Analysis of Discretionary Lane Changing Parameters on Freeways

Esmaiel Balal1, Ruey Long Cheu2, Thompson Gyan-Sarkodie3 and Jessica Miramontes4
1 Department of Civil Engineering, The University of Texas at El Paso, 500 W. University Ave, El Paso, TX 79968, USA, ebalalvarnosfaderani@miners.utep.edu
2 Department of Civil Engineering, The University of Texas at El Paso, 500 W. University Ave, El Paso, TX 79968, USA, rcheu@utep.edu
3 Department of Electrical & Computer Engineering, The University of Texas at El Paso, 500 W. University Ave, El Paso, TX 79968, USA, tsarkodi@utep.edu
4 Department of Civil Engineering, The University of Texas at El Paso, 500 W. University Ave, El Paso, TX 79968, USA, jlmiramontes@miners.utep.edu

ABSTRACT
A lane changing event, at the instant when a vehicle crosses the lane marker, involves up to five vehicles: the subject vehicle, preceding and following vehicles in the original lane, and the preceding and following vehicles in the target lane. Understanding the interactions of the subject vehicle with the surrounding vehicles is fundamental to the study of the safety and modeling of lane changing behavior. This research studies the statistical properties of 10 lane changing parameters. These parameters describe the gaps, times to collision between vehicles and the subject vehicle’s speed. The parameter values were extracted from the vehicle trajectory data in the Next Generation Simulation data sets. The results show that (i) all the parameters are positively correlated with each other; (ii) the gaps and distance are best described by the log-normal distribution; (ii) the times to collision are best described by the Laplace distribution; (iii) the speed may be described by the log-logistic distribution and the normal distribution. The results suggest that using one or few selected parameters may be sufficient to quantify the risk of a lane changing event.

1. INTRODUCTION
A vehicle’s two-dimensional motion on a highway surface may be decomposed into the longitudinal and lateral movements. The longitudinal movement in the same lane, in the presence of a vehicle ahead (the preceding vehicle) and/or a vehicle behind...
(the following vehicle), is termed car-following. On the other hand, the lateral movement, which is always accompanied with a longitudinal movement, is known as lane changing. Although car-following has been studied by researchers in more than 50 years, relatively fewer investigations on lane changing have been made. The reason could be due to the facts that (i) a lane change involves two-dimensional motions; and (ii) there are relatively more vehicles involved in a lane changing event. The above reasons have made the study of lane change more complex and challenging.

Lane changing model is as important as car-following model as the fundamental building blocks in microscopic traffic simulation tools [1–5]. Both models govern the second-to-second behavior of vehicles in a simulated road network. The microscopic driving behavior is also related to macroscopic property of traffic flow [6, 7]. Therefore, accurate modeling of a driver’s lane changing process is essential in producing realistic microscopic and macroscopic output of traffic conditions.

In general, there are two types of lane change in freeways: mandatory and discretionary. Mandatory lane change is also known as forced or necessary lane change. It usually occurs when a vehicle is trying to move from the left or center lane to the rightmost lane in order to exit the freeway. Mandatory lane change also happens when a vehicle has just entered the freeway from an on-ramp and is trying to move to the center or left lane to travel at a faster speed or to avoid a downstream exit lane.

Discretionary lane change is also known as free lane change or desired lane change. It normally occurs when a driver is following another vehicle at a speed slower than his/her desired speed and therefore seeks to increase its speed by moving to an adjacent lane. Obviously, the motivations and resulting driving behavior for the two types of lane change are different. Therefore a driver is expected have different decision rules or parameters for the two types of lane change.

Some researchers model lane change as a two-step process: (1) the decision to change lane; and (2) the execution of the decision [8]. As discussed above, the decision to change lane and the actual maneuver are different for mandatory and discretionary lane changes.

The objective of this research is to perform statistical analysis of selected parameters that describe the microscopic interaction of vehicles at a critical point of a discretionary lane changing maneuver on freeways. The critical point is the instant when the front center of a subject vehicle crosses the lane markers. Lane changing events were extracted from the well-known Next Generation Simulation (NGSIM) data sets collected at I-80 Freeway in Emeryville, California and U.S. Highway 101 in Los Angeles, California. The parameter values were computed from vehicle trajectory information. The statistical properties and probability distributions of these parameters were analyzed to draw conclusions on the vehicle interactions.

The organization of this paper is as follows. After this introduction, the lane changing models in popular microscopic traffic simulation tools are reviewed. The review also covers other lane changing models developed through traffic flow research. The literature review enabled the authors to identify the important decision parameters
used in the different models. The selected parameters are defined in the next section. Then, the steps in data processing are described. The results of statistical analysis are next presented, followed by conclusions.

2. LITERATURE REVIEW

This section reviews the lane changing models found in popular microscopic traffic simulation tools and in published articles, with the purpose of identifying the drivers’ decision parameters. These publications describe the models in varying level of details. The use of many different terms (in describing the types of lane change and their decision parameters), some with unspecified unit, have made the review challenging. The authors believe that these difficulties arise because research in lane change is still at its infancy. Nevertheless, the lane changing models are summarized here with the authors’ interpretations of the available information. This review focuses on discretionary lane change and the parameters used to make a decision. The authors have found that, in some models, it was impossible to distinguish discretionary lane change from mandatory lane change. The authors have also tried to separate a lane change into a two-step process of decision and execution. Again, in some models it was impossible to distinguish the two steps. Therefore, all the parameters that affect both decision and implementation of a discretionary lane change are reviewed to the best possible extent. The parameters are highlighted in italic font and not necessarily following the original terms used.

2.1. Lane Changing Models in Microscopic Traffic Simulation Tools

The lane changing model in FRESIM [1] is described in its predecessor INTRAS’s development report [9]. There are two types of lane change in FRESIM: free lane change and forced lane change. A free lane change is sought when a subject vehicle is traveling below its desired speed and it can gain speed by moving to an adjacent lane. A binary decision to change lane is generated according to a pre-defined probability and assigned to the subject vehicle. Once a decision has been made to change lane, the subject vehicle must check that the lead time to collision and lag time to collision in the target lane satisfy their respective “non-collision constraint”.

VISSIM [2] classifies lane changes into free lane change and necessary lane change. In the case of a free lane change, the lane changing model checks if the available lag time to collision between the subject vehicle and the following vehicle in the target lane satisfies the “desired safety distance” and “minimum time headway”. For a lane change in a queue, the model also checks the lead time to collision between the subject vehicle and the preceding vehicle in the target lane. PARAMICS [3] does not distinguish between mandatory lane change and discretionary lane change. The lane changing model in PARAMICS is based on the gap acceptance theory [10]. A vehicle is allowed to move from its original lane to the target lane if both (i) the front gap (in distance unit) between the subject vehicle and the preceding vehicle in the target lane; and (ii) the rear gap between the subject vehicle and the following vehicle in the target lane exceed their respective threshold value. AIMSUN [4] describes a vehicle’s lane changing decision making process in terms of necessity,
desirability and possibility to change lane. The necessity to change lane includes the need to overtake the existing leader to travel at a faster speed. If it is necessary to change lane, the logic checks if an adjacent lane’s speed is desirable. If this condition is met, the logic next looks for a gap in the target lane to make a safe maneuver. To distinguish between discretionary and mandatory lane changes, AIMSUN divides a freeway segment upstream of an off-ramp into three zones, where discretionary lane changes take place in the most upstream zone. TransModeler [5] uses the discrete choice approach to model a driver’s lane changing decision. It considers three types of lane change: discretionary, mandatory and forced lane changes. A discretionary lane change is considered when a driver is dissatisfied with the current speed. There are two discretionary lane change models: neighboring lane model and target lane model. The neighboring lane model, as its name suggests, has the target lanes adjacent to the original lane. In contrast, the target lane model moves the subject vehicle by more than one lane. In the neighboring lane model, the logit model calculates the probabilities of a driver selecting the left or right adjacent lane. Among the attributes in the utility function are average speed gain and slow vehicle ahead. Once a target lane has been selected, the subject vehicle seeks a suitable gap in the target lane to merge into. The gap acceptance attributes considered are lead time to collision and lag time to collision. The coefficients of the gap acceptance parameters have been calibrated with NGSIM data.

2.2. Other Lane Changing Studies
Gipps [11] is perhaps one of the earliest to document a lane change study in a signalized street. The driver’s decision making framework consists of the possibility, necessity and desirability to change lane. He then proposed a lane changing model encompassing mandatory and discretionary lane changes. The decision parameters for discretionary lane change included the subject vehicle’s safe speed, relative speed between the original lane and the target lane, and time to collision between preceding and following vehicle. The decision making framework is later used in AIMSUN. McDonald et al. [12], Blackstone et al. [13] and Wu et al. [14] described different motivations for lane change: pressure from the rear (fast approaching vehicle) and to gain speed. The pressure from the rear motivates the subject vehicle to move to a slower lane, while the intention to gain speed makes the subject vehicle to want move to a faster lane. Fuzzy rules are constructed to make use of lag times to collision of the original and target lanes as inputs. They did not distinguish between mandatory and discretionary lane changes. Das et al. [15] proposed a three-stage fuzzy logic discretionary lane changing model. The first stage is to decide if a subject vehicle should change lane. There are two inputs to the fuzzy rules: the subject vehicle’s speed and rear gap in the target lane. The second stage of the model looks for a suitable gap. The inputs to the fuzzy rules are speeds in the original and target lanes, front gap and rear gap in the target lane, and front gap in the original lane. The rules in the third stage consider the front gap and rear gap in the target lane and speed in the target lane. After observing video recordings of 73 lane
changing maneuvers in arterials in Sydney, Australia, Hidas [16] classified lane changes into free, forced and cooperative lane changes based on the front gap and rear gap in the target lane. Regardless of the type of lane change, the logic proposed by Hidas [16] makes use of front gap and rear gap in the target lane. Kesting et al. [17] used a linear combination of accelerations of the subject vehicle, the follower in the original lane and the follower in the target lane to form an incentive criterion for a lane change. This lane changing model is based purely on acceleration rates. Yeo et al. [18] proposed an oversaturated freeway flow algorithm which consists of a lane change model. The algorithm has two types of lane change: mandatory and discretionary. The purpose of a discretionary lane change is for the subject vehicle to increase speed or to improve its position in the traffic stream. The parameters for discretionary lane change are relative speed, speed of vehicles in the target lane, and speed of the subject vehicle. Moridpour et al. [19, 20] proposed two fuzzy logic models for heavy vehicles changing into a slower lane and a faster lane, respectively. The models used the following parameters: front gap (in the current lane), rear gap (in the current lane), lag space gap (in the target lane), speeds (in the current lane and target lane), relative speeds (in the current lane, between the subject vehicle and the vehicle ahead and behind), lag relative speed (in the target lane). If-then rules are applied to the inputs to make inference about a heavy vehicle driver’s lane changing decision. The models have been validated using the NGSIM data sets. Schakel et al. [21] combined incentives to follow a route, to gain speed and to keep right into a single lane change desire value, from which three types of lane change (free, synchronized and cooperative) are distinguished. The proposed lane changing model, which is based on the gap acceptance concept, includes seven parameters: relax headway, route desire, anticipated speed, speed desire, keep-right desire, combine desires, and gap-acceptance. The model has been calibrated with loop detector data collected at the A20 Motorway near Rotterdam, Netherlands. Hill and Elefteriadou [22] studied the lane changing behavior of drivers in instrumented vehicles driving on I-4 Freeway in Orlando, Florida, and I-95 Freeway in Jacksonville, Florida. The time for a lane changing maneuver, desired speed, front gap (in the target lane) and rear gap (in the target lane) were recorded for 321 discretionary lane changes. They found that the Gamma distribution provided the best fit for the rear gap. However, the Johnson SI distribution provided the best fit for the front gap. Other relatively more theoretical lane changing models have been reviewed by Zheng [23].

2.3. Lane Changing Parameters

After conducting the above review, the lane changing parameters are summarized in Table 1. This table does not include parameters that are used only in one model. The six parameters listed in Table 1 may be derived from the NGSIM vehicle trajectory data. From Table 1, it appears that most of the microscopic traffic simulation model use times to collision among the parameters, while gaps are more popular among the other lane changing models. Speed appears in all except three models listed in the table.
3. PARAMETERS STUDIED

This section defines the parameters that describe the interactions of vehicles during the critical instant of a lane change execution. These parameters are also referred to as lane changing parameters in this paper. Consider a typical lane changing scenario as depicted in Figure 1. Figure 1 shows the critical instant when front center of the subject vehicle $S$ crosses the lane markers between the original lane and the target lane. The vehicle in front of $S$ in the original lane is called the preceding vehicle before lane change, denoted as $PB$.

\begin{table}
\centering
\caption{Summary of lane changing parameters reviewed}
\begin{tabular}{lccccccc}
\hline
Simulation model and/or reference & Lead time & Lead time & Speed & Relative \\
& & to collision & to collision & & speed \\
& & (time) & (time) & & \\
FRESIM [1, 9] & Yes & Yes & Yes & \\
VISSIM [2] & Yes & Yes & \\
PARAMICS [3, 10] & Yes & Yes & Yes & \\
AIMSUN [4] & Yes$^\#$ & Yes$^\#$ & Yes & \\
TransModeler [5] & Yes & Yes & Yes & \\
McDonald et al. [12], & Yes & Yes & Yes & \\
Brackstone et al. [13], & & & & \\
Wu et al. [14] & & & & \\
Das et al. [15] & Yes & Yes & Yes & \\
Hidas [16] & Yes & Yes & Yes & \\
Yeo et al. [18] & Yes & Yes & Yes & \\
Moridpour et al. [19, 20] & Yes & Yes & Yes & \\
Schakel et al. [21] & Yes & Yes & Yes & \\
Hill and Elefteriadou [22] & Yes & Yes & \\
\hline
\multicolumn{7}{l}{$^\#$Gap (distance) between preceding and following vehicles is the sum of front gap and rear gap}
\end{tabular}
\end{table}

Figure 1. Vehicles and their positions during a lane change
The vehicle behind S in the original lane is called the following vehicle before lane change, denoted as FB. After the lane change, the subject vehicle inserts itself in the target lane between the preceding vehicle (denoted as PA) and the following vehicle (denoted as FA). The longitudinal positions of S, PB, FB, PA, FA, measured with reference to the center of each vehicle, are represented by $Y_S$, $Y_{PB}$, $Y_{FB}$, $Y_{PA}$, $Y_{FA}$, respectively. The lengths of S, PB, FB, PA, FA are denoted as $L_S$, $L_{PB}$, $L_{FB}$, $L_{PA}$, $L_{FA}$ respectively.

The following lane changing parameters are defined in this study:

**Front gap before lane change** (in meters):

$$G_{PB} = \left( Y_{PB} - \frac{1}{2} L_{PB} \right) - \left( Y_S + \frac{1}{2} L_S \right), \quad G_{PB} \geq 0 \quad (1)$$

**Rear gap before lane change** (in meters):

$$G_{FB} = \left( Y_S - \frac{1}{2} L_S \right) - \left( Y_{FB} + \frac{1}{2} L_{FB} \right), \quad G_{FB} \geq 0 \quad (2)$$

**Front gap after lane change** (in meters):

$$G_{PA} = \left( Y_{PA} - \frac{1}{2} L_{PA} \right) - \left( Y_S + \frac{1}{2} L_S \right), \quad G_{PA} \geq 0 \quad (3)$$

**Rear gap after lane change** (in meters):

$$G_{FA} = \left( Y_S - \frac{1}{2} L_S \right) - \left( Y_{FA} + \frac{1}{2} L_{FA} \right), \quad G_{FA} \geq 0 \quad (4)$$

**Lead time to collision before lane change** (in seconds):

$$T_{PB} = \frac{G_{PB}}{V_S - V_{PB}}, \quad -\infty \leq T_{PB} \leq +\infty \quad (5)$$

**Lag time to collision before lane change** (in seconds):

$$T_{FB} = \frac{G_{FB}}{V_{FB} - V_S}, \quad -\infty \leq T_{FB} \leq +\infty \quad (6)$$

**Lead time to collision after lane change** (in seconds):

$$T_{PA} = \frac{G_{PA}}{V_S - V_{PA}}, \quad -\infty \leq T_{PA} \leq +\infty \quad (7)$$

**Lag time to collision after lane change** (in seconds):

$$T_{FA} = \frac{G_{FA}}{V_{FA} - V_S}, \quad -\infty \leq T_{FA} \leq +\infty \quad (8)$$

**Distance** (in meters):

$$D = \left( Y_{PA} - \frac{1}{2} L_{PA} \right) - \left( Y_{FA} + \frac{1}{2} L_{FA} \right), \quad D \geq 0 \quad (9)$$

The speed of the subject vehicle $V_S$ (in meter/second) is also analyzed.
In defining the gap \((G)\) and time to collision \((T)\), the subscript \(P\) denotes the preceding vehicle, \(F\) denotes the following vehicle; while \(B\) represents the lane before lane change (the original lane), and \(A\) represents the lane after lane change (the target lane). The gaps and times to collision before a lane change are added in the analysis so as to study the proximity of the three associated vehicles \((S, PB, FB)\) immediately before the subject vehicle leaves its original lane. The relative speed in Table 1 is not studied separately as it has been included in the calculation of times to collision. The \(G_{PA}\) and \(G_{FA}\) as defined in eqns (3) and (4) are similar to the gaps in [22], except that in [22] the gaps was most likely measured when the subject vehicle began its lateral move.

4. DATA PROCESSING

The data used in this analysis was taken from the well-known NGSIM database. This database includes vehicle trajectory data collected at a segment of I-80 Freeway in Emeryville, California [24] and a segment of U.S. Highway 101 in Los Angeles, California [25]. For each freeway segment, vehicle motions were captured by several video cameras position on top of a tall building. The video images were post-processed to extract vehicle trajectory data at 0.1 second intervals.

The I-80 data was collected over a 1650 ft segment, in the northbound direction between Powell Street on-ramp and Ashby Street off-ramp. This segment of the freeway has six through lanes between the ramps. The available data was collected on April 13, 2005 from 4:00–4:15 p.m., 5:00–5:15 p.m. and 5:15–5:30 p.m. In this study, the data from 4:00–4:15 p.m. was used because it has the highest number of lane changes among the three 15-minute periods.

The U.S. 101 data was collected over a 2100 ft segment, in the southbound direction between Ventura Boulevard on-ramp and Cahuenga Boulevard off-ramp. This segment of the freeway also has six through lanes between the ramps. The available data was collected on June 15, 2005 from 7:50–8:05 a.m., 8:05–8:20 a.m. and 8:20–8:35 a.m. In this study, the data from 7:50–8:05 a.m. was used because it has the highest number of lane changes among the three 15-minute periods. The vehicle trajectory data was processed as follows:

- Only passenger cars were selected as the subject vehicles. Trucks and motorcycles, which were expected to have different lane changing behavior, and also have small sample sizes, were not considered.
- Only the subject vehicles originally travelled in lanes 2, 3 and 4 were considered. They were assumed to make discretionary lane changes. Vehicles in lanes 5 and 6 were not considered so as to eliminate the possibility of drivers executing mandatory lane changes after entering from the upstream on-ramp or to exit at the downstream off-ramp. Vehicles in lane 1 were not considered so as to eliminate the interference caused by any high occupancy vehicle lane. Although this procedure did not necessarily filter out all the mandatory lane changes in lanes 2, 3 and 4 within the data collection segment, this was the best guess that could be made from the NGSIM data.
• Vehicles making multiple lane changes were excluded in the analysis, so as to exclude a possible mandatory move.

• For each identified subject vehicle, the time \( t \) when the lane changing event occurred was taken as the time when the front center of the subject vehicle crossed the lane markers. This time reference \( t \) was used because (i) the width of vehicle is not provided in the NGSIM data; and (ii) it was impossible to determine from the NGSIM data when exactly a driver psychologically made his/her decision to change lane.

• At time instant \( t \), the preceding and following vehicles before and after the lane change were identified. The positions of up to the five vehicles involved in a lane changing event at \( t-0.4 \), \( t-0.3 \), \( t-0.2 \), \( t-0.1 \) and \( t \) seconds were extracted. According to Punzo et al. [26], vehicle speed derived from NGSIM vehicle trajectory data must be calculated from the vehicle’s coordinates, so that the positions and speeds are consistent. Therefore, the front center coordinates of the vehicles and vehicle lengths provided in the NGSIM data were used to calculate the centroids which were then used to compute the studied parameters.

• For each lane changing event, the parameter values were calculated at \( t-0.4 \), \( t-0.3 \), \( t-0.2 \), \( t-0.1 \) and \( t \) seconds respectively. The average values from \( t-0.4 \) to \( t \) seconds were used as the representative value of a parameter. The reasons for taking the average value over 0.5 second are (i) to reduce the error caused by using instantaneous values in the NGSIM data; (ii) to be more consistent with the human perception time; and (iii) to be consistent with other researches that used NGSIM data [27]. Average value of a longer duration was not considered as it will loss the “instantaneous” property of the calculated values.

• Because the calculated time to collision were in the theoretical range of \((-\infty, +\infty)\), some times to collision were unreasonably large. Therefore, they were truncated to \([-200,200]\) seconds.

Because a lane changing event involves up to five vehicles (as shown in Figure 1), not all the five vehicles may be captured by the video cameras. Therefore, it may not be possible to calculate all the 10 parameters for a lane changing event from the available NGSIM data. For example, if a subject vehicle changed lane near the downstream end of a segment, the preceding vehicles (\(PB\) and \(PA\)) may already have left the camera view. In this case, it is impossible to calculate the parameters associated with these two vehicles. Therefore, the computed parameters have different sample sizes. The NGSIM data was processed by means of MATLAB [28]. The descriptive statistics of each parameter are summarized in Table 2.

5. RESULTS AND DISCUSSIONS

5.1. Descriptive Statistics

Table 2 lists the descriptive statistics of the 10 parameters analyzed. The gaps and distance are processed to 0.001 m precision, times to collision are processed to 0.1 second precision while speed is processed to 0.01 m/s precision. For the same parameter, the mean and maximum values obtained from the I-80 data set are smaller than the
corresponding values in the U.S. 101 data set. For example, for \( G_{FA} \), the rear gap after lane change, the I-80 data set has a mean of 13.500 m while the U.S. 101 data set has a mean of 28.300 m. This is because the traffic condition in the I-80 data set was more congested than the traffic condition in the U.S. 101 data set. In the I-80 data set, the traffic volume was 8144 vph (1357 vphpl) and the average space mean speed was 17.86 mph. This is equivalent to a density of 75.98 veh/mi/ln. In the U.S. 101 data set, the volume was 8642 vph (1440 vphpl) and the average space mean speed was 25.66 mph, equivalent to a density of 56.12 veh/mi/ln.

### 5.2. Correlation Analysis

A correlation analysis was performed for all the parameters in each data set. The purpose of the correlation analysis was to examine if there is any strong relationship between any two parameters so that some of the parameters may be excluded in the analysis of lane changing events in the future. In a correlation analysis, all the variables must have the same sample size and be paired. The data was then filtered such that only the lane changes which yielded all the parameter values were used in the correlation analysis. This filtering resulted in sample sizes of 314 for the I-80 data set and 394 for the U.S. 101 data set. The correlation coefficients, or \( r \) values, calculated by Minitab [29], are presented in Table 3. All the \( r \) values are significantly different from 0, with \( p \)-values all less than 0.001.

### Table 2. Descriptive statistics of lane changing parameters

#### (a) I-80 Data 4:00 p.m. to 4:15 p.m.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( G_{PB} )</th>
<th>( G_{FB} )</th>
<th>( G_{PA} )</th>
<th>( G_{FA} )</th>
<th>( T_{PB} )</th>
<th>( T_{FB} )</th>
<th>( T_{PA} )</th>
<th>( T_{FA} )</th>
<th>( D )</th>
<th>( V_s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>m</td>
<td>m/s</td>
</tr>
<tr>
<td>Sample size</td>
<td>493</td>
<td>463</td>
<td>382</td>
<td>501</td>
<td>495</td>
<td>435</td>
<td>492</td>
<td>501</td>
<td>455</td>
<td>592</td>
</tr>
<tr>
<td>Min value</td>
<td>0.759</td>
<td>0.118</td>
<td>0.201</td>
<td>0.016</td>
<td>–198.0</td>
<td>–177.5</td>
<td>–100.3</td>
<td>–190.1</td>
<td>4.670</td>
<td>1.16</td>
</tr>
<tr>
<td>Max value</td>
<td>67.185</td>
<td>81.136</td>
<td>53.269</td>
<td>109.070</td>
<td>192.9</td>
<td>197.0</td>
<td>157.6</td>
<td>185.5</td>
<td>152.510</td>
<td>21.87</td>
</tr>
<tr>
<td>Mean</td>
<td>16.716</td>
<td>16.086</td>
<td>11.880</td>
<td>13.500</td>
<td>–0.3</td>
<td>0.6</td>
<td>1.4</td>
<td>–0.1</td>
<td>28.840</td>
<td>8.63</td>
</tr>
<tr>
<td>Std deviation</td>
<td>11.269</td>
<td>12.650</td>
<td>9.569</td>
<td>11.160</td>
<td>–0.3</td>
<td>0.6</td>
<td>1.4</td>
<td>–0.1</td>
<td>28.840</td>
<td>8.63</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.872</td>
<td>2.233</td>
<td>1.738</td>
<td>2.880</td>
<td>–0.2</td>
<td>–0.2</td>
<td>1.2</td>
<td>–0.2</td>
<td>2.536</td>
<td>0.59</td>
</tr>
</tbody>
</table>

#### (b) U.S. 101 Data 7:50 a.m. to 8:05 a.m.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>( G_{PB} )</th>
<th>( G_{FB} )</th>
<th>( G_{PA} )</th>
<th>( G_{FA} )</th>
<th>( T_{PB} )</th>
<th>( T_{FB} )</th>
<th>( T_{PA} )</th>
<th>( T_{FA} )</th>
<th>( D )</th>
<th>( V_s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>m</td>
<td>m/s</td>
</tr>
<tr>
<td>Sample size</td>
<td>503</td>
<td>442</td>
<td>355</td>
<td>510</td>
<td>475</td>
<td>415</td>
<td>469</td>
<td>482</td>
<td>428</td>
<td>587</td>
</tr>
<tr>
<td>Min value</td>
<td>1.820</td>
<td>2.279</td>
<td>0.130</td>
<td>0.998</td>
<td>–183.1</td>
<td>–153.8</td>
<td>–164.6</td>
<td>–142.6</td>
<td>14.620</td>
<td>4.30</td>
</tr>
<tr>
<td>Max value</td>
<td>173.680</td>
<td>83.232</td>
<td>103.980</td>
<td>171.690</td>
<td>168.4</td>
<td>161.6</td>
<td>168.7</td>
<td>199.0</td>
<td>164.190</td>
<td>27.79</td>
</tr>
<tr>
<td>Mean</td>
<td>27.050</td>
<td>18.661</td>
<td>18.730</td>
<td>28.300</td>
<td>–3.3</td>
<td>1.4</td>
<td>–1.1</td>
<td>–2.5</td>
<td>45.160</td>
<td>15.54</td>
</tr>
<tr>
<td>Std deviation</td>
<td>18.380</td>
<td>13.762</td>
<td>14.650</td>
<td>25.940</td>
<td>43.6</td>
<td>40.1</td>
<td>33.5</td>
<td>43.0</td>
<td>23.310</td>
<td>3.46</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.885</td>
<td>1.724</td>
<td>2.225</td>
<td>2.521</td>
<td>0.5</td>
<td>–0.1</td>
<td>–0.2</td>
<td>0.7</td>
<td>1.647</td>
<td>–0.10</td>
</tr>
</tbody>
</table>
The minimum $r$ value in the U.S. 101 data set is 0.832. This indicates that all the 10 parameters in the U.S. 101 data set are strongly correlated. However, this is not the case for the I-80 data set. The $r$ values smaller than 0.700 in the I-80 data set are highlighted in red in Table 3(a). From this table, it can be observed that $TPB$ (lead time to collision before lane change) does not have a strong correlation with any other parameters. In addition, $TFA$ (lag time to collision after lane change) is also not strongly correlated with gaps and distance $D$. The differences in the correlation matrices between the two data sets are indications that drivers in these two sites have different lane changing behavior.

Figure 2 shows the scatter plots of the parameters produced by Minitab [29]. The scatter plots of two parameters in each data set are presented in a 10 by 10 matrix. The diagonal elements of the matrix indicate the parameter names in the horizontal and vertical axles. The scatter plots visualize the correlations as listed in Table 3.

### Table 3. Correlation matrices of lane changing parameters

#### (a) I-80 Data 4:00 p.m. to 4:15 p.m.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$G_{PB}$</th>
<th>$G_{FB}$</th>
<th>$G_{PA}$</th>
<th>$G_{FA}$</th>
<th>$TP_{PB}$</th>
<th>$TP_{FB}$</th>
<th>$TP_{PA}$</th>
<th>$TP_{FA}$</th>
<th>$D$</th>
<th>$V_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_{PB}$</td>
<td>1</td>
<td>0.996</td>
<td>0.992</td>
<td>0.971</td>
<td>0.613</td>
<td>0.791</td>
<td>0.882</td>
<td>0.677</td>
<td>0.973</td>
<td>0.894</td>
</tr>
<tr>
<td>$G_{FB}$</td>
<td>0.996</td>
<td>1</td>
<td>0.996</td>
<td>0.985</td>
<td>0.573</td>
<td>0.794</td>
<td>0.876</td>
<td>0.689</td>
<td>0.970</td>
<td>0.915</td>
</tr>
<tr>
<td>$G_{PA}$</td>
<td>0.992</td>
<td>0.996</td>
<td>1</td>
<td>0.990</td>
<td>0.538</td>
<td>0.799</td>
<td>0.873</td>
<td>0.696</td>
<td>0.964</td>
<td>0.928</td>
</tr>
<tr>
<td>$G_{FA}$</td>
<td>0.971</td>
<td>0.985</td>
<td>0.990</td>
<td>1</td>
<td>0.458</td>
<td>0.802</td>
<td>0.856</td>
<td>0.722</td>
<td>0.942</td>
<td>0.954</td>
</tr>
<tr>
<td>$TP_{PB}$</td>
<td>0.613</td>
<td>0.573</td>
<td>0.538</td>
<td>0.458</td>
<td>1</td>
<td>0.426</td>
<td>0.605</td>
<td>0.284</td>
<td>0.675</td>
<td>0.358</td>
</tr>
<tr>
<td>$TP_{FB}$</td>
<td>0.791</td>
<td>0.794</td>
<td>0.799</td>
<td>0.802</td>
<td>0.426</td>
<td>1</td>
<td>0.958</td>
<td>0.971</td>
<td>0.783</td>
<td>0.895</td>
</tr>
<tr>
<td>$TP_{PA}$</td>
<td>0.882</td>
<td>0.876</td>
<td>0.873</td>
<td>0.856</td>
<td>0.605</td>
<td>0.958</td>
<td>1</td>
<td>0.897</td>
<td>0.871</td>
<td>0.893</td>
</tr>
<tr>
<td>$TP_{FA}$</td>
<td>0.677</td>
<td>0.689</td>
<td>0.696</td>
<td>0.722</td>
<td>0.284</td>
<td>0.971</td>
<td>0.894</td>
<td>1</td>
<td>0.659</td>
<td>0.853</td>
</tr>
<tr>
<td>$D$</td>
<td>0.973</td>
<td>0.970</td>
<td>0.964</td>
<td>0.942</td>
<td>0.675</td>
<td>0.783</td>
<td>0.871</td>
<td>0.659</td>
<td>1</td>
<td>0.871</td>
</tr>
<tr>
<td>$V_S$</td>
<td>0.894</td>
<td>0.915</td>
<td>0.928</td>
<td>0.954</td>
<td>0.358</td>
<td>0.895</td>
<td>0.893</td>
<td>0.853</td>
<td>0.871</td>
<td>1</td>
</tr>
</tbody>
</table>

#### (b) U.S. 101 Data 7:50 a.m. to 8:05 a.m.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$G_{PB}$</th>
<th>$G_{FB}$</th>
<th>$G_{PA}$</th>
<th>$G_{FA}$</th>
<th>$TP_{PB}$</th>
<th>$TP_{FB}$</th>
<th>$TP_{PA}$</th>
<th>$TP_{FA}$</th>
<th>$D$</th>
<th>$V_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_{PB}$</td>
<td>1</td>
<td>0.974</td>
<td>0.995</td>
<td>0.993</td>
<td>0.964</td>
<td>0.873</td>
<td>0.951</td>
<td>0.947</td>
<td>0.932</td>
<td>0.982</td>
</tr>
<tr>
<td>$G_{FB}$</td>
<td>0.974</td>
<td>1</td>
<td>0.971</td>
<td>0.985</td>
<td>0.909</td>
<td>0.871</td>
<td>0.903</td>
<td>0.879</td>
<td>0.974</td>
<td>0.921</td>
</tr>
<tr>
<td>$G_{PA}$</td>
<td>0.995</td>
<td>0.979</td>
<td>1</td>
<td>0.994</td>
<td>0.967</td>
<td>0.874</td>
<td>0.951</td>
<td>0.948</td>
<td>0.938</td>
<td>0.975</td>
</tr>
<tr>
<td>$G_{FA}$</td>
<td>0.995</td>
<td>0.985</td>
<td>0.994</td>
<td>1</td>
<td>0.944</td>
<td>0.855</td>
<td>0.927</td>
<td>0.920</td>
<td>0.946</td>
<td>0.968</td>
</tr>
<tr>
<td>$TP_{PB}$</td>
<td>0.964</td>
<td>0.909</td>
<td>0.967</td>
<td>0.944</td>
<td>1</td>
<td>0.901</td>
<td>0.987</td>
<td>0.995</td>
<td>0.876</td>
<td>0.973</td>
</tr>
<tr>
<td>$TP_{FB}$</td>
<td>0.873</td>
<td>0.871</td>
<td>0.874</td>
<td>0.855</td>
<td>0.901</td>
<td>1</td>
<td>0.945</td>
<td>0.900</td>
<td>0.915</td>
<td>0.832</td>
</tr>
<tr>
<td>$TP_{PA}$</td>
<td>0.951</td>
<td>0.903</td>
<td>0.951</td>
<td>0.927</td>
<td>0.987</td>
<td>0.945</td>
<td>1</td>
<td>0.990</td>
<td>0.891</td>
<td>0.949</td>
</tr>
<tr>
<td>$TP_{FA}$</td>
<td>0.947</td>
<td>0.879</td>
<td>0.948</td>
<td>0.920</td>
<td>0.995</td>
<td>0.900</td>
<td>0.990</td>
<td>1</td>
<td>0.850</td>
<td>0.963</td>
</tr>
<tr>
<td>$D$</td>
<td>0.932</td>
<td>0.974</td>
<td>0.938</td>
<td>0.946</td>
<td>0.876</td>
<td>0.915</td>
<td>0.891</td>
<td>0.850</td>
<td>1</td>
<td>0.867</td>
</tr>
<tr>
<td>$V_S$</td>
<td>0.982</td>
<td>0.921</td>
<td>0.975</td>
<td>0.968</td>
<td>0.973</td>
<td>0.832</td>
<td>0.949</td>
<td>0.963</td>
<td>0.867</td>
<td>1</td>
</tr>
</tbody>
</table>

The minimum $r$ value in the U.S. 101 data set is 0.832. This indicates that all the 10 parameters in the U.S. 101 data set are strongly correlated. However, this is not the case for the I-80 data set. The $r$ values smaller than 0.700 in the I-80 data set are highlighted in red in Table 3(a). From this table, it can be observed that $TP_{PB}$ (lead time to collision before lane change) does not have a strong correlation with any other parameters. In addition, $T_{FA}$ (lag time to collision after lane change) is also not strongly correlated with gaps and distance $D$. The differences in the correlation matrices between the two data sets are indications that drivers in these two sites have different lane changing behavior.

Figure 2 shows the scatter plots of the parameters produced by Minitab [29]. The scatter plots of two parameters in each data set are presented in a 10 by 10 matrix. The diagonal elements of the matrix indicate the parameter names in the horizontal and vertical axles. The scatter plots visualize the correlations as listed in Table 3.
The overall results of the correlation analysis indicate that, at least eight of the 10 parameters are highly correlated. This implies that an analyst may be able to use a few representative parameters to describe a lane changing event, for example, $T_{PB}$, $T_{FA}$ plus one or two of the remaining parameters which are highly correlated. The high correlation coefficients are not surprising, because drivers of the subject vehicle tend to

Figure 2. Matrix plots of lane changing parameters

(a) I-80 Data

(b) U.S. 101 Data
position their vehicles in the middle of the preceding and following vehicles in the original lane in steady state car-following, and then insert their vehicles in the middle of the preceding and following vehicles in the target lane.

5.3. Probability Distributions
The processed data as reported in Table 2 had been fitted with probability distributions using @RISK [30]. For each parameter, at least 10 distributions have been considered. The Akaike Information Criterion (AIC) has been used to select the distributions that provide the best fit to the observed data. AIC is an indicator for the goodness of fit that takes into account the number of estimated distribution parameters. For each lane changing parameter, the top three distributions that best fit the observed data are listed in Table 4. All the distributions listed in Table 4 provide good fit to the data, with p-values all smaller than 0.01.

To illustrate how good the probability distributions fit the data, the following figures are provided. Figure 3 plots the histogram distribution of $G_{FA}$ taken from the I-80 data set and the fitted log-normal probability density function. Figure 4 plots the histogram distribution of $T_{FA}$ taken from the U.S. 101 data set and the fitted Laplace probability density function.

From the results of distribution fitting presented in Table 4, the lane changing parameters studied have different probability distributions that provide the best fit. It is preferably to have one probability distribution that can describe the gaps ($G_{PB}$, $G_{FB}$, $G_{PA}$, $G_{FB}$), times to collision ($T_{PB}$, $T_{FB}$, $T_{PA}$, $T_{FB}$), distance ($D$) and speed ($V_S$) respectively. From the table it is obvious that the Laplace distribution provides the best fit to all the times to collision. To select one probability distribution for the gaps, a numeric scoring system was used, in which the distributions that provide the best, second best and third best fits were assigned scores of three, two and one, respectively. The distribution that has the highest total score was recommended. Both the log-logistic and log-normal distributions have the same total score. The log-normal distribution is recommended because it appears in the top three lists for all the gap parameters. As for distance $D$, there is no clear winner. The log-normal distribution is the third best fit for the I-80 data but the fourth best fit for the U.S. 101 data. The inverse Gaussian distribution is the third best fit for the U.S. 101 data but the fourth best fit for the I-80 data. The log-normal distribution is preferred to describe $D$ because it is more commonly known. As for the subject vehicle’s speed $V_S$, the logistic distribution is selected as it appears in the top three distribution lists of the two data sets. The inconsistency in the distributions for $D$ and $V_S$ may be empirical evidence that not all the drivers are using these two parameters for making lane changing decisions.

The recommended probability distributions are listed in Table 4. The distribution parameters, calculated from the method of moment, are also listed in the table. Note that, the fitted distributions for $G_{PA}$ and $G_{FA}$ are different from the Gamma and Johnson SI distributions, respectively, found in [22].

The log-normal distribution has a probability density function of

$$f(x | \lambda, \zeta) = \frac{1}{\sqrt{2\pi}\zeta} e^{-\frac{1}{2}\left(\frac{\ln x - \lambda}{\zeta}\right)^2} \quad X > 0$$

where $\lambda$ ($\lambda > 0$) is the location parameter while $\zeta$ ($\zeta > 0$) is the scale parameter.
### Table 4. Fitted probability distributions of lane changing parameter

(a) I-80 Data 4:00 p.m. to 4:15 p.m.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$G_{PB}$</th>
<th>$G_{FB}$</th>
<th>$G_{PA}$</th>
<th>$G_{FA}$</th>
<th>$T_{PB}$</th>
<th>$T_{FB}$</th>
<th>$T_{PA}$</th>
<th>$T_{FA}$</th>
<th>$D$</th>
<th>$V_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>m</td>
<td>m/s</td>
</tr>
<tr>
<td>Best fit</td>
<td>Log-logistic</td>
<td>Log-logistic</td>
<td>Inverse gauss</td>
<td>Pearson 5</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Log-logistic</td>
<td>Log-logistic</td>
</tr>
<tr>
<td>2nd best fit</td>
<td>Pearson 5</td>
<td>Pearson 5</td>
<td>Log-normal</td>
<td>Log-logistic</td>
<td>Logistic</td>
<td>Normal</td>
<td>Logistic</td>
<td>Normal</td>
<td>Log-normal</td>
<td>Logistic</td>
</tr>
<tr>
<td>3rd best fit</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Log-normal</td>
<td>Gamma</td>
</tr>
<tr>
<td>Recommended</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Log-normal</td>
<td>Logistic</td>
</tr>
<tr>
<td>Log-normal location parameter $\lambda$</td>
<td>2.629</td>
<td>2.537</td>
<td>2.225</td>
<td>2.342</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>3.233</td>
<td>–</td>
</tr>
<tr>
<td>Laplace scale parameter $\xi$</td>
<td>0.612</td>
<td>0.694</td>
<td>0.707</td>
<td>0.722</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.508</td>
<td>–</td>
</tr>
<tr>
<td>Logistic location parameter $\mu$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–0.3</td>
<td>0.6</td>
<td>1.4</td>
<td>-0.1</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Logistic scale parameter $b$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>13.6</td>
<td>49.7</td>
<td>12.8</td>
<td>12.4</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Logistic location parameter $\mu$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>8.63</td>
</tr>
<tr>
<td>Logistic scale parameter $s$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.57</td>
</tr>
</tbody>
</table>
### (b) U.S. 101 Data 7:50 a.m. to 8:05 a.m.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$G_{PB}$</th>
<th>$G_{FB}$</th>
<th>$G_{PA}$</th>
<th>$G_{FA}$</th>
<th>$T_{PB}$</th>
<th>$T_{FB}$</th>
<th>$T_{PA}$</th>
<th>$T_{FA}$</th>
<th>$D$</th>
<th>$V_S$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>m</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>m</td>
<td>m/s</td>
</tr>
<tr>
<td>Best fit</td>
<td>Log-logistic</td>
<td>Inverse Gaussian</td>
<td>Pearson 5</td>
<td>Pearson 5</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Gamma</td>
<td>Normal</td>
</tr>
<tr>
<td>2nd best fit</td>
<td>Pearson 5</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Log-logistic</td>
<td>Logistic</td>
<td>Normal</td>
<td>Log-logistic</td>
<td>Weibull</td>
<td>Weibull</td>
</tr>
<tr>
<td>3rd best fit</td>
<td>Log-normal</td>
<td>Pearson 5</td>
<td>Inverse Gaussian</td>
<td>Log-logistic</td>
<td>Logistic</td>
<td>Laplace</td>
<td>Log-normal</td>
<td>Laplace</td>
<td>Log-normal</td>
<td>Logistic</td>
</tr>
<tr>
<td>Recommended</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Log-normal</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Laplace</td>
<td>Log-normal</td>
<td>Logistic</td>
</tr>
<tr>
<td>Log-normal location parameter $\lambda$</td>
<td>3.108</td>
<td>2.709</td>
<td>2.692</td>
<td>3.038</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>3.108</td>
<td>–</td>
</tr>
<tr>
<td>Log-normal scale parameter $\zeta$</td>
<td>0.616</td>
<td>0.659</td>
<td>0.691</td>
<td>0.781</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>0.616</td>
<td>–</td>
</tr>
<tr>
<td>Laplace location parameter $\mu$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–3.3</td>
<td>1.4</td>
<td>–1.1</td>
<td>–2.5</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Laplace scale parameter $b$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>65</td>
<td>63.9</td>
<td>55.3</td>
<td>66.7</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Logistic location parameter $\mu$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>15.54</td>
<td>–</td>
</tr>
<tr>
<td>Logistic scale parameter $s$</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1.91</td>
<td>–</td>
</tr>
</tbody>
</table>
Figure 3. Observed and fitted probability distributions of rear gap after lane change from I-80 data

Figure 4. Observed and fitted probability distributions of lag time to collision after lane change from U.S. 101 data
The Laplace distribution has a probability density function of

\[ f(x \mid \mu, b) = \frac{1}{2b} e^{-\frac{|x-\mu|}{b}} \quad -\infty \leq x \leq \infty \]  

(11)

which is symmetrical about its mean \( \mu \), \( \mu \) is known as the location parameters while \( b \) (\( b > 0 \)) is known as the scale parameter. The part of the Laplace distribution with \( X \geq \mu \) has the same shape as the exponential distribution.

The logistic distribution has a probability density function of

\[ f(x \mid \mu, s) = \frac{e^{\frac{x-\mu}{s}}}{s \left(1 + e^{\frac{x-\mu}{s}} \right)^2 } \quad X > 0 \]  

(12)

The logistic distribution has two parameters: \( \mu \) the location parameter and \( s \) (\( s > 0 \)) the scale parameter.

6. APPLICATIONS

The fitted distributions describe the interactions between vehicles at the critical instants of lane changing events. Such distributions may be used to quantify the risk taking behavior of the drivers of subject vehicles.

For example, \( T_{FA} \) is the time-to-collision between the subject vehicle and the following vehicle in the target lane. Small and positive \( T_{FA} \) value indicates a risky maneuver by the subject vehicle. An unsafe maneuver or severe conflict may be arbitrarily defined as \( 0 < T_{FA} \leq 0.5 \) second. Using the fitted Laplace distribution, \( P(0 < T_{FA} \leq 0.5) = 0.0196 \) at the I-80 site and \( 0.0009 \) at the U.S. 101 site. Comparing the probabilities, one can infer that drivers at the I-80 site perform relatively more risky lane change maneuvers. This may be due to the drivers’ behavior, the relatively more congested traffic, or the combination of both.

In a similar way, an unsafe maneuver or severe conflict may be arbitrarily defined as \( G_{FA} \leq 1.0 \) m. Correspondingly \( P(G_{FA} \leq 1) = 5.893 \times 10^{-4} \) at the I-80 site and \( 5.012 \times 10^{-5} \) at the U.S. 101 site. The probabilities indicate that there is a higher chance at the I-80 site (compared to the U.S. 101 site) that the subject vehicles moves into the target lane in close proximity in front of the following vehicle.

One can also analyze \( G_{FB} \), by calculating the probability \( P(G_{FB} < g) \) that the subject vehicle’s driver is willing to get closer to the leader in the original lane with a gap smaller than \( g \) during a lane change.

7. SUMMARY, LIMITATIONS AND FUTURE RESEARCH

This research has studied 10 parameters that describe vehicle interactions when the subject vehicle crosses the lane markers during a lane change, and analyzed the probability distributions of these parameters, using the NGSIM data. It is found that, overall,

- the parameters related to gap (in distance unit) and distance may be described by the log-normal distribution;
the parameters related to time to collision may be described by the Laplace distribution;

the parameter related to speed may be described by the logistic distribution;

The distributions fitted to the NGSIM data collected at the I-80 Freeway in Emeryville, California, and the U.S. Highway 101 in Los Angeles, California were compared. Although the same distribution was fitted to the same lane changing parameter, the fitted distribution parameter values were different for the two sites. This indicates that drivers behaved differently at the two data collection sites.

Correlation analysis was also performed for the 10 parameters studied. It was found that at least eight out of the 10 parameters are highly correlated. This implies that a few representative parameters may be sufficient to describe the interactions of vehicles.

The major limitations of this study are:

- The subject vehicles are cars. The probability distributions of the parameters for other types of vehicles are likely to be different, as suggested by [19, 20]. However their sample sizes are much smaller and therefore were not studied at the time of writing.
- The lane changes took place in moderate to congested traffic flow. It is yet to see if the correlations and probability distributions of the parameters in relatively free-flowing traffic are similar.
- The two data collection sites are located in California. The probability distributions are yet to be verified with data in other states.
- For each lane changing event, the parameter values were taken at the time instant $t$ when the front center of the subject vehicle crossed the lane markers. The driver of the subject vehicle usually makes his/her decision to change lane a fraction of a second to a few seconds prior to $t$. However, it is, impossible to determine when he/she psychologically makes this decision and measure the decision parameters at this point in time.
- A successful lane changing event may be preceded by several unused (or unsafe) lane change opportunities. This is synonymous to the gap acceptance scenario where there are more rejected gaps than accepted gaps. The distributions of the same parameters without an observed lane change are yet to be studied.

The above analysis of 10 lane changing parameters has led to several research ideas. From the correlation analysis, it appears that many parameters are highly correlated. Therefore it is highly possible to use fewer parameters to perform risk assessment of a lane changing event. From the reduced number of parameters, it may be possible to construct a joint probability distribution to describe the risk taking behavior of a subject vehicle. Based on this behavior, probabilistic lane changing models may then be developed.

REFERENCES


