Australasian Transport Research Forum 2023 Proceedings
29 November – 1 December, Perth, Australia
Publication website: http://www.atrf.info

A safety-oriented microscopic traffic model

Seyedehsan Seyedabrishami[1], Mohsen Ramezani[1]

[1] School of Civil Engineering, The University of Sydney

Email for correspondence (presenting author): seyedehsan.seyedabrishami@sydney.edu.au

Abstract

The car following model (CFM) and lane changing model (LCM) are fundamental components of microscopic traffic models. While these models are primarily designed to ensure collision-free traffic operations, predicting potential crashes is equally critical for effective pre- and post-crash management. In this study, we have innovatively developed CFM and LCM models to model possible crashes microscopically. The models aim to accurately recreate the distribution of the distance between adjacent vehicles in different lanes, which enables the likelihood of forecasting potential crashes. To this end, we use linear regression models initially calibrated for CFM and LCM through an iterative optimization model that minimizes the error of distances between adjacent vehicles. The histograms of simulated distances between vehicles demonstrate the close similarity between model outputs and real-world observations. Additionally, the model accurately reconstructs the distribution of vehicle speeds in real-world traffic flow. This model can be used to predict potential crashes in transport infrastructure where there is no crash data or where newly introduced systems are in operation.

1. Introduction

The majority of safety research comprises passive studies, which rely on historical observations of crash and near-crash indices. Consequently, researchers require crashes or hazardous situations to calibrate descriptive and prescriptive safety models that identify the most influential factors affecting safety in the study area. However, the need to forecast crashes before they occur has been less studied. Safety simulation models can serve as a viable solution to identify high-risk segments or spots for crashes. The majority of existing microscopic traffic models utilize the collision-free car following model (CFM) and lane changing model (LCM). These models assume a minimum longitudinal and lateral gap between vehicles, which prevents any crash occurrence in traffic simulations. Nonetheless, neglecting the inevitability of crash occurrences can lead to misestimation of traffic safety characteristics.

Existing traffic simulation models tend to simulate traffic behavior using CFM and LCM, generating particular interactions such as intersection conflicts and rear-end conflicts between vehicles to assess safety conditions. Previous research in road safety simulation demonstrates that most researchers utilize the existing CFM and LCM models to measure safety performance. Given that CFM and LCM are mostly collision-free, these studies focus on traffic conflicts, surrogate safety measures, and near-crashes as the safety performance measures (Young et al., 2014).

This paper contributes to the literature by introducing and calibrating new CFM and LCM models that are not collision-free and using them to develop a microscopic simulation model for evaluating traffic safety on a straight stretch of road. The simulation model results in accurate distances between adjacent vehicles.

2. Problem description

Microscopic traffic simulation models usually lack the capability of predicting crashes, rendering them unsuitable for evaluating the safety performance of a road network that lacks crash data or features newly built infrastructures. Additionally, identifying high-risk road segments necessitates observing crashes that occur, with no possibility of predicting them before their occurrence.

This research aims to develop a safety-oriented microscopic traffic model, with the accurate determination of distances between adjacent vehicles being the most crucial factor for calibrating, validating and evaluating this model. If the model can determine the distribution of distances between adjacent vehicles accurately, it is expected to display crashes in situations where the distance between two adjacent vehicles reaches zero. The CFM and LCM models do not make any assumptions regarding collision-free conditions or minimum distances.

Therefore, the primary challenge is developing appropriate CFM and LCM that can be incorporated into a microscopic simulation model to accurately generate realistic traffic flow beahviours while also producing the distance between vehicles, especially short distances that may lead to crashes. Furthermore, the simulation model must exhibit satisfactory performance in modelling other traffic flow characteristics, such as speed and acceleration.

3. Methodology

This section outlines the various steps involved in developing the proposed simulation model.

3.1 Conceptual model

To construct the simulation model, two main components were required: the CFM and LCM. The former is responsible for the longitudinal movement of each vehicle, while the latter handles the lateral movements and any lane changes. In our proposed model, the output of each CFM and LCM is the acceleration of the vehicle in the X and Y directions, respectively.

The difference of the distance, speed and acceleration of adjacent vehicles with vehicle and the speed and the acceleration of all adjacent vehicles and the vehicle will be assumed to be influential on the acceleration of the vehicle in current time frame in both directions X and Y. After testing different combinations of influential variables, to predict a vehicle's acceleration in the X direction, the CFM utilizes information from the previous timeframe, including the vehicle's acceleration and speed in the X direction, differences in acceleration with its preceding vehicle, the acceleration of the following vehicle, and differences in speed with the left alongside vehicle. Meanwhile, the LCM predicts a vehicle's acceleration in the Y direction based on information from the previous frame, including its acceleration and speed in the Y direction, the difference between its Y coordinate and the preceding vehicle, the acceleration of the preceding vehicle in the Y direction, and the distance between the vehicle and the left alongside vehicle. It is assumed that the CFM and LCM function are linear and ordinary least square regression models can be applied to estimate the parameters.

$$\begin{aligned} xAcc_t = 1.003 * xAcc_{t-1} + 5.4E-06 * xSpd_{t-1} - .003 * \Delta xAccP_{t-1} \\ - 0.003 * xAccF_{t-1} - 1.08E-05 * \Delta xSpdLA_{t-1} \end{aligned} \tag{1}$$

$$\begin{aligned} yAcc_t = 1.001 * yAcc_{t-1} - 0.009 * ySpd_{t-1} + 1.65E-05 * \Delta yP_{t-1} \\ + 0.0018 * yAccP_{t-1} - 6.18E-06 * \Delta yLA_{t-1} \end{aligned} \tag{2}$$

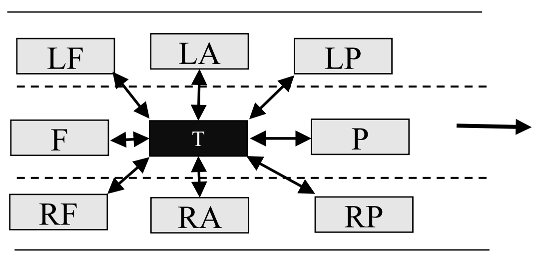
Where:

$xAcc_t$ and $yAcc_t$ = the acceleration of vehicle in the X and Y directions at frame t, respectively,

$xSpd_{t-1}$ and $ySpd_{t-1}$ = the speed of vehicle in the X and Y directions at frame $t-1$, respectively,

$\Delta xAccP_{t-1}$= the difference of acceleration between target vehicle and preceding vehicle in the X direction at time frame $t-1$,
$xAccF_{t-1}$= the acceleration of the following vehicle in X direction at frame time $t-1$,
$\Delta xSpdLA_{t-1}$= the difference in speed between the left alongside vehicle and the target vehicle at time frame $t-1$,
ΔyP_{t-1}= the difference in the Y-coordinate between the preceding vehicle and the target vehicle at time frame $t-1$,
$yAccP_{t-1}$= the acceleration of preceding vehicle in Y direction at time frame $t-1$,
ΔyLA_{t-1}= The difference in the Y-coordinate between the left alongside vehicle and the target vehicle at time frame $t-1$,
The estimated coefficients from regression are the initial parameters that will be optimised in an iterative optimisation algorithm to minimize the error function. The error function is the root mean square of distance between target vehicle and its surrounding vehicles as it is shown in Figure 1. The parameters shall be meticulously fine-tuned to effectively mitigate the error computed across a sequence of 50 frames during the simulation process. The underpinning rationale of adopting this method is predicated upon the well-founded observation that collisions frequently arise due to the convergence of vehicles from various directions in close proximity. Hence, the imperative lies in possessing a simulation framework that impeccably estimates inter-vehicular distances with utmost precision.

Figure 1: Definition of Distance between Target Vehicle with Surrounding Vehicles (P: Preceding, F: Following, LP: Left Preceding, LA: left Alongside, LF: Left Following, RP: Right Preceding, RA: Right Alongside, RF: Right Following, T: Target)



Lane changes occur when the sum of a vehicle's movements over consecutive frames exceeds a predefined threshold, then the lane of the vehicle is changed.
For each frame, the acceleration of the target vehicle is updated based on its CFM and LCM, and accordingly, its new location and speed are computed. This process is repeated for all vehicles included in the model for each timeframe until they leave the section, i.e., when their X and Y coordinates fall outside of the predefined ranges.

3.2 Calibration

The present models incorporate several key parameters, namely, the coefficients within CFM and LCM, the frequency of recalibration contingent upon the time frames, and the duration of each time frame. Notably, these parameters remain modifiable, affording the modeler a high degree of flexibility when deploying the simulation in diverse geographical contexts. Such adjustability facilitates the adaptability and applicability of the simulation to various locations, thereby enhancing its utility as a versatile tool for traffic analysis and forecasting. Ensuring an accurate reconstruction of real-world traffic, the simulation must be properly calibrated. The calibration process involves several steps, as outlined in the following procedure:
Step 1: Identify the first frame of each vehicle, which represents the point of entry onto the road section. This frame contains initial data such as the vehicle's location, speed, and acceleration.

Step 2: Specify the simulation update time, which is set at 0.04 seconds, as well as the total simulation time. Additionally, establish the initial parameters of the linear regression models for the car following and lane changing models. Set the lane change threshold to 2 meters.
Step 3: While the simulation time is less than the total simulation time, perform the following tasks:

- 3.1. Check the location of each vehicle to ensure that it is within the road segment. If not, remove it from the simulation process.
- 3.2 Update the location, speed, and acceleration of each vehicle in the first frame data using the car following and lane changing models.
- 3.3 Determine which vehicles are adjacent to each other based on their location, height, and width, and identify their position in relation to each other (e.g., same lane, left or right lane, preceding, following, or alongside).
- 3.4 If the sum of a vehicle's movement in sequential frames in Y direction exceeds 2 meters, change the lane. Otherwise, keep the vehicle in its existing lane.
- 3.5 Every 50 frames do the following process,
    - 3.5.1 Calculate the root mean square error of the distances between adjacent vehicles. This error represents the difference between the distances obtained from the simulation model and the real-world data.
    - 3.5.2 Adjust the coefficients of the car following and lane changing models to minimise the root mean square error, and update the coefficients accordingly using Nelder-Mead as the solver which is a downhill simplex method, is a commonly used optimization algorithm for nonlinear optimization problems.

Step 4: End the calibration process.
By following these steps, the simulation can be effectively calibrated to accurately replicate real-world traffic conditions, facilitating the forecasting of potential accidents and enabling better pre- and post-accident management.
This study employed the method of batch optimization to iteratively update the parameters of both CFM and LCM models. This optimization technique was applied using a subset of sequential time frames extracted from the available data. The utilization of batch optimization necessitates the availability of a comprehensive training dataset, as the principal objective of this investigation centers on the construction of precise traffic data tailored explicitly to the analysed location. It is important to note that the focus of this research lies in developing location-specific traffic models, rather than seeking generalizability to other geographical contexts.

3.3 Data description

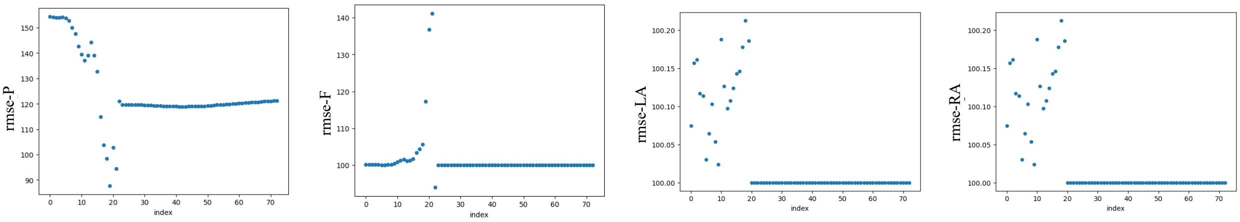
To obtain realistic traffic data, we utilized the HighD dataset [Krajewski et al., 2018], which is a collection of real-world vehicle trajectories captured by a drone on German highways. Computer vision algorithms were employed to extract the trajectories from the recordings, and neural networks were used to detect and localize vehicles in each frame. Bayesian smoothing was then applied to smooth the trajectories, resulting in a dataset of 110,500 vehicles.
The dataset contains valuable information about each vehicle, including its location, width, height, X and Y speed and acceleration, and availability of left and right preceding, following, and alongside vehicles in a given frame that is every 0.04 second. For each target vehicle, up

to eight surrounding vehicles in the front, back, left, and right lanes may impact the safety conditions in that segment during a specific timeframe.

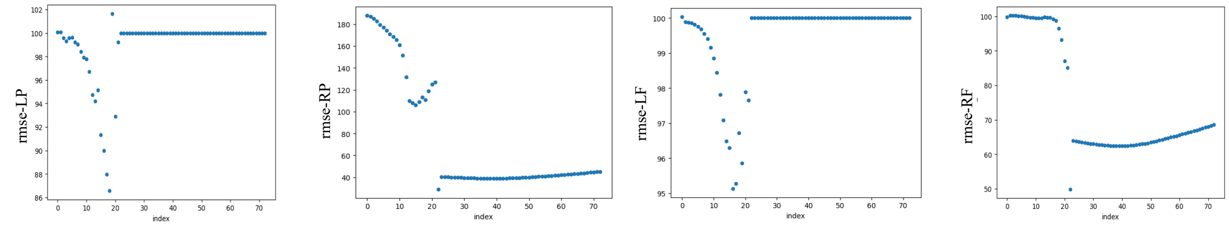
3.4 Verification and validation

To ensure the accuracy and effectiveness of the model, an analysis is conducted to evaluate the impact of the model's parameters. The initial parameters of the car following model (CFM) and lane changing model (LCM) were obtained from real-world data. During the optimisation process of adjusting the CFM and LCM coefficients, the root mean square error of distances is recorded at each iteration of the optimisation. The resulting error trends are depicted in Figure 2, which displays the variation of error terms for all eight distances.

Figure 2: Variation of the root mean square error (meter) of distances throughout the optimization iterations



(a) Preceding (b) Following (c) Left Alongside (d) Right Alongside

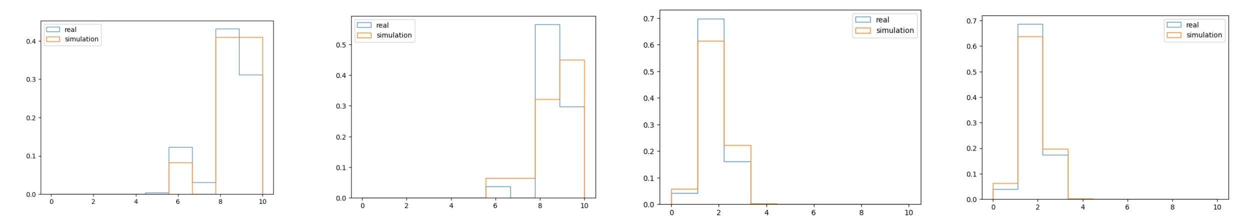


(e) Left Preceding (f) Right Preceding (g) Left Following (h) Right Following

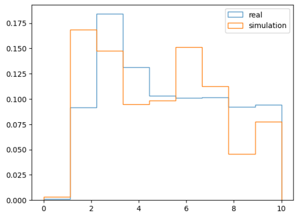
Figure 2 demonstrates that the error terms gradually decrease as the iteration progresses. However, due to the implementation of a batch optimisation technique, the optimisation algorithm adjusts the coefficients every 50 frames, and the permanent decrease in all errors may not be apparent. Due to our utilization of a combination of normalized error terms for each distance in the error function, we observe a general decrease in errors across most distances. However, it is worth noting that errors related to left proceeding and left following distances do not exhibit the same decreasing trend. This could potentially be attributed to the scarcity of samples involving such vehicle configurations in our dataset.

The results of the simulation were validated by comparing the histograms of vehicle distance with surrounding vehicles to the real-world histograms. Figure 3 displays eight histograms for eight different positions around a vehicle, with the X-axis indicating distances ranging from 0 to 10 meters (every bin is 1 meter), and the Y-axis representing the relative frequency for each distance interval.

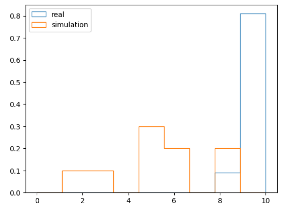
Figure 3: Comparison of simulation and real-world distance (meter) between adjacent vehicles



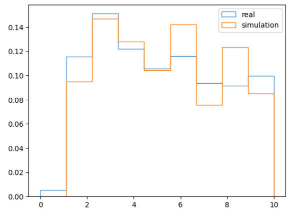
(a) Preceding (b) Following (c) Left Alongside (d) Right Alongside



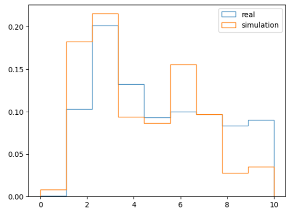
(e) Left Preceding



(f) Right Preceding



(g) Left Following



(h) Right Following

As demonstrated in Figure 3, the histograms of distances between vehicles in the simulation are highly similar to those of the ground truth measurements, indicating that the simulation model's performance is accurate enough to effectively monitor traffic safety performance. While most of the histograms' trends are similar between the simulation and observation, the right preceding simulation histogram is slightly different from the observation. However, this difference is not considered a significant concern, as it has a minimal impact on the target vehicle. These histograms depict the lateral distance distribution, where the left and right alongside distances range from 0 to 4 meters, while the remaining distances range from 0 to 10 meters. The models reveal that the distribution of close distances is similar, indicating that such close distances are more prone to potential crashes.

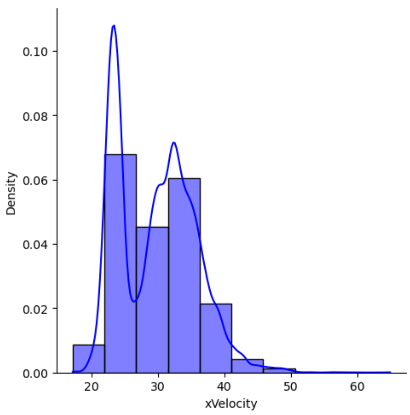
Overall, the R-squared values between the observed distances in reality and the simulated distances for the 8 distances with the surrounding vehicles range from 0.83 to 0.94. These high R-squared values, along with the visual comparison of histograms, demonstrate that the developed model accurately replicates real-world traffic conditions. This indicates that the model can be effectively utilized to monitor traffic safety performance with confidence.

4. Results and discussion

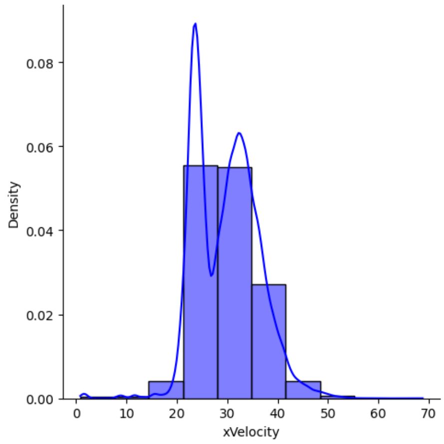
The simulation results indicate that the proposed CFM and LCM are effective in generating traffic flow that closely resembles the relative distances between vehicles in the real world. However, it is also essential that other traffic parameters, such as the speed and acceleration of vehicles in the subsequent time frames, are similar to real-world observations to achieve a reliable simulation model. Figure 4 depicts the distributions of vehicle speeds in the simulation and reality. Although the calibration process is primarily aimed at minimizing the distances between adjacent vehicles, the speed distribution of vehicles is also somewhat similar. The speed range varies from 20 to 50 kilometers per hour, and both distributions have similar peak and non-peak values. Furthermore, the simulation outputs are smoother than real-world observations, as expected.

Since the primary objective of this microscopic traffic simulation model is to predict the likelihood of crashes over a long-term interval, all calibration and validation indices used in this simulation model are aggregate measures.

Figure 4: Comparison of simulation and real-world speed distribution



(a) Real



(b) Simulation

5.Conclusion and future directions

This study introduces a novel microscopic traffic simulation model that employs newly proposed and calibrated CFM and LCM to simulate the distances between adjacent vehicles during the simulation. Unlike previous models that utilize collision-free CFM and LCM, our integrated models aim to reconstruct vehicle relative distances that may potentially forecast crashes in reality. The comparison of aggregate indices, such as the histograms of vehicle distances and the distribution of vehicle speeds, indicates that our simulation model is suitable for this purpose.

During the simulation, the location, speed, and acceleration of vehicles are updated using CFM and LCM at predefined intervals.

In order to improve the performance of CFM and LCM, future directions include the development of simulation software that incorporates nonlinear models or machine learning algorithms, which take into account more variables to predict the next frames' variables. This will likely enhance the accuracy and precision of the models.

In addition, efforts are underway to locate data that contain crash events, with the goal of testing whether the developed models can effectively predict such events. Initially, the objective is to forecast the number of crashes over a long-term interval, such as one year. This involves simulating traffic conditions for a year to estimate the probability of a crash occurrence. In the final stage, the aim is to predict crashes on a daily basis, which requires the use of more rigorous calibration indices. Instead of utilizing histograms, the discrepancies between each vehicle's trajectory in the simulation and the observation should be minimized.

Overall, these future directions are geared towards enhancing the performance and applicability of the CFM and LCM models, with the ultimate goal of improving traffic safety management.

6. References

Young, W., Sobhani, A., Lenne, M., Sarvi, M. 2014, ‘Simulation of safety: A review of the state of the art in road safety simulation modelling’, Accident Analysis and Prevention, Vol. 66, pp. 89-103.

Krajewski R., Bock J., Kloeker L., Eckstein L., (04-07 November 2018), ‘The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems’, [conference presentation], United State of America, accessed 09 December 2018, doi:10.1109/ITSC.2018.8569552