Research Article

Development of Macroscopic Cell-Based Logistic Lane Change Prediction Model

**Christina Ng ¹, Susilawati Susilawati ¹, Md Abdus Samad Kamal ²,
and Irene Mei Leng Chew¹**

¹*School of Engineering, Monash University Malaysia, Jalan Lagoon Selatan, Bandar Sunway, 47500, Subang Jaya, Selangor, Malaysia*

²*Division of Mechanical Science and Technology, Graduate School of Science and Technology, Gunma University, 1-5-1 Tenjincho, Kiryu 376-8515, Japan*

Correspondence should be addressed to Susilawati Susilawati; susilawati@monash.edu

Received 8 June 2021; Revised 2 August 2021; Accepted 13 August 2021; Published 2 September 2021

Academic Editor: Jing Zhao

Copyright © 2021 Christina Ng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper aims at developing a macroscopic cell-based lane change prediction model in a complex urban environment and integrating it into cell transmission model (CTM) to improve the accuracy of macroscopic traffic state estimation. To achieve these objectives, first, based on the observed traffic data, the binary logistic lane change model is developed to formulate the lane change occurrence. Second, the binary logistic lane change is integrated into CTM by refining CTM formulations on how the vehicles in the cell are moving from one cell to another in a longitudinal manner and how cell occupancy is updated after lane change occurrences. The performance of the proposed model is evaluated by comparing the simulated cell occupancy of the proposed model with cell occupancy of US-101 next generation simulation (NGSIM) data. The results indicated no significant difference between the mean of the cell occupancies of the proposed model and the mean of cell occupancies of actual data with a root-mean-square-error (RMSE) of 0.04. Similar results are found when the proposed model was further tested with I80 highway data. It is suggested that the mean of cell occupancies of I80 highway data was not different from the mean of cell occupancies of the proposed model with 0.074 RMSE (0.3 on average).

1. Introduction

The autonomous vehicle (AV) is able to detect its surroundings, identify obstacle, and signages that, in turn, enable it to develop its navigation paths and move along the road network without human control. Current AV research projects have focused on developing AV technologies to improve vehicle safety, particularly when performing its fundamental tasks, including car following, lane-keeping, and lane-changing [1–6]. Due to its complex process, which involves interactions of host vehicles and its following and leading vehicles, little attention has been given to the AV lane change prediction model. The existing lane-change algorithm (e.g., trajectory planning and maneuver planning algorithm) focused on maximizing individual vehicles' benefit [1–6]. In addition to that, these algorithms required

detailed microscopic traffic variables (i.e., relative speed and positions) of the surrounding subject vehicles, gaps between the host and following and leading vehicles, which often depends not only on the behavior and movement of the surrounding vehicles but also on the macroscopic traffic dynamics.

To assist AV control systems in predicting lane change occurrence under various traffic conditions, a macroscopic cell-based lane change prediction model that considers lateral vehicle movement from one lane to the other is introduced [7]. This cell-based macroscopic model discretizes the road section into a uniform cell form in which macroscopic traffic variables (e.g., traffic density and average vehicle speed) can be easily generated. These macroscopic traffic variables have been considered as traffic dynamic indicators that can be further used to predict lane change

occurrence. By understanding the behaviour of the surrounding traffic characteristics, the aggregate behaviour of lane change vehicles in any traffic stream is much easier to be observed than the disaggregate behaviour of individual vehicles. This way is also carried out due to the lack of studies that include the dynamics of lane change to compute multilane traffic flow under macroscopic conditions. To compute the possible solution of multilane traffic that fulfils the conservation law of traffic, a suitable macroscopic traffic flow model that can describe drivers' lane change behaviour in multilane highways is therefore needed.

Therefore, this study proposes a macroscopic cell-based lane change prediction model in a complex urban environment and integrates it into the cell transmission model (CTM). It has been done by refining the CTM flow and occupancy formulations to facilitate flow formulation when vehicles in the cell transfer from one cell to another in a longitudinal manner and update cell occupancy after a lane change.

The remainder of this paper is organized as follows. Section 2 introduces the CTM and development of the lane change prediction model, followed by basic terminology of the cell transmission model. This is then followed by the proposed methodology, leading to the development of the proposed macroscopic traffic simulation tool. Particularly, this section will guide us on how to develop the binary logistic lane change model and formulate the lane change occurrence. It will be followed by model development to integrate it into the cell transmission model by refining the CTM flow and occupancy formulations. Section 4 provides the steps undertaken in aggregating the real data in the preparation for the validation with the simulated CTM. Numerical results comparing the CTM and the field data are given in Section 5. Last, the discussion and conclusion of the paper are provided in Section 6.

2. Cell Transmission Model and CTM Based Lane Change Model

CTM is one the most widely used cell-based macroscopic traffic flow models developed by [8, 9]. It provides a convergent approximation to a simplified version of the Lighthill, Witham, and Richard (LWR) hydrodynamic model [10]. This model has been recognized as the most straightforward means, where only a few parameters are needed to explain the evolution of traffic features and dynamics [8]. It is known as the numerical estimation model of the LWR model. While fewer parameters are needed compared to the microscopic models, some limitations still exist. In the original CTM model, traffic flow has been modelled uniformly in a single stream over multilane highway traffic, assuming uniform and lateral distribution [8, 9]. The assumption as such may inaccurately represent the feature of the heterogeneous traffic flow distribution (i.e., lane-specific flow, density, speed, and vehicle types) observed in real-world traffic, which varies over time and space [11, 12]. Hence, the complex traffic feature of multilane roadways, specifically the lane change behavior, may not be well captured when modeling them as a single-lane traffic

stream [13, 14]. It is also known that frequent LC events result in traffic flow instabilities as a result of the drop in capacity as well as the formation and propagation of stop-and-go waves [15–20].

To fill the gaps, [21], [22], and [23] introduced the multiple lanes LWR model in which single links consist of two or more lanes in parallel. This extension enables the inflows from the neighboring lanes and outflows to neighboring lanes to be included in the conservation equation at each point that allows lane-changing occurrences. A multicommodity model was developed by [24] based on the theory of the LWR [10, 25] model. In this model, a new concept of lane change intensity was introduced, in addition to a new definition of the fundamental diagram and the introduction of an entropy condition. Other methods use the gas-kinetic traffic flow model [26], as well as a continuum and noncontinuum models [27]. The subject of modeling the lane flow distribution on specific sites of the traffic network has also been addressed in certain works [28]. During the free flow and congested regimes, [28] established a link between total density and lane densities. Following that, the number of lane changes is evaluated as a function of numerous incentives, with the findings indicating that the densities in the origin and target lanes are the most significant [29]. Duret et al. [30] examined real data from a three-lane highway in free-flow settings and came up with a simple linear model that took the lane distribution of traffic flow into account. Other previous research that looked at lane change in macroscopic models can be found in [24, 26, 27, 31, 32], and others that focused on specific locations of the road can be found in [28, 30, 33].

A few studies have considered the lane change features in the development of the CTM model. Wang et al. [34], for instance, assigned a fixed percentage of left-turn flow (i.e., 30%) when formulating the diverge movement to simulate oversaturated arterials. The improved form of the CTM model, which introduced a novel conditional cell at the intersection, has proven the reliability with the actual traffic flow conditions found in Highway Capacity Manual (HCM). Carey et al. [31] developed analytic network models that include lane-changing by extending the single-lane discretized CTM to two or more lanes and allowing traffic movement between lanes. This extension focused on mandatory lane changing by assuming that lane-changing motivation was influenced by driver's relative position to the entry and exit links and the corresponding levels of urgency. Carey et al. [31] introduced the fixed value of w_i^t (i.e., the number of vehicles that wish to change lane at cell i , time step t) as a variable to determine the cell occupancies in the following time step. While the theoretical development of the CTM has been continuously improved over years, majority of them only focused on improving the accuracy and computational effort of the previous method, but the accuracy and stability of the method have not been validated with actual field data [35].

Pan et al. [36] developed a mesoscopic multilane CTM considering the simultaneous simulation of both discretionary and mandatory lane change behaviours. The model realistically simulates the dynamics of heterogeneous lane

flow distribution based on a lane-specific fundamental diagram. Different priority levels are recognized based on varying lane change motivations and respective levels of urgency. A recursive lane change demand estimating technique was devised, which takes into account the role of urgency on the longitudinal probability distribution function of lane change movements. Most present models assumed that drivers are unaware of downstream traffic conditions before deciding to make a lane shift, which is impractical. However, [36] proposed a model that posits that drivers can make lane change decisions, with its effect decreasing exponentially as the distance from their current location increases. Meanwhile, their empirical analysis included some major lane change implications (e.g., capacity loss and flow balancing impact of discretionary lane shift).

Limitations still exist when incorporating lane change into CTM, particularly when fixed values were assigned to estimate the influence of lane change on inflow and outflow traffic. This assumption might not empirically represent actual lane change events as the surrounding traffic environment (i.e., speed and density between lanes) [13, 37–39] or some unknown factors (i.e., driving attitude) might affect the occurrence of a lane change.

3. Development of Cell Transmission Model Based Logistic Lane Change Model

Before introducing the extended CTM with lane changing, a basic principle and basic formulation of single-lane CTM are presented in Section 2.

3.1. Basic Principles of Cell Transmission Model. The original CTM [8, 9] is derived based on the improved version of the LWR model, which describes a simplified form of relationship based on the macroscopic fundamental diagram.

This relationship is represented as follows:

$$f = \min\{V_f k, Q, W(k_{\text{jam}} - k)\}, \quad (1)$$

where f is the traffic flow, V_f is the free-flow speed (free-flow represents the speed of vehicles when density is zero), W is the backward propagation speed of disturbances when traffic is congested, Q is the capacity flow into the next section, and k_{jam} is the jam density of traffic.

CTM discretized road section into homogenous cells with equal-length cells and time intervals. The cell length, L , is the distance that a typical vehicle traveled at free-flow speed in a single time interval ($L = Vt$). Under free-flow condition and no congestion, all the vehicles in a cell are expected to advance to the next with each time interval. This can be represented by the cell evolution given as follows:

$$n_{i+1}(t+1) = n_i(t), \quad (2)$$

where $n_i(t)$ is the number of vehicles in cell i at time t .

The boundary conditions of the cells are provided to control the flow of vehicles entering and exiting the first and last cell of all lanes, respectively. These are, respectively, given as the entrance and exit gates. The entrance gates could

prevent excessive vehicles from entering and overflowing the first cell, while the exit gates could control the accumulated queue in the last cell. A detail of the cell representation and their respective notation can be observed in Figure 1.

Each cell i denoted by the position, $C_i \mathcal{L}_i$ (where $\mathcal{L}$ refers to a specific lane) in every time steps, has its properties characterized as follows:

- (1) $n_i(t)$, the number of vehicles in each cell (also known as the occupancy). Based on the directional movement of vehicles in each cell, n is updated in every time step, after the flows (in and out) of each cell are known.
- (2) N is the maximum number of vehicles that can be stored in the cell. It is expressed as $N_i(t) = k_{\text{jam}} \times L$, where k_{jam} is empirically identified in the actual data that have been processed in the form of discretized cells.
- (3) $Q_i(t)$ is the maximum flow into the downstream cell $i+1$ at time t . It is expressed as $Q_i(t) = st$, where s is the saturation flow rate (veh/h). Considering the change in the saturation flow rate observed empirically from the datasets, a random value of $Q_i(t)$ fixed at a specific range based on the quartiles is applied differently for each cell. For now, it is assumed that the behavior of $Q_i(t)$ is the same for the rest of the time steps.
- (4) $k_i(t)$ is the density found in cell i at time t , and it is expressed as $n_i(t)/L$ (veh/km).
- (5) w/v_f is the backward propagation speed ratio over the free-flow speed found in each cell and is determined as $(1 - (k_i + k_{i+1}/k_{\text{jam}}))$.

3.2. The Improved Cell Transmission Model with Lane Change. Figure 2 illustrates trajectories of vehicles in cells in Lane 1 and 2 in the time windows (T1, T2, and T3). The lateral movement between lanes can be illustrated as follows: considering L1 and L2 as the origin and target lane, respectively, the mean lane change occurred from a cell in lane 1 (L1) to a cell in lane 2 (L2). As shown in Figure 2, the end of a curve in a cell denotes a lane change to a different lane, whereas the beginning of a new curve denotes the arrival of a vehicle from a different lane. Figure 2 demonstrates that no lane change happens in windows $C_1 T_1$ and $C_1 T_2$, and a lane change happens in $C_1 T_3$. A vertical red line is represented as a reference line at the instant when lane-changing starts to occur.

Density at the instant when the vehicle just makes a lane changing is identified from the number of intersecting points crossing the reference line. In Figure 2, the density for $C_1 T_3$ in lane 1 is 5 vehicles, including the LC vehicle, while for lane 2, the density is 2 vehicles. Due to the lane change occurrence, the density of $C_1 T_3$ in lane 1 reduces to 4 vehicles, while for lane 2, it is increased to 3 vehicles. Figure 1 also illustrates the macroscopic schematic diagram of lane change between cells in the multilane road, which also explains why the traditional CTM model cannot incorporate the addition or reduction of inflow traffic due to lane change

FIGURE 1: Cell representation of CTM.

FIGURE 2: Observation of the occurrence of lane change vehicles and the relevant information using vehicle trajectory for the preparation of data towards modeling lane change behavior.

occurrence. This lateral traffic between cells has not been well represented in the traditional CTM model, and its affects have not been well explored and studied previously. Therefore, the extended CTM model improves three main formulations in traditional CTM models to accommodate the change in inflow and cell occupancy.

To incorporate the vehicle's lateral movement or for changing the lane from one lane to another in the existing

CTM model, the improved cell transmission formulation is introduced. The formulation consists of three main components divided as follows: (i) the flow model, (ii) the lane change model, and (iii) the occupancy model. The flow model shows the formulation of how the vehicles in the cell are transferred from one cell to another in a longitudinal manner, the lane change model provided the formulation of vehicle lane change event, and the occupancy model describes how cell occupancy is updated considering the lane change component.

3.2.1. The Flow Model. Three components limit the flow of vehicles from one cell to another, $f_i(t)$: vehicles at the upstream cell $i-1$ waiting to enter i , the capacity of the successor cell i , and the space available in the successor cell i when a queue forms. This is in accordance with the original concept defined by [8, 9], which is described mathematically as follows:

$$f_i(t) = \min \left\{ n_{i-1}(t), Q_i(t), \left(\frac{w}{v_f} \right) [N_i(t) - n_i(t)] \right\}. \quad (3)$$

Here, a slight variation is applied to the formulation for the inflows and outflows of the first and the last cell, respectively. For the inflow at the first cell, the flow is identified based on the minimum between the actual demand, $d_i(t)$, entering the first cell at time step t , and the space available in the successor cell i . This is given as

$$f_{\text{inflow}}(t) = \min \left\{ d_i(t), \left(\frac{w}{v_f} \right) [N_i(t) - n_i(t)] \right\}. \quad (4)$$

For the outflow leaving the last cell, $f_{\text{outflow}}(t)$, the formulation is represented by the minimum between the number of vehicles in the upstream cell, $i - 1$, and the actual outflow, $q_{\text{out}}(t)$, leaving the last cell at time step t since the empty space in the successor cell is unknown. This is given as

$$f_{\text{outflow}}(t) = \min\{n_{i-1}(t), q_{\text{out}}(t)\}. \quad (5)$$

3.2.2. The Lane Change Model. The lane change model was developed macroscopically based on the US-101 NGSIM microscopic trajectory data, which are generalized with refined binary logistic regression [7, 33]. From these data, the main contributing factors of traffic characteristics were aggregated over the varying cell size configurations to identify the parameters for average traffic density, flow, and speed values. A general equation of the basic logistic regression model is given as follows:

$$\text{Logit}(P) = \ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m, \quad (6)$$

$$P = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m)}, \quad (7)$$

where P is the probability of an event occurring; $x_1, x_2, \dots$, and x_m refer to the explanatory variables; $\beta_0, \beta_1, \beta_2, \dots, \beta_m$ are model parameters (or coefficients) that may be determined using the so-called maximum likelihood estimation approach [40]; m is the number of variables to be considered; and $\text{Logit}(P)$ is a logit transformation of $P(x)$ by the natural log of the odds (defined by the ratio of occurrence probability over the nonoccurrence probability).

The so-called maximum likelihood estimation approach can be used to calculate the beta values (i.e., the model's parameters) for fitting the logistic regression model. In this study, the value of P is defined as the estimated probability of lane change, $P(\text{LC})$, estimated from the actual LC status (0 ("LC") or 1 ("NLC")) with their corresponding independent variables observed from the field data. In this case, the independent variables, $x_1, x_2, \dots$, and x_m , used are density difference (Δk) and speed difference (Δv), and lane change direction δ , which were the main contributing factors of traffic characteristics at the macroscopic level.

Δk and Δv are defined as the differences between the origin and the target cell. Both are, respectively, computed using equations (8) and (9). Aggregated cell density at the instant when the vehicle makes lane change is identified from the number of intersecting points crossing the reference line. Figure 2 shows the density of $C_1 T_3$ in lane 1 = 5 vehicles, including the lane change vehicle, while for lane 2 = 2 vehicles. The speed differences of the vehicles from the origin to the target cell were also obtained at the instant when lane change occurred. From Figure 2, the reference line separates the path of the vehicle from the origin to the target cell. The instantaneous speed of the vehicles before and after lane change can be obtained manually from the gradient of the trajectory, which is plotted for displacement, x against time, t . The slope at the left section of the reference

line in the origin lane is referred to as the instantaneous speed of the vehicle before making a lane change, whereas, the slope at the right section of the reference line in the target lane is referred to as the instantaneous speed of the vehicle after lane change. A positive density difference (DP) meant that the origin lane had a higher density over the target lane, while a positive speed difference (SP) meant that the origin lane had a higher speed over the target lane,

$$\Delta k = k_{i,\text{origin}} - k_{i,\text{target}}, \quad (8)$$

$$\Delta v = v_{i,\text{origin}} - v_{i,\text{target}}. \quad (9)$$

A fixed-size CTM-based model was developed and validated with observed data considering the speed, density differences, and lane change direction. Figure 3 summarizes the steps to identify the coefficients of the parameters for the base model using actual datasets.

Given the probability of lane change, formulated by equation (7) obtained from the observed data, the occurrence of lane change is determined from the receiver operating curve (ROC) as a classification measurement [41]. The deciding factor that determines the flow of vehicles from one lane to another, $q_{\text{LC}}(t)$, is decided based on the comparison of the predicted probability with the cutoff value C_v , obtained from the binary logistic regression. Optimum C_v , based on the ability of a model to discriminate the correct prediction between the presence and absence of LC, was used as the deciding factor [41]. Refer to Ng et al. [7] for details about the identification of optimum C_v . The following condition thus provide the decision as per discussed:

$$q_{\text{LC}} = \begin{cases} 1, & \text{if } C_v < P(\text{LC} = 1), \\ 0, & \text{if } C_v > P(\text{LC} = 1). \end{cases} \quad (10)$$

For checking,

$$\Delta \sum_{t, \text{lane}} q_{\text{LC}} \approx \Delta \sum_{t, \text{lane}} P(\text{LC} = 1). \quad (11)$$

Note that the independent variables, speed difference (Δv) and density difference (Δk), are each defined by the difference in the origin over the destination lane. Thus, the formulation of $P(\text{LC})$ in the cell transmission model differs between LC for the left to right and right to left lanes.

Inflows from the on-ramp refer to the rate of lane-change defined by the duration of time per lane change event (sec/LC) taken estimated from the US-101 data. Given that the inflow from the on-ramp was at an average of 140 veh per 15 minutes, a rate of 1 veh per 6 seconds time step is, therefore, assumed in this study.

3.2.3. The Occupancy Model. The occupancy for the number of vehicles in the cell at the following time step, $t + 1$, is updated based on the conservation of traffic flow of cell i [8, 9] in which the number of vehicles in cell i in the next time step, $t + 1$, is equal to the number of vehicles in cell i at the current time step, t plus the number of vehicles that entered, and minus the number of vehicles that left,

FIGURE 3: Steps to identify the coefficients of the parameters for the base model.

$$n_i(t+1) = n_i(t) + f_i(t) - f_{i+1}(t), \quad (12)$$

where i indicates the cell i ; $i+1$ and $i-1$ represent the downstream and upstream of cell i , respectively; $f_i(t)$ is the inflow of vehicles entering cell i during time t , while $f_{i+1}(t)$ is the outflow of vehicles leaving cell i into cell $i+1$ at time t . However, the equation above considers only the longitudinal movement of vehicle flows and ignores the lateral movement between lanes. By considering the lateral movement, the equation is updated as follows:

$$n_i(t+1) = n_i(t) + f_i(t) - f_{i+1}(t) + \sum f_{LC,i,in}(t) - \sum f_{LC,i,out}(t). \quad (13)$$

Figure 4 shows an example of the flow surrounding a single cell. The occupancy of the cells is updated as follows:

$$n_i(t+1) = n_i(t) + f_i(t) - f_{i+1}(t) + f_{i,(\mathcal{L}-1,\mathcal{L})}(t) + f_{i,(\mathcal{L}+1,\mathcal{L})}(t) - f_{i,(\mathcal{L},\mathcal{L}-1)}(t) - f_{i,(\mathcal{L},\mathcal{L}+1)}(t). \quad (14)$$

Note that some movements may not be available depending on the road layout, particularly at the leftmost and rightmost lanes. However, the theory behind adding inflows and minus outflows still holds.

Generally, the first step in developing the cell transmission model starts by filling the initial state's occupancy at the beginning of the time step, t_0 . To implement this method in a real-life scenario, the occupancy of the initial state can be obtained from the actual field data, by counting the vehicles in each discretized cell during the period t_0 . Meanwhile, the flow of vehicles between cells is then computed for t_0 . For the next time step, t_1 , the occupancy of cells is then updated based on the discussion made in the previous section. For the remaining time steps, the formulation of flow and occupancy is then duplicated as many times as desired to compare with the actual traffic flow conditions.

FIGURE 4: Schematic diagram of flows within a cell.

4. Preparation of Field Data for Validation with CTM

This section demonstrates how the raw microscopic data from the field are postprocessed to extract the relevant information needed to be used as the input for the CTM model.

4.1. US-101 NGSIM Data. In this study, a dataset containing a series of individual microscopic trajectories from the well-known and widely accessible US-101 NGSIM database was used to extract the information needed as an input used in developing the cell transmission model. In this study, the dataset collected at the US Highway 101 (Hollywood Freeway) consists of six lanes in a single direction. An on-off ramp connected by a short auxiliary lane will be used to validate the simulated CTM. The lane numbering of the study area can be observed in Figure 5 below.

FIGURE 5: (a) US-101 study area and (b) outputs of processed data.

For validation, the dataset is resorted and discretized into a cell form defined by space over the time length. The cell configuration of the cell size and the start and endpoint of the cells were assured matched between the processed US-101 data with the simulated CTM. Here, the cell size configuration of the study area is fixed at $t=6$ sec, $L=110$ m (eq. 361 ft) [42, 43]. In the original field dataset, $Local_X_f(t)$ defines the position of the lane starting from lane 1 (in ft), while $Local_Y_f(t)$ defines (in ft) the position of cell starting approximately 150 meters before Ventura Blvd. To avoid the noise in the initial and end-stage of processing, the first cell was set at the $Local_Y_f(t)$ coordinates from 60 ft and ended at 2,028 ft. The lane position and cell position given by the specified cell size is redefined as below.

4.2. Extraction of Input Data. To simulate the CTM model, some information from US-101 is used as the input data. These are: (i) number of vehicles in the initial state condition, (ii) demand entering the first cell, (iii) inflow at on-ramp, and (v) outflow leaving the last cell. The raw data consist of vehicle IDs available for every time frame of the study time. With each time step, the number of vehicles entering and leaving each cell can be counted by tracing the vehicle ID in all the time steps. An algorithm using MATLAB has been developed to process the vehicle ID. This algorithm can recognize the difference in vehicle ID found in each cell and at every time step, which can further assist in removing the duplication of vehicle ID between time steps to give information on the demand entering and outflows leaving each cell.

Since the highway was not empty at the initial time of the study period, it was necessary to estimate the initial occupancy of each link by using the same method as above. Otherwise, there would be fewer vehicles in the simulation than under actual traffic conditions. Initial occupancy estimates were obtained by measuring the number of vehicles

on the highway at the start time of the studied period. An equivalent unit of the passenger car unit where $1HV \approx 3.5$ passenger cars, 1 motorcycle ≈ 0.5 passenger cars, is used in simulating the traffic flow [44].

5. Numerical Results

To test the performance of the proposed model against the observed data, this section presented the integration of the results of the lane change model-based logistic regression into the CTM model for various initial demand and initial occupancy scenarios. The first scenario to test was that a scenario of 5 vehicles/time step and 0 vehicles/cell as initial demand and occupancy, respectively, for a cell size fixed at $t=6$ sec, $L=110$ m. The selection of this cell size was first proposed following the logistic regression lane change model by [37], for which the author then suggested to use longer time steps to produce stronger probability. With increasing cell size, [33] validated the presence of multiple lane changes in the same dataset for a cell size greater than the time horizon, $t=6$ sec and cell length, $L=110$ m, which results in the reliability of the model accuracy.

Figure 6 shows the changes in the total cell occupancies against increasing time steps for different lanes. These occupancies were taken based on the occupants of all cells within the same lane compiled for a period of 348 time steps. In Figure 6, it is observed that the model was able to capture the traffic behavior in the process of the formation, propagation, and dissipation of queues within lanes and cells. The increasing trend represents the increased number of vehicles in the traffic lanes where the queues are formed, while the declining trend shows the dissipation of these queues. The steep slope in the first 10–20 seconds represents the initialization period or the warm-up period of the simulation model for the traffic to reach equilibrium.

FIGURE 6: Comparison of the total occupancies over time for different lanes.

5.1. Validation of CTM with US-101. Next, a comparison based on RMSE and analysis of variance (ANOVA) test was made to compare the cases of the CTM with or without the LC models in terms of prediction accuracy according to the actual dataset. Specifically, comparisons based on cell occupancies were made between a few cases of experiments.

Case 1: CTM without LC model

Case 2: CTM with fixed percentage LC

Case 3: CTM with the percentage of LC fixed at 60% (RL) and 35%(LR)

Case 4: CTM with logistic regression

The network topology including lane and cell position used for the above cases is shown in Appendix A1. Boundary cells are applied on each end side, respectively, before and after the first and the last cell of the simulation model.

To avoid a bias from training data when comparing both CTMs, the same dataset fixed by the same demand at the same start and ending times were used. In this study, the discretized form of the real-world data from US-101 in the cell form of the same size is adopted as the ground truth for evaluation. The simulated occupancies for different cases of CTM (Case 1–4) are compared with the observed occupancies from discretized US-101 data to compute the root-mean-square errors (RMSE) as shown in the following equation:

$$\text{RMSE} = \sqrt{\sum \frac{(y_{\text{pred}} - y_{\text{ref}})^2}{N}}, \quad (15)$$

where y_{pred} are the simulated occupancies of cells from the CTM model and y_{ref} are the observed occupancies of the cells. N is the total number of observations.

It should be noted that the integrated binary logistic lane change model has been validated with the NGSIM data itself in the macroscopic cell-based form [7]. The accuracy of the lane change model was accessed through the area under the curve (AUC) of the receiver operating curve as the

performance measures to which its findings show that the lane change model fits the test data reasonably well with an AUC of 0.76.

Case 1: CTM without LC model

In Case 1, only the flow model, which allows vehicles to flow from one cell to another is considered to be given the same input demand, initial state, and outflow for every time step. In this case, all vehicles were assumed to travel within the same lane, and no vehicles were expected to change lanes.

Appendix A2 is provided with the results of Case 1 on the trend of the cell occupancy for all cells and compares the CTM and US-101 data in the 50th, 100th, and 200th steps, respectively. The average occupancies over the 348-time step are also provided, whereby the x -axis represents the different cell positions that are compared between the CTM (based on Case 1) and the US-101 data.

To observe further if there is a difference in the occupancies estimated between the CTM and the US-101, an analysis of variance (ANOVA) was conducted. The ANOVA tests whether or not the mean occupancies are the same as

$$H_o: \mu_{CTM} = \mu_{NGSIM} H_a: \text{not all } \mu_i \text{ are equal}, \quad (16)$$

where μ_i is the mean probability value at any approach i . Suppose a type I error is controlled at $\alpha = 0.05$, then $F(0.95, 1, 65) = 3.99$ with 1 and 65 as the degrees of freedom associated with the factor level and the error term of the given data. The decision rule is, thus,

$$\begin{aligned} \text{if } F^* \leq 3.99, & \text{ conclude } H_o, \\ \text{if } F^* > 3.99, & \text{ conclude } H_a. \end{aligned} \quad (17)$$

Case 2 and Case 3: CTM with fixed LC percentage

Case 2 and 3 assigned a fixed percentage of LC to the CTM model. In Case 2, a fixed probability of 30% is allocated for the lateral movement between lanes from both left to right and right to left lanes [25]. This percentage has been

FIGURE 7: Case 4—comparison of the occupancies between CTM and US-101 at the (a) 50th, (b) 100th, and (c) 200th time step.

FIGURE 8: Case 4—comparison of the average occupancies/total time duration in each cell found between CTM and US-101.

FIGURE 9: Case 4—comparison of the occupancies between CTM and US-101 in the specific cell: (a) C2L3, (b) C3L3, and (c) C4L3.

TABLE 1: Comparison of the statistical test between Cases 1–4.

Case	Average number of occupancy	RMSE overall (average)	ANOVA test	
			<i>F</i>	<i>p</i> value
1	6.08	0.0375 (0.211)	0.136	0.71
2	5.96	0.0396 (0.222)	0.491	0.49
3	5.99	0.0387 (0.216)	1.354	0.25
4	6.25	0.0428 (0.238)	1.025	0.30

FIGURE 10: (a) Layout of I80 Highway; (b) cell arrangement and position of I80-Highway in CTM.

TABLE 2: Parameters for the variables considered in the binary logistic lane change model.

Variables	Coefficient	Standard error	Odds ratio	Significance
Intercept	-3.0839	0.0899	0.0458	***
Density difference, Δk	0.2277	0.0232	1.2556	***
Speed difference, Δv	-0.0909	0.0077	0.9131	***
Direction of lane change, δ	1.5473	0.0983	4.6988	***

***Statistical significance < 0.

FIGURE 11: Comparison of the occupancies between CTM and I80 at the (a) 50th, (b) 75th, and (c) 100th-time step.

FIGURE 12: Comparison of the average occupancies/total time duration in each cell found between CTM and I80.

(a)

(b)

(c)

FIGURE 13: Comparison of the occupancies between CTM and I80 in specific cells: (a) C2L1, (b) C3L2, and (c) C4L3.

TABLE 3: Statistical results comparing CTM with I80.

Ratio	0.971
F (observed value)	0.971
F (critical value)	1.089
DF1	2124
DF2	2124
p value (two-tailed)	0.503
Alpha	0.05

FIGURE 14: Lane and cell position for US-101 highway.

FIGURE 15: Validation of CTM with US-101 for case 1. (a) 50th, (b) 100th, and (c) 200th time step of all cells.

FIGURE 16: Validation of CTM with US-101 for case 2.

given in the past studies to develop the CTM model. The decision to change the lane (either 0 or 1) is predicted based on a random probability assigned to the specific cells.

In Case 3, 60% of the lane change is assigned for vehicles making lane change from the right to the left lanes, while 35% is assigned for LC in the opposite directions. As defined by the right-hand traffic rule, the leftmost lane has the lane with the highest speed, which reduces gradually towards the right. In this case, it is assumed that the study area is located within the ramp, and not studied far off before these areas, a higher proportion of the discretionary lane change (i.e., for speed comfort) is expected for LC from the left to right lanes. The comparison of occupancies between CTM and US-101 for Case 2 and Case 3 is shown in Appendix A3 and A4, respectively.

5.1.1. CTM with Logistic Regression. In Case 4, the same formulation of traffic flows is applied to the CTM model. However, instead of assigning a fixed percentage of the lane change as before, a lane change model that has been generalized with real data is added into the flow model to allow lane change to be predicted based on the difference of speed and density between lanes. Figures 7–9 show the comparison of the cell occupancies in Case 4.

Table 1 summarizes the estimated average occupancy, the RMSE, and the results of ANOVA tests for all cases. It is observed that the average occupancy per cell for all four cases was underestimated from the actual US-101 data, which was estimated at an average of 6.93 veh/cell. When comparing all cases with the US-101, Case 4 exhibited the closest estimation with the actual data by approximately 10.9%. This is followed by Case 1 (14.0%), Case 3 (15.7%), and Case 2 (16.3%). Such a difference is further validated by its small error found in the RMSE. In general, the overall RMSE between the CTM and US-101 was found at around 0.04 (0.2 on average). This indicates a good prediction of these models in accurately predicting the cell occupancies.

From the results conducted using the test of ANOVA, it is observed that the p value in all cases is greater than 0.05, and the $F_{crit} = 3.99$ were found greater than all the F_s estimated for all cases. This shows that the null hypothesis that all means are equal cannot be rejected. Thus, the estimated occupancies estimated from the CTM are consistent with the hypothesis that the population means are equal between groups. In other words, the estimated occupancies do not differ much between them.

To validate the lane change behavior estimated from logistic regression in the CTM model, the total difference in

FIGURE 17: Validation of CTM with US-101 for case 23. (a) C2L3, (b) C3L3, and (c) C4L3.

TABLE 4: Lane and cell position of US-101 highway data.

Lane position (ft) based $Local_X_f(t)$, given width of each lane ≈ 3.5 metre (12 ft)		Cell position (ft) based $Local_Y_f(t)$	
Lane	Coordinate	Cell	Coordinate
1	[0.6, 12.6]	1	[60, 421]
2	[12.6, 24.6]	2	[421, 782]
3	[24.6, 36.6]	3	[782, 1143]
4	[36.6, 48.6]	4	[1143, 1504]
5	[48.6, 60.6]	5	[1504, 1865]
6	[60.6, 72.6]	6	[1865, 2028]

LC is obtained by summing the number of LC in each direction of each lane for the whole simulation run. The estimated number of LC predicted from the US-101 data is then compared with the predicted total number of LC found in the CTM simulation run. In the CTM simulated results, an average difference of 73% is predicted with the left LC higher over the right LC. A similar observation was found in the number of LC estimated from US-101 with a difference of 72% between the two LC directions. In the outcomes of both the proposed CTM model and the field data, the estimated number of LC vehicles for the right lane to the left lane is higher than the other directions.

5.2. Comparison with I80 Highway. In addition to comparing the simulated occupancies estimated between the CTM with the US-101 data, the CTM-based logistic regression is also applied on a different highway. Here, the I80 highway also found in NGSIM data is used. This highway has a similar layout and approximately the same length as that of the US-101 highway, minus the off-ramp plus and the auxiliary lanes, plus an additional main lane. Figures 10(a) and 10(b) show the layout and the CTM topology of this highway, given the same cell size (set at $t = 6$, $L = 110$) as in the previous cases is used.

The computation of the CTM-based logistic regression is carried out in the same way as before, excluding the

input, initial state, and output traffic flow, which is based on the actual conditions of the I80 highway. Table 2 provides the parameters of each variable, and its significances are generalized based on the US-101 dataset, and are applied in this highway to see whether the model can accurately predict the same outcome with the actual I80 highway. Under the variable “direction of lane change”, the lane change between the left lane change and right lane change was taken into consideration to reduce the biases of the lane change prediction between them. The direction of lane change was introduced as a categorical variable, where a value of 0 or 1 refers to the left lane change and right lane change, respectively. With the direction of lane change as a categorical variable, the outcomes of left and right merge can be identified from the probabilities estimated in the model.

Figures 11–13 show a comparison of the estimated occupancies between the simulated CTM and the actual I80 data. The average occupancy simulated from CTM was underestimated from the occupancy of the actual data by approximately 5.2%. The average occupancy estimated from the simulated CTM was found at 7.38 veh/cell, while the actual I80 has an average of 7.76 veh/cell. Comparing both results, the overall RMSE was estimated at 0.074 (0.3 on average), and still reasonably small in error. Furthermore, as shown in Table 3, the ANOVA test based on F statistics 0.97 observed a relationship that there are no significant differences between the occupancies estimated from the simulated CTM and the occupancies observed from the actual I80 data.

For evaluation and validation, a split-sample method was used in which the data were randomly divided into two sections: (i) training data to develop the model and (ii) test data to measure its performance. Since the sample size was large enough, 80% of the samples were used to train the model, while 20% of the samples were split into an independent section for test evaluation of the logistic regression model. The splitting was done once and randomly to ensure that the characteristics were similar to the training dataset. The degree to which the predicted lane change probabilities agree with the actual outcomes gives an overall true correct prediction at 66.41%. Even though the lane change model-based logistic regression has been trained in a different layout, the tested model still managed to estimate the occupancies as close as the actual data of another location. This implies that in order to estimate the traffic state well in various locations, the use of the traffic variables (i.e., speed and density differences) used to identify the prediction of lane change is viable.

6. Conclusion

A good traffic simulation model should be robust and reliable if variations in traffic behavior are adequately taken into account. This paper concerns whether the existing cell transmission model, with the addition of lane change model-based logistic regression, can be further improved to realistically replicate the simulated model with the actual traffic conditions. The use of logistic regression to model the lane change model has been

done in previous studies, but the integration of this model into traffic simulation models has not been done before.

Given the potential of a generalized lane change model that can predict lane change events based on the difference in speed and density between lanes, a new methodology that provide the method to integrate the logistic lane change model into the cell transmission model has been proposed. The logistic lane change model has been modified to the extent that it can produce reasonable estimations of the lane change probability following any change in cell sizes. The modified lane change model, which is generalized, simple, and easy to construct, makes them even easier to be integrated into the formulation of the CTM model.

To demonstrate the usefulness of the proposed methodology, the reliability of the improved CTM model is tested by comparing the outcome generated from the model with the real data. The findings from this study suggested that the improved model can potentially simulate the traffic state as close as the real one. Particularly, the simulated occupancies based on the CTM have reasonably followed the patterns of actual observation with an overall RMSE were found at around 0.03–0.04 (0.2 on average) between the CTM and US-101. This indicates a good prediction of the model in accurately predicting the cell occupancies, as well as proven its reliability on another set of locations with an overall RMSE estimated at 0.074 (0.3 on average). However, results have shown that the simulated CTM, in general, underestimated their occupancies. This could be due to the presence of human error involved in the trajectory dataset when it is discretized into a cell-based form. The discretization of the trajectory data involved traces of vehicle IDs to identify the lane-by-lane input demands and outflows at every simulation time interval. The involvement of large datasets further complicates the identification of vehicle IDs due to the mismatch of missing vehicles or random data errors in the video processing of the trajectory. A similar work by [45], who has also developed CTM for arterial traffic modeling using NGSIM Lankershim data, also mentioned the lack of necessary traffic data. From this simulation, it was realized that some of the inputs to a simple macroscopic model were still inaccessible or directly measurable. Though detailed trajectory NGSIM data have been used to get most of the model inputs, the time period of these data is still short, and the network coverage is small, which does not allow a more comprehensive simulation study that includes more traffic phenomena over a large scope and time period to be carried out.

Moreover, traffic data measured from point detectors (like loop detectors) may not be available in most of the freeways, which causes model inputs such as initial occupancy not available. In this case, the boundary flow of the highway and maximum exiting flows cannot be measured as well. In order to construct a freeway traffic model, detectors are, therefore, needed to provide necessary and accurate measurements.

The model accuracy can be improved if the collection of the demand inflow and outflows is less strenuous to obtain. Thus, fitting the inflows and outflows with a well-fitted distribution is a way to go to the next chapter. This would

have involved developing a different set of models that will be overloaded to discuss within this chapter. Furthermore, other possible outcomes such as lane change rates, delays, queues, and travel time may also be useful to perform the comparison between the simulated and real data. This will be further discussed in the future research.

Appendix

A1. Lane and Cell Position for US-101 Highway

Lane and cell position for US-101 highway is shown in Table 4 and is explained in Figure 14.

A2. Validation of CTM with US-101 for Case 1

Comparison of the occupancies between CTM and US-101 at the: (a) 50th, (b) 100th, and (c) 200th time step of all cells.

Comparison of the average occupancies/total time duration in each cell found between CTM and US-101.

Comparison of the occupancies between CTM and US-101 in the specific cell: (a) C2L3, (b) C3L3, and (c) C4L3 (Figure 15).

A3. Validation of CTM with US-101 for Case 2

Comparison of the occupancies between CTM and US-101 at the: (a) 50th, (b) 100th, and (c) 200th time step.

Comparison of the average occupancies/total time duration in each cell found between CTM and US-101.

Comparison of the occupancies between CTM and US-101 in the specific cell: (a) C2L3, (b) C3L3, and (c) C4L3 (Figure 16).

A4. Validation of CTM with US-101 for Case 3

Comparison of the occupancies between CTM and US-101 at the: (a) 50th, (b) 100th, and (c) 200th time step.

Comparison of the average occupancies/total time duration in each cell found between CTM and US-101.

Comparison of the occupancies between CTM and US-101 in the specific cell: (a) C2L3, (b) C3L3, and (c) C4L3 (Figure 17).

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

CN, SS, and KM conceptualized the study. CN and SS developed the methodology. CN and SS performed formal analysis. CN and SS performed validation. CN, SS, and KM performed investigation. CN wrote the original draft. CN, SS, and KM reviewed and edited the article. CN performed

visualization. SS, KM, and IC supervised the study. SS performed project administration; . SS, KM, and IC were responsible for funding acquisition.

Acknowledgments

This work was supported through NGSIM data provided by the Federal Highway Administration (FHWA) of the US Department of Transportation. The authors wish to thank Mr. Chua Boon Liang for his technical contribution and Mr. Joseph Ng for reviewing this paper. This research was funded by the Ministry of Education Malaysia (MOHE) under the Fundamental Research Grant Scheme (FRGS) [Project code FRGS/1/2019/TK01/MUSM/03/1] and was supported in part by the Japan Society for the Promotion of Science (JSPS) Grant-in-Aids for Scientific Research (C) 20K04531.

References

- [1] Y. Ali, M. C. J. Bliemer, Z. Zheng, and M. M. Haque, "Co-operate or not? exploring drivers' interactions and response times to a lane-changing request in a connected environment," *Transportation Research Part C: Emerging Technologies*, vol. 120, Article ID 102816, 2020.
- [2] A. Ji and D. Levinson, "A review of game theory models of lane changing," *Transportmetrica: Transport Science*, vol. 16, no. 3, pp. 1628–1647, 2020.
- [3] L. Li, R. Jiang, Z. He, X. Chen, and X. Zhou, "Trajectory data-based traffic flow studies: a revisit," *Transportation Research Part C: Emerging Technologies*, vol. 114, pp. 225–240, 2020.
- [4] Q. Chen, R. Gu, H. Huang, J. Lee, X. Zhai, and Y. Li, "Using vehicular trajectory data to explore risky factors and unobserved heterogeneity during lane-changing," *Accident Analysis & Prevention*, vol. 151, Article ID 105871, 2021.
- [5] Y. Ma, Z. Lv, P. Zhang, and C.-Y. Chan, "Impact of lane changing on adjacent vehicles considering multi-vehicle interaction in mixed traffic flow: a velocity estimating model," *Physica A: Statistical Mechanics and Its Applications*, vol. 566, Article ID 125577, 2021.
- [6] C. Dong, H. Wang, Y. Li, X. Shi, D. Ni, and W. Wang, "Application of machine learning algorithms in lane-changing model for intelligent vehicles exiting to off-ramp," *Transportmetrica: Transport Science*, vol. 17, no. 1, pp. 124–150, 2021.
- [7] C. Ng, S. Susilawati, M. A. S. Kamal, and I. M. L. Chew, "Development of a binary logistic lane change model and its validation using empirical freeway data," *Transportation Business: Transport Dynamics*, vol. 8, no. 1, pp. 49–71, 2020.
- [8] C. F. Daganzo, "The cell transmission model, part II: network traffic," *Transportation Research Part B: Methodological*, vol. 29, no. 2, pp. 79–93, 1995.
- [9] C. F. Daganzo, "The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 269–287, 1994.
- [10] M. Lighthill and G. Whitham, "On kinematic waves I. Flood movement in long rivers," *Proceedings of the Royal Society of London - Series A: Mathematical and Physical Sciences*, vol. 229, no. 1178, pp. 281–316, 1955.
- [11] F. Van Wageningen-Kessels and R. E. Wilson, "Lane distribution estimation for heterogeneous traffic flows," *Transportation Research Procedia*, vol. 10, pp. 13–20, 2015.

- [12] C. Mallikarjuna and K. R. Rao, "Heterogeneous traffic flow modelling: a complete methodology," *Transportmetrica*, vol. 7, no. 5, pp. 321–345, 2011.
- [13] G.-L. Chang and Y.-M. Kao, "An empirical investigation of macroscopic lane-changing characteristics on uncongested multilane freeways," *Transportation Research Part A: General*, vol. 25, no. 6, pp. 375–389, 1991.
- [14] J. A. Laval and C. F. Daganzo, *Multi-lane Hybrid Traffic Flow Model: Quantifying the Impacts of Lane-Changing Maneuvers on Traffic Flow*, Institute of Transportation Studies University of California, Berkeley, CA, USA, 2004.
- [15] B. Coifman, S. Krishnamurthy, and X. Wang, "Lane-change maneuvers consuming freeway capacity," in *Traffic and Granular Flow'03*, pp. 3–14, Springer, Berlin, Germany, 2005.
- [16] H. Yeo and A. Skabardonis, "Understanding stop-and-go traffic in view of asymmetric traffic theory," in *Transportation and Traffic Theory 2009: Golden Jubilee*, pp. 99–115, Springer, Berlin, Germany, 2009.
- [17] M. J. Cassidy and J. Rudjanakanoknad, "Increasing the capacity of an isolated merge by metering its on-ramp," *Transportation Research Part B: Methodological*, vol. 39, no. 10, pp. 896–913, 2005.
- [18] S. Ahn and M. J. Cassidy, "Freeway traffic oscillations and vehicle lane-change maneuvers," in *Transportation and Traffic Theory 2007. Papers Selected for Presentation at ISTT-T17 Engineering and Physical Sciences Research Council (Great Britain)*, Springer, Berlin, Germany, 2007.
- [19] R. L. Bertini and M. T. Leal, "Empirical study of traffic features at a freeway lane drop," *Journal of Transportation Engineering*, vol. 131, no. 6, pp. 397–407, 2005.
- [20] Z. Huang, Z. Zhang, H. Li, L. Qin, and J. Rong, "Determining appropriate lane-changing spacing for off-ramp areas of urban expressways," *Sustainability*, vol. 11, no. 7, p. 2087, 2019.
- [21] J. Laval and C. Daganzo, "A hybrid model of traffic flow: impacts of roadway geometry on capacity," in *TRB 2003 Annual Meeting CD-ROM*, Addison Wesley, Boston, MA, USA, 2003.
- [22] J. A. Laval and C. F. Daganzo, "Lane-changing in traffic streams," *Transportation Research Part B: Methodological*, vol. 40, no. 3, pp. 251–264, 2006.
- [23] J. A. Laval and L. Leclercq, "Microscopic modeling of the relaxation phenomenon using a macroscopic lane-changing model," *Transportation Research Part B: Methodological*, vol. 42, no. 6, pp. 511–522, 2008.
- [24] W.-L. Jin, "A multi-commodity Lighthill-Whitham-Richards model of lane-changing traffic flow," *Procedia - Social and Behavioral Sciences*, vol. 80, pp. 658–677, 2013.
- [25] P. I. Richards, "Shock waves on the highway," *Operations Research*, vol. 4, no. 1, pp. 42–51, 1956.
- [26] D. Helbing, "Modeling multi-lane traffic flow with queuing effects," *Physica A: Statistical Mechanics and Its Applications*, vol. 242, no. 1-2, pp. 175–194, 1997.
- [27] S. Hoogendoorn and P. Bovy, "Multiclass macroscopic traffic flow modelling: a multilane generalisation using gas-kinetic theory," in *Proceedings of the 14th International Symposium on Transportation and Traffic Theory*, Transportation Research Institute, Jerusalem, Israel, July 1999.
- [28] V. L. Knoop, A. Duret, C. Buisson, and B. V. Arem, "Lane distribution of traffic near merging zones influence of variable speed limits," in *Proceedings of the 13th International IEEE Conference on Intelligent Transportation Systems*, September 2010.
- [29] V. L. Knoop, S. P. Hoogendoorn, Y. Shiomi, and C. Buisson, "Quantifying the number of lane changes in traffic," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2278, no. 1, pp. 31–41, 2012.
- [30] A. Duret, S. Ahn, and C. Buisson, "Lane flow distribution on a three-lane freeway: general features and the effects of traffic controls," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 157–167, 2012.
- [31] M. Carey, C. Balijepalli, and D. Watling, "Extending the cell transmission model to multiple lanes and lane-changing," *Networks and Spatial Economics*, vol. 15, no. 3, pp. 507–535, 2015.
- [32] X. An, J. Zhao, P. Li, and X. Ma, "Effect of lane allocation on operational efficiency at weaving areas based on a cellular automaton model," *IET Intelligent Transport Systems*, vol. 13, no. 5, pp. 851–859, 2019.
- [33] C. Ng, S. Susilawati, M. A. S. Kamal, and I. C. M. Leng, "Macroscopic lane change model-A flexible event-tree-based approach for the prediction of lane change on freeway traffic," *Smart Cities*, vol. 4, no. 2, pp. 864–880, 2021.
- [34] P. Wang, L. S. Jones, and Q. Yang, "A novel conditional cell transmission model for oversaturated arterials," *Journal of Central South University*, vol. 19, no. 5, pp. 1466–1474, 2012.
- [35] H. M. Imran Kays, T. Hossain Shimu, M. Hadiuzzaman, S. M. Muniruzzaman, and M. Mizanur Rahman, "H-CTM for simulating non-lane-based heterogeneous traffic," *Transportation Letters*, vol. 11, no. 7, pp. 382–390, 2017.
- [36] T. L. Pan, W. H. K. Lam, A. Sumalee, and R. X. Zhong, "Modeling the impacts of mandatory and discretionary lane-changing maneuvers," *Transportation Research Part C: Emerging Technologies*, vol. 68, pp. 403–424, 2016.
- [37] M. Park, K. Jang, J. Lee, and H. Yeo, "Logistic regression model for discretionary lane changing under congested traffic," *Transportmetrica: Transportation Science*, vol. 11, no. 4, pp. 333–344, 2015.
- [38] P. G. Michalopoulos, D. E. Beskos, and Y. Yamauchi, "Multilane traffic flow dynamics: some macroscopic considerations," *Transportation Research Part B: Methodological*, vol. 18, no. 4-5, pp. 377–395, 1984.
- [39] T.-Q. Tang, S. C. Wong, H.-J. Huang, and P. Zhang, "Macroscopic modeling of lane-changing for two-lane traffic flow," *Journal of Advanced Transportation*, vol. 43, no. 3, pp. 245–273, 2009.
- [40] D. W. Hosmer Jr, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, John Wiley & Sons, vol. 398, Hoboken, NJ, USA, , 2013.
- [41] D. Hosmer and S. Lemeshow, "Interpretation of the fitted logistic regression model," *Applied logistic regression*, vol. 2, pp. 47–90, 2000.
- [42] M. N. Sharath and N. R. Velaga, "Enhanced intelligent driver model for two-dimensional motion planning in mixed traffic," *Transportation Research Part C: Emerging Technologies*, vol. 120, Article ID 102780, 2020.
- [43] L. Ye and T. Yamamoto, "Modeling connected and autonomous vehicles in heterogeneous traffic flow," *Physica A: Statistical Mechanics and Its Applications*, vol. 490, pp. 269–277, 2018.
- [44] U. Ahmed, *Passenger Car Equivalent Factors for Level Freeway Segments Operating under Moderate and Congested Conditions*, Marquette University, Milwaukee, WI, USA, 2010.
- [45] D. Su, *Modeling, Estimation and Control of Traffic*, UC Berkeley, Berkeley, WI, USA, 2014.