Modeling Recurrent Safety Critical Events among Commercial Truck Drivers: A Bayesian Hierarchical Jump Power Law Process

Miao Cai
Saint Louis University, Saint Louis, MO 63104

Amir Mehdizadeh
Auburn University, Auburn, AL 36849

Qiong Hu
Auburn University, Auburn, AL 36849

Mohammad Ali Alamdar Yazdi
Johns Hopkins University, Baltimore, MD 21202

Alexander Vinel
Auburn University, Auburn, AL 36849

Karen C. Davis
Miami University, Oxford, OH 45056

Hong Xian
Saint Louis University, Saint Louis, MO 63104

Fadel M. Megahed
Miami University, Oxford, OH 45056

Steven E. Rigdon
Saint Louis University, Saint Louis, MO 63104

July 4, 2020

Abstract

Many transportation safety studies aim to predict crashes based on aggregated road segment data. As an increasing number of naturalistic driving studies have been initiated in the recent decade, safety-critical events (SCEs) such as hard brakes are widely used as a proxy measure of driving risk. Different from real crashes, multiple SCEs can occur in a driving shift and they do not interrupt the state of driving. Motivated by a growing need of analyzing recurrent SCEs and the feature that multiple segments are nested within a shift for commercial truck drivers, we proposed a Bayesian hierarchical non-homogeneous Poisson process with power law process intensity function and an innovative Bayesian hierarchical jump power law process. We specified the parameterization, intensity functions, and likelihood functions for the two models and presented the estimation results for correctly and wrongly specified models based on simulated data. The two models are then applied to a naturalistic driving data of over 13 million driving records and 8,407 SCEs generated by 496 commercial truck drivers. Supplementary materials including simulated data and parameter estimation for reproducing the work, are available as an online supplement.

Keywords: trucking; safety-critical events; reliability; power law process

1. INTRODUCTION

Commercial truck drivers "form the lifeblood of [the U.S.] economy" (The White House, 2020), generating annual revenues exceeding $700 billion from the transportation of 10.8 billion tons of freight (John, 2019). The industry typically requires drivers to be on the road for an extended period of time, incentivizing drivers with hourly, per-mile or per-stop pay schedules. Furthermore, the industry is heavily regulated through the "hours of service" regulation (Federal Motor Carrier Safety Administration, 2020), which dictates the total number of driving hours permitted, minimum length of off-duty rest periods and allowable weekly total hours of driving/rest. Consequently, a major difference between commercial (large) truck drivers and commuters is the complex operational environment required of commercial drivers. Specifically, commercial drivers have to abide by government regulations while managing industry practices that attempt to optimize both productivity and safety.

The safety of truck drivers is of critical importance not only to trucking operators, but also to the general public. Truck crashes have a two-fold penalty (Tsai et al., 2018) (a) direct losses arising from injuries, fatalities and property damage affecting the truck driver and other commuters on the road, and (b) indirect losses in efficiency associated with slowing/damaging transferred goods and the impact to travel time for other commuters. Alarmingly, despite the regulatory oversights and continued advancements in safety technologies, the rates of truck-involved crashes in the U.S. have increased over the past decade. The involvement rate per 100 million large-truck miles traveled increased from 1.32 in 2008 to 1.48 in 2016 for fatal crashes, and from 21 in 2008 to 31 in 2015, most recent data, for injury crashes (NHTSA's National Center for Statistics and Analysis, 2019).

Traditional trucking safety studies utilize one or more *road segments* as the unit of analysis (Mehdizadeh et al., 2020) and attempt to model the occurrence or the number of crashes in a fixed time period (Lord and Mannering, 2010; Savolainen et al., 2011; Mannering and Bhat, 2014). Thus, the developed models use/resemble a case-control study design (Mehdizadeh et al., 2020). Limitations of those studies include a small number of observed crashes, difficulty in selecting control groups, and an undercount of less severe crashes (Mehdizadeh et al., 2020). More importantly, those studies cannot capture driver behavioral factors which contribute to 90% of traffic crashes (Federal Highway Administration, 2019).

To address the deficiencies in traditional safety studies, large-scale naturalistic driving studies (NDSs) have received significant attention in recent years (see e.g., Guo, 2019; Mehdizadeh et al., 2020; Cai et al., 2020). Those studies capitalize on advances in communication, computing and on-board vehicular sensing technologies which have allowed for the continuous recording of real-world driving data (e.g., timestamps capturing driving location, speed, rest brakes, etc.). In addition to their ability to capture continuous data on possible explanatory variables, NDSs allow for using more frequent near-crash safety critical events (SCEs) as proxies for crash data. It is well-established that increases in SCEs (e.g., hard braking events or the activation of forward-collision mitigation systems) are positively correlated with crash rates (Dingus et al., 2006; Guo et al., 2010; Gordon et al., 2011; Cai et al., 2020). Consequently, SCEs are the preferred choice for outcome variables in NDSs since they are more frequent (Cai et al., 2020) and hence, provide larger statistical power.

As with most point process data, SCEs can be divided into models attempting to (a) quantify the likelihood of observing one or more SCEs through binary classification (Ghasemzadeh and Ahmed, 2017, 2018) and count data models (Kim et al., 2013), respectively, and (b) estimate the time(s) of observing SCEs (Li et al., 2018; Liu and Guo, 2019; Liu et al., 2019; Guo et al., 2019). Miao's attempt at discussing the existing models in 2-3 sentences with a special emphasis on their limitations, i.e., why they cannot support the type of emphasis in Figure 2. These models are limited in three aspects. First, naturalistic driving studies follow the drivers or vehicles for a certain amount of time, which resembles a prospective cohort study (Mehdizadeh et al., 2020). However, many studies apply the methodologies of case-control studies to this type of data by including all events and match them with selective non-events (e.g., Das et al., 2019; Ghasemzadeh and Ahmed, 2018), which reduces the statistical power to detect potentially existing effects and fails to account for the fact that the driving data are nested within drivers. Second, SCEs are inherently different from crashes: there can be at most one crash in a continuous on-duty period and the drivers have to stop once a crash occurs; in contrast, there can be multiple SCEs in the same period and the drivers can continue driving even if SCEs occur. Classification and count data models are inefficient use of these data since the timing of events is not used and the recurrent-event nature of SCEs is ignored. Third, as required by Federal Motor Carrier Safety Administration (2017), commercial truck drivers, especially long-haul truck drivers, must take at least one break in long-distance transporting. Researcher would assume that some level of fatigue alleviation and reliability change occur at these breaks, but none of the previous studies to our notice had taken these breaks into account.

Steve's attempt at explaining the two main questions

In this article we address two types of questions regarding the safety behavior of commercial truck drivers. First, does the occurrence of SCEs tend to increase during a shift (a continuous period for which the driver is on duty, but not necessarily driving), due to fatigue or some other reason? If so, how is this effect manifested, and for which type of SCE does this occur? Second, what is the effect of rest breaks on safety behavior? To what extent does safety change after a rest break?

Fadel stopped here. In my opinion, in addition to addressing the issues above, the next 2-3 paragraphs in the introduction need to:

- briefly introduce the case study hypotheses/research questions that are being examined - in my opinion, the current next paragraph in the manuscript does not do this very clearly → see the paragraph by Steve in blue.

- Explain why PLP and JPLP can be used to evaluate the research hypotheses. Highlight the maintenance analogy the fact that we are borrowing models from Reliability (driving) and possibly maintenance (rest is a maintenance act where we assume the outcome to be not-as-good-as-new) Q: are we the first to this analogy? To my limited search, I think we are. A few papers do use recurrent event models, but I don't think they Second, regardless on whether we are the first, the human aspect requires on to account for heterogeneous units of analysis and hence the use of hierarchical models

- I also personally believe that Figure 2 should be moved to the introduction since it clearly communicates the hypotheses being examined. Figure 2 has been moved in the Intro and combined with the introduction of PLP and JPLP. Need Steve's help on introducing the NHPP and PLP for truck NDS data.

- Then add, with some minor cleaning up the final paragraph that discusses how the remainder of the paper is organized. I read through the last paragraph of Intro. It reads fine for me. Feel free to change it if you think it can be improved.

Fadel stopped here and I commented out two paragraphs below - one I incorporated some of it in the paragraph starting at line 60 i.e. “Existing approaches to SCE modeling” and the other primarily because it was not clear to me and it needs to be rewritten and added to the yellow highlights in the aforementioned paragraph.

To Steve: we need some justification for assuming a NHPP and PLP here. To answer these questions, we first introduced a Bayesian hierarchical non-homogeneous Poisson process with the power law process (PLP) intensity function to model SCEs within shifts. This model accounts for driver-level unobserved heterogeneity by specifying driver-level random intercepts for the rate parameter in PLP. On the other hand, to account for the feature that multiple breaks nested within a shift among commercial truck drivers, we then propose a Bayesian hierarchical jump power law process (JPLP) to take potential reliability changes at the time of rests into consideration. Figure 1 presents the intensity functions of PLP and JPLP. The intensity function of PLP is a smooth curve with concave-down trend when the shape parameter is greater than one (the first segment of the two curves overlaps), while the JPLP has a piecewise form. Whenever the driver takes a break, the intensity function of a JPLP jump by a certain amount κ.

The structure of this article is as follows. In Section 2, we define our terminology and notation for shifts, segments, and events for naturalistic driving data generated by commercial truck drivers. In Section 3, we specify our proposed PLP and JPLP models, their intensity functions and likelihood functions. In Section 4, simulation studies are conducted

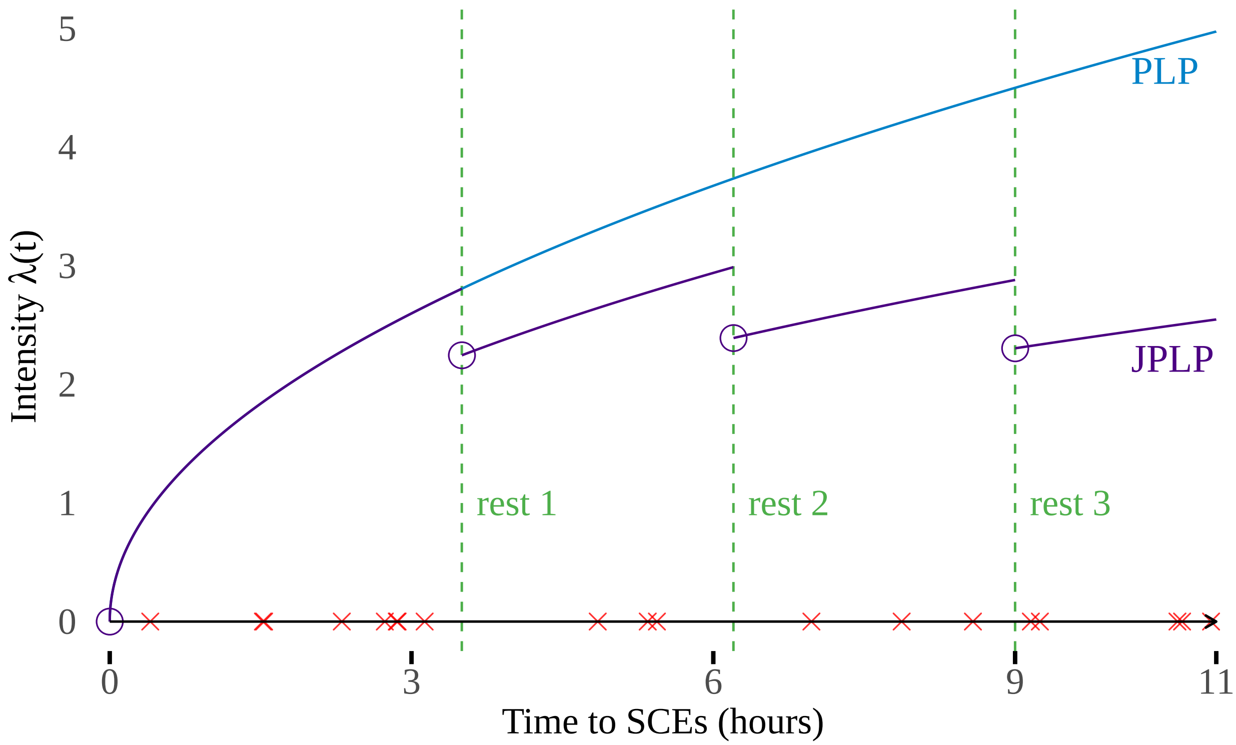

Figure 1: Simulated intensity function of PLP and JPLP. The x-axis shows time in hours since start and y-axis shows the intensity of SCEs. The crosses mark the time to SCEs and the vertical dotted lines indicates the time of the rests. Parameters for this simulated data set: shape parameters $\beta = 1.2$, rate parameter $\theta = 2$, jump parameter $\kappa = 0.8$.

to demonstrate the validity of our code and the consequences if the models are not specified correctly. In section 5, we present the results of real data analyses for 496 commercial truck drivers using PLP and JPLP. Strengths, possible limitations, and future research directions are discussed in Section 6. A simulated data set, description on data structure, and Stan and R code for Bayesian PLP and JPLP estimation are provided in the supplementary material.

2. TERMINOLOGY AND NOTATION

Figure 2 presents a time series plot of speed data for a sample truck driver (including two shifts and six segments nested within the shifts) and arrows suggesting shifts and segments. We use $d = 1, 2, \ldots, D$ as the notation for different drivers. A shift $s = 1, 2, \ldots, S_d$ is on-duty periods with no breaks longer than 10 hours for driver d. As the Federal Motor Carrier Safety Administration (2017) requires, a shift must be no more than 14 hours with no more than 11 hours of driving, and this leads to the phenomena that multiple segments $r = 1, 2, \ldots, R_{d,s}$ are separated by breaks longer than 30 minutes but less than 10 hours for each driver d and shift s.

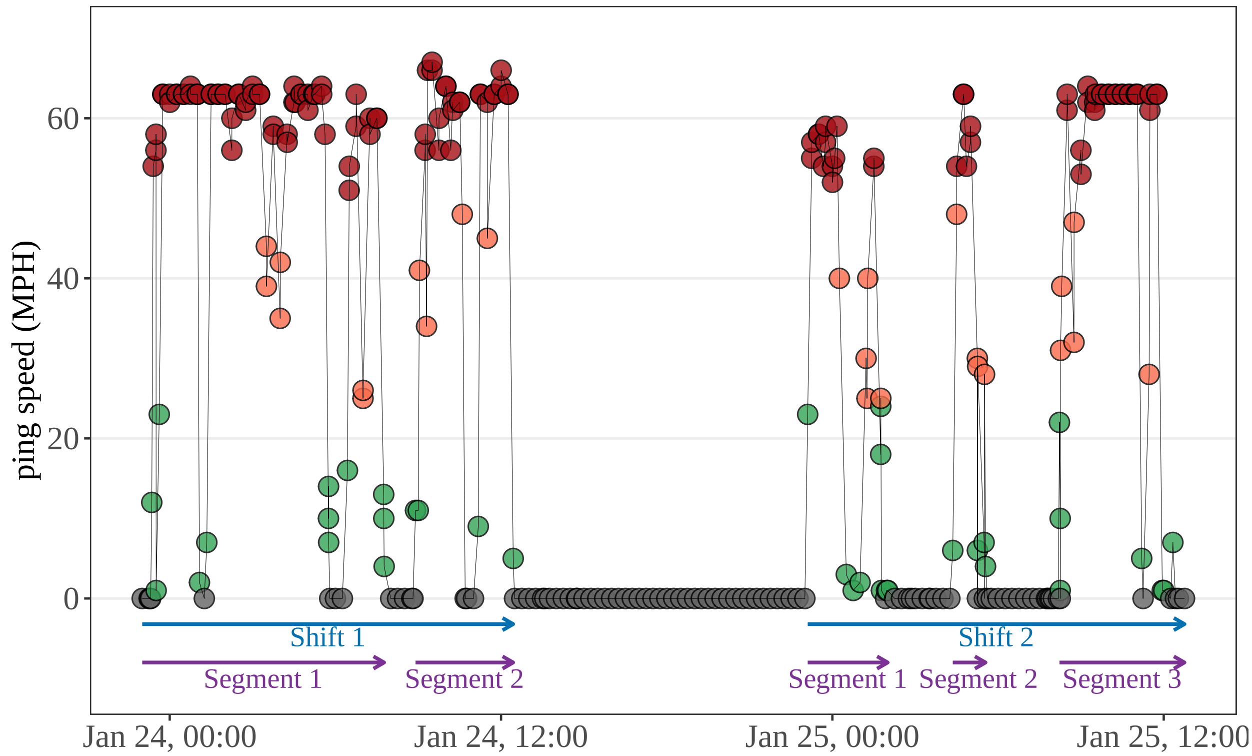

Figure 2: Naturalistic truck driving real-time ping data (points) and the aggregation process from pings to shifts and segments (arrows).

SCEs can occur any time in the segments whenever preset kinematic thresholds are trigger by the driver. We use $i = 1, 2, \ldots, I_{d,s}$ as notations for the i-th SCE for driver d in shift s. For each SCE, $t_{d,s,i}$ is the time to the i-th SCE for driver d measured from the beginning of the s-shift and the rest times between segments are excluded from calculation. $n_{d,s,r}$ is the number of SCEs for segment r within shift s for driver d. $a_{d,s,r}$ is the end time of segment r

within shift s for driver d.

3. MODELS

3.1 Non-homogeneous Poisson Process (NHPP) and Power Law Process

We assume the time to a SCE t follows a non-homogeneous Poisson process, whose intensity function $\lambda(t)$ is non-constant. The intensity function is assumed to have the following function form:

$$\lambda_{\mathrm{PLP}}(t) = \beta\theta^{-\beta}t^{\beta-1}, \tag{1}$$

where the shape parameter β indicates reliability improvement ($\beta < 1$), constant ($\beta = 1$), or deterioration ($\beta > 1$), and the scale parameter θ determines the rate of events. Here we assume the intensity function of a power law process because it has a flexible functional form, relatively simple statistical inference, and is a well-established model (Rigdon and Basu, 1989, 2000).

3.2 Bayesian Hierarchical Power Law Process (PLP)

The Bayesian hierarchical power law process is parameterized as:

$$
\begin{aligned}
t_{d,s,1}, t_{d,s,2}, \cdots, t_{d,s,n_{d,s}} &\sim \text{PLP}(\beta, \theta_{d,s}, \tau_{d,s}) \\
\beta &\sim \text{Gamma}(1,1) \\
\log \theta_{d,s} &= \gamma_{0d} + \gamma_1 x_{d,s,1} + \gamma_2 x_{d,s,2} + \cdots + \gamma_k x_{d,s,k} \\
\gamma_{01}, \gamma_{02}, \cdots, \gamma_{0D} &\sim \text{i.i.d. } N(\mu_0, \sigma_0^2) \\
\gamma_1, \gamma_2, \cdots, \gamma_k &\sim \text{i.i.d. } N(0, 10^2) \\
\mu_0 &\sim N(0, 5^2) \\
\sigma_0 &\sim \text{Gamma}(1,1),
\end{aligned}
\tag{2}
$$

where $t_{d,s,i}$ is the time to the i-th event for driver d in shift s, $\tau_{d,s} = a_{d,s,R_{d,s}}$ is the length of time of shift s (truncation time) for driver d, and $n_{d,s} = \sum_{r=1}^{n_{d,s}}$ is the number of SCEs in shift s for driver d. The likelihood function of event times generated from a PLP for driver d in shift s is given in Rigdon and Basu (2000, Section 2.3.2, Page 60):

$$
\begin{aligned}
L_{d,s}(\beta, \gamma_{0d}, \gamma | \mathbf{X}_d, \mathbf{W}_s) &= \Big(\prod_{i=1}^{n_{d,s}} \lambda_{\text{PLP}}(t_{d,s,i})\Big) \exp(-\int_0^{\tau_{d,s}} \lambda_{\text{PLP}}(u) du) \\
&= \begin{cases} \exp\Big(-(\tau_{d,s}/\theta_{d,s})^{\beta}\Big), & \text{if } n_{d,s} = 0, \\ \Big(\prod_{i=1}^{n_{d,s}} \beta \theta_{d,s}^{-\beta} t_{d,s,i}^{\beta-1}\Big) \exp\Big(-(\tau_{d,s}/\theta_{d,s})^{\beta}\Big), & \text{if } n_{d,s} > 0, \end{cases}
\end{aligned}
\tag{3}
$$

where $\mathbf{X}_d$ indicates driver specific variables (e.g. driver age and gender), $\mathbf{W}_s$ represents shift specific variables (e.g. precipitation and traffic), and $\theta_{d,s}$ is the function of parameters

$\gamma_{0d}, \gamma_1, \gamma_2, \ldots, \gamma_k$ and variables $x_{d,s,1}, x_{d,s,2}, \ldots, x_{d,s,k}$ given in the third line of Equation 2. The full likelihood function for all drivers are:

$$L = \prod_{d=1}^{D} \prod_{s=1}^{S_d} L_{d,s}(\beta, \gamma_{0d}, \gamma | \mathbf{X}_d, \mathbf{W}_s) \tag{4}$$

where $L_{d,s}(\beta, \gamma_{0d}, \gamma | \mathbf{X}_d, \mathbf{W}_s)$ is given in Equation 3.

3.3 Bayesian Hierarchical Jump Power Law Process (JPLP)

Since the Bayesian hierarchical PLP in Subsection 3.2 does not account for the rests ($r = 1, 2, \ldots, R_{d,s}$) within shifts and associated potential reliability improvement. In this subsection, we proposes a Bayesian hierarchical JPLP with an additional jump parameter κ.

Our proposed JPLP has the following piecewise intensity function:

$$\begin{aligned}
&\lambda_{\text{JPLP}}(t|d, s, r, \beta, \gamma_{0d}, \gamma, \mathbf{X}_d, \mathbf{W}_s) = \\
&\qquad \begin{cases}
\kappa^0 \lambda(t|\beta, \gamma_{0d}, \gamma, \mathbf{X}_d, \mathbf{W}_s), & 0 < t \leq a_{d,s,1}, \\
\kappa^1 \lambda(t|\beta, \gamma_{0d}, \gamma, \mathbf{X}_d, \mathbf{W}_s), & a_{d,s,1} < t \leq a_{d,s,2}, \\
\cdots & \cdots \\
\kappa^{R-1} \lambda(t|\beta, \gamma_{0d}, \gamma, \mathbf{X}_d, \mathbf{W}_s), & a_{d,s,R-1} < t \leq a_{d,s,R},
\end{cases} \\
&\qquad = \kappa^{r-1} \lambda(t|d, s, r, \kappa, \beta, \gamma_{0d}, \gamma, \mathbf{X}_d, \mathbf{W}_s),\ a_{d,s,r-1} < t \leq a_{d,s,r},
\end{aligned} \tag{5}$$

where the introduced parameter κ is the amount of intensity function change once the driver takes a break, and $a_{d,s,r}$ is the end time of segment r within shift s for driver d. By definition, the end time of the 0-th segment $a_{d,s,0} = 0$, and the end time of the last segment for the

d-driver within the s-th shift $a_{d,s,R_{d,s}}$ equals the shift end time $\tau_{d,s}$. We assume that this κ is constant across drivers and shifts.

The Bayesian hierarchical JPLP model is parameterized as:

$$
\begin{aligned}
t_{d,s,1}, t_{d,s,2}, \cdots, t_{d,s,n_{d,s}} &\sim \text{JPLP}(\beta, \theta_{d,s}, \tau_{d,s}, \kappa)\\
\beta &\sim \text{Gamma}(1, 1)\\
\log \theta_{d,s} &= \gamma_{0d} + \gamma_1 x_{d,s,1} + \gamma_2 x_{d,s,2} + \cdots + \gamma_k x_{d,s,k}\\
\kappa &\sim \text{Uniform}(0, 2)\\
\gamma_{01}, \gamma_{02}, \cdots, \gamma_{0D} &\sim \text{i.i.d. } N(\mu_0, \sigma_0^2)\\
\gamma_1, \gamma_2, \cdots, \gamma_k &\sim \text{i.i.d. } N(0, 10^2)\\
\mu_0 &\sim N(0, 5^2)\\
\sigma_0 &\sim \text{Gamma}(1, 1),
\end{aligned}
\tag{6}
$$

The notations are identical with those in Equation 2 except for the extra κ parameter. Here we set the prior distribution for κ as uniform(0, 2), which is assuming that the intensity function can jump down to 0 or increase by up to 100% at the time of rests. Similarly, the likelihood function of event times generated from a JPLP for driver d on shift s is:

$$
\begin{aligned}
L^*_{d,s}(\kappa, \beta, \gamma_{0d}, \gamma | \mathbf{X}_d, \mathbf{W}_s) &= \Big(\prod_{i=1}^{n_{d,s}} \lambda_{\text{JPLP}}(t_{d,s,i})\Big) \exp(-\int_0^{\tau_{d,s}} \lambda_{\text{JPLP}}(u)du)\\
&\begin{cases}
\exp\Big(-\int_0^{\tau_{d,s}} \lambda_{\text{JPLP}}(u)du\Big), & \text{if } n_{d,s} = 0,\\
\Big(\prod_{i=1}^{n_{d,s}} \lambda_{\text{JPLP}}(t_{d,s,i})\Big) \exp\Big(-\int_0^{\tau_{d,s}} \lambda_{\text{JPLP}}(u)du\Big), & \text{if } n_{d,s} > 0,
\end{cases}
\end{aligned}
\tag{7}
$$

where the piecewise intensity function $\lambda_{\text{JPLP}}(t_{d,s,i})$ is given in Equation 5.

However, since the intensity function depends on the segment r for the same driver d on shift s, it is hard to write out the specific form of Equation 7. Instead, we can rewrite the likelihood function at segment level, where the intensity function λ_{JPLP} is fixed for driver d on shift s and segment r:

$$L^*_{d,s,r}(\kappa, \beta, \gamma_{0d}, \gamma|\mathbf{X}_d, \mathbf{W}_r) = \begin{cases} \exp\Big(-\int_{a_{d,s,r-1}}^{a_{d,s,r}} \lambda_{\text{JPLP}}(u)du\Big), & \text{if } n_{d,s,r} = 0, \\ \Big(\prod_{i=1}^{n_{d,s,r}} \lambda_{\text{JPLP}}(t_{d,s,r,i})\Big) \exp\Big(-\int_{a_{d,s,r-1}}^{a_{d,s,r}} \lambda_{\text{JPLP}}(u)du\Big), & \text{if } n_{d,s,r} > 0, \end{cases} \tag{8}$$

where $t_{d,s,r,i}$ is the time to the i-th SCE for driver d on shift s and segment r measured from the beginning of the shift, $n_{d,s,r}$ is the number of SCEs for driver d on shift s and segment r. Compared to the PLP likelihood function given in Equation 4 where $\mathbf{W}_s$ are assumed to be fixed numbers during an entire shift, the rewritten likelihood function for JPLP in Equation 8 assumes external covariates $\mathbf{W}_r$ vary between different segments in a shift. In this way, JPLP can account for the variability between different segments within a shift.

Therefore, the overall likelihood function for drivers $d = 1, 2, \ldots, D$, their corresponding shifts $s = 1, 2, \ldots, S_d$, and segments $r = 1, 2, \ldots, R_{d,s}$ is:

$$L^* = \prod_{d=1}^{D} \prod_{s=1}^{S_d} \prod_{r=1}^{R_{d,s}} L^*_{d,s,r}, \tag{9}$$

where $L^*_{d,s,r}$ is a likelihood function given in Equation 8, in which the intensity function λ_{JPLP}

has a fixed functional form provided in the last line of Equation 5 for a certain driver d in a given shift s and segment r.

4. SIMULATION STUDY

4.1 Simulation setting

We conducted a simulation study to evaluate the performance of our proposed NHPP and JPLP under different simulation scenarios. We performed 1,000 simulations to each of the following three scenarios with different number of drivers $D = 10, 25, 50, 75, 100$:

1. Data generated from a PLP and estimated assuming a PLP (PLP),
2. Data generated from a JPLP and estimated assuming a JPLP (JPLP),
3. Data generated from a JPLP, but estimated assuming a PLP (PLP $\leftarrow$ JPLP).

The scenario "data generated from a PLP, but estimated assuming a JPLP" is not considered here since it is not theoretically possible: if the data is generated from a NHPP with PLP intensity function, then there are no breaks within shifts and it is pointless to estimate the data assuming a JPLP.

Specifically, for each driver, the number of shifts is simulated from a Poisson distribution with the mean parameter of 10. We assume there are three predictor variables x_1, x_2, x_3 for

θ ($k = 3$). x_1, x_2, x_3 and the shift time $\tau_{d,s}$ are generated from the following process:

$$\begin{aligned} x_1 &\sim \text{Normal}(1, 1^2) \\ x_2 &\sim \text{Gamma}(1, 1) \\ x_3 &\sim \text{Poisson}(2) \\ \tau_{d,s} &\sim \text{Normal}(10, 1.3^2) \end{aligned} \tag{10}$$

The parameters and hyperparameters are assigned the following values or generated from the following process:

$$\begin{aligned} \mu_0 = 0.2,&\ \sigma_0 = 0.5, \\ \gamma_{01}, \gamma_{02}, \cdots, \gamma_{0D} &\sim \text{i.i.d. } N(\mu_0, \sigma_0^2) \\ \gamma_1 = 1,&\ \gamma_2 = 0.3,\ \gamma_3 = 0.2 \\ \theta_{d,s} &= \exp(\gamma_{0d} + \gamma_1 x_1 + \gamma_2 x_2 + \gamma_3 x_3) \\ \beta = 1.2,&\ \kappa = 0.8. \end{aligned} \tag{11}$$

After the predictor variables, shift time, and parameters are generated, the time to events $t_{d,s,1}$, $t_{d,s,2}$, $\cdots$, $t_{d,s,n_{d,s}}$ and $t^*_{d,s,1}, t^*_{d,s,2}, \cdots, t^*_{d,s,n_{d,s}}$ are generated from PLP and JPLP:

$$\begin{aligned} t_{d,s,1}, t_{d,s,2}, \cdots, t_{d,s,n_{d,s}} &\sim \text{PLP}(\beta, \theta_{d,s}, \tau_{d,s}) \\ t^*_{d,s,1}, t^*_{d,s,2}, \cdots, t^*_{d,s,n_{d,s}} &\sim \text{JPLP}(\beta, \theta_{d,s}, \tau_{d,s}, \kappa) \end{aligned} \tag{12}$$

The parameters are then estimated using the likelihood functions given in Equation 4 and 9 using the probabilistic programming language `Stan` in R (Carpenter et al., 2017; Stan

Development Team, 2018), which uses efficient Hamiltonian Monte Carlo to sample from the posterior distributions. For each simulation, one chain is applied, with 2,000 warmup and 2,000 post-warmup iterations drawn from the posterior distributions.

4.2 Simulation results

The simulation results are shown in Table 1. For the five sets of drivers $D = 10, 25, 50, 75, 100$ in each of the three scenarios, the mean of posterior mean estimates, mean of estimation bias $\Delta = |\hat{\mu} - \mu|$, and mean of standard error estimates for parameters $\beta, \kappa, \gamma_1, \gamma_2, \gamma_3$ and hyperparameters μ_0 and σ are calculated.

When the models were specified correctly, the biases converges to 0 as the number of drivers increases; the standard errors converges to 0 roughly proportional to the square root of the number of drivers ($\sqrt{D}$), which is consistent with the central limit theorem. When the models are not specified correctly, there are still a fair amount of biases when the number of drivers increases and the speed of converging to zero is not consistent with either the other two correctly specified simulation scenarios or the central limit theorem. The Gelman-Rubin diagnostic $\hat{R}$ were all lower than 1.1 and no low effective sample size (ESS) issues were reported in `Stan`, suggesting that steady posterior distributions were reached while estimating for the simulated data sets.

5. REAL DATA ANALYSIS

5.1 Data description

A naturalistic truck driving data set was provided to the research team by a national commercial trucking company in North America. The data set includes 496 regional drivers

Table 1: Simulation results for PLP, JPLP, and PLP ← JPLP

sim_scenaio	D	estimate	γ_1	γ_2	γ_3	β	κ	μ_0	σ_0
PLP	10	bias Δ	0.0203	0.0095	0.0067	0.0102		0.0282	0.0527
PLP	25	bias Δ	0.0066	0.0046	0.0012	0.0045		0.0015	0.0220
PLP	50	bias Δ	0.0040	0.0033	0.0005	0.0017		0.0068	0.0077
PLP	75	bias Δ	0.0034	0.0004	0.0007	0.0017		0.0026	0.0091
PLP	100	bias Δ	0.0009	0.0009	0.0003	0.0006		0.0034	0.0042
PLP	10	s.e.	0.0777	0.0696	0.0413	0.0589		0.2401	0.1722
PLP	25	s.e.	0.0459	0.0414	0.0247	0.0360		0.1392	0.0916
PLP	50	s.e.	0.0316	0.0286	0.0172	0.0254		0.0960	0.0610
PLP	75	s.e.	0.0258	0.0232	0.0139	0.0207		0.0784	0.0497
PLP	100	s.e.	0.0220	0.0198	0.0119	0.0179		0.0667	0.0420
JPLP	10	bias Δ	0.0331	0.0218	0.0092	0.0226	0.0149	0.0401	0.0696
JPLP	25	bias Δ	0.0158	0.0081	0.0039	0.0131	0.0084	0.0202	0.0219
JPLP	50	bias Δ	0.0037	0.0012	0.0039	0.0057	0.0032	0.0014	0.0111
JPLP	75	bias Δ	0.0060	0.0012	0.0006	0.0058	0.0028	0.0057	0.0097
JPLP	100	bias Δ	0.0048	0.0003	0.0008	0.0043	0.0023	0.0004	0.0041
JPLP	10	s.e.	0.0992	0.0834	0.0498	0.0828	0.0573	0.2556	0.1854
JPLP	25	s.e.	0.0586	0.0477	0.0288	0.0512	0.0360	0.1453	0.0960
JPLP	50	s.e.	0.0406	0.0334	0.0201	0.0366	0.0256	0.0999	0.0647
JPLP	75	s.e.	0.0331	0.0272	0.0164	0.0298	0.0208	0.0812	0.0519
JPLP	100	s.e.	0.0287	0.0233	0.0141	0.0258	0.0179	0.0699	0.0442
PLP ← JPLP	10	bias Δ	0.1923	0.0645	0.0434	0.1843		0.1234	0.1599
PLP ← JPLP	25	bias Δ	0.1769	0.0514	0.0374	0.1740		0.0866	0.1053
PLP ← JPLP	50	bias Δ	0.1718	0.0531	0.0355	0.1734		0.0854	0.0977
PLP ← JPLP	75	bias Δ	0.1686	0.0511	0.0346	0.1724		0.0874	0.0960
PLP ← JPLP	100	bias Δ	0.1674	0.0512	0.0349	0.1713		0.0811	0.0925
PLP ← JPLP	10	s.e.	0.1041	0.0946	0.0559	0.0580		0.2952	0.2078
PLP ← JPLP	25	s.e.	0.0609	0.0546	0.0329	0.0354		0.1671	0.1095
PLP ← JPLP	50	s.e.	0.0423	0.0383	0.0230	0.0250		0.1167	0.0743
PLP ← JPLP	75	s.e.	0.0344	0.0310	0.0186	0.0204		0.0946	0.0601
PLP ← JPLP	100	s.e.	0.0297	0.0266	0.0160	0.0177		0.0810	0.0514

who move freights in regional routes that may include several surrounding states. A total of 13,187,289 ping records were generated between April 2015 and March 2016, with a total traveled distance of 20,042,519 miles in 465,641 hours (average speed 43 miles per hour). Each ping records the date and time (year, month, day, hour, minute, and second), latitude and longitude (specific to five decimal places), driver identification number, and speed at that time point. These pings were then aggregated into 64,860 shifts and 180,408 segments.

On the other hand, 8,407 kinematic SCEs were recorded independent of the pings, including 3941 (46.9%) headway, 3576 (42.5%) hard brakes, 869 (10.3%) collision mitigation, and 21 (0.2%) rolling stability. Historic weather data, including precipitation probability, precipitation intensity, and wind speed, were queried from the `DarkSky` Application Programming Interface, which provides historic real-time and hour-by-hour nationwide historic weather conditions for specific latitude-longitude-date-time combinations (The Dark Sky Company, LLC, 2020). The weather data were then merged back to pings data and aggregated to shift- and segment-level by taking the mean.

5.2 Real data analysis results

We applied the hierarchical Bayesian PLP and JPLP models to this data as specified in Equations 2 and 6. Since we have four types of SCEs, we then applied the JPLP to the four different types of SCEs separately. Collision mitigation and rolling stability were combined as one type because the later one is very scare and will yield very unstable estimates if modeled alone. Samples of the posterior distributions were drawn using the probabilistic programming language `Stan` in R (Carpenter et al., 2017; Stan Development Team, 2018). The convergence of Hamiltonian Monte Carlo was checked using Gelman-Rubin diagnostic statistics $\hat{R}$ (Gelman et al., 1992), effective sample size (ESS), and trace plots.

Table 2 presents the posterior mean, 95% credible interval (CI), Gelman-Rubin diagnostic statistics $\hat{R}$, and ESS for the sample 496 regional drivers using PLP and JPLP. In both the PLP and JPLP models, the posterior means of the shape parameters β are less than one and the 95% credible intervals exclude one, indicating SCEs occur in the early stages of the shifts. In JPLP, the reliability jump parameter κ was close to 1, suggesting that within-shift rests have very minor effects on the reliability of SCEs. Figure 3 shows the random intercepts γ_{0d} in both the PLP and JPLP present a fair amount of variability across different drivers. Besides, the random intercepts γ_{0d} are on average larger in JPLP than those in PLP models, while variability of random intercepts is similar in the two models. These patterns are consistent with the parameter estimates of μ_0 and σ_0 in Table 2. All the Gelman-Rubin diagnostic statistics $\hat{R}$ are less than 1.1 and the ESSs are greater than 1,000. The trace plots of important variables in the Appendix figure 4 indicate that all four chains for the parameters are well mixed. All these evidence suggests that a steady state posterior distribution have been reached for the two models.

Table 2: Posterior mean, 95% credible interval, Rhat, and effective sample size (ESS) of PLP and JPLP models for 496 commercial truck drivers

Parameters	Power law process				Jump power law process			
	mean	95% CI	$\hat{R}$	ESS	mean	95% CI	$\hat{R}$	ESS
β	0.968	(0.948, 0.988)	1.000	6,500	0.962	(0.940, 0.985)	1.001	3,798
κ					1.020	(0.995, 1.045)	1.000	5,400
μ_0	3.038	(2.397, 3.688)	1.001	2,979	3.490	(2.899, 4.091)	1.001	3,079
σ_0	0.974	(0.897, 1.058)	1.000	9,581	0.982	(0.905, 1.066)	1.000	9,050
Age	0.003	(-0.005, 0.012)	1.001	2,250	0.004	(-0.005, 0.012)	1.001	2,566
Race: black	-0.113	(-0.329, 0.103)	1.002	1,951	-0.130	(-0.342, 0.087)	1.001	2,277
Race: other	-0.343	(-0.707, 0.021)	1.001	2,833	-0.361	(-0.729, 0.010)	1.001	3,334
Gender: female	-0.071	(-0.441, 0.300)	1.001	3,069	-0.071	(-0.435, 0.296)	1.001	4,162
Mean speed	0.019	(0.016, 0.023)	1.000	20,229	0.015	(0.013, 0.018)	1.000	19,827
Speed variation	0.026	(0.017, 0.034)	1.000	24,825	0.017	(0.013, 0.022)	1.000	13,127
Preci. intensity	-3.608	(-6.181, -0.935)	1.000	22,025	-2.136	(-3.785, -0.368)	1.000	24,397
Preci. prob.	0.397	(0.168, 0.628)	1.000	21,416	0.121	(-0.050, 0.296)	1.000	25,329
Wind speed	0.018	(0.008, 0.029)	1.000	32,980	0.010	(0.001, 0.018)	1.000	33,093

95% CI: 95% credible interval; ESS: effective sample size;
PLP: power law process; JPLP: jump power law process;
Preci. intensity: precipitation intensity; Precip. prob.: precipitation probability.

We further estimated the JPLP models for different types of SCEs (headway, hard brakes, and collision mitigation and rolling stability), and the results are presented in Table 3. Headway and hard brake are similar: they have very close posterior mean and 95% credible

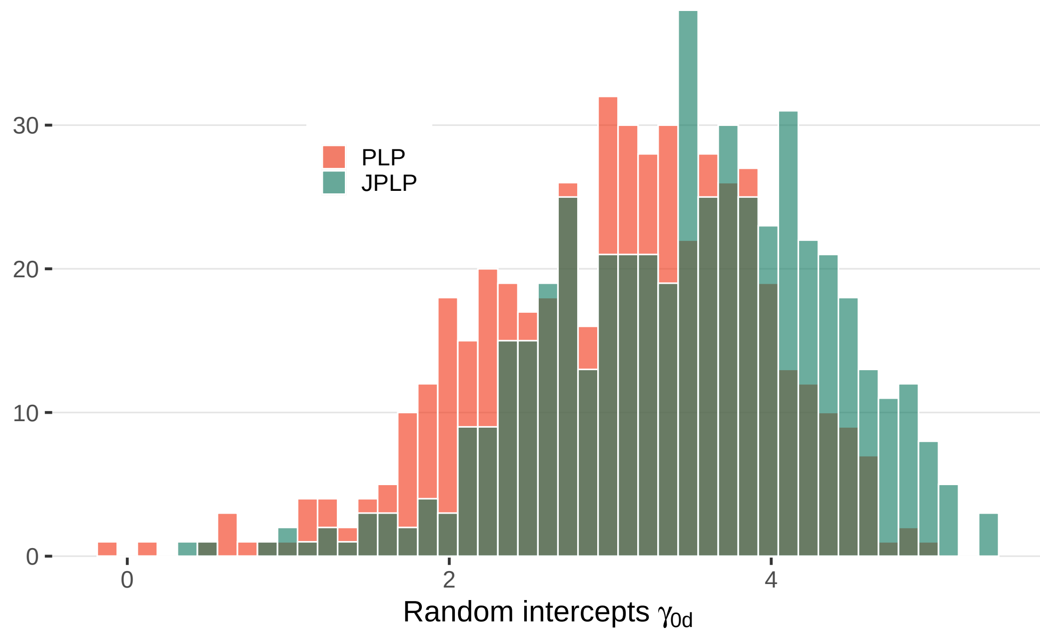

Figure 3: Histogram of random intercepts γ_{0d} across the 496 drivers.

intervals for parameters β and κ, although the hyperparameters for random intercepts are quite different. The results that $\beta < 1$ and $\kappa > 1$ suggest that headway and hard brake tend to occur in the early stages of driving shifts, and taking short breaks will slightly increase the intensity of these two events. In contrast, collision mitigation and rolling stability show a different pattern: they tend to occur in later stages of driving shifts, and taking short breaks will reduce the intensity of the event. The variability across drivers σ_0 is more evidence for headway than hard brake, collision mitigation and rolling stability.

6. DISCUSSION

In this article, we proposed a Bayesian hierarchical NHPP with PLP intensity function and an innovative Bayesian hierarchical JPLP to model naturalistic truck driving data. Our motivation comes from more popular use of naturalistic driving data sets in the recent decade and real-life truck driving characteristics of multiple segments nested within shifts.

Table 3: Parameter estimates and 95% credible intervals for jump power law process on 496 truck drivers, stratified by different types of safety-critical events

Parameters	Headway	Hard brake	Collision mitigation & Rolling stability
β	0.989 (0.956, 1.023)	0.922 (0.889, 0.955)	1.020 (0.950, 1.096)
κ	1.034 (0.998, 1.071)	1.034 (0.996, 1.072)	0.890 (0.821, 0.964)
μ_0	7.096 (6.083, 8.139)	3.470 (2.770, 4.199)	4.729 (3.836, 5.666)
σ_0	1.564 (1.411, 1.730)	1.073 (0.973, 1.182)	0.922 (0.786, 1.074)
Age	-0.006 (-0.020, 0.009)	0.011 (0.001, 0.021)	0.002 (-0.009, 0.012)
Race: black	0.184 (-0.170, 0.546)	-0.312 (-0.565, -0.064)	0.113 (-0.153, 0.386)
Race: other	0.306 (-0.340, 0.967)	-0.539 (-0.968, -0.106)	0.100 (-0.373, 0.605)
Gender: female	0.266 (-0.343, 0.870)	-0.217 (-0.654, 0.230)	-0.181 (-0.675, 0.309)
Mean speed	-0.026 (-0.031, -0.021)	0.043 (0.039, 0.047)	0.039 (0.032, 0.046)
Speed variation	-0.009 (-0.017, -0.002)	0.017 (0.010, 0.024)	0.013 (-0.002, 0.027)
Preci. intensity	-0.771 (-4.306, 3.188)	-1.912 (-3.924, 0.269)	-0.676 (-6.329, 6.297)
Preci. prob.	0.694 (0.376, 1.015)	-0.495 (-0.724, -0.263)	0.808 (0.206, 1.423)
Wind speed	0.003 (-0.009, 0.015)	0.019 (0.005, 0.034)	0.000 (-0.025, 0.026)

Preci. intensity: precipitation intensity; Precip. prob.: precipitation probability.

The proposed JPLP accounts for the characteristics of multiple rests within a shift among commercial truck drivers. The intensity functions, parameterization forms, and likelihood functions are specified separately. Simulation studies showed the consistency of the Bayesian hierarchical estimation if the models are specified correctly, as well as the persistent bias when the models are not specified correctly. A case study of 496 commercial truck drivers demonstrates a fair amount of variability exist across drivers. Headway and hard brake tend to occur in early stages while collision mitigation and rolling stability tend to occur in later stages.

The SCEs generated from naturalistic truck driving data are different from previous reliability problems and models in two aspects. Most previous studies either assume minimal repair or perfect repair (also known as renewal process). A minimal repair assumes the unit after the repair is exactly the same as it is before the repair (for example, our first proposed NHPP with a PLP intensity function), while a perfect repair assumes the unit is a completely new unit after repair (Rigdon and Basu, 2000). In the scenario of truck driving, although it is reasonable to assume that the drivers experience a perfect repair when they take a break of longer than 10 hours, researchers would not expect the reliability to be either perfectly repaired or minimally repaired during a short break of around half an hour. Instead, a partial repair is a more proper assumption here.

On the other hand, even though some studies proposed partial repair reliability models such as the modulated PLP (Lakey and Rigdon, 1993; Black and Rigdon, 1996), none of these previously proposed models fit for the data here since the repairs are independent of the SCEs in naturalistic driving data. These previous models are based on high-tech devices such as aircraft manufacturing, which need immediate repair once a failure is detected. However in the case of naturalistic truck driving data sets, SCEs such as hard brakes and headway do not seriously influence the driving and drivers generally will keep driving even if SCEs occur. The repair (breaks) can be considered as independent of SCEs in our study. Although our case study is based on naturalistic truck driving data sets, the JPLP can applied to any type of drivers who drive for a long distance with at least one break.

Our work can be extended in several aspects in the future. First, the assumption of proportion reliability jump may not hold. Other proper assumptions include reliability jumping for a fixed-amount jump or jumping dependent on the length of the rest. Additionally, in our proposed JPLP, the length of breaks within shifts are ignored to simply the parameterization and likelihood function. In truck transportation practice, longer breaks certainly have larger effects on reliability jump, hence the relationship between reliability jump and the length of breaks can have more complex functional forms, so it would be of interest to test different forms of reliability change as a function of the length of break.

Appendix A TRACE PLOTS

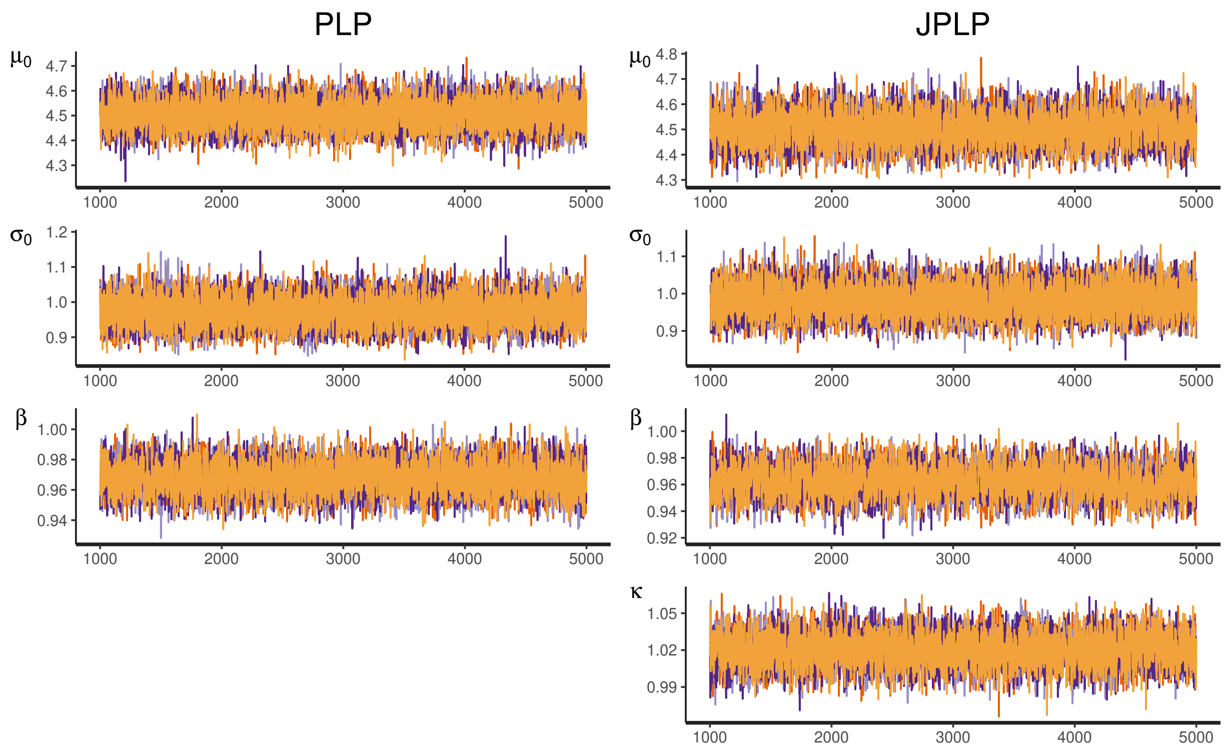

Figure 4: Trace plots of select parameters in PLP (left column) and JPLP (right column) for all types of SCEs

SUPPLEMENTARY MATERIALS

Since the real truck driving data provided by our industry partner are confidential and cannot be made publicly accessible, we provided a simulated data set that is similar to our real data in data structure but has fewer drivers to be completed within a reasonable amount of time. The online supplementary materials contain the R code to simulate PLP and JPLP data, explanation on the data structure, and Stan and R code for Bayesian hierarchical PLP and JPLP estimation. The supplementary materials can be access at `https://for-blind-external-review.github.io/JPLP/`.

ACKNOWLEDGEMENTS

The authors are grateful to our industry partner for providing us the naturalistic truck driving data for research in this study.

FUNDING

The work was supported by the National Science Foundation (CMMI-1635927 and CMMI-1634992), the Ohio Supercomputer Center (PMIU0138 and PMIU0162). We also thank the DarkSky company for providing us five million free calls to their historic weather API.

REFERENCES

Black, S. E. and Rigdon, S. E. (1996). Statistical inference for a modulated power law process. *Journal of Quality Technology*, 28(1):81–90.

Cai, M., Alamdar Yazdi, M. A., Hu, Q., Mehdizadeh, A., Vinel, A., Davis, K. C., Xian, H., Megahed, F. M., and Rigdon, S. E. (2020). The association between crashes and safety-critical events: synthesized evidence from crash reports and naturalistic driving data among commercial truck drivers. *Transportation Research Part C: Emerging Technologies*, Under review.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., and Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1).

Das, A., Ghasemzadeh, A., and Ahmed, M. M. (2019). Analyzing the effect of fog weather conditions on driver lane-keeping performance using the shrp2 naturalistic driving study data. *Journal of safety research*, 68:71–80.

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., et al. (2006). The 100-car naturalistic driving study. phase 2: Results of the 100-car field experiment. Technical report, United States. Department of Transportation. National Highway Traffic Safety.

Federal Highway Administration (2019). Human factors. U.S. Department of Transportation, `https://highways.dot.gov/research/research-programs/safety/human-factors`. [Updated December 2, 2019; accessed July 03, 2020].

Federal Motor Carrier Safety Administration (2017). Summary of hours of service regulations. [Online; accessed 20-February-2019].

Federal Motor Carrier Safety Administration (2020). Summary of hours of service regulations. United States Department of Transportation. `https://www.fmcsa.dot.gov/regulations/hours-service/summary-hours-service-regulations`. [Accessed July 03, 2020].

Gelman, A., Rubin, D. B., et al. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472.

Ghasemzadeh, A. and Ahmed, M. M. (2017). Drivers' lane-keeping ability in heavy rain: preliminary investigation using shrp 2 naturalistic driving study data. *Transportation research record*, 2663(1):99–108.

Ghasemzadeh, A. and Ahmed, M. M. (2018). Utilizing naturalistic driving data for in-depth analysis of driver lane-keeping behavior in rain: Non-parametric mars and parametric logistic regression modeling approaches. *Transportation research part C: emerging technologies*, 90:379–392.

Gordon, T. J., Kostyniuk, L. P., Green, P. E., Barnes, M. A., Blower, D., Blankespoor, A. D., and Bogard, S. E. (2011). Analysis of crash rates and surrogate events: unified approach. *Transportation Research Record*, 2237(1):1–9.

Guo, F. (2019). Statistical Methods for Naturalistic Driving Studies. *Annual Review of Statistics and Its Application*, 6:309–328.

Guo, F., Kim, I., and Klauer, S. G. (2019). Semiparametric Bayesian Models for Evaluating Time-Variant Driving Risk Factors Using Naturalistic Driving Data and Case-Crossover Approach. *Statistics in Medicine*, 38(2):160–174.

Guo, F., Klauer, S. G., Hankey, J. M., and Dingus, T. A. (2010). Near crashes as crash surrogate for naturalistic driving studies. *Transportation Research Record*, 2147(1):66–74.

John, S. (2019). 11 incredible facts about the $700 billion US trucking industry. Business Insider: Markets Insider. `https://markets.businessinsider.com/news/stocks/trucking-industry-facts-us-truckers-2019-5-1028248577`. [Published online June 3, 2019; accessed July 03, 2020].

Kim, S., Chen, Z., Zhang, Z., Simons-Morton, B. G., and Albert, P. S. (2013). Bayesian Hierarchical Poisson Regression Models: An Application to a Driving Study With Kinematic Events. *Journal of the American Statistical Association*, 108(502):494–503.

Lakey, M. J. and Rigdon, S. E. (1993). Reliability Improvement Using Experimental Design. In *Annual Quality Congress Transactions-American Society for Quality Control*, volume 47, pages 824–824. American Society for Quality Control.

Li, Q., Guo, F., Kim, I., Klauer, S. G., and Simons-Morton, B. G. (2018). A Bayesian Finite Mixture Change-Point Model for Assessing the Risk of Novice Teenage Drivers. *Journal of Applied Statistics*, 45(4):604–625.

Liu, Y. and Guo, F. (2019). A Bayesian Time-Varying Coefficient Model for Multitype Recurrent Events. *Journal of Computational and Graphical Statistics*, pages 1–12.

Liu, Y., Guo, F., and Hanowski, R. J. (2019). Assessing the Impact of Sleep Time on Truck Driver Performance using a Recurrent Event Model. *Statistics in Medicine*, 38(21):4096–4111.

Lord, D. and Mannering, F. (2010). The Statistical Analysis of Crash-Frequency Data: A Review and Assessment of Methodological Alternatives. *Transportation Research Part A: Policy and Practice*, 44(5):291–305.

Mannering, F. L. and Bhat, C. R. (2014). Analytic Methods in Accident Research: Methodological Frontier and Future Directions. *Analytic Methods in Accident Research*, 1:1–22.

Mehdizadeh, A., Cai, M., Hu, Q., Yazdi, A., Ali, M., Mohabbati-Kalejahi, N., Vinel, A., Rigdon, S. E., Davis, K. C., and Megahed, F. M. (2020). A Review of Data Analytic Applications in Road Traffic Safety. Part 1: Descriptive and Predictive Modeling. *Sensors*, 20(4):1107.

NHTSA's National Center for Statistics and Analysis (2019). 2017 data: large trucks. U.S. Department of Transportation. National Highway Traffic Safety Administration. Traffic Safety Facts. DOT HS 812 663. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812663. [Published online January 2019; accessed July 03, 2020].

Rigdon, S. E. and Basu, A. P. (1989). The Power Law Process: A Model for the Reliability of Repairable Systems. *Journal of Quality Technology*, 21(4):251–260.

Rigdon, S. E. and Basu, A. P. (2000). *Statistical Methods for the Reliability of Repairable Systems*. Wiley New York.

Savolainen, P. T., Mannering, F. L., Lord, D., and Quddus, M. A. (2011). The Statistical Analysis of Highway Crash-injury Severities: A Review and Assessment of Methodological Alternatives. *Accident Analysis & Prevention*, 43(5):1666–1676.

Stan Development Team (2018). RStan: the R interface to Stan. R package version 2.18.2.

The Dark Sky Company, LLC (2020). Dark Sky API — Overview. https://darksky.net/dev/docs. [Online; accessed 20-February-2020].

The White House (2020). Remarks by President Trump celebrating America's truckers. https://www.whitehouse.gov/briefings-statements/remarks-president-trump-celebrating-americas-truckers/. [Issued on April 16, 2020; accessed July 03, 2020].

Tsai, Y.-T., Swartz, S. M., and Megahed, F. M. (2018). Estimating the relative efficiency of highway safety investments on commercial transportation. *Transportation Journal*, 57(2):193–218.