Journal of Southeast University (English Edition) Vol. 42, No. 1, pp. 65-73 Mar. 2026 ISSN 1003-7985

Relationship between the extreme value distribution of bending moments and traffic characteristics for simply supported bridges based on WIM data

SUGIRA Jean Claude[1], ZHOU Xiaoyi[1,2], LI Xiaoya[3], LI Shutao[4], RUAN Xin[5], WANG Hao[6]

(1. School of Transportation Engineering, Southeast University, Nanjing 211189, China; 2. State Key Laboratory of Safety, Durability and Healthy Operation of Long Span Bridges, Nanjing 211189, China; 3. Zhejiang Scientific Research Institute of Transport, Hangzhou 310005, China; 4. CCCC Highway Consultants Co., Ltd., Beijing 100010, China; 5. School of Civil Engineering, Tongji University, Shanghai 200092, China; 6. School of Civil Engineering, Southeast University, Nanjing 211189, China)

Abstract: Extreme traffic loads significantly challenge the safety and cost-effectiveness of highway bridges, especially under site-specific traffic conditions. Conventional assessments often rely on overly conservative load models, leading to excessive structural design. In this study, a framework for the prediction of maximum bending moments in simply supported bridges is developed by integrating weigh-in-motion (WIM) data, traffic microsimulation, and generalized extreme value (GEV) regression modeling to establish relationships between the GEV parameters (μ, σ, ξ) and traffic factors—heavy vehicle proportion, bridge span length, vehicle speed, headway, and traffic volume. Using one-year WIM data from 7.4 million vehicles, the developed models for μ and σ exhibit high predictive accuracy (R^2>0.95) and are validated through leave-one-out cross-validation. The prediction of ξ is less accurate ($R^2 \approx 0.6$), requiring further improvement. Applying these models to a 1 000-year return level yields a reliable, data-driven extrapolation, supporting optimized bridge design and safety assessment under varying traffic conditions.

Key words: site-specific factors; extreme value; traffic load; weigh-in-motion (WIM); generalized extreme value (GEV) parameters; Monte Carlo simulation

DOI: 10.3969/j.issn.1003-7985.2026.01.006

Received 2025-06-14, **Revised** 2025-09-15.

Biographies: SUGIRA Jean Claude (1995—), male, Ph.D. candidate; ZHOU Xiaoyi (corresponding author), male, doctor, professor, xiaoyizhou@seu.edu.cn.

Foundation items: The National Natural Science Foundation of China (No. 52278149), the Natural Science Foundation of Jiangsu Province (No. BZ2024015), the Opening Project of State Key Laboratory for Track Technology of High-Speed Railway (No. 2023YJ375), the Opening Project of Zhejiang Engineering Centre of Road and Bridge Intelligent Operation and Maintenance Technology (No. 202402G).

Citation: SUGIRA Jean Claude, ZHOU Xiaoyi, LI Xiaoya, et al. Relationship between the extreme value distribution of bending moments and traffic characteristics for simply supported bridges based on WIM data[J]. Journal of Southeast University (English Edition), 2026, 42(1):65-73. DOI:10.3969/j.issn.1003-7985.2026.01.006.

Accurate prediction of the effects of extreme traffic loads on highway bridges is essential to ensuring structural integrity, optimizing design efficiency, and reducing construction costs. Conventional bridge design standards often adopt conservative load models, resulting in overdesigned structures that unnecessarily increase construction costs, particularly in regions characterized by dense or highly variable traffic conditions[1]. The progress in weigh-in-motion (WIM) technology has facilitated the detailed characterization of site-specific traffic conditions, providing a data-driven basis for improving the accuracy of load effect predictions[2]. However, the development of reliable, general models that clearly connect site-specific traffic conditions to extreme load effects remains a challenge. Such models can reduce unnecessary design expenses and potential structural failure caused by underestimated loads. Recent progress in probabilistic load modeling has significantly improved the statistical representation of traffic-induced load effects. The extreme value theory (EVT), particularly the generalized extreme value (GEV) distribution, provides a rigorous framework for estimating the tail behavior of maximum load effects[3]. Previous studies have proven the effectiveness of GEV distribution in modeling extreme load effects for short- and medium-span bridges, achieving high statistical fitting.

The recent progress in the prediction of bridge load effects has improved the modeling of extreme traffic conditions by employing traffic microsimulation and probabilistic approaches. Microsimulation estimates complex traffic conditions[4], and WIM-based models reduce structural overdesign by 10%-15% by incorporating site-specific vehicle compositions (10%-40% trucks)[5]. Gao et al.[6] combined generalized Pareto distribution models with WIM and structural health monitoring (SHM) data, demonstrating that traffic volume, traffic flow rates, and vehicle weights shape the tail of load effect distributions. Lu et al.[7] employed multilane microsimulation to demonstrate that traffic density and vehicle composition (15%-50% trucks) increase the GEV-modeled extreme loading probability in short-span bridges. Wang et al. [8] linked traffic flow rates and vehicle gaps to extreme load effects on long-span bridges using a Gaussian mixture model (GMM) and GEV fits. These studies have verified that traffic flow parameters, such as volume, headway, and vehicle mix, determine the extreme value distributions.

However, most of the existing studies have focused on either numerical simulations or statistical observations. The development of generalized analytical relationships between traffic flow parameters and extreme value distribution parameters regarding bridge traffic load effects has received less attention. This is particularly evident in complex traffic conditions, where analytical methods remain underdeveloped. The current reliance on either conservative design standards or case-specific simulations indicates the need for strong predictive tools that consider site-specific traffic variations. In this study, we propose a novel framework for the prediction of site-specific extreme traffic load effects by integrating WIM-based traffic microsimulation with EVT. Based on this framework, we derive analytical expressions that link GEV parameters to key traffic parameters, including heavy vehicle (HV) proportion, bridge span length, vehicle speed, headway, and traffic volume. Using Monte Carlo simulation and advanced regression modeling, our approach is validated against a one-year WIM dataset of 7.4 million vehicles obtained from Chinese highway bridges. By providing a practical, data-driven tool for bridge engineers, the proposed framework overcomes conservative existing standards, improves structural reliability, and supports cost-effective designs, particularly for bridges with high traffic variability. The derived equations lay the foundation for a generalized, predictive bridge design, marking a significant advancement over the existing methodologies. Apart from providing a practical tool for bridge engineers to optimize designs while ensuring safety, the proposed framework overcomes existing conservative standards, has the potential to significantly reduce costs, and improves structural reliability, particularly for bridges with high traffic variability. Additionally, the predictive equations derived in this study provide a foundation for general, data-driven bridge design, marking a significant advancement over previous methodologies.

1 Methodology

1.1 Data preprocessing

WIM data preprocessing is essential to ensuring the accuracy and reliability of subsequent analyses by avoiding common measurement errors inherent to WIM systems. This step involves three key tasks: outlier detection, data validation, and vehicle categorization. Outlier detection is performed using z-score filtering and interquartile range (IQR) analysis. For a given variable x_i (e.g., gross vehicle weight or speed), the z-score is calculated as

$$A = \frac{x_i - \bar{x}}{\sigma_x} \tag{1}$$

where $\bar{x}$ is the mean value and σ_x is the standard deviation of the variables. Data points with $|z_i|$ are identified as outliers (z_i is the z-score for each individual data point x_i). Alternatively, the IQR method defines outliers as values that satisfy

$$x_i < Q_1 - 1.5r \text{ or } x_i > Q_3 + 1.5r \tag{2}$$

where Q_1 and Q_3 are the first and third quartiles, respectively, and $r = Q_3 - Q_1$. Data validation involves filtering out implausible or erroneous records; these include vehicles with invalid axle counts (<2 or >8), gross vehicle weights (W) below a defined minimum threshold, and speeds exceeding a reasonable upper limit. Additionally, inconsistencies where the sum of individual axle weights does not match the total W, i.e., $\sum P_{ij} \neq W_i$, are either corrected or excluded. Vehicle classification is mostly based on the number of axles when detailed axle configurations are not available. WIM data preprocessing ensures a high-quality dataset suitable for traffic microsimulation and structural load effect analyses, effectively addressing the data quality problem frequently encountered in WIM measurements[9].

1.2 Traffic microsimulation

To replicate the site-specific traffic variability, a Monte Carlo simulation (MCS) framework is employed to generate synthetic traffic data over a prolonged period. This simulation incorporates annual traffic growth, enabling a realistic projection of the future traffic behavior. The traffic arrival rate in year y, adjusted for volume growth, is expressed as

$$\Lambda_y = \lambda_o (1 + g)^y \tag{3}$$

where λ_o is the baseline arrival rate, veh/h; and g is the annual growth rate.

Vehicle arrival is modeled using the Poisson distribution, where the probability of observing k arrivals within time t is given as

$$P(N(t) = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \tag{4}$$

where $N(t)$ denotes the number of arrivals. Consequently, the interarrival time (or headway) is exponentially distributed as follows:

$$f(h) = \lambda e^{-\lambda t} \qquad h \geqslant 0 \tag{5}$$

where h denotes the headway, s.

A complex traffic composition, such as free-flow, platooning, and congestion headway, can be modeled using the HyperLang distribution, which is a mixture of Erlang distributions as follows:

$$f(h)=\sum_{m=1}^{M}p_m f_m(h),\quad \sum_{m=1}^{M}p_m=1 \tag{6}$$

where M is the number of mixture components; p_m is the mixing probability; and $f_m(h)$ is the Erlang distribution for component m.

$$f_m(h)=\frac{\lambda_m^{k_m}h^{k_m-1}e^{-\lambda_m h}}{(k_m-1)!}\qquad h\geqslant 0 \tag{7}$$

where λ_m is the rate parameter; and k_m is the shape parameter (number of phases). The expected headway and variance are given as

$$E[h]=\sum_{m=1}^{M}p_m\frac{k_m}{\lambda_m} \tag{8}$$

$$\mathrm{Var}[h]=\sum_{m=1}^{M}p_m\left(\frac{k_m}{\lambda_m^2}+\left(\frac{k_m}{\lambda_m}\right)^2\right)\left(\sum_{m=1}^{M}p_m\frac{k_m}{\lambda_m}\right)^2 \tag{9}$$

The estimation of model parameters $\{p_m,\lambda_m,k_m\}_m^M=1$ and observed headways h_i is based on the maximum likelihood estimation (MLE), where the log-likelihood function

$$\ell(\boldsymbol{\theta})=\sum_{i=1}^{N}\ln\left(\sum_{m=1}^{M}p_m\frac{\lambda_m^{k_m}h^{k_m-1}e^{-\lambda_m h_i}}{(k_m-1)!}\right) \tag{10}$$

is maximized; here, N is the total number of vehicles and

$$\boldsymbol{\theta}=\{p_m,\lambda_m,k_m\}_m^M=1 \tag{11}$$

The individual vehicle characteristics can be simulated by sampling parameters, such as gross vehicle weight W, axle weights, axle spacings, and speeds, from empirical distributions (gamma distribution, beta distributions, lognormal distribution, etc.) and fitting their values to observed data. The selection of the best-fit distributions is based on goodness-of-fit tests, such as the Kolmogorov-Smirnov (K-S) test. This traffic microsimulation framework enables the generation of realistic, heterogeneous traffic scenarios that model complex vehicle interactions and time-varying behavior, thus overcoming the limitations of simple arrival models.

1.3 Load effect calculation

The load effect calculation estimates the maximum bending moments induced by vehicular traffic on a simply supported bridge; this calculation considers the vehicle positions, axle loads, dynamic amplification, and significant crossing events (SCEs), where multiple vehicles simultaneously cross a bridge and affect its structural response[10-11]. We adopted a time-stepping approach based on the influence line theory to calculate the structural response at each time step. The position of the j-th axle of a vehicle at time t is given as

$$x_{ij}(t)=(t-t_i)\frac{v_i}{3.6}+\sum_{k-1}^{j-1}s_{ik} \tag{12}$$

where t_i is the arrival time, s; v_i is the speed, km/h; s_{ik} is the k-th axle spacing, m.

The total time required for vehicle i to fully cross a bridge of span length L is

$$t_{\mathrm{cross},i}=(L+2w_i)\frac{3.6}{v} \tag{13}$$

where $w_i=\sum s_{ik}$ is the vehicle's wheelbase, m. Any vehicle j ($j>1$) contributes to the SCE of the vehicle if

$$t_j\leqslant t_i+t_{\mathrm{cross},i} \tag{14}$$

Thus, the number of vehicles involved in the SCE for vehicle i is

$$n_i=1+\sum_{j=i+1}^{N}1\qquad t_j\leqslant t_i+t_{\mathrm{cross},i} \tag{15}$$

The total bending moment at time t is

$$M(t)=\sum_{i=1}^{N}\sum_{j=1}^{n_i}\alpha P_{ij}I(x_{ij}(t)) \tag{16}$$

where P_{ij} is the axle load, kN; $I(x_{ij}(t))$ is the influence line ordinate at position $x_{ij}(t)$; and α is the dynamic amplification factor (DAF). In this study, a DAF was employed to model vehicle-bridge interaction effects, such as bridge vibrations[12], vehicle suspension dynamics, and road surface irregularities, based on code provisions and previous studies[13-17].

1.4 GEV parameter estimation

Traffic-induced load effects can be characterized by modeling the daily maximum bending moments using the GEV distribution[18]. This distribution is suitable for modeling block maxima (e. g., daily maxima) and provides a unified framework that includes the Gumbel, Fréchet, and Weibull extreme value distribution types via its shape parameter[19].

The cumulative distribution function of the GEV distribution is given as

$$F(x;\mu,\sigma,\xi)=\begin{cases}\exp\left(-\left(1+\xi\left(\dfrac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right)\right) & \xi\neq 0\\ \exp\left(-\exp\left(-\dfrac{x-\mu}{\sigma}\right)\right) & \xi=0\end{cases} \tag{17}$$

The probability density function is

$$f(x;\mu,\sigma,\xi)=\frac{1}{\sigma}t(x)^{\xi+1}e^{-t(x)} \tag{18}$$

where

$$t(x)=\begin{cases}\left(1+\xi\left(\dfrac{x-\mu}{\sigma}\right)\right)^{-\frac{1}{\xi}} & \xi\neq 0\\ \exp\left(-\left(\dfrac{x-\mu}{\sigma}\right)\right) & \xi=0\end{cases} \tag{19}$$

The return level x_T, which represents the level that is expected to be exceeded once every T year (e.g., a 100-year return level), is calculated as

$$x_T = \mu + \frac{\mu}{\xi}\left(\left(-\ln\left(1 - \frac{1}{T}\right)\right)^{-\xi} - 1\right) \quad \xi \neq 1 \tag{20}$$

The GEV distribution parameters can be estimated using the maximum likelihood estimation (MLE) method. The log-likelihood function for a sample of daily maxima $x_1, x_2, \cdots, x_n$ is given as

$$l(\mu, \sigma, \xi) = -n\ln\sigma - \left(1 + \frac{1}{\xi}\right)\sum_{i=1}^{n}\ln\left(1 + \xi\left(\frac{x_i - \mu}{\sigma}\right)\right) - \sum_{i=1}^{n}\left(1 + \xi\left(\frac{x_i - \mu}{\sigma}\right)\right)^{-1/\xi} \tag{21}$$

The variance-covariance matrix of the estimated parameters $\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma}, \hat{\xi})$ is approximated as

$$\mathrm{Var}(\hat{\boldsymbol{\theta}}) \approx \left(\frac{-\partial^2 l}{\partial\boldsymbol{\theta}\,\partial\boldsymbol{\theta}^{\mathrm{T}}}\right)^{-1} \tag{22}$$

1.5 Regression modeling

Here, we present the methodology for deriving analytical formulae for the GEV parameters, which are expressed as functions of five site-specific traffic characteristics and these GEV distribution parameters. The proposed approach combines linear and nonlinear regression models, such as quadratic, exponential, and polynomial models. Additionally, machine learning approaches, such as the multilayer perceptron (MLP), were employed to model complex dependencies.

1.6 Model fitting and selection

The selection of fitting models involves the optimization of regression models and the evaluation of their suitability. For μ and σ, the coefficients of quadratic, exponential, and power-law models are estimated by minimizing the sum of squared errors between the predicted and observed GEV parameters. For the shape parameter ξ, the MLP and polynomial models are used for fitting using a similar optimization approach. Model selection employs fivefold cross-validation to assess generalizability across the WIM dataset. The goodness-of-fit is evaluated using the coefficient of determination R^2,

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{23}$$

where y_i, $\hat{y}_i$, and $\bar{y}$ are the observed, predicted, and the mean of observed values, respectively. Additionally, the Akaike information criterion was used to balance model fitting and complexity[20]. For ξ, polynomial, exponential, and MLP models were compared, and alternative approaches, such as random forest and gradient-boosting regressors, were employed to ensure model robustness[21]. The model with the best cross-validation performance was selected. Its performance is presented in detail in Section 2.

2 Application of the Proposed Framework to Simply Supported Bridges

In this section, we present the results obtained by applying the proposed framework to a simulated 5-year traffic load condition, which was generated using WIM data. The results were obtained by employing all methodological stages: data preprocessing, traffic microsimulation, load effect calculation, GEV parameter estimation, regression modeling, and model selection.

2.1 Traffic microsimulation

2.1.1 Data preprocessing results

The WIM dataset was preprocessed according to procedures described in Section 1 to ensure data quality. The outlier detection using z-score filtering ($|Z_i| > 3$) and IQR analysis indicated 2.7% of records as outliers, mainly in W and speed.

The total number of records removed from the dataset was about 200 000. The vehicles were classified according to their number of axles (Table 1).

Table 1 Statistical and traffic composition of the simulated trucks

Truck type	Number of axles	Proportion/%	Mean W/t
Type Ⅰ	2	88.1	22.2
Type Ⅱ	3	1.7	28.6
Type Ⅲ	4	3.8	33.3
Type Ⅳ	5	0.3	47.4
Type Ⅴ	6	6.1	56.1

2.1.2 Synthetic data generation

The MCS was used to simulate synthetic traffic over a 2 000 d period. Eq. (3) was used to model the vehicle arrival rate. Using a 2% annual growth rate in traffic volume, the calculated increase over a 5-year simulation period was from 1 000 to 1 486 veh/h; we assumed that the W distributions remained stationary because they were obtained from empirical distribution data fitted to the WIM dataset without incorporating the increase in vehicle weights[22-23]. In addition, the vehicle arrival rate was modeled as a Poisson process, where headways followed the HyperLang distribution to accurately model a congested traffic flow or a free flow[24-25]. This was achieved using varying distribution parameters ($p_1 = 0.65w$, $\lambda_1 = 0.5\ \mathrm{s}^{-1}$, $k_1 = 2$ and $p_2 = 0.65$, $\lambda_2 = 0.5\ \mathrm{s}^{-1}$, $k_2 = 3$) with a mean headway $E[h]$=4.1 s and a variance $\mathrm{Var}[h]$=12.34 s^2. The vehicle characteristics were obtained from statistically validated distributions: GMMs for W, axle spacing, and speeds and Gamma distribution models for axle weights.

The selection of the distributions was based on K-S tests ($p>0.05$). The model validation against the WIM dataset exhibited high fidelity (Pearson's correlation coefficient $r = 0.95$ and Spearman's rank correlation coefficient $\rho = 0.93$), verifying simulation robustness for the subsequent load effect analyses.

2.2 Load effect calculation results

The maximum bending moments in simply supported bridges with span lengths in the 10-100 m were calculated using Eq. (20). Synthetic traffic data for more than 40 million vehicles within a 5-year period were used. Additionally, various traffic conditions, including vehicle speeds (10-100 km/h), HV proportions (10%-100%), headways (3-10 s), and traffic volumes (200-2 000 veh/h), were considered. For each case, the relationship between the maximum load effect and each traffic parameter was analyzed. An example is presented in Fig. 1, which effectively demonstrates the direct relationship between the maximum daily bending moment and HV proportion. We clearly observe that when the HV proportion increases, the loads on the bridge increase, leading to an increase in the corresponding bending moments across the bridge span. Fig. 2 shows that the relationship between the maximum daily load effect and the HV proportion is linear; when the HV proportion increases by 5%, the bending moments increase by 12%-15% for bridge spans in the 10-50 m range and 8%-10% for bridge spans over 70 m.

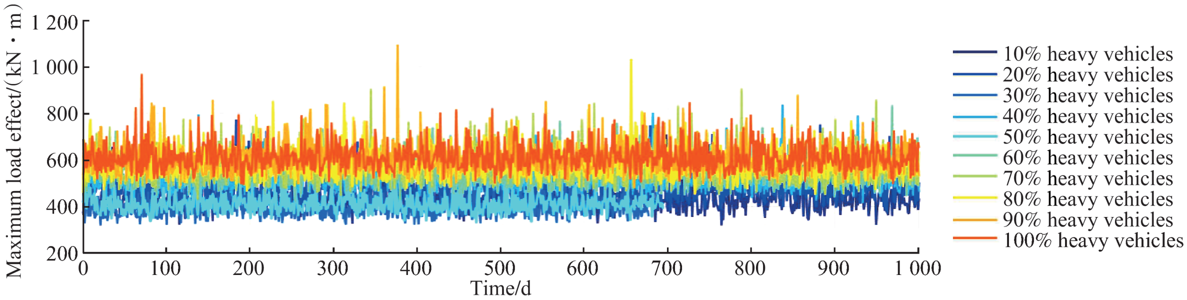

Fig. 1 Variation of the maximum load effect with HV proportion

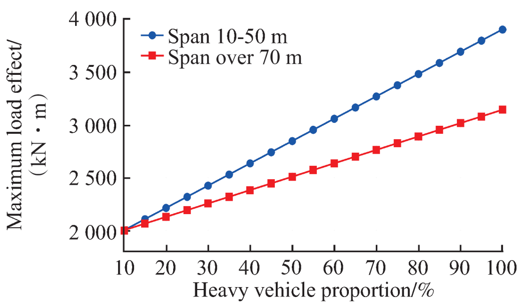

Fig. 2 Relationship between maximum load effect and HV proportion and headway

Additionally, the results show that when the vehicle speeds are in the 10-40 km/h range, the load effect increases by 15% compared with higher speeds; this is mainly attributed to the prolonged multivehicle loading during SCEs. Additionally, the analysis of the WIM dataset showed that HVs (e.g., Types Ⅳ and Ⅴ, with a mean W of 47.4 and 56.1 t) typically travel at slower speeds (25-35 km/h) compared with those of light vehicles (50-70 km/h).

This slow speed of HVs probably increases the load effect, because their W increases the bending moments during multivehicle loading scenarios. Fig. 3 shows that the span length significantly affects the load effect, causing a 35% increase in the bending moments per 10 m for 10-30 m spans, 28% for 30-60 m spans, and 22% for spans over 60 m; this is attributed to the increased multivehicle loading on long spans during congestion. Fig. 4

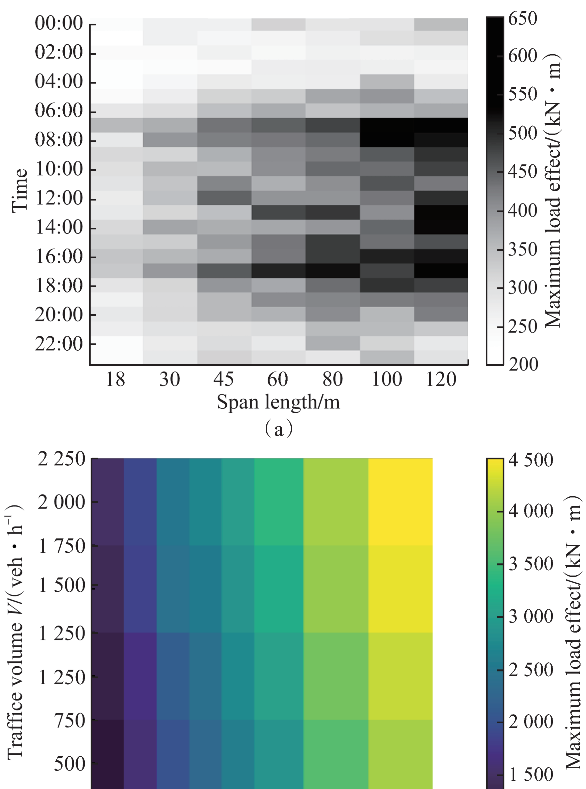

Fig. 3 Variation of the maximum load effect vs. span length and time. (a) Heatmap of the maximum load effect; (b) Variation of the maximum load effect vs. span length

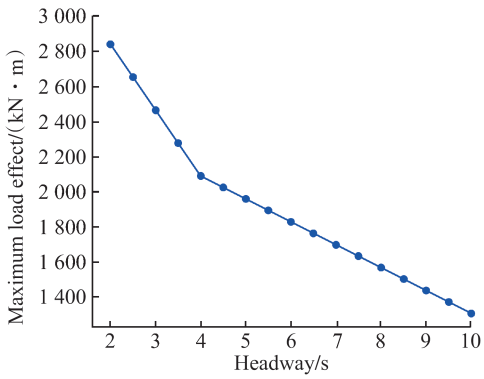

Fig. 4 Relationship between maximum load effect and headway

shows that when the headway H is below 4 s, the load effect increases by 8%-10% per 0. 5 s decrease, peaking at a 42% increase when the headway decreases from 10 to 2 s. When the traffic volume V increases from 200 to 1000 veh/h, the load effect increases by 30%; small additional increases (5%-8%) are observed when $V>$ 1 500 veh/h.

The daily maximum load effect is mainly observed during peak hours (e. g. , 07:00 and 17:00); this is expected and is consistent with the existing literature, as it is commonly known that increased traffic volumes lead to an increased load effect on the bridge. This is because when a high number of vehicles simultaneously cross a bridge, the load effect increases, thereby increasing the bending moments.

2. 3 GEV parameter estimation and regression modeling results

Regression models were developed to link the GEV parameters to traffic parameters. For example, quadratic, power-law, exponential, rational, and logarithmic functions were used to establish the relationship for each distribution parameter. Individual nonlinear regression fitting models (quadratic, power-law, exponential, and logarithmic) were used to establish the relationship between the GEV location parameter μ and bridge span length (Fig. 5(b)). We observe that the quadratic and power-law models provide promising results compared with the logarithmic model. The same approach was adopted for each traffic parameter in the validation of the derived equations for μ and scale parameter σ, where $R^2>0.95$. Regarding the multivariate model, which considers all traffic parameters, the analysis showed strong and accurate relationships for σ and μ ($R^2>0.95$; the root mean square error (RMSE) was small compared with standard values). The developed multivariate models for σ and μ are given in the following equations:

$$\sigma = \frac{2(HLP)^{0.5}}{VS+1} \tag{24}$$

$$\mu = 1\,022e^{0.2V+0.2S} + \frac{1\,344H}{\sqrt{LP+1}} \tag{25}$$

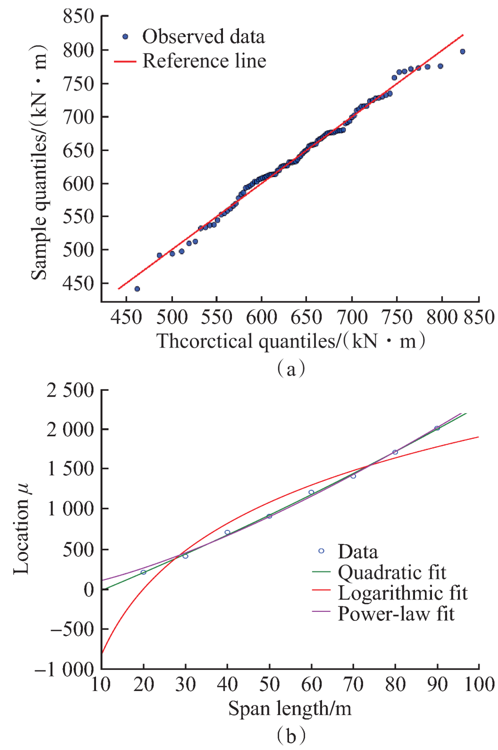

Fig. 5 Validation and modeling of the GEV distribution with the span length. (a) Q-Q plot for the GEV goodness-of-fit test (60% HV proportion); (b) Regression models used for data fitting

Compared with σ and μ, the relationship of ξ against all individual traffic parameters was weaker; the regression models exhibited low R^2 values. Advanced methods such as machine learning and deep learning techniques (including neural networks and random forests) regarding ξ also exhibited poor performance ($R^2<0.6$). Regarding the multivariate model, the comparative performance of an MLP is shown in Fig. 6. Compared with the polynomial fit with a composite feature, the MLP exhibited better performance ($R^2=0.6$). The derived equation is as follows:

$$x = HLP\ln(V+S+1) \tag{26}$$

which yields

$$\xi(x) = -0.4x^3 + 0.6x^2 + 0.4x + 0.4 \tag{27}$$

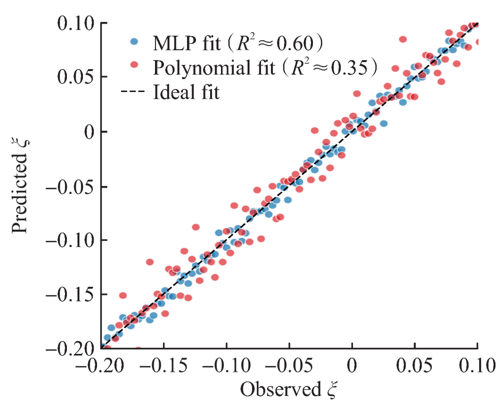

Fig. 6 Q-Q plot of MLP vs. a polynomial fit for shape parameter ξ

The equations derived in this study enable the precise estimation of the maximum bending moments on simply supported bridges by capturing the impact of traffic parameters on extreme load effects. For example, μ, which determines the tendency of the extreme value distribution, is strongly correlated with the traffic density and the proportion of HVs (e. g. , trucks constituting 15%-50% of traffic). This relationship indicates that when the traffic volume and the proportion of HVs increase, the expected maximum load significantly shift effects, allowing engineers to better predict critical design scenarios for bridges under site-specific traffic conditions. The high accuracy of the models used, as indicated by the strong correlation coefficients, demonstrates their reliability in practical applications, providing a data-driven alternative to conservative design assumptions that often lead to overdesigned structures. However, the models for ξ exhibited low predictive accuracy, indicating challenges in modeling the tail behavior of extreme load distributions.

3 Discussion

The results presented in Section 2, which are based on WIM data, constitute a strong framework for predicting extreme traffic load effects on simply supported highway bridges by providing useful theoretical and practical knowledge for bridge design optimization. The high accuracy of the models for μ and σ (R^2>0. 999, RMSE=6. 98) indicates the advantages of integrating traffic microsimulation with GEV regression modeling. These models can accurately capture the nonlinear interactions among traffic volume, vehicle speed, headway, span length, and HV proportion; additionally, they are consistent with theoretical predictions indicating that varying traffic conditions significantly affect the load effects. The obtained results, such as a 15% reduction in the load effect at speeds above 60 km/h and a 42% increase as the headway decreases from 10 to 2 s, indicate the critical role of vehicle clustering and temporal spacing in load effects. This is consistent with the vehicle-bridge interaction theory, where a reduced headway increases the multivehicle load effect. The exponential increase of μ with traffic volume and speed, which is modeled in Eq. (25), theoretically validates the hypothesis that increased traffic volumes lead to extreme load magnitudes, where the spatial parameters act as moderators. The scale model described in Eq. (26) indicates that variability is stabilized by increased traffic volumes, providing a theoretical basis for understanding the load dispersion under varying traffic conditions.

However, the predictive model for ξ exhibited notably lower accuracy (R^2≈0. 53-0. 63) compared with those for μ and σ, despite employing advanced techniques such as MLP and polynomial regression. ξ is a highly important parameter in the EVT for determining the heaviness of the distribution tail and directly controlling the decay rate of the extreme value distribution. Therefore, its accurate estimation is critical for a reliable extrapolation to long return periods.

The inherent challenge in modeling ξ stems from its fundamental sensitivity to the most rare and extreme events in the dataset. Although μ and σ are affected by the central mass and overall dispersion of the daily maxima, ξ is exclusively determined by the behavior of the uppermost tail. In the context of bridge traffic loads, these extreme events are often the result of low-probability, high-impact combinations of traffic scenarios, such as the simultaneous occurrence of multiple heavily loaded trucks traveling at critical speeds with a minimal headway on the bridge. Even when using a large WIM dataset of 7. 4 million vehicles, the number of such "perfect storm" events remains statistically sparse, leading to high uncertainty and volatility in the estimation of ξ. This uncertainty has profound implications in the extrapolation of long-term load effects. An overestimation of ξ (positive value, Fréchet-type tail) leads to an unbounded and rapidly increasing tail, potentially resulting in overly conservative design loads. Conversely, an underestimation of ξ (negative value, Weibull-type tail) leads to a bounded upper limit, which could nonconservatively underestimate the risk of extreme events. The sensitivity analysis showed that a variation of only ±0. 05 in the value of ξ can alter the 1 000-year return level estimate by 15% or more. Therefore, the relative inaccuracy in predicting ξ based on traffic parameters introduces a significant source of uncertainty in the very long-term predictions that are most crucial in the design and safety assessment.

To address this issue, we propose several future research directions: (1) Instead of relying solely on block (daily) maxima, a peaks-over-threshold method using the generalized Pareto distribution could be employed. This approach uses all data points that exceed a high threshold, making more efficient use of the available extreme value data and potentially yielding more robust estimates of the tail behavior. (2) Pooling WIM data from multiple bridge sites with varying traffic conditions would dramatically enlarge the dataset of extreme events. This would provide a robust statistical basis for analyzing the relationship between traffic parameters and ξ, potentially revealing trends that are obscured in a single-site analysis. (3) Enhancing ML models by incorporating physical constraints or prior knowledge about the system (e. g. , the maximum plausible load based on structural capacity or vehicle weight) could prevent the model from generating unrealistic extrapolations, even if the tail fit is imperfect.

Another limitation is that in this study, the predictive models were developed using a WIM dataset obtained from a single Chinese highway bridge; this limits model generalizability because of the spatial and temporal variability of traffic loads. To improve model applicability, the proposed framework was designed to be adaptable, allowing recalibration with site-specific WIM data obtained from different locations. In future work, the model will be validated using multisite WIM datasets to ensure robustness across diverse traffic conditions and bridge types.

4 Conclusions

A novel data-driven framework for predicting extreme traffic load effects on simply supported highway bridges was developed using a one-year WIM dataset comprising 7.4 million vehicles. By integrating detailed traffic microsimulation with GEV regression modeling, analytical relationships between key site-specific traffic parameters and the GEV distribution parameters was derived. The models for μ and σ parameters exhibited high predictive accuracy ($R^2>0.95$ and low RMSE), verifying their robustness and reliability in estimating the extreme bending moments in highway bridges. These models were successfully validated by employing leave-one-out cross-validation.

The main contribution of this work is the provision of a site-specific predictive tool that overcomes conservative traditional bridge design methodologies. This approach provides cost-effective and structurally optimized solutions, especially in high-traffic environments. The practical implementation of the framework was demonstrated on a 40 m span bridge example, yielding a recalibrated 1 000-year return level of around 2 800 kN·m, thus verifying the applicability of the model in real-world applications.

In this study, we established and demonstrated the fundamental relationships between site-specific traffic parameters and extreme load effect distribution parameters, rather than developing a fully generalized model. Consequently, the scope of the work was intentionally focused, utilizing WIM data obtained from a single site and examining the bending moments of a simply supported bridge. Furthermore, the low predictive accuracy ($R^2\approx0.6$) for ξ indicates a limitation in accurately modeling the tail behavior of the distribution, which is highly sensitive to rare traffic events. Therefore, although the proposed framework is highly reliable for estimating central load effects, extrapolation to very long return periods (e.g., beyond 1 000 years) should be treated with caution. In future work, a more generalized and robust model can be developed based on WIM data obtained from multiple sites and improved techniques for estimating ξ. Overall, this study is an important step toward data-driven bridge design (providing a pathway for improving structural reliability), accurate assessment of extreme traffic loads, and cost reduction in infrastructure planning and management.

References

[1] ZHOU J Y, CAPRANI C C, ZHANG L W. On the structural safety of long-span bridges under traffic loadings caused by maintenance works[J]. Engineering Structures, 2021, 240: 112407.

[2] KAWAKATSU T, AIHARA K, TAKASU A, et al. Data-driven bridge weigh-in-motion [J]. IEEE Sensors Journal, 2023, 23(15): 17064-17077.

[3] OBRIEN E, HAJIALIZADEH D, ENRIGHT B, et al. Factors affecting the accuracy of characteristic maximum load effects[M]//Bridge Traffic Loading. London: CRC Press, 2021: 143-184.

[4] ZHOU J Y, WU W R, CAPRANI C C, et al. A hybrid virtual-real traffic simulation approach to reproducing the spatiotemporal distribution of bridge loads[J]. Computer-Aided Civil and Infrastructure Engineering, 2024, 39 (11): 1699-1723.

[5] LYDON M, TAYLOR S E, ROBINSON D, et al. Recent developments in bridge weigh in motion (B-WIM) [J]. Journal of Civil Structural Health Monitoring, 2016, 6(1): 69-81.

[6] GAO X, DUAN G X, LAN C G, et al. Bayesian updates for an extreme value distribution model of bridge traffic load effect based on SHM data[J]. Sustainability, 2021, 13(15): 8631.

[7] LU N W, WANG H H, WANG K, et al. Maximum probabilistic and dynamic traffic load effects on short-to-medium span bridges [J]. Computer Modeling in Engineering & Sciences, 2021, 127(1): 345-360.

[8] WANG X J, RUAN X, CASAS J R, et al. Probabilistic model of traffic scenarios for extreme load effects in long-span bridges [J]. Structural Safety, 2024, 106: 102382.

[9] LI M, HUANG T L, LIAO J J, et al. WIM-based vehicle load models for urban highway bridge [J]. Latin American Journal of Solids and Structures, 2020, 17 (5): e290.

[10] RIZQIANSYAH A, CAPRANI C C. Hierarchical Bayesian modeling of highway bridge network extreme traffic loading[J]. Structural Safety, 2024, 111: 102503.

[11] HUI Y X, WANG W W, LIU X N. Comfort evaluation of prefabricated and assembled pedestrian cable-stayed bridges [J]. Journal of Southeast University (English Edition), 2023, 39(1): 26-32.

[12] YUAN Z J, WANG H, MAO J X, et al. Influence study of main cable displacement-controlled device type of long-span suspension bridges on structural mechanical properties[J]. Journal of Southeast University (English Edition), 2025, 41(1): 27-36.

[13] XING C X, SHU Y W, ZHU X J, et al. Multistage seismic damage constitutive model and parameter calibration of reinforced concrete columns[J]. Journal of Southeast University (English Edition), 2024, 40(4): 386-395.

[14] LING T Y, DENG L, HE W, et al. Determination of dynamic amplification factors for small- and medium-span

highway bridges considering the effect of automated truck platooning loads[J]. Mechanical Systems and Signal Processing, 2023, 204: 110812.
[15] PRADEEP A. Dynamic amplification effect on bridges under traffic loads: Parametric study and novel design integration approach[D]. Montreal, Québec, Canada: Concordia University. 2025.
[16] KALIN J, ŽNIDARIČ A, ANŽLIN A, et al. Measurements of bridge dynamic amplification factor using bridge weigh-in-motion data[J]. Structure and Infrastructure Engineering, 2022, 18(8): 1164-1176.
[17] XIE W, BAO Y Y, TIE N, et al. Sensitivity analysis of parameters influencing ED and self-centering capacity in self-centering bridge bents with ED beams using validated numerical model[J]. Journal of Southeast University (English Edition), 2025, 41(3): 338-347.
[18] DAI B R, WU D J, LI Q. Investigation of multiple-presence factor for traffic loads on road-rail bridges based on a novel extreme value analysis approach[J]. Structural Safety, 2022, 96: 102199.
[19] DAI B R, XIA Y, LI Q. An extreme value prediction method based on clustering algorithm[J]. Reliability Engineering & System Safety, 2022, 222: 108442.
[20] XIA H W, NI Y Q, WONG K Y, et al. Reliability-based condition assessment of in-service bridges using mixture distribution models[J]. Computers & Structures, 2012, 106: 204-213.
[21] HASSAN R, HEKMATI ATHAR S P, TAHERI M, et al. Regression model for structural health monitoring of a lab scaled bridge[C/OL]//Proc SPIE 11594, NDE 4.0 and Smart Structures for Industry, Smart Cities, Communication, and Energy. 2021. https://doi.org/10.1117/12.2592037.
[22] SVECHNIKOV E, SNIJDER H H, MALJAARS J, et al. Probabilistic design life selection procedure for a steel bridge[J]. ce/papers, 2023, 6(3/4): 477-482.
[23] LIU Y Y, ZHOU J Y, SU J X, et al. Residual capacity assessment of in-service concrete box-girder bridges considering traffic growth and structural deterioration[J]. Structural Engineering and Mechanics, 2023, 85(4): 531-543.
[24] SAHA P, ROY R. A descriptive study of vehicle class-wise headways using mixed traffic data[J]. Journal of the Institution of Engineers (India): Series A, 2022, 103(4): 1287-1298.
[25] KHAKIFIROOZ M, FATHI M, DU L L. Reliability assessment of vehicle-to-vehicle communication networks through headway distribution and information propagation delay[J]. Transportation Research Interdisciplinary Perspectives, 2024, 24: 101053.

基于WIM数据的简支桥梁弯矩极值分布与交通特性的关系

吉安[1], 周小燚[1,2], 李晓娅[3], 李书韬[4], 阮欣[5], 王浩[6]

(1.东南大学交通工程学院,南京 211189;2.长大桥梁安全长寿与健康运维全国重点实验室,南京 211189;3.浙江省交通科学研究院,杭州 310005;4.中交公路规划设计院有限公司,北京 100010;5.同济大学土木工程学院,上海 200092;6.东南大学土木工程学院,南京 211189)

摘要: 极端交通荷载对高速公路桥梁安全与经济性构成挑战,特定交通条件下影响更显著。传统设计方法采用保守载荷模型,易导致结构过度设计。本文结合动态称重(WIM)数据、交通微观模拟与广义极值分布(GEV)回归,提出基于交通特征参数的简支梁桥弯矩极值预测模型。以某高速WIM记录的740万辆车数据为基础,分析GEV分布的位置参数μ、尺度参数σ、形状参数ξ与重型车比例、跨径、车速、车距、流量5类交通特征的关系。结果表明,模型可高精度预测弯矩极值分布,μ与σ的R^2超过0.95,并通过留一法交叉验证验证了模型稳健性。但ξ预测精度较低,R^2约为0.6,尾部行为表征需进一步研究。数值分析显示,该模型能快速预测车辆荷载效应极值,为1 000年重现期值的可靠性评估提供数据驱动方案。

关键词: 场地特定因素;极值;交通载荷;动态称重;广义极值分布(GEV)参数;蒙特卡洛模拟

中图分类号: TU391