Bayesian hierarchical modelling of bridge traffic loading across a road network

Akbar Rizqiansyah
PhD Student, Dept. of Civil Engineering, Monash University, Melbourne, Australia

Colin C. Caprani
Assoc. Professor, Dept. of Civil Engineering, Monash University, Melbourne, Australia

ABSTRACT: The prediction of extreme traffic loading is a crucial part of bridge design and assessment. It provides a basis of how much action, and conversely the strength required for a bridge in its lifetime. However, there remain large uncertainties in the predictions of extreme loading due to the complex nature of the underlying traffic and its load effects. So far, efforts to model extreme load effects have been done primarily using standard extreme value methods, such as the generalized extreme value distribution and generalized Pareto distribution. However, these efforts provide techniques for fitting data from a single bridge only, requiring extrapolations for predictions to other bridge spans, which carries large uncertainties. Single fit methods also fail to take advantage of information contained in other spans that could be used to reduce estimation uncertainties – the 'shrinkage' effect. In this paper, a modern Bayesian hierarchical model is developed using the generalized extreme value distribution, covering intermediate spans where data is not available at the time of fitting. First, simple Bayesian model was explained and used with a simple and considered realistic traffic model. The simple model is then expanded to a Bayesian hierarchical model to simultaneously fit a range of spans. The final model shows accurate predictions for intermediate spans not used for the fitting process and provides reduced uncertainties compared to the single-span fits. This work provides a basis for estimating load effects across an entire road network at once; something not previously feasible.

1. INTRODUCTION

Bridges form an integral part of the road network. They are relied on daily to carry trucks ferrying heavy goods all over the world. Failure of bridges means disconnection of communities, loss of assets, and worst, loss of life. Hence, it is crucial for bridges to be designed and assessed so it could withstand loads it will be subjected to throughout its lifetime. However, the task of predicting these maximum loads is often difficult due to the uncertain and complex nature of the factors creating the loads. Although most states and countries impose legal limits on the characteristics of their trucks, e.g., number of axles, weights, and weight distribution, theoretically creating an upper bound on the maximum traffic load a bridge will be subjected to, in reality, these limits can be exceeded by heavier than legal trucks. Of interest is the extreme bridge traffic loading, i.e., loadings caused by extremely heavy trucks causing extreme loading values that occurs infrequently.

The problem of predicting extreme bridge traffic loading have been explored by numerous authors using various techniques. Most authors rely on statistical methods to predict values beyond the available data to cover bridges' lifespan. One of the most popular distributions used is the Generalized Extreme Value (GEV) distribution based on block maxima (Coles 2001), and has since been applied in traffic loading problem and expanded upon into the Cumulative Distribution Statistics (Caprani 2005), the GEV mixture model (Dai et al. 2022), applied in Bayesian method to predict account for growth in traffic (Yu et al. 2019), among others. Another popular method uses the Generalized Pareto Distribution (Crespo-Minguillón and Casas 1997; Nesterova et al. 2019) based on peak-over-threshold method. However, past studies only fit their model using data available for a handful of bridge spans, resulting in prediction only for a handful of bridge spans. No studies so far have been conducted on the prediction of bridge spans where data were not available during the model fitting. In contrast, the road network consists of bridges of varying spans, where data are unavailable for fitting and predictions were never made. Hence, a model that could infer extreme bridge traffic loading for any arbitrary bridge span is needed.

In this paper, a Hierarchical Bayesian (HB) model for bridge traffic loading is introduced. First, a simple Single-Span Bayesian (SSB) model is developed where parameters of different bridge spans are independent of one another. The SSB model is then expanded into the Hierarchical Bayesian model where functional relationships between parameters of different spans are proposed, allowing simultaneous fitting and predictions to be made for intermediate spans. The accuracy of both models is then compared against conventional Maximum Likelihood Estimator (MLE), and the accuracy of prediction made by the HB model tested.

2. BAYESIAN BRIDGE TRAFFIC LOADING MODEL

Bayesian statistics is underpinned by the use of Bayes rule:

$$P(\theta|x) = \frac{L(x|\theta)\, P(\theta)}{\int L(x|\theta)\, P(\theta) d\theta} \quad (1)$$

where $P(\theta|x)$ is the posterior distribution of the parameter θ given the observed data x (i.e., the estimate of θ), $\int L(\theta|x)P(\theta)d\theta$ is the normalization constant such that the posterior distribution sums up to 1, $L(x|\theta)$ is the likelihood function of the data given the parameter, and $P(\theta)$ is the prior distribution of the parameter.

The calculation of the posterior distribution $P(\theta|x)$ is difficult as the normalizing constant $\int L(x|\theta)P(\theta)d\theta$ is often mathematically intractable. In this paper, the No-U-Turn Sampler (NUTS) algorithm was used to compute the posterior distribution (Hoffman and Gelman 2014) as implemented in the python PyMC package (Salvatier et al. 2016).

This paper proposes a Bayesian bridge traffic loading model based on the Generalized Extreme Value (GEV) distribution that have been used by authors in the past (Caprani 2005; O'Brien et al. 2015) with the likelihood function:

$$L(x|\mu,\sigma,\xi) = \prod_{i=1}^{N} \frac{1}{\sigma} t(x_i|\mu,\sigma,\xi) e^{t(x_i|\mu,\sigma,\xi)} \quad (2)$$

$$t(x_i|\mu,\sigma,\xi) = (1 + \xi(x_i - \mu)/\sigma)^{-1/\xi} \quad (3)$$

The distribution is upper bounded when $\xi < 0$, lower bounded when $\xi > 0$, and support all real numbers from $-\infty$ to ∞ when $\xi = 0$. It has been recognized by multiple authors that bridge traffic loading is an upper bounded problem (O'Brien et al. 2015; O'Connor et al. 2002; van der Spuy and Francois 2020). As such, the proposed Hierarchical Bayesian model's prior distribution was selected such that ξ is strictly negative to ensure the existence of the upper bound. Further, a reparameterization of the GEV distribution with $\xi < 0$ was proposed to aid in assigning informative priors. The mean (E) and standard deviation (S) of GEV are:

$$E = \mu + \sigma(\Gamma(1 - \xi) - 1)/\xi \quad (4)$$

$$S = \sqrt{\sigma^2(\Gamma(1 - 2\xi) - \Gamma(1 - \xi)^2)/\xi^2} \quad (5)$$

where $\Gamma(\cdot)$ is the Gamma Function. The upper bound z_b of the distribution is:

$$z_b = \mu - \sigma/\xi \quad (6)$$

and the centered upper bound, z_{bc} defined as the upper bound of the GEV distribution that is scaled and translated such that the mean and standard deviation are zero and one respectively:

$$z_{bc} = (z_b - E)/S = [(\mu - \sigma/\xi\ ) - E]/S \quad (7)$$

The original location (μ) and scale (σ) parameters in terms of E and S (Equation 9-10) can then be obtained:

$$z_{bc} = \Gamma(1-\xi)/\sqrt{\Gamma(1-2\xi)-\Gamma(1-\xi)^2} \quad (8)$$

$$\mu = E + S\, z_{bc}(\Gamma(1-\xi)-1)/\Gamma(1-\xi) \quad (9)$$

$$\sigma = -\xi\, S\, z_{bc}/\Gamma(1-\xi) \quad (10)$$

The new reparameterization allows prior predictive checks (Gabry et al. 2019) against the sample mean and standard deviation, which would aid in assigning prior distributions.

Thus, the Bayesian bridge traffic loading model can be characterized as:

$$X \sim GEV(\mu_L, \sigma_L, \xi_L) \quad (11)$$

$$\mu_L = E_L + S_L\, z_{bc,L}\,(\Gamma(1-\xi_L)-1)/\Gamma(1-\xi_L) \quad (12)$$

$$\sigma_L = -\xi_L\, S_L\, z_{bc,L}/\Gamma(1-\xi_L) \quad (13)$$

$$z_{bc,L} = \Gamma(1-\xi_L)/\sqrt{\Gamma(1-2\xi_L)-\Gamma(1-\xi_L)^2} \quad (14)$$

where μ_L,σ_L,ξ_L are the GEV parameter for bridge of span L. Note that the standard deviation S_L are lower bounded to 0.001 to ensure positive value.

First, a basic Bayesian model, the Single-Span Bayesian (SSB) model is developed. In the SSB model, the parameter μ_L, σ_L, ξ_L are independent for each bridge span L. The prior distributions for the SSB model are:

$$E_L \sim \text{Normal}(\mu_{E_L}, \sigma_{E_L}^2) \quad (15)$$

$$S_L \sim \text{Normal}(\mu_{S_L}, \sigma_{S_L}^2) \quad (16)$$

$$\xi_L \sim \text{Negative Lognormal}(\mu_{\xi_L}, \sigma_{\xi_L}^2) \quad (17)$$

where the Negative Lognormal (NLN) distribution is simply the Lognormal distribution multiplied by negative one. The use of NLN prior restricts the ξ_L parameter to negative values, inline with the upper boundedness of real traffic.

The SSB model is expanded into the Hierarchical Bayesian (HB) model by introducing a hierarchical structural in the mean and standard deviation. Instead of separate E_L and S_L for each span, in the HB model the mean and standard deviation are a function of L. The following functions are proposed:

$$E_L = E(L) = \beta_{-1}L^{-1} + \beta_0 + \beta_1 L \quad (18)$$

$$S_L = S(L) = \alpha_{-1}L^{-1} + \alpha_0 + \alpha_1 L \quad (19)$$

$$\xi_L = \xi = -\exp(\gamma_0) \quad (20)$$

with the prior distributions:

$$\beta_i \sim \text{Normal}(\mu_{\beta_i}, \sigma_{\beta_i}^2) \quad (21)$$

$$\alpha_i \sim \text{Normal}(\mu_{\alpha_i}, \sigma_{\alpha_i}^2) \quad (22)$$

$$\gamma_0 \sim \text{Normal}(\mu_{\gamma_0}, \sigma_{\gamma_0}^2) \quad (23)$$

Note the ξ parameter in the HB model is constant across all spans and is strictly negative in line with the upper bounded property of traffic loading. The use of functions in terms of L for E and S introduces a hierarchy in the model which allows interpolation for any L outside of those used for fitting, and allows fitting of multiple bridge spans simultaneously. Thus, the proposed HB model can be used across a road network to simultaneously fit bridges with available traffic loading data, and infer traffic loading for the remainder of bridges of arbitrary spans without needing to collect data from those bridges.

3. APPLICATION TO SIMULATED TRAFFIC DATA

3.1. Traffic simulation description

A traffic simulation was performed with the aim of assessing the accuracy of the proposed model and the conventional MLE method against known true parameter values. The simulation generates 5 axle trucks (Table 2), with its axle weight and spacing distributed as a multi-modal beta distribution, with parameters in Table 1 and probability density function:

$$f(x) = \sum_{i=1}^{4} p_i \frac{x^{\alpha_i - 1}(1-x)^{\beta_i - 1}}{\mathrm{B}(\alpha_i, \beta_i)} \quad (24)$$

where p_i is the weight of mode i, α_i and β_i are the shape parameters of mode i, and $\mathrm{B}(\cdot)$ is the Beta Function. The distribution was then scaled such that the minimum and maximum is as in Table 2. Note that axle weight and spacings were generated independently. The multi-modality of the distribution mimics real trucks characteristics (Grave 2002), and the use of beta distribution ensures an upper and lower bound. The true upper

Table 1: Parameters of Multimodal Beta distribution.

Mode (i)	p_i	α_i	β_i
1	0.65	5	15
2	0.25	20	15
3	0.06	5	5
4	0.04	10	3

Table 2: 5 axle trucks used in simulation.

Axle Weight (kN)			Axle Spacing (m)		
Axle	Min	Max	Axle	Min	Max
1	68.4	102.6	1-2	0.77	1.43
2	108.0	162.0	2-3	3.50	6.50
3, 4, 5	46.8	70.2	3-4-5	2.10	3.90

bound of the load effect is then simply the truck with the heaviest axle weights and shortest axle spacings. Each truck was then individually run over a single lane simply supported bridge. The maximum bending moment at midspan (LE_M) and shear at both supports (LE_V) were calculated, and the monthly maximum for the bridge recorded. Ten simply supported bridges of span 10, 15, 20, 25, 27, 30, 35, 40, 45, and 50 m were simulated for a period of 3 years, consisting of 10 economic months each with 25 economic days (O'Brien et al. 2021). A random number of trucks were generated daily, with an average daily traffic of 10 000 trucks and a standard deviation of 50 trucks.

3.2. Prior selection

A prior predictive check (Gabry et al. 2019) (Figure 1) was performed to check the suitability of the proposed $E(L)$ and $S(L)$ functions for the HB model (Equation 18-19). The red lines were generated according to the priors in Table 3 and the $E(L)$ and $S(L)$ functions. Note that the observed data were not used directly to obtain the parameters in Table 3, rather trial and error was performed until the observed mean and standard deviation are covered by the prior predictive (red line). For both LE_M and LE_V, the proposed priors and functions adequately covered the sample mean and standard deviation of the sampled monthly maxima.

3.3. Results

The proposed models were fitted to both LE_V and LE_M data. However, for conciseness, only parameter estimates from LE_V and parameter predictions from LE_M will be discussed.

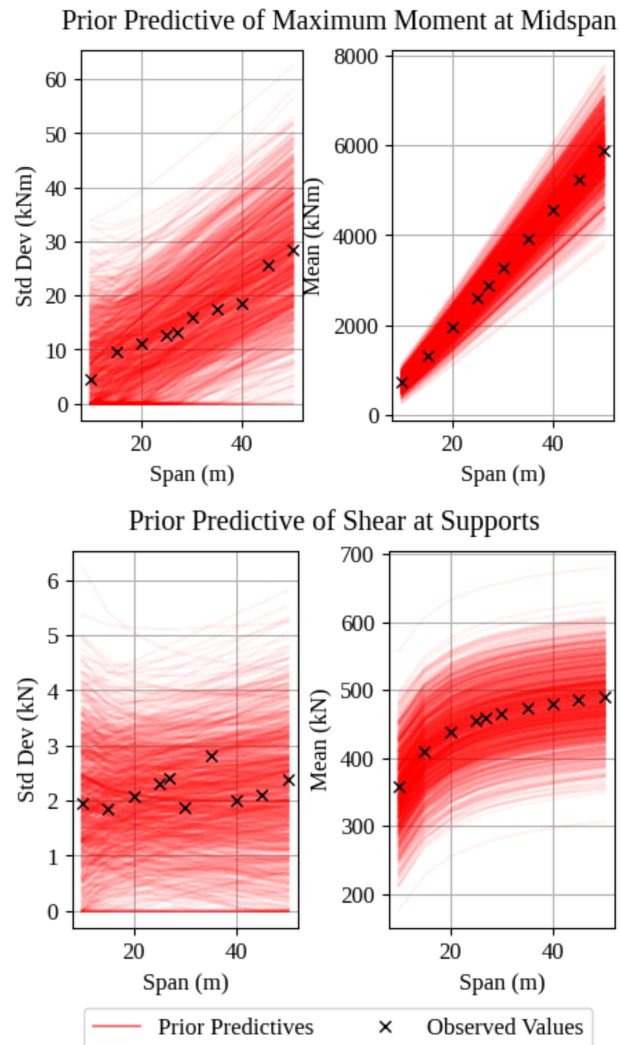

Figure 1:Prior predictive checks of HB model.

Table 3: Prior of HB model.

Maximum Shear at Supports					
	μ_θ	σ_θ		μ_θ	σ_θ
β_{-1}	-1600	160	α_{-1}	0.38	7.5
β_0	510	5.1	α_0	1.90	0.75
β_1	0.15	0.025	α_1	0.007	0.014
γ_0	-3.0	0.45			
Maximum Moment at Midspan					
	μ_θ	σ_θ		μ_θ	σ_θ
β_{-1}	1300	150	α_{-1}	13.1	65.0
β_0	-725	72.5	α_0	-2.25	6.5
β_1	130	13.0	α_1	0.60	0.15
γ_0	-3.0	0.45			

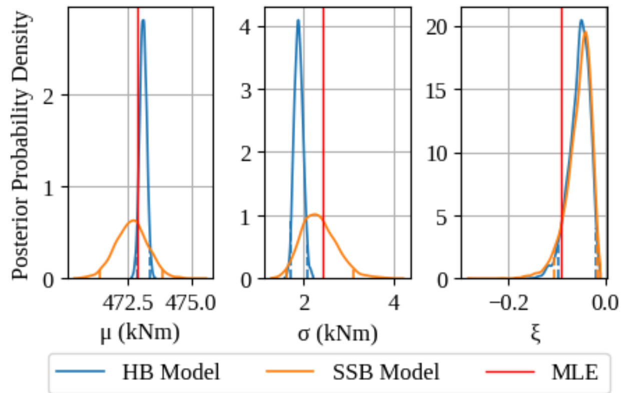

Figure 2: Marginal posterior density of GEV parameters of maximum shear at supports for 35 m simply supported bridge.

3.3.1. *Estimated parameters of maximum shear at supports*

Figure 2 shows the estimated GEV parameter posterior distribution of LE_V for a 35 m bridge by the HB and SSB model, with the MLE estimate in red vertical line. Both Bayesian models provided the full parameter distribution, an advantage over the conventional MLE method. Moreover, it can be seen that the use of data from multiple bridge spans simultaneously in the HB model creates lower uncertainties around the parameter estimates compared to the SSB model. This reduced uncertainty – or *shrinkage*, is the result of pooling of data into groups (i.e., bridge spans) in the Hierarchical model (McElreath 2020).

Figure 3 illustrates the posterior distributions and point estimates across 3 other spans. For brevity, the marginal distribution in Figure 2 is represented by the forest plot in Figure 3. The dot represents the mean of the estimate, the thick line is the interquartile range, and the thin line is the 95% highest density interval (HDI). Across all spans, estimates provided by HB, SSB and MLE are close to each other, often with overlapping HDI and point estimate, suggesting good agreement across all models. The true upper bounds were captured by both HB and SSB models, indicating good accuracy of the proposed Bayesian model. Further, the shrinkage effect provides lower estimate uncertainties for the HB model compared to the SSB model across all

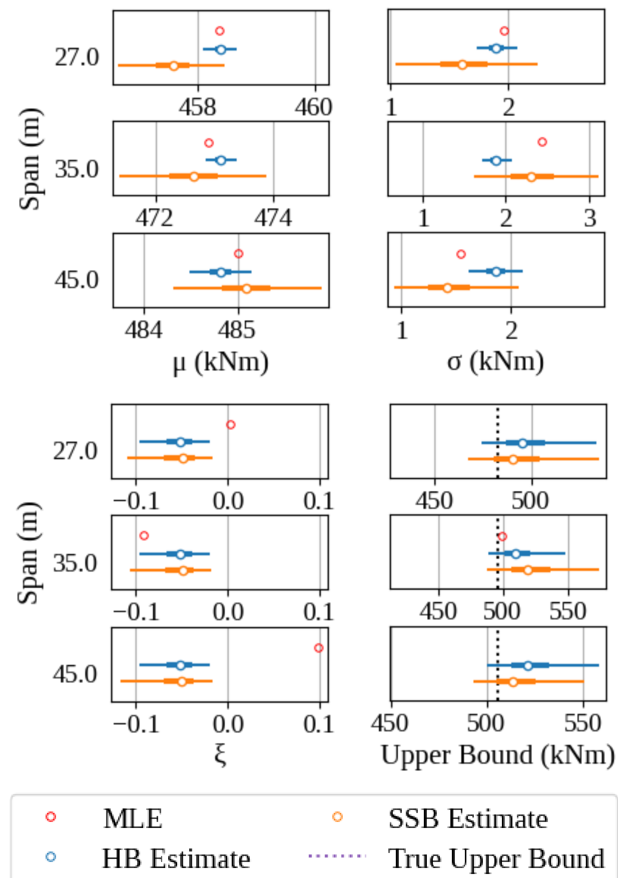

Figure 3:Estimated GEV parameters and upper bound. The dot is the mean of estimate, the thick line is the interquartile range (Q1 – Q3), and the thin line is the 95% highest density interval. The upper bound estimate is infinity for MLE when $\xi \geq 0$.

spans and parameters, illustrating the ability of the HB model to use more information across multiple spans to reduce uncertainty compared to the SSB model.

Additionally, note from Figure 3 the ξ parameter estimated by MLE is above zero for span 27 and 45 m, suggesting an infinite upper bound despite the deliberate upper boundedness set in the traffic simulation. In contrast, the proposed Bayesian models do not suffer from this problem as the prior was set to cover strictly negative values. As real traffic has been shown to be an upper bounded problem (O'Brien et al. 2015; O'Connor et al. 2002; van der Spuy and Francois 2020), the use of MLE could provide a misleading inference as it could imply an unbounded maximum load effect despite the upper bound that must exist in the physical process that is real traffic.

Furthermore, as the maximum load effect between different spans were produced by similar traffic, a degree of correlation in the maximum load effect (Bocchini and Frangopol 2011), and hence the GEV parameters should exist between different bridge spans. Table 4 shows the correlation coefficient of the μ and σ parameter posterior densities as estimated by HB model (lower triangular) and SSB model (upper triangular) across different spans. Correlations between different spans are correctly picked up by the HB model as multiple bridge spans were fitted simultaneously. Of note is the decreasing correlation as the difference in span increases. For example, a bridge of span 45 m has a strong correlation of 0.737 and 0.892 in the μ and σ parameter with another bridge of span 35 m but shows almost zero correlation with a bridge of span 15 m. In contrast, as the SSB model estimated each bridge in isolation, it is unable to detect this relationship between bridges, and estimated close to zero correlation across all parameters of all spans.

Table 4: Correlation coefficient in estimated μ and σ parameters between different bridge spans. The lower triangular is correlation from HB model, and the upper triangular is correlation from SSB model.

Correlation coeff. of μ posterior distribution

Span (m)	*15.0*	*27.0*	*35.0*	*45.0*
15.0	-	-0.015	-0.038	-0.018
27.0	0.554	-	-0.002	0.012
35.0	0.264	0.864	-	-0.03
45.0	-0.090	0.310	0.737	-

Correlation coeff. of σ posterior distribution

Span (m)	*15.0*	*27.0*	*35.0*	*45.0*
15.0	-	0.02	-0.032	-0.002
27.0	0.651	-	-0.013	0.004
35.0	0.277	0.868	-	0.005
45.0	-0.053	0.557	0.892	-

3.3.2. Prediction of intermediate spans

To examine the predictive capability of the HB model, the bridge spans were grouped into two: a fitting group and an intermediate spans group (Table 5). The HB model was fitted on the fitting group, and predictions made on the intermediate

Table 5: Fitting and intermediate span groups.

Group	*Fitting*	*Intermediate*
Span (m)	10,20,25,30,40,50	27,35,45

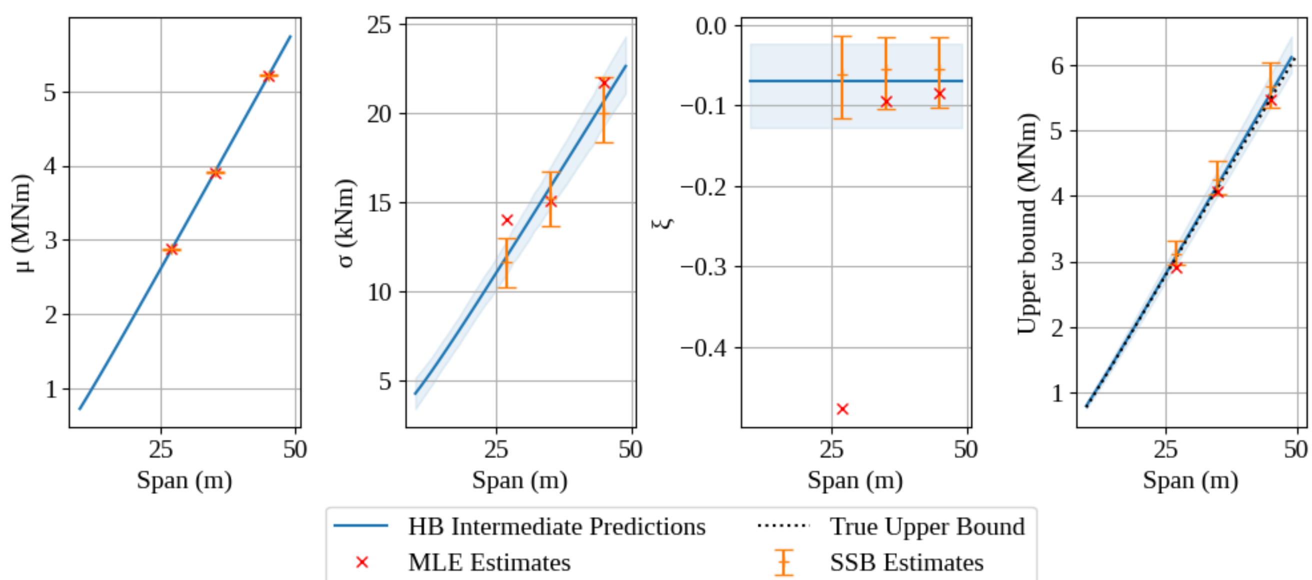

Figure 4: Predicted GEV parameters by the HB model and the estimated parameter by SSB and MLE. Light blue area indicates 95% highest density interval prediction made by HB model. Note that the spans fitted for SSB and MLE were never used for fitting by the HB model.

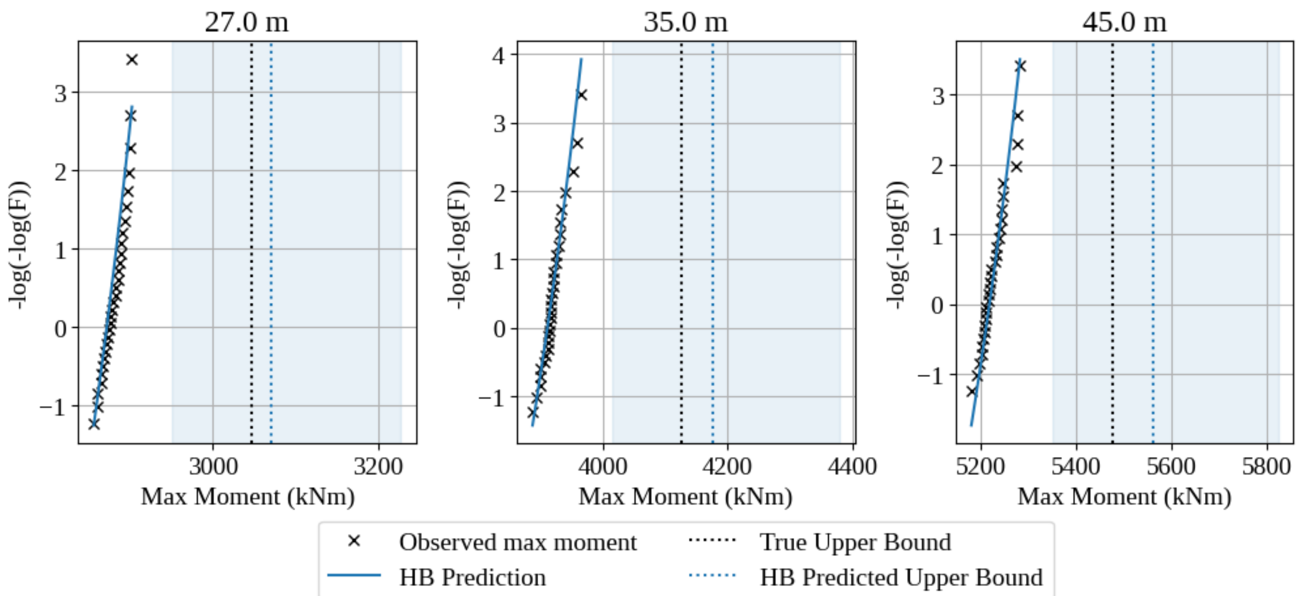

Figure 5: Gumbel probability paper plot of predicted intermediate spans by the HB model. Light blue area indicates 95% highest density interval of predicted upper bound.

span group. The results of the prediction by the HB model are then compared against GEV parameters estimated by SSB and MLE models fitted on the intermediate spans.

Figure 4 shows the GEV parameters predicted by the HB model across intermediate spans between 10 m and 50 m. The results of the HB prediction show good agreement with the SSB and MLE estimates fitted on the intermediate span group, with the 95% HDI and MLE point estimates often overlapping, despite the fact that the HB model never saw the intermediate spans during the fitting process.

Figure 5 shows the predicted maximum load effect for the intermediate spans by the HB model, and the observed sample on Gumbel probability paper. The result shows good agreement with the observed data. Moreover, the true upper bound were captured by the predicted upper bound 95% HDI (light blue region) across all 3 intermediate spans, showing predictions made by the HB model is accurate.

4. CONCLUSION

In this paper, a Hierarchical Bayesian model for bridge traffic loading was developed by introducing a functional relationship between the mean and standard deviation of different bridge spans. The No-U-Turn Sampler was used to sample the posterior distribution of the model. Information from past studies indicating the upper bounded nature of traffic loading was used to limit the shape parameter to strictly negative values, creating an engineering informed prior, and a reparameterization of the Generalized Extreme Value distribution is proposed to allow prior predictive checks to be performed against sample mean and standard deviation, allowing a better informed priors compared to non-informative priors used in past studies.

The Hierarchical Bayesian model with the informed priors is shown to be accurate when tested against 3 years simulated traffic across 10 different spans, ranging from 10 m to 50 m. The model accurately estimated and predicted the true upper bound of the traffic load for any arbitrary bridge spans. The results are also in good agreement with conventional methods while offering the following advantages: (1) Bayesian estimation offers full posterior distribution as opposed to point estimate from Maximum Likelihood Estimation, (2) simultaneous fitting of multiple bridge spans results in lower uncertainties for parameter estimate due to shrinkage, (3) correlation between parameters from different spans are obtained in the posterior distributions, and (4) allows accurate prediction of intermediate bridge spans not used during the fitting process.

Given the above advantages, the Hierarchical Bayesian model provides a framework for further studies on extreme bridge traffic loading across a road network, where weigh-in-motion data from multiple bridges across the network can be combined, parameters fit simultaneously, and predictions made on bridges where weigh-in-motion data are not available. This provides a basis for estimating load effects across an entire road network at once; something not previously feasible.

5. REFERENCES

Bocchini, P., and D. M. Frangopol. 2011. "Generalized bridge network performance analysis with correlation and time-variant reliability." *Structural Safety*, 33 (2): 155–164.

Caprani, C. C. 2005. "Probabilistic Analysis of Highway Bridge Traffic Loading." Dublin, Ireland: University College Dublin.

Coles, S. 2001. *An introduction to statistical modeling of extreme values*. Springer series in statistics. London ; New York: Springer.

Crespo-Minguillón, C., and J. R. Casas. 1997. "A comprehensive traffic load model for bridge safety checking." *Structural Safety*, 19 (4): 339–359

Dai, B., Y. Xia, and Q. Li. 2022. "An extreme value prediction method based on clustering algorithm." *Reliability Engineering & System Safety*, 222: 108442.

Gabry, J., D. Simpson, A. Vehtari, M. Betancourt, and A. Gelman. 2019. "Visualization in Bayesian workflow." *J. R. Stat. Soc. A*, 182 (2): 389–402.

Grave, S. 2002. "Modelling of site-specific traffic loading on short to medium span bridges." thesis. Trinity College (Dublin, Ireland). Department of Civil, Structural and Environmental Engineering.

Hoffman, M. D., and A. Gelman. 2014. "The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo."

McElreath, R. 2020. *Statistical Rethinking: A Bayesian Course with Examples in R and STAN*.

Nesterova, M., F. Schmidt, and C. Soize. 2019. "Probabilistic analysis of the effect of the combination of traffic and wind actions on a cable-stayed bridge." *Bridge Structures*, 15 (3): 121–138. IOS Press.

O'Brien, E. J., A. S. Nowak, and C. C. Caprani. 2021. *Bridge Traffic Loading: From Research to Practice*. CRC Press.

O'Brien, E. J., F. Schmidt, D. Hajializadeh, X.-Y. Zhou, B. Enright, C. C. Caprani, S. Wilson, and E. Sheils. 2015. "A review of probabilistic methods of assessment of load effects in bridges." *Structural Safety*, 53: 44–56.

O'Connor, A., C. Caprani, and A. Belay. 2002. "Site-Specific Probabilistic Load Modelling for Bridge Reliability Analysis." *Proceedings of the 3rd. International Conference on Current and Future Trends in Bridge Design*, 8. Dublin, Ireland.

Salvatier, J., T. V. Wiecki, and C. Fonnesbeck. 2016. "Probabilistic programming in Python using PyMC3." *PeerJ Comput. Sci.*, 2: e55. PeerJ Inc.

van der Spuy, P., and P. Francois. 2020. "Derivation of a traffic load model for the structural design of highway bridges in South Africa." Thesis. Stellenbosch : Stellenbosch University.

Yu, Y., C. S. Cai, W. He, and H. Peng. 2019. "Prediction of bridge maximum load effects under growing traffic using non-stationary bayesian method." *Engineering Structures*, 185: 171–183.