Models of Driver Acceleration Behavior Prior to Real-World Intersection Crashes

John M. Scanlon, Member, IEEE, Rini Sherony, Member, IEEE, and Hampton C. Gabler, Member, IEEE

Abstract—Drivers involved in intersection collisions are at high risk of serious or fatal injury. Intersection advanced driver assistance systems (I-ADAS) are emerging active safety systems designed to help drivers safely traverse intersections. The effectiveness of I-ADAS is expected to be greatly dependent on pre-crash vehicle acceleration during intersection traversals. The objective of this paper was to develop pre-crash acceleration models for non-turning drivers involved in straight crossing path crashes and left-turning drivers in left turn across path opposite direction and lateral direction crashes. This paper used 348 event data recorder pre-crash records taken from crashes investigated as part of the National Automotive Sampling System/Crashworthiness Data System. The acceleration models generated from this pre-crash data were evaluated using a leave-one-out cross-validation procedure. Previously developed non-crash models from the literature were compared with the pre-crash models. Our hypothesis was that drivers involved in crashes would accelerate more aggressively than the “typical” driving population. This result suggests that drivers in pre-crash scenarios tend to accelerate more aggressively than drivers in normal scenarios (p<0.001). This has important implications for the design of I-ADAS. Specifically, higher acceleration results in less available time for I-ADAS to detect and respond to an imminent collision.

Index Terms—Driver assistance system, autonomous vehicles, intersection, active safety system, driver behavior, I-ADAS.

I. INTRODUCTION

Intersection crashes are among the most serious crash modes in the U.S. In 2015, crashes in intersections accounted for 1.3 million police-reported crashes and 5,251 fatal crashes [1], [2]. These crashes correspond to approximately 20% of all police-reported crashes and 16% of all fatal crashes.

Automakers and regulatory agencies are pursuing “Vision Zero”, the ambitious goal of eliminating all roadway fatalities [3], [4]. Crash avoidance systems have been and are being developed to target the most frequent and most harmful crash modes. These systems, often called driver assistance systems, aim to help drivers navigate vehicles and avoid crashes. Common driver assistance systems include, but are not limited to, forward crash avoidance systems, electronic stability control, lane departure warning, lane keeping assist, blind spot detection, rear cross-traffic prevention systems, and backup cameras. Intersection Advanced Driver Assistance Systems (I-ADAS) have recently begun to be deployed within the vehicle fleet [5]–[7]. The objective of this technology is to prevent and reduce the severity of intersection crashes. I-ADAS operates by detecting approaching vehicles while the vehicle is entering the intersection and either alerting the driver of an imminent crash and/or taking some automated evasive action.

Pre-crash vehicle kinematics are expected to highly influence the effectiveness of I-ADAS in real-world intersection crash scenarios. Pre-crash behavior can be divided into three primary phases [8]: the approach phase, the traversal phase, and any crash avoidance action that the driver may take (e.g., braking or steering [9]). During the intersection traversal phase, the ability of I-ADAS to detect an impending collision and deliver a timely warning will be greatly dependent on (a) the speeds that drivers enter the intersection and (b) acceleration magnitude.

The three most common intersection crash modes are straight crossing path (SCP) collisions, left turn across path lateral direction (LTAP/LD) collisions, and left turn across path opposite direction (LTAP/OD) collisions [10]. Approximately one-half of SCP intersection crashes in the U.S. involve a non-turning vehicle performing a rolling stop or complete stop prior to accelerating through the intersection [8], [11], [12]. For LTAP/OD crashes, the left turning vehicle slows (71.6%) or stops (28.4%) prior to accelerating through the intersection [13]. Similarly in LTAP/LD crashes, the left turning vehicle typically slows or stops prior to entering the intersection [14], [15]. To help drivers avoid intersection crashes, it is crucial that I-ADAS deliver a warning based on how drivers typically navigate across intersections in these crash modes. A timely warning is essential for ensuring that there is adequate time for an avoidance maneuver. However, warning drivers too early could be perceived as a false-positive alert, which can affect driver response in a crash-imminent scenario [16]–[18] or lead to the driver turning off these systems altogether [19].

Several past studies have characterized and developed models of “typical” driver acceleration behavior. Two general methodologies were used to develop these models. One method is to mathematically model vehicle acceleration patterns as a function of relevant predictor variables, such as...
turning behavior, current speed, or distance from the intersection [20]–[24]. A second method is to construct a vehicle dynamics models that estimates vehicle acceleration based on vehicle body forces [25], [26]. The existing literature, however, focuses almost entirely on “normal” driving, i.e., non-crashes. The study of intersection traversal during crashes has been greatly hindered by a lack of detailed data on vehicle kinematics and driver behavior during these serious events.

There is reason to suspect pre-crash intersection traversals would vary from “normal” driving behavior. Driver aggressiveness and risk-taking behavior have been correlated with higher crash incidence [27]–[30]. A large body of literature has examined characteristics of high-risk driving behaviors, which includes speeding [31], [32], maintaining a narrow following distance [33], [34], driving while impaired [35], [36], gap acceptance [37], and driving without a seat belt [33]. Additionally, an elevated number of high g-force events, e.g., hard braking, rapid acceleration, and faster turning, have been associated with elevated crash risk [38], [39]. Accelerometers measuring high g-force events have also been used as an in-vehicle evaluation tool of risky driving behavior [30], [40]. Some mechanisms that can explain this higher crash risk are (a) less time for crash avoidance action from threat detection to impact [41] and (b) loss of control [42].

Our goal in this paper was to provide models of driver acceleration, which I-ADAS developers can use to project how quickly a driver will traverse a given intersection. Crash avoidance systems have to be “tuned” based on a number of factors, including how early on-board sensors could detect an oncoming vehicle, computational latency times for detecting conflicts, driver perception-reaction times, and any latency to activate actuators (e.g., applying brakes or steering). A vital part of this tuning is being able to predict a vehicle’s path through time. Determining acceleration traversal patterns can also be helpful for selecting appropriate sensor requirements for detecting oncoming vehicles [11], [41], [43], [44]. An ideal I-ADAS sensor should be capable of identifying potential collision partners at the first clear line-of-sight opportunity. Additionally, not only is this important for designers of these systems, but these models could also prove helpful for evaluators considering a future test that assesses intersection crash avoidance systems. These tests should be performed under crash conditions that replicate actual scenarios that occur in the crash population. Lastly, traversal models permit studies forecasting the population-wide feasibility of I-ADAS [8]. An important component of these studies is the simulation of vehicle behavior as it actually occurs during real-world crashes.

Event data recorders (EDR) can provide unique insights into pre-crash driver behavior. EDRs are the “black boxes” now installed in almost all new U.S. passenger vehicles, which directly measure and record pre-crash vehicle and driver behavior in the event of a collision [45]. EDRs record several pre-crash data elements which are key to understanding intersection traversal, including vehicle speed, accelerator application, brake application, yaw rate, and steering wheel angle [46]. For example, consider the EDR record of pre-crash speed from a real world crash shown in Figure 1.

**Fig. 1.** An EDR-recorded speed profile of a vehicle involved in a SCP intersection crash (NASS/CDS 2013-76-166) with the scene diagram prepared by the investigator of the crash.

The red shaded vehicle performed a complete stop, and then accelerated through the intersection. The driver that stopped attempted an evasive braking maneuver just prior to impacting the left side of a vehicle attempting to travel straight through the intersection.

This study analyzed the pre-crash kinematics of non-turning and left-turning vehicles accelerating into intersections using this EDR data. Our hypothesis was that drivers involved in crashes would accelerate more aggressively than the “typical” driving population. The objective of this study was to develop pre-crash acceleration models for non-turning drivers in SCP crashes and left-turning drivers in LTAP/LD and LTAP/OD crashes. Two research questions were proposed in this study. First, how representative are previously developed acceleration models of pre-crash acceleration behavior? Second, can a more accurate acceleration profile model be developed to represent this pre-crash data?

### II. METHODS

#### A. Data Source

This study was based on crash investigation records from the National Automotive Sampling System / Crashworthiness Data System (NASS/CDS). Case years 2000 to 2015 were used in this studies analyses. This database is compiled annually by the National Highway Traffic Safety Administration (NHTSA), and for a case to be included, the crash must have been police-reported and one of the involved vehicles had to have been towed from the scene due to damage.

NASS/CDS uses a complex sampling scheme in order to generate nationally representative estimates that represent the
The entire U.S. crash population [47]. Three stages of sampling are used to select crashes for inclusion in the database. First, the U.S. was divided into primary sampling units, or PSUs. This was done by dividing up the country into 1,195 areas based on whether the area is "central city, a county surrounding a central city, an entire county or a group of contiguous counties". PSU’s were then divided into 12 strata based on geographic region and type, such as urban vs. rural areas. A total of 24 PSUs were then selected from these 12 strata, where each strata represents an approximately equal proportion of crashes. Second, police jurisdictions were selected from each of the PSU’s. It is impractical and expensive to sample all crashes within a PSU. As an alternative, a sample of jurisdictions within each PSU were selected from which to draw crashes. Third, crashes were selected from each of the jurisdictions. Crashes were classified based on vehicle type, occupant injuries, vehicle towing, and vehicle model years. Each crash was then assigned a national weighting factor ("RATWGT" variable) that adjusts for crash severity and location. This study used these national weighting factors in order to make nationally representative estimates.

B. EDR Module Selection

Several criteria were used to select the set of EDR records to be included in this study. First, only non-turning vehicles in SCP intersection crashes and left-turning vehicles involved in LTAP/LD and LTAP/OD crashes were considered for this study. These intersection crash modes and vehicle turning behaviors were identified using the "ACCTYPE" and "RELINTER" variables within the NASS/CDS database. Second, the EDR must have recorded an airbag deployment or a total delta-v greater than 5 mph. EDRs lock deployment events into EDR memory, whereas non-deployment events can be overwritten by future impacts. A 5-mph threshold would likely have resulted in substantial vehicle damage, and accordingly, would have been unlikely to be a result of another event not described in the NASS/CDS database. Third, the first impact must have been described as the most severe impact (highest delta-v) experienced during the crash sequence as indicated by the "ACCSSEQ1" variable. These criteria helped to additionally ensure that the pre-crash record being analyzed corresponded to the time period preceding the initial impact. Previous analysis of the NASS/CDS database showed that the first event was the most severe impact in 97% of intersection crashes [9]. Fourth, the "Complete File Recorded" flag was examined for each record to ensure that the pre-crash data was fully collected. Chrysler EDRs do not contain this indicator, but instead have a "Pre-Crash Recorder Status" flag that records whether all pre-crash data (e.g., speed, braking) was collected for each time point. Fifth, only EDRs that recorded pre-crash speed and braking data were included in this study.

To ensure similitude between the model dataset and all intersection crashes, the compositions of each dataset was compared with all similar crash events within the NASS/CDS database. Several potential confounding factors were compared. A Rao-Scott chi-squared test was used to identify differences in the composition of each dataset. This analysis can be found in the Appendix.

C. Extracting EDR-Recorded Intersection Traversal Accelerations

The goal of this study was to model the acceleration patterns of vehicles as they are attempting to safely pass through the intersection and prior to crash avoidance braking. Because these are EDR pre-crash records, evasive braking commonly occurs [9], which is accounted for using the methods in the following section.

Velocity time points from accelerating vehicles were extracted from the EDR records in two steps. The overall process is depicted in Figure 2. First, any time points in which evasive braking occurred were removed from the time series. Evasive braking was determined by whether the brakes were applied at the last EDR-recorded time point. In the event of an evasive braking maneuver being indicated, all immediately prior time points where the service brake was engaged were removed.

Second, the remaining data were used to determine if and at what EDR-recorded time points the driver was accelerating into the intersection. Some of the EDRs in our sample recorded accelerator pedal position as a percentage of the fully-depressed position. When available, accelerator pedal application was used to determine when the driver was performing an acceleration maneuver. If the accelerator pedal was pressed at the last EDR-recorded time point prior to evasive braking, the driver was assumed to have been accelerating into the intersection. The time points with accelerator pedal application were then extracted as the acceleration time series.
If accelerator pedal application was not recorded by the EDR, the speed profile was instead used to extract the acceleration time series. If an increase in velocity was observed between the two velocity time points prior to evasive braking application, the driver was assumed to have accelerated into the intersection, and all acceleration time points were extracted. However, without accelerator pedal application data, there is some uncertainty in whether the driver was accelerating at the earliest EDR-recorded acceleration time point. For example, the driver could have begun depressing the accelerator pedal just prior to the second acceleration time point but following the first acceleration time point. Accordingly, the first acceleration time point was removed from the acceleration time series for EDRs that did not record accelerator pedal application.

All model development and evaluation performed in this study used these extracted pre-crash traversal acceleration time series. The first data point in the extracted acceleration time series was used as an “initial condition”. All subsequent acceleration time points were used to evaluate the predictive power of the model.

D. Previously Developed Acceleration Profile Models

The studies discussed below have reported non-crash vehicle acceleration models developed for traffic engineers and intersection geometry design. The general form of each model can be seen in Figure 3.

The constant acceleration model is the simplest representation of driver acceleration into intersections. Using this model, the driver is assumed to maintain a constant acceleration until some target final speed is reached. This simplistic model has its drawbacks. First, this model assumes an instantaneous jerk at the initiation and conclusion of the acceleration event. Second, the model fails to account for changes in acceleration magnitude throughout the traversal that occur due to natural driving tendencies and changing gears [26]. Several studies have consistently found this model to be a poor approximation of real-world driver behavior [20]–[22], and as such, the constant acceleration model was not considered in this study’s analyses.

Bham and Benekohal [21] developed a two-phase model that was evaluated in this study. Their model was based on acceleration data previously collected by Ohio State University using aerial photogrammetry [48]. The model assumed the driver would begin to accelerate at some higher constant value (1.10 m/s$^2$) until some vehicle speed had been reached (12.97 m/s), and then the vehicle would instantaneously begin to accelerate at some lower constant value (0.37 m/s$^2$). This relationship is shown in Equation 1. Integrating Equation 1 yields the velocity relation used in this study, as shown in Equation 2.

$$a(t_i) = a_{1,2},$$
$$v(t_{i+1}) = a(t_i) \cdot \Delta t + v(t_i)$$

Where,

- $a(t_i)$ = acceleration at the current time step in m/s$^2$,
- $a_1$ = acceleration before 12.97 m/s = 1.10 m/s$^2$,
- $a_2$ = acceleration before 12.97 m/s = 0.37 m/s$^2$,
- $v(t_{i+1})$ = velocity at the next time step in m/s,
- $v(t_i)$ = velocity at the current time step in m/s,
- $\Delta t$ = time step in s.

Wang et al. [22] developed a linear-decreasing model and a quadratic-decreasing model which were assessed in this study. Like the two-phase model, both of these models assume drivers will initially accelerate at a higher rate followed by lower rates. The acceleration data for developing these models were taken from a naturalistic driving study that was performed in the Atlanta urban area. In this study, 100 vehicles...
were equipped with GPS systems and the coordinates of these vehicles were tracked. The researchers fit regression models to the acceleration vs. velocity relation of drivers as they were accelerating. Unique models were created for non-turning and left-turning vehicles. The fit models are shown in Equations 3-6.

\[
\text{Straight – Linear Decreasing: } a(t) = 1.883 - 0.021 \cdot v(t) \\
\text{Left turn – Linear Decreasing: } a(t) = 1.646 - 0.017 \cdot v(t) \\
\text{Straight – Quad. Decreasing: } a(t) = [1.381 - 0.011 \cdot v(t)]^2 \\
\text{Left turn – Quad. Decreasing: } a(t) = [1.289 - 0.009 \cdot v(t)]^2
\]

Where, 
\[
a(t) = \text{Acceleration at time point } t \text{ in m/s}^2, \\
t = \text{time in s,} \\
v(t) = \text{Velocity in km/hr.}
\]

Akçelik and Biggs [20] developed a polynomial acceleration model which was also examined in this study. This model assumed that drivers begin to accelerate with zero jerk, ramps up vehicle acceleration to some maximum value, and then gradually decreases vehicle acceleration with zero jerk at acceleration offset. The data used to develop this model was collected in Sydney, Australia using the chase-car method. Data were collected in urban, suburban, and rural traffic. Using this technique, an instrumented vehicle randomly selects and follows vehicles in traffic. The equipped “chasing” car documents the movements of the lead vehicle. The general form of the polynomial model can be seen in Equation 7. The polynomial model provides a highly tunable option for representing a given acceleration profile. However, implementing this model requires several parameters that cannot be readily determined from pre-crash acceleration events, including vehicle speed at the end of the acceleration event, total acceleration time, and total acceleration distance. Each of these model inputs can only be determined if the driver completes the acceleration event, which very rarely occurs in intersection crash scenarios. As such, this model was not considered for this study’s analyses.

\[
a(t) = r \cdot a_m \cdot \theta \cdot (1 - \theta^m)^2
\]

Where, 
\[
a(t) = \text{acceleration at the current time step in m/s}^2, \\
t = \text{current time in s,} \\
r, m = \text{model parameters,} \\
a_m = \text{maximum acceleration in m/s}^2, \\
\theta = \text{time ratio, } t / t_a, \\
t_a = \text{acceleration time in s}
\]

E. Development of a New Pre-Crash Acceleration Model

A linear-decreasing model and a quadratic-decreasing acceleration model were developed in this study to represent driver acceleration behavior prior to real-world intersection crashes. These models were selected for several reasons. First, these models can be developed from only partial acceleration profiles (i.e., the beginning and end of the traversal event is not required). Second, several prior studies have consistently found that driver’s accelerate at a higher degree early in the acceleration traversal [20]-[22]. Third, this model is computationally efficient, and can be readily applied for future applications. The general form of the linear-decreasing and quadratic-decreasing models are shown in Equations 8 and 9, respectively.

\[
a(t) = c_1 - c_2 \cdot v(t) \\
a(t) = [c_1 - c_2 \cdot v(t)]^2
\]

Where, 
\[
a(t) = \text{Acceleration at time point } t \text{ in m/s}^2, \\
v(t) = \text{Velocity at time point } t \text{ in m/s}, \\
t = \text{time in s,} \\
c_1, c_2 = \text{model parameters.}
\]

The model parameters, \(c_1\) and \(c_2\) shown in Equation 8 and 9 were computed using a simulated annealing global optimization algorithm [49]. An alternative, and perhaps preferable, approach would be to compute parameters using regression methods. However, acceleration rates were not directly measured from EDR models and could not be determined from the raw sampling rate and resolution that vehicle speed is typically collected. An overall error score shown Equation 10 was developed as an objective function to be minimized by the optimizer. The form of this equation was selected for several reasons. First, the squared residual allowed for an unsigned measure of observed versus predicted speed. Second, summing the squared residuals and normalizing by the total number of data points helped account for differences in sampling rates and acceleration durations. Third, each acceleration trace was weighted using the NASS case weighting factor. Accordingly, the resulting model parameters were selected based on real-world frequency.

\[
\text{Overall Error} = \frac{1}{\sum w_{case} \sum_{i=1}^{N_{case}} \sum_{j=1}^{N_{traj,i}} [(v_{EDR,i,j} - v_{EDR,i,j})^2 \cdot w_{case,i,j}] / N_{traj,i}}
\]

Where, 
\[
\text{Overall error} = \text{Measure used to evaluate models in (m/s)^2,} \\
i = \text{index for case,} \\
j = \text{index for time point in trajectory,} \\
v_{EDR} = \text{EDR recorded vehicle speed in m/s,} \\
v_{sim} = \text{Model predicted vehicle speed in m/s,} \\
N_{case} = \text{Total number of cases,} \\
N_{traj} = \text{Number of time points in acceleration trajectory,} \\
w_{case} = \text{NASS case weight.}
\]
F. Model Evaluations

The accuracy of each developed model was compared with the previously developed non-crash acceleration profile models using a leave-one-out cross-validation procedure. Using this technique, one observation is withheld as a “testing dataset”, while the remaining n-1 observations (n = total number of observations) are used as a “training dataset” to develop the pre-crash acceleration models. The generated pre-crash models and previously developed models were then used to predict subsequent velocity time points in the testing dataset. This process was repeated n-times until velocity time point predictions had been generated for every acceleration profile. Using the overall error score shown in Equation 10, the overall ability of each model to accurately predict the observed EDR velocity time points was evaluated. Additionally, a regression line was fit through the observed versus predicted values. The slope of the line and $R^2$ were used to assess the predicted capacity of the models.

In addition, a trajectory error score, shown in Equation 11 was calculated for each case in the dataset. Trajectory error is similar to the overall error score but is only calculated for a single trajectory and additionally retains whether each model tended to overestimate or underestimate acceleration behavior. The trajectory “error sign” was retained by taking the sign of the difference between the observed and predicted final velocity points.

$$\text{Trajectory Error} = \sum_{j} \left( \frac{(v_{\text{sim},j} - v_{\text{EDR},j})^2}{N_{\text{traj}}} \cdot \frac{1}{N_{\text{traj}}} \cdot \text{sgn}(v_{\text{sim},N_{\text{traj}}}, v_{\text{EDR},N_{\text{traj}}}) \right)$$  

Where,

- Trajectory Error = Measure used to evaluate models in $m^2/s^2$
- $j$ = index for time point in trajectory,
- $v_{\text{EDR}}$ = EDR recorded vehicle speed in m/s,
- $v_{\text{sim}}$ = Model predicted vehicle speed in m/s,
- $N_{\text{traj}}$ = Number of time points in acceleration trajectory.

Differences between the least squares means of trajectory error scores were examined to assess each model’s ability to predict pre-crash speeds. All statistical analyses were performed using the “SURVEYREG” procedure in the SAS 9.3 software package (SAS Institute, Cary, NC) [50]. For a given crash mode, model type was used as the independent variable. Trajectory error, shown in Equation 11, was used as the dependent variable. An alpha level of 0.05 was used to determine significance.

As a second form of validation, a basic form of the cross validation method was also used to compare the models generated in this study. Odd numbered case years were used to “train” the dataset. Even numbered case years were used to test the dataset. Overall error was used to evaluate the different models.

### TABLE I

<table>
<thead>
<tr>
<th>Pre-crash Time Window</th>
<th>Sampling Rate</th>
<th>EDR Sampling Velocity Resolution</th>
<th>EDR Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 s</td>
<td>1 Hz</td>
<td>0.16 kph [1.0 mph]</td>
<td>3</td>
</tr>
<tr>
<td>5 s</td>
<td>1 Hz</td>
<td>1.6 kph [1.0 mph]</td>
<td>216</td>
</tr>
<tr>
<td>5 s</td>
<td>1 Hz</td>
<td>2.0 kph [1.2 mph]</td>
<td>26</td>
</tr>
<tr>
<td>5 s</td>
<td>2 Hz</td>
<td>0.16 kph [0.1 mph]</td>
<td>7</td>
</tr>
<tr>
<td>5 s</td>
<td>2 Hz</td>
<td>1.0 kph [0.6 mph]</td>
<td>27</td>
</tr>
<tr>
<td>5 s</td>
<td>2 Hz</td>
<td>1.6 kph [1.0 mph]</td>
<td>29</td>
</tr>
<tr>
<td>5 s</td>
<td>5 Hz</td>
<td>1.6 kph [1.0 mph]</td>
<td>2</td>
</tr>
<tr>
<td>5 s</td>
<td>10 Hz</td>
<td>1.6 kph [1.0 mph]</td>
<td>26</td>
</tr>
<tr>
<td>2.5 s</td>
<td>2 Hz</td>
<td>1.6 kph [1.0 mph]</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>348</td>
</tr>
</tbody>
</table>

### TABLE II

<table>
<thead>
<tr>
<th>Crash Mode (Turn)</th>
<th>Unweighted Case Count</th>
<th>Weighted Case Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCP (Straight)</td>
<td>138</td>
<td>53,708</td>
</tr>
<tr>
<td>LTAP/OD (Left)</td>
<td>125</td>
<td>58,530</td>
</tr>
<tr>
<td>LTAP/LD (Left)</td>
<td>85</td>
<td>26,493</td>
</tr>
<tr>
<td>Total</td>
<td>348</td>
<td>138,731</td>
</tr>
</tbody>
</table>

### III. Results

#### A. Case Selection

A total of 348 EDR records were extracted for this study’s analyses from case years 2000 to 2015. Table I categorizes the dataset by intersection crash mode. The largest dataset of EDRs came from SCP crashes followed by LTAP/OD and LTAP/LD. A total of 1,308 data points from these cases were used in this study for development and validation of the model. Each of these data points are velocity measurements which occurred after the time point of first acceleration.

A summary of the EDR modules used in this study’s analyses can be found in Table II. A variety of module types were used in this study. Most of the records contained 5 s of pre-crash data, but a few records (12 of 348) contained only 2.5 s. Most modules, in general, can record 5 s of pre-crash data, and modules with longer recordings also have a higher likelihood of containing an extractable acceleration profile. All of the records used collected vehicle speed at a resolution of at least 2.0 kph (1.2 mph). Additionally, most (245 of 348) of the EDRs used in this study collected vehicle speed at 1 Hz.

A summary of the dataset composition with respect to driver, vehicle, and roadway characteristics is available in the Appendix. A comparison between the model dataset and all similar events in NASS/CDS is also provided. Overall, the model dataset compared well with all similar events in NASS/CDS, which suggests that the model dataset is a
TABLE III
LEAST-SQUARES MEANS OF TRAJECTORY ERROR AND MEAN DIFFERENCES BETWEEN THE VARIOUS MODELS

<table>
<thead>
<tr>
<th>Trajectory Error (m/s)² (p-value)</th>
<th>SCP</th>
<th>Bham 2-phase</th>
<th>Wang Linear</th>
<th>Wang Quad.</th>
<th>Scanlon Linear</th>
<th>Scanlon Quad.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bham 2-phase</td>
<td>-3.17</td>
<td>-1.27</td>
<td>-1.54</td>
<td>0.00</td>
<td>-0.08</td>
<td></td>
</tr>
<tr>
<td>Wang Linear</td>
<td>-1.27</td>
<td>-1.27</td>
<td>-1.54</td>
<td>0.00</td>
<td>-0.08</td>
<td></td>
</tr>
<tr>
<td>Wang Quad.</td>
<td>1.54</td>
<td>-0.26</td>
<td>1.27</td>
<td>1.91</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scanlon Linear</td>
<td>0.00</td>
<td></td>
<td>1.53</td>
<td>1.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scanlon Quad.</td>
<td>-0.08</td>
<td></td>
<td></td>
<td>-0.08 (0.290)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| LTAP/OD                          |                    |              |             |            |                |               |
| Model                            |                    |              |             |            |                |               |
| Bham 2-phase                     | -2.50              | -1.26        | -1.37       | 0.01       | 0.02           |               |
| Wang Linear                      | -1.26              | -0.11        | 1.27        | 1.28       |                 |               |
| Wang Quad.                       | -1.37              |              | 1.38        | 1.39       |                 |               |
| Scanlon Linear                   | 0.01               |              |             | 0.01 (0.710) |                |               |
| Scanlon Quad.                    | 0.02               |              |             |            |                |               |

| LTAP/LD                          |                    |              |             |            |                |               |
| Model                            |                    |              |             |            |                |               |
| Bham 2-phase                     | -2.61              | -1.36        | -1.45       | -0.11      | -0.09          |               |
| Wang Linear                      | -1.36              | -0.09        | 1.24        | 1.27       |                 |               |
| Wang Quad.                       | -1.45              |              | 1.35        | 1.36 (0.001) |                |               |
| Scanlon Linear                   | -0.11              |              |             | 0.02 (0.470) |                |               |
| Scanlon Quad.                    | -0.09              |              |             |            |                |               |

B. Leave-one-out Cross-Validation Procedure

Table III shows least squares means of trajectory errors with comparisons between models for each crash mode. In general, the non-crash models tended to underestimate pre-crash vehicle speeds ($p < 0.001$). Using the leave-one-out method, the pre-crash models developed in this study were not found to not significantly underestimate or overestimate pre-crash acceleration patterns. When comparing the models, the new linear-decreasing and quadratic-decreasing models generated in this study were able to predict pre-crash acceleration behavior significantly ($p < 0.001$) better than each of the previously developed models for every crash mode. However, no differences were observed between the linear-decreasing and quadratic-decreasing models generated in this study. It should also be noted that each of the previously developed models performed significantly different from one another ($p < 0.001$) for every crash mode variation.

Figure 4 presents a comparison of the simulated versus observed velocity time points for this study’s models and each of the previously developed models that were evaluated. Overall error scores for each model/crash mode variation can be found in Table 4. A lower overall error scores corresponded to higher model accuracy. Each predicted velocity was generated using the Leave-one-out cross-validation procedure. All of the non-crash models being evaluated had greater overall error scores than the pre-crash models. The two-phase model underestimated the final velocity time point for 85%-97% of cases depending on the crash mode. The Wang linear-decreasing underestimated 75%-84% of these final velocity time points, and the Wang quadratic-decreasing underestimated 76%-86% of final velocity time points. The new pre-crash linear-decreasing models developed in this study only underestimated 45-55% of final velocity time points, and the new quadratic-decreasing model underestimated 47-60% of final velocity time points.

As an additional form of evaluation, a regression line was fit through the predicted versus observed vehicle speeds, which is shown in Figure 4. Predicted and observed speeds were presented with respect to the speed at the start of the recorded acceleration. The intercept was fixed to go through the origin. It should be noted that fitting this regression does violate the assumption of independent observations. However, analyzing the slope of this fit allows for interpreting the degree to which the models underestimate or overestimate the observed pre-crash data. Based on the slope, the non-crash two-phase model was found to underestimate pre-crash speeds by approximately 40-48%. The non-crash Wang models underestimated speeds by 29-34%. The pre-crash models developed in this study underestimated speed by approximately 7-10%. As discussed, based on the trajectory error analysis, the pre-crash model underestimation was not found to be significantly significant for either the linear- or quadratic-decreasing in any of the three crash modes.

TABLE IV
OVERALL ERROR SCORES FOR THE VARIOUS MODELS

<table>
<thead>
<tr>
<th>Overall Error (m/s)²</th>
<th>Model</th>
<th>SCP</th>
<th>LTAP/OD</th>
<th>LTAP/LD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bham 2-phase</td>
<td>3.19</td>
<td>2.62</td>
<td>2.63</td>
<td></td>
</tr>
<tr>
<td>Wang Linear</td>
<td>1.46</td>
<td>1.52</td>
<td>1.46</td>
<td></td>
</tr>
<tr>
<td>Wang Quad.</td>
<td>1.64</td>
<td>1.59</td>
<td>1.54</td>
<td></td>
</tr>
<tr>
<td>Scanlon Linear</td>
<td>0.91</td>
<td>0.89</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td>Scanlon Quad.</td>
<td>0.92</td>
<td>0.89</td>
<td>1.05</td>
<td></td>
</tr>
</tbody>
</table>

*greyed cells represent pre-crash models developed in current study.
Fig. 4. Simulated speeds versus EDR-recorded speed for every data point used in this study. Three previously generated models are shown in the figure, including the Bham two-phase, Wang linear-decreasing, and Wang quadratic-decreasing models. The Scanlon linear-decreasing and quadratic-decreasing models generated in this study are additionally presented in the figure. Data points are sized based on their NASS case weight.

C. Basic Cross-Validation Procedure
The Appendix contains the analysis using a basic cross-validation procedure, where a portion of the dataset was used to either test or validate the model. The results were consistent with the findings of the leave-one-out procedure, in that, the pre-crash models developed were found to provide a better representation of pre-crash acceleration patterns.

D. Overall Model Fits
Optimum model parameters for the linear-decreasing and quadratic-decreasing models were then computed for each dataset of intersection crash modes. The resulting models can are presented in Equations 12-17 listed in Table V. Figure 5 provides some example simulated trajectories for the three new linear-decreasing pre-crash models and each non-crash model. The quadratic-decreasing models produced similar trajectories to the linear-decreasing models and are not shown.

The non-crash and crash models give dramatically different predictions of time to collision. Consider a crash that is going to occur 20-m after an I-ADAS vehicle begins to accelerate from rest into an intersection. The Bham and Benkohal
TABLE V
EQUATIONS FOR OVERALL PRE-CRASH MODEL FITS

<table>
<thead>
<tr>
<th>Crash Mode (Turn)</th>
<th>Model</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCP (Straight)</td>
<td>Linear</td>
<td>$a(t) = 2.782 - 0.154 \cdot v(t)$ (12)</td>
</tr>
<tr>
<td></td>
<td>Quad.</td>
<td>$a(t) = [1.745 - 0.090 \cdot v(t)]^2$ (13)</td>
</tr>
<tr>
<td>LTAP/OD (Left)</td>
<td>Linear</td>
<td>$a(t) = 2.924 - 0.247 \cdot v(t)$ (14)</td>
</tr>
<tr>
<td></td>
<td>Quad.</td>
<td>$a(t) = [1.791 - 0.099 \cdot v(t)]^2$ (15)</td>
</tr>
<tr>
<td>LTAP/LD (Left)</td>
<td>Linear</td>
<td>$a(t) = 2.167 - 0.057 \cdot v(t)$ (16)</td>
</tr>
<tr>
<td></td>
<td>Quad.</td>
<td>$a(t) = [1.489 - 0.025 \cdot v(t)]^2$ (17)</td>
</tr>
</tbody>
</table>

Where,
- $a(t)$ = Acceleration at time point $t$ in m/s²,
- $v(t)$ = Velocity at time point $t$ in m/s,
- $t$ = time in s.

IV. DISCUSSION

The three non-crash models evaluated in this study underestimated pre-crash velocity. This finding provides support for our hypothesis that drivers involved in crashes may tend to accelerate differently than drivers in typical scenarios. The existing models were developed using non-crash data from “normal” driving. Additionally, some of the drivers, for example LTAP/OD crashes, may have seen the oncoming car, and attempted to rapidly accelerate to avoid an imminent collision. Gap selection errors due to misjudgment of available space and/or velocity of the oncoming driver is frequently cited as a contributing factor in real-world LTAP/OD crashes [13].

The findings of this paper have important implications for the design and evaluation of an I-ADAS. As previously stated, the acceleration patterns of vehicles as drivers enter and traverse intersections directly influences the amount of time available for crash avoidance action. The higher-than-normal acceleration behavior of driver’s prior to real-world crashes suggests that there is a need for I-ADAS design to account for this potentially more aggressive driving behavior. Using the Wang et al. non-crash models and our new pre-crash model, we can compare drivers in crashes versus drivers in “normal” intersection traversal. For non-turning drivers, the new SCP pre-crash model predict that drivers typically involved in crashes will travel 20-m from a complete stop 0.54-0.77 s faster than a driver in a “normal” driving scenario. Likewise, for left-turning drivers, drivers in LTAP/OD and LTAP/LD crash scenarios will traverse 20-m from a stop 0.83-0.94 s and 0.71-0.78 s faster, respectively. Although only less than a second sooner, current active safety systems begin to warn drivers a mere 2-3 s prior to a collision occurring [51]. Given a typical crash avoidance braking magnitude of 0.6 g [9], a driver could slow their vehicle by 10 mph if allotted an additional 0.8 s to react. This additional reduction in speed could not only be crucial for avoiding the crash but could also potentially reduce injuries in crashes which could not be avoided.

There do appear to be some differences in pre-crash acceleration behavior between the different crash modes. However, differences between the different crash modes was not tested in this study. In general, the acceleration of non-turning drivers in SCP crashes appears to be higher during the early traversal phase than left-turning drivers in LTAP/LD and LTAP/OD crashes. This finding is consistent with prior work comparing non-turning drivers with left-turning drivers [22]. The early acceleration of vehicles in LTAP/OD crashes also appears to be slightly greater than drivers in LTAP/LD crashes. As discussed previously, one reason for higher acceleration prior to real-world crashes may be that the driver saw the oncoming vehicle and then attempted to quickly accelerate through the gap in traffic. Gap selection, where the left-turning drivers sees the oncoming vehicle but still attempts to turn, commonly occurs in U.S. LTAP/OD crashes [13].

This study found insufficient statistical evidence to suggest that the linear-decreasing or the quadratic-decreasing model is a better representation of pre-crash traversal acceleration behavior. In fact, there appear to be only slight differences between the two models for predicting pre-crash data. Previous work by Wang et al. [22] found that the linear-decreasing relationship of acceleration versus velocity tends to overestimate vehicle acceleration at higher speeds. In the current study, incomplete acceleration traces (limited EDR recording

Fig. 5. Simulated distance versus time trajectories for a vehicle accelerating from rest. Each acceleration trace corresponds to one of the models evaluated in this study.

two-phase model predicts that the driver will reach the impact location in 6.03 s. The Wang et al. linear-decreasing and quadratic-decreasing models estimate that the crash will occur 4.89-4.97 s, respectively, following initial acceleration for non-turning vehicles. In contrast, the pre-crash models developed in this study predict that these crashes will occur much sooner. For SCP crashes, both models expect the crash to occur in 4.20-4.35 s – over 0.5 seconds earlier than the non-crash models predict.

For turning vehicles, the Wang et al. linear-decreasing and quadratic-decreasing models predict that the crash will occur 5.19-5.25 s, respectively. By contrast for LTAP/OD, the crash model anticipates the crash to occur in 4.31-4.36 s. For LTAP/LD, the crash model predicts the crash will occur in 4.47-4.48 s – substantially earlier than the non-crash models.

IV. DISCUSSION

The three non-crash models evaluated in this study underestimated pre-crash velocity. This finding provides support for our hypothesis that drivers involved in crashes may tend to accelerate differently than drivers in typical scenarios. The existing models were developed using non-crash data from “normal” driving. Additionally, some of the drivers, for example LTAP/OD crashes, may have seen the oncoming car, and attempted to rapidly accelerate to avoid an imminent collision. Gap selection errors due to misjudgment of available space and/or velocity of the oncoming driver is frequently cited as a contributing factor in real-world LTAP/OD crashes [13].

The findings of this paper have important implications for the design and evaluation of an I-ADAS. As previously stated, the acceleration patterns of vehicles as drivers enter and traverse intersections directly influences the amount of time available for crash avoidance action. The higher-than-normal acceleration behavior of driver’s prior to real-world crashes suggests that there is a need for I-ADAS design to account for this potentially more aggressive driving behavior. Using the Wang et al. non-crash models and our new pre-crash model, we can compare drivers in crashes versus drivers in “normal” intersection traversal. For non-turning drivers, the new SCP pre-crash model predict that drivers typically involved in crashes will travel 20-m from a complete stop 0.54-0.77 s faster than a driver in a “normal” driving scenario. Likewise, for left-turning drivers, drivers in LTAP/OD and LTAP/LD crash scenarios will traverse 20-m from a stop 0.83-0.94 s and 0.71-0.78 s faster, respectively. Although only less than a second sooner, current active safety systems begin to warn drivers a mere 2-3 s prior to a collision occurring [51]. Given a typical crash avoidance braking magnitude of 0.6 g [9], a driver could slow their vehicle by 10 mph if allotted an additional 0.8 s to react. This additional reduction in speed could not only be crucial for avoiding the crash but could also potentially reduce injuries in crashes which could not be avoided.

There do appear to be some differences in pre-crash acceleration behavior between the different crash modes. However, differences between the different crash modes was not tested in this study. In general, the acceleration of non-turning drivers in SCP crashes appears to be higher during the early traversal phase than left-turning drivers in LTAP/LD and LTAP/OD crashes. This finding is consistent with prior work comparing non-turning drivers with left-turning drivers [22]. The early acceleration of vehicles in LTAP/OD crashes also appears to be slightly greater than drivers in LTAP/LD crashes. As discussed previously, one reason for higher acceleration prior to real-world crashes may be that the driver saw the oncoming vehicle and then attempted to quickly accelerate through the gap in traffic. Gap selection, where the left-turning drivers sees the oncoming vehicle but still attempts to turn, commonly occurs in U.S. LTAP/OD crashes [13].

This study found insufficient statistical evidence to suggest that the linear-decreasing or the quadratic-decreasing model is a better representation of pre-crash traversal acceleration behavior. In fact, there appear to be only slight differences between the two models for predicting pre-crash data. Previous work by Wang et al. [22] found that the linear-decreasing relationship of acceleration versus velocity tends to overestimate vehicle acceleration at higher speeds. In the current study, incomplete acceleration traces (limited EDR recording
duration and final speed not reached due to being impacted) limited our ability to test this theory on the pre-crash data. In general, it is important to note that these models are only intended to simulate vehicle accelerations into the intersections.

A number of additional factors that were not considered in this study could influence vehicle acceleration rates. These include, but are not limited to, driver characteristics, vehicle characteristics, and roadway characteristics. It is important to recognize that I-ADAS systems do not currently consider factors such as intersection characteristics or driver characteristics. Future systems may be capable of adjusting based on specific factors. One factor not considered was roadway speed limit. Intuitively, we know that the acceleration of a driver speeding up to 45 mph will be different than that of a driver speeding up to 25 mph. What is not clear is if the differences lie in the acceleration magnitudes, acceleration duration, or both. Additional work on a larger dataset would help elucidate the effect of speed limit on intersection traversal kinematics. A second factor not considered was the size of the intersection. Drivers may be more likely to accelerate at higher rates if attempting to traverse a multilane high-speed road when compared to traveling through a smaller 4-way stop-controlled intersection. A third effect not considered was driver age and gender. Driver age and gender have both been correlated to more aggressive driving behavior and risk-taking [27], [34], [37], [38], [52], [53]. In particular, the young, male driving population has been consistently found to exhibit aggressive driving behavior. A fourth effect not considered was vehicle body type. There is obvious interplay between vehicle mass and engine power, among other factors, which dictate an individual vehicle’s acceleration capacity. This study only considers cars, light trucks, and vans, and does not distinguish between the groups in the developed models. A fifth effect not considered was other drivers on the roadway. For example, if the driver was following some other vehicles on the roadway, their ability to accelerate may be restricted [21]. A summary of the dataset composition can be found in the Appendix.

There are several limitations that should be considered with regards to the dataset used in the current study. First, because NASS/CDS exclusively contains tow away crashes, the dataset may be biased toward more severe crashes. Additionally, for an EDR record to be considered in this study, the vehicle must have experienced either an airbag deployment and/or a delta-v greater than 5-mph. This may limit the generalizability of the results to less severe crashes. Second, it is important to recognize that the drivers included in this study did not complete their intersection traversal due to either (a) being impacted or (b) initiating crash avoidance braking. As such, caution should be used if attempting to use these models to simulate a full acceleration profile up to some desired speed.

V. CONCLUSIONS

This paper presents novel acceleration profile models that are representative of pre-crash acceleration behavior. These models were developed using EDR-recorded pre-crash data taken from vehicles involved in U.S. intersection crashes. A comparison of non-crash acceleration behavior in normal driving with pre-crash driver behavior suggests that drivers in intersection crashes may tend to accelerate more aggressively than drivers in non-crash traversals. These findings
have important implications for designers of I-ADAS and for evaluators of active safety technology. Accounting for how drivers accelerate prior to real-world crashes is an important component for ensuring the effectiveness of I-ADAS in the U.S. vehicle fleet.

**ACKNOWLEDGMENT**

The authors give special thanks to Katsuhiko Iwazaki of Toyota for sharing their technical insights and expertise throughout the project. The authors would also like to thank Dr. Ada Tsoi and Dr. Kristofer Kusano for providing assistance in interpreting the EDR data.

**REFERENCES**


John M. Scanlon (M’14) received the Ph.D. degree in biomedical engineering from Virginia Tech in 2017. He is currently a Research Engineer with the Virginia Tech Center for Injury Biomechanics. He has authored extensively on driver behavior, event data recorders, injury biomechanics, and the potential effectiveness of driver assistance systems within the vehicle fleet. His interests are in crash reconstruction, crash causation analytics, and the research, development, and evaluation of automated vehicles. He is a member of the Association for the Advancement of Automotive Medicine, the Biomedical Engineering Society, and the Society of Automotive Engineers.

Rini Sherony (M’04) received the master’s degree in electrical engineering. She is currently a Senior Principal Engineer with the Toyota’s Collaborative Safety Research Center, and also a part of Toyota Motor North America, Ann Arbor, MI, USA. She has been with Toyota for the past 18 years, where she was involved in active safety research, system design, evaluation/planning, and data analysis. She is the lead for active safety, active and passive safety integration collaboration research, and data analysis. Her responsibilities include the development of standardized test procedures, test targets, testing, benefit estimation for active safety systems, such as pre-collision, pedestrian pre-collision system, and lane departure warning system. She is actively involved in SAE Active Safety committees developing pedestrian target standards. She is a member of the Association for the Advancement of Automotive Medicine.

Hampton C. Gabler (M’81) received the Ph.D. degree in mechanical and aerospace engineering from Princeton University. He is currently the Samuel Herrick Professor of Engineering with the Virginia Polytechnic Institute and State University (Virginia Tech). He is the Chair for Graduate Studies with the Department of Biomedical Engineering and Mechanics, Virginia Tech, and an Associate Director of the Center of Injury Biomechanics. He has active research programs underway in active safety systems, autonomous vehicle safety, vehicle crashworthiness, injury biomechanics, and event data recorders. He has authored over 150 technical papers and a book in these areas of research. Dr. Gabler is a fellow of the Society of Automotive Engineers (SAE), the Association for the Advancement of Automotive Medicine, and the American Institute for Medical and Biological Engineering. He is an Associate Editor of the Journal Traffic Injury Prevention and SAE International Journal of Transportation Safety.