Hierarchical Point Process Models for Recurring Safety Critical Events Involving Commercial Truck Drivers: A Reliability Framework for Human Performance Modeling

Abstract

Factors that lead to an increased risk of a crash in commercial trucks are investigated. Since crashes are rather rare, we use safety critical events (SCEs) such as hard breaks as a proxy for crashes. While many previous studies have focused on a crashes on a fixed road segment, we follow 496 commercial truck drivers who drove over 13 million miles over one year and incurred 8,386 SCEs. This naturalistic driving study, which is analogous to a prospective cohort study, is many times larger than any study done to date. Such a study design has advantages over the study of crashes on a fixed road segment. We address two questions related to trucking safety: whether the occurrence of SCEs tends to increase during a shift, and (2) the effect of rest breaks on SCEs. We apply point process models, similar to those employed for studying the reliability of repairable systems and find that the intensity for hard breaks decreases throughout a shift, while rest breaks reduce the likelihood of activation of the automated collision mitigation system. Properties of the approach are investigated through a simulation study. Supplementary materials, including simulated data and code, 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-delivery pay schedules. Furthermore, the industry is heavily regulated through the *hours of service* regulations (Federal Motor Carrier Safety Administration, 2020b), which dictate 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 truck drivers and commuters is the complex operational environment required of commercial drivers, who must abide by government regulations while managing industry practices that attempt to optimize both productivity and safety.

Truck safety is important not only to trucking operators, but also to the general public. Truck crashes pose a two-fold risk (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. Despite the regulatory oversights and continued advancements in safety technologies, the rates of truck-involved crashes in the U.S. have increased from 1.32 per 100 million miles traveled increased in 2008 to 1.48 in 2016 for fatal crashes, and from 21 in 2008 to 31 in 2015, for injury crashes (NHTSA, 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 frequency of crashes in a fixed time period (Lord and Mannering, 2010; Savolainen et al., 2011; Mannering and Bhat, 2014). Such models resemble a case-control study design (Mehdizadeh et al., 2020). Limitations of these 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, these 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). These 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. We study three categories of SCEs: hard breaks, headway (an unsafe gap between vehicles), and the activation of an automatic system for collision mitigation. It is now well-established that the occurrence of 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 higher statistical power.

Models for SCEs can be divided into those 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), and (b) model the times of observing

SCEs (Li et al., 2018; Liu and Guo, 2019; Liu et al., 2019; Guo et al., 2019). NDSs have three advantages over the analysis of road segments using a case-control design. First, NDSs follow drivers for an extended time period, i.e., their application resembles a prospective cohort study (Mehdizadeh et al., 2020). Studies on road segments utilize a case-control design to this type of data by including all events and matching them with selective non-events (e.g., Ghasemzadeh and Ahmed, 2018; Das et al., 2019), which reduces the power to detect potentially small effects and fails to account for the fact that the driving data are nested within drivers. Second, the occurrence of multiple SCEs in an extended time period is not unusual. Thus, binary classification models are inefficient since they cannot distinguish between cases where one or more than one SCEs occur. Furthermore, count models fail to consider the time stamps associated with each SCE, which can be helpful in designing interventions. Third, based on the *hours of service* regulations, breaks are required for intermediate and long trips. The underlying hypothesis is that these breaks would improve the driver's safety performance, which cannot be considered in a case-control study.

The overarching motivation of this study is to examine how large NDS data sets can be modeled to account for both the timing of an observed event and the effect of rest breaks on SCE occurrence. This study is performed in collaboration with a leading shipping freight company in the U.S. Teaming with industry provides the following unique settings: (a) the company's fleet used a commercially available driving event monitoring system, which means that the SCE data were collected routinely as a part of the fleet's operations; (b) the truck drivers included in this study were all employed by the company at the time of data collection, i.e., a consistent operational and safety policy governs the drivers' behavior; and (c) the routes are subject to company policies, delivery windows and government regulations, i.e., naturally follow realistic commercial driving patterns. Based on this setup, we have naturalistic driving data generated by 496 regional commercial large-truck drivers, capturing over 20 million miles driven and over 8,300 SCEs. Note that existing trucking NDS datasets are much smaller, with the largest reported values of approximately 200 drivers (Federal Motor Carrier Safety Administration, 2020a) and 0.414 million miles driven (Sparrow et al., 2016). This study, which involves nearly 50 times as many miles driven as the second largest NDS, is able to detect small effects that other studies might miss.

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 driving safety performance, and consequently, to what extent does safety change after a rest break? To facilitate the modeling of these two sets of questions, we introduce and capitalize on the following analogies:

- The first set of questions can be considered as a degradation process, where continued driving results in a degraded safety performance similar to the way continued operation degrades a repairable system in the field of *reliability*.
- A rest-break can be considered a preventive maintenance activity, which can be either time-based (e.g., every two hours) or condition-based (e.g., if drivers stop for coffee to increase their alertness). The *maintenance* act improves the driver's reliability by reducing the degradation, which we hypothesize will reduce the likelihood of an SCE.
- A potential difference between a product's and a driver's reliability is that products of similar vintage are typically assumed to be homogeneous. On the other hand, the

modeling of drivers should be personalized (i.e., assuming heterogeneity of the sampling units), which can be accounted for using hierarchical modeling approaches.

With these research questions and analogy in mind, we introduce a Bayesian hierarchical non-homogeneous Poisson process with the power law process (PLP) intensity function to model SCEs within shifts. This model can account for driver-level unobserved heterogeneity by specifying driver-level random intercepts for the rate parameter in PLP. To account for the feature that multiple breaks are 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 an illustration of using PLP and JPLP in modeling the intensity function of SCEs.

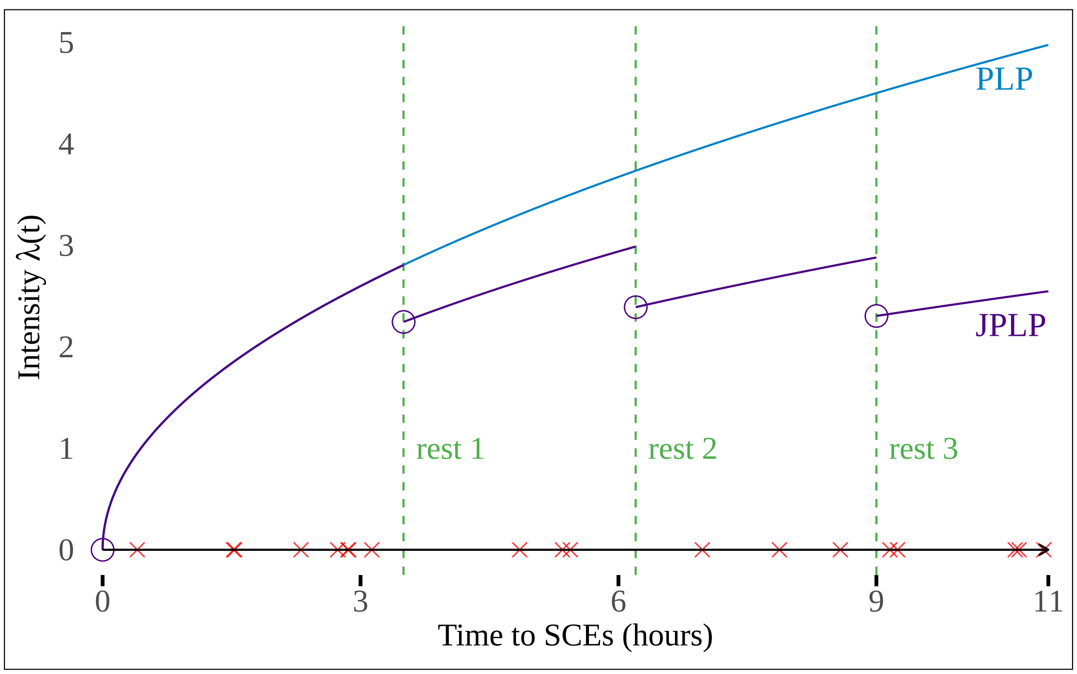

Figure 1: An illustration of a simulated intensity function of PLP and JPLP. The x-axis shows time in hours since start and y-axis shows the intensity of SCEs. The crosses mark the time to SCEs and the vertical dotted lines indicates the time of the rests.

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, we present the results of real data analyses for 496 commercial truck drivers using the PLP and JPLP. In Section 5, simulation studies are conducted to investigate the properties of the JPLP model and the consequences if the models are not specified correctly. In Section 6, strengths, possible limitations, and future research directions are discussed. To facilitate the replication of our analysis, we provide a simulated data set, description of the data structure, and `Stan` and `R` (Carpenter et al., 2017; Stan Development Team, 2018) code for Bayesian PLP and JPLP estimation as supplementary material, which we host online on a `GitHub` page.

2. TERMINOLOGY AND NOTATION

NDSs collect data by periodically recording data (location, speed, etc) at a certain frequency, which ranges from every few seconds to approximately 15 minutes (Cai et al., 2020). These data points, called *pings*, are then aggregated into segments, and then shifts. Figure 2 presents a time series plot of speed data for a sample driver (including two shifts and six segments nested within the shifts), as well as the aggregation process to shifts and segments suggested by arrows. We use $d = 1, 2, \ldots, D$ as the index for drivers. A shift $s = 1, 2, \ldots, S_d$ is an on-duty period with no breaks longer than 10 hours for driver d. Per the *hours of service* regulations, a shift may be no more than 14 hours with no more than 11 hours of driving. 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.

SCEs can occur any time in the segments whenever preset kinematic thresholds are

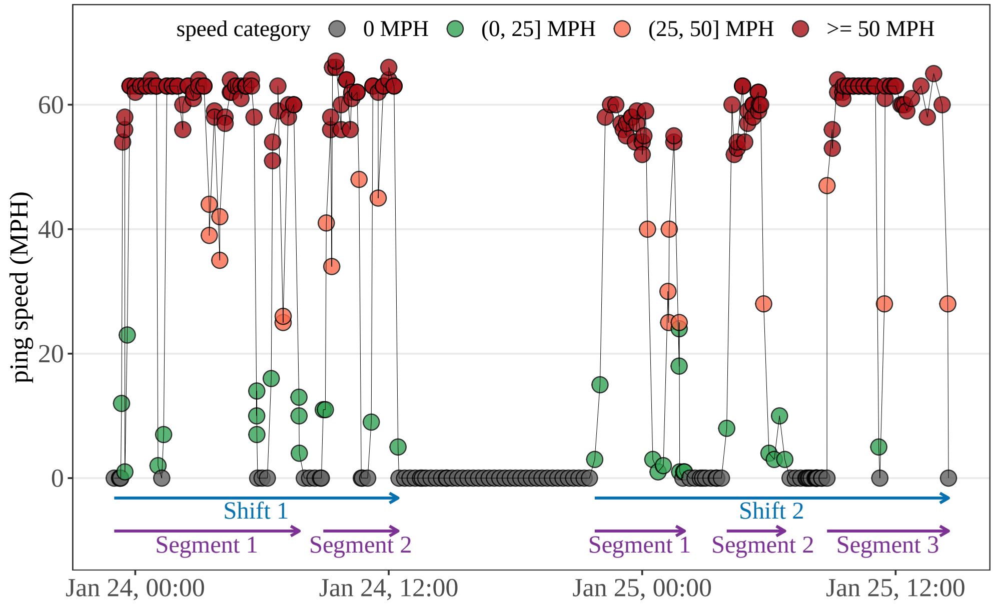

Figure 2: Time series plot of naturalistic truck driving sample ping data (points) and the aggregation process from pings to shifts and segments (arrows).

triggered while driving. We use $i = 1, 2, \ldots, I_{d,s}$ as the index 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 shift s and the rest times between segments are excluded from calculation. The number of SCEs for segment t within shift s is denoted $n_{d,s,r}$. Finally, the end time of segment r within shift$s$ for driver d is denoted $a_{d,s,r}$.

3. MODELS

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

We assume that the times to SCEs t follow a non-homogeneous Poisson process, whose intensity function $\lambda(t)$ is non-constant, having the functional form

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

where the parameter β indicates reliability improvement ($\beta < 1$), constant ($\beta = 1$), or deterioration ($\beta > 1$), and the parameter θ is a scale parameter. The power law process is an established model (Rigdon and Basu, 1989, 2000) with a flexible functional form.

3.2 Bayesian Hierarchical Power Law Process (PLP)

The Bayesian hierarchical power law process is parameterized as

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

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 priors for the parameters and hyperparameters are taken to be the relatively non-informative distributions

$$\begin{aligned}
\beta &\sim \text{Gamma}(1, 1) \\
\gamma_1, \gamma_2, \ldots, \gamma_k &\sim \text{i.i.d. } N(0, 10^2) \\
\mu_0 &\sim N(0, 5^2) \\
\sigma_0 &\sim \text{Gamma}(1, 1).
\end{aligned} \qquad (3)$$

The likelihood function of event times generated from a PLP for driver d in shift s is

$$L_{d,s}(\beta, \gamma_{0d}, \gamma | \mathbf{X}_d, \mathbf{W}_s) = \begin{cases} \exp\left(-\left(\tau_{d,s}/\theta_{d,s}\right)^{\beta}\right), & \text{if } n_{d,s} = 0, \\ \left(\prod_{i=1}^{n_{d,s}} \beta\ \theta_{d,s}^{-\beta}\ t_{d,s,i}^{\beta-1}\right) \exp\left(-(\tau_{d,s}/\theta_{d,s})^{\beta}\right), & \text{if } n_{d,s} > 0, \end{cases} \qquad (4)$$

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 second line of Equation (2). The full likelihood function for all drivers can be computed using

$$L = \prod_{d=1}^{D} \prod_{s=1}^{S_d} L_{d,s}(\beta, \gamma_{0d}, \gamma | \mathbf{X}_d, \mathbf{W}_s) \tag{5}$$

where $L_{d,s}(\beta, \gamma_{0d}, \gamma | \mathbf{X}_d, \mathbf{W}_s)$ is provided in Equation (4). Models like this are often used in the literature for repairable systems reliability (Rigdon and Basu, 2000). Here, a *failure* can be thought of as the occurrence of a SCE.

3.3 Bayesian Hierarchical Jump Power Law Process (JPLP)

Since the Bayesian hierarchical PLP does not account for rest breaks within shifts, and their associated potential performance improvement, we propose a Bayesian hierarchical JPLP with an additional jump parameter κ. Our proposed JPLP has the piecewise intensity function

$$\begin{aligned}
\lambda_{\text{JPLP}}(t|d, s, r, \beta, \gamma_{0d}, \gamma, \mathbf{X}_d, \mathbf{W}_s) &= \begin{cases} \kappa^0 \lambda(t|\beta, \gamma_{0d}, \gamma, \mathbf{X}_d, \mathbf{W}_s), & 0 < t \leq a_{d,s,1}, \\ \kappa^1 \lambda(t|\beta, \gamma_{0d}, \gamma, \mathbf{X}_d, \mathbf{W}_s), & a_{d,s,1} < t \leq a_{d,s,2}, \\ \vdots & \vdots \\ \kappa^{R-1} \lambda(t|\beta, \gamma_{0d}, \gamma, \mathbf{X}_d, \mathbf{W}_s), & a_{d,s,R-1} < t \leq a_{d,s,R}, \end{cases} \\
&= \kappa^{r-1} \lambda(t|d, s, r, \kappa, \beta, \gamma_{0d}, \gamma, \mathbf{X}_d, \mathbf{W}_s) \quad a_{d,s,r-1} < t \leq a_{d,s,r},
\end{aligned} \tag{6}$$

where κ is the proportional change in intensity function after 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 zeroth segment $a_{d,s,0} = 0$ and the end time of the last segment for driver d in shift s equals the shift end time ($a_{d,s,R_{d,s}} = \tau_{d,s}$). We assume that κ is constant across drivers and shifts.

The Bayesian hierarchical JPLP model is parameterized as

$$\begin{aligned}
\left(t_{d,s,1}, t_{d,s,2}, \ldots, t_{d,s,n_{d,s}}\right) \mid (\beta, \theta_{ds}) &\sim \text{JPLP}(\beta, \theta_{d,s}, \kappa, \tau_{d,s}) \\
\log \theta_{d,s} &= \gamma_{0d} + \gamma_1 x_{d,s,1} + \gamma_2 x_{d,s,2} + \cdots + \gamma_k x_{d,s,k} \\
(\gamma_{01}, \gamma_{02}, \ldots, \gamma_{0D}) \mid (\mu_0, \sigma_0^2) &\sim \text{i.i.d. } N(\mu_0, \sigma_0^2),
\end{aligned} \tag{7}$$

With the exception of the κ parameter, the above formulation is identical with that presented in Equation (2). We set the prior distribution for κ as uniform(0, 2), which allows the intensity function to change by a factor that ranges from 0 to 2 at rest breaks. The priors and hyperpriors for the JPLP are assigned as

$$\begin{aligned}
\beta &\sim \text{Gamma}(1, 1) \\
\kappa &\sim \text{Uniform}(0, 2) \\
\gamma_1, \gamma_2, \ldots, \gamma_k &\sim \text{i.i.d. } N(0, 10^2) \\
\mu_0 &\sim N(0, 5^2) \\
\sigma_0 &\sim \text{Gamma}(1, 1).
\end{aligned} \tag{8}$$

The likelihood function of event times generated from a JPLP for driver d on shift s is then

$$\begin{aligned}
L^*_{d,s}(\kappa, \beta, \gamma_{0d}, \gamma | \mathbf{X}_d, \mathbf{W}_s) &= \left(\prod_{i=1}^{n_{d,s}} \lambda_{\text{JPLP}}(t_{d,s,i})\right) \exp\left(-\int_0^{\tau_{d,s}} \lambda_{\text{JPLP}}(u)\ du\right) \\
&= \begin{cases} \exp\left(-\int_0^{\tau_{d,s}} \lambda_{\text{JPLP}}(u)\ du\right), & \text{if } n_{d,s} = 0, \\ \left(\prod_{i=1}^{n_{d,s}} \lambda_{\text{JPLP}}(t_{d,s,i})\right) \exp\left(-\int_0^{\tau_{d,s}} \lambda_{\text{JPLP}}(u) du\right), & \text{if } n_{d,s} > 0, \end{cases}
\end{aligned} \tag{9}$$

where the piecewise intensity function $\lambda_{\text{JPLP}}(t_{d,s,i})$ is given in Equation (6). Since the intensity function depends on the segment r for a given driver d on shift s, it is easier to present the likelihood function at a segment level, which can be computed as

$$
\begin{aligned}
& L^*_{d,s,r}(\kappa, \beta, \gamma_{0d}, \gamma | \mathbf{X}_d, \mathbf{W}_r) \\
& = \begin{cases}
\exp\left(-\int_{a_{d,s,r-1}}^{a_{d,s,r}} \lambda_{\text{JPLP}}(u)du\right), & \text{if } n_{d,s,r} = 0, \\
\left(\prod_{i=1}^{n_{d,s,r}} \lambda_{\text{JPLP}}(t_{d,s,r,i})\right) \exp\left(-\int_{a_{d,s,r-1}}^{a_{d,s,r}} \lambda_{\text{JPLP}}(u)du\right), & \text{if } n_{d,s,r} > 0,
\end{cases}
\end{aligned}
\tag{10}
$$

where the intensity function λ_{JPLP} is fixed for driver d on shift s and segment r, $t_{d,s,r,i}$ denotes the time to the i^{th} SCE for driver d on shift s and segment r measured from the beginning of the shift, and $n_{d,s,r}$ is the number of SCEs for driver d on shift s and segment r.

Compared to the PLP likelihood function given in Equation (5), where $\mathbf{W}_s$ are assumed to be fixed numbers during an entire shift, the rewritten likelihood function for JPLP in Equation (10) allows the external covariates $\mathbf{W}_r$ to vary between different segments within a shift. In this way, the JPLP can account for the variability between different segments within a shift. Thus, 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

$$
L^* = \prod_{d=1}^{D} \prod_{s=1}^{S_d} \prod_{r=1}^{R_{d,s}} L^*_{d,s,r}(\kappa, \beta, \gamma_{0d}, \gamma | \mathbf{X}_d, \mathbf{W}_r), \tag{11}
$$

where $L^*_{d,s,r}$ is a likelihood function given in Equation (10), in which the intensity function λ_{JPLP} has a fixed functional form provided in the last line of Equation (6) for a certain driver d in a given shift s and segment r.

4. DATA ANALYSIS

4.1 Data description

The NDS dataset was generated by an on-board sensor monitoring system based on the routes driven by 496 regional large-truck drivers between April 2015 and March 2016. Note that a regional driver's job typically entails moving freight within a geographic region encompassing several surrounding states. As such, they are typically on the road for five days or more, returning home on a weekly basis. A total of 13,187,289 ping records were generated, 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 (with precision of five decimal places), driver identification number, and speed at that time point. The geographic distribution of non-zero-speed (active) pings is depicted in Figure 3, which shows that most pings correlate with the U.S.'s population density. These pings were then aggregated into 64,860 shifts and 180,408 segments.

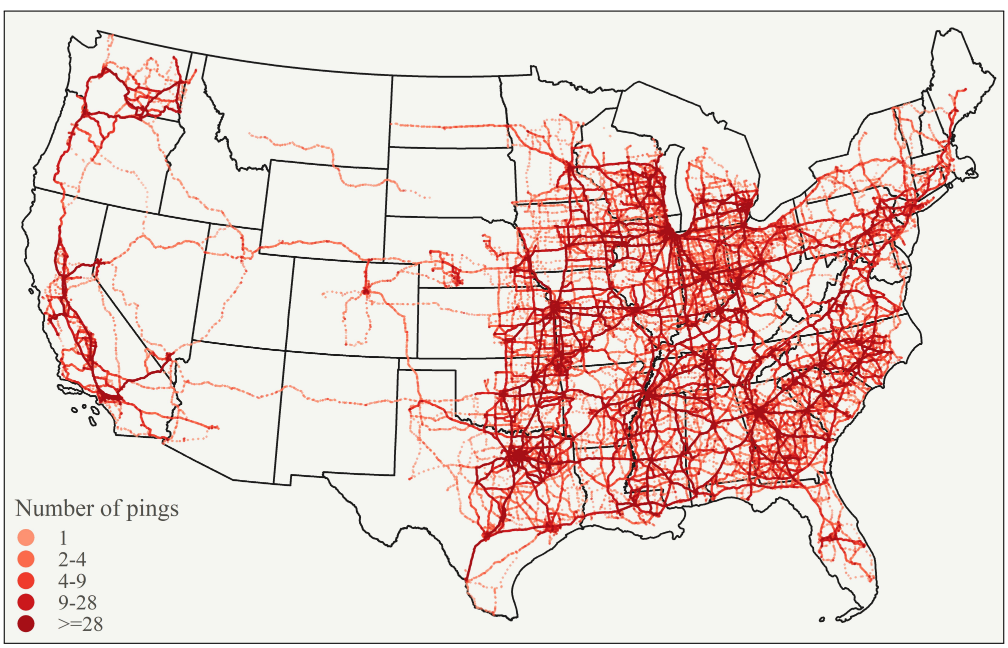

Figure 3: Active pings captured from the 496 regional commercial truck drivers.

Independent of the active, *safe driving*, ping data, 8,386 kinematic SCEs were captured. This corresponds to an overall SCE rate of approximately 0.42 per thousand miles driven. The observed SCEs were divided into three categories: (a) 3,941 (47%) headway events, where the sensor-based monitoring system captures sustained tailgating for ≥ 118 seconds at an unsafe gap time ≤ 2.8 seconds; (b) 3,576 (42.6%) hard brakes, where the truck decelerates at a rate ≥ 9.5 miles per hour per second; and (c) 869 (10.4%) collision mitigation events, which corresponded to instances when a truck's collision mitigation system was initiated. Among the shifts with at least one SCEs (N=6,112), 21.3% (N=1,302) of them have at least two SCEs in one shift. Note that headway represents the least severe SCE since no intervention is needed. On the other hand, collision mitigation is the most severe since the truck's system overrides the driver's control by automatically applying the brakes.

To complement the datasets provided by the company, we queried historical weather data (precipitation probability, precipitation intensity, and wind speed) using the `DarkSky` Application Programming Interface (API), which provides 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 the ping data set and aggregated to shift- and segment-level. Table 1 presents the summary statistics of the driver-, shift-, segment-level variables in our data set.

4.2 Model Inference

We applied the hierarchical Bayesian PLP and JPLP models to the aggregated data for all types of SCEs as specified in Equations (2) and (7). We then applied the JPLP to the three different types of SCEs separately. Samples from the posterior distributions were drawn

Table 1: Summary statistics of driver-, shift-, and segment-level variables

Variable	**Statistics**
Median [IQR] of *driver*-level variables (N = 496)	
Age	47 [36, 55]
Race (N (percent))	
White	246 (49.6%)
Black	206 (41.5%)
Other	44 (8.9%)
Male	460 (92.7%)
Distance	34422.9 [13707.5, 68660.9]
Driving hours	808.1 [337.8, 1626.4]
Mean speed	43.1 [40.8, 44.7]
Mean (S.D.) of *shift*-level variables (N = 64,860)	
Speed S.D.	22.6 (4.3)
Preci. intensity	0.0 (0.0)
Preci. prob.	0.1 (0.2)
Wind speed	3.6 (2.5)
Mean (S.D.) of *segment*-level variables (N = 180,408)	
Speed S.D.	18.6 (7.8)
Preci. intensity	0.0 (0.0)
Preci. prob.	0.1 (0.2)
Wind speed	3.6 (2.9)

Abbreviations:
IQR: interquartile range; S.D.: standard deviation;
Preci. intensity: precipitation intensity;
Precip. prob.: precipitation probability.

using the language `Stan`. Four chains are applied to each estimation, with 1,000 warmup and 4,000 post-warmup iterations drawn from the posterior distributions. The convergence of the Hamiltonian Monte Carlo was checked using the 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 the JPLP, the reliability jump parameter κ was close to 1, suggesting that within a shift, rests have very minor effects on the intensity of SCEs.

Table 2: Posterior mean, 95% credible interval, $\hat{R}$, and effective sample size (ESS) of PLP and JPLP models for 496 commercial truck drivers using aggregated SCEs.

Parameters	**Power law process**				**Jump power law process**			
	mean	95% CI	$\hat{R}$	ESS	mean	95% CI	$\hat{R}$	ESS
$\hat{\beta}$	0.968	(0.948, 0.988)	1.000	6,500	0.962	(0.940, 0.985)	1.001	3,798
$\hat{\kappa}$					1.020	(0.995, 1.045)	1.000	5,400
$\hat{\mu}_0$	3.038	(2.397, 3.688)	1.001	2,979	3.490	(2.899, 4.091)	1.001	3,079
$\hat{\sigma}_0$	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

Abbreviations:

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

In Figure 4, we present the histograms for estimates of the random intercepts. The visualization indicates that there is considerable variability across drivers. The random intercepts γ_{0d} are on average larger for the JPLP model than for the PLP model, 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.

In terms of the convergence of the Hamiltonian Monte Carlo, all the Gelman-Rubin diagnostic statistics $\hat{R}$ are less than 1.1 and the ESSs are greater than 1,000. Furthermore, the trace plots of important variables $(\beta, \kappa, \mu_0, \sigma_0)$, presented in our `GitHub` page (see supplementary materials), indicate that all four chains are well mixed. Thus, the evidence suggests

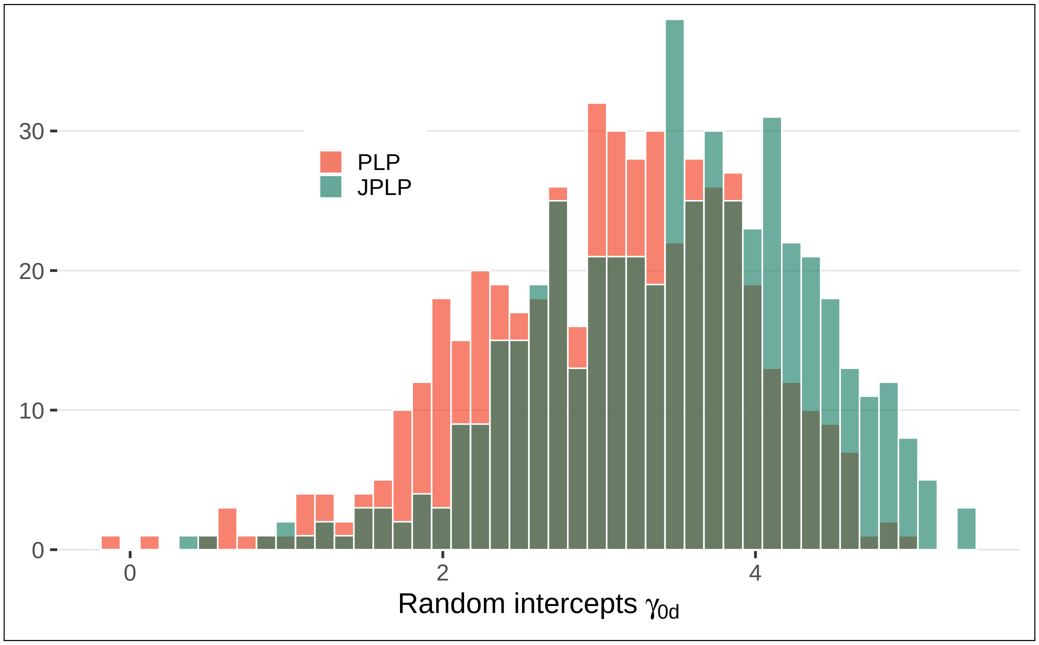

Figure 4: Histogram of random intercepts γ_{0d} across the 496 drivers.

that a steady state posterior distribution has been reached for the PLP and JPLP models.

To estimate the effect of rest breaks on the three different SCEs, we estimated the JPLP models for each type of SCE and the results are presented in Table 3. Headway and hard brakes are similar: the posterior means and 95% credible intervals for parameters β and κ are nearly identical, although the hyperparameters for random intercepts are quite different. The $\hat{\beta} < 1$ and $\hat{\kappa} > 1$ suggest that headway and hard brakes tend to occur in the early stages of driving shifts, and taking short breaks will slightly increase the intensity of these two events (although the credible intervals contain 1). Collision mitigation shows a different pattern, though: it tends to occur in later stages of driving shifts, and taking short breaks will reduce the intensity of this event. The estimate $\hat{\sigma}_0^2$ of the variance of the random intercepts across drivers is larger for headway than hard brake or collision mitigation.

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**
$\hat{\beta}$	0.989 (0.956, 1.023)	0.922 (0.889, 0.955)	1.020 (0.950, 1.096)
$\hat{\kappa}$	1.034 (0.998, 1.071)	1.034 (0.996, 1.072)	0.890 (0.821, 0.964)
$\hat{\mu}_0$	7.096 (6.083, 8.139)	3.470 (2.770, 4.199)	4.729 (3.836, 5.666)
$\hat{\sigma}_0$	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)
Precip. intensity	-0.771 (-4.306, 3.188)	-1.912 (-3.924, 0.269)	-0.676 (-6.329, 6.297)
Precip. 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)

Abbreviations:

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

5. SIMULATION STUDY

5.1 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 for 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).

For each driver, the number of shifts is simulated from a Poisson distribution with a mean parameter of 10. We assume there are three predictor variables, x_1, x_2, x_3, which are simulated from: $x_1 \sim N(1, 1^2)$, $x_2 \sim \text{Gamma}(1, 1)$, and $x_3 \sim \text{Poisson}(2)$. The shift time $\tau_{d,s}$ is generated from $\tau_{d,s} \sim \text{Normal}(10, 1.3^2)$ to emulate the observed shift distribution.

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

$$
\begin{aligned}
\mu_0 &= 0.2,\ \sigma_0 = 0.5, \\
\gamma_{01}, \gamma_{02}, \ldots, \gamma_{0D} &\sim \text{i.i.d. } N(\mu_0, \sigma_0^2) \\
\gamma_1 &= 1,\ \gamma_2 = 0.3,\ \gamma_3 = 0.2 \\
\theta_{d,s} &= \exp(\gamma_{0d} + \gamma_1 x_1 + \gamma_2 x_2 + \gamma_3 x_3) \\
\beta &= 1.2,\ \kappa = 0.8.
\end{aligned} \tag{12}
$$

After the predictor variables, shift time, and parameters are generated, the time to events are generated from either the PLP or the JPLP.

The likelihood functions given in Equations (5) and (11) and the prior distributions are then used by `Stan` 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.

5.2 Simulation results

The simulation results are shown in Table 4. For the five sets of drivers $D = 10, 25, 50, 75, 100$ in each of the three scenarios, 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 bias seems converge to 0 as the number of drivers increases; the standard errors converge 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 is still a fair amount of bias when the number

Table 4: Biases Δ and standard errors (S.E.) for PLP, JPLP, and PLP ← JPLP simulations

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

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 the parameters of the simulated data sets.

6. DISCUSSION

6.1 Contributions to statistical modeling

In this article, we proposed a Bayesian hierarchical NHPP with PLP intensity function and a Bayesian hierarchical JPLP to model naturalistic truck driving data. Our motivation comes from the desire to determine the factors that affect the risk of SCEs, and therefore crashes. The proposed JPLP accounts for the characteristics of multiple rests within a shift among commercial truck drivers. A case study of 496 commercial truck drivers demonstrates a considerable amount of variability across drivers. Headway and hard brakes tend to occur in early stages of a shift, while collision mitigation tends to occur in later stages. A simulation study analyzes the Bayesian hierarchical estimation when the models are specified correctly or incorrectly.

The models we have studied are based on ones that have been widely applied to the reliability of repairable systems. The NHPP model implies that a minimal repair is done at each failure, i.e., the reliability of the system is restored to its condition immediately before the failure. For the case of repairable systems, the time required for repair is usually not included in the cumulative operating time. In our case, the NHPP implies that the occurrence of an SCE does not change the intensity of the process. There is no repair time to account for because the driver continues to drive immediately after the SCE. A rest break for drivers is analogous to a preventive maintenance for a repairable system, whereby a system's reliability is (possibly) improved by performing maintenance. Our JPLP model is similar to the modulated power law process (Lakey and Rigdon, 1993; Black and Rigdon, 1996), except their model assumed that the reliability can be improved at every repair.

Our models differ in several respects from the repairable systems models. Our models involve the use of covariates, such as weather conditions and driver demographics. In addition, the heterogeneity of drivers required a hierarchical model. In fact, one important finding is that driver-to-driver variability accounts for much of the variability of SCEs. Finally, the size of the data (496 drivers, with over 13 million pings) is much larger than would normally be encountered in a reliability setting.

6.2 Contributions to trucking safety research and practice

From our analysis, we obtained three novel and interesting results. First, based on our large NDS dataset, we showed that headway and hard brakes, which do not involve an automated intervention, tend to occur early in the shift. On the other hand, the collision mitigation system events are more likely toward the end of the shift. From a behavioral safety perspective, the implication of this result is two-fold: (a) drivers may exhibit a somewhat aggressive driving behavior early in their shift since they assume that they can accommodate for the increased risk with their attentiveness/alertness; and (b) slower reaction times and/or less alertness can be found later at the shift, which can potentially explain the observed increases in forward-collision mitigation system events. Note that this result could not be observed in the majority of past studies since classification approaches cannot account for the time to a SCE. From a practical perspective, this finding can be used to improve behavioral-based safety (BBS) and defensive driving training modules, which attempt to help drivers conceptualize the implications of risky driving decisions.

The second result was obtained by examining the reliability jump parameter κ, which makes the post-break intensity proportional to the pre-break intensity. The grouping of

SCEs was consistent with the first result, where the collision mitigation system had a $\hat{\kappa} < 1$ indicating that a rest break reduced the likelihood of a collision mitigation event. On the other hand, a rest break did not decrease the likelihood of the other two SCEs. Thus, our research provides evidence that breaks of at least 30 minutes can reduce the occurrence of the most severe SCEs, which have a stronger association with trucking crashes (Cai et al., 2020). On the other hand, such breaks may increase (or at least do not decrease) the other two SCEs, which may be explained by the justification provided for finding one. The hours of service regulations require a break after eight hours of continuous driving, but most of the observed breaks occurred well before this. Due to the effectiveness of these breaks in reducing automated interventions, we suggest that trucking operators should consider our finding in improving their dispatching and rest-break scheduling policies for trucking operators. Furthermore, our finding can be used to inform future improvements to the *hours of service* regulations.

Third, our hierarchical model showed that much of the variability in SCEs can be explained by the heterogeneity of the drivers. This result supports the need for a personalized modeling approach in modeling driver behavior. Furthermore, this finding is consistent with the conclusions obtained from occupational safety studies dedicated to manufacturing and warehousing tasks (Baghdadi et al., 2019; Maman et al., 2020).

6.3 Limitations and future work

Our work can be extended in several aspects in the future. First, the assumption of proportional intensity jump at rest breaks 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 simplify the parameterization and likelihood function. In truck transportation practice, longer breaks may have larger effects on reliability improvement and the effectiveness of rest-breaks may decrease over the course of the shift. The effect of breaks may also diminish across time, to the point where the termination of the shift is advisable. This corresponds to the situation of condition-based maintenance where the decision to perform maintenance (take a break) is dependent on the system's (driver's) current state. Eventually, though, the system must be replaced or overhauled (the driver must end the shift and sleep before returning to work.) It may also be the case that the effect of a rest break varies across drivers.

This manuscript sets the foundation for extending reliability and maintenance models for personalized human performance modeling. The hierarchical nature of our proposed approach accounts for the heterogeneity of human operators, and our models suggest there is a large amount of driver-to-driver variability. Moreover, the models support the use of covariates, which can accelerate/decelerate the degradation in an operator's performance in many occupational settings (Cavuoto and Megahed, 2016). The JPLP, which accounts for multiple driving segments with rest breaks, can be applied not only to commercial truck drivers, but also to other applications where a deterioration in human performance can occur (e.g., occupational fatigue management and neuromuscular disorders).

SUPPLEMENTARY MATERIALS

Because we are unable to make the driving data set publicly accessible, we provide instead a simulated dataset that is similar to the real data. This allows us to mask any company sensitive data, yet allows industrial and academic researchers to replicate our work. We limited the simulated dataset to a smaller number of drivers to ensure that the computations can be completed in a reasonable amount of time, without the need for high performance computing resources. The online supplementary materials contain the `R` code used to simulate PLP and JPLP data, explanations on the data structure as well as `Stan` and `R` code. The material is organized using an R Markdown document, which is hosted on the following `GitHub` page https://for-blind-external-review.github.io/JPLP/. The supplementary material will be moved to a permanent location after the peer-review process.

FUNDING

To ensure the anonymity of the author(s), we removed the agencies funding our work. This information would be provided after the peer-review process is completed.

REFERENCES

Baghdadi, A., Cavuoto, L. A., Jones-Farmer, A., Rigdon, S. E., Esfahani, E. T., and Megahed, F. M. (2019). Monitoring worker fatigue using wearable devices: A case study to detect changes in gait parameters. *Journal of Quality Technology*, pages 1–25.

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).

Cavuoto, L. and Megahed, F. (2016). Understanding fatigue and the implications for worker safety. In *ASSE Professional Development Conference and Exposition*. American Society of Safety Engineers.

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 (2020a). Naturalistic driving study (data analysis). United States Department of Transportation. https://www.fmcsa.dot.gov/research-and-analysis/research/naturalistic-driving-study-data-analysis. [Published March 10, 2020 and accessed July 03, 2020].

Federal Motor Carrier Safety Administration (2020b). 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.

Maman, Z. S., Chen, Y.-J., Baghdadi, A., Lombardo, S., Cavuoto, L. A., and Megahed, F. M. (2020). A data analytic framework for physical fatigue management using wearable sensors. *Expert Systems with Applications*, page 113405.

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 (2019). Traffic safety facts 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.

Sparrow, A. R., Mollicone, D. J., Kan, K., Bartels, R., Satterfield, B. C., Riedy, S. M., Unice, A., and Van Dongen, H. P. (2016). Naturalistic field study of the restart break in US commercial motor vehicle drivers: truck driving, sleep, and fatigue. *Accident Analysis & Prevention*, 93:55–64.

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.