{f_{L_P}^{*}(z; y)}{\int f_{L_P}^{*}(z; y)} \qquad (21)$$

Save for Davison [35], the statistical literature on predictive likelihood does not generally consider its implementation. Numerical instability is a feature of predictive likelihood function maximization; the details of the algorithm used to address these problems is given by [9] and [36].

17

4. APPLICATION

Theoretical application

Distributions that reflect the usual relationship between loading events are described by Caprani [9] and given in Table 1. Samples from these distributions are used to simulate the statistical analysis of bridge traffic loading. These examples are used here to assess the accuracy of the composite distribution predictive likelihood. Two random data samples of 1000 points are generated and predictive likelihood analyses are performed. The results are assessed against the exact distribution of lifetime load effect, calculated from equation (10) with the problem parameters given in Table 1. We consider the distribution of load effect for a 100-year design life. The final design load effect will therefore be the 90th percentile value of this distribution (that is, the load effect with approximately 10% probability of exceedance in 100 years).

Figure 3 shows the results of the application of predictive likelihood to the two generated data sets. It can be seen that different calculated distributions of lifetime load effect are estimated for the different samples. However, this is not surprising as in any analysis the data is the only incontrovertible evidence and different data will result in a different prediction. Both distributions compare reasonably well with the exact 100-year return level distribution determined from equation **Error! Reference source not found.** (10): the mode of the exact distribution is well approximated by the predictive likelihood distribution. The tails are not approximated as well. Considering that the predictive likelihood distribution modes are to the left of the exact mode, predictive likelihood does not appear to be unduly sensitive to the particular sample obtained. The estimation of the design load effect, represented by the 90th-percentile of the lifetime

distributions of load effect, determined through the use of predictive likelihood, are conservative compared with that of the exact distribution. Of course, the particular value of design load effect is sample-dependent.

Application to Bridge Loading

WIM data, taken from the A6 motorway near Auxerre, France, is used to assess the implications of predictive likelihood on the estimation of characteristic bridge traffic loading. Weight and dimensional data were collected for 36 373 trucks travelling in the two slow lanes of the 4-lane motorway. The statistical models of the traffic characteristics provided input to Monte-Carlo simulations of traffic at the measured site. The distribution for headways, in particular, was found to be important and is modelled as described by OBrien and Caprani [37]. In this paper, we do not consider traffic growth as part of the problem, and so the statistical model is based on a stationary process.

A 1000-day sample period of two-lane bi-directional truck traffic is generated and 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.

The load effects resulting from the 1000-day simulation of Auxerre traffic are analysed using predictive likelihood and the results are given in Table 2. Unfortunately the information matrices exhibited considerable numerical instability and so the modification of predictive likelihood for parameter variability could not be made to the results presented. However, for some stable problems, Caprani [9] finds that the allowance for parameter variability only slightly affects the final predictive distribution. In any case, the basic Fisherian predictive likelihood distribution is informative about the distribution of lifetime load effect.

Sample predictive distributions of 100-year lifetime-maximum load effect are presented in Figure 4 and Figure 5. Also shown is a GEV fit to the predictive distribution. The GEV distribution is reasonable as it is sufficiently flexible and by virtue of the stability postulate [8] is the exact form of distribution of the lifetime load effect (or return level, RL). Further, the load effect with 10% probability of exceedance in 100 years is 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 fully objective means. In such cases, the approach is to fit the upper tail more closely than either the lower tail or the mode. Due to the numerical nature of the predictive distributions themselves, such GEV fits may be considered as a smoothing process. 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 differences are generally negligible, the maximum difference being about 3% for Load Effect 2, 40 m bridge length.

Comparison of the predictive likelihood results with the 1000-year CDS results are given in Figure 6. Of significance is the fact that the usual method of extrapolation to a 1000-year return period results in general non-conservative results (with the exception of Load Effect 2, 40 m bridge length), compared with either of the predictive likelihood-based results. However, the differences are not substantial. Further, in the light of the results of the theoretical example, it may be surmised that the predictive likelihood results are closer to the actual lifetime load effect than those of the more usual CDS extrapolation technique.

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 no longer equivalent. This has implications for the specification of acceptable probabilities and the manner in which practitioners estimate the associated design levels.

5. SUMMARY

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. This method is applied to problems for which the results are known and the result found to be good. The method is then applied to the results of bridge load simulations. Predictive likelihood generally gives larger lifetime load effect values than the usual return period approach. This is as a result of inclusion of sources of variability within the predictive likelihood distribution. The differences in lifetime load effects are considerable, yet within reason, and are also dependent on the influence line and bridge length. This is to be expected from the physical nature of the problem.

The application of predictive likelihood is shown to require strict definition of acceptable safety levels, as the more usual return period definition does not yield the same results in general. This will have implications for practitioners and code definitions. Also, it is shown that in comparison to the return period approach, which generates a single predictand, the predictive likelihood distribution represents a considerable increase in the information gained from a sample. This increase in information represents more confidence about the result in comparison with the return period approach. Therefore predictive likelihood is a valuable tool in estimating distributions of extremes of stochastic processes.

6. REFERENCES

[1] Cooper DI. Development of short span bridge-specific assessment live loading. Safety of Bridges, Ed. P. C. Das, Thomas Telford, 64-89, 1997.

[2] Nowak AS. Load model for bridge design code. Canadian Journal of Civil Engineering 1994;21;36-49.

[3] 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.

[4] Coles SG. An Introduction to Statistical Modeling of Extreme Values. London: Springer-Verlag, 2001.

[5] Davison AC and Hinkley DV. Bootstrap Methods and Their Application. Cambridge: Cambridge University Press, 1997.

[6] Bjørnstad JF. Predictive likelihood: a review. Statistical Science 1990;5(2):242-254.

[7] Melchers RE. Structural Reliability Analysis and Prediction, 2nd End., Chichester: John Wiley & Sons, 1999

[8] Gumbel EJ. Statistics of Extremes. Columbia University Press, 1958.

[9] Caprani CC. Probabilistic Analysis of Highway Bridge Traffic Loading. Ph.D. Thesis, School of Architecture, Landscape, and Civil Engineering, University College Dublin, Ireland, 2005. Available at www.colincaprani.com.

[10] Jenkinson A F. The frequency distribution of the annual maximum (or minimum) of meteorological elements. Quarterly Journal of the Royal Meteorological Society 1955;81:158–171.

[11] Castillo E. Extreme Value Theory in Engineering. New York: Academic Press, 1988.

[12] Fisher RA and Tippett LHC. Limiting forms of the frequency distribution of the largest or smallest number of a sample. Proc. Cambridge Philosophical Society 1928;XXIV, Part II:180-190.

[13] Mood AM, Graybill FA and Boes DC. Introduction to the Theory of Statistics, 3rd edn., McGraw-Hill, Statistics Series, 1974.

[14] Ang AHS and Tang WH. Probability Concepts in Engineering Planning and Design, Vols I and II. Wiley and Sons, 1975.

[15] Nowak AS. Probabilistic basis for bridge design codes. International Conference on Structural Safety and Reliability, San Francisco, 2019-2026, 1989.

[16] Heywood RJ and Nowak AS. Bridge live load models. International Conference on Structural Safety and Reliability, San Francisco, 2147-2154, 1989.

[17] Nowak AS, Hong YK and Hwang ES. Bridge load models. ICASP 6, 880 – 887, 1991.

[18] Nowak AS and Hong YK. Bridge live load models. Journal of Structural Engineering, ASCE, 1991;117(9):2757-2767.

[19] Nowak AS. Live load model for highway bridges. Structural Safety 1993;13:53-66.

[20] Bruls A, Croce P, Sanpaolesi L and Sedlacek G. 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.

[21] Flint AR and Jacob BA. Extreme traffic loads on road bridges and target values for their effects for code calibration. Proceedings of IABSE Colloquium, Delft, The Netherlands, IABSE-AIPC-IVBH, pp. 469-478, 1996.

[22] Grave SAJ, OBrien EJ and O'Connor AJ. The determination of site-specific imposed traffic loadings on existing bridges. Bridge Management 4, eds. M.J. Ryall et al, Thomas Telford, 442-449, 2000.

[23] Bailey SF and Bez R. A parametric study of traffic load effects in medium span bridges. Proceedings of the Fourth International Conference on Short and Medium Span Bridges, Halifax, Canada, 503-514, 1994.

[24] Bailey SF and Bez R. Site specific probability distribution of extreme traffic action effects. Probabilistic Engineering Mechanics 1999;14(1):19-26.

[25] Cooper DI. The determination of highway bridge design loading in the United Kingdom from traffic measurements. Pre-Proceedings of the First European Conference on Weigh-in-Motion of Road Vehicles, eds. B. Jacob et al., E.T.H., Zürich, 413-421, 1995.

[26] Crespo-Minguillón C and Casas JR. A Comprehensive traffic load model for bridge safety checking. Structural Safety 1997;19:339-359.

[27] Bhattacharya B, Li D, Chajes MJ and Hastings J. Reliability-based load and resistance factor rating using in-service data. Journal of Bridge Engineering, ASCE, Vol. 10, No. 5, pp. 530-543. 2005.

[28] Fisher RA. Statistical Methods and Scientific Inference. Edinburgh: Oliver and Boyd, 1956.

[29] Mathiasen PE. Prediction functions. Scandinavian Journal of Statistics 1979;6:1-21.

[30] Butler RW. Approximate predictive pivots and densities. Biometrika 1989;76:489-501.

[31] Lindsey JK. Parametric Statistical Inference. Oxford: Oxford University Press, 1996.

[32] Edwards AWF. Likelihood: An Account of the Statistical Concept of Likelihood and its Application to Scientific Inference, 2nd edn. Baltimore: Johns Hopkins University Press, 1992.

[33] Thomasian AJ. The Structure of Probability Theory with Applications. New York: McGraw-Hill, 1969.

[34] Caprani CC, OBrien EJ and Mclachlan GJ. Characteristic traffic load effects from a mixture of loading events on short to medium span bridges. Structural Safety Vol. 30(5), pp. 394-404.2008.

[35] Davison AC. Approximate predictive likelihood. Biometrika 1986;73:323-332.

[36] Caprani CC. OBrien, E.J., Statistical Computation for Extreme Bridge Traffic

Load Effects. Proc. 8th Int. Conf. on Computational Structures Technology. Ed.

B.H.V. Topping. Stirling, Scotland: Civil-Comp Press. Paper No. 139, 17 pp.,

2006.

[37] OBrien EJ and Caprani CC. Headway modeling for traffic load assessment of

short-to-medium span bridges. The Structural Engineer, 2005;83(16):33-36.

Table 1: Parameters of mechanisms for theoretical example.

Table 2: Table of predictive likelihood results.

Figure 1: Sample extrapolation procedure showing data and GEV fit.

(a) Sample predictive likelihood analysis;

(b) Parent distribution and distribution of maxima for $G(y; 300, 20, 0.1)$;

Figure 2: Empirical description of predictive likelihood analysis.

Figure 3: Predictive likelihood result for Study 1.

Figure 4: Characteristic load effect prediction for Load Effect 1, 30 m bridge length (see text for details).

Figure 5: Characteristic load effect prediction for Load Effect 3, 40 m bridge length (see text for details).

Figure 6: Differences in the CDS return period characteristic load effect prediction relative to the GEV fit to the predictive likelihood results.

Table 1

Table 1: Parameters of mechanisms for theoretical example.

	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 2

Table 2: Table of predictive likelihood results.

Load Effect	Bridge Length (m)	Characteristic Load Effect			Percentage difference[a]	
		PL[b]	GEV[c]	CDS[d]	GEV	CDS
1	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	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	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;

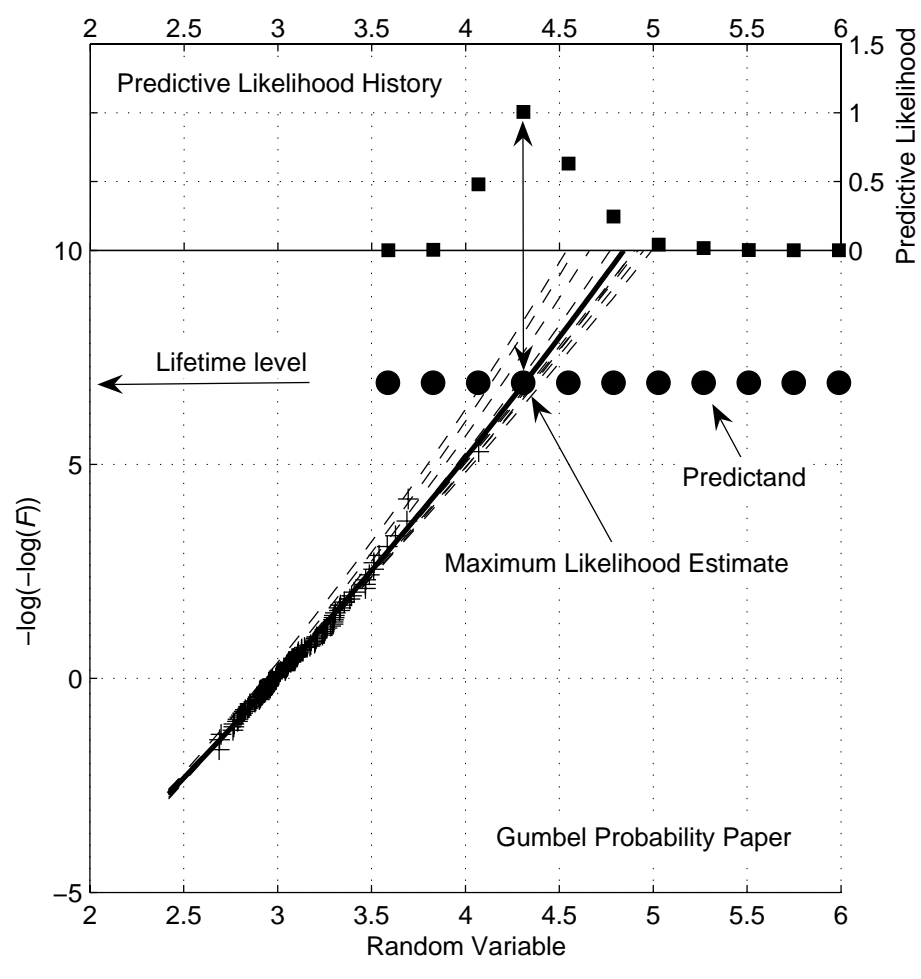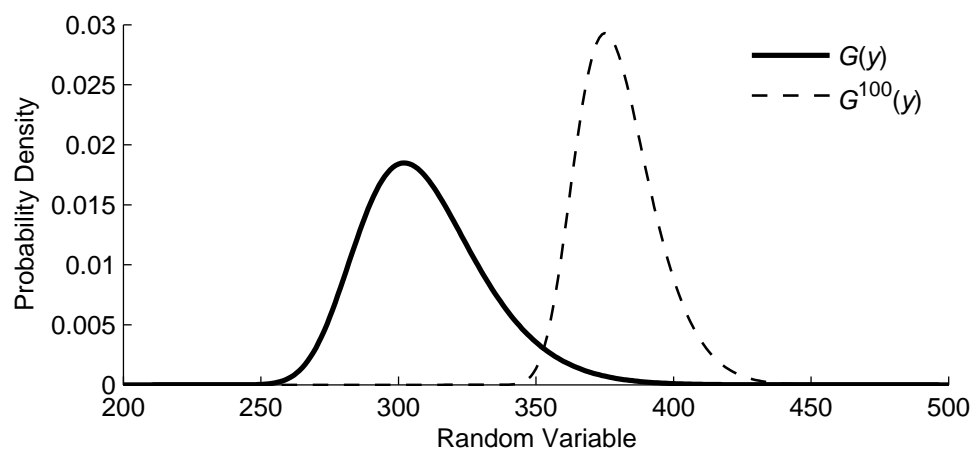[d] 1000-year return level based on CDS extrapolation.

Figure 1
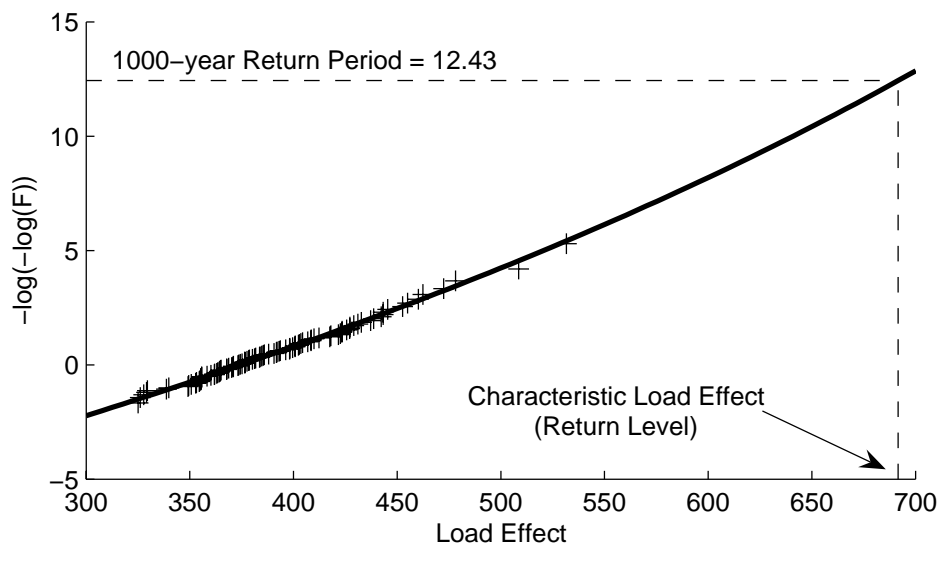
Figure 2b
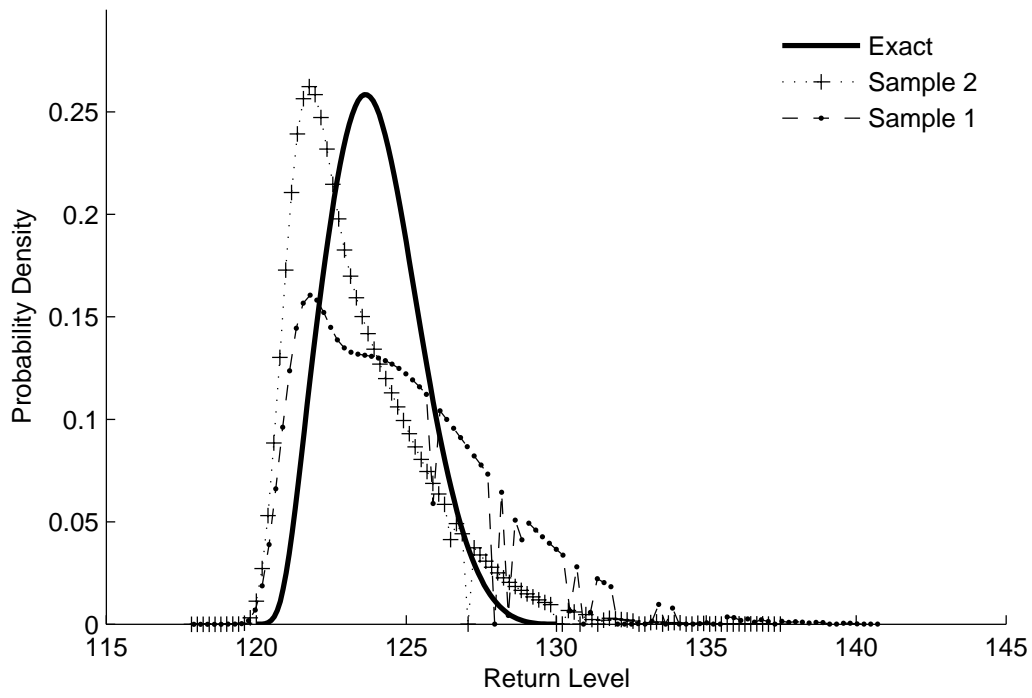
Figure 3

Figure 4

Figure 5

Figure 6