

Estimating Characteristic Bridge Traffic Load Effects Using Bayesian Statistics

Cathal Leahy

Research Engineer, Roughan & O'Donovan Innovative Solutions, Dublin, Ireland

Eugene J. OBrien

Professor of Civil Engineering, University College Dublin, Dublin, Ireland

Bernard Enright

Lecturer, Dublin Institute of Technology, Dublin, Ireland

Richard Power

Corporate Secretary, Roughan & O'Donovan Innovative Solutions, Dublin, Ireland

ABSTRACT: This paper investigates the use of Bayesian updating to improve estimates of characteristic bridge traffic loading. Over recent years the use Weigh-In-Motion technologies has increased hugely. Large Weigh-In-Motion databases are now available for multiple sites on many road networks. The objective of this work is to use data gathered throughout a road network to improve site-specific estimates of bridge loading at a specific Weigh-In-Motion site on the network. Bayesian updating is a mathematical framework for combining prior knowledge with new sample data. The approach is applied here to bridge loading using a database of 81.6 million truck records, gathered at 19 sites in the US. The database represents the prior knowledge of loading throughout the road network and a new site on the network is simulated. The Bayesian approach is compared with a non-Bayesian approach, which uses only the site-specific data, and the results compared. It is found that the Bayesian approach significantly improves the accuracy of estimates of 75-year loading and, in particular, considerably reduces the standard deviation of the error. With the proposed approach less site-specific WIM data is required to obtain an accurate estimate of loading. This is particularly useful where there is concern over an existing bridge and accurate estimates of loading are required as a matter of urgency.

1. INTRODUCTION

Accurate modeling of bridge traffic loading is critically important in bridge engineering. At the design stage it allows the design of bridges which are fit for purpose, while reducing the waste associated with overdesign. In the assessment of existing structures, estimates of traffic loading are possibly more important. Where there is concern over an existing bridge, this information will determine whether a bridge needs to be repaired or replaced. If a bridge is saved as a result, significant cost savings can be made.

The most accurate method for modelling bridge traffic loading is to use Weigh-In-Motion (WIM) data. WIM systems measure truck weights and axle configurations as they pass along a road at normal highway speeds. Statistical methods are generally used to extrapolate from the relatively short WIM measurement period to the return period used for bridge design/assessment. In Europe, a return period of 1000 years is used for bridge design (EC1 2003) and a design life of 75 years is used in the US (AASHTO 2012).

A common extrapolation approach is to fit a statistical distribution to the measured data and to extrapolate to the required return period. The

generalized extreme value distribution is a popular choice in the literature. This distribution can be separated into a family of three extreme value distributions, comprising of the Gumbel (type I), Fréchet (type II) and Weibull (type III) distributions. The type I Gumbel distribution is used by some authors (Cooper 1997; O'Connor and Eichinger 2007) but the type III Weibull distribution is perhaps the most common (Bailey and Bez 1999; Miao and Chan 2002; OBrien et al. 2010). These extreme value distributions must be fitted to block maxima, with maximum daily or weekly values often used.

As an alternative to direct extrapolation from WIM data, Monte Carlo simulations can also be used. With these simulations, statistical distributions are fitted to the properties of the measured traffic and a new stream of traffic is simulated (Enright and OBrien 2013; Sivakumar et al. 2008). Hundreds or thousands of years of traffic can be simulated in this way and the characteristic load effects can be obtained directly from the simulated traffic, without the need for extrapolation. These long run simulations are time consuming but they allow for loading scenarios, such as truck meeting or overtaking events, which did not occur during the WIM measuring period.

WIM data is becoming increasingly more common and this is reflected in recent bridge loading literature. Enright et al. (2013) used 21 million trucks from 11 sites to examine long span bridge loading in Alabama. Nowak & Rakoczy (2012) simulated bridge loading using 35 million truck records from 32 sites in the US and Fu & You (2011) examine multi-truck loading events using 63 million truck records gathered in New York.

This paper develops a method for using the large amounts of data which are now available for many road networks to help estimate loading at a new site on the same network. Bayesian updating is used as it provides a mathematical framework for combining prior knowledge about a certain statistical property with new sample data (Ang and Tang 2007). This is useful in

engineering where data from various sources must be combined in the decision-making process and has been used for many civil engineering applications. For example, Hong and Prozzi (2006) use this approach to update the parameter distributions for a pavement performance model. Enright and Frangopol (1999) use it to predict the deterioration of concrete bridges as it allows the results of bridge inspections to be combined with prior information to obtain improved predictions. Other studies (Beck and Au 2002; Igusa et al. 2002) also apply a Bayesian approach to improve the accuracy of structural engineering models.

In this work 81.6 million trucks from 19 sites in the US are used to test the approach. These truck records provide the prior knowledge about loading in the US. This prior data is then combined with new site-specific data, using the Bayesian approach, in order to obtain improved estimates of characteristic loading at the new site. The Bayesian approach is compared with a Weibull extrapolation approach, which uses only site-specific data, and the improvements in accuracy examined. The proposed approach reduces the amount of WIM data required to make accurate estimates of loading. This is very useful where there is concern over the safety of an existing bridge and an estimate of loading is required in order to make a prompt decision about the future of the bridge.

2. METHODOLOGY

2.1. Bayesian updating

A random variable (maximum weekly GVW) is described by a parameter vector θ (e.g. for a Weibull distribution with three parameters, $\theta = (\mu, \sigma, \xi)$). However, these parameters are themselves random variables due to the uncertainty associated with parameter estimation. These random variables have an assumed 'prior' distribution $f'(\theta)$. When a new set of sample data, x , is acquired this prior distribution can be updated to obtain the 'posterior' distribution $f''(\theta)$. The posterior distribution is calculated as:

$$f''(\theta) = \frac{P(x|\theta)f'(\theta)}{\int_{-\infty}^{\infty} P(x|\theta)f'(\theta)d\theta} \quad (1)$$

where $P(x|\theta)$ = conditional probability, or likelihood $L(\theta)$, of the new data x , given θ .

The denominator is a normalizing constant which ensures that $f''(\theta)$ is a complete probability density function (PDF). As a result Eq. (1) can be rewritten as:

$$f''(\theta) = kL(\theta)f'(\theta) \quad (2)$$

This is the approach used here as it avoids the computationally intensive calculation of the denominator $\int_{-\infty}^{\infty} P(x|\theta)f'(\theta)d\theta$. The likelihood function is calculated for the new set of data as:

$$L(\theta) = \prod_{i=1}^n f(x_i; \theta) \quad (3)$$

The posterior $f''(\theta)$ is often used to obtain the expected value of θ , which gives a point estimator of the parameter vector. However the approach used in this work includes the uncertainty associated with estimating the value of θ . The posterior distribution of the parameter vector can be used to calculate the cumulative distribution function (cdf) for the underlying random variable X (in this case the maximum weekly GVW), and this cdf can be used to calculate characteristic values:

$$P(X \leq z) = \int P((X \leq z)|\theta) f''(\theta) d\theta \quad (4)$$

The prior distribution of the parameter vector is obtained by fitting a Weibull distribution to each site-year of prior data. The data from the Colorado site is used to simulate

the new site-specific data and is not included in the prior data. This new site-specific data is modelled by randomly sampling 25 non-consecutive weeks of data from the Colorado site. Bayesian updating is then used obtain the posterior distribution of the parameter vector and hence the cdf for loading at that site. For the Bayesian approach, the Weibull distribution must be fitted to the entire distribution, rather than the tail of the maximum daily values, which is sometimes used (OBrien et al. 2010). With the aim of examining the trend in this tail region, the Weibull distribution is fitted to maximum weekly GVWs, rather than maximum daily. GVW is used here to assess the proposed method but the same approach could easily be applied to bridge load effects required for assessment.

The Generalized Extreme Value (GEV) distribution is given by:

$$F(x; \mu, \sigma, \xi) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\} \quad (5)$$

defined on $\{x : 1 + \xi(x - \mu)/\sigma > 0\}$, where:

μ = location parameter, $-\infty < \mu < \infty$

σ = scale parameter, $\sigma > 0$

ξ = shape parameter

The Weibull distribution is a subset of the GEV distribution, where the shape parameter ξ is negative. The location and scale parameters are similar to mean and standard deviation and the shape parameter describes the behaviour in the tail of the distribution.

2.2. WIM data

The prior data used here is WIM data from the United States Federal Highway Administration's Long-Term Pavement Performance (LTPP) Program. Initially the LTPP collected WIM data with inconsistent quality control measures (Walker and Cebon 2012). In 1999 a plan was developed which, among other things, improved and centralized quality control. This led to a significant improvement in WIM

data reliability and since 2003 ‘research quality’ WIM data is being collected at 28 of the Specific Pavement Studies LTPP sites. The 81.6 million truck records were collected over the period 2005-2012 and are detailed in Table 1.

Table 1. WIM Data

Site	State	Years Data	Trucks/ Weekday	MMW GVW ²
1	Arizona	5.7	575	63
2	Arizona	5.7	4,988	93
3	Arkansas	6.0	5,526	100
4	California	5.0	5,939	68
5 ¹	Colorado	6.5	1,473	73
6	Delaware	5.5	930	66
7	Illinois	7.4	3,139	98
8	Indiana	4.5	1,489	69
9	Kansas	6.6	1,851	81
10	Louisiana	4.9	506	77
11	Maine	4.4	835	67
12	Maryland	6.8	1,030	51
13	Minnesota	6.2	316	67
14	New Mexico	4.7	716	61
15	New Mexico	4.7	2,934	89
16	Pennsylvania	5.6	5,315	90
17	Tennessee	5.6	5,474	89
18	Virginia	6.0	1,082	63
19	Wisconsin	5.3	987	73

¹Site used to test the Bayesian approach

² MMW GVW: mean maximum weekly gross vehicle weight (a good measure of intensity of loading at a site (OBrien and Enright 2012))

Bridges in the US are designed and assessed using the HL-93 load model (AASHTO 2012). This model applies to “normal vehicular use” (all legal trucks, illegal overloads and un-analyzed routine permits) and does not include special permit vehicles. In order to describe a methodology which can be used to examine “normal vehicular use”, special permit vehicles are removed from the database. These vehicles are identified in the WIM data using filtering rules, as described by Enright et al. (2015).

2.3. Estimating the prior distribution using kernel density estimators (KDE)

The Colorado site is used as the test data for the Bayesian approach. The other 18 sites are used as the prior data. Four years of data are used from each site as this amount is available for each site and gives each site an equal weighting. A Weibull distribution is fitted to the maximum weekly GVWs for each site-year of prior data – see Figure 1. This gives 72 sets of Weibull parameters (μ, σ, ξ). These parameter values are shown in Figure 2. In this figure, values of σ are plotted against μ for different ranges of the shape factor, ξ .

Figure 1 also shows the Colorado site used to test the Bayesian approach. It can be seen that the loading at the Colorado site agrees well with the general trend of the prior data. As a result it is expected that the proposed method will perform well with this test site.

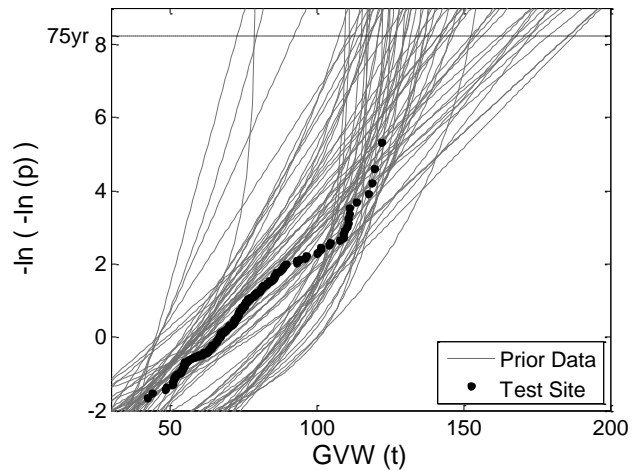


Figure 1. Weibull distributions of the prior data alongside the Colorado site used to test the proposed approach.

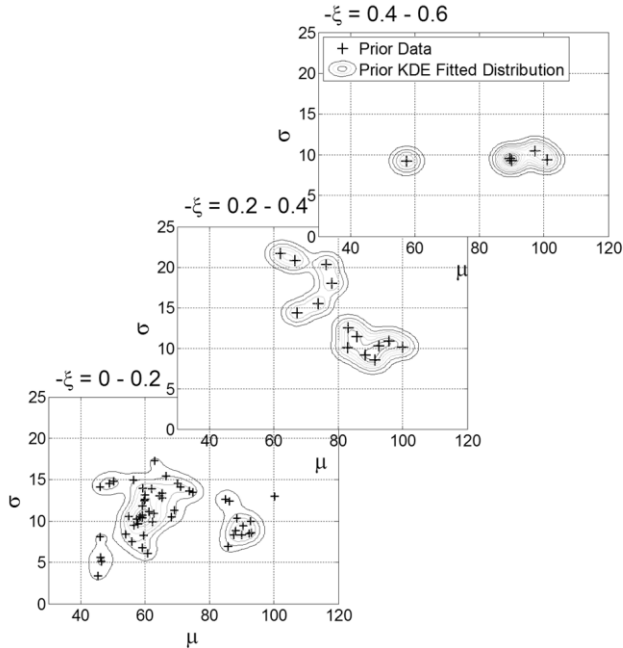
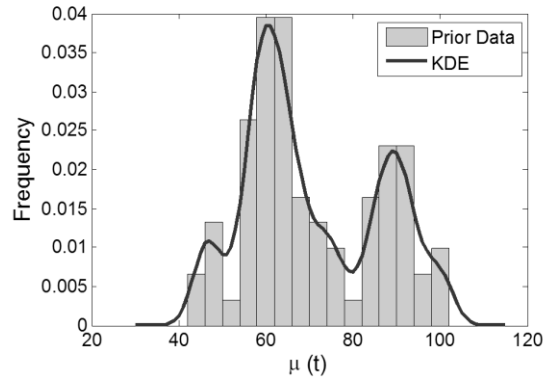


Figure 2. Prior data used with kernel density estimated distribution for different ranges of shape parameter ξ .

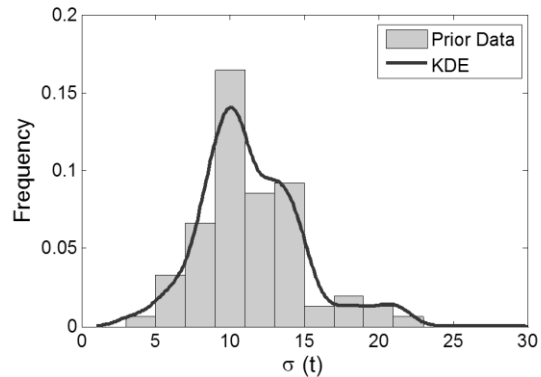
A continuous distribution, rather than discrete points, is needed for the prior distribution. In order to fit a distribution to the 72 data points, kernel density estimators (KDEs) are used. With the KDE method, each data point is replaced with a component density. These densities are known as kernel functions and are added to obtain an estimate of the overall distribution. Any distribution can be used for the kernel functions but the normal distribution is a popular choice and is used here. The product kernel method, as described by Scott (1992), is used. The kernel function is essentially a trivariate normal distribution where all variables are independent. The estimated distribution is shown in Figure 2. For visualization purposes, the distribution is plotted for different ranges of the shape parameter ξ .

In the Bayesian calculations, the weighting of the prior data, with respect to the new data, is determined by the bandwidth of the kernel function. The bandwidth, in this case, refers to the variance of the trivariate normal distribution, where each of the three variables has its own independent standard deviation. If the bandwidth

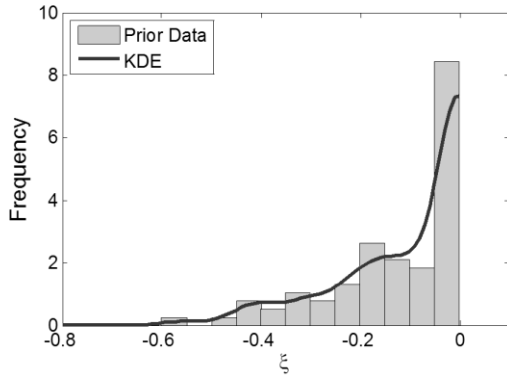
is too large, over-smoothing will occur and the trends in the prior data will be lost. On the other hand, if the bandwidth is too small the posterior distribution will be restricted to match very closely to the prior data. Optimal bandwidths exist for sample data which matches a certain theoretical distribution, but this is not the case here. In this work a trial and error approach, as recommended by Silverman (1986), is used and an optimal value selected by visual inspection of the marginal distributions – see Figure 3. Bandwidth values of three tonnes, one tonne and 0.04 are used for μ , σ and ξ , respectively. The selection of these values is somewhat subjective but when applying this approach bandwidth selection should be carefully considered as it does influence the results.



(a) Location parameter



(b) Scale parameter



(c) Shape parameter

Figure 3. Histograms of marginal distributions of prior data with the kernel density estimated distribution.

2.4. Assessing accuracy of proposed Bayesian approach

The Bayesian approach is assessed to determine its accuracy in estimating the 75-year load. 75 years is the design life used in US bridge design (AASHTO 2012). The benchmark 75-year load is first calculated by fitting a Weibull distribution to all the data available at the Colorado site (6.5 years) and extrapolating to the 75-year load – see Figure 4. This approach assumes that this extrapolation from 6.5 year of data represents the true 75-year loading.

The accuracy of the method is assessed by randomly selecting 25 non-consecutive maximum weekly values from the Colorado data. The Bayesian approach, using the prior data, is then applied to this sample dataset and the 75-year load estimated. The accuracy of the estimate is then determined by comparison with the benchmark.

A non-Bayesian approach, where no prior data is used, is then applied to the same sample data. A Weibull distribution is fitted directly to the 25-week sample. The accuracy in estimating the 75-year value can then be compared with the Bayesian approach – see Figure 4. This process is then repeated for 100 different random datasets in order to assess the standard deviation of the error.

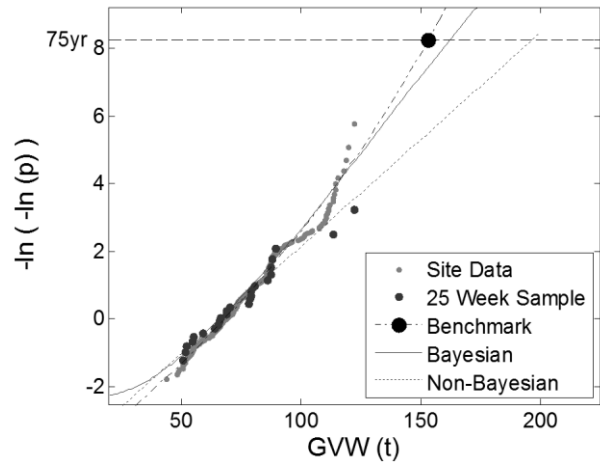


Figure 4. Example of the Bayesian and Weibull approach applied to a 25-week sample from the Colorado site (“Site data” includes the weekly maxima from 6.5 years of measurements at the site).

3. RESULTS

The errors for both the Bayesian and non-Bayesian approaches are shown in Figure 5 and Table 2. A mean error of zero with a small standard deviation is preferable. It can be seen for the non-Bayesian approach, where no prior data is used, that there is large variation in the errors. This variation is not desirable when estimating characteristic loading on bridges as it may result in an estimate of loading which is significantly greater or less than the true value. This variation is significantly reduced with the Bayesian approach, with the standard deviation reducing from 19.3% to 4.4%. The mean error is also improved from 7.8% to 2.3%.

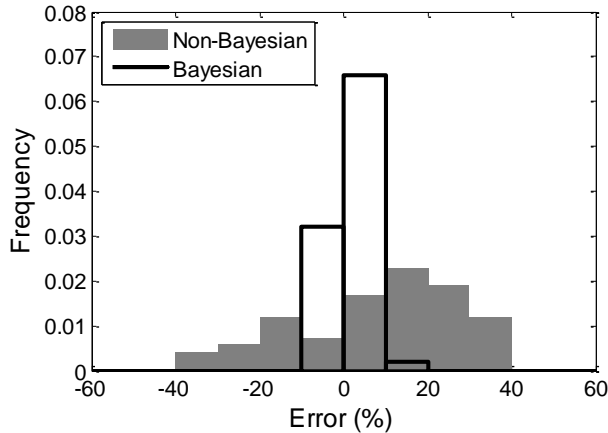


Figure 5. Errors in estimating 75-year load for the Bayesian and non-Bayesian approaches, for 100 sample datasets.

Table 2. Error with Bayesian and non-Bayesian approaches

	Mean	Standard Deviation
Bayesian	2.3%	4.4%
Non-Bayesian	7.8%	19.3%

4. CONCLUSION

A Bayesian methodology is proposed for obtaining additional value from the large amount of WIM data which are available for multiple sites on many road networks. The approach uses this existing, or ‘prior’, network data to help predict site-specific loading at a new site on the network. The prior data used here is from 18 WIM sites in the US. This WIM data is used to determine the prior distribution of the parameters of a Weibull distribution and this prior distribution is updated using the new site-specific data. The Bayesian approach is compared with the alternative of fitting a Weibull distribution directly to the site-specific data without using any prior data.

The errors in estimating characteristic loading are compared for the proposed Bayesian approach and the non-Bayesian approach. The Bayesian approach achieves significant reductions in the variation of the error while also improving the mean of the error. The proposed

method allows accurate estimates of loading to be achieved with less site-specific WIM data. This is particularly useful where there is concern over the safety of an existing bridge and accurate estimates of loading are required to make a prompt decision about the future of the bridge.

The method performs well with the Colorado site used to test it. However, this site has similar loading to the general trend in the prior data. It is possible that the Bayesian approach may not perform as well at a site with unusually low or high levels of loading as the method may tend to push estimates of loading closer to that of an average site.

The proposed Bayesian approach examined here used a large WIM database as the prior data. The WIM measurements were gathered at 18 WIM sites which were subject to strict quality control measures. The results of the method depend on the extent of good quality prior data which is available. It is acknowledged that such a large database of good quality prior data may often not be available when assessing bridge traffic loading and that the improvement in accuracy may not be as significant if less prior data is available. It is also noted that the Bayesian approach does not consider future traffic growth and further work may be needed to allow for such growth in the method.

ACKNOWLEDGMENTS

The research presented in this paper was carried out as part of the CEDR Transnational Road Research Programme Call 2013. The funding for the research was provided by the national road administrations of Denmark, Germany, Ireland, Netherlands, UK and Slovenia. The authors also gratefully acknowledge the Federal Highway Administration’s Long-Term Pavement Performance Program for access to an extensive WIM database and the Irish National Roads Authority’s Fellowship Programme for its financial support.

5. REFERENCES

- AASHTO. (2012). *AASHTO LRFD Bridge Design Specifications*. American Association of State Highway and Transportation Officials, Washington D.C.
- Ang, A. H.-S., and Tang, W. H. (2007). *Probability concepts in engineering: emphasis on applications in civil & environmental engineering*. Wiley, New York.
- Bailey, S. F., and Bez, R. (1999). "Site specific probability distribution of extreme traffic action effects." *Probabilistic engineering mechanics*, Elsevier, 14(1), 19–26.
- Beck, J. L., and Au, S.-K. (2002). "Bayesian updating of structural models and reliability using Markov chain Monte Carlo simulation." *Journal of Engineering Mechanics*, American Society of Civil Engineers, 128(4), 380–391.
- Cooper, D. I. (1997). "Development of short span bridge-specific assessment live loading." *Safety of Bridges*, P. C. Das, ed., Thomas Telford, London, 64–89.
- EC1. (2003). *Eurocode 1: Actions on structures, Part 2: Traffic loads on bridges. European Standard EN 1991-2:2003*. European Committee for Standardization, Brussels.
- Enright, B., Carey, C., and Caprani, C. C. (2013). "Microsimulation Evaluation of Eurocode Load Model for American Long-Span Bridges." *Journal of Bridge Engineering*, American Society of Civil Engineers, 18(12), 1252–1260.
- Enright, B., and O'Brien, E. J. (2013). "Monte Carlo simulation of extreme traffic loading on short and medium span bridges." *Structure and Infrastructure Engineering*, Taylor & Francis, 9(12), 1267–1282.
- Enright, B., O'Brien, E. J., and Leahy, C. (2015). "Identifying and modelling permit trucks for bridge loading." *Published online ahead of print in Proceedings of the ICE - Bridge Engineering*.
- Enright, M. P., and Frangopol, D. M. (1999). "Condition prediction of deteriorating concrete bridges using Bayesian updating." *Journal of Structural Engineering*, American Society of Civil Engineers, 125(10), 1118–1125.
- Fu, G., and You, J. (2011). "Truck load modeling and bridge code calibration." *Applications of Statistics and Probability in Civil Engineering*, Faber, Kohler, and Nishijima, eds., Taylor & Francis Group, London, 406–413.
- Hong, F., and Prozzi, J. A. (2006). "Estimation of pavement performance deterioration using Bayesian approach." *Journal of infrastructure systems*, American Society of Civil Engineers, 12(2), 77–86.
- Igusa, T., Buonopane, S. G., and Ellingwood, B. R. (2002). "Bayesian analysis of uncertainty for structural engineering applications." *Structural Safety*, Elsevier, 24(2), 165–186.
- Miao, T. J., and Chan, T. H. T. (2002). "Bridge live load models from WIM data." *Engineering Structures*, Elsevier, 24(8), 1071–1084.
- Nowak, A., and Rakoczy, P. (2012). "WIM Based Simulation Model Of Site Specific Live Load Effect On The Bridges." *6th International Conference on Weigh-In-Motion*, B. Jacob, A. H. McDonnell, F. Schmidt, and W. Cunagin, eds., Dallas, 352–358.
- O'Connor, A., and Eichinger, E. (2007). "Site-specific traffic load modelling for bridge assessment." *Proceedings of The Institution of Civil Engineers. Bridge Engineering*, 160(4), 185–194.
- O'Brien, E. J., and Enright, B. (2012). "Using Weigh-In-Motion Data to Determine Aggressiveness of Traffic for Bridge Loading." *Journal of Bridge Engineering, ASCE*, 18(3), 232–239.
- O'Brien, E. J., Enright, B., and Getachew, A. (2010). "Importance of the tail in truck weight modelling for bridge assessment." *ASCE Journal of Bridge Engineering*, 15(2), 210–213.
- Scott, D. W. (1992). *Multivariate density estimation: theory, practice, and visualization*. John Wiley & Sons.
- Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. Chapman & Hall, London.
- Sivakumar, B., Ghosn, M., and Moses, F. (2008). *Protocols for Collecting and Using Traffic Data in Bridge Design - NCHRP Report 683*. Transport Research Board, Washington D.C.
- Walker, D., and Cebon, D. (2012). "The Metamorphosis of LTPP Traffic Data." *6th International Conference on Weigh-In-Motion*, B. Jacob, A. H. McDonnell, F. Schmidt, and W. Cunagin, eds., Dallas, 242 – 249.