CONGESTION SCENARIO-BASED VEHICLE CLASSIFICATION DETECTION MODELS BASED ON TRAFFIC FLOW CHARACTERISTICS AND OBSERVED EVENT DATA

Heng Wei[1], Qingyi Ai[1], Hao Liu[1], Zhixia Li[2], and Haizhong Wang[3]

[1] Dept. of Civil & Architectural Engineering & Construction Management, The University of Cincinnati, Cincinnati, Ohio 45221-0071; Tel: 513-556-3781; Fax: 513-556-2599; Email: heng.wei@uc.edu

[2] Traffic Operations and Safety (TOPS) Laboratory, Dept. of Civil & Environmental Engineering, University of Wisconsin-Madison, 1241 Engineering Hall, 1415 Engineering Drive, Madison, WI 53706

[3] School of Civil & Construction Engineering, Oregon State Univ., 101 Kearney Hall, Corvallis, OR 97331

ABSTRACT

While the existing applied length-based vehicle classification model has been to estimate vehicle lengths accurately with dual-loop traffic monitoring station data under free traffic condition, it produces considerable errors against congested traffic. In this study, both ground-truth vehicle trajectory and simultaneous loop event data are used to characterize the impact of congested traffic on vehicle classification. Eight scenarios are synthesized to define the vehicles' stopping locations over two single loops of the dual-loop station. Under the synchronized traffic flow, acceleration or deceleration is considered in the new developed Vehicle Classification under Synchronized Traffic Model (VC-Sync model) to reflect the speed variation between loops. As a result, the error of the vehicle classification is reduced from 33.5% to 6.7%, compared to the existing applied model. Under the stop-and-go traffic condition, a Stop-on-Both-Loops-only (SBL) was developed along with the VC-Sync model to simplify the complexity of congested traffic situation in vehicle length estimation. The error is reduced by using the SBL model from 235% to 17.1%, compared to the existing applied model. Capability of identifying traffic phases is a critical prerequisite to applying the new vehicle classification models under congestions. An innovative method for identifying the traffic phases has been therefore proposed based on the existing traffic stream models along with the new findings of the authors' empirical data analysis. As a result, a heuristic traffic phase identification model has developed and successfully applied in the case study for evaluating the new length-based vehicle classification models with dual-loop data.

INTRODUCTION

This paper presents a scenario-based vehicle classification modeling method to estimate vehicle length via revealing possible scenarios of congested traffic impact on accuracy of vehicle length detection at a dual-loop station. The modeling effort addresses two issues: 1) identifying a sound solution to the problem of distinguishing congestion conditions that could be measured by loop data based on traffic flow characteristics and new findings resulting from analysis of the video-based vehicular event data; and 2) developing scenario-based models for improving vehicle length estimation under congested traffic flows with evaluation of its improved accuracy by comparing the results with the existing applied model.

A dual-loop detector consists of two single loop detectors placed with a fixed distance between single loops (e.g. 20 ft or 6 m), as shown by Figure 1. A vehicle can be detected by the dual-loop detector as electrical pulses of current are deduced in the loops when the vehicle enters and leaves the loop detection area. Each event of the electrical pulse is recorded as a timestamp. Normally, four timestamps, t_1, t_2, t_3, and t_4 are recorded when a vehicle is operating through the loop detector area, as illustrated by Figure 1. This feature enables measuring traffic speed over the detection area, which is one of the key factors in estimating the vehicle length. Vehicle types are then identified in three or four "bins" based on the detected vehicle lengths.
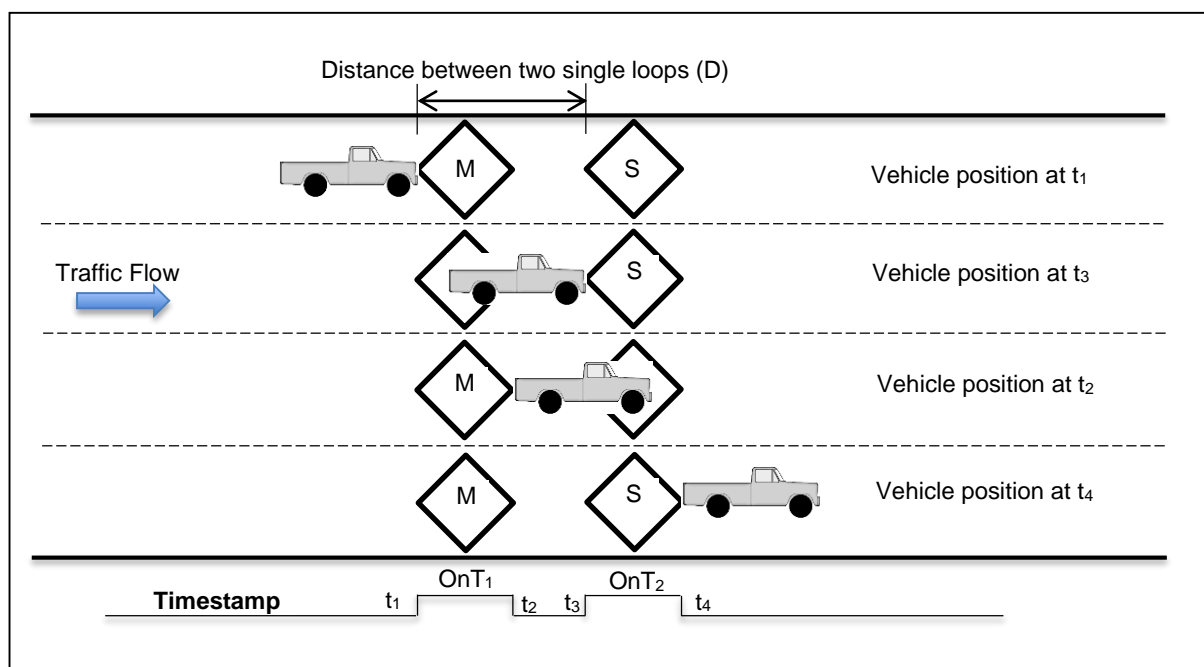
FIGURE 1. Layout of a Dual-loop Detector on Highway

In the existing applied vehicle classification model (which was then proven to be good for free traffic flow), no variation of a vehicle's speed on both single loops is assumed (Nihan et al. 2002). The existing model is described as follows:

$$speed = \frac{D}{t} \tag{1}$$

$$vehicel\_length = speed \times \frac{OnT_1 + OnT_2}{2} - loop\_length \tag{2}$$

Where,

D = distance between two loops (ft);

$t = t_3\text{-}t_1$;

$OnT_1 = t_2\text{-}t_1$;

$OnT_2 = t_4\text{-}t_3$; *and*

t_1, t_2, t_3, and t_4 are timestamps when a vehicle enters or leaves the upstream loop (M loop) or downstream loop (S loop) (Figure 1).

Under congested traffic, however, a vehicle's speed changes frequently and even fiercely as it is traveling through the loops. In order to improve the accuracy of the vehicle length estimation against congested traffic, the authors extracted the ground-truth vehicle event data from video by using the software VEVID (Wei et al, 2005), which was finally complied into high-resolution vehicular trajectory data. Meanwhile, simultaneous event data is derived from the dual-loop data. The sampling dual-loop station is located in the freeway I-71/I-74 in Columbus, Ohio (Ai, 2013). Both datasets were used to define scenarios of vehicles' maneuvers as traversing through the loops and model the traffic conditions based on applied traffic stream characteristics and relevant theories. Finally, new models suitable for congested flows were developed and evaluated with the ground-truth data.

LITERATURE REVIEW

Greenshield (1935) firstly proposed the traffic stream theory addressing the relationships among flow rate, speed, and density, in which speed and density is assumed to be linearly correlated. Greenberg (1959) revised the model of the speed and density to fit a logarithmic curve, based on a hydrodynamic analogy and assumption regarding the traffic flow as a perfect fluid and one-dimensional compressible flow. Underwood (1961) used exponential expression for such a model. The discontinuities of the relationships between traffic variables have been disclosed by researchers. Edie (1961) quantified the linear relationship between density and the logarithm of velocity above the "optimum velocity" for uncongested traffic and velocity and the logarithm of spacing (the inverse of density) for congested traffic. Multiple curves are often applied to depict the "discontinuities". For instance, Koshi (1983) proposed a reverse lambda shape to describe the flow-density relationship. May (1990) developed the "two-regime" models to describe the relationship of flow and density. Hall (1986) proposed an inverted 'V' shape to represent the flow-occupancy relationship. Polus et al. (2002) proposed three regimes of traffic flows (free, dense, and unstable flows), and traffic breakdown was explained as the change from dense flow to unstable flow.

Kerner et al. (1994 and 2010) defined traffic flows in three categories: free flow, synchronized flow, and stop-and-go flow. The free flow has high travel speed and low traffic volume and density. The congested traffic flow is further classified into synchronized flow (S) and wide moving jam (J). The synchronized flow has relative low speed and high volume and density. A wide moving jam is a moving jam that maintains the mean velocity of the downstream front of the jam as the jam propagates. They also disclosed the double Z-characteristic shape for relating speed and density. The empirical double Z-characteristic shape is used to depict the phase transitions between two different phases. F→S (free flow to synchronized flow) and S→J (synchronized flow to jam flow) transitions can be illustrated by a double Z shape (or termed Z-characteristic) for the F→S→J (free to synchronized to jam conditions) transitions. The double Z-characteristic consists of a Z characteristic for an F→S transition and a Z-characteristic for an S→J transition, as well as the phases associated with the critical speeds required for the phase transitions. The synchronized traffic defined by Kerner is also described as the traffic oscillation by other researchers (Bertini and Leal, 2005;, Zielke et al., 2008; Ahn and Cassidy, 2007; Daganzo, 2002; and Mauch and Cassidy, 2002). Treiber M. and Kesting A. (2011) studied the convective instability in congested traffic flow, and they classified congested traffic flow into five classes according to the stability which lead to significantly different sets of traffic patterns (Blandin et al., 2013)

It is necessary to determine what traffic variables and thresholds of the selected traffic variables will be used to describe the traffic phases and identify the transitions between them. Habib-Mattar et al. (2009) found out that the congestion would occur if the situation, where the speed is less than 37 mph and the density is greater than 64 vpmpl, lasts at least five minutes. Chow et al.'s study (2010) indicates that if

the speed drop is greater than 5 mph during a 5-minute period, the traffic flow is at the congestion situation. Lorenz et al. (2001) defined a traffic breakdown as the traffic condition in which the average speed of all lanes on a highway section decreases to below 90 km/h for at least a 15-minute period, and then Elefteriadou et al. (2003) changed the speed threshold as of below 80 km/h. On the other hand, other studies indicated that speed alone is insufficient to ensure the identification of congestion. Congestion may not be detected by the speed-based algorithm only, and "perhaps the optimal speed thresholds are different above a certain occupancy threshold" (Wieczorek et al. 2010). Zhang et al. (2009) used four features to characterize an oscillatory traffic pattern: the occurrence of oscillation, the offset of the oscillation patterns different lanes, the oscillation period, and the oscillation amplitude in flow levels. They set the extreme jam density of 240 vpmpl, flow speed of 50 mph, and wave speed of 10 mph. Deng et al. (2013) proposed a three detector approach to identify traffic states using multiple data sources, including loop detector counts, AVI Bluetooth travel time readings and GPS location samples. However, it is always not easy in practice to obtain the all three sensor data for the traffic flow on a certain highway segment.

Since the event dual-loop data records individual vehicles' timestamps over the loops, it is usually applied in traffic analysis to derive traveling features of the vehicles (Chen et al., 1987; Turner et al., 2000; Coifman, 2004; Nihan et al., 2002 and 2006; and Cheevarunothai et al., 2005). The traffic parameters, such as traffic volume, speed, and occupancy or density can be extracted or calculated from the event dual-loop detector datasets, which further enable calculating vehicle lengths. The existing applied model of estimating vehicle lengths via dual-loop data (Nihan et al., 2006) is based on the assumption that vehicles drive across the dual-loop detection area at a constant speed. The model has been validated well against light traffic. Under light traffic condition, vehicles operate at a relatively high and stable speed, which can be considered at a constant speed. According to Kerner's Three Phases Theory, during uncongested traffic flow, it is reasonable that vehicle speeds are regarded as constant. However, during congested traffic, especially stop-and-go traffic, vehicle speeds become very unstable and are not constant. When the existing model is used to estimate vehicle lengths, the accelerations and decelerations of vehicles will distort the outputs of the model. Accuracy of vehicle classification drops greatly under very congested traffic (Fekpe et al., 2004). It is reported that observed errors in truck misclassification ranged from 30 to 41 percent for off-peak hours, and from 33 to 55 percent for peak hours (Nihan et al. 2006). Li (2009) developed a method of Bayesian inference for vehicle speed and length estimation using dual-loop data. But the congested traffic flow features were not addressed in the method and it was only tested using the traffic flow data with the average speed of 56 mph.

DATA COLLECTION

The selected dual-loop detector station, numbered as V1002, is located in the interstate freeway I-70/71 at West Mound Street downtown Columbus, and has 6 dual-loop detectors in both directions of the highway. A video camera was placed on the top of The Franklin County Juvenile Parking Garage that is close to the station to videotape the traffic flow on I-70/71 over the dual-loop detector station, as shown by Figure 2.
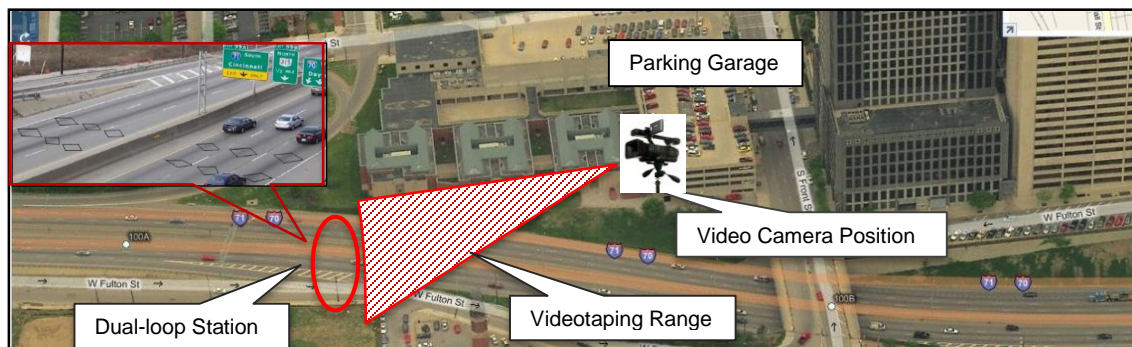
FIGURE 2. Video Data Collection and Loop Station at Study Site

Three-day traffic videotaping was conducted on July 14 - 16, 2009. A total of 26 hour traffic video data were collected, including light traffic and congestion traffic flows. The concurrent event dual-loop data was obtained from the Traffic Management Center (TMC) at the Ohio Department of Transportation (ODOT). The event loop data is the raw data from the dual-loop station, which records the timestamps of each vehicle as it enters and leaves each loop. The scanning frequency of the loop is 60 Hz, that is, occupied status of a loop is automatically updated 60 times per second.

The ground-truth data used in this study is the vehicle trajectory data extracted from the collected traffic video footage. The software VEVID (Wei et al., 2005) was employed to extract the ground-truth vehicle trajectory data from the video.

A QSTARZ^{TM} BT-Q1200 Ultra GPS Travel Recorder was adopted as the data logger to collect GPS data. The GPS travel data logger was equipped in a probe car running roundly along freeway segments of the I-70/I-71 which cover the selected station. The probe vehicle's speed and location information can be collected by the data logger by second. Some parameters which represent characteristics of very congested traffic can be derived from the statistical analysis of the collected GPS data, which includes range of acceleration or deceleration rate and average minimum speed to maintain a vehicle's moving.

DISTINGUISHING TRAFFIC FLOW STATES OR CONDITIONS
Traffic Flow Condition Determined by "Phase Representative Variables"

Flow rate has been conventionally used as one of measurable variables to depict the characteristics of the traffic flow in previous studies; however, application of the flow rate along may be problematic to identifying the traffic conditions (or phases) when the length-based vehicle classification is practiced with dual-loop data. Firstly, any flow rate value may be explained by two or more traffic phases (e.g., uncongested or congested traffic), which may cause a wrong identification of traffic condition. Secondly, the flow rate is an aggregated outcome from the dual-loop based vehicle classification model and supposed to be produced after the traffic phase is identified. That leads to an illogic procedure in practice. Timestamps and occupancies of a vehicle entering and leaving the loops are direct outputs of the loop data. Speed and density can be estimated as a mathematical function of the timestamps and occupancies. According to Kerner's empirical double Z-characteristic shape (as shown in Figure 3), the speed and density are two variables that can be used to determine the boundaries of each traffic flow phase. The speed and density/occupancy are accordingly identified as the "phase representative variables" in this study.
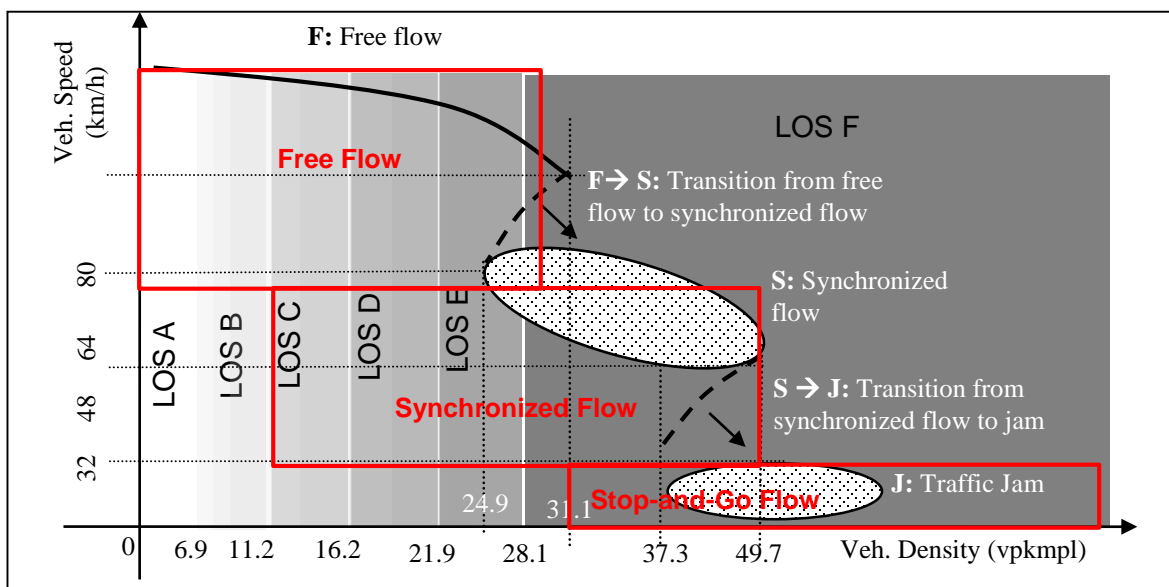

FIGURE 3: Classified traffic flow states (based on Kerner's Z-curve & data in this study)

In Kerner's study (2010) speed and density were applied to depict the empirical double Z-characteristic shape for the phase transitions between two different phases. The original Z-characteristic shape was enhanced and simplified in the study, as illustrated in Figure 3. It conceptually provides a profile of all the possible phases of traffic flows that could be justified by speed and density (or occupancy). Density can be estimated from the loop data by Equation (3) if the average vehicle length of the traffic flow for varying time of a day could be predetermined based on the historical traffic data.

$$K_i = \frac{1000 \times O_{cc}}{L_V + L_{eff}} \tag{3}$$

Where, K_i = density of the traffic flow (vpkmpl) for time period i of a day;
O_{cc} = loop occupancy measurement (%);
L_v = average vehicle length (m); and
L_{eff} = effective detector length (m).

To simplify the procedure of the traffic condition identification, the F→S transition was merged into the free flow phase and S→J transition into the synchronized phase. Equation (4) was proposed to facilitate the development of a computing algorithm that will be used to determine the traffic flow phase $F(t_i)$ of any time period i.

$$F(t_i)=\begin{cases} \textbf{FF}, \text{IF } [u \geq 80 \text{ \& } k \leq 28.1] \text{ OR IF } [\bar{v}(t) - \bar{v}(t+1) \leq \Delta v \text{ \& } var(v) < v^*] \\ \textbf{SF}, \text{IF } [32 \leq u < 80 \text{ \& } 11.2 \leq k \leq 49.7] \text{ OR IF } [(\bar{v}(t) - \bar{v}(t+1) > \Delta v \text{ or } var(v) \geq v^*) \text{ \& } \\ \quad (\overline{occ}(t) - \overline{occ}(t+1) \leq \Delta occ) \text{ \& } \overline{(occ}(t) \leq occ*)] \\ \textbf{TJ}, \text{IF } [0 \leq u < 32 \text{ \& } k \geq 31.1] \text{ OR IF } [(\bar{v}(t) - \bar{v}(t+1) > \Delta v \text{ or } var(v) \geq v^*) \text{ \& } \\ \quad (\overline{occ}(t) - \overline{occ}(t+1) > \Delta occ) \text{ or } \overline{(occ}(t) > occ*)] \\ \textbf{SU}, \text{IF others} \end{cases} \tag{4}$$

Where: k = density, vehicle/km/lane;
u = speed, km/h;
i = time periond i;
FF = Free flow phase;
SF = Synchronized flow phase;
TJ = Traffic jam phase;
SU = special or unreasonable case;
t = a short period of time (5 minutes in this study);
$\bar{v}(t)$ = the average speed in time interval t, km/h;
$\bar{v}(t+1)$ = the average speed in the successive time interval $t+1$, km/h;
$var(v)$ = the variation of all vehicles' speed during time interval t;
Δv = predefined threshold of spot speed difference in successive time intervals, km/h;
v^* = predefined threshold of the speed variation range in successive time intervals, km/h;
$\overline{occ}(t)$ = the average occupancy during time interval t; and $\overline{occ}(t+1)$ = the average occupancy in the successive time interval $t+1$;
Δocc = the predefined occupancy bandwidth during the time interval t; and
occ^* = the maximum average occupancy during the time interval t.

In this study, the percentage of types of vehicles and their average lengths are obtained from the sample dual-loop data at the dual-loop station V1002. The sample size is 13,722. The 3-bin scheme standard adopted by ODOT is used. The sample data indicates that the percentages of small vehicle (length ≤ 8.5 m), medium vehicle (8.5 m <length< 14.0 m), and large vehicle (length ≥ 14.0 m) are 86%, 4%, and 10%, respectively. Their mean lengths are estimated as 5.0 m, 11.1 m, and 22.6 m, respectively. At V1002, L_{eff} is 2.6 m, and then, L_v = 0.86×5.0+0.04×11.1+0.10×22.6 = 7.0 m. The assumed "phase representative variables" are evaluated against the real-world dual-loop data and the VEVID-based vehicular trajectory

data. In light of the statistical analysis performed on the collected ground-truth and loop data, the thresholds of Δv is determined as 16.1 km/h, and v^* is determined as 127.7 km/h^2 (or the standard deviation is 11.3 km/h), Δocc is defined as 0.3, and occ^* is 0.35. To better understand the relationship between each defined traffic phase and the associated level of service (LOS), the LOS is overlaid in Figure 3 with their corresponding density ranges as defined in the Highway Capacity Manual 2010 (TRB, 2010).

MODELING SCENARIOS OF CONGESTED VEHICLE MANEUVERS OVER LOOPS

Under the synchronized traffic, vehicles speeds may change rapidly and frequently. In other words, a vehicle may drive over the upstream and downstream loops at different speeds as it increases or decreases its speed after leaving the upstream loop. Under this circumstance, the vehicle's acceleration or deceleration, which is not considered in the existing applied model, should not be ignored and is assumed to affect measurement of the vehicle length in great part. The characteristics of vehicle movement in the stop-and-go traffic flow are much different from the free or synchronized flow traffic. Vehicles are operating at a high, relatively constant speed under the free flow traffic, and the free flow traffic will transit to the synchronized traffic flow when the traffic speed drops significantly. The synchronized traffic flow will change into stop-and-go traffic when the traffic speed becomes very slow with more frequently acceleration or deceleration involved, and from time to time vehicles have to experience one or more stops. Under the stop-and-go traffic phase, a vehicle may stop within the dual-loop detection area for at least one time. The existing applied vehicle classification model produced more errors under the stop-and-go traffic, especially for large vehicles (See Figure 4), and the sample error even reaches 235%. It is observed from the comparison of the video-based vehicular event data and result from the existing applied model that the vehicle traveling features against stop-and-go traffic, such as acceleration or deceleration, and situation of vehicle stopping on loops, actually affect the estimation of vehicle lengths. An updated length-based vehicle classification model is therefore developed to improve the accuracy of vehicle length estimation under the stop-and-go traffic.
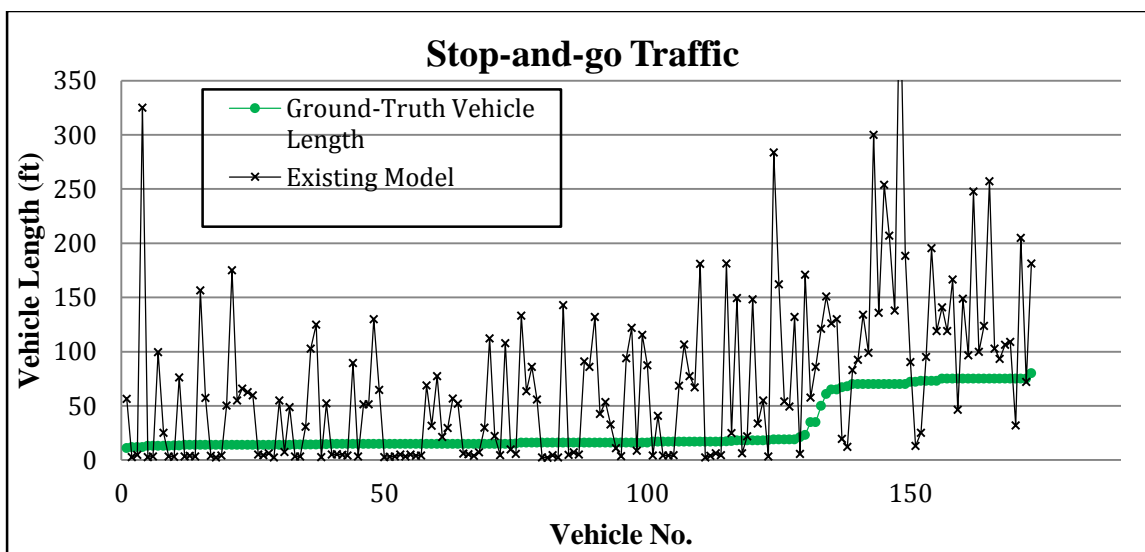

FIGURE 4. Vehicle Length Estimation of the Existing Applied Model under Stop-and-Go Traffic

After careful analysis of synchronizing the ground-truth vehicular trajectory data and the dual-loop data, eight possible scenarios were synthesized based on possible stopping locations of the detected vehicles within the detection area, as illustrated by Figure 5. Those eight scenarios are briefly described as follows.

Scenario 1: the vehicle drives across the dual-loop detection area without a stop, which is a typical synchronized flow feature; Scenario 2: the vehicle stops merely on the M loop and then leaves the dual-

loop detection area without another stop; <u>Scenario 3</u>: the vehicle runs across the M loop and stops only on the S loop; <u>Scenario 4</u>: the vehicle comes into the dual-loop detection area and stops only on both the M and S loops, and leaves the detection area without another stop; <u>Scenario 5</u>: the vehicle stops on the M loop and then move on, and then stops on the S loop and finally leaves the detection area without another stop; <u>Scenario 6</u>: the vehicle stops firstly on the M loop and then stops on both the M and S loops and finally leaves the detection area; <u>Scenario 7</u>: the vehicle stops firstly on both of the M and S loops, and then stops only on S loop; and <u>Scenario 8</u>: the vehicle stops firstly only on the M loop and then stops on both of the M and S loop, and finally stops only on the S loop. Eventually the vehicle leaves the dual-loop detection area without another stop.
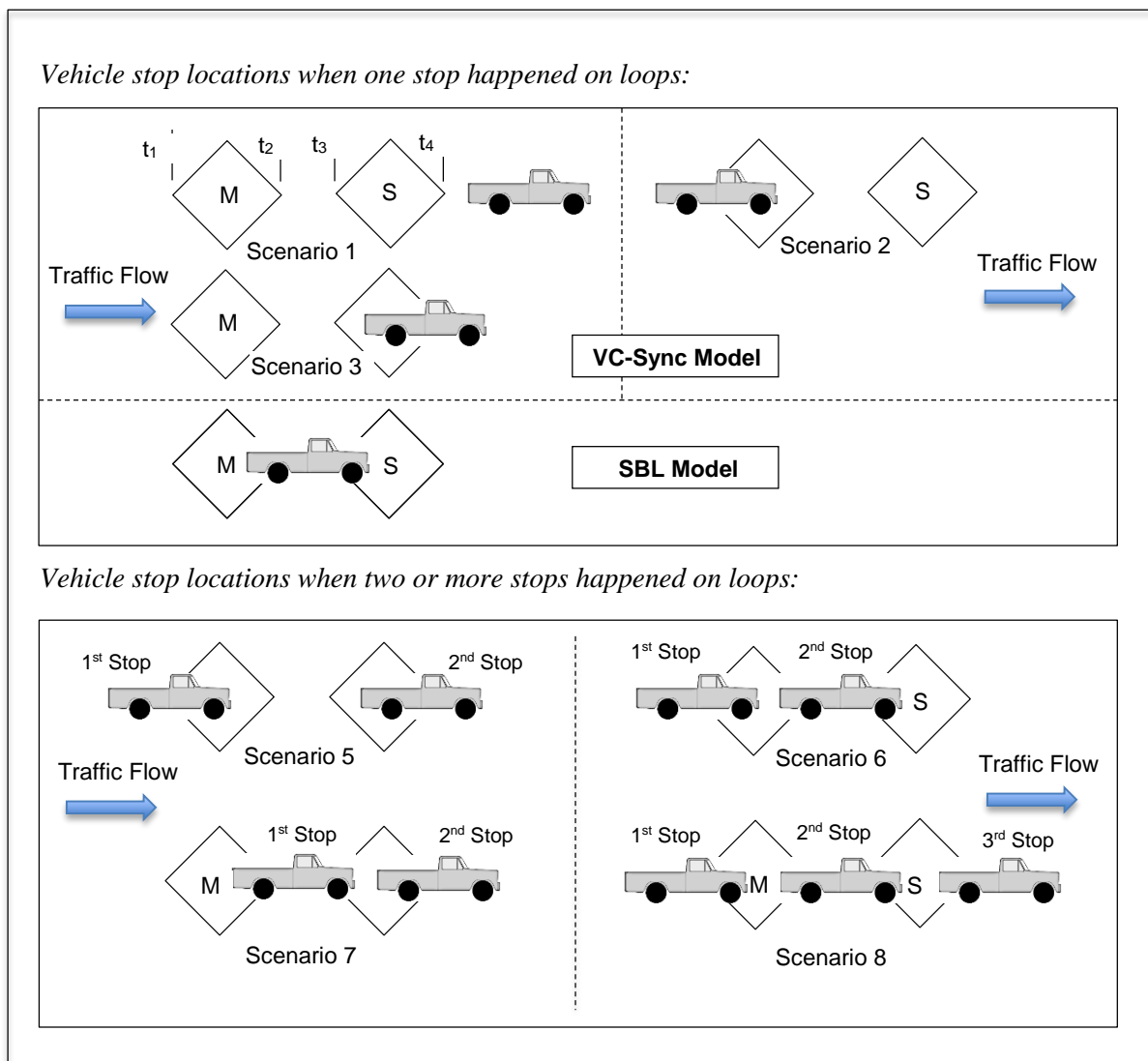
FIGURE 5: Scenarios of Vehicle Stopping on Dual-loops under Congestion

Statistical analysis of the sample data indicates that Scenarios 1 through 4 happened much more frequently than other scenarios (Figure 6 and Table 1). Scenarios 1 through 4 were hence focused in the study, and other scenarios will be considered in the future once sufficient sample data will be gained.
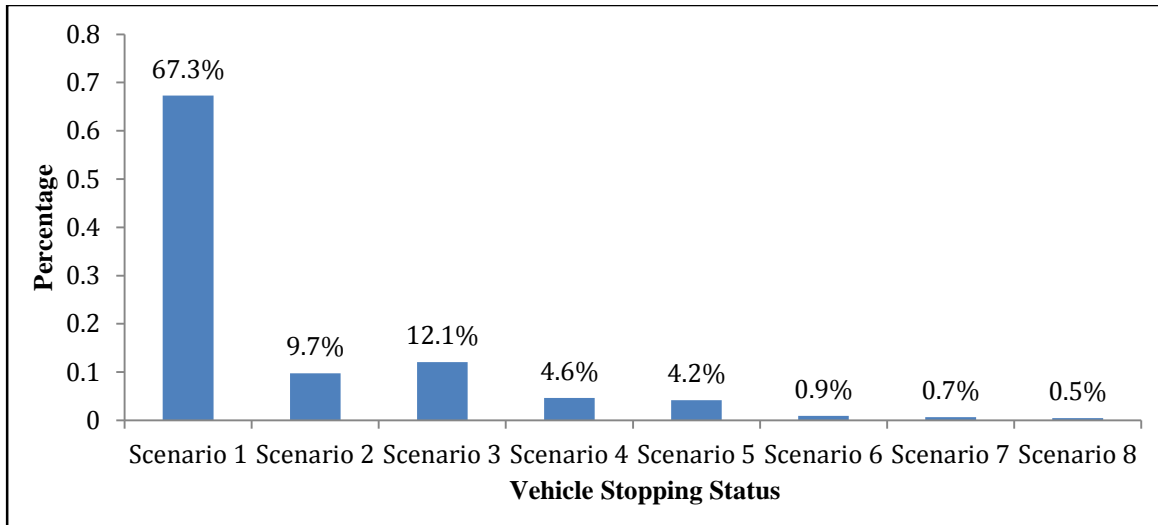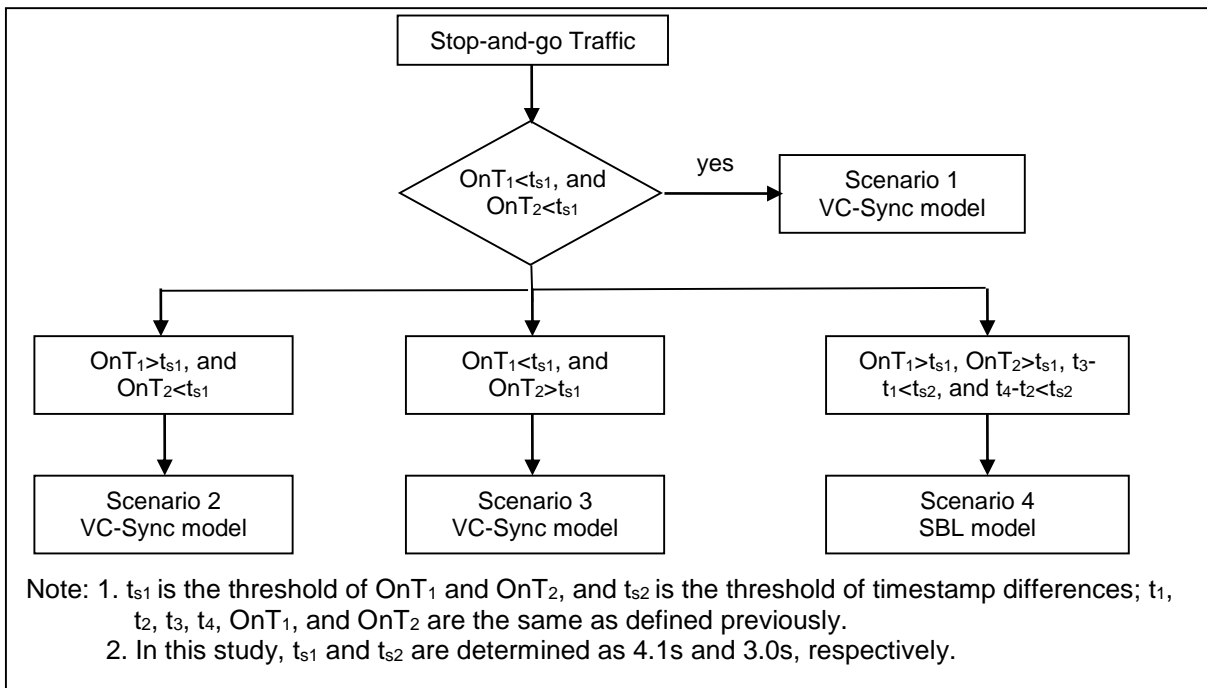
FIGURE 6. Percentage of Vehicle Stopping Status in Congested Traffic

TABLE 1. Vehicle Stopping Status Statistics

Scenario	1	2	3	4	5	6	7	8
Percentage	67.3%	9.7%	12.1%	4.6%	4.2%	0.9%	0.7%	0.5%

Under the stop-and-go traffic flow, a detected vehicle's stopping status can be estimated based on its corresponding dual-loop data, i.e., the time stamps. An algorithm, as illustrated by Figure 7, was developed using On-times and difference of On-times to determine the scenario that the detected vehicle has fallen in. Based on the determined scenario, a suitable vehicle classification model can be applied to estimate the vehicle length.

Note: 1. t_{s1} is the threshold of OnT_1 and OnT_2, and t_{s2} is the threshold of timestamp differences; t_1, t_2, t_3, t_4, OnT_1, and OnT_2 are the same as defined previously.
  2. In this study, t_{s1} and t_{s2} are determined as 4.1s and 3.0s, respectively.

FIGURE 7. Scenario Identification Algorithm

In this algorithm, timestamp t_1, t_2, t_3, t_4, OnT_1, and OnT_2 are adopted as the variables. t_{s1} is defined as the threshold of OnT_1 and OnT_2, and t_{s2} is defined as the threshold of the differences the timestamps. For a vehicle operating under stop-and-go traffic condition:

     (1) If both of OnT_1 and OnT_2 are less than t_{s1}, it indicates that the vehicle did not make a stop within the dual-loop detection area, which means this vehicle falls into Scenario 1.

     (2) If OnT_1 is larger than t_{s1}, and OnT_2 is less than t_{s1}, it indicates that the vehicle spent much longer time on the upstream loop, and this vehicle will be identified into Scenario 2.

     (3) If OnT_1 is less than t_{s1}, and OnT_2 is larger than t_{s1}, it indicates that the vehicle spent much longer time on the downstream loop, and this vehicle will be identified into Scenario 3.

     (4) If both of OnT_1 and OnT_2 are larger than t_{s1}, and t_3-t_1<t_{s2} and t_4-t_2<t_{s2} (t_1, t_2, t_3, and t_4, are the same as defined previously), the vehicle can be identified as falling into Scenario 4.

In this study, in light of the statistical analysis on the dual-loop data under stop-and-go traffic, the thresholds are determined as: $t_{s1} = 4.1$s, and $t_{s2} = 3.0$s. A flow chart of the scenario identification algorithm is illustrated by Figure 7.

LENGTH-BASED VEHICLE CLASSIFICATION MODELS UNDER CONGESTION

Vehicle Classification Model under Synchronized Flow (Scenarios 1 through 3)

Scenario 1 is a typical case of the synchronized traffic. Its flow density is higher than the free flow, and the freedom of maneuvers is greatly restricted. The travel speed is lower than the free flow, and higher than the stop-and-go flow. A new model, **V**ehicle **C**lassification under **S**ynchronized Traffic Model (VC-Sync model), was proposed to estimate vehicle lengths under the synchronized traffic flow. In the model, a vehicle is assumed to pass the detection area at a constant acceleration rate *a* (*a* can be either positive or negative) without a stop. The length of the vehicle passing over the dual-loop detection area can be calculated by the equations as follows:

$$L_v = v_0 \cdot OnT_1 + \frac{1}{2} a (OnT_1)^2 - L_s \tag{5}$$

$$v_0 = \frac{D}{t} - \frac{a \cdot t}{2} \tag{6}$$

$$a = \frac{D}{t} \left[\frac{2 \cdot (OnT_1 - OnT_2)}{(OnT_2)^2 - (OnT_1)^2 + (OnT_1 + OnT_2) \cdot t} \right] \tag{7}$$

Where,

     L_v = length of the detected vehicle (ft);

     L_s = length of each single loop which makes up a dual-loop detector (ft);

     v_o = speed of the vehicle at the moment it is to enter the upstream loop (M loop) (ft/s);

     a = vehicle acceleration (ft/s^2);

     D = distance between two loops (ft);

     $t = t_3$-t_1; OnT_1=t_2-t_1; and OnT_2=t_4-t_3. t_1, t_2, t_3, and t_4 are timestamps when a vehicle enters or leaves the upstream loop (M loop) or downstream loop (S loop) (Figure 1)

Scenarios 2 and 3 can be viewed as special cases of Scenario 1. Scenarios 2 is approximately equivalent to the situation in which a vehicle stops merely at the front edge of the upstream loop and then goes across the detection area without a further stop. This situation can be explained that a vehicle under the synchronized traffic is traversing through the detection area with acceleration and an initial speed of zero. Similarly, Scenario 3 is approximately equivalent to the situation in which a vehicle goes across the detection area without a stop and only stops at the end edge of the downstream loop. This situation can be

interpreted that a vehicle under the synchronized traffic is traversing through the detection area with deceleration and a final speed of zero.

Vehicle Classification Model under Stop-and-Go Flow (Scenario 4)

The stop-and-go traffic has much slower speeds, involving more frequent acceleration or deceleration maneuvers. Under the stop-and-go condition, a vehicle may stop within the detection area for at least once. Based on the ground-truth data, a statistical analysis was conducted to identify the pattern of vehicle stopping locations. As a result, a **S**top-on-**B**oth-**L**oops-only (SBL) model was developed to estimate the vehicle lengths under Scenario 4. For simplicity, it is assumed that the detected vehicle stops right in the middle of the dual loop. After stopping for a period of time t_s, the vehicle restarts to leave the dual-loop detection area at an acceleration rate a. The SBL model is expressed by Equation (8):

$$L_v = f_1 \cdot t_{dec} \cdot D \cdot \frac{1}{t} + \frac{1}{2} f_2 \cdot a \cdot t_{acc}^2 - L_s \qquad (8)$$

Where, $t_{dec} + t_{acc} = OnT_1 - t_s$ and $t_s = t_2 - t_3 - f_3 * t_{acc}^2 / v_{min}$;

    L_v = length of vehicle (ft);
    L_s = length of each single loop (ft);
    t_{dec} = time period as a vehicle enters the M loop until it stops (s);
    t_{acc} = time period as a vehicle starts to move and leaves the M loop (s);
    a = the average acceleration of vehicles as they start to move under stop-and go traffic (ft/s^2);
    t_s = time period for a vehicle to stop on both loops (s);
    v_{min} = average minimum speed remaining without stop (ft/s);
    $f_1, f_2, and f_3$ = adjusting factors for different vehicle types (in this study, $f_1 = f_2 = f_3 = 1$); and
    D, t, t_2, t_3, OnT_1, and OnT_2 = *as the* same as defined previously.

    In order to make the SBL model applicable to estimating vehicle lengths in practice, the vehicle's acceleration rate (a) and average minimum non-stop speed (v_{min}) need to be predetermined. In reality, however, it's extremely difficult to simply derive the acceleration rate of a detected vehicle from its corresponding dual-loop raw data under the stop-and-go condition. The GPS data collected by using GPS data loggers is therefore used to obtain a and v_{min}. Based on the collected GPS data, the variables involved in the SBL model were eventually determined as follows: the average acceleration rate is 2.5 ft/s^2 and the average minimum speed v_{min} is 7 ft/s.

    Finally, the simulated vehicle lengths from the new developed models were compared with the results from the existing model while the ground-truth event data was used as a benchmark. The relative error is reduced from 33.5% of the existing model to 6.7% of the VC-Sync model under Scenarios 1 through 3 (see Figure 8). Under the stop-and-go traffic condition as represented by Scenario 4, the relative error was reduced from 235% of the existing model to 17.1% of the SBL model (Figure 9 and Table 2).
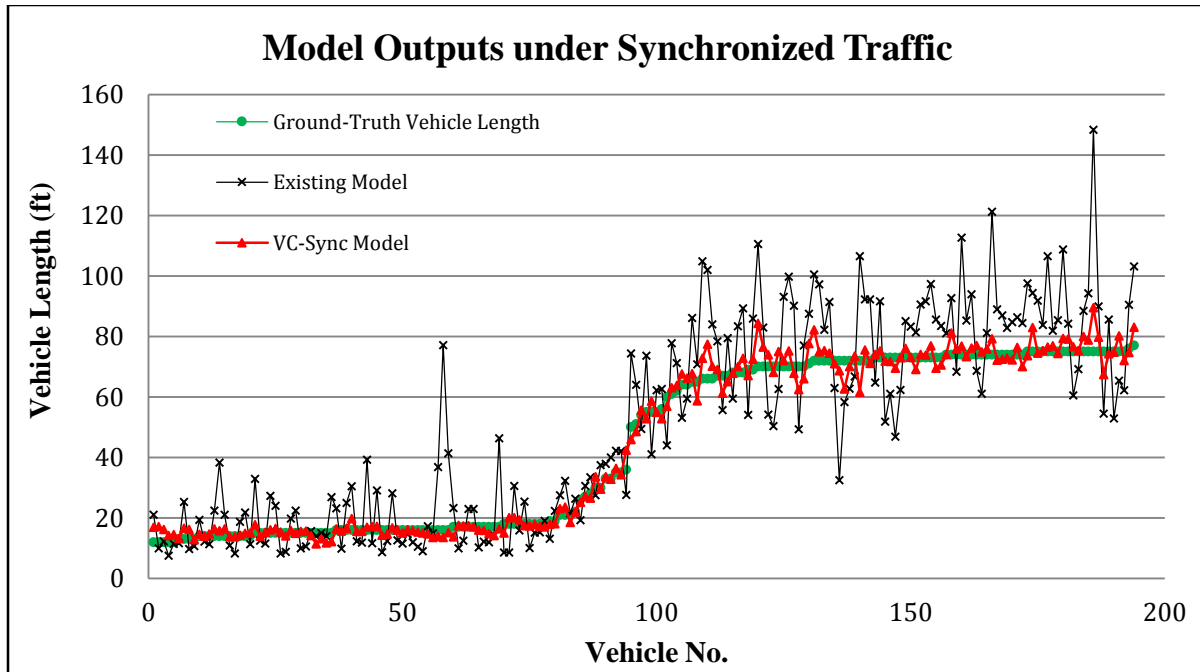
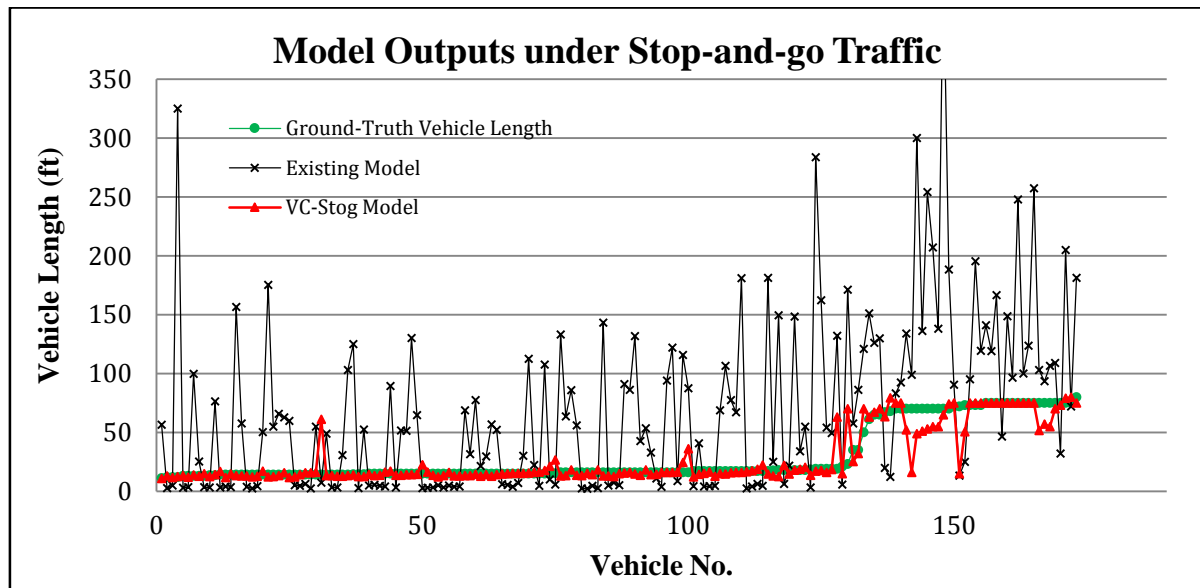**FIGURE 8. Estimated Vehicle Lengths under Synchronized Traffic**

FIGURE 9. Estimated Vehicle Lengths under Stop-and-go Traffic

TABLE 2. Relative Errors Produced by Classification Models

Traffic Flow Condition	Vehicle Classification Model	Error Produced
Synchronized flow	VC-Sync Model	6.7%
	Existing Model	33.5%
Stop-and-Go flow	SBL Model	17.1%
	Existing Model	235%

CONCLUSION

The scenario-based vehicle classification models against both synchronized and stop-and-go traffic flows were developed by fully considering the impact of congested traffic flows. On the basis of watching synchronizing the ground-truth vehicular trajectory data and the dual-loop data, eight possible scenarios were synthesized based on possible stopping locations of the detected vehicles within the detection area. Those eight scenarios reflect the situations of vehicle stopping over loops, which were observed to occur with high possibility in the dual-loop detection area. This synthesized method simplifies the modeling of the vehicles' movements to reveal the impact of traffic on the identification of vehicle lengths at the dual-loop station. Under the synchronized traffic flow, acceleration or deceleration is considered in the VC-Sync model to reflect the speed variation between both loops, which were not conventionally considered in the existing applied models. As a result, the error of the vehicle length estimation is reduced from 33.5% by using the existing model to 6.7% by using the VC-Sync model. Under the stop-and-go traffic condition, the stopping status was synthesized into typical scenarios in the SBL model, which makes it easier to identify the variables involved in the associate vehicle length modeling. As a result, the error is reduced by using the SBL model from 235% to 17.1%, compared with the existing applied model.

Capability of identifying traffic phases is a critical support to applying the length-based vehicle classification models. This paper presents an innovative method for identifying the traffic phases that was developed based on integrated analysis of the existing traffic stream models and the new findings from the authors' empirical data analysis and modeling efforts. As a result, a heuristic traffic phase identification model has developed and successfully applied in the case study for evaluating the new length-based vehicle classification models with dual-loop data.

ACKNOWLEDGEMENT

REFERENCES

Ahn, S. and Cassidy, M. J. (2007). "Freeway traffic oscillation and vehicle lane-change maneuvers." *Transportation and Traffic Theory 2007*, Elsevier, pp. 691-710.

Ai, Q. (2013). Length-Based Vehicle Classification Using Dual-loop Data under Congested Traffic Conditions. Ph.D. Dissertation. The University of Cincinnati. December 2013.

Athol, P. (1965). "Interdependence of Certain Operational Characteristics within a Moving Traffic Stream." Highway Research Record 72, pp. 58-87.

Bertini, B. A. and Leal, M. T. (2005). "Empirical study of traffic features at a freeway lane drop." Journal of Transportation Engineering 131 (6), pp. 397-407.

Cheevarunothai, P., and Wang, Y. (2006). "Identification and Correction of Dual-Loop Sensitivity Problems." Compendium of Papers CD-ROM, 85th Transportation Research Board Annual Meeting, Washington, DC, January 2006.

Cheevarunothai, P., Wang, Y., and Nihan, L. N. (2005). "Development of Advanced Loop Event Analyzer (ALEDA) for Investigations of Dual-Loop Detector Malfunctions", Presented at The 12 world Congress on Intelligent Transportation Systems, San Francisco.

Chen, L., and May, A. (1987). "Traffic detector errors and diagnostics", In Transportation Research Record: Journal of the Transportation Research Board, No. 1132, TRB, National Research Council, Washington, D.C., pp. 82–93.

Coifman, B. (2004). *Research Reports: An Assessment of Loop Detector and RTMS Performance*. Report No.: UCB-ITS-PRR-2004-30. ISSN 1055-1425.

Daganzo, C. F. (2002). "A behavioral theory of multi-lane traffic flow. Part I: Long Homogeneous freeway sections. Transportation Research Part B 36 (2), pp. 131-158.

Deng, W., Lei, H., and Zhou X. (2013). "Traffic state estimation and uncertainty quantification based on heterogeneous data sources: A three detector approach." Transportation Research Part B 57(2013), pp. 132-157.

Edie, L.C. (1961). "Car following and steady-state theory for non-congested traffic." *Operations Research* 9, pp. 66-76.

Elefteriadou, L., and Lertworawanich, P. (2003). "Defining, Measuring and Estimating Freeway Capacity." Compendium of Papers CD-ROM, 82nd Transportation Research Board Annual Meeting, Washington, DC, January 2003.

Fekpe, E., Gopalakrishna, D., and Middleton, D. (2004). *Highway Performance Monitoring System Traffic Data for High-Volume Routes: Best Practices and Guidelines.* Final Report. Office of Highway Policy Information Federal Highway Administration, U.S. Department of Transportation, Washington, D.C., 2004.

Greenberg, H. (1959). "An Analysis of Traffic Flow." Operation Research, Vol. 7, p. 78-85.

Greenshields, B. D. (1935). "A Study in Highway Capacity." HRB Proc., Vol. 14, p. 468.

Habib-Mattar, C., Polus, A., and Cohen, M. A. (2009). "Analysis of the Breakdown Process on Congested Freeways." Compendium of Papers CD-ROM, 88th Transportation Research Board Annual Meeting, Washington, DC, January 2009.

Hall, F.L. and Gunter, M.A. (1986). "Further analysis of the flow-concentration relationship." *Transportation Research Record: Journal of the Transportation Research Board, No. 1091*, TRB, National Research Council, Washington, D.C., pp. 1-9.

Kerner, B. S. (1998). "Experimental Features of Self-Organization in Traffic Flow." *Physical Review Letters*, Volume 81, Number 17, October 1998.

Kerner, B. S. (2004). *The Physics of Traffic: Empirical Freeway Pattern Features, Engineering Applications, and Theory*. Springer, 2004.

Kerner, B. S., and Klenov, S. L. (2010). "Explanation of Complex Dynamics of Congested Traffic in NGSIM-Data with Three-Phase Traffic Theory." Compendium of Papers CD-ROM, 89th Transportation Research Board Annual Meeting, Washington, DC, January 2010.

Kerner, B. S., and Konhäuser, P. (1994). "Structure and Parameters of Clusters in Traffic Flow." *Physical Review E*, Volume 50, Number 1, July 1994.

Koshi, M., Iwasaki, M. and Okhura, I. (1983). "Some findings and an overview on vehicular flow characteristics." *Proceedings*, 8*th International Symposium on Transportation and Traffic Flow Theory* (Edited by Hurdle, V.F., Hauer, E. and Steuart, G.F.) University of Toronto Press, Toronto, Canada, pp. 403-426.

Li, B. (2010). "Bayesian inference for vehicle speed and vehicle length using dual-loop detector data." Transportation Research Part B 44 (2010), pp. 108-119.

Lorenz, R. M., and Elefteriadou, L. (2001). "Defining Freeway Capacity as a Function of the Breakdown Probability." Compendium of Papers CD-ROM, 80th Transportation Research Board Annual Meeting, Washington, DC, January 2001.

Mauch, M., and Cassidy, M. J. (2002). "Freeway traffic oscillations: observations and predictions." Transportation and Traffic Theory in the 21st Century: Proceedings of the Symposium of Traffic and Transportation Theory. Elsevier, pp. 653-672.

May, A.D. (1990). *Traffic Flow Fundamentals*. Prentice-Hall. Englewood Cliffs, New Jersey. *Data: Quick Detection of Malfunctioning Loops and Calculation of Required Adjustments.* Research Report, TNW2006.

Nihan, N. L., Zhang, X., and Wang, Y. (2002). *Evaluation of Dual-loop Data Accuracy Using Video Ground Truth Data.* Research Report, TNW2002.

Polus, A. and Pollatschek, M. (2002). "Stochastic Nature of Freeway Capacity and Its Estimation." *Canadian Journal of Civil Engineering*, 29, 2002, pp. 842-852.

Transportation Research Board (TRB) (2010). *Highway Capacity Manual 2010*. Fifth Edition.

Treiber, M. and Kesting, A. (2001). "Evidence of convective instability in congested traffic flow: A systematic empirical and theoretical investigation." Transportation Research Part B 45(2011), pp. 1362-1377.

Turner, S., Albert, L., Gajewski, B., and Eisele, W. (2000). "Archived Intelligent Transportation System Data Quality: Preliminary Analysis of San Antonio Transguide Data", *Transportation Research Record: Journal of the Transportation Research Board, No. 1719*, TRB, National Research Council, Washington, D.C., pp. 77-84.

Underwood, R. T. (1961). "Speed Volume, and Density Relationships: Quality and Theory of Traffic Flow." Yale Bureau of Highway Traffic, p. 141-188.

Wei, H. (2008). Report on Characterize Dynamic Dilemma Zone and Minimize Its Effect at Signalized Intersections. Research report for Ohio Transportation Consortium.

Wei, H., Meyer, E., Lee, J. and Feng, C.E. (2005). "Video-Capture-Based Approach to Extract Multiple Vehicular Trajectory Data for Traffic Modeling," *ASCE Journal of Transportation Engineering* Volume 131, No. 7 (2005), pp. 496-505.

Wieczorek, J., Fernández-Moctezuma, R. J., and Bertini, R. L. (2010). "Techniques for Validating an Automatic Bottleneck Detection Tool Using Archived Freeway Sensor Data." Compendium of Papers CD-ROM, 89th Transportation Research Board Annual Meeting, Washington, DC, January 2010.

Zhang, H. M., and Shen, W. (2009). "A Numerical Investigation of the Stop-and-go Traffic Patterns Upstream of A Freeway Lane-drop." Compendium of Papers CD-ROM, 88th Transportation Research Board Annual Meeting, Washington, DC, January 2009.

Zielke, B., Bertini, R., and Treiber, M. (2008). "Empirical measurement of freeway oscillation characteristics: an international comparison." Transportation Research Record 2088, pp. 57-67.