

# Modelling Low Animal-Vehicle Collision Counts across Large Networks

Zili Li ([zili.li@uq.edu.au](mailto:zili.li@uq.edu.au)) University of Queensland  
& Kara Kockelman ([kkockelm@mail.utexas.edu](mailto:kkockelm@mail.utexas.edu)) UT Austin

Presented at the November 2018 North American Meetings  
of the Regional Science Association International (RSAI) in San Antonio,  
Texas and under review for publication in *Accident Analysis & Prevention*

---

## Abstract

Animal-vehicle collisions (AVCs) are common around the world and result in considerable loss of animal and human life, as well as significant property damage and regular insurance claims. For this reason, understanding their occurrence in relation to various contributing factors and being able to identify locations of high risk are valuable to AVC prevention and economic, social and environmental cost savings. However, many challenges exist in the study of AVC datasets. These include seasonality of animal activity, very low counts across many sections of extensive roadway networks, and computational burdens that come with discrete response analysis across large data sets. In recognizing these challenges, a modelling framework is proposed in this study both to infer the impact of segment specific contributing factors and to identify high-risk locations. The utility of the model is demonstrated by using the AVC dataset collected in the state-controlled roadway network system of Texas, U.S. (with more than 100,000 segments). It is demonstrated that the proposed model can not only deal with the stated challenges reasonably well, but also allow high-risk location identification to be done in relation to various impacting factors. In this way, useful information on the efficacy of various measures can be obtained for future preventative purposes.

*Keywords:* Animal-vehicle collisions, crash count modelling, seasonality, low-count data, transportation networks.

---

## Introduction

Animal-vehicle collisions (AVCs) are common around the world and result in considerable loss of animal and human life, as well as significant property damage and regular insurance claims (Bruinderink and Hazebroek 1996; Al-Ghamdi and AlGadhi 2004; Seiler 2005; Klöcker, Croft, and Ramp 2006; Mountrakis and Gunson 2009; Sullivan 2011; Mrtka and Borkovcová 2013). For this reason, there has been a continued research interest toward understanding the occurrence of AVCs and the effectiveness of various prevention measures (Gunson, Mountrakis, and Quackenbush 2011).

In terms of the occurrence of AVCs, special attention has been paid to both the spatial and temporal characteristics of AVCs. In the spatial dimension, the focus has been on identifying the relationship between the locations of AVCs and animal species' natural habitats (Hurley, Rapaport, and Johnson 2009; Gkritza, Baird, and Hans 2010; Neumann et al. 2012) and nearby landscapes (Malo, Suárez, and Díez 2004; Grilo, Bissonette, and Santos-Reis 2009; Danks and Porter 2010; Jensen, Gonser, and Joyner 2014). In the temporal dimension, the multiple seasonal patterns at different time resolutions have also become well recognized. For example, in the time span of one day, the daily activity pattern of both animals and humans can result in an intra-day pattern in the recorded AVCs (Haikonen and Summala 2001; Rowden, Steinhardt, and Sheehan 2008; Diaz-Varela et al. 2011; Neumann et al. 2012). Within a year, factors such as the migratory behavior of animals, variations in daylight and climate conditions can also play a role (Garrett and Conway 1999; Rodríguez-Morales, Díaz-Varela, and Marey-Pérez 2013; Hothorn et al. 2015; Niemi et al. 2017).

With regards to the effectiveness of various preventative measure for AVCs, warning signs (Ujvari et al. 1998), light-reflecting devices (Brieger et al. 2016), fencing and barriers (Leblond et al. 2007; Zuberogoitia et al. 2015), bypass or underpass (Huijser et al. 2009; McCollister and van Manen 2010), modification of nearby landscapes (Grosman et al. 2009), etc. are among those that has been investigated. There are also studies that focus on the impact of various roadway characteristics such as speed limit (Found and Boyce 2011; Meisingset et al. 2014), road width (Litvaitis and Tash 2008), shoulder width (Lao, Wu, et al. 2011), number of lanes (Lao, Zhang, et al. 2011), etc. on the occurrence of AVCs.

Moreover, due to the important role of the information regarding the exact locations with high exposure to AVC risk in guiding the implementation of various preventative measures, hotspot identification also forms an important research area in the domain of safety research. Different from all the studies mentioned previously that

use highly aggregated measures over the network for both AVC counts and other attributes, hotspot identification by definition is interested in road segments of smaller size. Consequently, disaggregation of the network on a higher level is applied, which normally lead to a large number of road segments. In order to deal with the increasing number of segments in the data, two common approaches are normally taken: First, a computationally simple approach is used for the reason of computational feasibility. For example, Kolowski and Nielsen (2008) use correlation coefficient to define the similarity between any road segment and a road segment with AVC occurrences, and classify hotspots according to the value of the correlation. Alternatively, a degree of smoothing can be applied for all the segments, in this direction kernel methods are popular (Ramp, Wilson, and Croft 2006; Snow, Williams, and Porter 2014; Bíl et al. 2016).

As suggested by Snow et al. (2014), these approaches of hotspot identification normally require a large amount of subjective inputs for the implementation, and can result in unreliable inference. More importantly, the relative importance of each attribute for identifying hotspots is generally treated as equal or unknown. For this reason, scenario evaluations based on specific attributes can be unreliable or impossible.

To overcome these issues, our study continues the exploration in the direction of hotspot identification for AVCs by using a Bayesian regression framework. In this framework, special attention is paid to the seasonality of animal activity (e.g., deer-crashes peaking in October and November), the very low counts (many zero-crash segments) across many sections of our extensive roadway networks, and the computational burdens that comes with discrete response analysis across large data sets (e.g., 100,000 reasonably homogeneous segments distinguished in the Texas Department of Transportation’s state-maintained network).

By using the proposed modelling framework, we demonstrate how hotspots can be identified and related to the seasonality pattern of AVCs based on the estimated model. Moreover, inference and scenario evaluations can also be made based on any segment specific attributes (e.g. speed limit changes). We believe this information can prove to be helpful to the road network operator in terms of AVC prevention purposes.

## 1. Animal-vehicle collisions in Texas

The dataset used in this study consists of information derived from two sources. First, the AVC records made available on the crash records information system maintained by the Texas Department of Transportation (<https://cris.dot.state.tx.us/public/Query/app/public/welcome>). Second, the information regarding segment-specific roadway design factors are obtained from the Texas Department of Transportation website, <https://www.txdot.gov/inside-txdot/division/transportation-planning/roadway-inventory.html>.

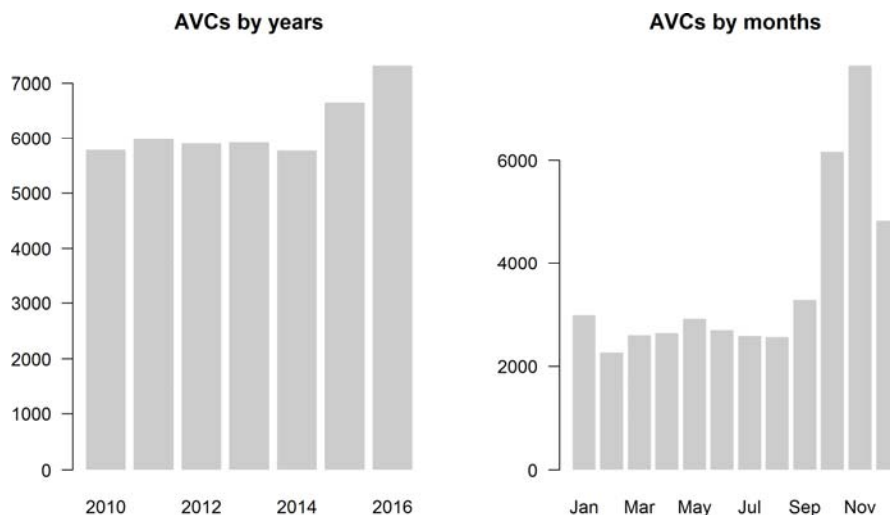


Figure 1: The number of AVCs by years (the left panel) and by months (the right panel) recorded during the period from 2010 to 2016 on the state-maintained roadway system in Texas, U.S.

In the AVC dataset, the AVCs recorded during the period from 2010 to 2016 on the state-maintained roadway system are used in this study. A total number of 43,319 AVCs for the entire period of seven years. Figure 1 shows the number of AVCs by year (the left panel) and by month (the right panel). It can be seen that there is a small increase in the number of total collisions in more recent years. More interestingly, the number of AVCs is shown to be far greater during the months from October to December than in other months. This monthly difference in the number of AVCs

clearly indicates a seasonal pattern in the occurrence of AVCs that may be attributed to seasonal patterns in animal activity (Bruinderink and Hazebroek 1996; Sullivan 2011; Niemi et al. 2017).

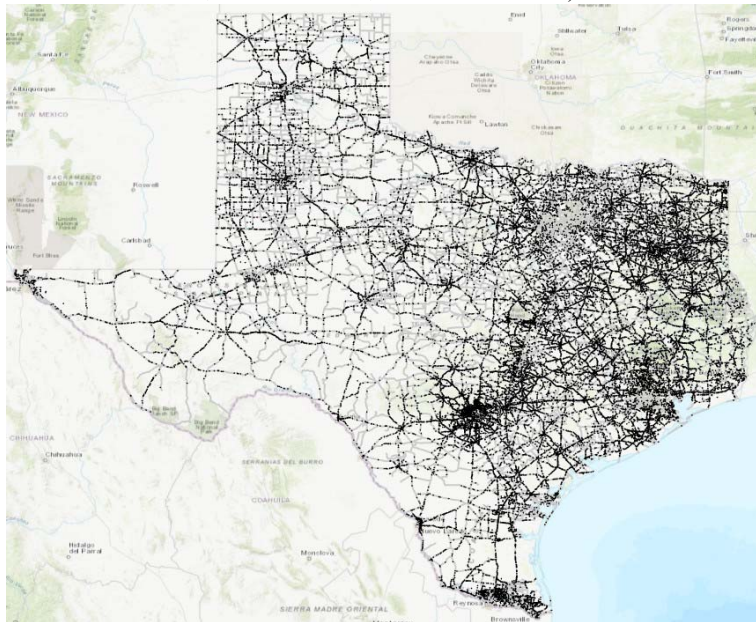


Figure 2: The state-maintained roadway network (the grey colored lines) in Texas, U.S. All the AVCs recorded from 2010 to 2016 are denoted as black dots.

Looking at the data from another perspective, the locations of the AVCs are mapped onto the 120,719 segments of state-maintained roadway and shown in Figure 2.<sup>1</sup> Here, it is clear that a larger portion of the AVCs are located on the east side of the state around several urbanized areas.

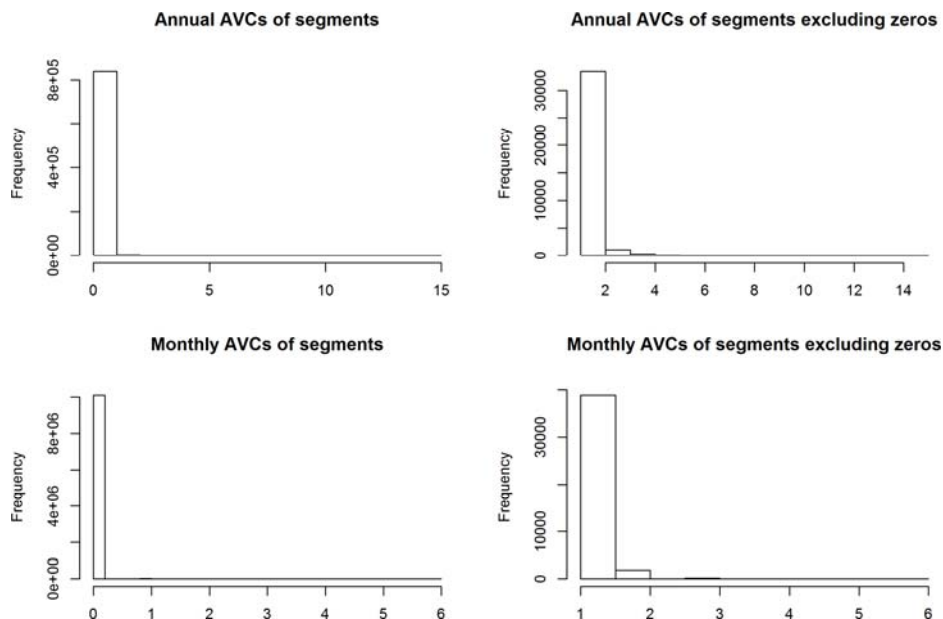


Figure 3: The total numbers of AVCs in a year (the upper panel) and in a month (the lower panel) during the period from 2010 to 2016 for all segments (the left panel) and for segments excluding those with zero AVC count (the right panel).

By disaggregating the AVCs over a large number of segments as shown in Figure 2, the upper panel of Figure 3

<sup>1</sup> A small percentage of AVC displayed are not occurred on the network system. The total number of off-system AVCs is 5930, which account for around 12 percent of the total AVCs recorded.

shows that AVCs only occur rarely at majority of the individual segments. Moreover, if the seasonality is to be considered, the further disaggregation in the time dimension can lead to the monthly AVC counts to be very low for all the segments. In the lower panel of Figure 3, it shows that only around 0.4% (40,953/10,140,396) of the monthly segment-wise AVC counts are non-zero. Among the non-zero monthly segment-wise AVCs (the right of the lower panel), around 5% (2,084/40,953) of them have the observed AVCs of more than one with the maximum being at six. Consequently, two challenges emerge from making inference on the effect of segment specific factors based on the highly disaggregated data. Firstly, the computation involved would increase dramatically due to the number of observations (120,719 segments  $\times$  7 years  $\times$  12 months = 10,140,396). Secondly, the segment-wise AVC counts at a monthly interval is both sparser and highly variable compared with that at a higher aggregation level in the time and spatial dimensions.

Furthermore, Figure 4 shows the AVCs recorded on the segments in a small area. In this figure, the AVCs are denoted as circles (one for each occurrence), whereas segments are denoted as solid black lines with the two ends indicated by cross bars. It is clear that in this area, a majority of segments have zero AVCs recorded. Consequently, the traditional way of allowing spatial correlation between spatial units may become inappropriate. That is, the spatial correlation is expected to be close to zero due to the existence of many segments with zero AVC counts in between the others with non-zero AVC counts. If a non-zero spatial correlation is imposed, the actual effect of any segment specific factor would be biased downwards due to the smoothing effect.

Moreover, the zero counts of AVCs shown in Figure 4 may indicated a degree of heterogeneity between the segments, which should be treated individually. For example, in the segments located in the bottom left (or top left) of Figure 4, it is possible that some of them are elevated bypasses. Due to this unique but unknown characteristics, the occurrence of AVCs would be highly unlikely due to the potentially zero number of animal road crossing. This implies that the effect of a specific design factor should in large part only have an impact on the occurrence of AVCs on a selected number of segments in a large network. In this case, the need to identify the heterogeneity individually for a large number of segments further complicate the computation.



Figure 4: The recorded AVCs for the period from 2010 to 2016 shown as circles for segments in a small area in Texas, U.S. The end points of segments are indicated by cross bars.

In summary, the preliminary data analysis presented in this section reveals several aspects that are important in the modelling of AVCs and making inferences on the effect of segment specific design factors on the occurrences of AVCs. These include: the seasonal pattern of AVCs, the sparse and highly variable nature of the information contained in the monthly and segment-wise disaggregated AVC counts, and the potential heterogeneity between segments.

## 2. The modelling framework

To deal with the important aspects identified in the previous section, we start by first assigning a binomial distribution for the number of AVCs at each segment  $s$  and each month  $t$ , so the probability of having  $k$  number of AVCs,

$$P(k, n_{s,t}, p_s) = \binom{n_{s,t}}{k} p_s^k (1 - p_s)^{n_{s,t}-k}.$$

In which,  $n_{s,t}$  is the number of Bernoulli trials (animal road crossing) and it can differ among segments ( $s$ ) and month ( $t$ ). Also  $p_s$  is the Bernoulli trial probability, which may differ among segments.

Differing from the common choice of using negative binomial distribution for modelling collision counts in the literature, the use of binomial distribution is more appropriate in present case. More specifically, the negative binomial

distribution describes the probability of the number of failures (animal road crossings without causing collisions) before the observed number of successes (animal road crossings leading to collisions) is observed. By definition, the Bernoulli trial probability can only lie in the open interval between zero and one, and the number of failures, hence the total number of Bernoulli trials is strictly greater than zero. In contrast, the binomial distribution describes the probability of the number of AVCs given the number total trials (animal road crossings). And in this case, the total number of trials can take value zero, which can be the case for segments that are elevated bypasses, for example. Consequently, the use of the binomial distribution allows the model to behave similarly to a zero-inflated regression model. By allowing segments to have the number of animal crossings taking the value of zero, it can both accommodate the potential heterogeneity between segments (as suggested in Figure 4) to a certain degree and allows the inference on the trial probabilities to be made only based on those segments with non-zero animal road crossings.

Using the binomial distribution with two parameters  $n_{s,t}$  and  $p_s$ , the number of AVCs,  $k$  can be thought as the result of repeated  $n_{s,t}$  Bernoulli trials. Each trial represents an animal road crossing with probability of  $p_s$  causing an AVC. For this reason, the collision probability,  $p_s$  can be regarded as a quantity that is determined by segment specific characteristics such as the design factors. A convenient way is to write,

$$p_s = \frac{1}{1 + \exp(-\psi_s)}, \quad (1)$$

where  $\psi_s$  is a quantity that determines the probability of AVC for each animal road crossing at segment  $s$ , and can be represented as a linear combination of unknown parameters and segment specific design factors. In addition, we have also argued that there may be heterogeneity in this probability due to the omission of important segment specific influencing factors. In order to accommodate this possibility while at the same time recognizing the fact about the limited observations available in the dataset for each segment (only one set of values for the design factors of each segment), a two-component clustering is used for the intercept, so

$$\psi_s = \beta_0 I_s + \boldsymbol{\beta}' \mathbf{x}_s, \quad (2)$$

where  $\mathbf{x}_s$  is a column vector containing the values of the design factors of segment  $s$ , and  $\boldsymbol{\beta}$  is a conformable parameter vector. More importantly,  $I_s$  is an indicator for the non-zero constant effect at segment  $s$ ,  $\beta_0$ . In this way, we assume that given an animal road crossing, the probability of that crossing leading to an AVC is segment specific and depend on the design factors of each segment. And the constant effect of each segment is either zero or a fixed quantity of  $\beta_0$ . If  $\beta_0$  is significantly different from zero, it would suggest that there are segments in the network having their collision probability being affected by some important but ignored factors.

The segment specific indicators  $I_s$  for the constant effect can be thought as segment specific random parameter in a hierarchical model. The only difference is that in our model, due to the sparse information on a highly disaggregated level, we assume the non-zero impact of any design factor is the same for all the segments having the corresponding indicator taking value one. Thus the heterogeneity is only restricted to differ from zero to one non-zero value. This allow us to obtain a more reliable estimates for the non-zero effect based on a larger number of observations from a group of segments.

Given the collision probability being defined as in (1) and (2), the total number of AVCs occurred at segment  $s$  is then dependent on the number of animal road crossings of that segment,  $n_{s,t}$ . Moving to the specification for  $n_{s,t}$ , we can now envision that the total number of animal road crossing for each segment may firstly differ among segments and secondly be depend on animal activity level around a year with the potential seasonality shown in Figure 1. To achieve this while at the same time being able to recognize the fact of having very low AVC counts for a large number of segments, a Dirichlet process prior is used for the number of animal road crossing at each segment ( $n_{s,t}$ ) and at each month ( $t$ ) of a year.

In this way, the information about  $n_{s,t}$  can be shared among segments and months in a same cluster induced by the Dirichlet process. At the same time, a large number  $n_{s,t}$  can be assigned to a cluster representing zero animal crossings. So, the Dirichlet process can act as a non-parametric counterpart to zero-inflated count models. In this case, those segments and months with  $n_{s,t}$  being zero, would have the binomial probability of AVC to be equal to one, thus having no influence on the product probability for the observed sequence of AVCs and does not contribute to the estimation of parameters in (2).

In summary, the proposed modelling framework for AVCs can be written as follow:

$$\begin{aligned} P &\sim \mathbf{DP}(\alpha P_0), & q &\sim \mathbf{Beta}(a_0, b_0), \\ (\mu_{s,t}, \sigma_{s,t}) &\sim P, s = 1, \dots, S, t = 1, \dots, T, & I_s &\sim \mathbf{Bernoulli}(q), s = 1, \dots, S, \\ n_{s,t}^* &\sim \mathbf{N}_{(-0.5, \infty)}(\mu_{s,t}, \sigma_{s,t}^2), & [\beta_0, \boldsymbol{\beta}]' &\sim \mathbf{MVN}(\mathbf{0}, \boldsymbol{\Sigma}_0), \\ n_{s,t} &= \lfloor n_{s,t}^* \rfloor, & p_s &= 1/(1 + \exp(-\beta_0 I_s - \boldsymbol{\beta}' \mathbf{x}_s)), \\ & & k_{s,t} &\sim \mathbf{Binomial}(n_{s,t}, p_s). \end{aligned}$$

In which the total number of segments and the total number of months are denoted as  $S$  and  $T$ , respectively. The

number of animal road crossing,  $n_{s,t}$  and the collision probability,  $p_s$  at segment  $s$  and month  $t$  arise from the top-left and the top-right block, respectively. And the number of the observed AVCs at site  $s$ ,  $k_{s,t}$  is described by the binomial distribution with the parameters  $n_{s,t}$  and  $p_s$  at the bottom.

More specifically, in the top-left block, a discrete distribution  $P$  is drawn from the Dirichlet process with prior precision parameter  $\alpha$  and base distribution  $P_0$ . Then the cluster locations  $\mu_{s,t}$  and scales  $\sigma_{s,t}$  are generated from the discrete distribution  $P$  for each segment  $s$  and month  $t$ . Conditional on the cluster locations and scales, a real valued latent quantity  $n_{s,t}$  is drawn. Then, the total number of animal crossings ( $n_{s,t}$ ) at site  $s$  and  $t$  month is set to equal to the nearest integer,  $\lfloor n_{s,t}^* \rfloor$ .

For collision probabilities ( $p_s$ ) in the top-right block, the indicator probability,  $q$  is first drawn from a beta distribution with prior parameters  $a_0$  and  $b_0$ . Then the indicator variable for all  $S$  segments ( $I_s$ ) can be generated from a Bernoulli distribution with the just obtained indicator probability. Whereas the non-zero constant effect and the effect of segment specific design factors,  $[\beta_0, \boldsymbol{\beta}']'$  are drawn from a multivariate normal distribution with prior mean equals zero ( $\mathbf{0}$ ) and prior covariance ( $\boldsymbol{\Sigma}_0$ ) being chosen to be diffused and having only non-zero entries on the diagonal. Then the crash probabilities,  $p_s$  is set to be the linear combination of  $[I_s, \boldsymbol{x}'_s]'$  with parameters  $[\beta_0, \boldsymbol{\beta}']'$ .

Taking the two blocks together, the just determined collision probability,  $p_s$  and the number of animal crossing  $n_{s,t}$  are used as parameters for the binomial density function for determining the likelihood of the observed sequence of AVCs at all segments.

### 3. Posterior sampling

The unknown quantities in the proposed modelling framework can be divided into two parts: The first part contains the number of animal road crossings,  $n_{s,t}$  at each segments and months, and those used for describing the distributions of  $n_{s,t}$ . The second part includes the collision probability  $p_s$  of each animal road crossing on all segment and the other parameters involved in describing  $p_s$ . Given a set of starting values for all the unknown parameters, the posterior sampling can proceed in turn for the two parts. In the sampling of all the parameters, closed form posterior conditional distributions are available except in one step for the sampling of  $n_{s,t}$ . In this single step however, since the actual AVCs for each  $s$  and  $t$  is small or largely zero (Figure 3), the true value of which is expected to be also very small or equal to zero. As a results, a Metropolis step can be used without significantly compromising the speed of convergence and computational efficiency.

#### 3.1. Posterior sampling of $n_{s,t}$ and the related parameters

Conditioning on the starting value of  $n_{s,t}$  for all  $s$  and  $t$ , a blocked Gibbs sampler for the Dirichlet process (Ishwaran and James 2001) can be used. In this context, the kernel function used for representing clusters is a truncated normal density function with the truncation made at  $-0.5$  from below. For simplicity, the prior concentration parameter  $\alpha$  for the Dirichlet process is set to equal one, and the base distribution,

$$P_0(\mu, \sigma^2) = \mathbf{N}(\mu|\mu_0, \sigma^2) \mathbf{Gamma}(1/\sigma^2 | a_0, b_0),$$

which is conjugate to the form the kernel function. The parameters are chosen to be weakly informative ( $\mu_0 = 0$ ,  $a_0 = 2$  and  $b_0 = 100$ ). Due to the discrete nature and the limited number of distinct values of  $n_{s,t}$ , the maximum number of distinct clusters ( $C$ ) used in the blocked Gibbs sampling procedure is set at ten.

Using the above prior specification, the Gibbs sampler proceed in the following steps:

1. For each  $s = 1, \dots, S$  and  $t = 1, \dots, T$ , update  $\mu_{s,t}$  and  $\sigma_{s,t}$  by sampling from a multinomial distribution with

$$p(\mu_{s,t} = \mu_c^* \text{ and } \sigma_{s,t} = \sigma_c^* | \cdot) = \frac{w_c p(n_{s,t} | \mu_c^*, \sigma_c^*)}{\sum_{l=1}^C w_l p(n_{s,t} | \mu_l^*, \sigma_l^*)}.$$

In which  $w_l$  is the weight for the  $l$ th cluster. And the kernel function for each cluster,

$$p(n_{s,t} | \mu_l^*, \sigma_l^*) = \frac{\Phi(n_{s,t} + 1/2 | \mu_l^*, \sigma_l^{*2}) - \Phi(n_{s,t} - 1/2 | \mu_l^*, \sigma_l^{*2})}{1 - \Phi(-1/2 | \mu_l^*, \sigma_l^{*2})},$$

with

$$\Phi(a | \mu, \sigma) = \int_{-\infty}^a \mathbf{N}(z | \mu, \sigma^2) dz,$$

and  $\mu_l^*$  and  $\sigma_l^*$  are the  $l$ th pair of distinct parameters among the total number of  $C$  distinct pairs. For notation simplicity,  $P(A | \cdot)$  is used to denote the probability of condition  $A$  conditioning on all the rest of the parameters and the data.

2. Sample the cluster weights  $w_l$  for  $l = 1, \dots, C$  by first generate,



$$V_l \sim \mathbf{Beta}\left(1 + n_l, \alpha + \sum_{i=l+1}^C n_i\right), \text{ for } l = 1, \dots, C-1,$$

and set  $V_C = 1$ . In which,  $n_i$  is the number of  $\mu_{s,t}$  that is equal to  $\mu_i^*$ . Then set  $w_l = V_l \prod_{i<l} (1 - V_i)$  for  $l = 1, \dots, C$ .

- For each  $s = 1, \dots, S$  and  $t = 1, \dots, T$ , draw  $n_{s,t}^*$  by first sample

$$u_{st} \sim \mathbf{Uniform}\left(\Phi\left(n_{s,t} - \frac{1}{2} \mid \mu_{s,t}, \sigma_{s,t}^2\right), \Phi\left(n_{s,t} + \frac{1}{2} \mid \mu_{s,t}, \sigma_{s,t}^2\right)\right),$$

then set  $n_{s,t}^* = \Phi^{-1}(u_{st} \mid \mu_{s,t}, \sigma_{s,t}^2)$ .

- Update the parameters for defining the unique clusters using their conditional distributions:

$$1/\sigma_l^{*2} \sim \mathbf{Gamma}\left(a_0 + \frac{n_l}{2}, b_0 + \frac{1}{2} \sum_{\{(s,t): \mu_{s,t} = \mu_l\}} (n_{s,t}^* - \bar{n}_l^*) + \frac{n_l}{1 + n_l} \bar{n}_l^{*2}\right),$$

and

$$\mu_l^* \sim I_{\left[-\frac{1}{2}, \infty\right)} \mathbf{N}\left(\frac{\sum_{\{(s,t): \mu_{s,t} = \mu_l\}} n_{s,t}^*}{1 + n_l}, \frac{\sigma_l^{*2}}{1 + n_l}\right),$$

for  $l = 1, \dots, C$ . In which,  $\bar{n}_l^* = \sum_{\{(s,t): \mu_{s,t} = \mu_l\}} \frac{n_{s,t}^*}{n_l}$ .

An iteration of the above steps updates all the parameters involved for describing the distribution of  $n_{s,t}$ . Conditional on these parameters and the parameters used in describing collision probabilities  $p_s$ , for any non-negative integral value of  $n_{s,t}$ ,

$$P(n_{s,t} \mid \cdot) = \left[ \sum_{l=1, \dots, C} w_l \frac{\Phi\left(n_{s,t} + \frac{1}{2} \mid \mu_l, \sigma_l^2\right) - \Phi\left(n_{s,t} - \frac{1}{2} \mid \mu_l, \sigma_l^2\right)}{1 - \Phi\left(-\frac{1}{2} \mid \mu_l, \sigma_l^2\right)} \right] \prod_{t=1, \dots, T} \mathbf{Binomial}(k_{s,t} \mid n_{s,t}, p_s).$$

Utilizing the above expression, a Metropolis step can be used to sample  $n_{s,t}$  for all  $s$  and  $t$ .

### 3.2. Posterior sampling of $p_s$ and the related parameters

Conditioning on the values of  $n_{s,t}$  for all  $s$  and  $t$ , the posterior sampling for all the parameters describing  $p_s$  can be divided into two blocks: The first block is concerned with the parameters for defining  $p_s$ , which are the regression parameters  $\beta_0$  and  $\beta$ . And the second block is related to the indicators ( $I_s$ ) for all segments and the parameter describing the distribution of the indicators.

In the first block, conditioning on both  $n_{s,t}$  and  $I_s$  for all  $s$  and  $t$ , the posterior sampling of  $\beta_0$  and  $\beta$  is a standard linear regression problem with the observations following a binomial distribution. Consequently, the data augmentation approach of Polson, Scott, and Windle (2013) can be used. The detailed steps are as follow:

- Conditional on the current values of  $\beta_0$  and  $\beta$ , generate latent variable

$$\omega_{s,t} \sim \mathbf{PG}(n_{s,t}, \beta_0 I_s + \beta' \mathbf{x}_s),$$

for all  $s$  and  $t$ . In which  $\mathbf{PG}(b; z)$  is a Pólya-Gamma distribution with shape parameter  $b$  and tilting parameter  $z$ .

- Conditional on all the just generated  $\omega_{s,t}$ , draw

$$[\beta_0, \beta']' \sim \mathbf{MVN}(\mathbf{m}_\omega, \mathbf{V}_\omega),$$

where

$$\mathbf{V}_\omega = (\mathbf{X}\Omega\mathbf{X}' + \mathbf{B}_0^{-1})^{-1},$$

$$\mathbf{m}_\omega = \mathbf{V}_\omega (\mathbf{X}\boldsymbol{\kappa} + \mathbf{B}_0^{-1}\mathbf{m}_0)$$

with  $\mathbf{X} = [\mathbf{X}_1, \dots, \mathbf{X}_T]'$ ,  $\mathbf{X}_t = [[I_1, \mathbf{x}'_1]', \dots, [I_T, \mathbf{x}'_T]']$  for  $t = 1, \dots, T$ ,  $\boldsymbol{\kappa} = \left[k_{1,1} - \frac{n_{1,1}}{2}, \dots, k_{S,T} - \frac{n_{S,T}}{2}\right]'$ , and  $\mathbf{m}_0$  and  $\mathbf{B}_0$  are the prior mean (equals zero) and covariance matrix (diagonal matrix with large diagonal elements), respectively for the regression parameters,  $[\beta_0, \beta']'$ .

In the second block, conditioning on the just sampled  $\omega_{s,t}$  and  $[\beta_0, \beta']'$ , the indicators  $I_s$  can be drawn from their posterior distribution,

$$P(I_s = 1 \mid \cdot) = \frac{P_s^1 q}{P_s^0 q + P_s^1 (1 - q)},$$

with

$$P_s^0 = \prod_{t=1}^T \mathbf{N} \left( \frac{\kappa_{s,t}}{\omega_{s,t}} \middle| \boldsymbol{\beta}' \mathbf{x}_s, \frac{1}{\omega_{s,t}} \right),$$

$$P_s^1 = \prod_{t=1}^T \mathbf{N} \left( \frac{\kappa_{s,t}}{\omega_{s,t}} \middle| \beta_0 I_s + \boldsymbol{\beta}' \mathbf{x}_s, \frac{1}{\omega_{s,t}} \right).$$

Lastly, a complete iteration of the updates for all the parameters in the model is completed by sampling the parameters describing  $q$  from

$$\mathbf{Beta} \left( 1 + \sum_{s=1}^S I_s, 1 + \sum_{s=1}^S (1 - I_s) \right).$$

In which, the prior distribution for the indicators is assumed to be  $\mathbf{Beta}(1,1)$ .

#### 4. Results

The proposed model detailed in Section 3 is fitted using the AVCs data introduced in Section 2. A list of the considered segments specific design factors are shown in Table 1. Many aspects of the fitted model can be used for inference purposes. Firstly, the posterior means of the parameters for the number of animal road crossings are aggregated by months and shown in Figure 5. The relative magnitude of the aggregated quantities resembles a seasonal pattern similar to that shown in Figure 1 for the total number of AVCs by months. In other words, a seasonal pattern is captured in the model by the parameters for the number of animal road crossings ( $n_{s,t}$ ). Note however, the absolute magnitude shown in Figure 5 is generally smaller than that in Figure 1. This is because the total for each month are aggregated for all segments only in Figure 5 (one  $n_{s,t}$  for each segment and each month), whereas in Figure 1, the AVCs of each months are aggregated for all segments and over seven years.

Table 1: Descriptive information for the segment specific design factors considered in the model.

Name	Description	Range	Min.	Mean	Max.
SegLength	Length of the segments in miles	$\mathbb{R}^+$	0.001	0.66	30.11
ADT	Averaged daily traffic volume	$\mathbb{R}^+$	0.01	11.20	341.30
NumLanes	Number of the lanes	$\mathbb{N}^+$	1	2.77	14
RBWid	Roadbed width in feet	$\mathbb{R}^+$	10.00	47.11	318.00
MedWid	Median width in feet	$\mathbb{R}^+$	0.00	7.45	710.00
SouWidI	Inside shoulder width in feet	$\mathbb{R}^+$	0.00	5.19	60.00
SouWidO	Outside shoulder width in feet	$\mathbb{R}^+$	0.00	6.53	53.00
SurWid	Surface width in feet	$\mathbb{R}^+$	10.00	33.82	236.00
PeakFac	Percentage of peak period	(0, 100)	4.20	10.75	19.90
HWYDes	Controlled access (1 for yes)	{0, 1}	0	0.18	1
SPDMax	Speed limit in miles per hour	$\mathbb{R}^+$	5.00	56.15	85.00
RU	Urbanised area (1 for yes)	{0, 1}	0	0.29	1
MedType	Positive barrier as median (1 for yes)	{0, 1}	0	0.04	1

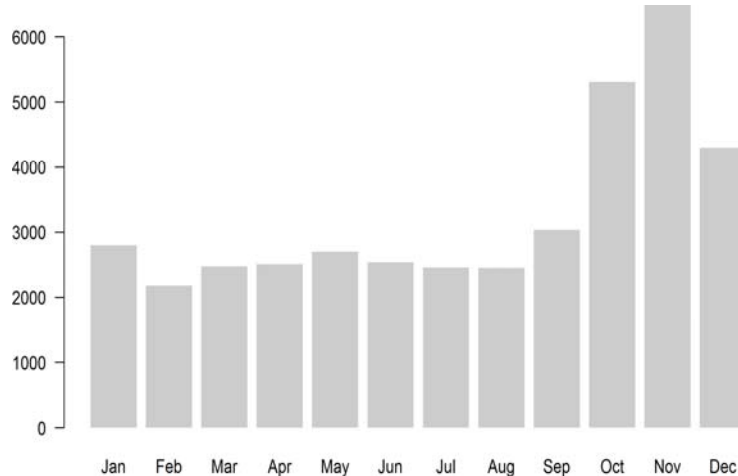




Figure 5: The sum of posterior means of the number of animal road crossings,  $n_{s,t}$  for each segment  $s$  and month  $t$  by months.

In terms of the number of non-zero  $n_{s,t}$ , the percentage shown in the second row of Table 2 suggests that on average the number of non-zero  $n_{s,t}$  is around 17%. Clearly, this is a larger percentage than that indicated in Figure 3, and should not come as a surprise. To explain, the quantity displayed in Table 2 is the estimated total number of animal crossings for the months of a year averaged over the entire seven years, whereas in Figure 3 the information shown is the number of actual AVCs observed for the months of each year in the seven year period. Despite this considerably higher percentage of the non-zero  $n_{s,t}$  in comparing to that of the observed AVCs, there are still over 80% of  $n_{s,t}$  taking value zeros. The over 80% of  $n_{s,t}$  that are equal to zero serve a role that is similar to that of the zero-class in zero-inflated regression models. By having this in the current Bayesian framework, it not only allows the parameters to be not biased downwards due to a large number of observations with zero AVC counts, but also accelerate the computation dramatically due to the fact that the sampling of  $\beta_0$  and  $\beta$  is now only dependent on the observations with non-zero  $n_{s,t}$ .

Moving to the part of posterior samples of parameters for determining the probabilities of AVCs,  $p_s$ . Table 2 shows the medians and 90% credible intervals for  $I_s$ ,  $\beta_0$  and  $\beta$ . Firstly, it can be seen that among all the indicators for a constant effect on the segments collision probability, around 43% of the segments that have non-zero  $n_{s,t}$  carry an inherent (unexplained by the included design factors) non-zero constant effect on the probability of having an AVC given an animal road crossing. This non-zero constant effect is positive and quite large (around 0.53) in comparing with the effect of other included design factors. In other words, there are around 7% ( $17\% \times 0.43$ ) segments in the networks are exposed to some unknown and increased risk of having AVCs given the same number of animal road crossings. Consequently, a practical and operationally meaningful information can be extracted by locating the segments that corresponds to high values of the posterior mean of non-zero constant effect, thus identifying the more problematic segments in the network.

In addition to the non-zero constant effect shown in Table 2, the estimated effects for a number of design factors can also be illuminating. Namely, speed limit (“SPDMax”) is positively associated with higher probability of causing an AVC by an animal road crossing. Segments in urbanized areas (“RU”) tend to have a lower probabilities of causing an AVC by an animal road crossings. Positive barriers (“MedType”) tend to decrease the probability. And the busier the segments (“PeakFac”) the lower the probabilities. Also, the estimated effect of “RBWid”, “MedWid” and “HWYDes” also accord with the expectation. Interestingly, the width of inside shoulder “SouWidI” is shown to be negative associated with the probability of causing an AVC by an animal road crossing. This may suggests a driving behavioral difference between driving on roads with and without inside shoulder.

Table 2: Descriptive information for the posterior samples of the regression parameters  $n_{s,t,q}$ ,  $\beta_0$  and  $\beta$ .

	Median	90% interval
Percentage of segments with non-zero $n_{s,t}$ (%)	17.1946	(17.1713,17.2168)
Probability of non-zero constant effect	0.4364	(0.3943,0.4794)
Constant	0.5371	(0.5243,0.5501)
SegLength	0.1446	(-0.2699,0.5692)
ADT	-0.0001	(-0.0006,0.0004)
NumLanes	0.0036	(-0.0176,0.0244)
RBWid	0.0019	(0.0004,0.0034)
MedWid	-0.0007	(-0.0011,-0.0002)
SouWidI	0.0049	(0.002,0.0078)
SouWidO	0.0004	(-0.0025,0.0032)
SurWid	0.0017	(-0.0003,0.0037)
PeakFac	-0.0079	(-0.0112,-0.0045)
HWYDes	0.0399	(0.0007,0.0814)
SPDMax	0.0028	(0.0021,0.0035)
RU	-0.2462	(-0.2774,-0.2151)
MedType	-0.0622	(-0.1053,-0.0179)

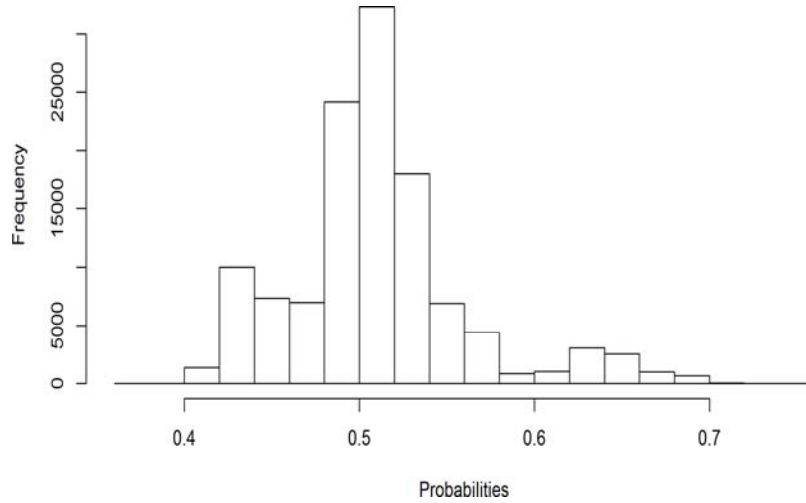


Figure 6: The empirical distribution of the collision probability,  $p_s$  evaluated based on a randomly chosen draw from the posterior distribution.

In terms of the insignificant parameters, the one reported in Table 2 for segment length (“SegLength”) may need further clarification. In a typical count regression model in terms of the mean of counts, it is within the expectation that the longer a segment the higher the AVC counts. In our case however, the effect indicated in Table 2 is not on mean of AVC but the probability of causing a single AVC given an animal crossing is occurred. In other words, only those design factors at the point of the road crossings instead of the total length of the segment should be expected to have a significant effect. Whereas for the potential effect of segment length, it is contained in the segment specific parameter,  $n_{s,t}$  for the number of animal road crossings.

Using the posterior samples for  $I_s$ ,  $\beta_0$  and  $\beta$ , the probabilities of causing an AVC by an animal road crossing on each segments can be evaluated according to (1) and (2). Figure 6 shows the probabilities evaluated based on a randomly chosen posterior sample. It can be seen that there are a large number of segments having the probability close to 0.5, meaning that the causing a AVC by an animal road crossing is close to random guessing. However, there are also a small group of segments with the probabilities being markedly higher at nearly 0.7, and another group having a lower probability at around 0.4. These differences may suggest that in general difference in road design factors do have influence on the collision probability and thus the number of AVCs.

In addition, the absolute quantities shown in Figure 6 may come as counter intuitive. In expectation, the occurrence of AVC should be rare, meaning that the probability should be much smaller. In order to make sense of the probabilities shown in Figure 6, it is important to recognize the difference between the information contained in the AVC dataset and that used in forming our expectation. To explain, the animal crossing inferred from the model are in fact those crossings that occurred when there are vehicles driving pass. The other crossings happened without having any vehicle on the segments would apparently not lead to AVCs thus not being recorded in the dataset. In this sense, the probability show in Figure 6 is in fact the probability of having an AVC giving both the presence of an animal road crossing and vehicles driving pass.

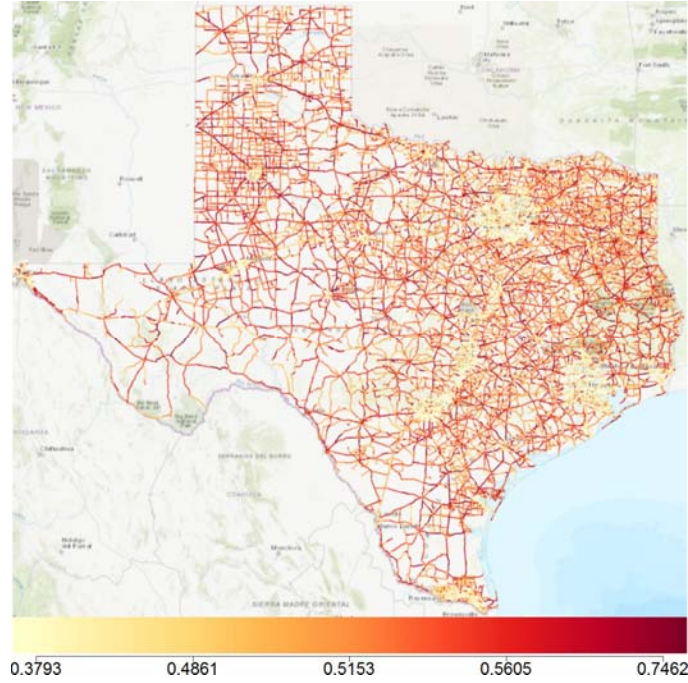


Figure 7: Posterior mean of collision probability  $p_s$  for all the segments in the state-controlled roadway network in Texas, U.S.

To explore the calculated collision probabilities ( $p_s$ ) in more details, Figure 7 shows them mapped onto the network. It can be seen clearly that there are several clusters of light colored segments, which correspond to the network around major urban areas. In contrast, the deep colored segments are mainly those major throughways spanning over the whole state. This pattern is a manifestation of the sampled parameters shown in Table 2. In particular, the negative effects of “HWYDes”, “RU” and the positive effect of “SPDMax”.

So far, the probabilities shown represent the likelihood of causing an AVC by an animal road crossing. This information can be informative in terms of analyzing the contributing effect of road design factors on the occurrence of AVC. However, one piece of information missing is about the number of animal road crossings for the segments. A segment with a high probability but zero animal crossing would apparently not be of a concern in comparing with one with both a high probability and a large number of animal road crossings. From the perspective of the road network operators, these segment having both a high probability and a large number of animal road crossings may be regards as the hotspots for AVC prevention purposes. Consequently, the identification of these hotspots could guide the resource to be used at the most needed place and thus archive a more effective measure for AVC prevention. For this reason, the posterior mean of the expected AVCs ( $n_{s,t}p_s$ ) may become useful and is shown in Figure 8 for two different months of a year.

By using  $n_{s,t}$  in the calculation for the expected AVCs, it can be seen in Figure 8 that a seasonality is induced. More specifically, the expected AVCs is higher (darker color) for more segments in November (the right panel) than that in February (the left panel). This corresponds well with the pattern shown in Figure 1 and allows the hotspot identification to be seasonal dependent. Another apparent feature shown in Figure 8 is that the segment with high mean AVCs (darker colored) are only a small percentage of the segments and scattered across the network. For this small percentage, it is of cause due to the fact of having over 80% of  $n_{s,t}$  being at zero (Table 2).

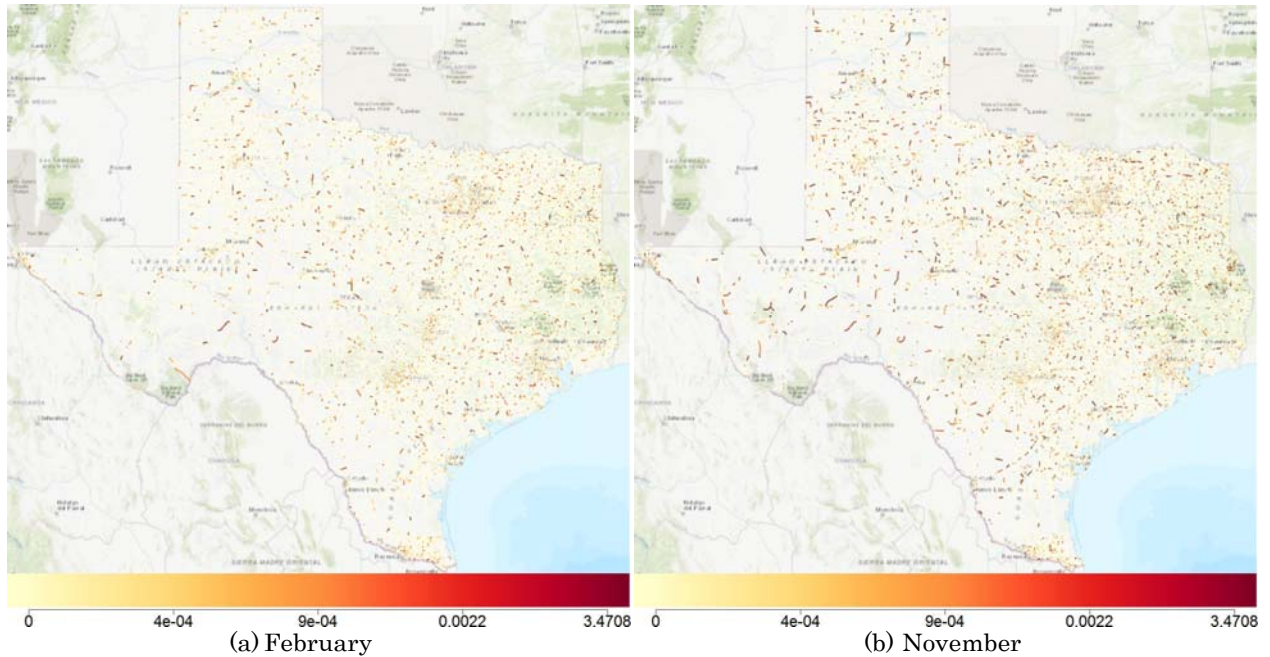


Figure 8: Posterior mean of expected AVCs for all segments in the state-controlled roadway network in Texas, U.S.

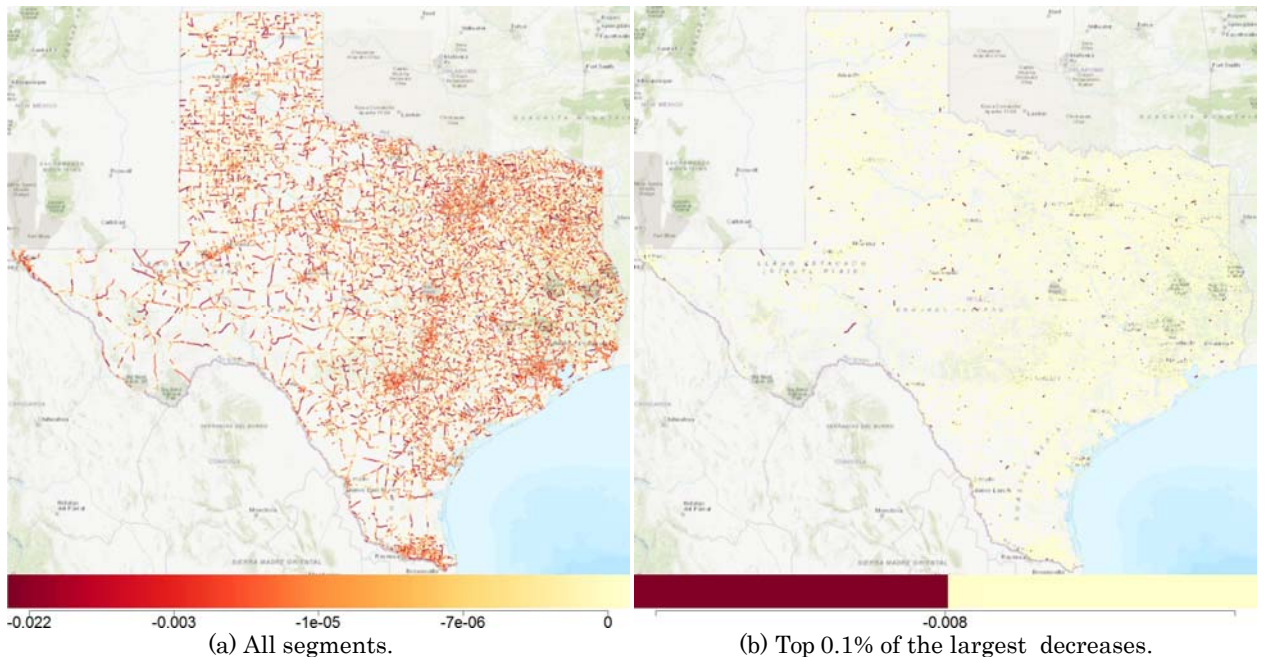


Figure 9: The changes in the posterior means of expected AVCs given a decrease of ten miles per hour in the speed limit for segments in the state-controlled roadway network in Texas.

As to the result of having segments with a high number of expected AVC scattered across the network, one may prefer a smooth changes from one segment to other nearby segments. However in doing so, we would then loose the advantage of having the segments with zero AVCs to play no role in the parameter estimation, and consequently causing a downward bias to the parameter estimate due to averaging the effect over a larger number of segments in the nearby area. Moreover, it would also contradict the goal of the present study on inferring the effect of segment specific design factors and shift the focus onto that at a higher level of spatial aggregation.

Given the focus remains on a segment level, another interesting aspect to explore from the perspective of the network operator, is the effect of speed limit on the occurrence of AVCs. For example, for the design factors included in the model (Table 2), speed limit can be regarded as one of the most cost effective and time efficient way to AVC



prevention. In this regard, a direct way of finding out the effectiveness of changing speed limit on AVC prevention would be calculating the marginal effect of speed limit. In this case, since there are around 43% of the segments with the constant effect being positive and quite large, we may focus on these at-risk segments and calculate the posterior mean of the changes in expected AVCs assuming a decreasing of speed limit by ten miles per hour. Figure 9 shows the obtained quantities. In the left panel, the quantities for all the sites are shown. It is interesting to see that although the calculated probability alone is quite small around urbanized areas (Figure 7), the decrease in expected AVCs is on the contrary relatively large.

If the segments shown in the left panel are too many for operational purposes, we can also shift the focus to a top quantile of the expected AVCs changes. In this case, the 0.1% top decreases in expected AVCs are shown in the right panel of Figure 9. In which, the shown segments would then correspond to those that are most relevant in terms of AVC prevention by means of controlling the speed limit.

## 5. Conclusion

In this study, we focus on modelling AVCs on highly disaggregated segments in a large network. Several advantages of the proposed modelling framework are demonstrated by using the AVC dataset collected on the state-control roadway network of Texas, U.S. with more than 100,000 segments. Firstly, the proposed model allows hotspot identification to be seasonal dependent. Secondly, the impact of segment specific attributes can be inferred directly on a highly disaggregated level. Third, straightforward scenario evaluations on the effect of any segment specific (e.g. speed limit) can be done using the proposed modelling framework.

By using the binomial distribution for the probability of AVC counts and allowing the parameter for the number of total trials in the binomial distribution to be governed by a Dirichlet process, the proposed model works in a way similar to zero-inflated regression models. By doing so, it made both the computation more efficient and the parameter estimate to be less biased due to averaging over a large number of segments with potentially zero animal road crossings. However, the drawback of this strategy is that the information on AVCs are only shared among segments in a same cluster as defined by the discrete probability distribution arose from the Dirichlet process, thus no smooth change on the spatial dimension is guaranteed. If one is instead interested in the effect of some attributes to be changing smoothly in the spatial dimension or on a higher level of spatial aggregation (such as, weather condition, green area ratio, counties), a distance based spatial correlation or attributes at a higher level of aggregation level may enter the model by specifying the mean of  $n_{s,t}$  as a linear combination of the attributes and means of adjacent segments.

In another direction for further development, the marginal effect shown in Figure 9 may be unsatisfactory. The magnitude of the effects of most design factors in Table 2 may be too small. In fact if we assume a decrease of speed limit by ten miles per hour for all the segments, a predicted decrease in the annual AVC occurrences would only be around 257. This change only accounts for around a four percent decrease in the total number of the average annual AVCs in the dataset. A main cause of this unsatisfactory phenomenon may come from ignoring the interaction/non-linear effect of the design factors. For this reason, exploring the interaction/non-linear effect could be an interesting direction for future development.

## References

- Al-Ghamdi, Ali S., and Saad A. AlGadhi. 2004. "Warning Signs as Countermeasures to Camel-Vehicle Collisions in Saudi Arabia." *Accident Analysis and Prevention*. <https://doi.org/10.1016/j.aap.2003.05.006>.
- Bíl, Michal, Richard Andrášik, Tomáš Svoboda, and Jiří Sedoník. 2016. "The KDE+ Software: A Tool for Effective Identification and Ranking of Animal-Vehicle Collision Hotspots along Networks." *Landscape Ecology*. <https://doi.org/10.1007/s10980-015-0265-6>.
- Brieger, Falko, Robert Hagen, Daniela Vetter, Carsten F. Dormann, and Ilse Storch. 2016. "Effectiveness of Light-Reflecting Devices: A Systematic Reanalysis of Animal-Vehicle Collision Data." *Accident Analysis and Prevention*. <https://doi.org/10.1016/j.aap.2016.08.030>.
- Bruinderink, G.W.T.A. Groot, and E. Hazebroek. 1996. "Ungulate Traffic Collisions in Europe." *Conservation Biology*. <https://doi.org/10.1046/j.1523-1739.1996.10041059.x>.
- Danks, Zachary D., and William F. Porter. 2010. "Temporal, Spatial, and Landscape Habitat Characteristics of Moose-Vehicle Collisions in Western Maine." *Journal of Wildlife Management*. <https://doi.org/10.2193/2008-358>.
- Diaz-Varela, Emilio R., Iban Vazquez-Gonzalez, Manuel F. Marey-Pérez, and Carlos J. Álvarez-López. 2011. "Assessing Methods of Mitigating Wildlife-Vehicle Collisions by Accident Characterization and Spatial Analysis." *Transportation Research Part D: Transport and Environment*. <https://doi.org/10.1016/j.trd.2011.01.002>.
- Found, Rob, and Mark S. Boyce. 2011. "Predicting Deer-Vehicle Collisions in an Urban Area." *Journal of Environmental Management*. <https://doi.org/10.1016/j.jenvman.2011.05.010>.
- Garrett, Larry C., and George A. Conway. 1999. "Characteristics of Moose-Vehicle Collisions in Anchorage, Alaska, 1991-1995." *Journal of Safety Research*. [https://doi.org/10.1016/S0022-4375\(99\)00017-1](https://doi.org/10.1016/S0022-4375(99)00017-1).
- Gkritza, Konstantina, Michael Baird, and Zachary N. Hans. 2010. "Deer-Vehicle Collisions, Deer Density, and Land Use in Iowa's Urban Deer Herd Management Zones." *Accident Analysis and Prevention*. <https://doi.org/10.1016/j.aap.2010.05.013>.

- Grilo, Clara, John A. Bissonette, and Margarida Santos-Reis. 2009. "Spatial-Temporal Patterns in Mediterranean Carnivore Road Casualties: Consequences for Mitigation." *Biological Conservation*. <https://doi.org/10.1016/j.biocon.2008.10.026>.
- Grosman, Paul D., Jochen A.G. Jaeger, Pascale M. Biron, Christian Dussault, and Jean Pierre Ouellet. 2009. "Reducing Moose-Vehicle Collisions through Salt Pool Removal and Displacement: An Agent-Based Modeling Approach." *Ecology and Society*. <https://doi.org/10.5751/ES-02941-140217>.
- Gunson, Kari E., Giorgos Mountrakis, and Lindi J. Quackenbush. 2011. "Spatial Wildlife-Vehicle Collision Models: A Review of Current Work and Its Application to Transportation Mitigation Projects." *Journal of Environmental Management*. <https://doi.org/10.1016/j.jenvman.2010.11.027>.
- Haikonen, Hannu, and Heikki Summala. 2001. "Deer-Vehicle Crashes: Extensive Peak at 1 Hour after Sunset." *American Journal of Preventive Medicine*. [https://doi.org/10.1016/S0749-3797\(01\)00352-X](https://doi.org/10.1016/S0749-3797(01)00352-X).
- Hothorn, Torsten, Jörg Müller, Leonhard Held, Lisa Möst, and Atle Mysterud. 2015. "Temporal Patterns of Deer-Vehicle Collisions Consistent with Deer Activity Pattern and Density Increase but Not General Accident Risk." *Accident Analysis and Prevention*. <https://doi.org/10.1016/j.aap.2015.04.037>.
- Huijser, Marcel P., John W. Duffield, Anthony P. Clevenger, Robert J. Ament, and Pat T. McGowen. 2009. "Cost-Benefit Analyses of Mitigation Measures Aimed at Reducing Collisions with Large Ungulates in the United States and Canada: A Decision Support Tool." *Ecology and Society*. <https://doi.org/10.1016/j.contraception.2009.11.002>.
- Hurley, Michael V., Eric K. Rapaport, and Chris J. Johnson. 2009. "Utility of Expert-Based Knowledge for Predicting Wildlife-Vehicle Collisions." *Journal of Wildlife Management*. <https://doi.org/10.2193/2008-136>.
- Ishwaran, Hemant, and Lancelot F. James. 2001. "Gibbs Sampling Methods for Stick-Breaking Priors." *Journal of the American Statistical Association*. <https://doi.org/10.1198/016214501750332758>.
- Jensen, Ryan R., Rusty A. Gonser, and Christian Joyner. 2014. "Landscape Factors That Contribute to Animal-Vehicle Collisions in Two Northern Utah Canyons." *Applied Geography*. <https://doi.org/10.1016/j.apgeog.2014.02.007>.
- Klöcker, Ulrike, David B. Croft, and Daniel Ramp. 2006. "Frequency and Causes of Kangaroo-Vehicle Collisions on an Australian Outback Highway." *Wildlife Research*. <https://doi.org/10.1071/WRO4066>.
- Kolowski, Joseph M., and Clayton K. Nielsen. 2008. "Using Penrose Distance to Identify Potential Risk of Wildlife-Vehicle Collisions." *Biological Conservation*. <https://doi.org/10.1016/j.biocon.2008.02.011>.
- Lao, Yunteng, Yao Jan Wu, Jonathan Corey, and Yin Hai Wang. 2011. "Modeling Animal-Vehicle Collisions Using Diagonal Inflated Bivariate Poisson Regression." *Accident Analysis and Prevention*. <https://doi.org/10.1016/j.aap.2010.08.013>.
- Lao, Yunteng, Guohui Zhang, Yao Jan Wu, and Yin Hai Wang. 2011. "Modeling Animal-Vehicle Collisions Considering Animal-Vehicle Interactions." *Accident Analysis and Prevention*. <https://doi.org/10.1016/j.aap.2011.05.017>.
- Leblond, Mathieu, Christian Dussault, Jean-Pierre Ouellet, Marius Poulin, Réhaume Courtois, and Jacques Fortin. 2007. "Electric Fencing as a Measure to Reduce Moose-Vehicle Collisions." *Journal of Wildlife Management*. <https://doi.org/10.2193/2006-375>.
- Litvaitis, John A., and Jeffrey P. Tash. 2008. "An Approach toward Understanding Wildlife-Vehicle Collisions." *Environmental Management*. <https://doi.org/10.1007/s00267-008-9108-4>.
- Malo, Juan E., Francisco Suárez, and Alberto Díez. 2004. "Can We Mitigate Animal-Vehicle Accidents Using Predictive Models?" *Journal of Applied Ecology*. <https://doi.org/10.1111/j.0021-8901.2004.00929.x>.
- McCollister, Matthew F., and Frank T. van Manen. 2010. "Effectiveness of Wildlife Underpasses and Fencing to Reduce Wildlife-Vehicle Collisions." *Journal of Wildlife Management*. <https://doi.org/10.2193/2009-535>.
- Meisingset, Erling L., Leif E. Loe, Øystein Brekkum, and Atle Mysterud. 2014. "Targeting Mitigation Efforts: The Role of Speed Limit and Road Edge Clearance for Deer-Vehicle Collisions." *Journal of Wildlife Management*. <https://doi.org/10.1002/jwmg.712>.
- Mountrakis, Giorgos, and Kari Gunson. 2009. "Multi-Scale Spatiotemporal Analyses of Moose-Vehicle Collisions: A Case Study in Northern Vermont." *International Journal of Geographical Information Science*. <https://doi.org/10.1080/13658810802406132>.
- Mrtka, Jiří, and Marie Borkovcová. 2013. "Estimated Mortality of Mammals and the Costs Associated with Animal-Vehicle Collisions on the Roads in the Czech Republic." *Transportation Research Part D: Transport and Environment*. <https://doi.org/10.1016/j.trd.2012.09.001>.
- Neumann, Wiebke, Göran Ericsson, Holger Dettki, Nils Bunnefeld, Nicholas S. Keuler, David P. Helmers, and Volker C. Radeloff. 2012. "Difference in Spatiotemporal Patterns of Wildlife Road-Crossings and Wildlife-Vehicle Collisions." *Biological Conservation*. <https://doi.org/10.1016/j.biocon.2011.10.011>.
- Niemi, Milla, Christer M. Rolandsen, Wiebke Neumann, Tuomas Kukko, Raisa Tiilikainen, Jyrki Pusenius, Erling J. Solberg, and Göran Ericsson. 2017. "Temporal Patterns of Moose-Vehicle Collisions with and without Personal Injuries." *Accident Analysis and Prevention*. <https://doi.org/10.1016/j.aap.2016.09.024>.
- Polson, Nicholas G., James G. Scott, and Jesse Windle. 2013. "Bayesian Inference for Logistic Models Using Pólya-Gamma Latent Variables." *Journal of the American Statistical Association*. <https://doi.org/10.1080/01621459.2013.829001>.
- Ramp, Daniel, Vanessa K. Wilson, and David B. Croft. 2006. "Assessing the Impacts of Roads in Peri-Urban Reserves: Road-Based Fatalities and Road Usage by Wildlife in the Royal National Park, New South Wales, Australia." *Biological Conservation*. <https://doi.org/10.1016/j.biocon.2005.11.002>.
- Rodríguez-Morales, Beatriz, Emilio Rafael Díaz-Varela, and Manuel Francisco Marey-Pérez. 2013. "Spatiotemporal Analysis of Vehicle Collisions Involving Wild Boar and Roe Deer in NW Spain." *Accident Analysis and Prevention*. <https://doi.org/10.1016/j.aap.2013.07.032>.
- Rowden, Peter, Dale Steinhart, and Mary Sheehan. 2008. "Road Crashes Involving Animals in Australia." *Accident Analysis and Prevention*. <https://doi.org/10.1016/j.aap.2008.08.002>.
- Seiler, Andreas. 2005. "Predicting Locations of Moose-Vehicle Collisions in Sweden." *Journal of Applied Ecology*. <https://doi.org/10.1111/j.1365-2664.2005.01013.x>.
- Snow, Nathan P., David M. Williams, and William F. Porter. 2014. "A Landscape-Based Approach for Delineating Hotspots of Wildlife-Vehicle Collisions." *Landscape Ecology*. <https://doi.org/10.1007/s10980-014-0018-y>.
- Sullivan, John M. 2011. "Trends and Characteristics of Animal-Vehicle Collisions in the United States." *Journal of Safety Research*. <https://doi.org/10.1016/j.jsr.2010.11.002>.
- Ujvari, M., Hans J. Baagøe, AB Aksel B Madsen, and HJ Baagøe. 1998. "Effectiveness of Wildlife Warning Reflectors in Reducing Deer-Vehicle Collisions: A Behavioral Study." *The Journal of Wildlife Management*. <https://doi.org/10.2307/3802562>.
- Zuberogoitia, Iñigo, Javier del Real, Juan José Torres, Luis Rodríguez, María Alonso, Vicente de Alba, Carmen Azahara, and Jabi Zabala. 2015. "Testing Pole Barriers as Feasible Mitigation Measure to Avoid Bird Vehicle Collisions (BVC)." *Ecological Engineering*. <https://doi.org/10.1016/j.ecoleng.2015.06.026>.