Accident Analysis and Prevention 129 (2019) 55–65

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Effects of globally obtained informative priors on bayesian safety performance functions developed for Australian crash data

Amir Pooyan Afghari[a], Md. Mazharul Haque[a,*], Simon Washington[b], Tanya Smyth[c]

[a] School of Civil Engineering and Built Environment, Queensland University of Technology, 2 George Street, Brisbane City, QLD, 4001, Australia
[b] Civil Engineering, University of Queensland, Building 49 Advanced Engineering Building, Staff House Road, St Lucia, QLD, 4072, Australia
[c] Department of Transport and Main Roads, Australia

ARTICLE INFO

Keywords:
Informative prior
Bayesian inference
Distribution fitting
Safety performance function
Crash prediction model
Random parameters model

ABSTRACT

The precision and bias of Safety Performance Functions (SPFs) heavily rely on the data upon which they are estimated. When local (spatially and temporally representative) data are not sufficiently available, the estimated parameters in SPFs are likely to be biased and inefficient. Estimating SPFs using Bayesian inference may moderate the effects of local data insufficiency in that local data can be combined with prior information obtained from other parts of the world to incorporate additional evidence into the SPFs. In past applications of Bayesian models, non-informative priors have routinely been used because incorporating prior information in SPFs is not straightforward. The previous few attempts to employ informative priors in estimating SPFs are mostly based on local prior knowledge and assuming normally distributed priors. Moreover, the unobserved heterogeneity in local data has not been taken into account. As such, the effects of globally derived informative priors on the precision and bias of locally developed SPFs are essentially unknown.

This study aims to examine the effects of globally informative priors and their distribution types on the precision and bias of SPFs developed for Australian crash data. To formulate and develop global informative priors, the means and variances of parameter estimates from previous research were critically reviewed. Informative priors were generated using three methods: 1) distribution fitting, 2) endogenous specification of dispersion parameters, and 3) hypothetically increasing the strength of priors obtained from distribution fitting. In so doing, the mean effects of crash contributing factors across the world are significantly different than those same effects in Australia. A total of 25 Bayesian Random Parameters Negative Binomial SPFs were estimated for different types of informative priors across five sample sizes. The means and standard deviations of posterior parameter estimates as well as SPFs goodness of fit were compared between the models across different sample sizes. Globally informative prior for the dispersion parameter substantially increases the precision of a local estimate, even when the variance of local data likelihood is small. In comparison with the conventional use of Normal distribution, Logistic, Weibull and Lognormal distributions yield more accurate parameter estimates for average annual daily traffic, segment length and number of lanes, particularly when sample size is relatively small.

1. Introduction

Safety Performance Functions (SPFs) have been extensively used in the transportation safety literature to correlate crashes with their contributing factors and ultimately estimate the effects of these factors on crash risk. These effects may vary from one region to another due to geographic factors (e.g. continental, national, or state-wide attributes) and/or unobserved behavioural factors (e.g. drivers' speeding patterns), requiring every region/jurisdiction to develop its own SPF based on local data—where the phrase "local data" here refers to the availability of a temporally and spatially representative and sufficient sample. However, estimating a SPF can be fairly challenging when sufficient local data are not available. In addition, parameter estimates are likely to be biased and inefficient in small sample sizes and so inferences made based on these parameters may not be correct. In such circumstances, alternative solutions may be used to overcome the problem of data insufficiency in developing SPFs.

A promising solution to tackle insufficient local data is to adopt a

* Corresponding author.
E-mail addresses: a.afghari@qut.edu.au (A.P. Afghari), m1.haque@qut.edu.au (Md. M. Haque), s.washington@uq.edu.au (S. Washington), tanya.l.smyth@tmr.qld.gov.au (T. Smyth).

https://doi.org/10.1016/j.aap.2019.04.023
Received 28 January 2019; Received in revised form 29 April 2019; Accepted 29 April 2019
Available online 17 May 2019
0001-4575/

SPF developed elsewhere and apply it on local data. The transferability of the adopted SPF, however, needs to be evaluated first. This transferability—commonly referred to as *spatial transferability*—indicates that a SPF developed in one region may not be appropriate in another region, and thus it may be required to adjust the adopted SPF to local conditions using a calibration factor (Harwood et al., 2000; Persaud et al., 2002; Sawalha and Sayed, 2006; Li et al., 2018; Wang et al., 2017). The use of calibration factor, however, has been criticized in the literature. Chen et al. (2012) showed that a simple multiplicative factor may produce biased estimates in high ranges of Average Annual Daily Traffic (AADT) and thus may not be appropriate for different strata of traffic volume.

An alternative approach to overcome inaccurate parameter estimates when sufficient local data are not available is Bayesian inference. The appealing property of Bayesian inference is that it provides a platform to utilize existing information on the effects of crash contributing factors obtained from the past knowledge (referred to as "*informative prior"*) by combining it with the data likelihood and maximizing the product of these two terms. Accordingly, model parameters are calibrated in a compromising manner depending on the strength of these two pieces of information. This property of Bayesian inference explicitly accounts for the shortcoming of the recalibration method in that model parameters are guided by prior information while influenced by the pattern of local data. Incorporating informative priors into the models, however, are not straightforward because there is usually no available information about regression parameters. The previous few attempts to employ informative priors in estimating SPFs are mostly based on local prior knowledge and assuming normally distributed priors (more on this later). Moreover, the unobserved heterogeneity in local data has not been taken into account. As such, the effects of globally derived informative priors on the precision and bias of locally developed SPFs are essentially unknown.

This study aims to investigate the effects of globally obtained informative priors on the bias and precision of SPFs developed for Australian crash data. Rather than forcing the priors to follow a particular distribution type, the effects of prior distribution types on the prediction ability and parameter estimates are examined. The random parameters model specification is applied to account for unobserved heterogeneity in crash data. To achieve these aims, a thorough literature review is first conducted to extract the information on the parameters of SPFs developed across the world. These parameters are then translated into informative priors with appropriate distribution fitting. A total of 25 Bayesian SPFs are estimated for different types of informative prior and for five different sample sizes. A conventional SPF (i.e. a random parameters Negative Binomial model) is also estimated via maximum simulated likelihood estimation for Australian data, which serves as the means of comparison among candidate models. Ultimately, the precision and bias of parameter estimates and goodness of fit are compared among these scenarios

2. Literature review

As stated succinctly, adopting a SPF developed elsewhere and adjusting it for local data is one way of tackling data insufficiency. Recalibration is the most common adjustment method in the literature to improve spatial transferability of local SPFs adopted from other parts of the world. In doing so, the adopted SPF is applied on local data and predictions are made on crash counts. Calibration factor is then computed as the ratio of sum of observed crash counts (in local data) to sum of predicted crash counts (predicted by adopted SPF) (Sacchi et al., 2012; Farid et al., 2016; Kaaf and Abdel-Aty, 2015). The calibration factor is then used as a multiplicative factor in the adopted SPF to adjust for spatial transferability. Several studies have adopted this approach and have shown that the calibrated SPF has superior predictability than the original adopted SPF. On the contrary, many studies have obtained opposite findings, have criticized recalibration approach and have argued that a simple multiplicative calibration factor implicitly uses global unobserved assumptions (e.g. behavioral patterns) for local data which may not be totally accurate (Geedipally et al., 2017; Saha et al., 2017).

Estimating SPFs in Bayesian inference and incorporating informative priors into the estimation process is another approach to deal with data insufficiency. Earlier research reported that informative prior can significantly contribute to the parameter estimation of regression analysis when there is limited data available (Xu et al., 2015). Datasets with low sample mean and small sample size are other examples which can significantly benefit from informative priors via Bayesian inference (Miranda-Moreno et al., 2013). On the contrary, the effect of informative priors can be minimal if there is sufficient local data available (e.g. large sample size), and thus parameter estimates tend to be inclined towards data likelihood estimates in such circumstances. In spite of informative priors' potential to improve the precision and bias of SPFs, "*non-informative"* (also referred to as "*vague*") priors have been routinely and extensively used in the transportation safety literature (Hauer, 1992; Hauer et al., 1991; Mitra and Washington, 2007; Park and Lord, 2007; Persaud et al., 1999; Washington et al., 2003; Miaou and Lord, 2003; Afghari et al., 2016; Sacchi and Sayed, 2016; Saha et al., 2017; Ahmed et al., 2011; Huang and Abdel-Aty, 2010; Rusli et al., 2018). A few studies have attempted to develop informative priors for the parameters of SPFs (regression parameters and dispersion parameter). For example, Washington and Oh (2006) utilized the analogy of stated preference data analysis and applied random selection and laws of large numbers to derive accident modification factor densities from expert opinions. Mitra and Washington (2007) conducted an in-depth investigation on over dispersion parameter of count models in which they endogenously used dispersion parameter as a function of major and minor traffic volumes at intersections. Lord and Miranda-Moreno (2008) assumed a hypothetical informative prior on the dispersion parameter of a Poisson-Gamma model and showed that such informative prior can significantly increase the precision of the dispersion parameter estimate caused by low sample mean and small sample size. Subsequently, Heydari et al. (2013); Miranda-Moreno et al. (2013) and Heydari et al. (2014) adopted the relationship between inverse dispersion parameter and segment length reported in the Highway Safety Manual (HSM) of the American Association of State Highway and Transportation Officials (AASHTO) and computed the dispersion parameter for multilane undivided roadway segments in New York. They utilized the mean and variance of these parameters to create informative priors about the dispersion parameter for New York data. They also employed regression coefficients reported in the HSM for undivided 4-lane highway segments and utilized them as the mean to create semi-informative priors for the coefficients of crash contributing factors in New York data. They found that informative priors are effective in increasing the precision of model parameters, in particular when the sample size is small and the sample mean is low. In another attempt, Yu and Abdel-Aty (2013) investigated four approaches to develop informative priors including two-stage Bayesian updating, maximum likelihood estimation, method of moments and expert experience. They found that informative priors on regression parameters and inverse dispersion parameter increases the precision of estimates and improves models goodness of fit. Xu et al. (2015) conducted a Bayesian meta-analysis for the coefficient of Average Annual Daily Traffic (AADT) and developed informative priors for AADT and subsequently applied it for estimating SPFs. They also found that employing the meta-analysis technique can increase the prediction precision of SPFs by 15% compared to models developed with limited data. Recently, Farid et al. (2017) estimated SPF for crash data in California and used the parameter estimates as informative priors in the SPF estimated for crash data in Florida (and vice versa) and found that informative priors increase the precision of SPFs from one state to another.

As briefly reviewed above, empirical studies have shown that informative priors significantly increase the precision of parameter estimates when the sample mean is low, when the sample size is small, or when there is limited data available. The informative priors used in those studies, however, are either hypothetical or locally developed. More specifically, they either have used arbitrary mean and variance for informative priors or employed mean values suggested in the American manual (i.e. HSM) for estimating SPFs for the American crash data. An implicit assumption of the latter approach is that unobserved factors, which are unique to a specific country (i.e. United States), are assumed to be fixed across different regions of that country. This implicit assumption, however, may not hold for the unobserved factors in other countries (e.g. Australia) and thus employing HSM values as informative priors for the parameters of SPFs developed for other countries may not be theoretically plausible and may ultimately lead to incorrect parameter estimates. This restrictive property of local priors has acute consequences when noting that the past studies have not taken unobserved heterogeneity into account. Globally developed informative priors, on the other hand, can play a valuable role in bridging this gap, in which the knowledge gained through SPFs developed across the world may enhance local estimates. As such, the fundamental research question we seek to answer in this paper is whether *globally obtained informative priors contribute to the local parameter estimates in SPFs.*

In addition, while the importance of prior distribution type has been noted in the earlier research (Hadayeghi et al., 2006), the informative priors are often assumed to follow a typical distribution without questioning much whether the distribution type has an effect on parameter estimates. In particular, the prior information for regression parameters is typically assumed to follow Normal distributions and the prior for dispersion parameter is assumed to follow a Gamma distribution. It is not known whether these assumptions are valid and whether prior distribution types influence the precision of parameter estimates and subsequent statistical inferences.

3. Methodology

3.1. Random parameters negative binomial safety performance function

Random Parameters Negative Binomial (RPNB) model is a widely accepted safety performance function to establish the relationship between crashes and their contributing factors while accounting for unobserved heterogeneity in crash data (Mannering et al., 2016). Let Y_i represent the total observed crash count at site i following a Negative Binomial distribution with mean μ_i and dispersion parameter Φ. The probability of observing y_i crashes at site i can be expressed as:

$$P(Y_i = y_i) = \frac{\Gamma(y_i + \varphi)}{y_i!\Gamma(\varphi)}\left(\frac{\varphi}{\varphi + \mu_i}\right)^{\varphi}\left(\frac{\mu_i}{\varphi + \mu_i}\right)^{y_i} \tag{1}$$

where $\Gamma(.)$ is the gamma function. Assuming an exponential function for the mean of the NB distribution (μ_i), the predicted mean of crash count at site i can then be expressed as:

$$\mu_i = \alpha_0 F_{1i}^{\alpha_{1i}} F_{2i}^{\alpha_{2i}} \exp\left(\sum \beta_i X_i + \varepsilon_i\right) \tag{2}$$

where F_{1i} and F_{2i} are the measures of exposure, X_i are explanatory variables at site i, and α_i and β_i are estimated regression parameters. $\exp(\varepsilon_i)$ is a Gamma-distributed disturbance term with mean 1 and variance $1/\varphi$. To account for unobserved heterogeneity arising from various sources (e.g. unobserved variables), model parameters are allowed to vary across sites.

Estimating the above mentioned SPF can be achieved via two main approaches: frequentist (e.g. maximum simulated likelihood estimation) and Bayesian (i.e. Bayesian inference). In the frequentist approach, the likelihood of the data, *l(y)*, is maximized across the entire sample (*N* being sample size):

$$l(y) = \prod_{i=1}^{N} \int \frac{\Gamma(y_i + \varphi)}{y_i!\Gamma(\varphi)}\left(\frac{\varphi}{\varphi + \mu_i}\right)^{\varphi}\left(\frac{\mu_i}{\varphi + \mu_i}\right)^{y_i} f(\boldsymbol{\beta})d\boldsymbol{\beta} \tag{3}$$

while in Bayesian approach, model estimation is achieved by maximizing the posterior probability as in the following:

$$P[\mu|Y] = \frac{P[Y|\mu]\pi[\mu]}{\int m(y)dy} \tag{4}$$

where P[μ|Y] is referred to as the posterior probability, P[Y|μ] is the likelihood of data which is the same as Eq. (3) and $\int m(y)\,dy$ is the marginal distribution of observed data. π[μ] is the *prior,* which in our context, is the distribution of the regression parameters (α_i and β_i) and the dispersion parameter (Φ) in SPFs. A model with non-informative prior will have *vague* distributions (with means of zero and wide standard deviations), and a model with informative prior will have distributions with known parameters (known from the past research) assigned to the corresponding regression parameters.

3.2. Developing informative priors

To derive informative prior, three different methods are explored in this study: (i) distribution fitting, (ii) endogenous specification of dispersion parameter, and (iii) developing hypothetical priors. These methods have led to three different prior types.

3.2.1. Distribution fitting

A way of imposing informative priors for regression parameters is to assign them a distribution with known parameters. To do so, it is required to fit a distribution to observed values of the regression coefficients obtained by various studies over the world. This distribution fitting process is applied in three steps: hypothesizing a family of distributions, estimating the parameters of the hypothesized distributions and evaluating the quality of fit by applying the goodness of fit statistical tests (Ricci, 2005).

Choosing a distribution (or a family of distributions) for a set of observations depends on fundamental characteristics of those observations such as whether they are continuous or discrete, symmetric or asymmetric (with respect to the mean value), bell-shaped or skewed and so forth. Since regression coefficients for crash contributing factors within the SPFs can take any values in the set of real numbers ($\Re$), continuous distributions are considered for such coefficients. Since there are no particular restrictions for the regression coefficients of crash contributing factors, both symmetric and asymmetric distributions are attempted in this study.

The second step in distribution fitting is to estimate the parameter (s) of a distribution that have been hypothesized for the observations. The Maximum Likelihood method is commonly used for distribution fitting in which the parameters of the hypothesized distribution are estimated in such a way that the likelihood of estimated values being equal to the observed values is maximized. Kolmogorov-Smirnov (KS) test is then applied to assess the goodness of fit of a hypothesized distribution. The following hypotheses are tested through this test:

H0. The data follows the hypothesized distribution

H1. The data does not follow the hypothesized distribution

The KS test statistic is then used to accept or reject the null hypothesis. Given that *N* is the number of data observations, the KS statistic is expressed as in the following:

$$D_n = \text{supremum } |F(X_i) - F_n(X_i)| \quad 1 \le i \le N \tag{5}$$

where D_n is the test statistic for the KS test, $F(X_i)$ and $F_n(X_i)$ are theoretical and empirical cumulative distribution functions, respectively. If the KS statistic is higher than the critical statistic (i.e. the KS statistic associated with the 5% confidence level), the null hypothesis is rejected, suggesting that the data does not follow the hypothesized distribution (Sheskin, 2003).

3.2.2. Endogenous specification of dispersion parameter

In addition to distribution fitting, informative prior for the dispersion parameter can be developed through an alternative model specification in which the dispersion parameter is defined as a function of covariates. This model specification was first introduced by Miaou and Lord (Miaou and Lord, 2003) and then by Mitra and Washington (Mitra and Washington, 2007). They endogenously correlated the dispersion parameter with the traffic volume of the road via a non-linear function as follows:

$$Log(\phi_i) = \beta_0 + \beta_1 \times [Log(AADT_i) - mean(Log(AADT))] \tag{6}$$

This specification of Φ is replaced in Eq. (3) and thus instead of introducing informative priors for the dispersion parameter, only informative priors for β_0 and β_1 are used in this model specification.

3.2.3. Hypothetical priors

To further test the effects of empirical prior types and their strength on parameter estimates, hypothetical strong prior types are developed from the fitted distribution of prior types. In this study, the strength of priors was hypothetically increased by reducing the variance of the distribution obtained from the *distribution fitting* method. The mean of the empirical distribution was kept same. If the empirical distribution does not affect parameter estimates but the hypothetical distribution does, it is an indication of week empirical priors. It may eventually suggest the need to collect more information on that prior.

3.3. Testing for sample size

According to the literature (Lord and Miranda-Moreno, 2008), the effects of informative priors on parameter estimates and prediction ability of regression models are acute in small samples and become less influential as the sample size (N) increases. As such, the effects of prior types are also tested for five different sample sizes with $N = 903$, 100, 50, 30 and 15. Samples in the reduced dataset are randomly drawn from the total observed data ($N = 903$) and this random selection of sites are repeated 200 times to obtain statistically reliable estimates.

3.4. Criteria for model comparison

The effects of informative priors on SPFs are investigated by comparing the precision and bias of parameter estimates and goodness of statistical fit of SPFs developed using non-informative priors with their counterparts using different types of informative priors. The precision and bias of parameter estimates are compared based on the standard deviation and mean of their posteriors, respectively and the statistical fit of SPFs is compared with the help of goodness of fit measures. Bayesian models developed for a unique sample size can be compared using Deviance Information Criterion (DIC), which is the hierarchical modelling generalization of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) (Geedipally et al., 2014):

$$DIC = \bar{D}(\Theta) + P_D \tag{7}$$

where

$$\bar{D}(\Theta) = E[-2\log L]$$

$$P_D = \bar{D}(\Theta) - D(\bar{\Theta})$$

In the formulation above, L is the likelihood of the model at convergence, Θ is the total number of parameters, P_D is the effective number of parameters reflecting model complexity and $D(\bar{\Theta})$ is the deviance evaluated at a posterior summary of Θ. The model with a lower DIC is generally preferred.

4. Dataset

SPFs in this study were developed for a random sample of state controlled roads in Queensland, Australia, which include highway segments and major arterials (divided and undivided). These road segments were reviewed at the beginning of the study and short segments were either removed (in case they were outliers) or merged into other segments. However, the data cleaning process resulted in a few road segments that were shorter than the minimum recommended length in the Highway Safety Manual (AASHTO, 2010) but were still independent and homogenous. As such, these segments were kept in the dataset as independent road segments across the network. The final clean dataset contained 903 highway segments with the total extent of 2,790 km.

Five years of crash data from 2010 to 2014 with a total count of 7525 crashes along these segments were analysed. Crash contributing factors utilized in this study consisted of traffic characteristics and roadway geometric factors collected from Queensland Transport and Main Roads Department in GIS formats. The traffic characteristics included Annual Average Daily Traffic (AADT) and percentage of heavy vehicle traffic and the roadway geometric factors included segment length, functional classification of roads (urban/rural), number of lanes, shoulder width, presence of divided median, median width, and degree of curvature of roadway horizontal alignment. Table 1 presents descriptive statistics of the variables used for this study.

To develop informative priors for regression parameters, a rigorous literature review was conducted. The literature mostly consisted of SPFs developed for the U.S. jurisdictions but also included other regions such as Canada, Europe and Asia. The collected information through this review was sufficient for developing priors for AADT, percentage of heavy vehicle traffic, road segment length, number of lanes, shoulder width, median width, and degree of curvature of horizontal alignment in this study. The criteria to select these factors among others were *availability*, *reliability* and *sufficiency* of the past research conducted around them; not only there should be available research from reliable sources (i.e. published in well-known journals), the number of studies considered such factors must be sufficient so that the informative priors can be developed based on the obtained values. For example, while rainfall is an important environmental crash contributing factor, we could not find sufficient research around the coefficient of this factor in SPFs. Dispersion parameter of the NB model was another factor for which informative prior was developed. Table 2 presents these eight factors (seven regression parameters and dispersion parameter), their descriptive statistics of estimated coefficients in SPFs and the region in which these studies were conducted.

Table 1
Descriptive Statistics of the Data Used in the Study.

Variable	Mean	Std. Dev.	Min	Max
Crash	8	13	0	150
AADT	24846	30125	26	146357
% of Heavy vehicle traffic	11.151	8.801	0	50.150
Length (meters)	3091.083	3899.834	80	26109
Number of lanes	2.829	1.720	1	8
Shoulder width (meters)	0.986	1.405	0	15
Median width (meters)	3.562	4.800	0	38
Degree of curvature	0.804	1.080	0.030	10.517
Categorical variables	Frequency		Sample share (%)	
Functional classification of road – rural	340		37.65	
Presence of divided median	451		49.95	

Table 2
Regression Coefficients of Crash Contributing Factors Used in SPFs around the World.

Variable	Average	Standard deviation	Number of observations	Region of study	Facility type	Reference
AADT (Vehicle/Day)	0.708	0.466	25	U.S. (Florida, Indiana, Michigan, Virginia, Texas, Washington, California, Colorado, Delaware, Maryland, Minnesota, Oregon, Ohio, Connecticut, Alabama), Canada (Toronto, Edmonton), Italy	Urban and rural major arterials, Urban and rural interstate highways, rural two-lane (divided and undivided) highways, Rural four-lane (divided and undivided) highways	(Abdel-Aty and Radwan, 2000; Anastasopoulos and Mannering, 2009; Geedipally et al., 2012; Joshua and Garber, 1990; Lord and Bonneson, 2007; Malyshkina and Mannering, 2010; Miaou and Lum, 1993; Milton and Mannering, 1998; Lord et al., 2008; Shankar et al., 1997; Zou et al., 2013; Schneider IV et al., 2010; Persaud et al., 2004; Manuel et al., 2014; Islam et al., 2014; Mehta and Lou, 2013; Caliendo et al., 2007; Montella et al., 2008; Elvik, 2008; Montella and Imbriani, 2015; Washington et al., 2014)
Percentage of heavy vehicle traffic	0.479	0.623	9	U.S. (Indiana, Washington, Virginia, five other states), South Korea	Rural interstate highways, rural major and minor arterials, Rural two-lane and four-lane (divided and undivided) highways	(Anastasopoulos and Mannering, 2009; Milton and Mannering, 1998; Shankar et al., 1997; Washington et al., 2014; Joshua and Garber, 1990; Malyshkina and Mannering, 2010; Miaou and Lum, 1993)
Length (km)	0.939	0.697	16	U.S. (Florida, Indiana, Michigan, Virginia, Texas, Washington, California, Colorado, Delaware, Maryland, Minnesota, Oregon, Ohio, Connecticut, Alabama), Canada (Toronto, Edmonton), Italy, Norway, South Korea	Urban and rural major arterials, Urban and rural interstate highways, rural two-lane and four-lane (divided and undivided) highways	(Abdel-Aty and Radwan, 2000; Anastasopoulos and Mannering, 2009; Geedipally et al., 2012; Joshua and Garber, 1990; Lord et al., 2008; Malyshkina and Mannering, 2010; Milton and Mannering, 1998; Persaud et al., 2004; Islam et al., 2014; Manuel et al., 2014; Schneider IV et al., 2010; Caliendo et al., 2007; Mehta and Lou, 2013; Shankar et al., 1997)
Number of lanes	0.460	0.425	5	U.S. (Indiana, Washington), Norway	Rural highways, major and minor arterials	(Malyshkina and Mannering, 2010; Milton and Mannering, 1998; Shankar et al., 1997; Elvik, 2008)
Shoulder width	−0.040	0.274	13	U.S. (Florida, Indiana, Michigan, Washington, Texas, California, Colorado, Delaware, Maryland, Minnesota, Oregon, Ohio, Connecticut), Canada (Toronto)	Urban and rural major arterials, Urban and rural interstate highways, rural two-lane highways (divided and undivided)	(Abdel-Aty and Radwan, 2000; Geedipally et al., 2012; Islam et al., 2014; Lee and Mannering, 2002; Lord and Bonneson, 2007; Malyshkina and Mannering, 2010; Persaud et al., 2004; Schneider IV et al., 2010; Zou et al., 2013; Anastasopoulos and Mannering, 2009; Lu et al., 2013; Miaou and Lum, 1993; Milton and Mannering, 1998)
Median width	−0.093	0.126	6	U.S. (Florida, Indiana, Michigan, Washington)	Urban major arterials, Rural interstate highways, Urban and rural (two-lane) highways	(Abdel-Aty and Radwan, 2000; Anastasopoulos and Mannering, 2009; Lee and Mannering, 2002; Geedipally et al., 2012)
Degree of curvature	−0.026	0.274	6	U.S. (Florida, Indiana, Ohio), Italy, South Korea	Urban major arterials, Rural interstate highways, Rural two-lane and four-lane divided highways	(Abdel-Aty and Radwan, 2000; Anastasopoulos and Mannering, 2009; Caliendo et al., 2007; Miaou and Lum, 1993; Washington et al., 2014; Schneider IV et al., 2010)
Dispersion parameter	0.554	0.347	9	U.S. (Florida, Indiana, Georgia, Washington, Alabama,)	Urban major arterials, Rural interstate highways, Rural two-lane highways (divided and undivided)	(Abdel-Aty and Radwan, 2000; Anastasopoulos and Mannering, 2009; Lee and Mannering, 2002; Ma et al., 2008; Malyshkina and Mannering, 2010; Mehta and Lou, 2013; Milton and Mannering, 1998)

Table 3
Regression Results of the Random Parameters Negative Binomial Model estimated for the complete dataset (*N = 903*) via Maximum Simulated Likelihood Estimation.

Variable	Mean	St. E.	p-value	95% Confidence Interval
Non-random parameters				
Constant	−7.403	0.438	0.000	[−8.261, −6.544]
Segment length	0.693	0.027	0.000	[0.640, 0.746]
Number of lanes	0.082	0.019	0.000	[0.045, 0.120]
Shoulder width	−0.092	0.020	0.000	[−0.131, −0.053]
Mean of random parameters				
AADT	0.449	0.034	0.000	[0.382, 0.515]
Percentage of heavy vehicle traffic	−0.043	0.004	0.000	[−0.052, −0.035]
Median width	−0.027	0.006	0.000	[−0.039, −0.015]
Degree of curvature	0.129	0.029	0.000	[0.070, 0.185]
Standard deviation of random parameters				
AADT	0.032	0.003	0.000	[0.026, 0.038]
Percentage of heavy vehicle traffic	0.031	0.002	0.000	[0.027, 0.036]
Median width	0.030	0.005	0.000	[0.021, 0.040]
Degree of curvature	0.060	0.021	0.010	[0.018, 0.102]
Dispersion parameter				
phi	2.262	0.161	0.000	[1.947, 2.577]

5. results and discussion

5.1. Baseline for comparison: best linear unbiased estimates (BLUE)

The RPNB model was first estimated for the complete dataset in the frequentist approach. The complete dataset (*N = 903*) is assumed to be a proper representative of the population and thus the parameters of the RPNB model estimated on the complete dataset are the Best Linear Unbiased Estimates (BLUE) of the population (Washington et al., 2003). The significant explanatory variables were included in the model following a backward stepwise variable selection procedure. Variables were tested for multicollinearity by computing the Pearson correlation coefficients, and the variables with high correlation coefficients were excluded from the models. In addition, the interactions of explanatory variables were attempted in the model but were not statistically significant. The RPNB parameters are compared with the parameters estimated for different prior types in Bayesian approach. The frequentist estimation of the RPNB model was achieved by Maximum Simulated Likelihood Estimation (MSLE) with 500 Halton draws. Table 3 presents the estimated parameters of the RPNB model in which AADT and segment length were used as the measures of exposure.

As reported in Table 3, the parameters for segment length and number of lanes are non-random (fixed) and positive, indicating their increasing effect on crash counts. On the contrary, the parameter of shoulder width is negative and non-random, indicating the fixed decreasing effect of shoulder width on crash counts. The parameters of AADT, degree of curvature, percentage of heavy vehicle traffic and median width are random implying their varying effects across the road network. The mean parameters of AADT and degree of curvature are positive indicating that these factors have increasing effect (on average) on total crash count. The mean parameters of percentage of heavy vehicle traffic and median width are negative indicating that they have decreasing effect (on average) on the total crash count. However, the standard deviation for these parameters show that the effects of AADT, degree of curvature of horizontal alignment, percentage of heavy vehicle traffic and median width on crashes vary across segments. This variation may capture the differences in the effects of these variables across different roadway facility types (e.g. urban/rural or divided/undivided). Finally, the dispersion parameter is significant with positive sign suggesting that the data are over-dispersed and the negative binomial model is a suitable model.

5.2. Global informative priors

To investigate the effects of informative priors on Bayesian SPFs, seven different types of commonly used continuous distributions were tested to develop prior information for each of the crash contributing factors listed in Table 2. These distributions included beta, exponential, gamma, logistic, lognormal and normal distributions. For the dispersion parameter, Gamma distribution was considered as the relevant theories suggest it is the suitable conjugate distribution for the dispersion parameter (Lord and Miranda-Moreno, 2008). Fig. 1, illustrates fitted distributions for AADT, segment length, number of lanes and dispersion parameter.

Based on Kolmogorov-Smirnov test, only the fitted distributions (for which H_0 was not rejected at 5% significance level) were selected for each variable. Table 4 presents the distribution fitting results for the seven crash contributing factors as well as the dispersion parameter. The last column in Table 4 presents the mean, the 2.5th percentile and the 97.5th percentile of the globally developed priors (obtained from fitted distributions). The difference between the mean of globally developed prior distributions and locally estimated parameters (obtained from the RPNB model reported in Table 3) shows that in average, the Australian dataset is significantly different from the other regions based on which the prior information was collected (i.e. United States, Canada, Europe and South Korea). While this difference is moderate for the parameters of some crash contributing factors such as AADT, length, median width and dispersion parameter, it is more substantial for factors such as percentage of HV traffic, number of lanes, shoulder width and horizontal degree of curvature. This acute difference in global prior information with local data likelihood clearly indicates that the mean effect of crash contributing factors on the total crash count is statistically different for Australian road network compared to other parts of the world. This difference may be due to the driving behaviour differences and unobserved spatial factors that are omitted from the SPFs. It also highlights the fundamental shortcoming of the previous attempts in developing local priors using fixed parameters models: if a variable has varying effects on crashes in a sample of local data (as is the case in this study), then defining a fixed parameter for that variable and employing a distribution obtained from another region as the prior for that fixed parameter may not be appropriate because this approach inherently assumes that the effect of the variable is fixed over the two regions. A closer look into the 2.5th percentile and the 97.5th percentile of globally developed priors, however, reveals that the 95% interval for such priors overlap the 95% confidence intervals of BLUE estimates obtained from local data (Table 3) for almost all of the parameter estimates (except for the dispersion parameter). This interesting finding is the primary clue reflecting the potential of globally developed priors in improving locally developed parameter estimates. In other words, globally developed parameter estimates may be transferable into local conditions although the mean of such priors are different from the local data likelihood.

Endogenous specification of the dispersion parameter was achieved based on the studies that first used this approach in the literature (Mitra and Washington, 2007; Miaou and Lord, 2003). However, such studies have mainly been attempted for urban intersections in which the crash occurrence is different than that of road segments. Since no similar specification is available for road segments, the estimated parameters obtained in Mitra and Washington (2007) were used as hypothetical parameters for this type of informative prior (Mitra and Washington, 2007):

$$Log(\phi_i) = 0.960 - 0.125 \times [Log(AADT_i) - mean(Log(AADT))]$$

β_0 ~ Normal (0.960, 0.254)
β_1 ~ Normal (-0.125, 0.309)

To test the effects of the strength of prior information, the standard

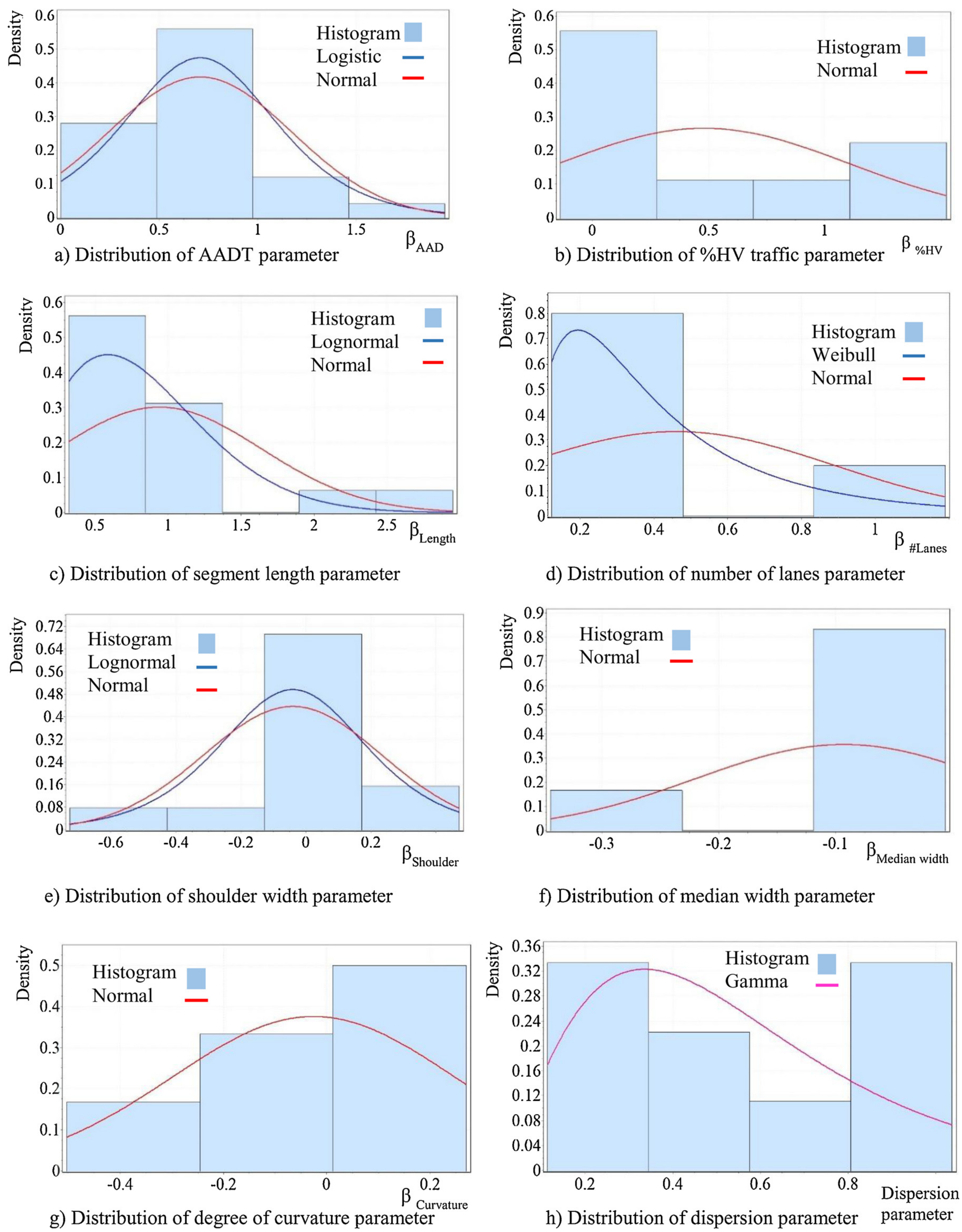

Fig. 1. Fitted distribution for the parameters of crash contributing factors.

deviations of the global informative priors were manipulated to generate an intuitive and concrete scenario for the means of comparison. In this research, the strength of each prior was increased by using a quarter of their standard deviations of the empirically fitted distributions reported in Table 4.

5.3. Bayesian models with non-informative and global informative priors

The RPNB model was then estimated in the Bayesian approach with non-informative priors and with various types of informative priors. For the model with non-informative priors, the priors for the regression parameters were assumed to be a Normal distribution with a mean of zero and a wide standard deviation of 100. The prior for the dispersion parameter was assumed to be a Gamma distribution with the shape and scale parameters of 1. For the models with informative priors, the following four scenarios were considered:

a: fitted Normal distributions for regression parameters and Gamma distribution for dispersion parameter

b: best fitted distributions for regression parameters and Gamma distribution for dispersion parameter

Table 4
Distribution Fitting Results of the Prior Information on Crash Contributing Factors.

Variable	Distribution (parameters)	Kolmogorov Smirnov (KS) Test			Empirical prior mean (95% Interval)
		Statistic	KS critical value ($KS_{\alpha=0.05}$)	Reject H_0	
AADT	Logistic (μ = 0.708, σ = 0.257)	0.104	0.264	No	0.708 (−0.233, 1.650)
	Normal (μ = 0.708, σ = 0.466)	0.126	0.264	No	
% of heavy vehicle traffic	Normal (μ = 0.479, σ = 0.623)	0.225	0.264	No	0.479 (−0.742, 1.700)
Segment length	Weibull (α = 1.781, β = 0.934)	0.162	0.264	No	0.760 (0.118, 1.944)
	Normal (μ = 0.760, σ = 0.275)	0.170	0.264	No	
Number of lanes	Lognormal* (μ = −1.074, σ = 0.752)	0.219	0.264	No	0.342 (0.078, 1.492)
Shoulder width	Logistic (μ=-0.041, σ = 0.151)	0.239	0.264	No	−0.041 (−0.512, 0.594)
	Normal (μ=-0.041, σ = 0.274)	0.251	0.264	No	
Median width	Normal (μ = −0.093, σ = 0.126)	0.389	0.264	Yes	−0.093 (−0.339, 0.154)
Degree of curvature	Normal (μ = −0.026, σ = 0.274)	0.218	0.264	No	−0.026 (−0.563, 0.511)
Dispersion parameter	Gamma (α = 2.540, β = 0.218)	0.168	0.264	No	0.436 (0.032, 1.625)

* Lognormal distribution is represented by the parameters of the corresponding normal distribution.

c: best fitted distributions for regression parameters and endogenous specification of the dispersion parameter

d: Hypothetical informative priors with the mean of best fitted distributions and quarter of their standard deviation

The Bayesian RPNB models for each of the above scenarios were estimated for five different sample sizes (N = 903, 100, 50, 30 and 15) to investigate the effects of prior types across various sample sizes (please note that for each sample size scenario, random sampling was repeated 200 times to obtain statistically reliable estimates). The MCMC simulation for the Bayesian RPNB models resulted in two Markov chains converging after 100,000 iterations. The convergence was ensured by visual monitoring as well as assessing the Gelman-Rubin statistics ($R_{\text{Gelman-Rubin}} \rightarrow 1$). Model updating was continued for 5000 iterations and the Bayesian inference was made accordingly. The effects of informative priors on Bayesian SPFs were investigated from two perspectives: precision and bias of parameter estimates and goodness of fit. Details of model comparison results are presented in the following.

5.4. Effects of global informative priors on the precision and Bias of parameter estimates

Table 5 illustrates the results of the Bayesian RPNB models with non-informative and informative priors. The mean of posterior parameter estimates in all scenarios with the smallest sample (*N = 15*) is significantly different from the BLUE estimates. However, their variances are substantially larger for the non-informative prior scenario compared to all informative prior scenarios. This finding is intuitive and indicates that non-informative priors lead to inaccurate estimates of parameters when the sample size is very small (*N = 15*) whereas informative priors help obtaining more accurate estimates of parameters in the same sample. As the sample size increases (*N = 30* and *50*), the means of posterior estimates become closer to the BLUE estimates and their variances become smaller for all scenarios; yet, the variances are still larger in the non-informative prior scenario compared to all informative prior scenarios.

The means of posterior parameter estimates in large samples (*N = 903* and *100*) with both non-informative and informative priors are very close to the BLUE estimates obtained from the RPNB model and their variances are very small. This finding indicates that non-informative and informative priors have negligible effects on the precision and bias of parameter estimates in large samples.

The parameter estimates of the percentage of heavy vehicle traffic, median width, shoulder width and degree of curvature are not sensitive to the developed informative priors in any of the scenarios. Overall, these findings suggest that globally developed informative priors –although their mean are significantly different from local estimates– can be helpful in improving the precision and bias of parameter estimates when the sample size is small.

5.5. Effects of global informative prior on the precision and Bias of dispersion parameter

Table 5 shows that the posterior estimates of the dispersion parameter in scenarios with informative prior obtained from distribution fitting has smaller standard deviation compared to the scenario with non-informative prior, irrespective of the sample size. This finding implies that global knowledge about over dispersion of crashes can increase the estimation precision of this parameter even in large sample sizes where the variance of local data likelihood is small. It is noteworthy that this finding shows the benefit of informative priors even for large sample sizes and is in contrast to the previous findings by Miranda-Moreno et al. (2013), presumably because they did not take unobserved heterogeneity of crash data into account. If not accounted for, the unobserved heterogeneity results in biased and inefficient parameter estimates (Mannering et al., 2016). Since explanatory variables are not completely independent, the bias and inefficiency of one parameter may result in bias and inefficiency of other parameters. As such, the sensitivity (or lack of sensitivity) of the regression parameters (measured in terms of bias or efficiency) to the informative priors in a fixed parameter model may be an artefact of not accounting for the unobserved heterogeneity in regression parameters. When the unobserved heterogeneity is accounted for, on the contrary, the true effects of informative priors on the regression parameters may be revealed (as it is the case in this study).

Although the standard deviation of the posterior estimates of dispersion parameter is larger in scenario with non-informative prior, the mean of the posterior estimate for dispersion parameter in this scenario is closer to the BLUE estimate compared to the scenarios with informative priors. This lack of suitability of the global informative prior in estimating the mean of local dispersion parameter might be due to the fact that these two pieces of information on dispersion parameter are statistically significantly different (i.e. the 95% interval of dispersion parameter obtained from the global prior does not overlap the 95% confidence interval of this parameter obtained from local data likelihood (Table 4). This finding supports our previous assertion that the unobserved factors in one region of the world may not be the same in another region.

5.6. Effects of global informative priors on goodness of fit

Another criterion to compare the Bayesian RPNB models with informative and non-informative priors is their goodness of fit. Table 6 shows the results of goodness of fit measures for the Bayesian RPNB with non-informative priors and the Bayesian RPNB models with informative priors across five different sample sizes. The Bayesian RPNB

Table 5
Parameter Estimates of AADT, Length and Dispersion Parameter in The Bayesian RPNB Model with Non-informative vs Informative Priors.

Sample size	Model type	AADT Mean (St. D.)	Length Mean (St. D.)	Number of lanes Mean (St. D.)	Dispersion parameter Mean (St. D)
15	B-RPNB **(Ninf)**	0.553 (1.437)	1.177 (1.679)	0.520 (1.861)	1.333 (1.042)
	B-RPNB **(a)**	0.661 (0.500)	0.770 (0.439)	0.325 (0.326)	0.811 (0.371)
	B-RPNB **(b)**	0.658 (0.358)	0.828 (0.420)	0.367 (0.266)	0.821 (0.373)
	B-RPNB **(c)**	0.628 (0.325)	0.794 (0.360)	0.349 (0.239)	3.182 (1.642)*
	B-RPNB **(d)**	0.707 (0.113)	0.410 (0.225)	0.367 (0.157)	0.486 (0.199)
30	B-RPNB **(Ninf)**	0.438 (0.350)	0.826 (0.388)	0.166 (0.312)	2.236 (1.202)
	B-RPNB **(a)**	0.591 (0.286)	0.786 (0.291)	0.210 (0.219)	1.153 (0.405)
	B-RPNB **(b)**	0.586 (0.228)	0.738 (0.267)	0.264 (0.156)	1.142 (0.405)
	B-RPNB **(c)**	0.550 (0.201)	0.739 (0.242)	0.248 (0.137)	3.554 (1.729)*
	B-RPNB **(d)**	0.678 (0.109)	0.486 (0.205)	0.321 (0.120)	0.739 (0.225)
50	B-RPNB **(Ninf)**	0.498 (0.196)	0.735 (0.200)	0.128 (0.156)	2.781 (1.247)
	B-RPNB **(a)**	0.519 (0.189)	0.733 (0.198)	0.137 (0.149)	1.410 (0.419)
	B-RPNB **(b)**	0.506 (0.160)	0.691 (0.187)	0.198 (0.098)	1.413 (0.417)
	B-RPNB **(c)**	0.508 (0.151)	0.719 (0.170)	0.197 (0.092)	3.844 (1.766)*
	B-RPNB **(d)**	0.642 (0.096)	0.553 (0.168)	0.271 (0.090)	0.922 (0.234)
100	B-RPNB **(Ninf)**	0.454 (0.102)	0.690 (0.117)	0.120 (0.085)	3.033 (1.182)
	B-RPNB **(a)**	0.462 (0.105)	0.710 (0.121)	0.108 (0.090)	1.717 (0.412)
	B-RPNB **(b)**	0.465 (0.096)	0.690 (0.114)	0.156 (0.065)	1.714 (0.414)
	B-RPNB **(c)**	0.477 (0.095)	0.717 (0.108)	0.147 (0.061)	3.833 (1.665)*
	B-RPNB **(d)**	0.578 (0.075)	0.613 (0.111)	0.209 (0.059)	1.228 (0.243)
903 (complete dataset)	B-RPNB **(Ninf)**	0.456 (0.052)	0.696 (0.028)	0.084 (0.025)	2.214 (0.410)
	B-RPNB **(a)**	0.462 (0.015)	0.695 (0.031)	0.084 (0.022)	2.027 (0.227)
	B-RPNB **(b)**	0.447 (0.021)	0.680 (0.026)	0.094 (0.023)	1.996 (0.290)
	B-RPNB **(c)**	0.461 (0.023)	0.692 (0.029)	0.090 (0.021)	2.984 (1.092)*
	B-RPNB **(d)**	0.488 (0.058)	0.677 (0.031)	0.110 (0.021)	1.835 (0.208)

B-RPNB: Bayesian Random Parameters Negative Binomial model.
Ninf: Non-informative priors.
a: fitted Normal distributions for regression parameters and Gamma distribution for dispersion parameter.
b: best fitted distributions for regression parameters and Gamma distribution for dispersion parameter.
c: best fitted distributions for regression parameters and endogenous specification of dispersion parameter.
d: hypothetical informative priors with the mean of best fitted distributions and quarter of their standard deviations.
* Posterior estimate of dispersion parameter.

models with informative priors—regardless of which scenario is used—have lower DIC values compared to the non-informative scenarios revealing the significant effect of informative priors in improving the statistical fit of SPFs. This improvement in fit is substantial in small sample sizes and is minimal in larger sample sizes.

5.7. Effects of prior type on the precision and Bias of parameter estimates and goodness of fit

To investigate the effects of informative prior type on parameter estimates and goodness of fit, scenarios were compared one by one. A comparison of scenario ***"c"*** with other informative scenarios (***"a"***, ***"b"*** and ***"d"***) in Table 5 shows that the standard deviation of posterior estimates of the dispersion parameter are substantially larger in scenario ***"c"*** compared to the other informative scenarios, irrespective of sample size. This finding indicates that endogenous specification of the dispersion parameter as a way of defining informative priors result in a less accurate estimate of dispersion parameter.

A comparison of posterior estimates of AADT, segment length and number of lanes in scenario ***"b"*** and their counterparts in scenario ***"a"*** shows that the former scenario has resulted in smaller standard deviation of posterior estimates, particularly in small samples ($N = 15, 30$ and 50). While this difference in standard deviations is minimal for segment length, it is more substantial for AADT and number of lanes. Please note that the only difference between these two scenarios is the distribution type for AADT, segment length and number of lanes. The outperformance of the former scenario shows that in contrast to the conventional assumption of Normal distribution, Logistic, Weibull and Lognormal distributions yield more accurate parameter estimates respectively for average annual daily traffic, segment length and number of lanes, particularly in small samples.

According to the measures of fit in Table 6, scenario ***"c"*** results in a lower DIC value compared to other scenarios showing its superiority in terms of statistical fit. Although this finding is in contrast to our earlier comment on the low precision (large variance) of dispersion parameter obtained from scenario ***"c"***, the outperformance of this scenario in statistical fit is presumably because endogenous specification of the dispersion parameter results in a mean of posterior estimate which is closer to the BLUE estimate ($\varphi = 2.262$) compared to the other scenarios. Scenario ***"a"*** and scenario ***"b"*** have almost equal DIC values indicating that prior distribution type on regression parameters does not influence the statistical fit in small samples. Lastly, to investigate the strength of the globally obtained priors for regression parameters and dispersion parameter, scenario ***"d"*** is compared with scenario ***"b"***. The only difference between these two scenarios is that scenario ***"d"*** has hypothetically smaller variances (a quarter of the obtained variances in scenario ***"b"***) for the regression parameters and dispersion parameter. This comparison reveals that increasing the strength of priors deteriorate the statistical fit of Bayesian RPNB models. To shed more light on this controversial finding, the 95% intervals of the hypothetical global priors and the 95% confidence intervals of the BLUE estimates (Australian data) are reviewed in Table 7.

The results show that the 95% intervals of hypothetical priors are mostly non-overlapped with the 95% confidence intervals obtained from Australian data. Cumming and Finch (2005) have shown that if the 95% confidence intervals of two independent samples (global informative priors and Australian data in our study) have less than 50% overlap, their mean are significantly different at 5% significance level. A comparison of the proportion overlaps between hypothetical priors and Australian data (Table 7) shows that the mean of hypothetical priors for all parameters are significantly different in these two sets. This finding implies that reducing the variance of informative priors

Table 6
Goodness of Fit Measures across Sample Size and Prior Types.

Sample Size	Model Type	DIC
15	B-RPNB **(Ninf)**	128.015
	B-RPNB **(a)**	88.961
	B-RPNB **(b)**	88.802
	B-RPNB **(c)**	86.451
	B-RPNB **(d)**	91.261
30	B-RPNB **(Ninf)**	183.807
	B-RPNB **(a)**	177.067
	B-RPNB **(b)**	177.614
	B-RPNB **(c)**	174.417
	B-RPNB **(d)**	177.337
50	B-RPNB **(Ninf)**	309.201
	B-RPNB **(a)**	293.204
	B-RPNB **(b)**	292.390
	B-RPNB **(c)**	289.434
	B-RPNB **(d)**	296.848
100	B-RPNB **(Ninf)**	577.050
	B-RPNB **(a)**	573.110
	B-RPNB **(b)**	573.560
	B-RPNB **(c)**	570.919
	B-RPNB **(d)**	576.598
903	B-RPNB **(Ninf)**	4811
	B-RPNB **(a)**	4741
	B-RPNB **(b)**	4784
	B-RPNB **(c)**	4607
	B-RPNB **(d)**	4820

B-RPNB: Bayesian Random Parameters Negative Binomial model.
Ninf: Non-informative priors.
a: fitted Normal distributions for regression parameters and Gamma distribution for dispersion parameter.
b: best fitted distributions for regression parameters and Gamma distribution for dispersion parameter.
c: best fitted distributions for regression parameters and endogenous specification of dispersion parameter.
d: hypothetical informative priors with the mean of best fitted distributions and quarter of their standard deviations.

Table 7
Comparison of 95% Interval in Hypothetical Prior with 95% Confidence Interval of RPNB Model.

Variable	95% Interval of hypothetical empirical prior	95% Confidence Interval of BLUE estimates	Proportion overlap
AADT	[0.473, 0.944]	[0.382, 0.515]	0.075
% of Heavy vehicle traffic	[0.403, 1.013]	[−0.052, −0.035]	0.000
Segment length	[0.412, 1.004]	[0.640, 0.746]	0.179
Number of lanes	[0.420, 0.996]	[0.045, 0.120]	0.000
Shoulder width	[0.549, 0.867]	[−0.131, −0.053]	0.000
Median width	[0.646, 0.768]	[-0.039, -0.015]	0.000
Degree of curvature	[0.574, 0.842]	[0.070, 0.185]	0.000

may increase the strength of such priors but may alter their mean effects to become significantly different from Australian data and so may have adverse effect on the statistical fit of SPFs.

6. Conclusions

Safety performance functions are typically calibrated on a spatially and temporally representative set of local data to estimate the effects of contributing factors to crashes. These estimates might be biased, inconsistent and inefficient if sufficient data are not available. Due to the differences in spatial and geographic characteristics of different regions, the effect of crash contributing factors may not be consistent across regions, cities, or countries. These inconsistencies have typically impeded the use of SPFs developed elsewhere for the local case. As demonstrated in this research, a promising approach to overcome data insufficiency is combining local data with globally derived priors. By applying Bayesian inference, such priors can significantly improve the bias and precision of local parameter estimates in small samples.

This research firstly demonstrates that on average, globally obtained priors (mostly from the United States) on parameters of safety factors were significantly different than those derived using Australian data—implying that the mean effect of crash contributing factors is inconsistent across cities, regions, and countries. This lack of consistency is likely due to the unobserved factors that are omitted from the SPFs, differences in driver behaviour, and differences in driving laws and regulations. However, the 95% confidence intervals of these two pieces of information (i.e. global priors and local data) are overlapping, indicating that common unobserved factors may exist in Australia and other parts of the world in spite of such differences. As such, globally obtained informative priors were intuitively found to have the potential to improve the parameter estimates and statistical fit of SPFs. In testing such potential, we found that the precision of parameter estimates in SPFs developed for Australian data were improved by incorporating proper types of informative priors. Appropriate informative priors are particularly helpful in small samples. This improved precision was reflected in smaller variance of posterior parameter estimates and improved statistical fit of the SPFs in small samples when using informative priors. It is seemingly more effective to employ "lumpy" empirical distributions as priors rather than assumed "well behaved' parametric distributions. Finally, we found that hypothetically increasing the strength of global informative priors may preclude them from improving precision and bias of Australian parameter estimates, and may have adverse effects on the statistical fit of SPFs developed for Australian data.

The practical implication of this study lies in the advantage of using global informative priors in accurately estimating Australian safety performance functions which leads to improved identification of crash blackspots and ultimately results in more effective allocation of resources. The findings imply that there is still benefit in using the information obtained from other parts of the world even if their transportation context is different from Australia. This benefit is substantial if Australian data are not sufficiently available to estimate the impact of observed and unobserved factors contributing to crashes.

This study is not without limitations. The aim of this research was to investigate the influence of global informative priors on SPFs developed for Australian data. As such, it is not known whether the inferences and conclusions made for this data retain if the analysis is repeated for data of another country. In addition, the effects of global informative priors on SPFs with that of non-informative priors were only tested. The effects of Australian informative priors could not be tested as very limited number of studies developed SPFs for Australian datasets. Future research may be dedicated to develop informative priors with the local studies of a country and compare the performance of SPFs using the proposed three types of priors.

Acknowledgment

This research was supported under Australian Research Council's *Discovery Project* funding scheme (DP140101653). The authors would like to acknowledge the department of Transport and Main Roads (TMR) in Queensland, Australia for their support for data collection of this study.

References

AASHTO, 2010. The Highway Safety Manual. Association of State Highway Transportation Professionals, Washington, DC.
Abdel-Aty, M.A., Radwan, A.E., 2000. Modeling traffic accident occurrence and involvement. Accid. Anal. Prev. 32, 633–642.

Afghari, A.P., Haque, M.M., Washington, S., Smyth, T., 2016. Bayesian latent class safety performance function for identifying motor vehicle crash black spots. Transportation Research Record: J. Transport. Res. Board 2601 (1), 90–98.

Ahmed, M., Huang, H., Abdel-Aty, M., Guevara, B., 2011. Exploring a Bayesian hierarchical approach for developing safety performance functions for a mountainous freeway. Accid. Anal. Prev. 43, 1581–1589.

Anastasopoulos, P.C., Mannering, F.L., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accid. Anal. Prev. 41, 153–159.

Caliendo, C., Guida, M., Parisi, A., 2007. A crash-prediction model for multilane roads. Accid. Anal. Prev. 39, 657–670.

Chen, Y., Persaud, B., Sacchi, E., 2012. Improving transferability of safety performance functions by Bayesian model averaging. Transportation Research Record: J. Transport. Res. Board 2280, 162–172.

Cumming, G., Finch, S., 2005. Inference by eye: confidence intervals and how to read pictures of data. Am. Psychol. 60 (2), 170.

Elvik, R., 2008. The predictive validity of empirical Bayes estimates of road safety. Accid. Anal. Prev. 40, 1964–1969.

Farid, A., Abdel-Aty, M., Lee, J., Eluru, N., Wang, J.-H., 2016. Exploring the transferability of safety performance functions. Accid. Anal. Prev. 94, 143–152.

Farid, A., Abdel-Aty, M., Lee, J., Eluru, N., 2017. Application of Bayesian informative priors to enhance the transferability of safety performance functions. J. Safety Res. 62, 155–161.

Geedipally, S.R., Lord, D., Dhavala, S.S., 2012. The negative binomial-Lindley generalized linear model: characteristics and application using crash data. Accid. Anal. Prev. 45, 258–265.

Geedipally, S.R., Lord, D., Dhavala, S.S., 2014. A caution about using deviance information criterion while modeling traffic crashes. Saf. Sci. 62, 495–498.

Geedipally, S.R., Shirazi, M., Lord, D., 2017. Exploring the need for having region-specific calibration factors. Transportation Research Record: J. Transport. Res. Board 2636, 73–79.

Hadayeghi, A., Shalaby, A.S., Persaud, B.N., Cheung, C., 2006. Temporal transferability and updating of zonal level accident prediction models. Accid. Anal. Prev. 38, 579–589.

Harwood, D.W., Council, F., Hauer, E., Hughes, W., Vogt, A., 2000. Prediction of the Expected Safety Performance of Rural Two-lane Highways. US Transportation Collection, Federal Highway Administration, Technical Report No. FHWA-RD-99-207.

Hauer, E., 1992. Empirical Bayes approach to the estimation of “unsafety”: the multivariate regression method. Accid. Anal. Prev. 24, 457–477.

Hauer, E., Persaud, B.N., Smiley, A., Duncan, D., 1991. Estimating the accident potential of an Ontario driver. Accid. Anal. Prev. 23, 133–152.

Heydari, S., Miranda-Moreno, L., Amador, L., 2013. Does prior specification matter in hotspot identification and before-after road safety studies? Transportation Research Record: J. Transport. Res. Board 2392, 31–39.

Heydari, S., Miranda-Moreno, L.F., Lord, D., Fu, L., 2014. Bayesian methodology to estimate and update safety performance functions under limited data conditions: a sensitivity analysis. Accid. Anal. Prev. 64, 41–51.

Huang, H., Abdel-Aty, M., 2010. Multilevel data and Bayesian analysis in traffic safety. Accid. Anal. Prev. 42, 1556–1565.

Islam, M., Ivan, J., Lownes, N., Ammar, R., Rajasekaran, S., 2014. Developing safety performance function for freeways by considering interactions between speed limit and geometric variables. Transportation Research Record: J. Transport. Res. Board 2435, 72–81.

Joshua, S.C., Garber, N.J., 1990. Estimating truck accident rate and involvements using linear and Poisson regression models. Transp. Plan. Technol. 15, 41–58.

Kaaf, K.A., Abdel-Aty, M., 2015. Transferability and calibration of Highway Safety Manual performance functions and development of new models for urban four-lane divided roads in Riyadh, Saudi Arabia. Transp. Res. Rec. 2515, 70–77.

Lee, J., Mannering, F., 2002. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. Accid. Anal. Prev. 34, 149–161.

Li, J., Wang, X., Abdel-Aty, M., Yu, R., 2018. Exploring the Transferability of Cross-country Safety Performance Functions for Urban Arterials With Pooled Data. Transportation Research Board Annual Meeting, Washington D.C., United States No. 18-4325.

Lord, D., Bonneson, J.A., 2007. Development of accident modification factors for rural frontage road segments in Texas. Transportation Research Record: J. Transport. Res. Board 2023, 20–27.

Lord, D., Miranda-Moreno, L.F., 2008. Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of Poisson-gamma models for modeling motor vehicle crashes: a Bayesian perspective. Saf. Sci. 46, 751–770.

Lord, D., Guikema, S.D., Geedipally, S.R., 2008. Application of the Conway–Maxwell–Poisson generalized linear model for analyzing motor vehicle crashes. Accid. Anal. Prev. 40, 1123–1134.

Lu, J., Gan, A., Haleem, K., Wu, W., 2013. Clustering-based roadway segment division for the identification of high-crash locations. J. Transp. Saf. Secur. 5, 224–239.

Ma, J., Kockelman, K.M., Damien, P., 2008. A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. Accid. Anal. Prev. 40, 964–975.

Malyshkina, N.V., Mannering, F.L., 2010. Zero-state Markov switching count-data models: an empirical assessment. Accid. Anal. Prev. 42, 122–130.

Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16.

Manuel, A., El-Basyouny, K., Islam, M.T., 2014. Investigating the safety effects of road width on urban collector roadways. Saf. Sci. 62, 305–311.

Mehta, G., Lou, Y., 2013. Calibration and development of safety performance functions for Alabama: two-lane, two-way rural roads and four-lane divided highways. Transportation Research Record: J. Transport. Res. Board 2398, 75–82.

Miaou, S.-P., Lord, D., 2003. Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and Bayes versus empirical Bayes methods. Transportation Research Record: J. Transport. Res. Board 1840, 31–40.

Miaou, S.-P., Lum, H., 1993. Modeling vehicle accidents and highway geometric design relationships. Accid. Anal. Prev. 25, 689–709.

Milton, J., Mannering, F., 1998. The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. Transportation 25, 395–413.

Miranda-Moreno, L.F., Heydari, S., Lord, D., Fu, L., 2013. Bayesian road safety analysis: incorporation of past evidence and effect of hyper-prior choice. J. Safety Res. 46, 31–40.

Mitra, S., Washington, S., 2007. On the nature of over-dispersion in motor vehicle crash prediction models. Accid. Anal. Prev. 39, 459–468.

Montella, A., Imbriani, L.L., 2015. Safety performance functions incorporating design consistency variables. Accid. Anal. Prev. 74, 133–144.

Montella, A., Colantuoni, L., Lamberti, R., 2008. Crash prediction models for rural motorways. Transportation Research Record: J. Transport.Res. Board 2083, 180–189.

Park, E.S., Lord, D., 2007. Multivariate Poisson-lognormal models for jointly modeling crash frequency by severity. Transportation Research Record: J. Transport. Res. Board 2019, 1–6.

Persaud, B., Lyon, C., Nguyen, T., 1999. Empirical Bayes procedure for ranking sites for safety investigation by potential for safety improvement. Transportation Research Record: J. Transport. Res. Board 1665, 7–12.

Persaud, B., Lord, D., Palmisano, J., 2002. Calibration and transferability of accident prediction models for urban intersections. Transportation Research Record: J. Transport. Res. Board 1784, 57–64.

Persaud, B.N., Retting, R.A., Lyon, C.A., 2004. Crash reduction following installation of centerline rumble strips on rural two-lane roads. Accid. Anal. Prev. 36, 1073–1079.

Ricci, V., 2005. Fitting Distributions With R. Contributed Documentation Available on CRAN 96.

Rusli, R., Haque, M.M., Afghari, A.P., King, M., 2018. Applying a random parameters Negative Binomial Lindley model to examine multi-vehicle crashes along rural mountainous highways in Malaysia. Accid. Anal. Prev. 119, 80–90.

Sacchi, E., Sayed, T., 2016. Bayesian estimation of conflict-based safety performance functions. J. Transp. Saf. Secur. 8 (3), 266–279.

Sacchi, E., Persaud, B., Bassani, M., 2012. Assessing international transferability of Highway Safety Manual crash prediction algorithm and its components. Transportation Research Record: J. Transport. Res. Board 2279, 90–98.

Saha, D., Alluri, P., Gan, A., 2017. A Bayesian procedure for evaluating the frequency of calibration factor updates in highway safety manual (HSM) applications. Accid. Anal. Prev. 98, 74–86.

Sawalha, Z., Sayed, T., 2006. Transferability of accident prediction models. Saf. Sci. 44, 209–219.

Schneider Iv, W., Savolainen, P., Moore, D., 2010. Effects of horizontal curvature on single-vehicle motorcycle crashes along rural two-lane highways. Transportation Research Record: J. Transport. Res. Board 2194, 91–98.

Shankar, V., Milton, J., Mannering, F., 1997. Modeling accident frequencies as zero-altered probability processes: an empirical inquiry. Accid. Anal. Prev. 29, 829–837.

Sheskin, D.J., 2003. Handbook of Parametric and Nonparametric Statistical Procedures. CRC Press, United States.

Wang, J.-H., Abdel-Aty, M.A., Lee, J., 2017. Evaluation of safety performance functions based on experimental design using a cross-validation method. Transportation Research Record: J. Transport. Res. Board 2636, 62–72.

Washington, S., Oh, J., 2006. Bayesian methodology incorporating expert judgment for ranking countermeasure effectiveness under uncertainty: example applied to at grade railroad crossings in Korea. Accid. Anal. Prev. 38, 234–247.

Washington, S., Karlaftis, M., Mannering, F., 2003. Statistical and Econometric Techniques for Transportation Data Analysis. CRC/Chapman Hall Press, New York, NY.

Washington, S., Haque, M.M., Oh, J., Lee, D., 2014. Applying quantile regression for modeling equivalent property damage only crashes to identify accident blackspots. Accid. Anal. Prev. 66, 136–146.

Xu, C., Wang, W., Liu, P., Li, Z., 2015. Calibration of crash risk models on freeways with limited real-time traffic data using Bayesian meta-analysis and Bayesian inference approach. Accid. Anal. Prev. 85, 207–218.

Yu, R., Abdel-Aty, M., 2013. Investigating different approaches to develop informative priors in hierarchical Bayesian safety performance functions. Accid. Anal. Prev. 56, 51–58.

Zou, Y., Zhang, Y., Lord, D., 2013. Application of finite mixture of negative binomial regression models with varying weight parameters for vehicle crash data analysis. Accid. Anal. Prev. 50, 1042–1051.