Accident Analysis and Prevention 156 (2021) 106088

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Decision-adjusted driver risk predictive models using kinematics information

Huiying Mao [a,b], Feng Guo [a,b,*], Xinwei Deng [a], Zachary R. Doerzaph [b,c]

[a] Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
[b] Virginia Tech Transportation Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA
[c] Biomedical Engineering and Mechanics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

ARTICLE INFO

Keywords:
Automobile crash risk
Decision-adjusted modeling
Predictive modeling
Telematics data
Naturalistic driving study

ABSTRACT

Accurate prediction of driving risk is challenging due to the rarity of crashes and individual driver heterogeneity. One promising direction of tackling this challenge is to take advantage of telematics data, increasingly available from connected vehicle technology, to obtain dense risk predictors. In this work, we propose a decision-adjusted framework to develop optimal driver risk prediction models using telematics-based driving behavior information. We apply the proposed framework to identify the optimal threshold values for elevated longitudinal acceleration (ACC), deceleration (DEC), lateral acceleration (LAT), and other model parameters for predicting driver risk. The Second Strategic Highway Research Program (SHRP 2) naturalistic driving data were used with the decision rule of identifying the top 1% to 20% of the riskiest drivers. The results show that the decision-adjusted model improves prediction precision by 6.3% to 26.1% compared to a baseline model using non-telematics predictors. The proposed model is superior to models based on a receiver operating characteristic curve criterion, with 5.3% and 31.8% improvement in prediction precision. The results confirm that the optimal thresholds for ACC, DEC and LAT are sensitive to the decision rules, especially when predicting a small percentage of high-risk drivers. This study demonstrates the value of kinematic driving behavior in crash risk prediction and the necessity for a systematic approach for extracting prediction features. The proposed method can benefit broad applications, including fleet safety management, use-based insurance, driver behavior intervention, as well as connected-vehicle safety technology development.

1. Introduction

Predicting crash risk and identifying high-risk drivers are critical for developing appropriate safety countermeasures, driver education programs, and use-based insurance systems. Predicting driver-level risk is challenging due to the numerous factors contributing to individual crash risk coupled with the rarity of crash events and the substantial heterogeneity among drivers. The rapid advancements in in-vehicle data instrumentation and connected vehicle technology will make high-frequency driving data collection almost ubiquitously available and cost-effective. This high-frequency, high-resolution telematics driving information provides a unique opportunity to improve the state-of-the-practice driver risk prediction models. However, the sheer size of the telematics data and considerable data noise also bring technical challenges for predictive feature engineering and model development. Thus, there is a need for a comprehensive driver risk prediction framework to effectively use the kinematic driving data.

The crash risk of individual drivers varies substantially (Guo et al., 2013; Ulleberg, 2001; Habtemichael and de Picado-Santos, 2013). The majority of drivers are relatively safe, while a small percentage of drivers contribute to a disproportionate volume of the total crashes. Guo and Fang (2013) found that 6% of drivers account for 65% of crashes and near-crashes. Based on a simulation study, Habtemichael and de Picado-Santos (2013) showed that limiting the risk behavior of 4% to 12% of high-risk drivers would reduce crashes by 9% to 27% in different traffic conditions. Predicting and identifying this small portion of high-risk drivers can provide important information for developing targeted safety countermeasures to improve transportation safety.

The quality and quantity of predictors are critical for the performance of risk models. Traditional predictive features include driver demographics, personality factors, crash/citation/violation history, and observable risky driving behavior (Cantor et al., 2010; Ouimet et al.,

* Corresponding author at: Department of Statistics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA.
E-mail address: feng.guo@vt.edu (F. Guo).

https://doi.org/10.1016/j.aap.2021.106088
Received 28 December 2020; Received in revised form 25 February 2021; Accepted 9 March 2021
Available online 16 April 2021

2014; Soccolich et al., 2011; Habtemichael and de Picado-Santos, 2013; Wu et al., 2014). One limitation of these predictors is the limited resolution. For example, classification based on age and gender will lead to large cohorts of drivers, which makes it impossible to distinguish the small percentage of high-risk drivers within a age-gender cohort. Personality measures, such as the NEO-Five personality test, require drivers to take a questionnaire survey. The scalability of personality surveys is limited due to user consent issues and high survey costs. Information on past crash history has been shown to be a good crash indicator but suffers from a rarity of crashes and the regression-to-the-mean effect; i. e., a unit with a high number of crashes in the past, when observed in a future period, tends to have a lower number of crashes (AASHTO, 2010). Observable risky driving behavior, such as distraction, also indicates driver risk, but the identification procedure can be challenging (Sagberg et al., 2015; Sun et al., 2017; Yin et al., 2017).

The surging prevalence of connected vehicle technology, smartphone apps, and third-party *in-situ* telematics data collection devices provides a novel sources of risk predictors (Guo and Fang, 2013; Bagdadi and Várhelyi, 2011)?. Telematics data can include a variety of content, such as vehicle kinematics information, vehicle operation information, driver behavior information, as well as driving environment and traffic information. Abnormal kinematic events, e.g., large deceleration or swerving, could be caused by hazardous traffic conditions or aggressive driving behavior, including, for instance, hard braking, harsh acceleration, and sharp turning. Other risk factors, such as emotional states of drivers can also lead to abnormal kinematics signatures (Roidl et al., 2014).

It is broadly suggested in the literature that kinematic signatures are useful for crash risk prediction. High G-force events, i.e., when acceleration or passed a given thresholds is one of the most commonly used measure for predicting driver risk and identifying crashes Hankey et al. (2016), Bagdadi and Várhelyi (2013). The Teenage Naturalistic Driving Study consists a thorough evaluation of novice teenage drivers' high G-force events and demonstrated that its prediction power for crashes and near-crashes. (Simons-Morton et al., 2009, 2012, 2013). The association between rear-end crash rates and hard deceleration are confirmed using different data sources Kim et al. (2016), Palat et al. (2019), Zhu et al. (2017) also showed that rapid acceleration and deceleration are associated with increased risk with consideration for contextual factors. Beside high G-force event, longitudinal jerk, i.e., derivative of acceleration, especially large negative jerks have the potential to detect aggressive driving behaviors and predict crash risk (Feng et al., 2017; Bagdadi and Várhelyi, 2011; Bagdadi, 2013). Other kinematic measures such mean and volatility are also used in risk prediction (af Wåhlberg, 2006; Wang et al., 2015). Studies also demonstrated by setting the acceleration limit for buses can reduce the risk of crash injuries (Karekla et al., 2020).

Limited research has been conducted to systematically address how to properly use such information to maximize driver risk prediction performance. For example, there is no standard on what defines a risky kinematic signature. A common approach is to use triggers when deceleration surpasses a preset acceleration threshold (i.e., high G-force events). The threshold values used in the literature are not consistent and were usually selected based on researchers' subjective judgement. For example, Simons-Morton et al. (2012) defined "elevated gravitational event" as when acceleration exceeds $0.35g$, deceleration exceeds $0.45g$, or lateral acceleration exceeds $0.05g$. Klauer et al. (2009) explored the prevalence of events when acceleration or deceleration exceeds $0.30g$ among driver groups with different crash and near-crash rates. Limited research attempts to define an optimal threshold for the high G-force events.

Large-scale naturalistic driving studies (NDSs), such as the Second Strategic Highway Research Program (SHRP 2) NDS, provide opportunities to fully investigate optimal thresholds that provide the most prediction power for identifying high risk drivers. One commonly used approach to measure the performance of prediction models is based on the area under the curve (AUC) of the receiver operating characteristic (ROC) curve (Simons-Morton et al., 2012; Guo and Fang, 2013). However, AUC and other general model selection methods do not necessarily provide the optimal solution for a specific objective, such as identifying the top 1% of riskiest drivers. The AUC criterion is with respect to the entire spectrum of possible decision settings (resulting true positive rate and false positive rate range from zero to one). Subsequently, the model with the largest AUC is not necessarily optimal for a specific objective, such as the one noted above of identifying a small percentage of the riskiest drivers. A specific objective is often best served by a dedicated prediction model that is optimized with respect to the objective rather than by generic criteria.

In this paper, we propose a decision-adjusted modeling framework and develop an optimal driver risk prediction model based on driving telematics data from the SHRP 2 NDS. The optimal kinematic thresholds were identified by specific decision rule rather than subjective selection. Under this framework, model estimation will be conducted to optimize a decision-based model evaluation criterion. This framework is applied to develop the optimal driver risk prediction models under different decision rules. We also focus on incorporating a broader set of fused telematics data in the risk prediction models, in which the parameters to be adjusted are the thresholds of kinematic signatures, such as elevated longitudinal and lateral acceleration. The proposed model is compared with a traditional driver-characteristics-based model as well as a model optimized using the generic AUC criterion.

The remainder of the paper is organized as follows. Section 2 details the proposed decision-adjusted predictive modeling framework. Section 3 introduces SHRP 2 NDS data, followed by a formal definition and a full exploration of kinematic signatures using different threshold values. Section 4 quantifies the model improvement by comparing our proposed model to a model using traditional features and a model selected by AUC. We conclude this work with discussion in Section 5.

2. Decision-adjusted modeling framework

Traditional statistical model selection methods are based on certain statistical criteria, such as the likelihood ratio test or the ROC curve. The resultant model may not be optimal for a specific decision goal. This study focuses on overcoming this limitation by proposing a decision-adjusted modeling framework that directly formulates the study's specific goal through a decision-based objective function in the model selection/optimization process. The model selection/optimization process involves model form determination, variable selection, and parameter tuning. The model form depends on the response variable type. For binary response data, potential models include logistic regression, decision tree, neural network, etc. Variable selection determines which covariates should be included. The parameter tuning refers to the hyperparameter tuning for the selected model form and to the critical value adjustment in building certain predictor variables. For example, in predicting driver risk using kinematic signatures, the critical value adjustment is with respect to the threshold values for defining a kinematic signature that will maximize the prediction power for a pre-specific percentage of high risk drivers.

2.1. Decision-adjusted driver risk prediction

The objective of risk prediction models could vary substantially according to specific application, which leads to varying decision criteria. For example, for fleet safety management, it is common that only a certain number of drivers can be coached or provided with advanced active safety systems due to limited resources. The specific study goal in this context will be to identify a targeted number of the riskiest drivers for receiving these safety enhancing approaches/technologies, which will allow fleets to maximize improvements with the limited resources. The decision-adjusted framework will provide better support for the specific objective than models based on generic criteria.

The structure of the proposed decision-adjusted modeling is illustrated in Fig. 1. The prediction model $\mathcal{M}$ can come from a variety of models, such as logistic regression, gradient boosting tree, support vector machine, or neural network. Data $\boldsymbol{X}$ is the matrix of predictors; Y is the observed response variable; e.g., an indicator variable of whether driver i experienced crashes or not. The $\boldsymbol{\delta}$ is the parameter that defines the predictive feature $\boldsymbol{R}$; $\boldsymbol{Z}$ is the predicted value, which depends on the decision parameter τ. η is the objective function to be optimized, typically a distance measure between predicted and observed response ($\boldsymbol{Z}$ and $\boldsymbol{Y}$). The optimization is achieved by adjusting tuning parameters σ and the predictive model $\mathcal{M}$.

The decision-adjusted framework is a flexible modeling approach that can accommodate a wide variety of decision rules and tuning parameters. We apply this framework to predict a preset percentage of highest risk drivers and determine the optimal threshold values for high G-force events. Specifically, let Y_i be an indicator variable of whether driver i was involved in any crashes, $i = 1, 2, \ldots, n$,

$$Y_i = \begin{cases} 1, & \text{driver} i \text{involvedinatleastonecrash}, \\ 0, & \text{driver} i \text{notinvolvedinanycrashes}; \end{cases}$$

$\boldsymbol{X}_i$ is a vector of traditional explanatory variables, such as age, gender, and other characteristics.

The kinematic driving information is incorporated in the driver risk prediction model in the form of high G-force event rates. The high G-force events are derived from kinematic signatures when acceleration (ACC), deceleration (DEC), or lateral acceleration (LAT) exceed certain thresholds. Determining the optimal threshold values for kinematic variables is a crucial component of the model selection/optimization process. $\boldsymbol{R}_i(\boldsymbol{\delta})$ is a vector of high G-force event rates, which depends on threshold value $\boldsymbol{\delta} = (\delta_{\mathrm{ACC}}, \delta_{\mathrm{DEC}}, \delta_{\mathrm{LAT}})$. That is,

$$\boldsymbol{R}_i(\boldsymbol{\delta}) = (\mathrm{RA}_i(\delta_{\mathrm{ACC}}), \mathrm{RD}_i(\delta_{\mathrm{DEC}}), \mathrm{RL}_i(\delta_{\mathrm{LAT}})).$$

With binary response, the model $\mathcal{M}$ (e.g., logistic regression) predicts the probability of a driver i being involved in a crash, $Pr(Y_i = 1)$, with X and kinematic metrics R($\boldsymbol{\delta}$). The predicted label for this driver is as follows:

$$Z_i = Z_i(\mathrm{X}, \mathrm{R}(\boldsymbol{\delta}), \mathcal{M}, \tau) = \begin{cases} 1, & \text{if} \hat{Pr}(Y_i = 1) > \tau, \\ 0, & \text{if} \hat{Pr}(Y_i = 1) \leq \tau, \end{cases}$$

where τ is a cutoff value usually determined by the specific study objective. For example, in the application of providing safety improvement education for a targeted number of the highest risk drivers, τ will be selected such that the number of drivers with label $Z = 1$ satisfies the requirement of the decision of interest.

With the above setup, the decision-adjusted modeling approach can be described as an optimization problem.

$$\max_{\boldsymbol{\delta}, \mathcal{M}} \ (\text{or} \min) \quad \eta(\boldsymbol{Z}(\mathrm{X}, \mathrm{R}(\boldsymbol{\delta}), \mathcal{M}, \tau), \boldsymbol{Y}), \tag{1}$$

where $\eta(\cdot)$ is the decision-adjusted objective function, which is a function of $\boldsymbol{Z} = (Z_1, Z_2, \cdots, Z_n)$ and $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)$. If the specific goal is to identify the top 10% of riskiest drivers, the objective function would be to maximize the prediction precision, the percentage of correct identification among those 10% drivers with the predicted label $Z = 1$. That is,

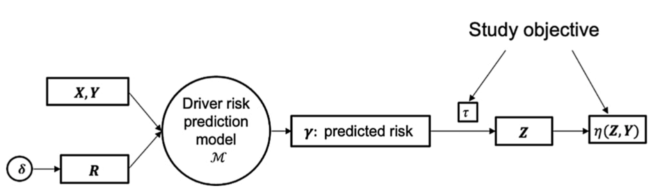

Fig. 1. The structure of decision-adjusted risk prediction model.

$$\max \quad \eta(\boldsymbol{Z}, \boldsymbol{Y}) = \frac{\boldsymbol{Z}'\boldsymbol{Y}}{\boldsymbol{1}'\boldsymbol{Z}} = \frac{\sum_{i=1}^{n} Y_i \cdot Z_i}{\sum_{i=1}^{n} Z_i},$$

such that

$$\boldsymbol{1}'\boldsymbol{Z} = \sum_{i=1}^{n} Z_i = 10\% \cdot n.$$

Similarly, for the example of identifying at least 50% of the drivers who would have crashes, the decision-adjusted objective function is

$$\min \quad \eta(\boldsymbol{Z}, \boldsymbol{Y}) = \boldsymbol{1}'\boldsymbol{Z} = \sum_{i=1}^{n} Z_i,$$

such that

$$\frac{\boldsymbol{Z}'\boldsymbol{Y}}{\boldsymbol{1}'\boldsymbol{Y}} = \frac{\sum_{i=1}^{n} Y_i \cdot Z_i}{\sum_{i=1}^{n} Y_i} \geq 50\%.$$

As to the last example of minimizing the average cost of running a driver education program, the objective becomes minimizing the expected average cost, which can be expressed as

$$\min \quad \eta(\boldsymbol{Z}, \boldsymbol{Y}) = \frac{1}{n}[\boldsymbol{Y}'(\boldsymbol{I} - \boldsymbol{Z}) \cdot C(-|+) + \boldsymbol{Z}'(\boldsymbol{I} - \boldsymbol{Y}) \cdot C(+|-)]$$

where $C(-|+)$ is the cost of applying safety countermeasures to a driver who would have no crashes, and $C(+|-)$ is the cost of crashes for a driver who is labeled as low-risk. There would be no cost for correctly identified drivers (i.e., $Z_i = Y_i$).

Optimizing the above driver risk prediction model will provide not only the maximum predictive power according to the modeling objective but also the optimal high G-force threshold.

3. Application

We applied the proposed method to the SHRP 2 NDS data, with the primary objective of identifying a targeted percentage of "high-risk" drivers. The corresponding decision-adjusted objective function is to maximize the prediction precision – i.e., the percentage of correctly identified high-risk drivers. Model performance was evaluated under the study objective of identifying a certain percentage of the riskiest drivers, and we elaborated the prediction precision and the optimal threshold values of high G-force event rates for the target percentage of high-risk drivers from 1% to 20%.

3.1. The SHRP 2 NDS data

The SHRP 2 NDS was a large-scale observational study with more than 3500 drivers from six data collection sites in Florida, Indiana, North Carolina, New York, Pennsylvania, and Washington. The study collected a wide range of information for drivers under natural driving conditions. A data acquisition system, including radar, multiple cameras, three dimensional accelerometers, and other equipment, was installed in each participant's vehicle (Campbell, 2012). The driving data were collected continuously from ignition-on to ignition-off. The data were collected asynchronously; for example, videos at 10 Hz, GPS at 1 Hz, and acceleration at 10 Hz. The SHRP 2 NDS provides a great opportunity to address questions about driver performance and behaviour as well as traffic safety (Dingus et al., 2016; Guo, 2019; Guo et al., 2017).

The crashes were identified thorough a comprehensive process. The kinematic time series data were screened through an automated algorithm. The driving segments that were potentially crashes were manually examined via the videos to confirm whether a crash had actually occurred (Hankey et al., 2016). This study includes 1149 crashes of severity Levels 1–3, where level 1 are the most severe crashes involving airbag deployment or potential injury, and level 3 are minor crashes in which the vehicle made physical contact with another object or departed the road.

The prediction models include 3440 drivers, among whom 810 (23.5%) experienced at least one crash during the study period. The study also collected demographic and personality information, such as age, gender, and sleep habits, at the beginning of the study. Table 1 shows the descriptive statistics for the traditional predictors used in our study. Personality factors come from survey data. For example, the driving knowledge survey is a questionnaire compiled from a number of Department of Motor Vehicle (DMV) driving knowledge tests, and the score represents the number of questions answered correctly (out of 19). The clock drawing assessment is used as a screening tool to help identify possible signs of dementia or other neurological disorders. Mileage history, number of violations/crashes, insurance status, and specific driving behavior are self-reported. The number of violations/crashes refers to the past 3 years, and the participants reported their specific driving behavior history for the past 12 months.

The longitudinal and lateral acceleration are measured by the gravity of Earth g; ($1g = 9.8m/s^2$). Let a_x and a_y denote the longitudinal (x-axis) acceleration and lateral (y-axis) acceleration while driving. Positive a_x indicates acceleration and negative indicates deceleration. Positive and negative a_y mean lateral acceleration to the left and right. A high ACC G-force event occurred if $a_x > \delta_{\mathrm{ACC}}$; a high DEC G-force event occurred if $a_x < -\delta_{\mathrm{DEC}}$; and a high LAT G-force event occurred if $|a_y| > \delta_{\mathrm{LAT}}$. The corresponding event rates per driving hour for driver i, RA_i,RD_i,RL_i, are calculated as

$$\begin{aligned}
\mathrm{RA}_i(\delta_{\mathrm{ACC}}) &= \frac{1}{T_i}\sum_{\text{driver } i} 1(a_x > \delta_{\mathrm{ACC}});\\
\mathrm{RD}_i(\delta_{\mathrm{DEC}}) &= \frac{1}{T_i}\sum_{\text{driver } i} 1(a_x < -\delta_{\mathrm{DEC}});\\
\mathrm{RL}_i(\delta_{\mathrm{LAT}}) &= \frac{1}{T_i}\sum_{\text{driver } i} 1(|a_y| > \delta_{\mathrm{LAT}}),
\end{aligned}$$

where $\sum_{\text{driver } i} 1(\cdot)$ represents the number of high G-force events and T_i is the total driving time for driver i. For an actual hard brake or other maneuver, there is typically a sequence of multiple data points beyond the threshold value. To identify the number of events rather than the number of data points, a moving average filter was applied to the raw data and several criteria were used to cluster data points from a potential high G-force event into one event, including criteria that the data points should be close to each other temporally and the smoothed data should be a local maximum pattern.

The high G-force events were identified from a total of 1,161,112 driving hours with a minimum threshold of $0.3g$ to ensure the majority of evasive maneuvers were included. Fig. 2 shows the number of ACC, DEC, and LAT G-force events by different thresholds. The number of high G-force events varies substantially with respect to different thresholds, from millions at $0.3g$ to a few thousand at $0.65g$.

Setting the thresholds too low or too high could result in non-informative explanatory variables in the driver risk prediction model. The G-force events based on low thresholds are likely to be dominated by normal driving behaviors, such as stopping before traffic signals, and may mask true risky behaviors. On the other hand, if the thresholds are set too high, high G-force events themselves become rare and provide little information about crash risk. Fig. 3 shows the Spearman's rank-order correlation for the occurrence of a crash and three types of high G-force event rates under different threshold values. All three kinematic event types show a clear pattern of initial increase followed by decrease in terms of correlation with crashes, which is consistent with the above postulation. There should be an optimal threshold value to best represent the risk behavior. The proposed decision-adjusted driver risk prediction framework provides an ideal solution for this optimization problem.

Table 1
Description of traditional predictors.

Covariates	Descriptive statistics
Demographic features	
Gender	F: 1796(52.21%); M: 1644(47.79%)
Age group	16–19: 534(15.52%); 20–24: 741(21.54%); 25–29: 280 (8.14%); 30–34: 164(4.77%); 35–39: 129(3.75%); 40–44: 117(3.4%); 45–49: 149(4.33%); 50–54: 165(4.8%); 55–59: 144(4.19%); 60–64: 149(4.33%); 65–69: 209 (6.08%); 70–74: 173(5.03%); 75–79: 262(7.62%); 80–84: 155(4.51%); 85–89: 59(1.72%); 90–94: 8(0.23%); 95–99: 2(0.06%)
Marital status	Single: 1514(44.01%); Married: 1384(40.23%); Divorced: 199(5.78%); Widow(er): 205(5.96%); Unmarried partners: 111(3.23%); NA: 27(0.78%)
Education level	Some high school: 272(7.91%); High school diploma or G.E.D.: 321(9.33%); Some education beyond high school but no degree: 980(28.49%); College degree: 913(26.54%); Some graduate or professional school, but no advanced degree (e.g., J.D.S., M.S. or Ph.D.): 352(10.23%); Advanced degree (e.g., J.D.S., M.S. or Ph.D.): 584 (16.98%); NA: 18(0.52%)
Income level	Under $29K: 576(16.74%); $30K to $40K: 396(11.51%); $40K to $50K: 321(9.33%); $50K to $70K: 572(16.63%); $70K to $100K: 611(17.76%);$100K to $150K: 502 (14.59%); $150K +: 245(7.12%); NA: 217(6.31%)
Work status	Not working outside the home: 1237(35.96%); Part-time: 945(27.47%); Full-time: 1214(35.29%); NA: 44(1.28%)
Having children or not	No: 2423(70.44%); Yes: 681(19.8%); NA: 336(9.77%)
Sleep duration	Sufficient: 1730(50.29%); Slightly insufficient: 1130 (32.85%); Markedly insufficient: 212(6.16%); Very insufficient or did not sleep at all: 17(0.49%); NA: 351 (10.2%)
Data collection site	FL: 768(22.33%); IN: 275(7.99%); NC: 558(16.22%); NY: 772(22.44%); PA: 263(7.65%); WA: 804(23.37%)
Personality factors	
Driving knowledge score	mean: 15.1; std. dev: 2.0; NA: 351(10.2%)
Clock drawing score	mean: 2.1; std. dev: 1.0; NA: 259(7.53%)
Barkley's ADHD Score	mean: 3.2; std. dev: 2.2; NA: 234(6.8%)
Sensation seeking score	mean: 14.6; std. dev: 6.9; NA: 242(7.03%)
Driving behavior	
Driving hours in the study	mean: 337.7; std. dev: 243.9
Mileage last year (mile)	mean: 12117.1; std. dev: 9530.8; NA: 251(7.3%)
Annual mileage	10K–15K miles: 1078(31.34%); 15K–20K miles: 458 (13.31%); 20K–25K miles: 219(6.37%); 25K–30K miles: 121(3.52%); 5K–10K miles: 901(26.19%); less than 5K miles: 412(11.98%); more than 30K miles: 194(5.64%); NA: 57(1.66%)
Driving experience (years)	mean: 25.6; std. dev: 22.5; NA: 23(0.67%)
Number of violations	0: 2452(71.28%); 1: 664(19.3%); 2 or More: 305(8.87%); NA: 19(0.55%)
Number of crashes	0: 2548(74.07%); 1: 667(19.39%); 2 or More: 197(5.73%); NA: 28(0.81%)
Number of crashes at fault	0: 458(13.31%); 1: 343(9.97%); 1 or More: 50(1.45%); NA: 2589(75.26%)
Insurance status	No: 47(1.37%); Yes: 3342(97.15%); NA: 51(1.48%)
Run red lights past 12 mo	Never: 2062(59.94%); Rarely: 1030(29.94%); Sometimes: 84(2.44%); Often: 7(0.2%); NA: 257(7.47%)
Drive sleepy past 12 mo	Never: 1479(42.99%); Rarely: 1409(40.96%); Sometimes: 279(8.11%); Often: 15(0.44%); NA: 258(7.5%)
Impatiently pass on right	Never: 824(23.95%); Hardly Ever: 1022(29.71%); Occasionally: 1045(30.38%); Quite Often: 194(5.64%); Frequently: 74(2.15%); Nearly All the Time: 21(0.61%); NA: 260(7.56%)
Brake aggressively	Never: 1926(55.99%); Hardly Ever: 1114(32.38%); Occasionally: 109(3.17%); Quite Often: 5(0.15%); NA: 286(8.31%)
Involved in racing	

(*continued on next page*)

Table 1 (*continued*)

Covariates	Descriptive statistics
Racing frequency	Never: 2963(86.21%); Often: 6(0.17%); Rarely: 167 (4.86%); Sometimes: 28(0.81%); NA: 273(7.94%) Frequently: 3(0.09%); Hardly Ever: 214(6.22%); Nearly All the Time: 1(0.03%); Never: 2904(84.42%); Occasionally: 41(1.19%); Quite Often: 0(0%); NA: 277 (8.05%)
Nod off while driving	No: 2570(74.71%); Yes: 503(14.62%); NA: 367(10.67%)
Nod off frequency	1–2 times per month: 41(1.19%); 1–2 times per week: 9 (0.26%); 3–4 times per week: 2(0.06%); Nearly every day: 1(0.03%); Never: 10(0.29%); Rarely: 426(12.38%); NA: 2951(85.78%)

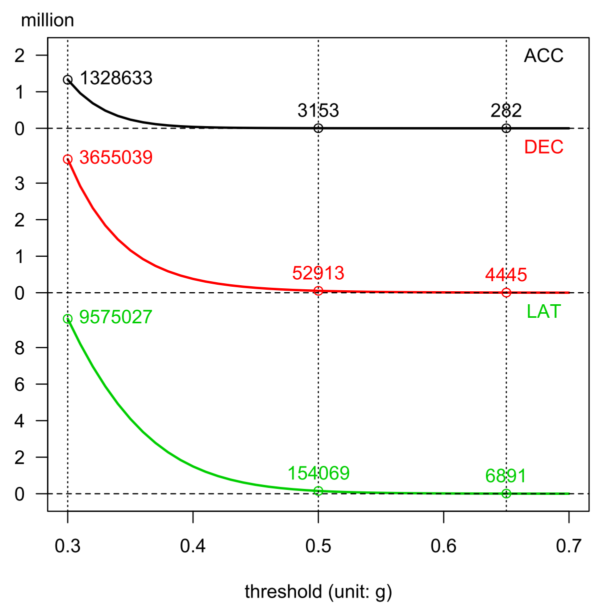

Fig. 2. Number of high G-force events (ACC, DEC, LAT) versus thresholds.

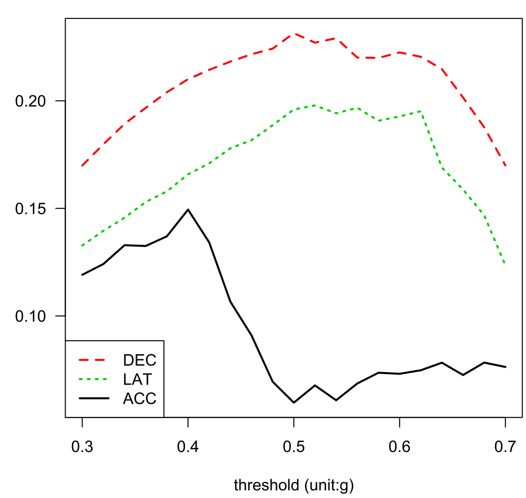

Fig. 3. Spearman's rank-order correlation between high G-force event rates and crash occurrence at the driver level.

3.2. *Decision-adjusted driver risk prediction through regularized logistic regression*

A variety of models, such as generalized linear, gradient boosting tree, and neural network, can be applied to driver risk prediction. In this study, we adopt a regularized logistic regression, the elastic net model, for model development and performance evaluation (Zou and Hastie, 2005; Friedman et al., 2010). This class of models controls model complexity through L_1 and L_2 penalties for regression coefficients. It addresses the multicolinality issue and can automatically select significant factors with proper model setup. Specifically, the risk for driver i with predictors $\boldsymbol{X}_i$ and kinematic metrics $\boldsymbol{R}_i(\boldsymbol{\delta})$ is formulated as

$$Pr(Y_i = 1|\boldsymbol{X}_i, \boldsymbol{R}_i(\boldsymbol{\delta})) = \frac{exp((\boldsymbol{X}_i, \boldsymbol{R}_i(\boldsymbol{\delta}))^T\boldsymbol{\beta})}{1 + exp((\boldsymbol{X}_i, \boldsymbol{R}_i(\boldsymbol{\delta}))^T\boldsymbol{\beta})}, \quad (2)$$

where $\boldsymbol{\beta}$ is the vector of regression coefficient. The model fitting is accomplished by maximizing a penalized log-likelihood function

$$\ell(\boldsymbol{\beta}) - \lambda[(1-\alpha)\frac{1}{2}||\boldsymbol{\beta}||_2^2 + \alpha||\boldsymbol{\beta}||_1], \quad (3)$$

where $\ell(\boldsymbol{\beta})$ is the log-likelihood function; and $\lambda[(1-\alpha)\frac{1}{2}||\boldsymbol{\beta}||_2^2 + \alpha||\boldsymbol{\beta}||_1]$ is the regularization part with $||\cdot||_2$ being the L_2 norm and $||\cdot||_1$ being the L_1 norm. The two hyperparameters, $0 \le \alpha \le 1$ and $\lambda \ge 0$, regularize the complexity of the logistic regression model. The LASSO regression and ridge regression are two special cases of elastic net models when $\alpha = 1$ and $\alpha = 0$ respectively.

The decision-adjusted approach optimizes the following objective functions with respect to the threshold values $\boldsymbol{\delta}$ as well as hyperparameters λ and α. For ROC-based model optimization, the η is the AUC of the ROC curve. For identifying a given percentage of the highest risk drivers, the η is the predictive precision, which measures the percentage of correct identification among drivers who were labeled as high-risk; $\frac{\boldsymbol{Z}'\boldsymbol{Y}}{\boldsymbol{1}'\boldsymbol{Z}} = \frac{\sum_{i=1}^{n} Y_i \cdot Z_i}{\sum_{i=1}^{n} Z_i}$, where $\boldsymbol{Z}$ is the prediction and $\boldsymbol{Y}$ is the observed true classification. Higher prediction precision indicates better prediction performance.

$$\max_{\boldsymbol{\delta},\lambda,\alpha} \quad \eta(\boldsymbol{Z}(\boldsymbol{X}, \boldsymbol{R}(\boldsymbol{\delta}), \widehat{\boldsymbol{\beta}}(\lambda, \alpha), \tau), \boldsymbol{Y}). \quad (4)$$

The traditional driver risk prediction model without telematics variables can be considered as a limiting case of the decision-adjusted model. As the threshold values $\delta_{\text{ACC}}, \delta_{\text{DEC}}, \delta_{\text{LAT}}$ approach infinity, all G-force event counts will converge to zero—$\boldsymbol{R}_i(\boldsymbol{\delta}) = \boldsymbol{0}, i = 1, 2, \ldots, n$—and the event rates become non-informative. The decision-adjusted driver risk prediction model degenerates to traditional models.

We compared three prediction models: $\mathcal{M}_0$, $\mathcal{M}_1$, and $\mathcal{M}_2$.

- $\mathcal{M}_0$: traditional driver risk prediction without high G-force events
- $\mathcal{M}_1$: driver risk prediction with high G-force event rates, optimized by AUC
- $\mathcal{M}_2$: decision-adjusted driver risk prediction for preset portion of highest risk drivers

The predictors and modeling strategy for the three models are listed

Table 2
Description of three comparison methods.

	Predictors	Modeling Strategy	Optimization Criterion
$\mathcal{M}_0$	Traditional predictors[a]	Elastic net	None
$\mathcal{M}_1$	Traditional covariates[a] + ACC, DEC, and LAT rates	Elastic net	AUC of ROC curve
$\mathcal{M}_2$	Traditional covariates[a] + ACC, DEC, and LAT rates	Elastic net	Predictive Precision of high risky drivers

[a] The traditional covariates are listed in Table 1.

in Table 2. The difference between $\mathcal{M}_0$ and $\mathcal{M}_1$ is that $\mathcal{M}_1$ also includes the occurrence rates of ACC, DEC, and LAT as predictors, along with other traditional covariates listed in Table 1. $\mathcal{M}_1$ selects the threshold values for ACC, DEC, and LAT by maximizing the AUC of the prediction model. Its α is set to be 1 (lasso), and its λ is chosen to minimize the 10-fold cross-validation error ($\lambda = 0.0074$). $\mathcal{M}_2$ uses the same set of predictors as $\mathcal{M}_1$. $\mathcal{M}_1$ and $\mathcal{M}_2$ differ in their modeling strategies. $\mathcal{M}_2$ tunes its threshold values and model hyperparameters to optimize the specific study objective compared to the AUC by $\mathcal{M}_1$.

We conducted a fine grid search of the parameter space to optimize $\mathcal{M}_1$ and $\mathcal{M}_2$. The candidate threshold values for $\delta_{\text{ACC}}, \delta_{\text{DEC}}, \delta_{\text{LAT}}$ range from $0.30g$ to $0.70g$, by $0.02g$ in equal steps. For the hyperparameters, $log(\lambda)$ ranges from -8.5 to -2.5, by 0.4 in equal space, and α ranges from 0 to 1, by 0.2 in equal steps. The coefficients for Model $\mathcal{M}_1$ are listed in Table 3, ordered by the magnitude. All predictors are standardized so the magnitude can be used to compare the relative impacts of risk factor. Coefficients for non-listed variables are shrunk to zero, and therefore the variables listed in the table are those with significant impacts on crash risk. The optimization process shows an optimal threshold for DEC rate of $\delta_{\text{DEC}} = 0.46g$, a LAT rate of $\delta_{\text{LAT}} = 0.50g$, and 39 other variables. R package 'glmnet' is used for the implementation of the elastic net (Friedman et al., 2009).

Table 3
The non-zero coefficients for model $\mathcal{M}_1$.

Variable	Coefficient
DEC rate when $\delta_{DEC} = 0.46g$	2.18
Driving hours in the study	2.17
LAT rate when $\delta_{LAT} = 0.50g$	0.74
Nod off while driving frequency. 1–2 times per week	0.48
Driving knowledge survey score	− 0.37
Racing frequency. Nearly all the time	0.36
Education level. NA	− 0.35
Nod off while driving frequency. Nearly every day	0.33
ACC rate when $\delta_{ACC} = 0.30g$	0.32
Age group. 80–84	0.27
Number of violations. 2 or more	0.26
Education level. Some high school	0.24
Sensation seeking scale survey score	0.22
Number of crashes at fault. 1 or more	0.19
Run red lights or not. NA	− 0.18
Drive sleepy or not. Often	0.12
Age group. 45–49	− 0.08
Education level. Some graduate or professional school, but no advanced degree	− 0.08
Marital status. Single	0.08
Insurance status.Yes	− 0.08
Barkley's ADHD Score	0.07
Brake aggressively. Occasionally	0.06
Work status. NA	− 0.06
Annual mileage. less than 5K miles	0.06
Nod off while driving frequency. 1–2 times per month	− 0.05
Annual mileage. 15K–20K miles	− 0.05
Having children at home or not. Yes	− 0.05
Age group.40–44	− 0.05
Work status. Part-time	0.05
Mileage last year	− 0.04
Education level. Some education beyond high school but no degree	0.04
Impatiently pass on the right. Never	0.04
Number of crashes at fault. NA	− 0.03
Data collection site. WA	− 0.03
Data collection site. PA	− 0.03
Racing frequency. Rarely	0.03
Nod off while driving frequency. Rarely	0.03
Income level.Under $29,000	0.02
Education level. College degree	− 0.02
Drive sleepy or not. Rarely	0.02
Run red lights or not. Rarely	0.01
Data collection site. NC	− 0.01

3.3. Prediction performance comparison

The performance of models are evaluated by the prediction precision as discussed in Eq. (4). Fig. 4 shows the relative improvement of prediction precision for $\mathcal{M}_1$ and $\mathcal{M}_2$ compared to $\mathcal{M}_0$ when the targeted percentage of high-risk drivers is between 1% and 20%. Table 4 lists the corresponding number of true positives and prediction precision for the three models. The relative improvement of prediction precision for $\mathcal{M}$ compared to $\mathcal{M}_0$ is defined as $\frac{\text{precision}(\mathcal{M})-\text{precision}(\mathcal{M}_0)}{\text{precision}(\mathcal{M}_0)}$.

The proposed method $\mathcal{M}_2$ performs the best among the three alternative models. The decision-adjusted model improves the prediction precision by 6.3% to 26.1% compared to the baseline model $\mathcal{M}_0$. It is also superior to $\mathcal{M}_1$ by 5.3% to 31.8%. The improvement is more prominent when the objective is to identify a small percentage of the riskiest drivers (e.g., $<$ 5%). $\mathcal{M}_1$ is better than $\mathcal{M}_0$ when the target percentage of high-risk drivers is greater than 4%. The benefit can be credited to the inclusion of high G-force event rates. The results confirm that using kinematic information can improve individual driver risk prediction, and the improvement is more significant when a decision-adjusted modeling approach is applied. When the target percentage is small ($<$ 4%), the prediction performance of $\mathcal{M}_1$ is worse than $\mathcal{M}_0$. This is because the model tuned for overall performance sacrifices the boundary cases. This illustrates the necessity of decision-adjusted modeling if the boundary cases are of main interest. Furthermore, the substantial improvement of $\mathcal{M}_2$ indicates that the decision-adjusted modeling is more powerful for highly imbalanced data scenarios—e.g., predicting a small percentage of high-risk drivers.

3.4. Optimal thresholds for high G-force events

The threshold values for ACC, DEC, and LAT have a substantial influence on the association between high G-force event rates and crash occurrence. Both $\mathcal{M}_1$ and $\mathcal{M}_2$ invovle optimization, with no threshold needed for the traditional predictor only model ($\mathcal{M}_0$).

Model $\mathcal{M}_1$ incorporates high G-force event rates and chooses the threshold values of high G-force event rates by maximizing the AUC of the ROC curve. The optimal threshold values for predicting high risk drivers are:

$$\boldsymbol{\delta} = (\delta_{\text{ACC}}, \delta_{\text{DEC}}, \delta_{\text{LAT}}) = (0.30g, 0.46g, 0.50g).$$

The profiles of the AUC values by threshold settings are shown in Fig. 5.

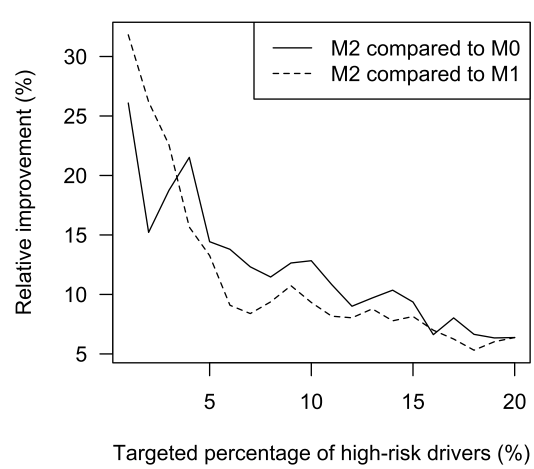

Fig. 4. Relative improvement of prediction precision of the three models.

Table 4
Prediction performance comparison of the three models.

Risky pct	Number of true positives			Prediction precision[a]		
	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_2$	$\mathcal{M}_0$	$\mathcal{M}_1$	$\mathcal{M}_2$
1%	23	22	29	67.6%	64.7%	85.3%
2%	46	42	53	66.7%	60.9%	76.8%
3%	64	62	76	62.1%	60.2%	73.8%
4%	79	83	96	57.7%	60.6%	70.1%
5%	97	98	111	56.4%	57%	64.5%
6%	116	121	132	56.3%	58.7%	64.1%
7%	138	143	155	57.3%	59.3%	64.3%
8%	157	160	175	57.1%	58.2%	63.6%
9%	174	177	196	56.3%	57.3%	63.4%
10%	187	193	211	54.4%	56.1%	61.3%
11%	203	208	225	53.7%	55%	59.5%
12%	222	224	242	53.9%	54.4%	58.7%
13%	237	239	260	53%	53.5%	58.2%
14%	251	257	277	52.2%	53.4%	57.6%
15%	267	270	292	51.7%	52.3%	56.6%
16%	287	286	306	52.2%	52%	55.6%
17%	299	304	323	51.2%	52.1%	55.3%
18%	316	320	337	51.1%	51.7%	54.4%
19%	331	332	352	50.7%	50.8%	53.9%
20%	345	345	367	50.2%	50.2%	53.4%

[a] $\text{Predictionprecision} = \frac{\text{numberoftruepositives}}{\text{numberofpredictedpositives}}$

The green area corresponds to a relatively small AUC value and the red to white area is for a relatively large AUC value. The two black dots on the heat map represent the optimal threshold setting $(0.30g, 0.46g, 0.50g)$, which generates a driver risk prediction model with an AUC value of 0.752. The color heterogeneity along the x-axis on the left heat map indicates that the resulting AUC value is sensitive to the selection of δ_{DEC}. On the right heat map, the major colors are yellow and red, which implies that given $\delta_{\text{DEC}} = 0.46g$, the driver risk prediction models are robust for ACC and LAT thresholds. For comparison, the AUC of the traditional driver risk prediction model, $\mathcal{M}_0$, is 0.742.

For the decision-adjusted model $\mathcal{M}_2$, the selected optimal thresholds vary according to the different decision rules under consideration. For example, when 20% is the target percentage of high-risk drivers, the optimal threshold setting is

$$\boldsymbol{\delta} = (0.62g, 0.48g, 0.60g), \quad log(\lambda) = -6.9, \quad \alpha = 0.6.$$

Under such a setting, $\mathcal{M}_2$ can correctly identify 367 high-risk drivers out of the 688. When the preset target percentage is 10%, $\mathcal{M}_2$ performs the best at a different parameter combination:

$$\boldsymbol{\delta} = (0.64g, 0.70g, 0.30g), \quad log(\lambda) = -8.1, \quad \alpha = 0.0.$$

Fig. 6 shows the point-and-whisker-plot for $\delta_{\text{ACC}}, \delta_{\text{DEC}}$, and δ_{LAT} in the "top three" optimal settings when the target percentage of high-risk driver ranges from 1% to 20%. The connected points represent the mean of the "top three" optimal settings, and the length of the half whiskers represents their one standard deviation. Take the target percentage of 20% as an example: the optimal model can identify 367 high-risk drivers correctly (Table 4), and we plot the mean and $(+/-)$ one standard deviation for the settings that can correctly identify 367, 366, and 365 high-risk drivers. The "top three" optimal settings are considered to ensure the robustness of the presented settings. The optimal thresholds for δ_{ACC}, δ_{DEC}, and δ_{LAT} are calculated and presented separately.

The pattern differs for the three types of high G-force events. The means the "top three" optimal δ_{ACC}s are about the same for different decision rules, and their whiskers are relatively long compared to the "top three" optimal δ_{DEC}s and δ_{LAT}s. This phenomenon indicates that the selection of δ_{ACC} is trivial for the driver risk prediction model. The prediction performance is insensitive to the thresholds for δ_{ACC}.

4. Summary and discussion

Accurate assessment and prediction of driving risk is of great importance for developing safety countermeasures and fleet safety management. In practice, the objectives of the prediction models typically vary by specific applications. The traditional risk assessment treats model development and decision making as two separate modules, which often lead to sub-optimal performance. The decision-adjusted modeling framework proposed in this paper combines decision objectives and modeling procedure in an integrated framework which leads to optimal models for the specific research application and objectives. The approach is especially beneficial in severely imbalanced data scenarios, i.e., the number of observations from one category dominates the others for the model response. In the application of identifying a small percentage of highest risk drivers, the decision-adjust modeling framework demonstrates superior performance over general model selection criteria, such as the AUC of ROC curve.

The high frequency, high resolution kinematic driving data provide a novel data source for driving risk assessment. The high G-force events have been widely used in insurance industry and driver risk prediction. The kinematics driving data reflect many factors related to safety, e.g., traffic condition and risky driving behavior. Subsequently, they can serve as a risk predictor for future crashes or crash surrogates for safety assessment as crashes are rare events Simons-Morton et al. (2013). The kinematic events can be detected automatically through continuous driving data thus are more efficient and scalable than surrogates requiring manual confirmation such as near-crashes Guo et al. (2010).

This study bridges a important gap in utilizing high G-force, namely how to identify the optimal threshold to define a high G-force events. The proposed decision adjusted framework provides systematic and

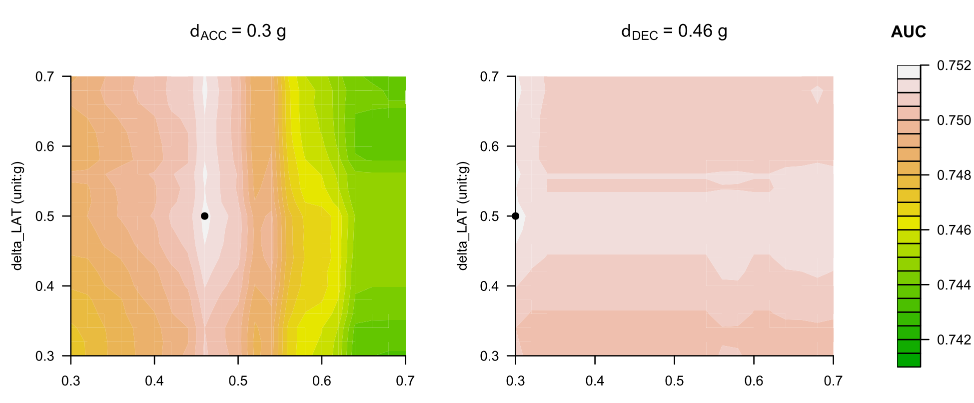

Fig. 5. Heat map of $\mathcal{M}_1$'s AUC values evaluated on $\delta_{ACC} = 0.3g$ (left) and $\delta_{DEC} = 0.46g$ (right). The dots represent the optimal threshold setting of $\mathcal{M}_1$, $(\delta_{ACC}, \delta_{DEC}, \delta_{LAT}) = (0.30g, 0.46g, 0.50g)$, which maximizes the AUC value among all threshold settings evaluated.

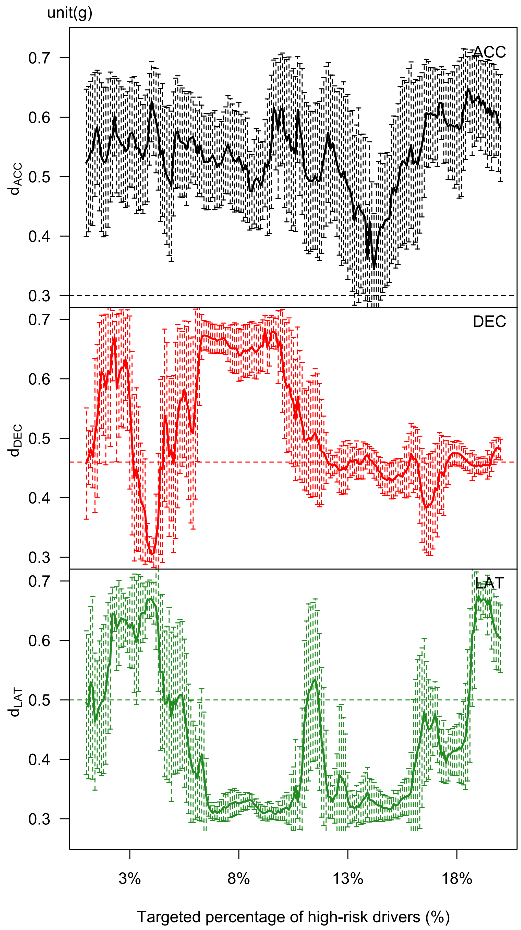

Fig. 6. Point-and-whisker-plot of the "top three" optimal threshold settings under different decision rules for $\mathcal{M}_2$. The dashed lines show $\mathcal{M}_1$'s optimal threshold of for reference.

comprehensive way to address the threshold issue. The results demonstrate that the optimal thresholds do vary according to different decision rules. In the context of predicting a small percentage of highest risk drivers, the proposed model framework provides superior results compared to subjectively defined thresholds or selected based on general model selection criteria.

Using data from the largest NDS conducted to date, the SHRP 2 NDS, we confirmed that high G-force events, including longitudinal and lateral acceleration, provide crucial information on individual driver risk prediction. With the rapid development in connected vehicle technology, high-frequency, high resolution telematics data will be ubiquitously available in the near-future. Telemetries information provides a cost-effective and scalable approach for large scale driving risk assessment applications, such as insurance, fleet safety management, ride-hailing driver and teenage driver risk management.

There are several limitation of the study. The optimization process does require substantial amount of events with low threshold, which might not be universally available. Thresholds based AUC of ROC curve may be valuable for general purpose, although it's not recommended for cases with severely imbalanced classes. The SHRP2 NDS drivers were voluntary samples thus the behavior might not represent the general driver population and caution should be used when extrapolating the results to other driver populations.

The decision-adjusted modeling framework provides a tailored solution to optimize models for a specific research objective. The results confirm that driving kinematic signatures provide useful information for crash risk prediction and that the thresholds for kinematic events are critical for driver-level crash risk prediction. The methodology and results of this paper could provide crucial information for driver risk prediction, safety education, use-based insurance, driver behavior intervention, as well as connected-vehicle safety technology development.

Conflict of interest

None.

Authors' contribution

Huiying Mao: Data curation; Formal analysis; Investigation; Methodology; Validation; Visualization; Roles/Writing – original draft. Xinwei Deng: Formal analysis; Investigation; Methodology; Supervision; Validation; Roles/Writing – original draft. Feng Guo: Conceptualization; Data curation; Formal analysis; Funding acquisition; Investigation; Methodology; Project administration; Resources; Supervision; Validation; Visualization; Roles/Writing – original draft. Zachary Doerzaph: Conceptualization, Funding acquisition; Investigation; Roles/Writing – original draft.

Acknowledgement

This study is partially funded by the National SafeD University Transportation Center at the Virginia Tech.

References

AASHTO, 2010. Highway Safety Manual, vol. 1. American Association of State Highway and Transportation Officials, Washington, DC.
af Wåhlberg, A.E., 2006. Speed choice versus celeration behavior as traffic accident predictor. J. Saf. Res. 37 (1), 43–51.
Bagdadi, O., 2013. Assessing safety critical braking events in naturalistic driving studies. Transport. Res. F 16, 117–126.
Bagdadi, O., Várhelyi, A., 2011. Jerky driving an indicator of accident proneness? Accid. Anal. Prev. 43 (4), 1359–1363.
Bagdadi, O., Várhelyi, A., 2013. Development of a method for detecting jerks in safety critical events. Accid. Anal. Prev. 50, 83–91.
Campbell, K.L., 2012. The shrp 2 naturalistic driving study: addressing driver performance and behavior in traffic safety. TR News 282.
Cantor, D.E., Corsi, T.M., Grimm, C.M., “Ozpolat, K., 2010. A driver focused truck crash prediction model. Transport. Res. E 46 (5), 683–692.
Dingus, T.A., Guo, F., Lee, S., Antin, J.F., Perez, M., Buchanan-King, M., Hankey, J., 2016. Driver crash risk factors and prevalence evaluation using naturalistic driving data. Proc. Natl. Acade. Sci. U.S.A. 113 (10), 2636–2641.
Feng, F., Bao, S., Sayer, J.R., Flannagan, C., Manser, M., Wunderlich, R., 2017. Can vehicle longitudinal jerk be used to identify aggressive drivers? An examination using naturalistic driving data. Accid. Anal. Prev. 104, 125–136.
Friedman, J., Hastie, T., Tibshirani, R., 2009. glmnet: Lasso and elastic-net regularized generalized linear models. R Package Version 1 (4).
Friedman, J., Hastie, T., Tibshirani, R., 2010. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 33 (1), 1.
Guo, F., 2019. Statistical methods for naturalistic driving studies. Annu. Rev. Stat. Appl. 6 (1) https://doi.org/10.1146/annurev-statistics-030718-105153.
Guo, F., Fang, Y., 2013. Individual driver risk assessment using naturalistic driving data. Accid. Anal. Prev. 61, 3–9.
Guo, F., Klauer, S.G., Fang, Y., Hankey, J.M., Antin, J.F., Perez, M.A., Lee, S.E., Dingus, T. A., 2017. The effects of age on crash risk associated with driver distraction. Int. J. Epidemiol. 46 (1), 258–265. https://doi.org/10.1093/ije/dyw234.
Guo, F., Klauer, S., Hankey, J., Dingus, T., 2010. Using near-crashes as a crash surrogate for naturalistic driving studies. Transport. Res. Record 2147 (66–74).
Guo, F., Simons-Morton, B.G., Klauer, S.E., Ouimet, M.C., Dingus, T.A., Lee, S.E., 2013. Variability in crash and near-crash risk among novice teenage drivers: a naturalistic study. J. Pediatr. 163 (6), 1670–1676.
Habtemichael, F.G., de Picado-Santos, L., 2013. The impact of high-risk drivers and benefits of limiting their driving degree of freedom. Accid. Anal. Prev. 60, 305–315.
Hankey, J.M., Perez, M.A., McClafferty, J.A., 2016. Description of the shrp 2 Naturalistic Database and the Crash, Near-Crash, and Baseline Data Sets. Virginia Tech Transportation Institute. Technical report.
Karekla, X., Gkiotsalitis, K., Tyler, N., 2020. The impact of a passenger-safety-driven acceleration limit on the operation of a bus service. Accid. Anal. Prev. 148, 105790.
Kim, S., Song, T.-J., Rouphail, N.M., Aghdashi, S., Amaro, A., Gonçalves, G., 2016. Exploring the association of rear-end crash propensity and micro-scale driver behavior. Saf. Sci. 89, 45–54.

Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J., 2009. Comparing Real-World Behaviors of Drivers With High Versus Low Rates of Crashes and Near Crashes. Technical Report.

Ouimet, M.C., Brown, T.G., Guo, F., Klauer, S.G., Simons-Morton, B.G., Fang, Y.J., Lee, S. E., Gianoulakis, C., Dingus, T.A., 2014. Higher crash and near-crash rates in teenaged drivers with lower cortisol response an 18-month longitudinal, naturalistic study. JAMA Pediatr. 168 (6), 517–522.

Palat, B., Saint Pierre, G., Delhomme, P., 2019. Evaluating individual risk proneness with vehicle dynamics and self-report data? Toward the efficient detection of at-risk drivers. Accid. Anal. Prev. 123, 140–149.

Roidl, E., Frehse, B., Höger, R., 2014. Emotional states of drivers and the impact on speed, acceleration and traffic violations – a simulator study. Accid. Anal. Prev. 70, 282–292.

Sagberg, F., Selpi, Bianchi Piccinini, G.F., Engstr”om, J., 2015. A review of research on driving styles and road safety. Hum. Fact. 57 (7), 1248–1275.

Simons-Morton, B.G., Cheon, K., Guo, F., Albert, P., 2013. Trajectories of kinematic risky driving among novice teenagers. Accid. Anal. Prev. 51, 27–32.

Simons-Morton, B.G., Ouimet, M.C., Wang, J., Klauer, S.G., Lee, S.E., Dingus, T.A., 2009. Hard braking events among novice teenage drivers by passenger characteristics. Proceedings of the.. International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, vol. 2009 236.

Simons-Morton, B.G., Zhang, Z., Jackson, J.C., Albert, P.S., 2012. Do elevated gravitational-force events while driving predict crashes and near crashes? Am. J. Epidemiol. 175 (10), 1075–1079.

Soccolich, S.A., Hickman, J.S., Hanowski, R.J., et al., 2011. Identifying High-Risk Commercial Truck Drivers Using a Naturalistic Approach. Virginia Tech. Virginia Tech Transportation Institute. Technical Report.

Sun, Q.C., Odolinski, R., Xia, J.C., Foster, J., Falkmer, T., Lee, H., 2017. Validating the efficacy of gps tracking vehicle movement for driving behaviour assessment. Travel Behav. Soc. 6, 32–43.

Ulleberg, P., 2001. Personality subtypes of young drivers. relationship to risk-taking preferences, accident involvement, and response to a traffic safety campaign. Transport. Res. F 4 (4), 279–297.

Wang, X., Khattak, A.J., Liu, J., Masghati-Amoli, G., Son, S., 2015. What is the level of volatility in instantaneous driving decisions? Transport. Res. C 58, 413–427.

Wu, K.-F., Aguero-Valverde, J., Jovanis, P.P., 2014. Using naturalistic driving data to explore the association between traffic safety-related events and crash risk at driver level. Accid. Anal. Prev. 72, 210–218.

Yin, S., Duan, J., Ouyang, P., Liu, L., Wei, S., 2017. Multi-cnn and decision tree based driving behavior evaluation. Proceedings of the Symposium on Applied Computing 1424–1429.

Zhu, X., Yuan, Y., Hu, X., Chiu, Y.-C., Ma, Y.-L., 2017. A bayesian network model for contextual versus non-contextual driving behavior assessment. Transport. Res. C 81, 172–187.

Zou, H., Hastie, T., 2005. Regularization and variable selection via the elastic net. J. R. Stat. Soc. Ser. B: Stat. Methodol.) 67 (2), 301–320.