Accident Analysis and Prevention 135 (2020) 105345

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Short-term prediction of safety and operation impacts of lane changes in oscillations with empirical vehicle trajectories

Meng Li, Zhibin Li*, Chengcheng Xu, Tong Liu

School of Transportation, Southeast University, Jiangsu Key Laboratory of Urban ITS, Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Dong Nan Da Xue Rd. #2, Nanjing, 211189, China

ARTICLE INFO

Keywords:
Freeway
Safety
Prediction
Machine learning
Lane change impact

ABSTRACT

Lane changes made during traffic oscillations on freeways largely affect traffic safety and could increase collision potentials. Predicting the impacts of lane change can help to develop optimal lane change strategies of autonomous vehicles for safety improvement. The study aims at proposing a machine learning method for the short-term prediction of lane-changing impacts (LCI) during the propagation of traffic oscillations. The empirical lane-changing trajectory records were obtained from the Next Generation Simulation (NGSIM) platform. A support vector regression (SVR) model was trained in this study to predict the LCI on the crash risks and flow change using microscopic traffic variables such as individual speed, gap and acceleration on both original lanes and target lanes. Sensitivity analyses were conducted in the SVR to quantify the contributions of correlative lane changing factors. The results showed that the trained SVR model achieved an accuracy of 72.81 % for the risk of crashes and 95.34 % in predicting the flow change. The sensitivity analysis explored the optimal speed and acceleration for the lane changer to achieve the lowest time integrated time-to-collision (TIT) value for safety maximization. Finally, we compared the LCI for motorcycles, automobiles and trucks as well as the LCI for both lane-changing directions (from left to right and from right to left). It was found that motorcycles conducted lane changes with smaller gaps and larger speed differences, which brings the highest crash risks. Passenger cars were found to be the safest when they conduct lane changes. Lane changes to the right had more negative impacts on traffic flow and crash risks.

1. Introduction

Traffic oscillations on freeways refer to the stop-and-go driving conditions which are usually presented as vehicle deceleration and acceleration alternately in congested traffic condition. Some researchers have provided empirical evidence that car-following and lane-changing behaviors are primarily responsible for the formation and growth of oscillations on multilane freeways (Zheng et al., 2011; Chen et al., 2012a,b). Once triggered, the propagation of oscillations or kinematic waves could bring adverse impact on traffic safety such as increasing the risk of traffic accidents (Zheng et al., 2010; Golob et al., 2004; Lee and Cassidy, 2008; Srivastava and Geroliminis, 2013; Li et al., 2014).

Vehicle lane-changing maneuvers are usually triggered by the heterogeneity of traveling speeds between different lanes. They could cause disturbances on the traffic flow and driver behaviors, which result in negative impacts on the safety. Previous studies have reported that lane changes cause frequent variation of gaps and speeds on target lanes and original lanes, which contributes to increased probability of collisions (Golob et al., 2004; Lee and Cassidy, 2008; Srivastava and Geroliminis, 2013), traffic flow disturbances (Ahn and Cassidy, 2007; Wang and Coifman, 2008) and capacity drop (Elefteriadou et al., 1995; Cassidy and Rudjanakanoknad, 2005; Laval et al., 2007; Chen and Ahn, 2018).

Previously, numerous studies have used different methods for predicting the lane change decisions in order to analyze the lane-changing impacts (LCI). Early studies applied theoretical and mathematical formulas to develop the microscopic lane change models for vehicles with factors measured at a variety of complexities (Yang and Koutsopoulos, 1996; Toledo et al., 2007; Laval and Leclercq, 2008). However, such models may contain multiple parameters that are difficult to observe and calibrate; for example, they usually assume impatient human factors to be constant in the models which is not realistic. Some other models developed for measuring LCI are with the kinematic wave theories (Laval and Leclercq, 2008; Laval and Daganzo, 2006; Jin, 2010, 2013). Such models are macroscopic (or hybrid) in nature and

* Corresponding author.
E-mail addresses: limeng@seu.edu.cn (M. Li), lizhibin@seu.edu.cn (Z. Li), xuchengcheng@seu.edu.cn (C. Xu), liut@seu.edu.cn (T. Liu).

https://doi.org/10.1016/j.aap.2019.105345
Received 28 March 2019; Received in revised form 10 September 2019; Accepted 18 October 2019
Available online 18 November 2019
0001-4575/

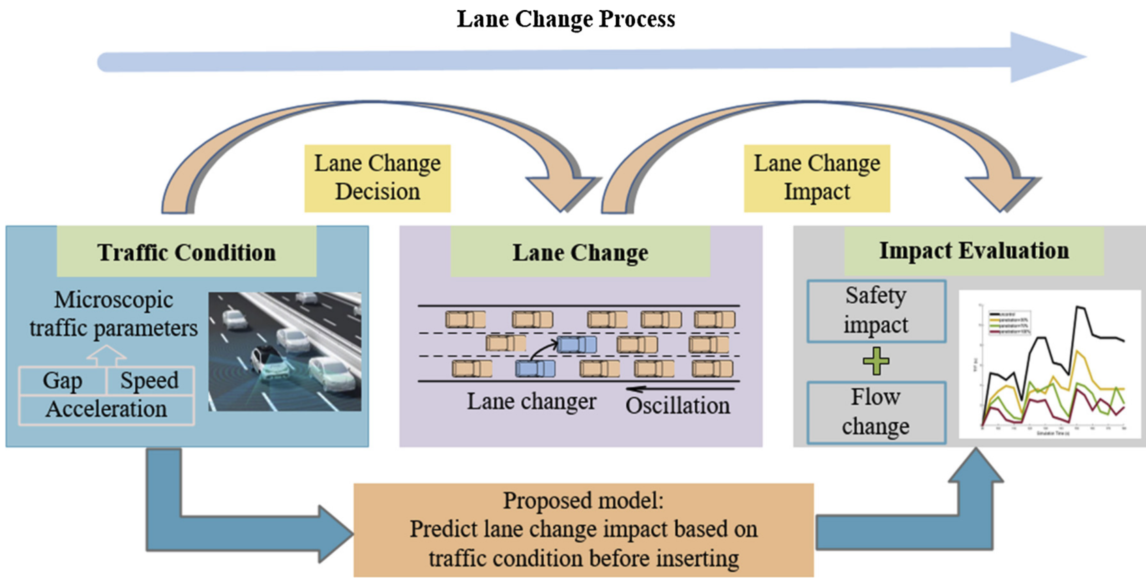


Fig. 1. Framework of Lane Change Impact Analysis.

cannot explore the influence of microscopic lane-changing behavioral factors. Later, some studies used the simulation techniques to estimate the LCI (Cheu et al., 2009; Pan et al., 2016; Li and Sun, 2017). However, they are not realistic since the lane-changing execution process in the simulation environment is different from the real word.

In recent years, empirical vehicle trajectory data are applied to evaluate the impacts of lane changes on traffic operations (Shvetsov and Helbing, 1999; Laval and Daganzo, 2006; Patire and Cassidy, 2011). For example, Zheng et al. (2011) calculated the oscillation propagation and flow change (FC) triggered by lane changes. Some other studies have analyzed the LCI on traffic safety using some surrogate safety models (Wang and Stamatiadis, 2013, 2014; Park et al., 2018). However, the aforementioned studies did not fully consider the microscopic traffic parameters before and after lane changes, and did not develop a prediction model considering the LCI on crash risk and flow simultaneously. In addition, many studies only considered the LCI on the target lanes, while those on the original lanes are ignored.

Our study aims at building a direct relationship between the real-time traffic parameters and the LCI (see Fig. 1) caused by the lane change behavior. Then, the LCI can be predicted directly based on traffic conditions before the inserting behavior. Thus, they can provide support for the development of optimal strategies, such as lane change controllers for autonomous vehicles, to prevent occurrence of traffic collisions in a proactive way. Statistical models contain straightforward formulas between the predictors and dependent variables, but may not work accurately in predictions. Recently, machine learning-based methods have been widely applied for prediction tasks due to their superior and reliable performances (Karlaftis and Vlahogianni, 2011; Pongnumkul et al., 2014; Marković et al., 2015); while the biggest issue is that they perform like a black box and do not report a direct relationship between the input and output variables. Such issue can be fixed by performing sensitivity analyses to traffic flow variables in our study (Li et al., 2008, 2012; Yu and Abdel-Aty, 2013).

The primary objective of the present study is to propose a machine learning-based method to predict the LCI on crash risks based on microscopic traffic parameters. The LCI on traffic flow is also considered to give a full evaluation of lane change behaviors. The delicate lane change records were extracted from the empirical vehicle trajectories in the Next Generation Simulation (NGSIM) database. Moreover, the support vector regression (SVR) model was developed to predict the short-period LCI. The rest of the paper is organized as follows. Section 2 presents a literature review on LCI. Section 3 shows the study sites and the data to be processed. Section 4 introduces the measurements of LCI on operational and safety and the SVR model. Section 5 describes the results of model analysis. Section 6 compares the LCI of different vehicle types and different lane-changing directions. Section 7 contains the conclusions and suggestions for future works.

2. Literature review

A number of studies have evaluated the safety impacts of lane-changing maneuvers on freeways and reported some negative effects. Golob et al. (2004) examined accidents that occurred on three types of weaving sections defined by differences of lane changes. They found there were significant differences in traffic accidents that occur within these types in terms of severity, location of primary collision, factors causing the accident, and time period in which accident is most likely to occur. Wang and Stamatiadis (2013, 2014) proposed an Aggregate Conflict Propensity Metric (ACPM) as a surrogate metric which can better evaluate the crash risk of lane change events by taking into account the driver's reaction time and braking capacity. Zheng et al. (2011) found lane change can also affect the formation and dissipation of traffic oscillations, which caused severe safety issues on freeways. Park et al. (2018) suggested that vehicle interactions between lane changer and adjacent vehicles are analyzed by two indexes called risk exposure level (REL) and risk severity level (RSL) to determine whether a lane change is in a dangerous situation. Those studies were just used for estimation the crash risks while the potential factors were ignored. Recently, Chen et al. (2018) established the relationship between lane-change related crashes and lane specific traffic and weather data. However, this research only considered macroscopic traffic parameters which are relatively fixed in a short time period. Thus, it could not be applied for a practical use (e.g. providing a driving recommendation to connected or autonomous vehicles).

Some studies focused on evaluating the impacts of lane changes on traffic operations. Some debates do exist because some scholars found that lane changes have negative impact on traffic efficiency in some scenarios while other researchers unveiled the opposite phenomenon. Specifically, many experiments such as Bertini and Leal (2005) showed a consistent reduction in discharge flow after the onset of congestion at bottlenecks caused by lane-drops. Cassidy and Rudjanakanoknad (2005) found that lane-changing maneuvers can cause capacity drop at a merge bottleneck. Other studies found that lane-changing maneuvers can enlarge the capacity drop phenomenon (Treiber et al., 2006; Laval

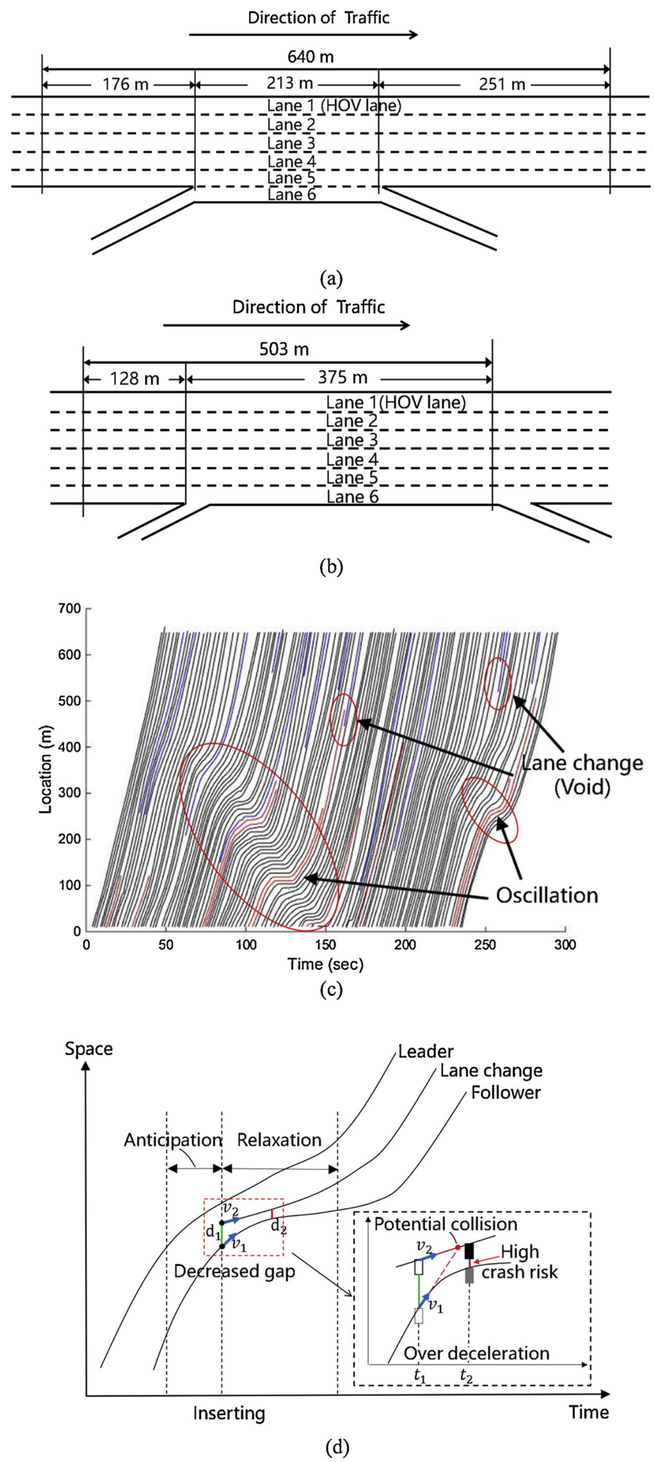


Fig. 2. (a) The sketch of US-101 study area; (b) The sketch of I-80 study area; (c) Vehicle trajectory map; (d) Three hypothetical vehicle trajectories to illustrate potential crash risks caused by lane changer.

et al., 2007). Coifman et al. (2006) found that lane-changing maneuvers cause additional delay in queues. Laval and Daganzo (2006) proposed a hybrid lane change model with four-parameters (which are free flow speed, wave speed, jam density, and capacity) and explained the mechanism of capacity drop. Jin (2010) suggested that a lane changer's impact to total density was doubled because it essentially uses two lanes during the LC execution.

On the other hand, some other studies have reported that the lane-changing maneuvers could increase traffic flow in some scenarios. It was found that the heterogeneity among different lanes is usually a reason motivating lane-changing (Shvetsov and Helbing, 1999; Patire and Cassidy, 2011). While lane-changing maneuvers, on the other hand, might have balancing effect, i.e., lane-changing could smooth out differences between adjacent lanes under certain situations. This balancing effect could be beneficial to the whole traffic system in achieving higher efficiency. Patire and Cassidy (2011) found that early in the rush, when flow was relatively low in the shoulder lane, drivers readily migrated toward that lane to escape the oncoming speed disturbances (SDs). Cheu et al. (2009) conducted simulation analysis and found that lane-changing could reduce the overall system queuing delay.

Many other aspects of LCI were also considered by previous studies, such as safety cost, equilibrium cost, control cost, travel efficiency cost, route cost, lane preference cost, etc. Researchers have postulated the linkage between LC's complex impact on surrounding traffic and some long-puzzling traffic phenomena, such as breakdown, capacity drop, and traffic oscillations (Cassidy and Rudjanakanoknad, 2005; Laval and Daganzo, 2006). In a more recent study, Zheng et al. (2013) detected three distinct effects of lane changer on the follower in the target lane: anticipation, relaxation, and the regressive effect on driver behavior (i.e., lane changer "neutralizes" the follower's behavior by encouraging a timid (aggressive) driver to become less timid (aggressive)). Wang et al. (2015) proposed a cost function to calculate LCI but only based on a central mathematical framework and numerical experiments.

In terms of safety measures, some surrogate safety measures (SSMs) were proposed to quantify the potential of crash risks (Hydén, 1987), which is suitable for evaluate traffic safety in short-term period or microscopic scope like trajectory data. There are mainly three types of SSMs. The first type is based on Time-to-collision (TTC), which is calculated based on predicted collision point. TTC and its advanced indexes, time exposed TTC (TET) and time integrated TTC (TIT) (Minderhoud and Bovy, 2001; Van Winsum et al., 2000) have been successfully applied in traffic and vehicle safety field. The second type is based on Post encroach time (PET), which is the time difference when two vehicles pass through one point (Allen et al., 1978). Its advanced indexes include gap time (GT), encoding time (ET) and time advantage (TAdv) (Hansson, 1975). The third type is derived from deceleration rate, which includes Maximum deceleration (Max D), deceleration-to-safety time (DST), deceleration rate to avoid crash (DRAC), and stopping distance index (SDI) (Gettman and Head, 2003; Hupfer, 1997; Cooper and Ferguson, 1976; Oh et al., 2006).

The major objective of our study is to develop a machine learning model to proactively predict the impact of lane change from both safety and operational aspects based on observable traffic parameters. The machine learning model is developed in a data-driven way. Unlike any regression model, it does not require pre-assumption of the mathematical form between lane change and traffic status. Thus, the machine learning model can improve the prediction accuracy of lane change impacts.

3. Lane change trajectory

3.1. Trajectory data

In this study, vehicle trajectory databases used for LCI analysis were obtained from FHWA's Next Generation Simulation (NGSIM) program (FHWA, 2008). The US-101 data was used to establish the LCI prediction models: it was collected from 07:50 a.m. to 08:35 a.m. on a 640 m segment on the south-bound direction of US-101 in Los Angeles, California on June 15th, 2005. The second database, on the I-80, was used for transferability tests: it was collected from 4:00 p.m. to 4:15 p.m. and from 5:00 p.m. to 5:30 p.m. on a segment of approximately 500 m in the San Francisco, California on April 13, 2005.

Fig. 2(a) and (b) provides a schematic illustration of these two study sites. Lanes are numbered in an increasing order from the left to the right. Each study site contains an on-ramp and an off-ramp, where intensive lane-changing activities are expected. The trajectory data includes rich information of individual vehicles such as speed, gap, acceleration in each time step (0.1 s), as well as lateral and longitudinal locations which can be used to identify lane-changing maneuvers. In US-101 dataset, the traffic volume varies from 5192 veh/h to 7972 veh/h during the 45 min. The component of traffic during investigation is 0.79 % motorcycle, 97.04 % passenger car and 2.17 % truck. In I-80 dataset, the traffic volume varies from 4968 veh/h to 6900 veh/h during the 45 min. The component of traffic during investigation is 0.68 % motorcycle, 95.37 % passenger car and 3.95 % truck.

An example of vehicle trajectories on a single lane on US-101 during 5 min from 7:50 to 7:55 were illustrated in Fig. 2(c). The blue lines indicate the vehicles departing from one lane, the red lines indicate the vehicles arriving to another lane and the black lines indicate the vehicles that have not changed lanes. We notice that the oscillations have caused large influence (i.e. abrupt vehicle deceleration and acceleration) on numerous vehicles, as it is observed with the presence of oscillation waves. Numerous lane changes were identified in the trajectory map. Further, when a lane changer arrives to the target lane, a "void" (i.e. the void ahead of the lane-changer due to the bounded acceleration in target lane) will occur between lane changer and its immediately preceding vehicle, which was found to be the main reason of capacity drops (Laval and Daganzo, 2006).

Crash risks are likely to occur during lane changes. We plotted three hypothetical vehicle trajectories to illustrate potential crash risks caused by a lane-changing maneuver in Fig. 2(d). Note that a lane-changing process is divided into two parts: anticipation (i.e., the transition of the follower in the target lane after the lane changer's intention is noticed and before the lane changer inserts) and relaxation (i.e., lane changer tends to accept shorter spacing than desired spacing upon merging and then gradually reaches the desired spacing) (Laval and Leclercq, 2008; Zheng et al., 2013; Zheng, 2014). When an urgent inserting of a lane changer occurs, the follower with a relatively high speed has to conduct an abrupt deceleration to prevent traffic accident. From t_1 and t_2 in Fig. 2(d), the gap between lane changer and his follower rapidly decreases, and the extremely small gap in t_2 brings a high crash risk. The safety impacts of lane change should be carefully estimated.

3.2. Data processing

Previous scholars found there exist some problems in NGSIM dataset (e.g. acceleration often exhibits unrealistically large magnitudes) (Coifman and Li, 2017). The trajectory data appear unfiltered and exhibited some noise artefacts (Khodayari et al., 2012; Thiemann et al., 2008). Therefore, we apply a moving-average filter (Zhou et al., 2017) to all raw trajectories, and obtain the numerical velocity/acceleration data from the first-/second-order finite differences of the position respectively. Further data pre-processing includes: i) setting negative vehicle movements to zero which is caused by moving average smoothing; ii) refining differentiated accelerations to a range of (3.41, 3.41), which is consistent with the range in NGSIM acceleration data

Lane change records were extracted when the lane label of the vehicles changed. The data collection process recorded the labels from the original and the target lanes for lane changes, as well as microscopic vehicle parameters. In total, we obtained 3586 lane change records, with 2376 on the US-101 and 1210 on the I-80. Traffic oscillation was a confounding factor in the accurate estimation of the flow change occurred due to lane changes. Oh and Yeo (2012) suggested using a 5 vehicles immediately preceding the lane changer to calculate the local traffic flow in an oscillation without the influence of lane change. With this approach, the flow changes triggered only by the lane changer could be calculated as the difference of the values of local traffic flow and the flow influenced by lane change and oscillation. The former value is calculated based on the 5 vehicles while the latter value is calculated based on the lane changer and its following vehicles (Oh and Yeo, 2012).

In our study, 5 adjacent vehicles immediately preceding the lane changer on the target lanes and the original lane are considered to calculate the LCI. Note that the 5 vehicles are not necessarily passenger cars. The impact of vehicle type on the result is limited as the majority of vehicles are passenger cars (more than 95 %). The impact of the vehicle types of the 5 vehicles preceding lane changer on the results has been reduced because the average headway of the 5 vehicles was used in the calculation. Accordingly, data without enough vehicles preceding the lane-changer in either original lane or target lane were excluded. The time interval after the inserting for the calculation of the LCI should be defined carefully. Oh and Yeo (2012) used 3 s, 6 s, 9 s as the time interval respectively to evaluate LCI in the relaxation period and found that the relaxation effect lasts to 9 s. Thus, our study adopted 9 s as the time interval for estimating LCI. If a vehicle had records of multiple lane changes within 9 s, such data did not support LCI estimation and were excluded.

3.3. Variables for LCI prediction

Lane-changing maneuvers involve interactions of multiple vehicles. In this section, the variables which may influence the drivers' lane change decisions and affect the LCI should be carefully determined. For this purpose, it was considered five vehicles including the lane changer which are directly involved in a typical lane-changing scenario (see Fig. 3). In the original lane, there is one vehicle that follows (FO) the lane changer and one that precedes (PO) the lane changer; in the target lane, the same definition was applied, where one follows (FT) the lane changer and one precedes (PT) the lane changer.

In order to make our algorithm more suitable for practical use, the selected variables for LCI estimation should be those measurable by in-vehicle equipped sensors. In the meantime, the selected traffic parameters should avoid multicollinearity. For example, the speed of individual vehicles and the speed difference between vehicles would not be selected simultaneously as they are highly correlated. Similarly, the gap and vehicle type have a strong inter-correlation and are not considered in the model at the same time.

Finally, eleven variables were selected and their mean value and

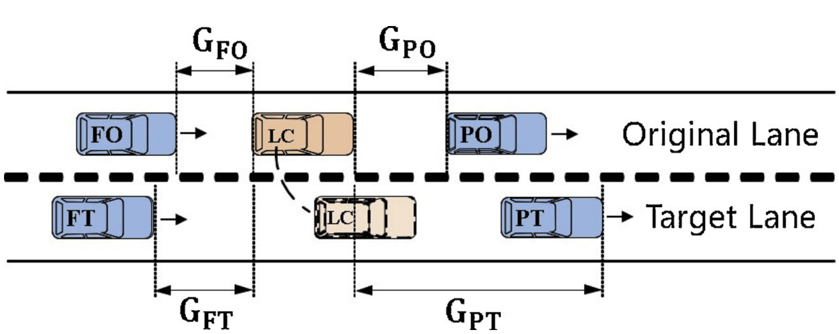


Fig. 3. Vehicles in a typical lane change.

Table 1
Descriptive information of variables related to lane changes.

Variable	Description	Mean	S.D.
G_{FO}(m)	Gaps between lane changer and vehicle FO	19.12	11.39
G_{PO}(m)	Gaps between lane changer and vehicle PO	14.09	10.07
G_{FT}(m)	Gaps between lane changer and vehicle FT	17.03	10.41
G_{PT}(m)	Gaps between lane changer and vehicle PT	13.88	11.39
VD_{FO}(m/s)	Speed of lane changer minus speed of FO	0.52	2.21
VD_{PO}(m/s)	Speed of PO minus speed of lane changer	0.75	2.28
VD_{FT}(m/s)	Speed of lane changer minus speed of FT	0.44	2.22
VD_{PT}(m/s)	Speed of PT minus speed of lane changer	0.32	1.98
V_{LC}(m/s)	Speed of lane changer	12.02	2.58
A_{LC}(m/s^2)	Acceleration of lane changer	0.22	1.49
VD_{Lane}(m/s)	Average speed in target lane minus that in original lane	1.07	2.31

standard deviation (S.D.) are listed in Table 1.

4. Methodology

4.1. Calculation of LCI on traffic safety

Traditional crash count-based safety models rely on crash data accumulating in long periods. They are not appropriate for collision risk assessment in vehicle trajectories that is the purpose of this study. Thus, it was used an index of time-to-collision (TTC) which represents the time required for two successive vehicles, occupying the same lane, to collide if they continue at their present speed when vehicle i moves faster than leading vehicle $i-1$ (Hayward, 1972; Svenson et al., 2012; Jiménez et al., 2013):

$$TTC_i(k) = \begin{cases} \dfrac{x_{i-1}(k) - x_i(k) - L}{v_i(k) - v_{i-1}(k)} \; if \; v_i(k) > v_{i-1}(k) \\ \infty \; if \; v_i(k) \le v_{i-1}(k) \end{cases} \quad (1)$$

where k = time step, L = length of vehicle i-1. Based on the TTC notation, an advanced surrogate safety measure, time integrated time-to-collision (TIT) was developed to assess the risks of collisions. Larger TIT value indicates higher collision risks (Minderhoud and Bovy, 2001).

The TIT expresses the entity of the TTC lower than the TTC threshold, denotes as TTC*, should be determined to distinguish the unsafe car following situations from the ones considered safe. According to previous studies, the threshold of 2 s was adopted in our study (Sultan et al., 2002). The reciprocal transformation was made considering the fact that a lower TTC means a higher collision risk:

$$\text{TIT}(k) = \sum_{i=1}^{N} \left[\frac{1}{TTC_i(k)} - \frac{1}{TTC^*} \right] \bullet \Delta k, \; \forall \; 0 < TTC_i(k) \le TTC^* \quad (2)$$

$$\text{TIT} = \sum_{k=1}^{T} TIT(k) \quad (3)$$

where TIT(k) = TIT value at time k, T = calculating period, k = time step, i = vehicle ID, N = number of total vehicles.

Previous studies have suggested that a typical lane change has direct safety impacts on three vehicles, which are the lane changer, the immediate follower in the original lane, and the immediate follower in the target lane (Zheng, 2014). Accordingly, the collision risks associated with the lane change were estimated by calculating the TIT values for the three vehicles (N = 3). In addition, the calculating period of TIT is set to be 9 s (T = 9 s) as discussed in Section 3.

TTC and TIT have been widely used for surrogate measures of crashes, especially in freeway environment (Li et al., 2016, 2017; Rahman and Abdel-Aty, 2018). They are considered appropriate surrogate safety measures because they take the full course of vehicles over space and time into account. As a result, TIT which can give a complete and comprehensive picture of the safety level on a particular stretch of road during a particular period of time was adopted in this study.

4.2. Calculation of LCI on traffic flow

The impacts of lane change on traffic operation should be evaluated for fully LCI estimation. In the congested traffic flow the LCI may be affected by other factors such as the propagation of oscillation waves (Oh and Yeo, 2012; Zheng et al., 2011). Such effects should be carefully controlled in order to estimate more accurate LCI. We divided traffic into two regions (see Fig. 4), where region A represents vehicles following the lane changer and region B represents vehicles proceeding the lane changer. In region B, the traffic operation is only affected by preceding oscillations and the flow rate in this region is denoted by q_{Normal}. In region A, the traffic operation is jointly affected by the lane change and the same oscillation in region B, and the flow rate in this region is denoted by q_{LC}. Then, the FC influenced by lane changer alone can be calculated by the difference of these two values:

$$FC = q_{LC} - q_{normal} \quad (4)$$

Note that the Eq. (4) is deduced based on an assumption that oscillation has equal impacts on traffic flow in a small segment. Thus, the same oscillation has equal influence (denoted by q_{normal}) on region A and region B. Then, this value is subtracted from q_{LC} to remove the impacts of oscillation.

FC is calculated for the target lane and the original lane separately. Several schematic trajectories in two lanes are illustrated in Fig. 5 to depict the calculation process of FC_{tar} (FC in target lane) and FC_{ori} (FC in original lane).

According to the relationship between headway and flow, the flow (q) can be estimated by the reciprocal of the time headway (h):

$$q = \frac{1}{\frac{1}{m}\sum_{i=1}^{m} h_i} = \frac{1}{\bar{h}} \quad (5)$$

The FC in target lane is first calculated. In target lane (see Fig. 5(a)), the trajectory of the lane changer is indicated by a red line. The flow in normal case ($q_{Normal,tar}$) is defined as the reciprocal of average time headway of 4 headways in the 5 vehicles preceding the lane changer (h_i):

$$q_{Normal,tar} = 1/\frac{1}{4}\sum_{i=1}^{4} h_i \quad (6)$$

Moreover, the flow rate, influenced by the lane-changing maneuver in the target lane, is defined as $q_{LC,tar}$, which can be calculated as

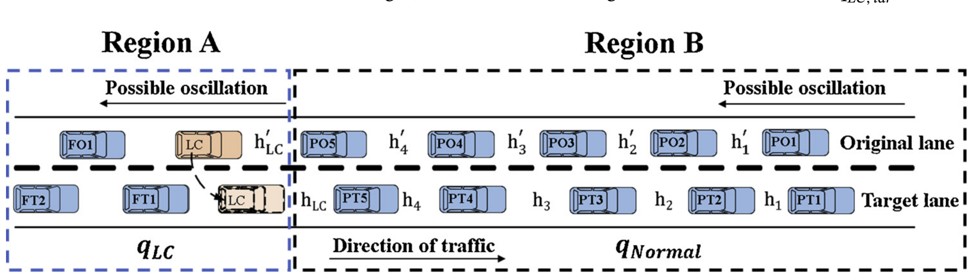
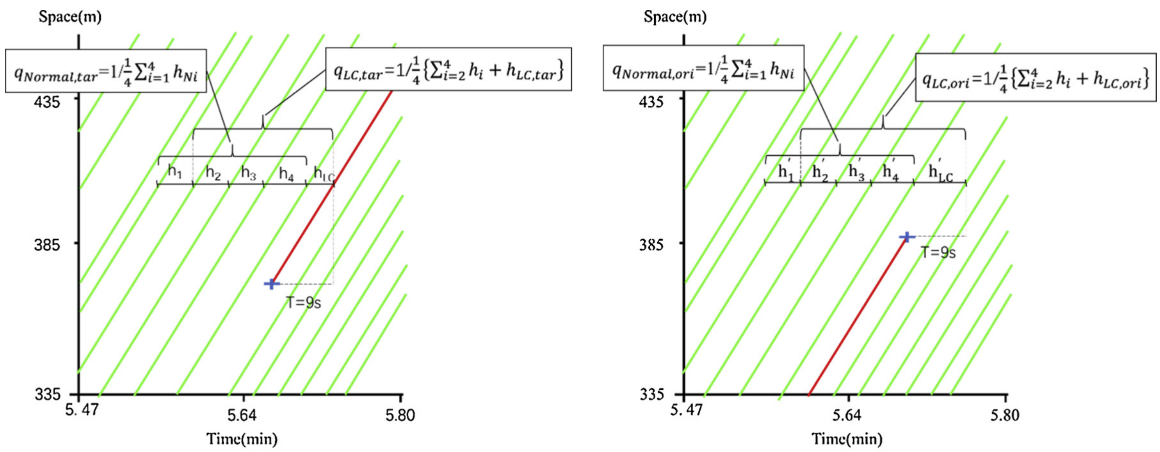


Fig. 4. Region division in traffic flow.



(a) (b)

Fig. 5. Sketches of trajectories (a) In target lane; (b) In original lane.

follows (Oh and Yeo, 2012):

$$q_{LC,tar} = 1/\frac{1}{4}\left\{\sum_{i=2}^{4} h_i + h_{LC,tar}\right\} \tag{7}$$

where $h_{LC,tar}$ is the time headway between lane changer and its immediately preceding vehicle.

Then, the FC triggered by the lane-changing maneuver alone in the target lane could be calculated as follows:

$$FC_{tar} = q_{LC,tar} - q_{Normal,tar} \tag{8}$$

The FC in original lane is then calculated. As illustrated in Fig. 5(b), in original lane, the flow rate in normal case ($q_{Normal,\,ori}$) can be calculated using the same formula (see Eq. (6)). The flow influenced by lane-changing maneuvers in original lane is defined as $q_{LC,ori}$, which can be calculated as follows:

$$q_{LC,ori} = 1/\frac{1}{4}\left\{\sum_{i=2}^{4} h_i + h_{LC,ori}\right\} \tag{9}$$

where $h_{LC,ori}$ is the time headway between the immediately following and immediately preceding vehicle of lane changer.

The calculation of $q_{Normal,ori}$ is similar to the calculation of $q_{Normal,tar}$. As a result, the change of flow triggered by lane change alone in the original lane could be calculated as follows:

$$FC_{ori} = q_{LC,ori} - q_{Normal,ori} \tag{10}$$

Eventually, the FC caused by lane-changing maneuvers defined as FC can be calculated as:

$$FC = FC_{tar} + FC_{ori} \tag{11}$$

4.3. Supporting vector regression

Recently, machine learning based car following models have been developed and successfully applied, which attempt to learn the car-following maneuvers from a large number of human driver car-following data. (Khodayari et al., 2012; Zhou et al., 2017; Yang et al., 2018). The machine learning models can extract the drivers' car-following behaviors and capture the potential relationships among the various variables impacting car-following behavior. The machine learning model was applied in this study to analyze lane change maneuvers.

Among various machine learning algorithms, Support Vector Machine (SVM) model has been gaining increasing popularity due to its good predictive performance. Some studies found that the SVM model showed better results than the outcomes estimated by statistical and other machine learning methods (Iranitalab and Khattak, 2017; Zhang et al., 2018). SVM can be divided into a classification support vector machine which is primarily used for the classification problems and a regression support vector machine (SVR) which is primarily used for the prediction of continuous variables.

The SVR is known for its good generalization performance and its ability to handle nonlinear problems, which has been successfully applied to a variety of real-world problems. It simultaneously minimizes the regularization error and empirical risk with a suitable penalty factor (Marković et al., 2015; Cheng et al., 2017; Alade et al., 2018). Though more complex models such as deep learning methods could be considered for the predicting task (Bao et al., 2019), the improvement of model performance is marginal but the model practicability and usability are reduced. As a result, SVR was employed in this study to predict the LCI.

More specifically, the SVR model treats the LCI modeling as a regression problem where the measurements of safety were fitted using empirical data. The variables representing traffic status before the inserting (see Table 1) were considered as the input vectors while LCI was considered as the output vectors in the SVR models in this study.

Let $\{(X_1, y_1), \ldots, (X_n, y_n)\}$ be training data, where X_i and y_i denote input and target data, respectively. The goal in ε-SVR is to determine a function f(X) that (a) has at most ε deviation from target data and (b) is as flat as possible (Smola and Schölkopf, 2004). In particular, we apply the ε-SVR with radial basis functions:

$$f(X, W) = \sum_{j=1}^{n} W_j \exp(-\gamma \, \|X - X_j\|^2) \tag{12}$$

where γ is a parameter and vectors X_i are inputs from the training data. The vector of unknown parameters W is determined to minimize the function:

$$\min \frac{1}{2}\,\|W\| + C\sum_{i=1}^{n} max(|y_i - f(X_i, W)| - \varepsilon, 0) \tag{13}$$

where parameter $C > 0$ controls the tradeoff between the flat ness of f(·) and amount up to which deviations greater than ε are tolerated. The dual of this optimization problem is solved using convex programming techniques (Smola and Schölkopf, 2004).

Data preprocessing and parameter selection are important for good SVR performance. First, all the variables were scaled linearly to the range [0,1]. Second, our ε-SVR included parameters γ, C and ε. We intended to find the values of these parameters that could maximize the predictive power of our models. Finally, the process was done via

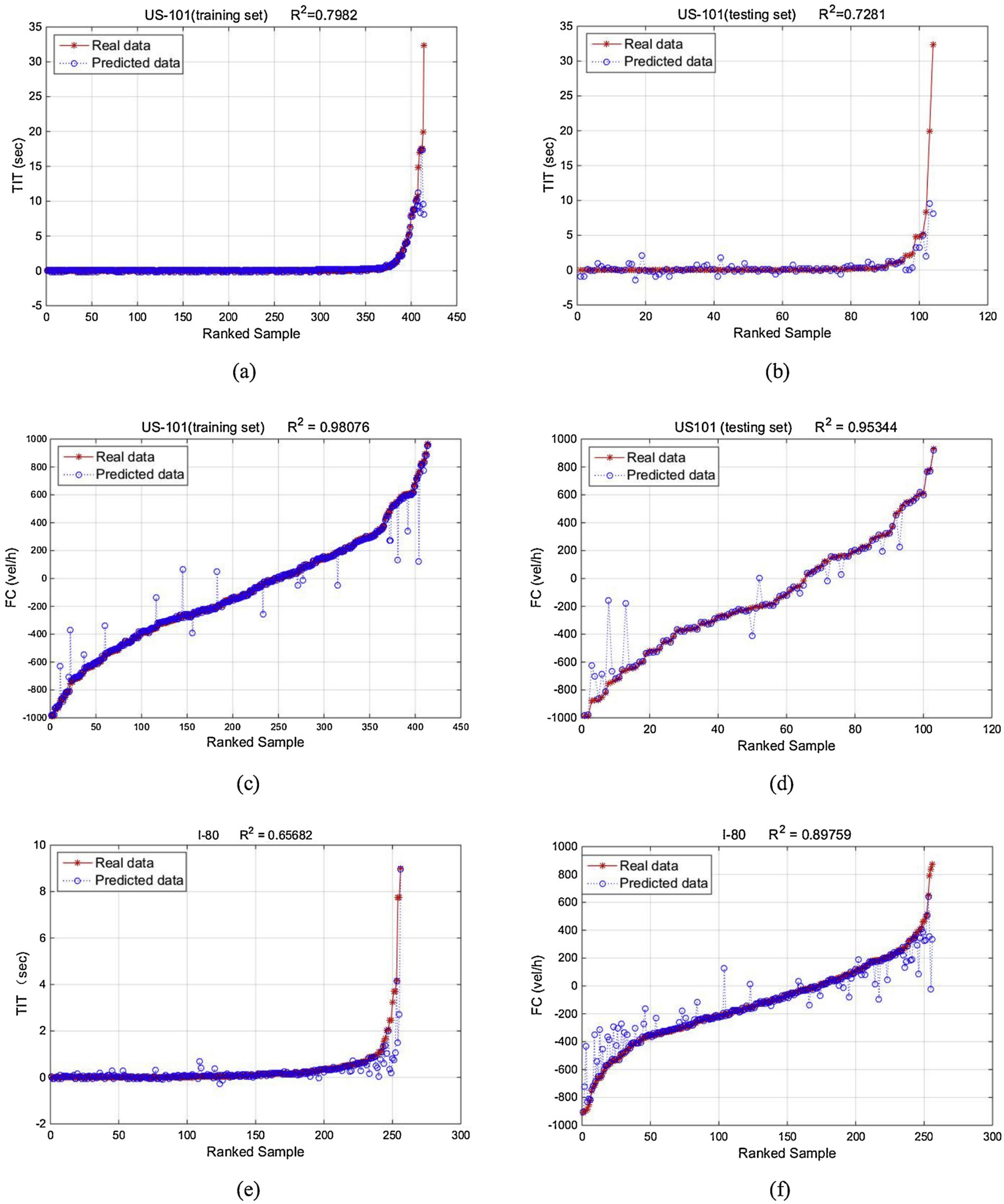


Fig. 6. Results of predicting LCI using SVR model (a) Predicting TIT in the training set of US-101; (b) Predicting TIT in the testing set of US-101; (c) Predicting FC in the training set of US-101 (d) Predicting FC in the testing set of US-101; (e) Predicting TIT in the dataset of I-80; (f) Predicting FC in the dataset of I-80.

combinations of cross-validation and grid-search techniques (Marković et al., 2015).

5. Results of model prediction

5.1. Predicting performance and transferability test

In this section, the SVR models are developed based on dataset on US-101 which contains 518 valid lane-changing records. The dataset on I-80, which includes 256 records, is employed to test for model transferability. The original dataset on US-101 was randomly separated into a training set and a testing set with a ratio of 4:1. We used data normalization to [0, 1] interval for forecasting. The LIBSVM tool developed by Chang and Lin in MATLAB was employed to specify the SVR model in this study (Chang and Lin, 2011). The LIBSVM tool provided a grid searching algorithm for determining the model parameters (C,γ). The prediction results of the LCI are shown in Fig. 6.

The small red stars displayed in the increasing sequence represent the observed values, while the values predicted by the SVR are marked as small blue circles. To evaluate the performance of the SVR models, the following measurements are calculated as follows:

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |observed_i - predicted_i| \tag{14}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N} (observed_i - predicted_i)^2 \tag{15}$$

$$NRMSE = \frac{\sqrt{MSE}}{observed_{max} - observed_{min}} \tag{16}$$

$$R^2(R\,square) = 1 - \frac{\sum_{i=1}^{n}(observed_i - predicted_i)^2}{\sum_{i=1}^{n}(observed_i - \bar{observed})^2} \tag{17}$$

The following measurements are widely used to evaluate the accuracy of the model prediction (Hadayeghi et al., 2010; Wang and Shi, 2013; Tastambekov et al., 2014; Bao et al., 2018; Guo et al., 2018). Mean Absolute Error (MAE) and Mean Square Error (MSE) provide relatively intuitive judgment of the accuracy of model predictions. R square was used to present the goodness of fit of SVR models (Marković et al., 2015). The NRMSE, which is deduced based on MSE, has the advantage of considering the range of raw data and is more reasonable for evaluating the predicting performance (Haworth et al., 2014). The results of the four measurements are shown in Table 2.

The measurements in Table 2 suggest that the overall predicting accuracy of the SVR models is reasonably good. For example, the MAE in the prediction of TIT triggered by lane change are 0.7944 and 0.1738, which are considered very small as the max value of TIT could be over 30. Note that the value of MAE and MSE in the validation dataset is even smaller than that in the test dataset when predicting TIT. The reason might be that the variance range of TIT value on I-80 is smaller than that on US-101, due to the different road layouts and traffic parameters, which brings less prediction error. In such case, the NRMSE considering data range is a more reasonable index and the results show that the NRMSE is slightly increased on I-80.

Similarly, the results of the FC prediction model show a slightly better performance with the data on US-101 than that on the I-80, which is consistent with the performance of the TIT prediction model. As compared to the FC model, the general accuracy of the crash risk model was obviously lower. We carefully checked the predicted results and found that the main reason is the relatively large errors in predicting some extremums (see Fig. 6(a), (b) and (e)). Nevertheless, the SVR models have shown performing an acceptable transferability when applied on different freeways.

5.2. Sensitivity analysis for traffic parameters

As a practical and widely used machine learning method, SVR was blamed for performing like a black-box; in other words, in this method the effects of explanatory variables on the dependent variable could not be directly seen. To solve this problem, Fish and Blodgett suggested performing a sensitivity analysis to explore the relationships between output and input variables (Fish and Blodgett, 2003). In this study, sensitivity analysis was conducted to explore the relationship between crash risks triggered by lane change and microscopic traffic parameters.

Table 2
Predictive performance of SVR models.

LCI	Dataset	MAE	MSE	NRMSE	R^2
TIT	Training set on US-101	0.2858	2.0806	0.0445	0.7982
	Testing set on US-101	0.7944	1.8839	0.0424	0.7281
	Validation set on I-80	0.1738	0.4001	0.0703	0.6568
FC	Training set on US-101	19.5799	4108.5	0.0327	0.9808
	Testing set on US-101	27.6807	7832.0	0.0466	0.9534
	Validation set on I-80	43.5174	10809.0	0.0583	0.8976

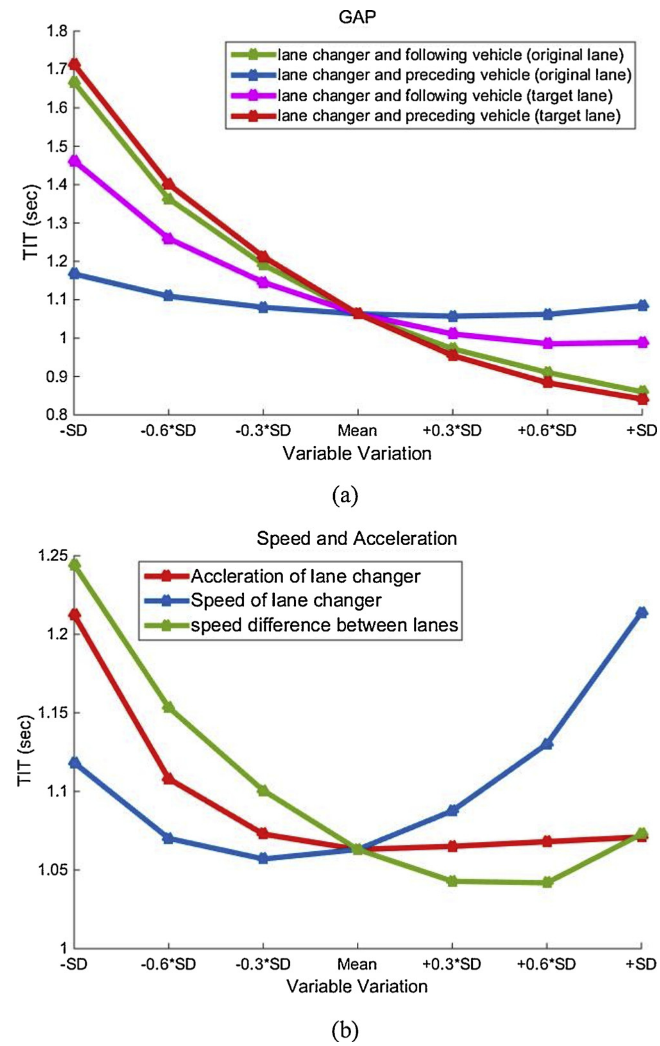


Fig. 7. Sensitivity analysis of relationship between TIT and (a) gaps between different vehicles; (b) speed and acceleration of lane changer, speed difference between target lane and original lane.

Each variable was set to perturb (i.e. decrease and increase) from the mean by one standard deviation, to ensure the variation is within the reasonable range (large perturbation may result in unrealistic variable values). During the sensitivity analysis, only one variable was allowed to change per time. The estimation results before and after the perturbation of the input variable was recorded (Li et al., 2012). In this way, the change of TIT solely caused by the variation of the selected variable is captured.

Note that there are some correlations among the speed differences between adjacent vehicles (e.g. the VD_{FO}, VD_{PO}, VD_{FT}, VD_{PT}). For example, a lane changer with fast speed contributes to a relatively large VD_{FO} and a relatively small VD_{PO}, thus, the sensitivity of the speed differences between adjacent vehicles was excluded in this study.

A sensitivity analysis was conducted for the LCI on safety performance. In Fig. 7(a), the TIT decreases with the enlargement of gaps between vehicles. It is evident that longer distance between vehicles contributes a safer traffic condition. We noticed that almost all curves illustrated in Fig. 7(b) show a concave trend. It means that is possible to obtain an optimal speed and acceleration for the lane changer to achieve the lowest TIT value for safety maximization. Such findings are particularly useful for the development of driving strategies for autonomous vehicles. For example, we can predict the real-time crash risk of lane change and automatically adjust the speed and acceleration of the vehicle to the optimal value before conducting the lane changing maneuver.

6. Comparison for vehicle class and lane change direction

Different types of vehicles have distinct sizes, dynamics and driving behaviors. Their lane-changing maneuvers could have different influences on traffic safety. Similarly, due to the heterogeneity of traffic between different travel lanes, the lane change direction also has various impacts on traffic system. In this section the LCI was compared for the vehicle class and lane change direction. Previously, the value of TTC* varies from 1 to 3 s (Nilsson et al., 1992; Hirst and Graham, 1997; Sultan et al., 2002; Li et al., 2014). Hence, a sensitivity analysis for five values of the TTC* from 1.0 s to 3.0 s with a step of 0.5 s was conducted for a more comprehensive comparison.

6.1. Comparison of LCI for different vehicle types

Vehicles on the study sites were classified into three types: motorcycles, passenger cars and trucks. The crash risks caused by lane changes of the three vehicle types were calculated with different TTC threshold (see Fig. 8 a). Note that some motorcycles may not drive in a lane based manner. However, there are no surrogate safety measures specifically proposed for motorcycles. As a result, to compare the safety of motorcycles with other vehicle types, we still kept using the TIT as the surrogate safety measure in the analysis. Different influences caused by three vehicle types were obtained. Results show that motorcycle is the most dangerous vehicle type for any TTC threshold when changing lanes, while passenger car and truck have less influence on traffic safety. Further, passenger cars were found to be the safest when they conducted lane changes than motorcycles and trucks. It is worth mentioning that the conclusions are quite consistent for different TTC thresholds.

To explore the potential reasons, we further plotted the speed and acceleration curves of three vehicle types, as illustrated in the Fig. 8(b) and (c). It was found that the motorcycles had the fastest speed and the largest acceleration/deceleration rate in congested traffic when changing lanes. Larger acceleration/deceleration by motorcycles might cause smaller gaps and larger speed differences that brought higher safety risks. Because of the larger sizes and the lower acceleration/deceleration abilities, trucks were not so aggressive and responded slowly to traffic changes, which contributed to fewer but still considerable crash risks. Further, automobiles were found to be the safest when they conducted lane changes than motorcycles and trucks.

LCI on traffic operation was also estimated in this part for full comparison (see Fig. 8d). Results showed that motorcycle lane changer could averagely increase traffic flow while trucks have the most negative effect on flow when change lanes (with average FC being 782.24 veh/h and −487.17 veh/h, respectively). The reason might be that with averagely larger speed, motorcycle lane changers are more likely to fill the gap between individual vehicles timely. Such behavior increases the probability of flow increase, vice versa.

6.2. Comparison of LCI for different lane change directions

Lane changes were divided into two directions, i.e. from left to right and from right to left. The TIT of different lane-changing directions calculated based on different TTC threshold was compared as shown in Fig. 9(a). Overall, the lane change to the left lane was found to be safer in terms of lower TIT values. Similar to the comparison between different vehicle types, the speed and acceleration curves for the two lane-changing directions are plotted in Fig. 9(b) and (c). It was found that the lane changes to the left lane had averagely larger speeds and less acceleration/deceleration rates than the lane change to the right lane. As a result, changing to left lane may perform better in filling the gap,

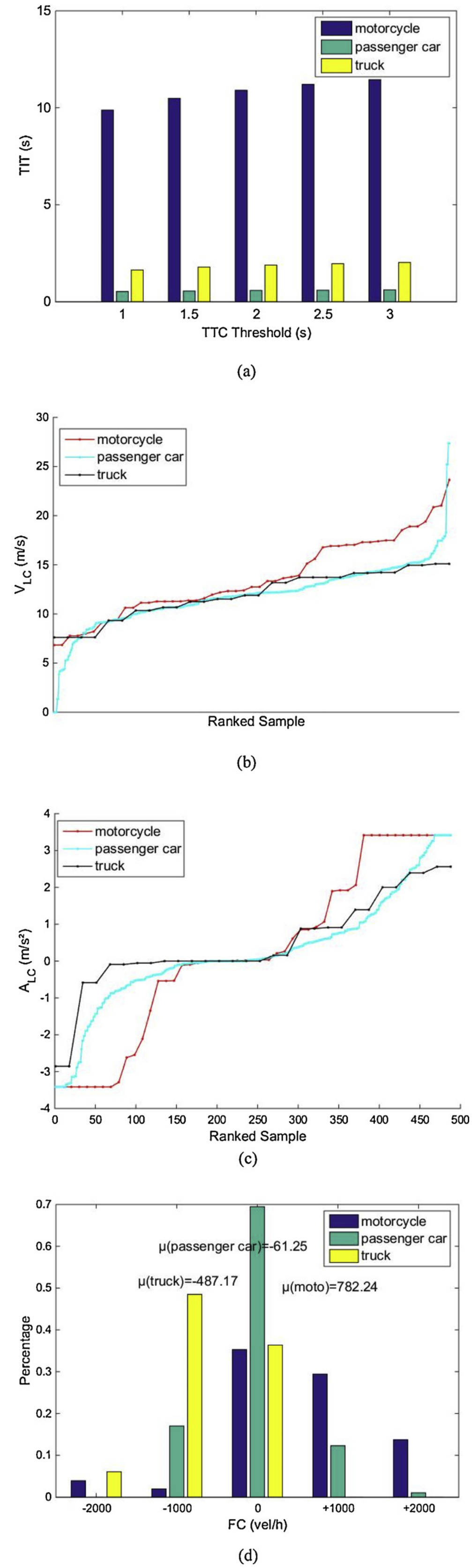


Fig. 8. Comparison between different vehicle types (a) Distribution of TIT; (b) Speed of lane changer; (c) Acceleration of lane changer; (d) Distribution of FC.

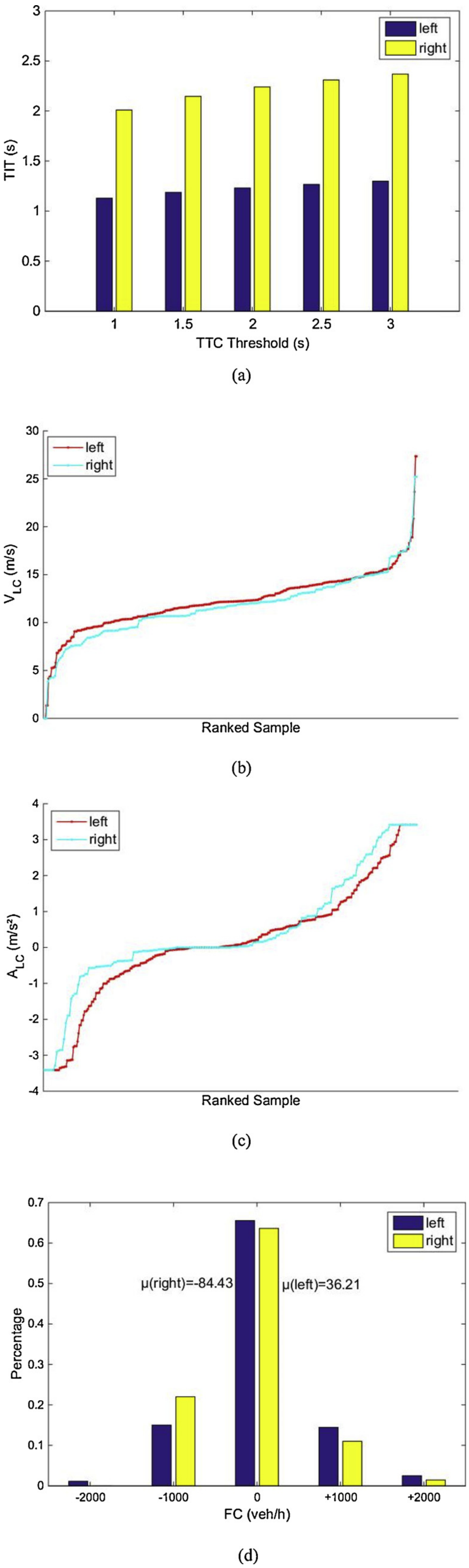


Fig. 9. Comparison between different lane-changing direction (a) Distribution of TIT; (b) Speed of lane changer; (c) Acceleration of lane changer; (d) Distribution of FC.

which tends to bring less adverse impacts on crash risks. We further calculated the speed difference between different lanes. The average speed difference between lanes of lane changes to the left is much small than that of lane changes to the right (which is 2.6367 and 4.5761 m/s, respectively). Thus, lane changes to the left are safer because of the smaller speed difference between lanes.

As for FC comparison, we found that the probability of increasing the flow for the lane change to the left lane was slightly larger than that caused by the lane change to the right lane (see Fig. 9d), which was intuitive because individual vehicle changed lane for a relatively fast speed benefiting the whole traffic operation (Note that vehicles in left lanes had higher speed in most cases).

7. Conclusions and discussion

This paper proposed a machine learning-based approach in order to predict the short-term impacts of lane-changing maneuvers on traffic safety measurements in traffic oscillations based on empirical vehicle trajectory data. Valid records of lane changes were extracted from the NGSIM database on two freeway sections including US-101 and I-80. We formulated an approach to calculate the LCI on crash risks while controlling the confounding effects of the oscillation propagation; in the meantime, the LCI on traffic operation were also calculated. The SVR models were estimated to predict the LCI based on a series of microscopic traffic parameters of individual vehicles before the inserting behaviors. The data on the US-101 was used to train the SVR models and to test the predicting performances, while the data on the I-80 was used to test for transferability. The relationship between the LCI on crash risks and the explanatory variables were explored via the sensitivity analyses in the SVR models. It was also compared the LCI for various vehicle types as well as different lane-changing directions.

The results have shown that the SVR models predicted well the LCI on traffic safety and flow based on the training and testing dataset on US-101. The predicting accuracy was slightly reduced when based on the validation dataset on I-80, but still indicated a good transferability of our proposed models. From the sensitivity analysis, it was found that almost all curves illustrated in Fig. 7(b) showed a concave trend, which means that it is possible to obtain an optimal speed and acceleration for the lane changer to achieve the lowest TIT value for safety maximization.

In addition, with the fastest speed and largest acceleration/deceleration rate in relatively congested region, motorcycles conducted lane changes with smaller gaps and larger speed differences, which brings the highest safety risks. Truck lane changes contributed to fewer but considerable crash risks. Automobiles were found to be the safest vehicle type when they conduct lane changes. As for FC caused by lane change, trucks' lane changes resulted in the largest flow reductions. And lane changes to the right were found to have more negative impacts on the traffic flow and crash risks if compared to those to the left. Previous study modelled lane-change related crashes with macroscopic traffic parameters and found that the lower traffic flow of the target lane may increase the chance of the occurrence of lane-change related crashes upstream (Cheng et al., 2017). This point is consistent with our study because the lower traffic flow usually appears on the right lanes and it found that lane changes to the right are more dangerous.

This study can support the development of lane change strategies in connected and automated driving environments in two ways. Firstly, when a running vehicle generates a need to change lane, the system can predict the LCI caused by the lane change in the current traffic status. If the TIT and FC all satisfy the predefined threshold (which are defined by the system preference), the lane change maneuver can be executed, vice versa. Secondly, the automated driving vehicle can proactively adjust its longitudinal driving action, such as reduce speed difference of the lane changer and surrounding vehicles, to improve the safety impact and avoid causing large flow reduction during changing lane. The model developed in our paper can provide quantitative evaluation in such decision making process.

Further, this study could promote a better understanding of the how various lane-changing maneuvers actually affect the traffic safety in the oscillation conditions. In order to estimate the impact of a lane change, it calculates the 11 microscopic traffic parameters in Table 1 as the input of the crash risk prediction model. The TIT value associated with the lane change at the traffic scenario can be predicted by the developed SVR model. A crash risk threshold can be applied to decide whether conducting the lane change or not. Some findings of the study are particularly useful for the development of the optimal lateral movement control strategies, especially for autonomous vehicles to reduce the negative safety impacts of lane changes.

Some other measurements such as weather, traffic volume and share of heavy vehicles could affect the results. However, due to the limitation of trajectory data, those measurements are relatively constant. It is hard to judge whether those factors have an impact on the results in our present study. Those factors can be included in our modeling framework to estimate their impacts when more trajectory data are available. In addition, we will propose the optimal lane change strategies based on the findings of the present study and conduct experimental tests of the safety impacts in the simulation environments.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was jointly supported by the National Natural Science Foundation of China (71871057, 51678132), and the Fundamental Research Funds for the Central Universities (2242018K41009).

References

Ahn, S., Cassidy, M.J., 2007. Freeway Traffic Oscillations and Vehicle Lane-change Maneuvers. In Transportation and Traffic Theory 2007. Papers Selected for Presentation at ISTTT17Engineering and Physical Sciences Research Council (Great Britain) Rees Jeffreys Road Fund Transport Research Foundation TMS Consultancy Ove Arup and Partners, Hong Kong Transportation Planning (International) PTV AG.
Alade, I.O., Bagudu, A., Oyehan, T.A., Rahman, M.A.A., Saleh, T.A., Olatunji, S.O., 2018. Estimating the refractive index of oxygenated and deoxygenated hemoglobin using genetic algorithm-support vector machine approach. Comput. Methods Programs Biomed. 163, 135–142.
Allen, B.L., Shin, B.T., Cooper, D.J., 1978. Analysis of traffic conflicts and collision. Transp. Res. Rec. 667, 67–74.
Bao, J., Liu, P., Qin, X., Zhou, H., 2018. Understanding the effects of trip patterns on spatially aggregated crashes with large-scale taxi GPS data. Accid. Anal. Prev. 120, 281–294.
Bao, J., Liu, P., Ukkusuri, S.V., 2019. A spatiotemporal deep learning approach for city-wide short-term crash risk prediction with multi-source data. Accid. Anal. Prev. 122, 239–254.
Bertini, R.L., Leal, M.T., 2005. Empirical study of traffic features at a freeway lane drop. J. Transp. Eng. 131, 397–407.
Cassidy, M.J., Rudjanakanoknad, J., 2005. Increasing the capacity of an isolated merge by metering its on-ramp. Transp. Res. Part B Methodol. 39, 896–913.
Chang, C.C., Lin, C.J., 2011. LIBSVM: a library for support vector machines. ACM Trans. Intel. Syst. Technol. (TIST) 2, 27.
Chen, D., Laval, J., Ahn, S., Zheng, Z., 2012a. Microscopic traffic hysteresis in traffic oscillations: a behavioral perspective. Transp. Res. Part B Methodol. 46 (10), 1440–1453.
Chen, D., Laval, J., Zheng, Z., Ahn, S., 2012b. A behavioral car-following model that captures traffic oscillations. Transp. Res. Part B Methodol. 46 (6), 0–761.
Chen, D., Ahn, S., 2018. Capacity-drop at extended bottlenecks: merge, diverge, and weave. Transp. Res. Part B Methodol. 108, 1–20.
Chen, Z., Qin, X., Shaon, M.R.R., 2018. Modeling lane-change-related crashes with lane-specific real-time traffic and weather data. J. Intell. Transp. Syst. Technol. Plan. Oper. 22 (4), 291–300.
Cheng, K., Lu, Z., Wei, Y., Shi, Y., Zhou, Y., 2017. Mixed kernel function support vector regression for global sensitivity analysis. Mech. Syst. Signal Process. 96, 201–214.
Cheu, R.L., Martinez, J., Duran, C., 2009. A cell transmission model with lane changing and vehicle tracking for port of entry simulations. Transp. Res. Rec. 2124, 241–248.
Coifman, B., Mishalani, R., Wang, C., Krishnamurthy, S., 2006. Impact of lane-change maneuvers on congested freeway segment delays: pilot study. Transp. Res. Rec.: J. Transp. Res. Board 1965, 152–159.
Coifman, B., Li, L., 2017. A critical evaluation of the Next Generation Simulation (NGSIM) vehicle trajectory dataset. Transp. Res. Part B Methodol. 105, 362–377.
Cooper, D.F., Ferguson, N., 1976. Taffic Studies at T-junctions. 2. A Conflict Simulation Record. Traffic Eng. Control 17 (Analytic).
Elefteriadou, L., Roess, R.P., McShane, W.R., 1995. Probabilistic nature of breakdown at freeway merge junctions. Transp. Res. Rec. 1484.
Fish, K.E., Blodgett, J.G., 2003. A visual method for determining variable importance in an artificial neural network model: an empirical benchmark study. J. Target. Meas. Anal. Mark. 11, 244–254.
Golob, T.F., Recker, W.W., Alvarez, V.M., 2004. Safety aspects of freeway weaving sections. Transp. Res. Part A Policy Pract. 38, 35–51.
Gettman, D., Head, L., 2003. Surrogate safety measures from traffic simulation models. Transp. Res. Rec. 1840, 104–115.
Guo, Y., Li, Z., Wu, Y., Xu, C., 2018. Exploring unobserved heterogeneity in bicyclists' red-light running behaviors at different crossing facilities. Accid. Anal. Prev. 115, 118–127.
Hadayeghi, A., Shalaby, A.S., Persaud, B.N., 2010. Development of planning level transportation safety tools using geographically Weighted Poisson Regression. Accid. Anal. Prev. 42, 676–688.
Hansson, A., 1975. Studies in Driver Behavior, With Application in Traffic Design and Planning: Two Examples (No. Bulletin 9 Monograph).
Haworth, J., Shawe-Taylor, J., Cheng, T., Wang, J., 2014. Local online kernel ridge regression for forecasting of urban travel times. Transp. Res. Part C Emerg. Technol. 46, 151–178.
Hayward, J.C., 1972. Near-miss determination through use of a scale of danger. Highway Res. Record 384, 24–34.
Hirst, S., Graham, R., 1997. The format and presentation of collision warnings. Ergon. Saf. Intell. Driv. Interfaces 203–219.
Hupfer, C., 1997. Deceleration to safety time (DST)—a useful figure to evaluate traffic safety. In ICTCT Conference Proceedings of Seminar (3).
Hydén, C., 1987. The Development of a Method for Traffic Safety Evaluation: The Swedish Traffic Conflicts Technique. Bull. Lund Inst. Technol. Department 70.
Iranitalab, A., Khattak, A., 2017. Comparison of four statistical and machine learning methods for crash severity prediction. Accid. Anal. Prev. 108, 27–36.
Jiménez, F., Naranjo, J.E., García, F., 2013. An improved method to calculate the time-to-collision of two vehicles. Int. J. Intell. Transp. Syst. Res. 11, 34–42.
Jin, W.L., 2010. A kinematic wave theory of lane-changing traffic flow. Transp. Res. Part B Methodol. 44, 1001–1021.
Jin, W.L., 2013. A multi-commodity Lighthill-Whitham-Richards model of lane-changing traffic flow. Procedia-Social Behav. Sci. 80, 658–677.
Karlaftis, M.G., Vlahogianni, E.I., 2011. Statistical methods versus neural networks in transportation research: differences, similarities and some insights. Transp. Res. Part C Emerg. Technol. 19, 387–399.
Khodayari, A., Ghaffari, A., Kazemi, R., Braunstingl, R., 2012. A modified car-following model based on a neural network model of the human driver effects. IEEE Trans. Syst. Man Cybern. A. Syst. Hum. 42, 1440–1449.
Laval, J.A., Daganzo, C.F., 2006. Lane-changing in traffic streams. Transp. Res. Part B Methodol. 40, 251–264.
Laval, J., Cassidy, M., Daganzo, C., 2007. Impacts of lane changes at merge bottlenecks: a theory and strategies to maximize capacity. Traffic and Granular Flow'05. Springer, Berlin, Heidelberg, pp. 577–586.
Laval, J.A., Leclercq, L., 2008. Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model. Transp. Res. Part B Methodol. 42, 511–522.
Lee, J., Cassidy, M.J., 2008. An Empirical and Theoretical Study of Freeway Weave Bottlenecks – Escholarship. University of California Transportation Center Working Papers.
Li, X., Lord, D., Zhang, Y., Xie, Y., 2008. Predicting motor vehicle crashes using support vector machine models. Accid. Anal. Prev. 40, 1611–1618.
Li, X., Sun, J.Q., 2017. Studies of vehicle lane-changing dynamics and its effect on traffic efficiency, safety and environmental impact. Phys. A Stat. Mech. Its Appl. 467, 41–58.
Li, Z., Liu, P., Wang, W., Xu, C., 2012. Using support vector machine models for crash injury severity analysis. Accid. Anal. Prev. 45, 478–486.
Li, Z., Ahn, S., Chung, K., Ragland, D.R., Wang, W., Yu, J.W., 2014. Surrogate safety measure for evaluating rear-end collision risk related to kinematic waves near freeway recurrent bottlenecks. Accid. Anal. Prev. 64 (1), 52–61.
Li, Y., Wang, H., Wang, W., Liu, S., Xiang, Y., 2016. Reducing the risk of rear-end collisions with infrastructure-to-vehicle (I2V) integration of variable speed limit control and adaptive cruise control system. Traffic Inj. Prev. 17 (6), 597–603.
Li, Y., Li, Z., Wang, H., Wang, W., Xing, L., 2017. Evaluating the safety impact of adaptive cruise control in traffic oscillations on freeways. Accid. Anal. Prev. 104, 137–145.
Marković, N., Milinković, S., Tikhonov, K.S., Schonfeld, P., 2015. Analyzing passenger train arrival delays with support vector regression. Transp. Res. Part C Emerg. Technol. 56, 251–262.
Minderhoud, M.M., Bovy, P.H., 2001. Extended time-to-collision measures for road traffic safety assessment. Accid. Anal. Prev. 33, 89–97.
Nilsson, L., Alm, H., Jansson, W., 1992. Collision avoidance systems: effects of different levels of task allocation on driver behaviour. VTI särtryck 182.
Oh, S., Yeo, H., 2012. Microscopic analysis on the causal factors of capacity drop in highway merging sections. In: The 91st Annual Meeting of the Transportation Research Board. Washington.
Oh, C., Park, S., Ritchie, S.G., 2006. A method for identifying rear-end collision risks using inductive loop detectors. Accid. Anal. Prev. 38 (2), 295–301.
Park, H., Oh, C., Moon, J., Kim, S., 2018. Development of a lane change risk index using vehicle trajectory data. Accid. Anal. Prev. 110, 1–8.
Pan, T.L., Lam, W.H., Sumalee, A., Zhong, R.X., 2016. Modeling the impacts of mandatory

and discretionary lane-changing maneuvers. Transp. Res. Part C Emerg. Technol. 68, 403–424.
Patire, A.D., Cassidy, M.J., 2011. Lane changing patterns of bane and benefit: observations of an uphill expressway. Transp. Res. Part B Methodol. 45, 656–666.
Pongnumkul, S., Pechprasarn, T., Kunaseth, N., Chaipah, K., 2014. Improving arrival time prediction of Thailand's passenger trains using historical travel times. In: Computer Science and Software Engineering (JCSSE), 2014 11th International Joint Conference on. IEEE. pp. 307–312.
Rahman, M.S., Abdel-Aty, M., 2018. Longitudinal safety evaluation of connected vehicles' platooning on expressways. Accid. Anal. Prev. 117, 381–391.
Shvetsov, V., Helbing, D., 1999. Macroscopic dynamics of multilane traffic. Phys. Rev. E 59, 6328.
Smola, A.J., Schölkopf, B., 2004. A tutorial on support vector regression. Stat. Comput. 14, 199–222.
Srivastava, A., Geroliminis, N., 2013. Empirical observations of capacity drop in freeway merges with ramp control and integration in a first-order model. Transp. Res. Part C Emerg. Technol. 30, 161–177.
Sultan, B., Brackstone, M., McDonald, M., 2002. Parameter analysis for collision avoidance systems. In: 9th World Congress on Intelligent Transport Systems ITS America. ITS Japan, ERTICO (Intelligent Transport Systems and Services-Europe).
Svenson, O., Eriksson, G., Gonzalez, N., 2012. Braking from different speeds: judgments of collision speed if a car does not stop in time. Accid. Anal. Prev. 45, 487–492.
Tastambekov, K., Puechmorel, S., Delahaye, D., Rabut, C., 2014. Aircraft trajectory forecasting using local functional regression in Sobolev space. Transp. Res. Part C Emerg. Technol. 39, 1–22.
Toledo, T., Koutsopoulos, H.N., Ben-Akiva, M., 2007. Integrated driving behavior modeling. Transp. Res. Part C Emerg. Technol. 15 (2), 96–112.
Treiber, M., Kesting, A., Helbing, D., 2006. Understanding widely scattered traffic flows, the capacity drop, and platoons as effects of variance-driven time gaps. Phys. Rev. E 74, 16123.
Thiemann, C., Treiber, M., Kesting, A., 2008. Estimating acceleration and lane-changing dynamics from next generation simulation trajectory data. Transport. Res. Rec.: J. Transport. Res. Board 2088, 90–101.
Van Winsum, W., Brookhuis, K.A., de Waard, D., 2000. A comparison of different ways to approximate time-to-line crossing (TLC) during car driving. Accid. Anal. Prev. 32 (1), 47–56.
Wang, C., Coifman, B., 2008. The effect of lane-change maneuvers on a simplified car-following theory. Ieee Trans. Intell. Transp. Syst. 9, 523–535.
Wang, C., Stamatiadis, N., 2013. Surrogate safety measure for simulation-based conflict study. Transp. Res. Rec. 2386, 72–80.
Wang, C., Stamatiadis, N., 2014. Derivation of a new surrogate measure of crash severity. Transp. Res. Rec.: J. Transp. Res. Board 2432, 37–45.
Wang, J., Shi, Q., 2013. Short-term traffic speed forecasting hybrid model based on chaos–wavelet analysis-support vector machine theory. Transp. Res. Part C Emerg. Technol. 27, 219–232.
Wang, M., Hoogendoorn, S.P., Daamen, W., van Arem, B., Happee, R., 2015. Game theoretic approach for predictive lane-changing and car-following control. Transp. Res. Part C Emerg. Technol. 58, 73–92.
Yang, Q., Koutsopoulos, H.N., 1996. A microscopic traffic simulator for evaluation of dynamic traffic management systems. Transp. Res. Part C Emerg. Technol. 4 (3), 113–129.
Yang, D., Zhu, L., Liu, Y., Wu, D., Ran, B., 2018. A novel car-following control model combining machine learning and kinematics models for automated vehicles. Ieee Trans. Intell. Transp. Syst. 20 (6), 1991–2000.
Yu, R., Abdel-Aty, M., 2013. Utilizing support vector machine in real-time crash risk evaluation. Accid. Anal. Prev. 51, 252–259.
Zhang, J., Li, Z., Pu, Z., Xu, C., 2018. Comparing prediction performance for crash injury severity among various machine learning and statistical methods. IEEE Access 6, 60079–60087.
Zheng, Z., Ahn, S., Monsere, C.M., 2010. Impact of traffic oscillations on freeway crash occurrences. Accid. Anal. Prev. 42 (2), 626–636.
Zheng, Z., Ahn, S., Chen, D., Laval, J., 2011. Freeway traffic oscillations: microscopic analysis of formations and propagations using wavelet transform. Procedia-Social Behav. Sci. 17, 702–716.
Zheng, Z., Ahn, S., Chen, D., Laval, J., 2013. The effects of lane-changing on the immediate follower: anticipation, relaxation, and change in driver characteristics. Transp. Res. Part C Emerg. Technol. 26, 367–379.
Zheng, Z., 2014. Recent developments and research needs in modeling lane changing. Transp. Res. Part B Methodol. 60, 16–32.
Zhou, M., Qu, X., Li, X., 2017. A recurrent neural network based microscopic car following model to predict traffic oscillation. Transp. Res. Part C Emerg. Technol. 84, 245–264.