# UNIVERSITY OF CINCINNATI

**Date:** 18-Oct-2010

**I,** Sudhir Reddy Itekyala **,**

**hereby submit this original work as part of the requirements for the degree of:**

Master of Science

**in** Civil Engineering

**It is entitled:**

Vehicle Classification under Congestion using Dual Loop data

**Student Signature:** Sudhir Reddy Itekyala

**This work and its defense approved by:**

**Committee Chair:** Heng Wei, PhD

*Heng Wei, PhD*

# Vehicle Classification under Congestion using Dual Loop data

A thesis submitted to the

Graduate School

of the University of Cincinnati

in partial fulfillment of the

requirements for the degree of

MASTER OF SCIENCE

In the School of Advance Structures

College of Engineering and Applied Science

2010

By

**Sudhir Reddy Itekyala**

B.TECH Civil Engineering, Jawaharlal Nehru Technological University, India, 2007

Committee

**Dr. Heng Wei, Chair**

**Dr. Anant Kukreti**

**Dr. Changjoo Kim**

# ABSTRACT

The growing congestion problem on Interstates has been identified as a serious problem for accurate data collection from automatic sensors like Inductive loop detectors (ILD). Traffic speed and vehicle classification data are typically collected by dual-loop detectors on freeways. During congestion, measurement of vehicle lengths which is based on detector ON and OFF timestamps (raw loop event data) often lead to misclassification of vehicle data. Accurate detection of raw event data and modified classification algorithm are increasingly important for higher data accuracy needs for agencies such as Advanced Traffic Management Systems (ATMS) and Advanced Traffic Information Systems (ATIS). Vehicle classification algorithm works on the assumption of constant vehicle speed in the detection area. This assumption is violated during congestion which induces errors in to vehicle length estimates leading to more inaccurate vehicle classification data.

This paper unlike in preceding works presents a model which is simple enough to be implemented using existing loop detector hardware. This new model assumes vehicle travels with constant acceleration over loop detection area and thus named as “Constant Acceleration based Vehicle Classification model (CAVC)”. This model first identifies traffic flow state and later uses Kinematic equations for estimating vehicle length values. Data is collected by videotaping dual loop station and also simultaneously collecting raw loop event data. Ground truth vehicle data is then extracted using Vehicle Video-Capture Data Collector (VEVID) [Wei et al. 2005] from video data. This improved model (CAVC model) is then validated using ground truth classification data and also compared with the results from existing vehicle classification model for different traffic flow states (under specific scenarios).

# ACKNOWLEDGMENTS

First I would like to thank my advisor Dr. Heng Wei for being my source of inspiration. He has helped me grow both professionally and personally; supported me in times of need; motivated me to do my best always.

I must also thank Dr. Anant Kukreti and Dr. Changjoo Kim for being in the committee and for their valuable feedback on my proposal which has stimulated my thoughts and improved the quality of work. I also would like to thank Qingyi Ai, Ph.D. student of Transportation engineering at University of Cincinnati (UC) for his assistance in modeling, data collection and extraction for OTC project.

I would also like to mention my special thanks to Dr. Benn Coifman from Ohio state university (OSU), Nick Hegemier and Matthew Graf from ODOT District 6, I.T.S for their help in the raw loop data collection. And thanks to Carolyn Bethel, Assistant Director, Public Facilities Management (Franklin County parking garage) and to Col. Africentric Ec Elementary School management who gave us permission to use their parking lots for setting up the camera for data collection.

I would like to thank my parents who are the source of my faith in life and my pillars of support. They have always encouraged me to explore and expand my interests. Everything I achieve, I achieve for them. Last, but not the least, I could not have completed this work without the unending support of my lab partners Zhixia Li**,** Vijay Krishna Nemalapuri, Viswanath Pokala, Zhuo Yao and also my friends Raja, Bharath, Himaja, Deepika, Aparna, Harika, Balu, Kethan, Amar and Vijay Kumar. They cheered me up at my lowest, and cheered me on through the highest points of my life.

*To my loving parents, Suvarna and Hanmanth Reddy and to my sister Divya*

# TABLE OF CONTENTS

## LIST OF FIGURES

# LIST OF TABLES

# CHAPTER 1: INTRODUCTION

Vehicle classification data on freeways is of considerable use to agencies involved in almost all aspects of transportation planning and engineering. This data along with other traffic flow data such as volume, occupancy, density, etc provides valuable information for evaluating existing roadways and also in planning and designing new infrastructure. Automatic traffic sensors are the main data providers for maintaining historical data base of a wide roadway networks. These are used to provide data continuously for longer period of time.

Inductive Loop Detector (ILD) is the major type of intrusive vehicle sensor for collecting traffic data. Dual-loop detectors which are formed by two consecutive single-loop detectors spaced several feet apart are usually used to collect length based vehicle classification data. The governing principle behind existing model for estimation vehicle lengths is that the vehicle observes constant speed over detection area. During congestion vehicles violate this basic assumption of constant speed leading to misclassification of vehicle data. Both synchronized and stop-and-go flow can be summarized as two forms of congestion [Neubert et al. 1999]. And thus the existing vehicle classification model needs to be modified to identify traffic flow states and should be able to estimate accurate vehicle lengths resulting in more accurate vehicle length based classification during congestion.

## 1.1 Goals and objectives

The goal of this research is to improve the vehicle classification model under congestion. This requires study and identification of vehicle motion under congested flow and its impact on the vehicle length calculation. Parameters which effect vehicle classification include speed (v), on-time (OnT) and acceleration (deceleration) (a). Use of Kinematic equations to modify the existing vehicle classification model to accurately estimate vehicle length is proposed.

Modified vehicle classification model is expected to identify the traffic flow state and also acceleration factors and use them for vehicle length calculation during both synchronized flow and also non-stopping scenario during stop-and-go flow. This new model will be evaluated using video data for corresponding raw loop event data which is collected simultaneously. Image processing software "Vehicle Video-Capture Data Collector" (VEVID) will be used to process video images to generate vehicle trajectory data and other parameters [Wei et al. 2005]. VEVID is developed by Dr. Heng Wei and further improved by his Ph.D. Student, Mr. Zhixia Li under his guidance.

## 1.2 Background

Vehicles are classified in different ways like length of vehicle, number of axles and number of units (including power and trailer units). Type of classification depends on the type of sensors used further depending on its limitations. The kind of sensor installed determines the raw data obtained from it which is used for obtaining traffic data such as vehicle classification, volume, speed, occupancy, etc. As explained earlier this research considers vehicle classification using dual loop detectors which work on the principle of change in inductance for detecting vehicle's presence.

As explained earlier existing vehicle classification model assumes that vehicle travels with constant speed over the "loop detection area". This is valid and common during free flow period where vehicles travels without any significant acceleration or deceleration patterns. Detector on-time (occupied time), speed pattern, and other traffic features such as headway and occupancy vary greatly during congestion. Irregular acceleration (deceleration) and stopping of vehicles during congestion cause errors in vehicle speed estimates and thus effect classification. Traffic flow is classified in to three states they are “free flow”, “synchronized flow” and “stop-and-go flow”. Hence following sections present detailed background of different vehicle detection systems and especially the principle and working of inductive loop detectors, description of existing vehicle classification model and traffic flow theory.

### 1.2.1 Vehicle detection systems

“Vehicle detection systems” or “Traffic detectors” form major data collectors (source) for maintaining historical database of traffic data. National Electrical Manufacturers Association (NEMA) defines a vehicle detection system as "A system for indicating the presence or passage of vehicles". Traffic detectors provide traffic flow data for freeway traffic management, data collection and traffic-responsive signal control, etc [Klein et al. 2006]. Type of traffic sensor used for a particular roadway section depends on many factors such as type of data needed, roadway characteristics and also on available funds.

Traffic detectors can be broadly classified into:

1) Non-intrusive detectors

2) Intrusive detectors

#### *1.2.1.1 Non-intrusive detectors*

Non-intrusive detectors are placed over pavement and installed without much interruption to the ongoing traffic. Examples of non-intrusive sensors are passive and active infrared devices, cameras associated with video image processors, ultrasound detectors, microwave radar and acoustic arrays. Following sections give an overview of some major non-intrusive traffic detectors.

**1) Video image processor (VIP)**

Video cameras are used for traffic management and surveillance based on their ability to transmit television images to a human operators. But in recent days the video image processing is done to automatically extract information required for traffic surveillance, management and data collection. A video image processor (VIP) system typically consists of one or more camera units connected to a microprocessor, which are then used to digitize and analyze the images (Traffic Detector Handbook) [Klein et al. 2006]. The software thus used to convert the imagery into traffic flow data.

**Figure 1: Video camera located on the pole for data collection**
(Source: http://www.mctraffic.org/technology.htm, accessed on August 26th, 2010)

Figure 1 here shows a camera installed on the freeway. A VIP can be used to replace in-ground inductive loops and thus providing detection of vehicles across multiple lanes. VIP's have high initial costs for installation but have an advantage of lower maintenance costs. VIPs can classify vehicles by their length and also detect vehicle presence, volume, lane occupancy, and speed for each class and lane.

**2) Microwave radar sensor**

Radar sensor is a device which is used for transmitting electromagnetic signals and receiving echoes from objects or targets within its coverage area [Booth et al. 1993]. Radar stands for **RA**dio **D**etection **A**nd **R**anging. Microwave sensors for traffic data collection usually operate at frequency intervals near 10.5, 24.0, and 34.0 GHz, these standards are to be satisfied on U.S. roadways as per Federal Communications Commission (FCC) regulations.

Figure 2 shows a typical Microwave radar operation. Microwave sensors which transmit a Continuous Wave (CW) in Doppler waveform can only detect vehicle passage. They can provide vehicle count and speed data but cannot detect stopped vehicles. Whereas microwave sensors that transmit a Frequency Modulated Continuous Wave (FMCW) can detect stopped vehicles and thus provide measurements of vehicle count, speed, lane occupancy, and vehicle length based classification.

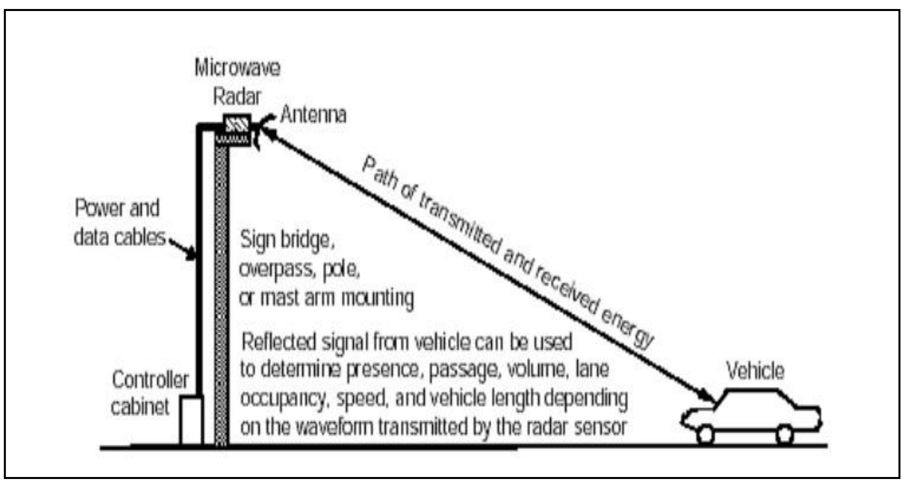


**Figure 2: Microwave radar operation**
(Source: Traffic Detector Handbook: Third Edition-Volume I, Klein et al. 2006)

**3) Infrared sensor**

Infrared sensors can be classified in to two types 1) Active infrared sensors (or Laser radar) and 2) Passive infrared sensors. Active infrared sensors illuminate detection zones with infrared energy which is transmitted by laser diodes and operate in infrared region of the electromagnetic spectrum at 0.85 mm [Klein et al. 2006]. Passive sensors transmit no such energy instead detect energy from sources like vehicles, road surfaces, and other objects. Figure 3 shows an image of vehicle with trailer detected by active infrared sensor

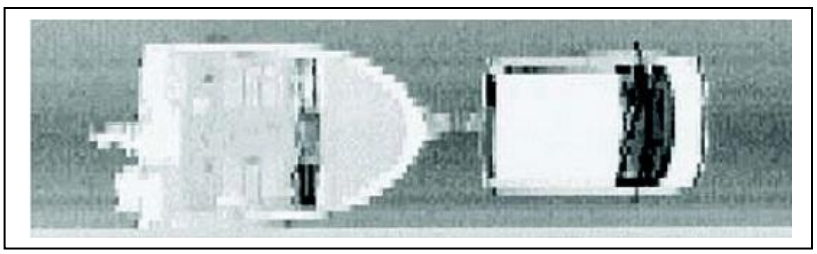

**Figure 3: Laser sensor image**
(Source: Traffic Detector Handbook: Third Edition-Volume I, Klein et al. 2006; Photograph courtesy of Schwartz Electro-Optics, now OSI Laserscan, Orlando, FL).

***1.2.1.2 Intrusive sensors***

Intrusive sensors are installed below pavement surface, examples of Intrusive sensors are Inductive loop detector, Weigh in Motion (WIM), Magnetic sensor, etc. They are installed in the roadway surface and buried below the pavement and covered with pavement materials.

**1) Inductive-loop detector:**

An inductive-loop detector senses the presence of a metal object by inducing currents in the object (further explained in section 1.2.3.1). This further reduces the loop inductance and is used to detect the presence of the vehicle. They typically consist of four parts a wire loop embedded in pavement (one or more turns), a lead-in wire connecting the loop to a pull box, a lead-in cable connecting the wire to the controller, and finally an electronic unit in the controller cabinet. They are installed by cutting a slot in the pavement and placing one or more turns of wire. The wires are later covered with sealant. Loop detectors are installed in different shapes and sizes depending on data needs.

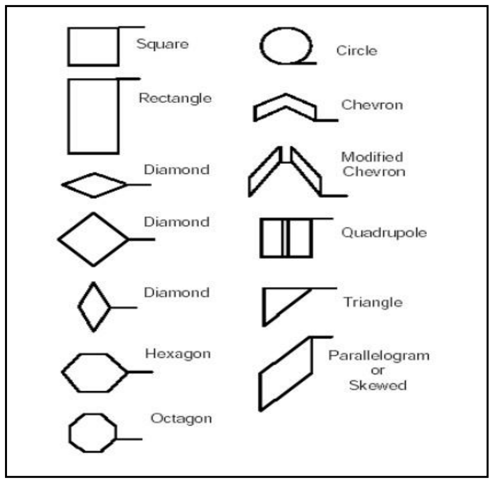


**Figure 4: Different loop detector shapes**

(Source: Traffic Detector Handbook, Chapter 4, Third Edition—Volume I, Klein et al. 2006)

Figure 4 shows different shapes commonly used for loop detector installation [Klein et al. 2006].Many loop configurations were designed to detect various sizes and shapes of vehicles, ranging from bicycles, motorcycles to trailer trucks and avoiding detection of vehicles in adjacent lanes.

**2) Weigh-in-motion (WIM)**

Weigh-in-motion (WIM) devices are designed to record truck axle weights and gross vehicle weights as the vehicle drives over sensor. WIM systems do not require the vehicle to stop which makes them much more efficient. The application of these sensors is more for trucks which can be weighed as they travel at highway speeds. But WIM can be only used to classify vehicles based on the overall weight and its main application is for following purposes (Training guide, http:// training.ce.washington.edu):

- Pavement and bridge design, monitoring, and research
- Size and weight enforcement
- Legislation and regulation
- Administration and planning

**3) Magnetic sensor**

Magnetic sensors are passive devices which detect the presence of a vehicle (ferrous metal) through the magnetic anomaly which they cause in the Earth's magnetic field [Klein et al. 2006]. Figure 5 shows dipoles on a vehicle and their effect on readings of magnetic compass and these determine the sensor output.

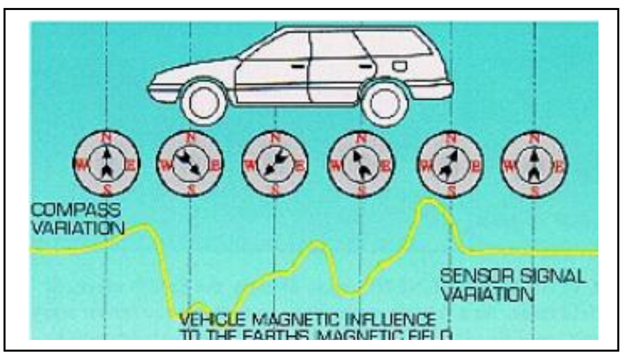


**Figure 5: Perturbation of Earth's magnetic field by ferrous metal vehicle**
(Source: Traffic Detector Handbook, Chapter 4, Third Edition—Volume I, Klein et al. 2006)

And finally the Table 1 compares different sensor technologies (traffic detectors) based on data output, communications bandwidth, and costs. The cost comparison is based on the prices in 1999 in dollars ($) and covers almost all major sensors technologies used in the field. Check marks indicate the sensors ability to collect the type of data mentioned.

**Table 1: Comparison of vehicle detection systems (traffic detectors)**

| Sensor Technologies | Count | Presence | Speed | Output State | Classification | Multiple lane, multiple detection zone data | Communication band width | Sensor purchase cost[a] (each in 1999) |
|---|---|---|---|---|---|---|---|---|
| Inductive loop | ✓ | ✓ | ✓ b | ✓ | ✓ c | | Low to moderate | Low[i] ($500-$800) |
| Magnetometer (two axis fluxgate) | ✓ | ✓ | ✓ b | ✓ | | | Low | Moderate[i] ($900-$6,300) |
| Magnetic Induction coil | ✓ | ✓ d | ✓ b | ✓ | | | Low | Low to moderate[i] ($385-$2000) |
| Microwave radar | ✓ | ✓ e | ✓ | ✓ e | ✓ e | ✓ e | Moderate | Low to moderate ($700-$2000) |
| Active infrared | ✓ | ✓ | ✓ f | ✓ | ✓ | ✓ | Low to moderate | Moderate to high ($3300-$6,500) |
| Passive infrared | ✓ | ✓ | ✓ f | ✓ | | | Low to moderate | Low to moderate ($700-$1200) |
| Ultrasonic | ✓ | ✓ | | ✓ | | | Low | Low to moderate (Pulse model: $600-$1900) |
| Acoustic array | ✓ | ✓ | ✓ | ✓ | | ✓ g | Low to moderate | Moderate ($3100-$8100) |
| Video image processor | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | Low to high[h] | Moderate to high ($5000-$26,000) |

(Source: Traffic Detector Handbook, Chapter 4, Third Edition—Volume I, Klein et al. 2006)

Where,

**a-** Costs related to installation, repair and maintenance not included.
**b-** Speed can be measured using two sensors at fixed distance (speed traps).
**c-** Uses special electronic units to classify vehicles.
**d-** With the help of special sensor layouts and signal processing software.
**e-** Only applicable with microwave radar sensors that are capable of signal processing and transmitting the proper waveform.
**f-** Available with infrared sensors capable of multi-detection zone feature.
**g-** With models that contain appropriate beam forming and signal processing.
**h-** Depends on type of data transmitted to the TMC (higher-bandwidth/ lower-bandwidth raw data or video image data)
**i-** The price includes intrusive sensor unit and receiver electronics unit.

### 1.2.2 Vehicle classification

Information on truck and freight movements is important considering the role of freight mobility on the economy. The classification data is also used by highway engineers for the geometric and structural design of roadways and bridges. Common uses of vehicle classification information include:

- Pavement design and management
- Scheduling the resurfacing, reconditioning, and reconstruction of highways
- Design inputs relative to the current and predicted capacity of highways
- Development of weight enforcement strategies
- Environmental impact analysis, including air quality studies
- Analysis of alternative highway regulatory and investment policies.

Many valuable statistics can be extracted from the vehicle classification data using the historical database. These include vehicle distance traveled (VDT) information, and also considering the sensitivity of pavement to truck volume and also truck VDT has important role to play in pavement design. Other common uses of VDT statistics by vehicle class include air quality emission monitoring, crash statistics by type of vehicle and general transportation trend monitoring. Dual loop sensors classify vehicles into more general classifications less than FHWA's 13 classes. This is because of following reasons:

- Length classifiers already in use are not accurate to measure small differences in vehicle lengths [Bonsall et al. 1987]. Hence broad vehicle length categories are used which reduce the total amount of error and in turn lead to more accurate classification.
- Also Inductive loop detectors cannot differentiate between multiple smaller units joined together and a single long vehicle.
- Length based classification also cannot precisely identify vehicle classes like tractor, trucks, five-axle, semi-trailer for their variety of lengths in use.

But still length based vehicle classification provides accurate enough data for transportation planning and engineering requirements [Kell et al. 1990]. Vehicle classification bins are most common approach for classifying vehicles depending upon vehicle lengths [Nihan et al. 2002]. Three or four classes are still sufficient for many planning and analytical purposes. Practically as the number of vehicle classes increase the chances of placing the vehicle in to wrong bin increases giving rise to need of more accurate vehicle detection systems. Loop detectors cannot detect the number of axles of the vehicle; this disadvantage leads to classifying vehicles only using length. With respect to engineering, planning and design applications this classification data in fewer bins (three or four) is sufficient.The data presented in this research is in the form of bins.

Vehicle classification into three bins (ODOT):

1. Small: Personal vehicles and smaller vehicles (length 28ft or less)
2. Medium: Small trucks and buses (28ft to 46ft)
3. Large: Larger trucks and buses (length greater than 46ft)

Vehicle classification into four bins (WSDOT):

1. Bin 1: Personal vehicles and smaller vehicles (length 26ft or less)
2. Bin 2: Small trucks and buses (26ft to 39ft)
3. Bin 3: Larger trucks and buses (39ft to 65ft)
4. Bin 4: Largest trucks and articulated buses (length greater than 65ft)

### 1.2.3 Loop detectors

Loop detectors are most common and widely used detectors on interstates for collecting traffic data. As explained earlier in Figure 4 (section 1.2.1.2) loop detectors are of different sizes and shapes depending on the type of data desired, lane widths, etc. The following sections present you the principle behind dual loop detection and also the existing length based vehicle classification model.

#### *1.2.3.1 Loop detection principle*

Operation of loop detectors involves passing of an alternating current through a loop which generates a fluctuating magnetic field around its wires. This extends above and below the road surface and when a metal object enters this field it produces eddy currents. Vehicle's own magnetic field couples with the loop field and reduce the loop's inductance which is measured by the loop detector circuitry. Amount of inductance change depends on the strength of the magnetic field cutting the vehicle [Bonsall et al. 1987]. When such inductance change is observed the loop is assumed to be occupied by controller and produces a value of "1" (occupied) otherwise "0" (unoccupied).

Typically loop detection area is classified based on the area covered in to two categories, 1) large-area detection 2) small-area detection. Large-area detection normally contains a detection zone covering an area of at least 20ft (6m) or more in a traffic lane [Klein et al. 2001]. They are primarily used for detecting the presence of a vehicle as long as the area is occupied. Whereas small-area detection is commonly implemented with a single short inductive loop.

Loops detectors used for traffic data collection can be classified as below:

1. Single loop detectors
2. Dual loop detectors

Single loop detectors are mainly installed for traffic flow counts. They provide data of individual vehicle actuations which are in the form of count and occupancy. Vehicle length estimation using single loop is not accurate and many assumptions are made for its calculation. Thus use of dual loops for vehicle classification is widely adopted.

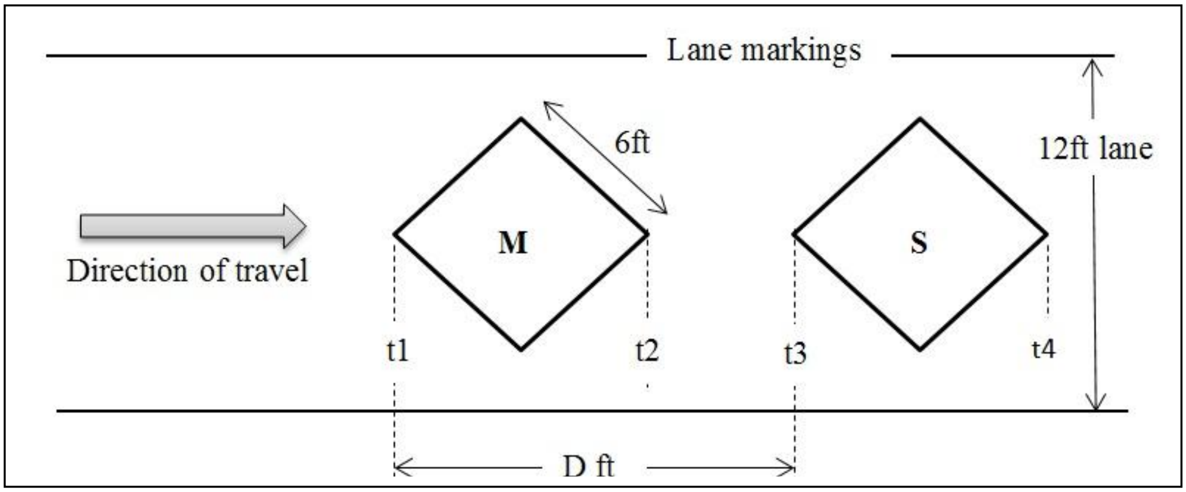


**Figure 6: Layout of typical dual loop detector**

As explained earlier a dual-loop detector consists of two consecutive single loops that are spaced few feet apart (usually 20ft). In Figure 6 we can see the layout of typical dual loop detector with corresponding loop dimensions. The raw data (time stamps $t_1$, $t_2$, $t_3$ and $t_4$) from loops is used as input to vehicle classification model to calculate vehicle speed, length and occupancy. Dual loops are more accurate in vehicle length calculations since each vehicle is observed twice on dual loops. Loop detectors as explained earlier can detect only the presence of vehicle (metal body) and thus vehicles with trailer in some cases are detected as two vehicles.

#### *1.2.3.2 Loop detector errors*

Loop detector data is erroneous due to several reasons such as malfunction of hardware, insufficient computing power of the cabinet controller, crosstalk between sensors, sensitivity, pulse breakups, omitted phase etc. Many researchers have tried to address these problems using algorithms to filter out the erroneous data. It is observed that loop detectors produce either reasonable accurate data or erroneous data all the time rather than fluctuating [Chen and May, 1987]. Several methodologies using occupancy measurements for intervals of 20 second to 5 minutes were proposed to identify detector errors. Selected samples are then tested for quality depending on threshold limits of flow (or volume, q), density (k), and speed (v).

Sensitivity of loops is a major type of detection errors. Basic operating principle behind the inductive loop (explained in section 1.2.3.1) is the interaction between the magnetic field produced by the loop and the conducting surface which in this case is the vehicle. According to Coifman, B. (2001) loop detector data does not provide sufficient information necessary to separate length due to sensitivity from the vehicle's actual length. According to Cheevarunothai et al. 2006 loop sensitivity problems can be divided into two categories:

1) Sensitivity discrepancies between the upstream loop (M) and downstream loop (S).
2) Unsuitable sensitivity levels of both the M and S loops.

Vehicle length distribution is used to find the appropriate sensitivity levels. A statistical approach was applied using “Short Vehicle” (SV) length distribution as observed by Wang and Nihan. 2004. They observed SV which corresponds to Bin-1 vehicle class follow a normal distribution with a mean of 15.21ft (4.64m) and a standard deviation of 2.20ft

(0.67m). And this length information for SVs was used to trace a correct sensitivity level for the loop. Thus in this study, they [Cheevarunothai et al. 2006] used the SV-length distribution reported by Wang and Nihan et al. 2004 as the ground-truth vehicle length distribution for SVs. Crosstalk which is another type of detection error induces considerable amount of misrepresentation of actual data because it produces false detection of vehicles. The common source of cross talk is two sensors set to similar frequencies. Crosstalk occurs when one detector activates another detector in an adjacent lane or happens when two loop's lead-in cables share a common conduit or when two loops are installed within a few feet of each other or when there are poor quality of splices and couplings [Bhagat and Woods, 1997; Kell et al. 1990].

Researchers like Bhagat and Woods have examined physical characteristics that cause cross talk. These studies thus provide some advice for correcting cross talk and also thus localizing the unit with cross-talk using a specialized loop tester that bypasses the controller and loop sensor. Coifman (1999) on the other hand presented a simple application of the error detection for detecting cross talk between sensors. He proposed that each loop of a dual loop station is set to a different operating frequency and a small amount of cross talk which affects the detection is detected through a differencing method. Thus loop detectors can be corrected for these detection errors during installation and also checked frequently for ensuring accurate data collection.

### 1.2.4 Existing vehicle classification model

Raw loop actuation data is the loop event data and is accurate to 1/60th of a second. Sixty data points for every second are collected. This data when processed using vehicle classification model results in vehicle length calculation. As shown in Table 2 ‘M’ represents up stream loop, ‘S’ represents the downstream loop. Each data point (vehicle) consists of status of the loop whether occupied or unoccupied represented by ‘1’ or ‘0’ respectively. The time stamp is represented by converting time period in to 1/60$^{th}$ of second units.

Figure 7 shows a time-space representation of loop detector output along with pulses. The highlighted area represents the path of vehicle over the loop detection area with reference to time on X axis. Vehicle as shown in figure records four different time stamps using both front and rear bumpers. Each loop as explained in earlier section are separated by a small distance (20ft) and together form dual loop detector (or speed-trap).

**Table 2: Raw loop event data**

| Vehicle No | Loop status | Upstream (M) | Downstream (S) |
|---|---|---|---|
| 1 | 1 | 3111959 | 3111973 |
| | 0 | 3111975 | 3111988 |
| 2 | 1 | 3112344 | 3112359 |
| | 0 | 3112380 | 3112394 |
| 3 | 1 | 3112447 | 3112461 |
| | 0 | 3112529 | 3112478 |
| 4 | 1 | 3112765 | 3112780 |
| | 0 | 3112830 | 3112845 |
| 5 | 1 | 3112909 | 3112925 |
| | 0 | 3112930 | 3112945 |

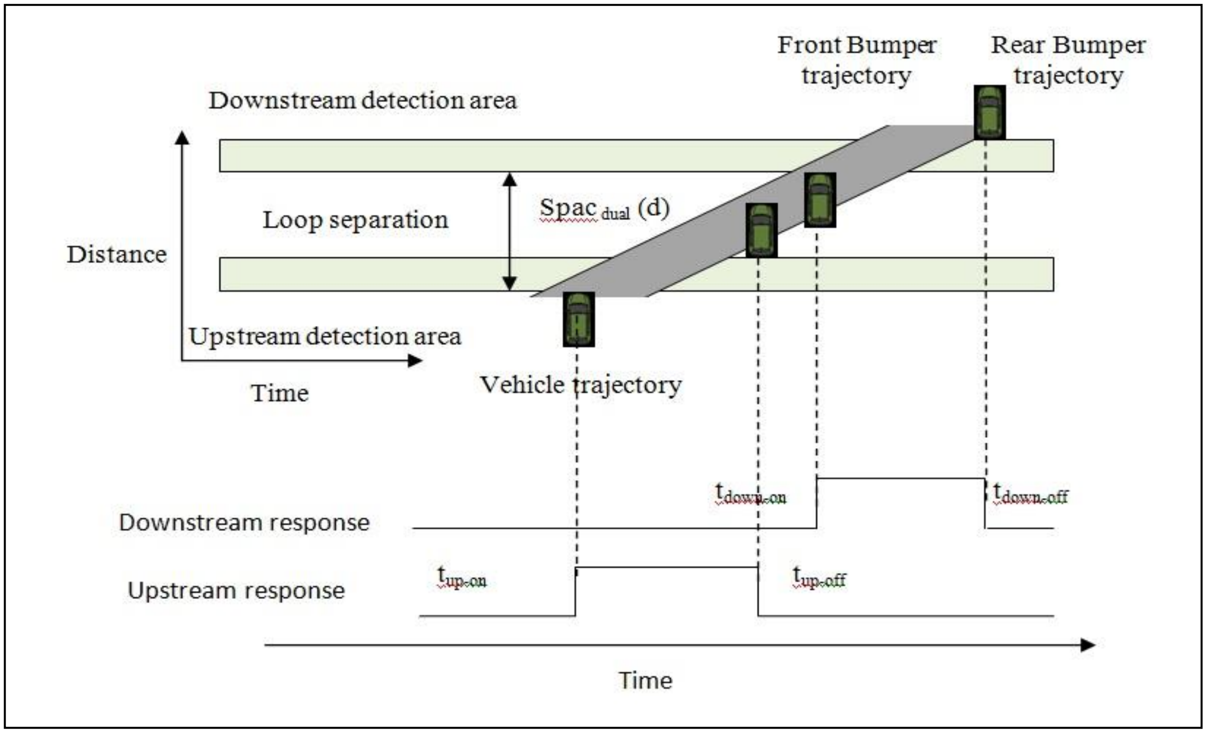


**Figure 7: Time-space representation of loop detector pulses**
(Source: Using Dual Loop Speed Traps to Identify Detector Errors by Benjamin Coifman, 1999)

Where,

$t_{up\text{-}on}$= $t_1$ = on-time stamp at upstream Loop (M loop);

$t_{up\text{-}off}$ = $t_2$ = off-time stamp at upstream Loop (M loop);

$t_{down\text{-}on}$ = $t_3$ = on-time stamp at downstream Loop (S loop);

$t_{down\text{-}off}$ = $t_4$ = off-time stamp at downstream Loop (S loop);

Spac $_{dual}$ = d=loop spacing;

Loop length = length of loop measured along the roadway;

Once the detection of vehicle over dual loop detection area produces four different time stamps the existing vehicle classification as shown in Figure 8 is used to calculate vehicle length using speed and on-time. Speed is calculated by dividing the distance between the loops (Spac $_{dual}$) by the difference between the On-time stamps at downstream loop ($t_{down\text{-}on}$) and at the upstream loop ($t_{up\text{-}on}$).

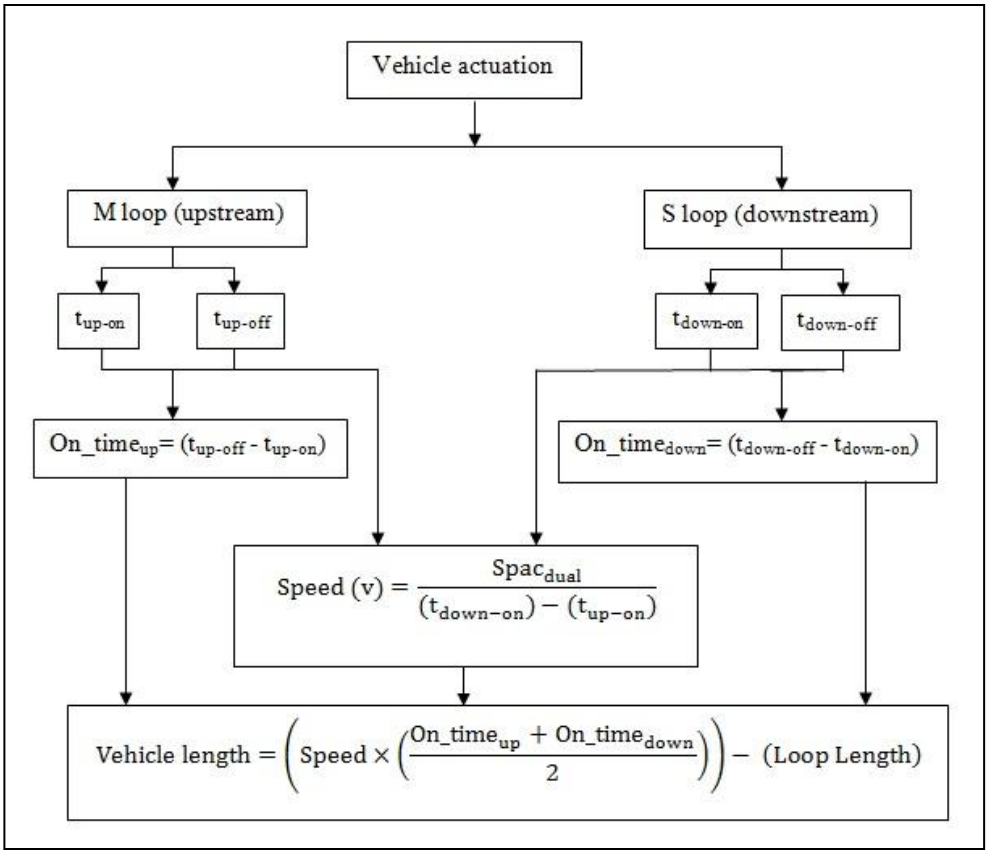


**Figure 8: Existing length based vehicle classification model**

Where,

Speed (v) = Speed of vehicle on the loop (normally constant during free flow);

Loop length (l) = Length of loop measured along the roadway;

A similar calculation for speed using rear bumper can also be done using off-time stamps at both M (upstream) and S (downstream) loops. Thus using dual loop produces much more accurate velocity measurements than the results from single loop detectors. Detection Zone is not a point phenomenon hence a vehicle which travels over the loop is "detected" for slightly longer period. This is calibrated such that the loops provide accurate data for detection of vehicles.

### 1.2.5 Traffic flow theory

In past few decades many researchers and scientists have developed a wide range of different mathematical models for traffic flow aiming to explain the complex nature of traffic. These models are based on the behavior of drivers and expected to show phenomena observed in real traffic. Traffic flow models are classified depending on the non-linear interaction and dynamics of vehicles. For example submicroscopic models take into account details such as perception thresholds, changing gears, acceleration characteristics of specific vehicle type, reaction to brake light, etc.

On the other hand Gas-kinetic model formulate a partial differential equation for the temporal evolution of vehicles density and velocity distribution. It is common to distinguish two classes of macroscopic models they are first order models such as the Lighthill-Whitham-Richard (LWR) model [Lighthill and Whitham, 1955; Richards, 1956]. These are based on a partial differential equation for the density or velocity-density relation or a fundamental diagram (flow-density relation). This model was developed to represent traffic flow by collective traffic flow parameters which include flow rate q (x,t), traffic speed v(x,t) and traffic density ρ (x,t) which are functions of space (x) and time (t). The LWR model is based on conversation equation as follows

$$\frac{\partial \rho}{\partial t} + \frac{\partial q(\rho)}{\partial x} = 0 \qquad (1)$$

LWR model is proposed based on fundamental assumption that the road section without any sources and exits have number of vehicles conserved and also the flow (q) is assumed to be the product of density (ρ) and speed (v). They also proposed that under steady conditions, the ratio of flow to density known as space-mean speed is observed to be nearly constant (for uncongested regime). But during congested conditions due to rise in densities the driving situations vary leading to constraint on drivers from choosing desired speed.

Second-order models on the other hand contain an additional partial differential equation for the average velocity and take into account the finite relaxation time to adapt the velocity to changing traffic conditions. If identical driver-vehicle units are assumed, macroscopic traffic models can be derived from microscopic car-following models [Payne,1971, 1979a; Nelson, 2000; Helbing et al. 2002], and thus the approximations needed in gas-kinetic derivations can be avoided. In recent times simultaneous micro-macro-simulation [Helbing et al. 2002], which can be performed based on empirical boundary conditions are developed. Mesoscopic or hybrid traffic models describe the dynamics of single vehicles depending on the aggregate quantities such as the density [Kates et al. 1998]. And finally queueing models restrict to the temporal change of numbers of vehicles as a function of entering and leaving flows [Kerner, 2001].

Mathematical models which represent traffic flow theory are explained below and also the significance and identification of three states of traffic flow are presented. Kerner (1994) based on the Queuing model proposed three-phase traffic theory:

1. Free flow,
2. Synchronized flow, and
3. Stop-and-go flow (or) wide moving jam

Similarly Polus et al. (2002) proposed three phases of traffic flow as free flow, dense flow, and unstable flow. The traffic breakdown is defined as the change from dense flow to unstable flow. According to Kerner (1998) the movement of downstream front of stop-and-go phase (wide jam) is a deterministic process and is determined by the drivers behavior who tend to escape from the jam (congested section). He defined this velocity of escape ($v_g$) using density ($\rho_{max}$) and average delay time ($t_{del}$) between vehicles following each other trying to escape from the jam as follows,

$$v_g = -1/(\rho_{max} \times t_{del}) \qquad (2)$$

On the other hand Kockelman (2001) considered that each class of driver will be able to control the *spacing* at which he follows the vehicle preceding them. From assumption Daganzo (1997) the spacing is represented as a linear function of speed and inverse of density which is nothing but average spacing between vehicles.  And Kockelman found that for speed determination such assumption is reasonable during congested period using the empirical data. Thus Kockelman proposed that total vehicle density as inverse of average spacing of vehicles, average spacing proportioned by weighted sum of class densities (eq. 3)

$$k = \frac{1}{s} = \frac{1}{\Sigma_i p_i s_i} = \frac{1}{\Sigma_i p_i (a_i + b_i v)} \qquad (3)$$

where,

$s_i$ = inter vehicle spacing,

i = class of vehicle,

$a_i$ and $b_i$ = constants defining the behavior of ith class,

k = density

$p_i$ = proportion of road vehicles of driver/vehicle ith class

Kerner (2005) defined deterministic microscopic traffic model called Speed Adaptation model (SA model) which also considers three-phase traffic theory. SA model is also helpful in showing spatiotemporal congested traffic patterns (from adequate empirical results). Various local driving situations and behaviors are incorporated into the model. SA model considers minimum free flow speed ($\mathrm{v}_{\mathrm{min}}^{(\mathrm{free})}$), space gaps (during jams or stop-and-go states) and vehicle accelerations depending on traffic state (free or synchronized or stop-and-go flow). Equations which form the basis of the model are as explained below.

$$\frac{\mathrm{dx}}{\mathrm{dt}} = \mathrm{v} \tag{4}$$

$$\frac{\mathrm{dv}}{\mathrm{dt}} = \begin{Bmatrix} \mathrm{a}^{\mathrm{free}} \text{ at } \mathrm{v} > \mathrm{v}_{\mathrm{min}}^{(\mathrm{free})} \text{ and } \mathrm{g} > \mathrm{g}_{\mathrm{max}}^{(\mathrm{jam})}, \\ \mathrm{a}^{\mathrm{syn}} \text{ at } \mathrm{v} \le \mathrm{v}_{\mathrm{min}}^{(\mathrm{free})} \text{ and } \mathrm{g} > \mathrm{g}_{\mathrm{max}}^{(\mathrm{jam})}, \\ \mathrm{a}^{\mathrm{jam}} \text{ at } 0 \le \mathrm{g} \le \mathrm{g}_{\mathrm{max}}^{(\mathrm{jam})}. \end{Bmatrix} \tag{5}$$

Where,

x = vehicle space coordinate,

t = time coordinate,

v = speed

$\mathrm{v}_{\mathrm{min}}^{(\mathrm{free})}$ = free flow minimum speed

$\mathrm{a}^{\mathrm{free}}$, $\mathrm{a}^{\mathrm{syn}}$and $\mathrm{a}^{\mathrm{jam}}$ vehicle acceleration (deceleration) in the free flow, synchronized flow and stop-and-go (wide moving jams),

g = space gaps,

$\mathrm{g}_{\mathrm{max}}^{(\mathrm{jam})}$ = maximum space gaps during jams.

Further Kener simplified and presented math functions with space gap (g) as variable to calculate average speed in synchronized traffic state $\left(V_{av}^{(syn)}\right)$. They are as follows,

$$V_{av}^{(syn)}(g) = \frac{g'(g)}{T_{av}^{syn}}, \tag{6}$$

$$V_{av}^{(syn)}(g) = V_1\left[\tanh\left(\frac{g'(g)}{T_{av}^{syn} \times V_1}\right) + cg'(g)\right] \tag{7}$$

where,

function $g'(g) = g - g_{max}^{(jam)}$,

$g$ = space gaps,

$V_1$ = steady state speed (constant),

$T_{av}^{syn}$ = average safety time gap in synchronized state (constant),

c = constant

Kerner explained that the onset of congestion in SA model is associated with an F → S transition and stop-and-go state (moving jam) occurs spontaneously in synchronized flow (observed from empirical results). SA model is thus shown useful to simulate an F → S transition and also the features the sequence of F → S → J transitions. SA model thus confirms the assumption of three-phase traffic theory regarding the fundamental hypothesis about the F→S→J transitions (based on mathematical models). Kener further modeled the transition from two effects one that the discontinuity of steady speed solutions (from Figure 9 (a), (c) and (e)) or their instability in area close to the maximum points of free flow ($v_{min}^{(free)}, \rho_{max}^{(free)}$). Another Speed adaptation (SA) effect is modeled through $K(v, v_l)(v - v_l)$ which adjusts the present speed to the speed of proceeding vehicle in synchronized flow. In Figure 9 a, c represent space-gap-speeds and b, d, e, f, g, h represent SA models for flow-density plane during different steady state models of SA model.

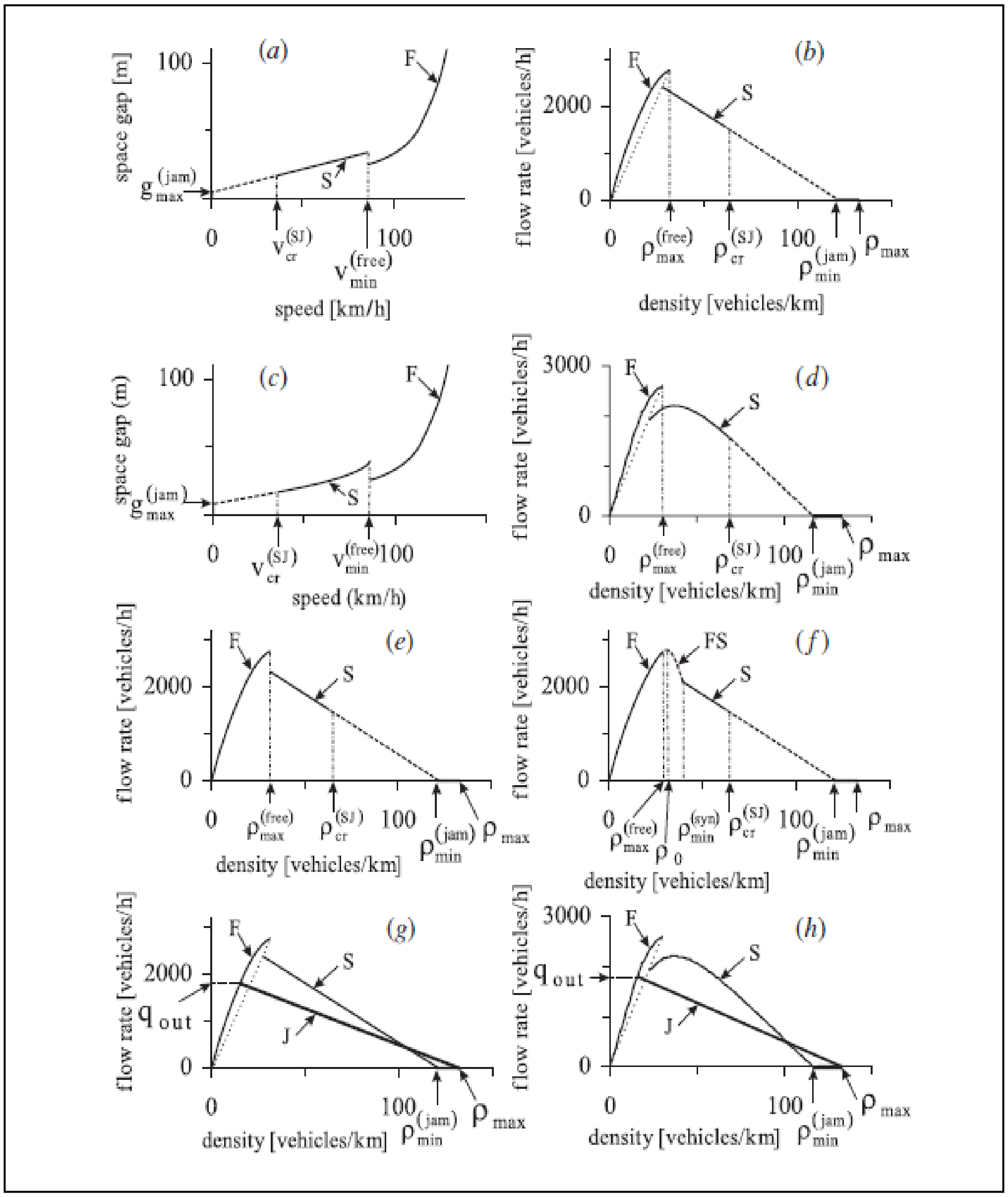


**Figure 9: Steady state Speed Adaptation (SA) model**

#### *1.2.5.1 Free flow*

Free flow traffic state is defined as the state in which vehicle density is low enough and it is easy for vehicles to pass or overtake each other on multi-lane roads (Kerner 1999). Speeds are usually much higher than in congested traffic and driver has the opportunity to travel at his desired speed. In free flow state average vehicle speed may be different for different lanes. Individual vehicle speeds depend on different parameters such as lane, geometric design and driver's perspective of his safe and comfortable speed. The low vehicle density in free flow makes it easy for vehicles to overtake each other and hence the average speed for each lane tends to be different.

In the Figure 10 (a) shows speeds in different lanes during free flow and we can observe that speeds usually range between 45 to 70 mph (70 to 110 kph). And also the densities are less than or around 32 vehicles per mile (20 vehicles per kilometer). If the density is high enough, the flow becomes congested and a transition from free flow to synchronized flow occurs.

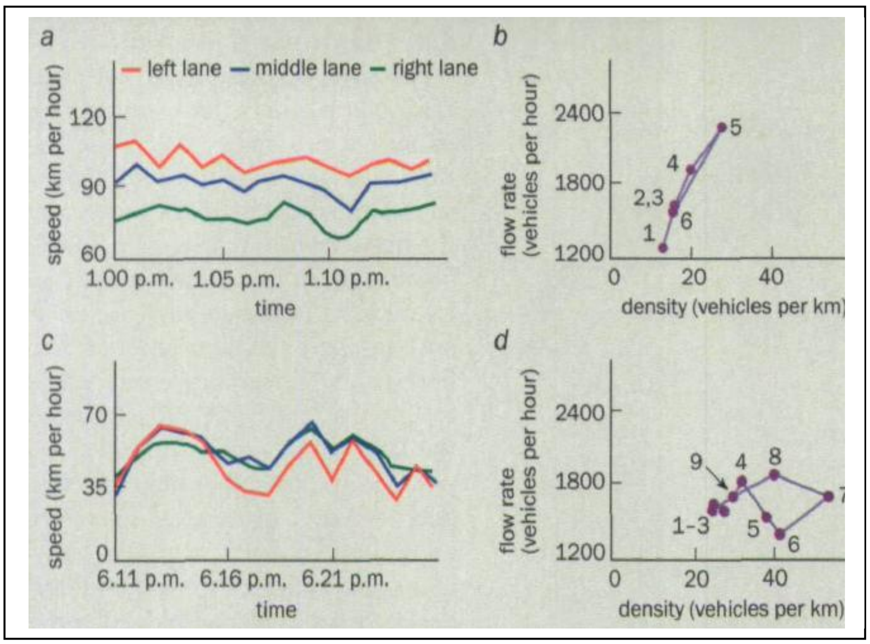


**Figure 10: Comparison between free flow (a, b) and synchronized flow (c, d)**
(Source: The physics of traffic; Kerner, 1999)

#### *1.2.5.2 Synchronized flow*

Synchronized flow can be considered as the state which is between the free flow and wide moving jam. In synchronized flow vehicles in different lanes move with almost the same speed. Density is high and overtaking is not so easy thus forcing individual vehicles to move with almost the same average speed in the different lanes.

Thus synchronized flow is characterized with low average vehicle speed, high flow rate (higher than free flow). In synchronized flow, fluctuations in speed can be noticeably smaller in amplitude compared to free flow as vehicles tend to travel as a group. In Figure 9 (c) Kerner, 1999 observed that speeds are as low as 18 mph (30 kph) and as high as 45 mph (70 kph). Similarly in another research by Habib-Mattar et al. (2009) defined the beginning of the unstable flow using speed and density. They observed speed drop during the transition from free flow to unstable flow for at least a 5-minute period and this is then accompanied with increase in density.

#### *1.2.5.3 Stop-and-go flow*

Kerner (1998) summarized that the in a density range where homogeneous states of traffic flow cannot exist due to instability or phase transitions leads to stop-and-go pattern [Kerner, 1998]. Shockwaves occur as a result of differences in flow and density which occur when there are constrictions in traffic flow. These constrictions lead to heavy congestion. The speed at which the growth of the ensuing queue results in shockwave, and is defined as the difference in flow divided by the difference in density.

The stop-and-go traffic is considered as extreme unstable flow. Stop-and-go flow (jam) emerges in two steps first with the phase transition of free flow to synchronized flow which is followed by step known as the "pinch effect”. During this process the synchronized flow compresses itself into a very high density state. Spontaneous local perturbations grow and lead to a traffic jam. Compared to synchronized flow, stop-and-go is characterized by very high density but both flow rate and vehicle speed are very low. In the Figure 11 we can observe the speed distributions during stop-and-go flow range between 0 MPH to 15 MPH.

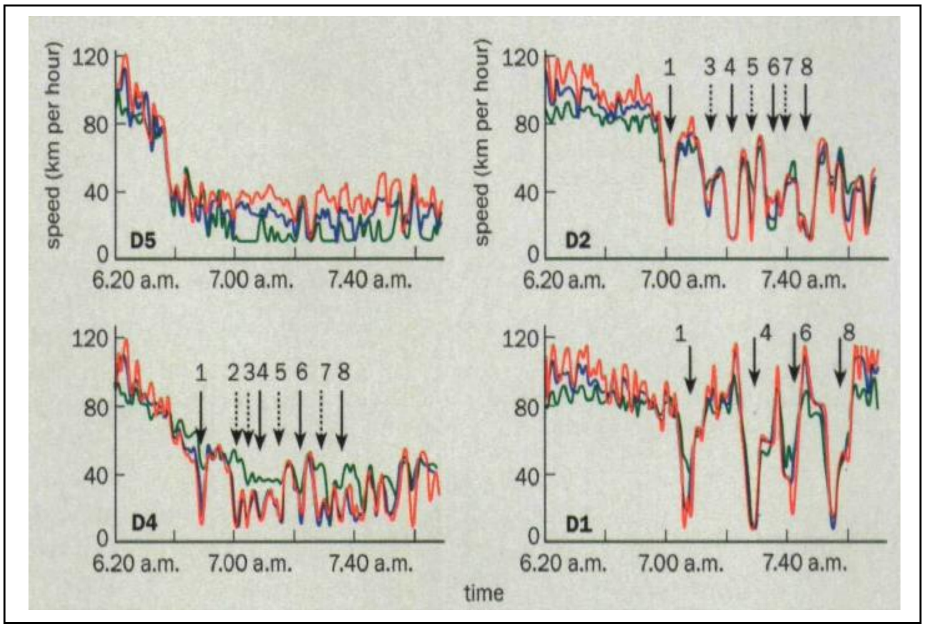


**Figure 11: Speed pattern observed during stop-and-go flow**
(Source: The physics of traffic; Kerner, 1999)

From Kerner’s SA model empirical observations (Figure 9) and also from above definition, traffic parameter threshold values are tabulated (Table 3). This presents a comparison of different traffic flow parameters thresholds values during different traffic flow states. The traffic flow parameters here include speed (v), dual loop on-time difference (OnT1-OnT2), density (k), flow (q) and gap (g).

**Table 3: Thresholds of traffic flow parameters.**

| **Traffic flow parameters** / **Traffic flow states** | **Speed**, v (miles per hour) | **Loop on-time difference** (OnT1- OnT2) (seconds) | **Density**, k (no. of vehicles per mile) | **Flow**, q (no. of vehicles per hour) | **Gap** between vehicles, g (ft) |
|---|---|---|---|---|---|
| **Free flow** | v > 45 | < (+/- (3.5/60) | k < 40 | 0 – 2400 | g > 82 |
| **Synchronized** | 45 > v > 15 | ≥ (+/- (3.5/60) | 40 < k < 120 | 1300 – 2200 | 82 > g > 40 |
| **Stop-and-go** | v < 15 | | k > 120 | q < 1300 | g < 40 |

## 1.3 Problem statement

As explained in earlier sections the existing vehicle length based classification algorithm assumes that vehicle travels at constant speed over dual loop detection area. The above assumption is true when the flow over the loops is free-flow. This results in identical on-times on both loops (M and S). Existing model can thus produce satisfactory results with high accuracy for vehicles which travel during free flow. But during congestion this assumption (constant speed) is violated as vehicles travel at non constant speeds. Not only during stop-and-go flow but also during synchronized flow, vehicles travel for longer periods of time and slight acceleration or deceleration have much effect on vehicle length estimation.

Thus miscalculation of vehicle length resulting from these time stamps is a source for inaccurate vehicle length based classification. Thus factors such as acceleration, deceleration of vehicles on loops cause vehicle trajectory to vary abruptly and need to be determined for accurate classification data. The inaccurate detection of these factors during congestion is the main cause for the failure of existing model. Modification of the existing model for congested flow can thus provide accurate vehicle classification data.

## 1.4 Significance and scope of the research

As explained in earlier sections this research deals with dual loop detectors which are widely used traffic detectors on interstates. They are used to collect speed, occupancy and vehicle classification data. Vehicle classification data is useful piece of information for many transportation planning, design and analysis requirements. Many researchers worked in correcting dual loop detector data which includes sensitivity problems, pulse breakups, cross talk between loops, etc. But very little importance was given to determine the effect of congestion on vehicle detection.

The existing vehicle classification algorithm (section 1.2.4) which works on the assumption of constant vehicle speed over dual loops is true for free flow but not during congestion. Free flow is the period when the driver has the ability to travel at his desired safe speed (often at speed limits) and will have the freedom to change lanes and overtake slow moving vehicles. On the other hand congested traffic flow occurs due to bottleneck creation during merging of lanes at exits and on ramps and also during accidents and extreme climatic conditions. This flow is featured by decreased gaps between vehicles in the queue, very low speeds and significant amount of accelerations (decelerations) in vehicle trajectory. This leads to failure of existing vehicle classification model during congestion and especially for vehicles which move over detection area with significant amount of acceleration (deceleration) values. Thus development of new vehicle classification model (CAVC model) for accurate data is required during congestion.

The scope of this research mainly aims in improving the model for synchronized flow and also for vehicles which does not stop during stop-and-go flow. Thus the research aims in improving accuracy by developing a model which calculates a single acceleration (deceleration) value and use it for estimation of more realistic vehicle length during congestion. Identification of vehicle flow state is proposed and is necessary to select the type of model (existing or CAVC) to be used. This methodology uses traffic parameters calculated for each vehicle to identify the flow state and then switch to appropriate classification model. Hence existing model is used during free flow and CAVC model is used during congested flow (both synchronized flow and for non-stopping vehicles during stop-and-go flow). Thus this research deals with vehicle classification issues which are inherent in existing model (algorithm) which impact classification data.

## 1.5 Outline of thesis

The thesis report presents detailed procedure followed in data collection, extraction and analysis and also presents the background, new vehicle classification model and results. The outline of the thesis below presents the scope of each chapter covered in the thesis:

<u>Chapter 1: Introduction</u>

After a brief introduction "Goals and objectives" are presented which discuss about possible deliverables from the research work presenting objectives of research.This chapter then gives a detailed "Background" of the research topic with focus on types of vehicle detection systems, vehicle classification system, loop detector principle with existing vehicle classification model and finally presents traffic flow theory.  This presents literature review and also fundamentals of the research topic useful for understanding the gap in the research.

This chapter also includes "Problem statement" discussing the problem being addressed in the research, and finally "Significance of the research" presents the importance of research work done.

Chapter 2: Data collection methodology

This chapter presents the data collection methodology which includes study sites surveyed and selected. Then the data collection procedure is explained for both forms of data (video and raw loop event data). Then finally the data extraction methodology using VEVID software is presented with sample results.

Chapter 3: Evaluation of existing classification model

Before we move on to the development and evaluation of new vehicle classification model we need to present results showing the failure of existing model during congestion. This evaluation is made with the help of ground truth data extracted (as explained in chapter 2). Chapter 3 serves thus presents results by categorizing them into free, synchronized and stop-and-go flow (flow identification is as discussed in sections 1.2.5 and 4.2). Thus this chapter ends concluding the failure of existing model during both forms of congestion that is synchronized and stop-and-go flow.

Chapter 4: Constant Acceleration based Vehicle Classification model (CAVC)

Once the existing model is proved to fail during congested flow condition a new vehicle classification model named as Constant Acceleration based Vehicle Classification model (CAVC) is presented. Finally presents a traffic flow identification methodology which is used to determine the type of model to be used for the extraction of vehicle classification data from raw loop data.The principle and the concept behind this identification is presented with a detailed flow chart of data processing yielding vehicle length based classification.

Chapter 5: Results and discussion

CAVC model is evaluating using vehicle trajectory data collected.  This chapter then ends with concluding remarks from the results and also talks about future research possible in this area. Finally the significance of this research and its contribution to the area of vehicle classification data collection is discussed.

# CHAPTER 2: DATA COLLECTION METHODOLOGY

One of the major phases in this research is the data collection effort; this chapter explains a detailed methodology which was followed in collecting and extracting data. Both raw loop event data and corresponding video data are collected. Video data collection was mainly aimed to collect ground truth data during different states of congestion. Raw loop event data (loop data) is also collected simultaneously from the study site (loop detector station). Both existing and improved vehicle classification models are then used to process loop data for extracting vehicle classification (Chapter 3 and 5) and then these results are evaluated using the ground truth data extracted from video files.

This chapter thus presents details of study sites selected for data collection such as location of dual loop detector station, place of setting up the camera (with images of the loop station). Later the chapter deals with extraction methodology of vehicle trajectory data using "Vehicle Video-Capture Data Collector" (VEVID). Setting up of VEVID reference system to extract the trajectory data and also the new GPS based method is used for this setting up of reference points is explained. The trajectory information here includes on-time, vehicle length and speed. In this research, consecutive three weekdays of data was collected during July 14th, 15th and 16th, 2009.

## 2.1 Study sites

A field study was initially conducted in Columbus aiming to find feasible dual-loop detector locations which could be videotaped for the ground truth data. A total 16 dual-loop detector stations near Columbus were observed for suitability in respect with possible camera location, loop data quality and the amount of congestion observed near the station. List of loop stations surveyed and their location:

1. V0001: Near Long St towards North I 71.
2. V0004: Near East 11th St on I 71.
3. V0007: Near Velma Rd on I 71.
4. V0010: Near East Weber Rd on I-71.
5. V0013: Near East North Broadway St on I 71.
6. V0016: Near E Cooke Road on I-71.
7. V0019: Near Morse Road on I-71.
8. V0022: Near E Dublin Granville Rd on I-71.
9. V0025: Near Schrock Rd on I-71.
10. V0027: Near exit 119A on I-71.
11. V0030: Near Park Rd on I-71.
12. V0033: Near Polaris Pkwy on I-71.
13. V1002: Near S Front St on I-70/71.
14. V1003: Near S 4th St on I-70/71
15. V1005: Near S Grant Ave on I 70/71
16. V1008: Near E town St on I 71.

After initial survey of the detector stations, two stations V1002 and V1003 for data collection are selected depending on suitable place to mount the camera and also the congestion patterns observed. These dual-loop stations are located on Interstate 70/71 which is within downtown Columbus, OH. At this location, recurrent congestion has been observed during both morning and evening peak hours. A camcorder was set nearby to videotape the traffic congestion. Following section gives a description of the dual loop study sites selected.

### 2.1.1 V1002 dual loop detector station

This dual-loop detector station is located on I-70/71 at S Front Street.  Both sides (i.e. I-71 north and south) have the detectors (Figure 12, 13). The location is within downtown area with very close on ramps and off ramps and observes recurrent stop-and-go traffic during peak hours. Detectors on both eastbound and westbound pavement are clearly visible.

<u>Camera Location</u>: Camera is set in the corner on the top (the 5$^{th}$ Floor) of the Franklin County parking garage. The distance between the detectors and the possible camera location is estimated to be about 70 meters (Figures 12). The parking garage is owned by the Franklin county Juvenile department and operated by Public facility management.

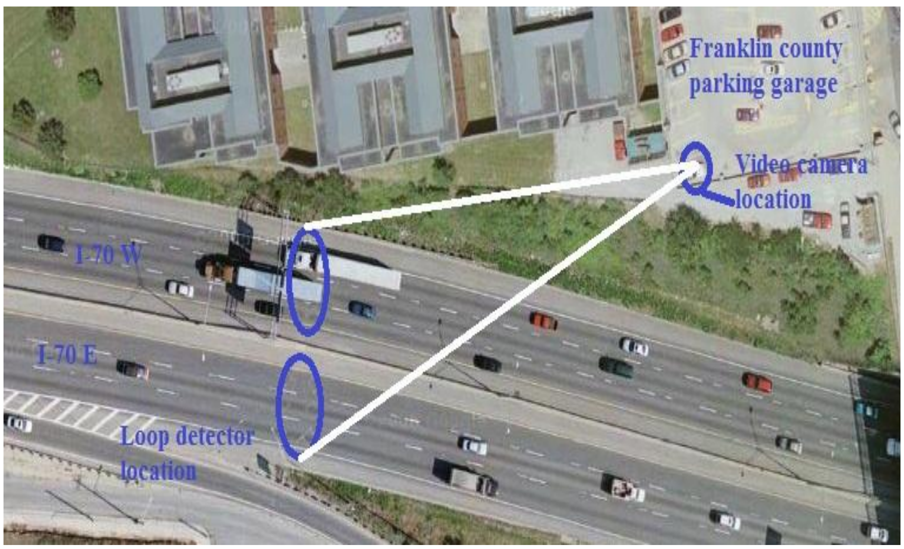


**Figure 12: V1002 detector station and neighboring area.**
(Source: Picture taken from the top of Franklin County parking garage)


**Figure 13: View of the V1002 detector station from parking garage.**
(Source: Google maps, accessed on July 5th 2010)

### 2.1.2 V1003 dual loop detector station

This loop detector station is located on I-70WB/I71SB near 4th Street (Figure 14). The location is within downtown area with very close on ramps and off ramps and experiences recurrent congestion during peak hours.


**Figure 14: V1003 detector station and neighboring area.**
(Source: Google maps, accessed on July 5th 2010)

Camera Location: Camera is set in the school on the right side of I-71N and it is close to the Interstate. The school is on a higher level compared to the pavement, and the ideal camera location is at the edge of the school (Figure 15). The estimated distance between the camera and the detectors is about 30 meters.


**Figure 15: View of V1003 detector station from the school.**
(Source: Picture taken from Col. Africentric Ec Elementary School parking lot)

## 2.2 Video data collection

As mentioned earlier the video data is collected simultaneously for extracting ground truth data. Selected dual loop locations are videotaped from an elevated position from which the vehicle traveling over the dual loops can be viewed, as shown in Figure 16. Data is collected from the V1002 and V1003 dual-loop stations located on Interstate 70/71. This is within downtown Columbus, OH which usually observes congestion. At this location, recurrent congestion has been observed during both morning and evening peak hours. A camera was installed on the top floor of a parking garage to videotape station V1002 next to Interstate

(Figure 17). Similarly an elevated position from school parking for V1003 dual loop station was selected.

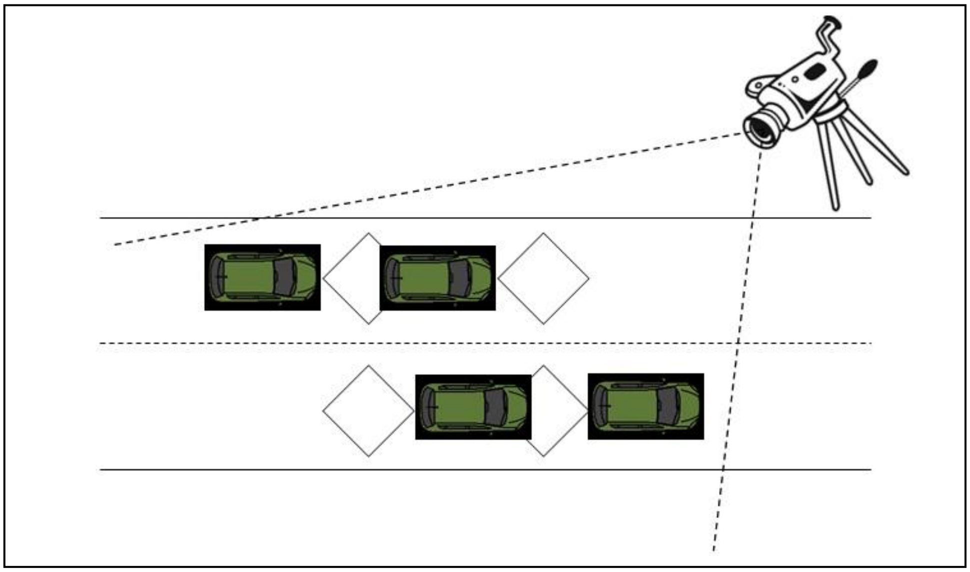

**Figure 16: Videotaping procedure demonstration**


**Figure 17: Videotaping at selected loop station**.
(Source: Picture taken from the top of Franklin County parking garage)

A vehicle with a GPS data logger was made to drive over the loop detector station at a constant speed. Cruise control feature has been used to maintain constant speed. This is used for setting reference system for extraction of trajectory data. The video tapes were converted into AVI files at a specified frame rate to be compatible for extraction using VEVID. VEVID needs each small segments of video file; hence the AVI files are split into multiple files of about 1 minute each.

## 2.3 Event dual loop data collection

Raw loop event data (loop data) at these stations is requested from ODOT (Ohio Department of Transportation) with the help from Dr. Benn Coifman from Ohio state university (OSU). Traffic management center (TMC) was also instrumental in the data collection of raw loop actuation data for the same period of time during the field data collection. The format of sample raw loop data is as shown in earlier section 1.2.4 (Table 2).

Raw loop actuation data is the loop event data and is accurate to 1/60th of a second. Sixty data points for every second are collected. Thus data is collected at proposed study sites along the freeway section which is experiencing congestion during peak hours of traffic. Particularly care was taken that traffic over loops experience congestion during data collection.

## 2.4 Vehicle trajectory data extraction

Once the video data was collected the next step is to extract the data. This data thus forms the ground truth data which is used for evaluating existing vehicle classification model and also new improved vehicle classification model. This section introduces VEVID software used for data extraction, setting up of reference system and final data collection procedure.

### 2.4.1 Introduction of VEVID

To extract ground-truth vehicle event data, the software VEVID (Vehicle Video-Capture Data Collector) is used to extract high-resolution vehicle trajectory data from videotapes. The software VEVID had been developed to extract accurate trajectory data [Wei et al. 2005], and the accuracy of its outputs has been proven [Wei, 2008]. A new GPS based setting up reference points in VEVID has been used in our research. Using VEVID, the timestamps, on-times, speeds, and lengths are extracted.

### 2.4.2 Setting reference system for data extraction

Reference points are required to be measured in the field. Traditional way of setting up reference points is done through marking intervals (approximately 20ft) along both sides of the roadway. This traditional method of setting reference points on Interstate poses safety concerns for data collection.

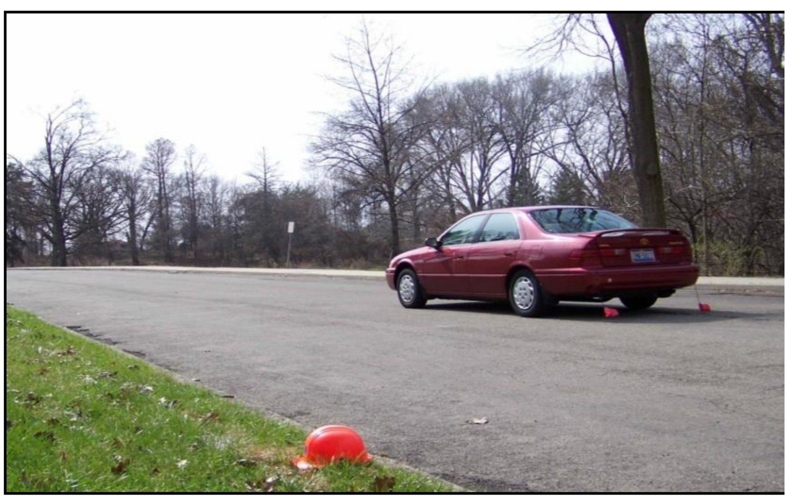

**Figure 18: Vehicle demonstrating VP-GPS method of referencing**

In order to solve this case a new method using Global Positioning System (GPS) device is used. **V**ideo-capture, **P**erspective drawing techniques using Cruise control function and **GPS**-based probe technology form the base of this new approach and thus named as VP-GPS approach by University of Cincinnati research team.

With the VP-GPS approach, no staff is needed in field while the accuracy of the reference points is well ensured. The new method (Figure 18) uses two small red flags symmetrically to the rear bumper of a car. Flags are hanged under the bumper and the flags ends are placed just above the road surface such that they do not touch the ground (Figure 18). The car with red flags runs to and fro passing the segment of the road, using the cruise control function to maintain a constant speed which is monitored by GPS travel logger (Figure 19).


**Figure 19: GPS travel logger instrument**
(Source: www.eFronteir.com, accessed July 14th 2010, 19)

This method of reference system marking has been used and its accuracy has been tested in field by the University of Cincinnati research team. Flags are used to locate the exact point on the ground. Vehicle positioning from video frames using video-capture and linear perspective drawing techniques are used to determine reference points using VEVID software.

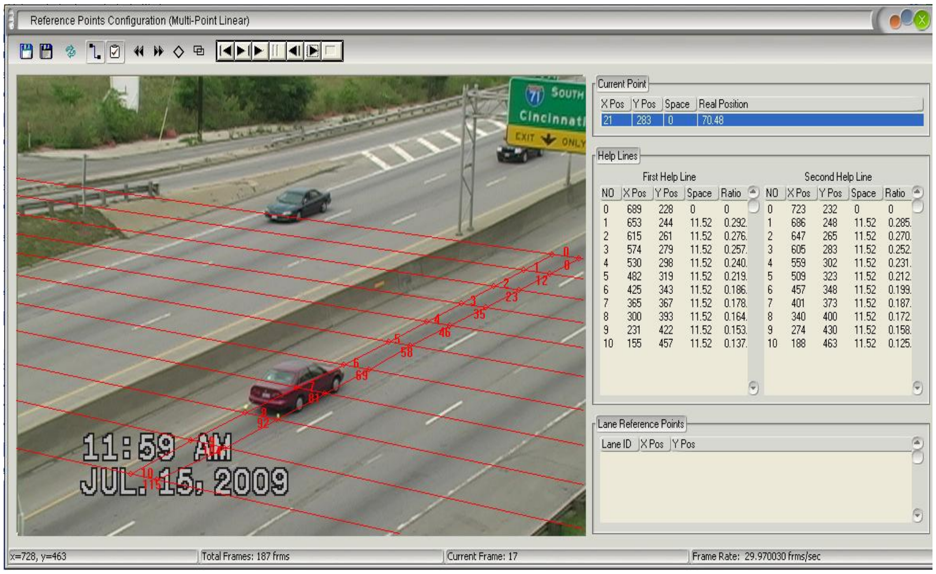


**Figure 20: VEVID interface showing reference system**

Speed probed by the testing vehicle is used to determine the reference spacing intervals, and then a real (ground) distance coordinate system is formed in VEVID. Figure 20 shows the VEVID interface with reference system and is ready for extraction of vehicle trajectory data.

### 2.4.3 Data extraction

Once the reference system is successfully established then process of manual extraction of trajectory data starts (Figure 21). The vehicle trajectory ground-truth data extracted using VEVID, is used to evaluate both existing vehicle classification model and also proposed new vehicle classification model (CAVC) for different traffic states. Table 4 below shows sample vehicle trajectory data extracted from raw video data using VEVID. Speed on M and S are calculated similar to existing loop algorithm. Frame ID (number) is used to calculate time

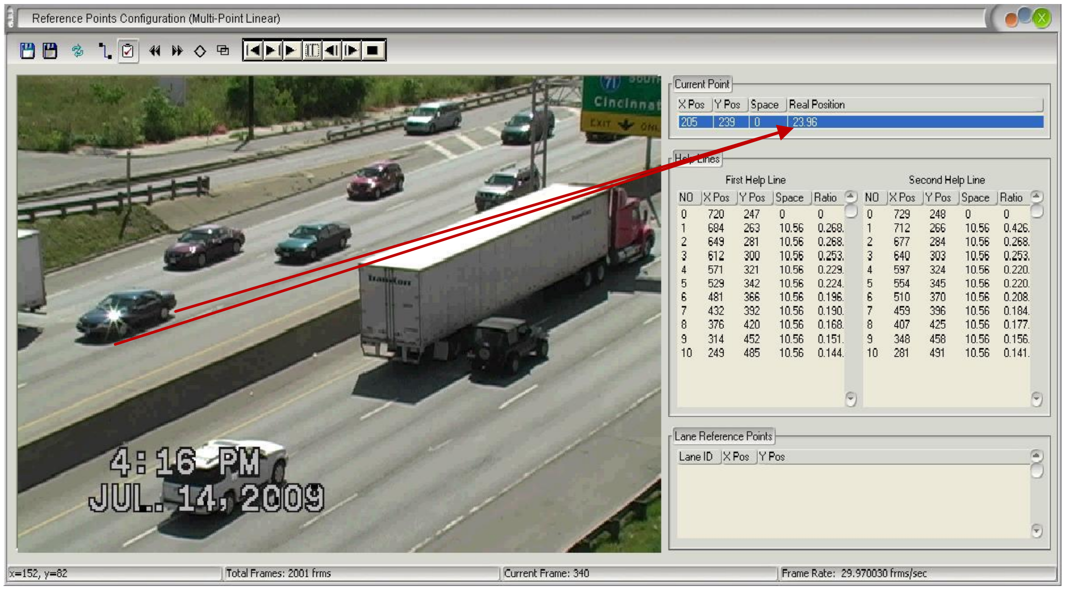


**Figure 21: Demonstration of vehicle length data extraction using VEVID**

**Table 4: Format of extracted trajectory data**

| Veh No. | Speed on M loop (mph) | Speed on S loop (mph) | Ontime1 (M loop) (Sec) | Ontime2 (S loop) (Sec) | Vehicle Length (ft) |
|---|---|---|---|---|---|
| 1 | 58.47078 | 54.03409 | 0.3 | 0.266667 | 15.4 |
| 2 | 54.20455 | 51.75 | 1.033333 | 1.033333 | 70.29 |
| 3 | 48.53571 | 47.1733 | 0.3 | 0.333333 | 15.73 |
| 4 | 55.02273 | 51.51989 | 0.3 | 0.333333 | 16.03 |
| 5 | 52.56818 | 47.30114 | 0.3 | 0.333333 | 15.28 |
| 6 | 49.34659 | 45.5 | 0.8 | 0.833333 | 49.81 |
| 7 | 51.34091 | 45.53693 | 0.4 | 0.4 | 19.57 |
| 8 | 49.62784 | 46.86364 | 0.366667 | 0.4 | 18.6 |
| 9 | 46.05195 | 46.79545 | 0.333333 | 0.4 | 18.09 |
| 10 | 47.62987 | 42.46875 | 0.366667 | 0.4 | 15.13 |
| 11 | 49.16761 | 48.78409 | 0.3 | 0.3 | 15.18 |
| 12 | 49.73864 | 55.15057 | 0.733333 | 0.8 | 49.9 |
| 13 | 50.69805 | 53.46307 | 0.3 | 0.333333 | 15.94 |
| 14 | 49.90909 | 47.76136 | 0.333333 | 0.333333 | 17.07 |

travelled between different vehicle motions. In VEVID we can achieve up to an accuracy of 1/30th of a second frame rate. Both vehicle location (obtained from reference system) and frame ID (or values) are used for calculating flow parameters like speed and ontime whereas individual vehicle length is obtained from the location of front bumper and rear bumper of the vehicle.

# CHAPTER 3: EVALUATION OF EXISTING VEHICLE CLASSIFICATION MODEL

The existing vehicle classification model as explained earlier is based on the assumption that individual vehicle speed over the dual loop detection area is constant. To evaluate this model, vehicle length and also vehicle classification during different traffic states is extracted using existing vehicle classification model and compared with ground truth data. The comparison is made in different traffic flow states separately. The ground truth data is collected from video data using VEVID software (as explained in earlier section 2.4.3). This is done by synchronizing the loop and video timestamps and extracting individual trajectory data which yields ground truth data. Different traffic states considered are “free-flow”, “synchronized flow” and “stop and go flow”. Thus existing vehicle classification data is compared with ground truth data for the above mentioned traffic states.

Ground truth data was extracted for free flow condition is based on start and end timestamps from both video segment and loop timestamps (total numbers of vehicles were matched). On the other hand for synchronized and stop-and-go flow conditions the trajectory data is extracted by matching individual vehicles in both data formats. This is done by verifying the vehicle detected on the loop (from raw data) with that from video segment (upon careful observation of timestamps).

## 3.1 Existing model during free flow

Existing classification model is compared with the ground truth vehicle classification data for free flow condition. Ground truth data for V1002 EB lane 3 for July 15 is collected using VEVID. A total of 661 vehicles are collected which forms our sample to evaluate the existing model for free flow state. Existing vehicle classification model is used to estimate vehicle lengths from raw loop data and later vehicles are classified based on their length.

**Table 5: Vehicle classification in to bins from existing model (free flow)**

| | | **Ground truth classification** | **Existing classification model** | **Misclassification** (Existing-Ground) (No. of vehicles) |
|---|---|---|---|---|
| | | No of vehicles | No.of vehicles | |
| **ODOT** | | | | |
| Small | <= 28ft | 564 | 573 | (573-564) = 9 |
| Medium | <=46ft | 23 | 25 | 2 |
| Large | >46 ft | 74 | 63 | -11 |
| **WSDOT** | | | | |
| Bin 1: | <= 26ft | 558 | 566 | 8 |
| Bin 2: | <= 39ft | 24 | 26 | 2 |
| Bin 3: | <= 65ft | 21 | 18 | -3 |
| Bin 4: | > 65ft | 58 | 51 | -7 |
| **Total** | | 661 | 661 | |

The classification of data as per standard ODOT and WSDOT vehicle length based classification bins can be seen in Table 5. The table shows a comparison between the classification results obtained from both existing and ground truth data. According to ODOT classification an error (misclassification) of -11 vehicles (negative sign represents less vehicles detected by model than actual ground truth classification) was found for large vehicles. Similarly as per WSDOT classification results we can see that Bin 3 and Bin 4 were underestimated (-3 and -7 vehicles) by existing vehicle classification model. This error is small compared to sample size taken and also considering the fact that the sample is not formed by identical matching of vehicles. This mismatch is due to the fact that data was collected between 2:24 PM and 2:45 PM on July 15$^{th}$ in lane 3 which yielded 661 vehicles (video). And only 642 vehicles (loop data) during that time, this under detection of loop data is due to detection errors (pulse break ups and omitted phases). More 19 vehicles were added from loop data beyond 2:45 PM and leading to samples which are not completely identical. Classification error is also due to observed lane changing (5 vehicles) and presence of trolleys (vehicles with multiple units) (3 vehicles). Hence the error is due to during vehicle detection and also the fact that the sample may not be exactly identical lead to this minor error. But in case of congested vehicles they are exactly matched form both the data and all the vehicles which observed lane changing behavior (or detection error) are eliminated from the analysis. And thus for congested case the samples are exactly identical. Hence we can conclude from observing the vehicle lengths and classification data (Table 4) that the existing model accurately estimates vehicle length during free flow. Thus existing model works well during free flow traffic condition.

## 3.2 Existing model under synchronized flow

Synchronized flow as previously explained (section 1.2.5.2) refers to slow moving traffic and considered as an intermediate state between free flow and stop-and-go flow (wide moving jams). Vehicles observe acceleration (deceleration) during this period resulting in change in speed over the detection area. This contradicts the basic assumption of constant speed of existing vehicle classification model. This leads to misclassification of vehicle data for synchronized flow state. Classification data is evaluated using ground truth data. The ground truth data in this case is collected as explained earlier by identifying individual vehicles using both timestamps. Thus each individual vehicle length from existing model is compared with the ground truth vehicle length.

For estimating required sample size for synchronized vehicles we assume a confidence level of 95% (z = 1.96) and error allowed (precision, e) as 2ft and standard deviation (σ) of 16.55 (from ground truth data).

$$n_0 = \left(\frac{Z^2 \times \sigma^2}{e^2}\right) \tag{8}$$

Where,

$n_0$ is the sample size,

z is the abscissa of the normal curve that cuts off an area at the tails,

e is the desired level of precision (in the same unit of measure as the variance),

$\sigma$ is the standard deviation of an attribute in the population.

The calculated minimum sample size for synchronized flow calculated using the formula mentioned above (eq. 8) is 264 vehicles. The sample size used in the study is 414 vehicles which more than required sample size. A total of 414 vehicles are collected from both V1002 (July 14$^{th}$, 2009) and V1003 (July 16, 2009) stations. Initial observations confirmed that vehicles were overestimated in most of the cases and many smaller vehicles were detected as medium or large vehicles. To explain this further a graphical representation of the length estimates of existing model and ground truth data are presented in Figure 22. By observing the data we can understand the impact of acceleration and deceleration on vehicle length estimate.

As we can see from the graph the results are much dispersed from the linear line (y=x), which shows the inaccuracy of the existing model in predicting the vehicle lengths. Two straight lines representing error ranges are drawn on either side linear line (y=x) to form upper and lower boundaries of the dispersion of data points. Two straight lines representing error ranges of $\pm 40\% \left(\frac{(\text{true length}\pm(0.4\times\text{true length}))-(\text{true length})}{(\text{true length})}\times 100\right)$ are drawn connecting origin (0, 0) and $\{100, (100 \pm 40)\}$. These lines ($y = \frac{7}{5}x$ and $y = \frac{3}{5}x$) show the dispersion of vehicle length estimation about the ground truth value.

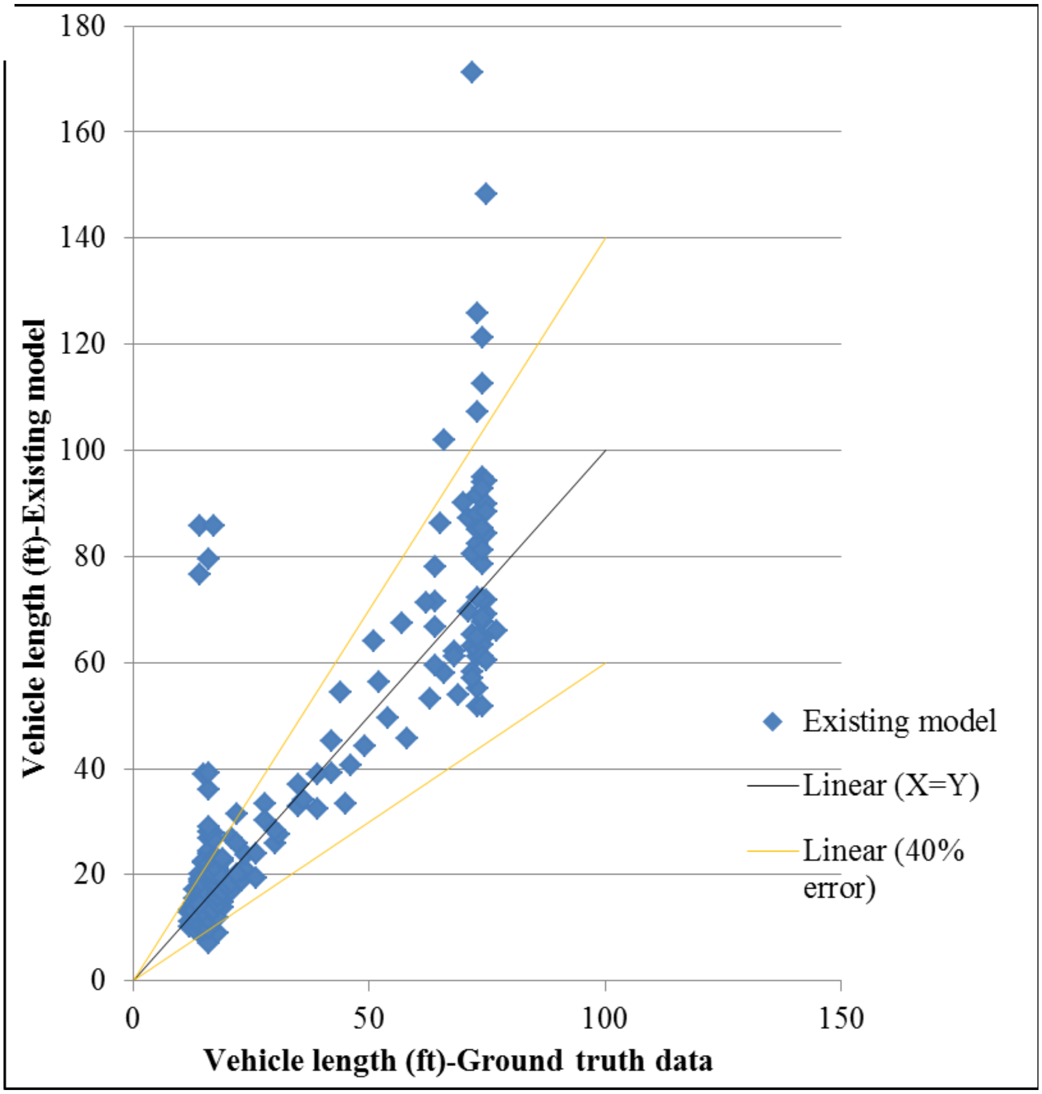


**Figure 22: Existing model vs Ground truth data (synchronized flow)**

From graph we can observe the impact of acceleration (deceleration) resulting in underestimating (overestimating) vehicle lengths (outliers). Observed outliers are for example

171ft, 148ft, 125ft, 121ft etc. On detailed analysis of these data points (outliers) yielded valuable information required for adjusting the model. For example a vehicle detected as 11ft was actually 18ft. The vehicle accelerated resulting in on-times difference between loops and underestimated vehicle length by about 7ft. For example vehicle detected as 148ft is actually 75ft long vehicle. This vehicle decelerated on the loop detection area. Similarly another vehicle which is detected as 171ft is actually 72ft long. This vehicle similar to previous example observed deceleration.

Finally classifying vehicles into standard vehicle classification bins during synchronized flow is as shown in the Table 6. We can observe that existing vehicle classification is not accurately predicting the classification. Especially the small and medium class (ODOT) and also all four classes of WSDOT classification have significant amount of misclassification (error). For small vehicle (ODOT) class of misclassification by -8 vehicles (negative represents that less vehicles are identified). Similarly for Bin 1, Bin 2 and Bin 3 we can observe -15, 11 and 14 vehicles of misclassification respectively.

**Table 6: Vehicle classification in to bins from existing model (synchronized flow)**

| | | **Ground truth** | **Existing model** | Misclassification (Existing-Ground) (No. of vehicles) |
|---|---|---|---|---|
| | | No of vehicles | No.of vehicles | |
| **ODOT** | | | | |
| Small | <= 28ft | 328 | 320 | (320-328) = -8.00 |
| Medium | <=46ft | 15 | 20 | 5.00 |
| Large | >46 ft | 71 | 74 | 3.00 |
| **WSDOT** | | | | |
| Bin 1: | <= 26 ft | 330 | 315 | -15.00 |
| Bin 2: | <= 39 ft | 8 | 19 | 11.00 |
| Bin 3: | <= 65 ft | 18 | 32 | 14.00 |
| Bin 4: | > 65 ft | 58 | 48 | -10.00 |
| **Total** | | 414 | 414 | |

Small vehicles have been identified as medium vehicles or large vehicles for ODOT classification. And similarly Bin 1 vehicles have been misclassified as either Bin 2 or Bin 3 class vehicles. Thus the existing model in most of the cases has overestimated small vehicles as large vehicles. Hence the existing vehicle classification model needs to be modified to improve the vehicle classification for synchronized traffic flow condition.

## 3.3 Existing model under stop-and-go flow

Stop-and-go traffic flow which is also known as “wide moving jam” is most complex scenario of the congested flow. During stop-and-go flow vehicles usually travel at lower speeds compared to synchronized flow and occasionally stop on the dual loop detection area. Vehicles during stop-and-go flow pose a challenging task as vehicles also undergo quick acceleration and deceleration motions. In this section the existing vehicle classification model is evaluated for the stop-and-go traffic data. Data is collected from stations V1002 and V1003 during stop-and go traffic flow.

A total of 130 vehicles are collected from both V1002 (July 14$^{th}$, 2009) and V1003 (July 16, 2009) stations. Similar to the procedure explained for synchronized flow minimum sample size required for stop-and-go flow is calculated to determine data sufficiency. But considering the complexity of stop-and-go flow the desired level of precision (e) is assumed as 3ft. Thus the sample size required for stop-and-go data for 95% confidence interval and standard deviation of 16.55ft (similar to section 3.2) is calculated as 117 vehicles. Thus collected sample of 130 vehicles is sufficient to validate the model.

From initial observation it is confirmed that vehicles were overestimated in most of the cases. Thus smaller vehicles are detected as medium and in some cases as even large vehicles (similar to synchronized flow, section 3.2). This is further explained using graphical representation of the length estimates from existing model and ground truth data (Figure 23). As we can see from the graph the results are much dispersed about the drawn linear line (y=x), which shows the inaccuracy of the existing model in predicting the vehicle lengths.

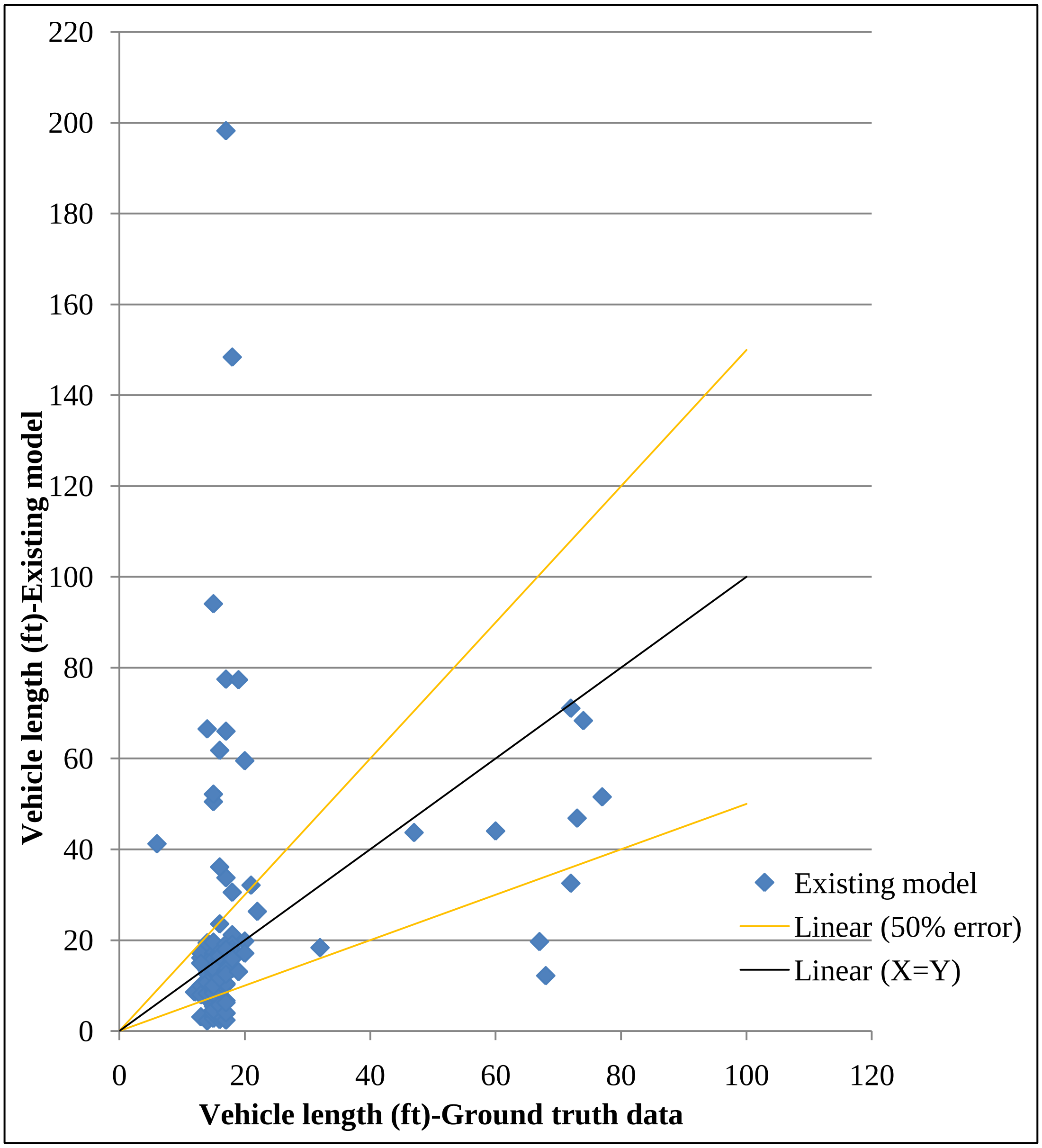


**Figure 23: Existing model vs Ground truth data (stop-and-go flow)**

Similar to synchronized flow (3.2) two straight lines ($y = \frac{3}{2}x$ and $y = \frac{1}{2}x$) representing ±50% error ranges are drawn on either side. The data has few outliers which represent the inaccuracy of existing classification model. For example 17ft vehicle has been detected as 198ft long vehicle. This can be explained as the vehicle was observed to stop on M loop and accelerated and also stopped on S loop. Similarly an 18ft vehicle has been detected as 148ft. In this case the vehicle decelerated and stopped on loop S only. In both previous cases vehicle lengths are overestimated due to factors such as acceleration (deceleration) and stopping patterns. Similarly in another data point a 67ft vehicle is underestimated as 19ft vehicle, in this case vehicle observed acceleration over loop detection area. In another example 68ft vehicle has been detected as 12ft vehicle and this is the case of vehicle stopped on M loop and then accelerated over rest of the detection area. In both the cases vehicle lengths have been underestimated by existing vehicle classification model.

**Table 7: Vehicle classification in to bins from existing model (stop-and-go flow)**

| | | **Ground truth classification** | **Existing classification model** | Misclassification (Existing-Ground) (No. of vehicles) |
|---|---|---|---|---|
| | | No of vehicles | No.of vehicles | |
| **ODOT** | | | | |
| Small | <= 28ft | 121 | 108 | -13 |
| Medium | <=46ft | 1 | 8 | 7 |
| Large | >46 ft | 8 | 14 | 6 |
| **WSDOT** | | | | |
| Bin 1: | <= 26ft | 120 | 106 | -14 |
| Bin 2: | <= 39ft | 1 | 6 | 5 |
| Bin 3: | <= 65ft | 2 | 9 | 7 |
| Bin 4: | > 65ft | 7 | 9 | 2 |
| **Total** | | 130 | 130 | |

Finally the vehicles are classified in to standard bins (Table 7) which compares the results from existing model with that of ground truth data. From the table we can observe that vehicle lengths have been overestimated resulting in more medium, Bin 2 and Bin 3 vehicles. We can observe this by a misclassification of -13 vehicles (Small) and -14 vehicles (Bin 1) which result in overestimation of 7 vehicles for medium class, 5 vehicles for Bin 2 and 7 vehicles for Bin 3. From the results and analysis we can conclude that the existing vehicle classification model fails during stop-and-go traffic condition and needs to be improved.

# CHAPTER 4: CONSTANT ACCELERATION BASED VEHICLE CLASSIFICATION MODEL (CAVC)

Existing model as explained in earlier sections assumes that vehicle runs across loop detectors at constant speed. This model is valid under light traffic (free flow) and results evaluated in chapter 3 show that the existing classification model is not accurate enough during congested traffic (synchronized flow and stop-and-go flow). It is difficult for vehicles to maintain constant speed during congested traffic and the acceleration (deceleration) will distort the results obtained from the existing model.

From analysis of existing model data and observations we can conclude that the acceleration effect on estimated vehicle length is different for different vehicle classes (or type of vehicle). Hence estimation of individual acceleration factors is necessary than proposing a constant acceleration value for all vehicle classes. For example a long vehicle (e.g. 65 ft) experiencing small amount of acceleration, the calculated vehicle length will be much different from the real length. This difference in length is much higher compared to that for a small vehicle with same amount of acceleration (deceleration).

To solve for vehicle estimate and produce accurate vehicle length based classification Kinematic equations are used in this research. We can draw relevant information useful for solving the problem of acceleration (deceleration) during congestion from these equations. The following section thus presents a new vehicle classification model named as Constant Acceleration based Vehicle Classification model (CAVC model). This model is used to solve both cases of congestion which are synchronized and stop-and-go. The principle behind new

CAVC model for the congested traffic flow condition and traffic state identification methodology will be discussed in next sections.

## 4.1 Principle behind CAVC model

As explained in earlier section the new model considers the acceleration factors which influence vehicle motion on roadway (especially during congestion). Figure 24 represents a dual loop detector with direction of traffic flow and upstream loop (M) and downstream loop (S). Where D is the distance between the loops and $t_1$, $t_2$, $t_3$, $t_4$ are time stamps for the vehicle travelling over the loops which is collected as raw loop data.

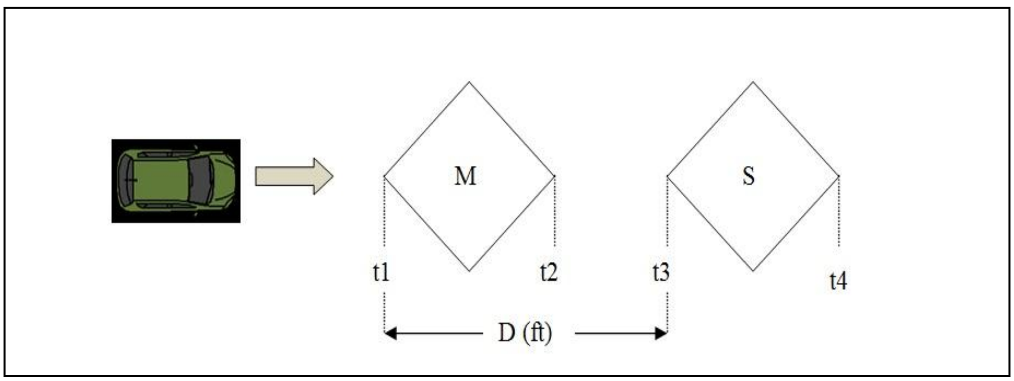


**Figure 24: Layout of dual loop detectors**

Where,

$t_1$ = loop M on-time stamp

$t_2$ = loop M off-time stamp

$t_3$ = loop S on-time stamp

$t_4$ = loop S off-time stamp

As we can see in from Figure 24 for each vehicle four time stamps are recorded. Two of them being on-time stamps and other two are off-time stamps. The front bumper of vehicle which initializes the detection results in on-time stamps over both M and S loops. Similarly the rear bumper results in recoding of off-time stamps while leaving both loops. Thus constitute two trajectories one front bumper and another rear bumper. We can thus only calculate only two speeds for the vehicle using these four different time stamps and their locations (from loop dimensions). Thus accurate determination of vehicle length using these two speeds which assume that vehicle travels with constant speed ignores the impact of acceleration (deceleration) during congestion.

This new model which is named as CAVC model assumes that vehicle runs across the dual-loop detector area at variable speed which is the result of constant acceleration (deceleration).

The following four equations forms the basis for CAVC model:

$$L_s + L_v = v_o \times OnT_1 + \frac{1}{2}a(OnT_1)^2 \tag{9}$$

$$L_s + L_v = v_t \times OnT_2 + \frac{1}{2}a(OnT_2)^2 \tag{10}$$

$$v_t = v_o + at \tag{11}$$

$$D = (v_o \times t) + \frac{1}{2}at^2 \tag{12}$$

Where,

$L_v$: Length of vehicle

$L_s$: Length of single loop; in this study, $L_s$ = 8.5ft (in our case)

$v_o$: Speed of vehicle when it is entering the upstream loop (M loop)

$v_t$: Speed of vehicle when it is entering the downstream loop (S loop)

a: Vehicle acceleration

D: Distance between loops

t: Time taken for vehicle to travel from M loop to S loop = (t3 – t1)

$OnT_1$: Total time for which the vehicle is detected on M loop = (t2-t1)

$OnT_2$: Total time for which the vehicle is detected on M loop = (t4-t3)

We here have four equations and four unknowns (variables). Variables here include ‘$L_v$’, ‘$v_o$’, ‘$v_t$’, and ‘a’. And OnT1, OnT2 and t are derived from raw loop event data (as shown above). On the other hand $L_s$ and D are constants and for our dual loop stations (V1002 and V1003) these values are 8.5 ft and 20 ft respectively. Solving the equations (9, 10, 11 and 12) for $L_v$ (length of vehicle) we derive following equation (model).

$$L_v = \left[OnT_1 \times \left(\left(\frac{D}{t}\right) + \left(\frac{(OnT_1 - OnT_2) \times (OnT_1 - t)}{(OnT_1 + OnT_2)(OnT_2 - OnT_1 + t)}\right)\right)\right] - L_S \tag{13}$$

Sample output of the CAVC model under congested traffic flow is shown in Table 8 and compared with the results from both existing model and ground-truth data. It can be seen that the accuracy of vehicle length calculation has been improved which results in improved vehicle classification.

**Table 8: Sample data of estimated length using different models**

| Vehicle length (ft) (Ground truth) | Vehicle length (ft) (Existing model) | Vehicle length (ft) (CAVC model) | (Existing-Ground) d1 (ft) | (CAVC-Ground) d2 (ft) |
|---|---|---|---|---|
| 16 | 39 | 17 | (39-16) = 23 | (17-16)= 1 |
| 16 | 4 | 14 | -12 | -2 |
| 65 | 54 | 67 | -11 | 2 |
| 15 | 8 | 14 | -7 | -1 |
| 65 | 113 | 77 | 48 | 12 |
| 70 | 85 | 73 | 15 | 3 |
| 65 | 86 | 68 | 21 | 3 |

**Note**: Negative d1 or d2 represent that the vehicle length calculated is less than ground truth data

## 4.2 Identification of traffic flow state

The next step is to classify the data in three traffic flow conditions as to use the appropriate vehicle classification model. To classify the traffic flow as free flow, synchronized and stop-and-go flow traffic flow threshold values are used as explained in chapter 1 (section 1.2.5). According to Coifman (2002) during free flow on-time difference ($OnT_1$-$OnT_2$) would be in the range of ± (3.5/60) sec. From the field observation it is observed that this range is valid and can be used as an indicator for congestion. All the vehicles with in this on-time difference range would fall in to free flow and the existing vehicle classification model can be used. The rest of the vehicles fall into congested traffic condition and are further classified into synchronized flow and stop-and-go flow using speed of vehicle.

As explained earlier (section 1.2.5.2) synchronized flow is characterized with low average vehicle speeds. From discussions of Kerner (1999) we can observe that speeds as low as 18 mph (30 kph) and as high as 45 mph (70 kph) constitute to synchronized flow. Based on the information gained from his [Kerner 1999] observations and other traffic flow models and also from the field data collected, a model is developed to classify the traffic flow as synchronized vehicle whose speed is less than 45 mph and greater than 15 mph (rounded off). Similarly for stop-and-go vehicles the speeds range from 0 mph to 15 mph. To calculate the vehicle speed for classifying the existing model approach is used. Two speeds are proposed to be calculated one using on-time stamps of both M and S loops ($v_f$) (front bumper) and another using off-time stamps of M and S loops ($v_r$) (rear bumper).

**Steps followed for identification of traffic state:**

**Step 1**: Dual loop data is provided as input to CAVC model (Figure 25). Comparison is made between M-loop on-time (OnT1) and N loop on-time (OnT2) and a check is applied on condition (OnT1-OnT2). Check If (OnT1-OnT2) $\geq \pm$ (3.5/60) seconds.

**Step 2**: If NO for previous condition in Step 1 another check for speed is made If ($v_f$ > 45 mph (AND) $v_r$ > 45 mph) is YES traffic flow state is identified as **free flow**. Where $v_f$ and $v_r$ are speeds calculated using on-time stamps (front bumper) and off-time stamps (rear bumper). If the result is NO vehicle is identified to experience congestion (congested traffic).

**Step 3**: If YES for condition in Step 1 the vehicle is classified as to experience congested traffic. Then another condition is applied for speed ($v_f \leq 15$ mph (OR) $v_r \leq 15$ mph). If the result for this if-condition is YES then the vehicle is identified to observe **stop-and-go flow**.

**Step 4**: If NO for condition in Step 3 then the vehicle is again checked for speed ($v_f \leq 45$ mph (OR) $v_r \leq 45$ mph). If the result for this condition is NO then it is a case of loop detection error (eg, Pulse break ups, detector stuck in the 'off' or 'on', etc). These vehicles have thus satisfied condition in Step 1 as a result of above mentioned loop detection errors and are eliminated through this condition in Step 4. If the result for previous condition (Step 4) is YES then vehicle will be identified to observe **synchronized flow**.

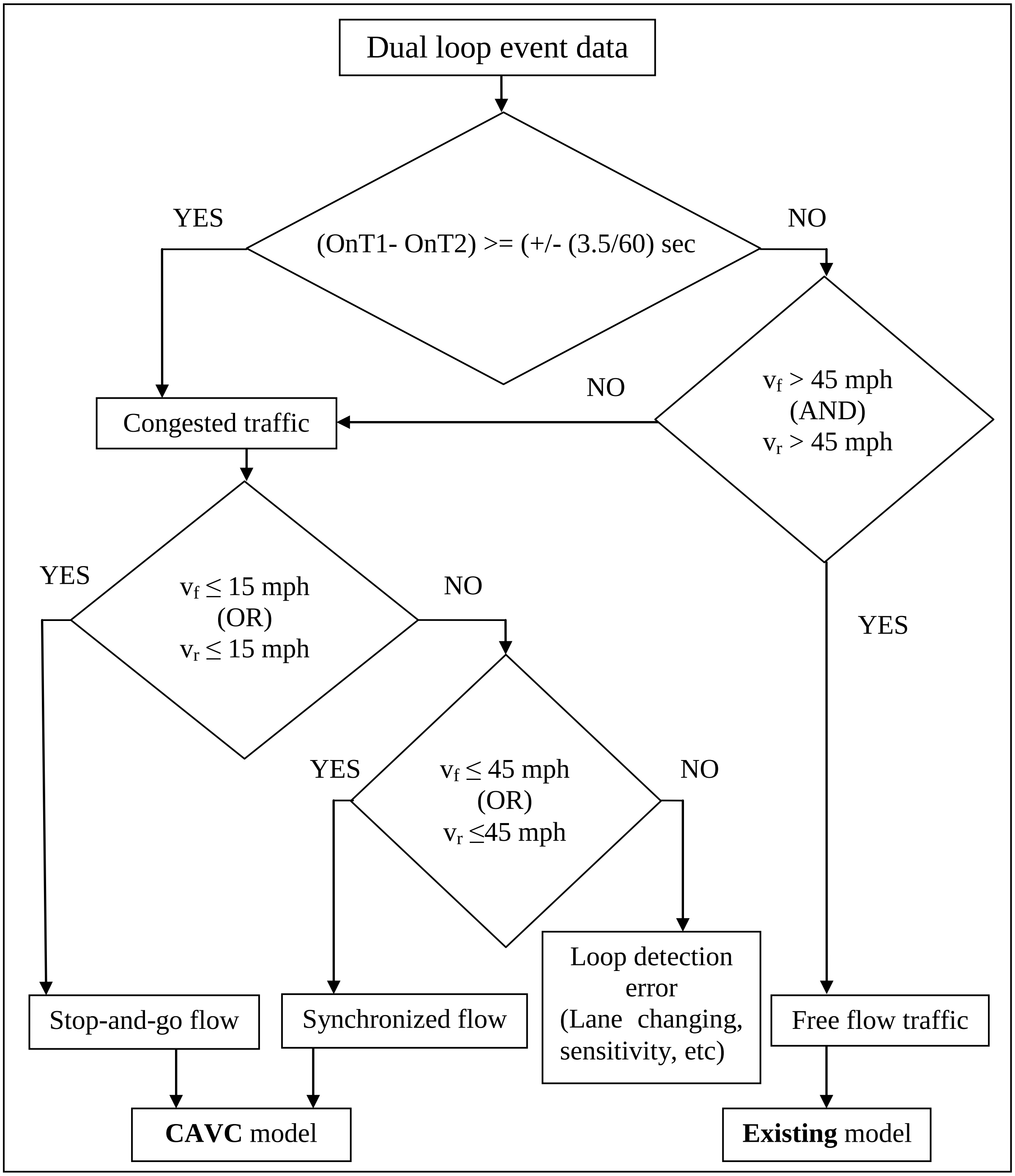


**Figure 25: Traffic flow state identification methodology used in CAVC model**

Traffic flow identification methodology used in CAVC model (Figure 25) thus includes the all previous steps to identify the traffic state. Once the state of traffic flow is identified as free flow or synchronized flow or stop-and-go flow then applicable models are applied. Then the existing model is applied on free flow and CAVC model is applied on both synchronized and stop-and-go flow.

Finally we have established a model for vehicle length calculation using kinematic equations considering the impact of acceleration. According to the CAVC model (Figure 25) the acceleration is assumed to be constant; this assumption has been observed to produce significantly good results during synchronized flow and also for vehicles in stop-and-go flow. In some cases such when vehicle observes multiple accelerations (acceleration is not fairly constant) and also during stopping of vehicles this assumption may not be completely valid. These extreme conditions are further discussed in chapter 5 with examples of complex vehicle trajectories as outliers.

# CHAPTER 5: RESULTS AND DISCUSSION

The Constant Acceleration based Vehicle Classification model (CAVC) is used to improve the vehicle classification and also accurately predict vehicle length. CAVC model in this chapter is evaluated for congested traffic flow condition. The new model is thus used on the corresponding raw loop data to estimate vehicle lengths and vehicle classification. CAVC model is used for both synchronized flow and stop-and-go flow conditions and evaluated using ground truth data. This model calculates acceleration values for individual vehicles and produces length estimates accurately. Vehicles are classified based on calculated vehicle lengths in to standard vehicle classification bins. Results are presented by comparing CAVC model with ground truth data.

## 5.1 CAVC model under synchronized flow

Sample data used for evaluating existing model (Chapter 3.2) is also used for evaluating classification results from CAVC model. For synchronized flow a sample of 414 vehicles have been collected (ground truth data). The results from the new model are evaluated based on classification bins and also using individual vehicle lengths plotted for corresponding ground truth data as in Figure 26. The graph plotted presents the vehicle length data points from both CAVC model and existing model comparing them with corresponding ground truth vehicle length data. Thus this graph displays the accuracy of vehicle length prediction using both old and new models. A linear straight line (y=x) is drawn to represent the accuracy of the results and we can see that most of the data points from new CAVC model fall close to the straight line. And we can also observe some data falling on both sides of the line (y=x) and to represent the percentage error range (upper and lower range) about the true vehicle length values two straight lines representing ±20% error ($y = \frac{6}{5}x$ and $y = \frac{4}{5}x$) are drawn. This is a significant improvement in results as observed for existing model (in section 3.2) for which the data points were beyond the ±40 % error range.

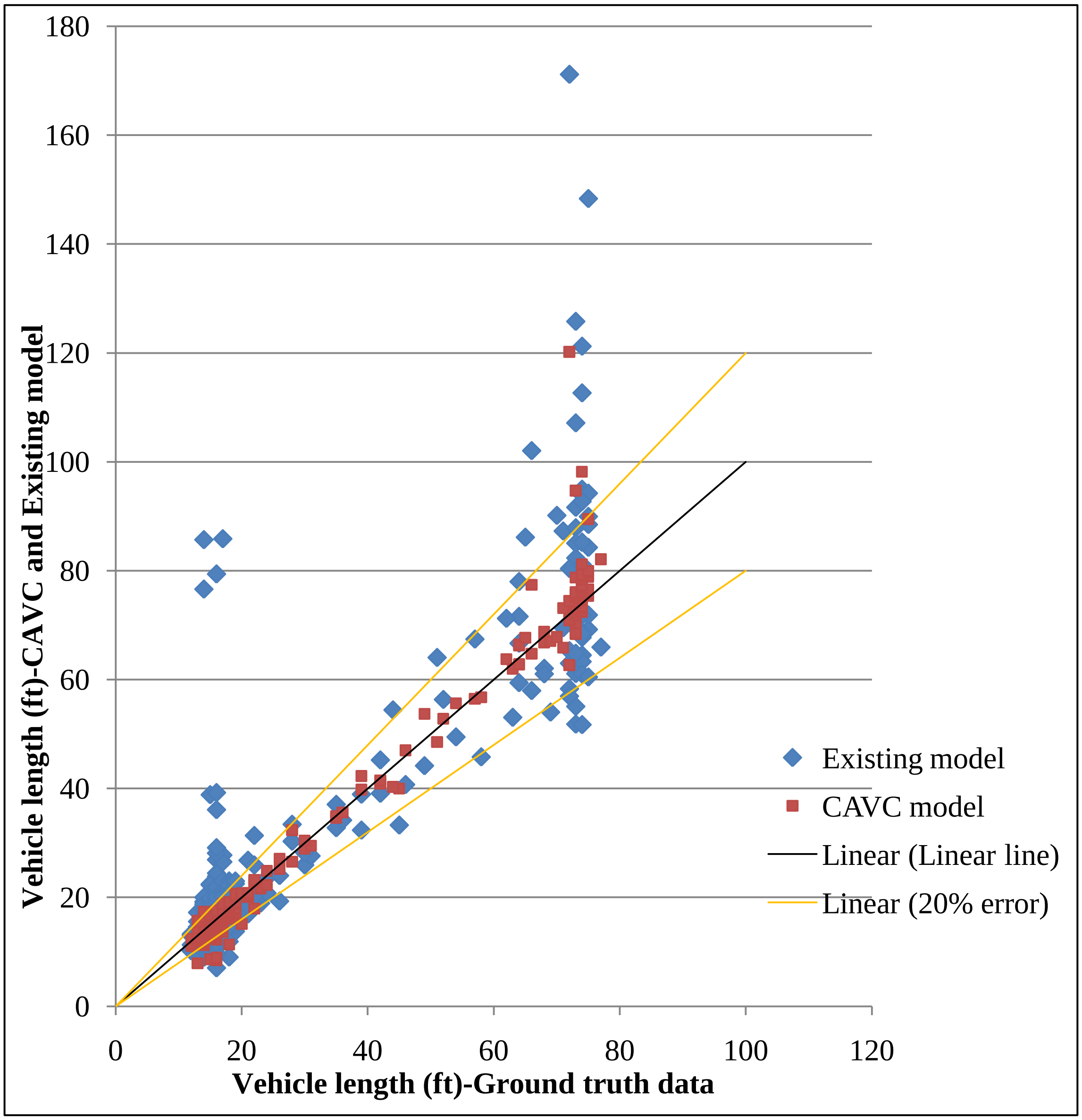


**Figure 26: CAVC and Existing model vs Ground truth data (synchronized flow)**

From the Figure 26 we can observe some outliers which are quite far away from linear straight line (y=x). For example 72ft vehicle has been detected as 120ft, in this case vehicle has been observed to experience multiple decelerations. Similarly in other outlier 73ft vehicle is detected as 94ft and 74ft vehicle is detected as 98 ft. These two vehicles also observed non constant deceleration patterns.

CAVC model as indicated previously is helpful to reduce the amount of error in vehicle length estimate. To statistically validate the results from CAVC model the difference between the vehicle estimate (CAVC model) and actual vehicle length (ground truth) is calculated which forms corresponding sample 'D' as shown in Table 9. This is used to prove that there is no significant difference between the vehicle lengths calculated using the developed CAVC model and ground truth values. Thus all 414 vehicles have been used to calculate this data set 'D' (CAVC-Ground). Theoretically if all vehicles lengths have been exactly estimated the values in D should be equal to zero. Hence the difference in length data (D) is statistically analyzed using standard t-test.

**Table 9: Explanation of sample D (CAVC-Ground) in ft**

| Veh No. | Vehicle length (ft) (Ground) | Vehicle length (ft) (CAVC) | (CAVC-Ground) D (ft) |
|---|---|---|---|
| 1 | 14 | 11.21 | (11.21-14) = -2.79 |
| 2 | 12 | 11.5 | -0.5 |
| 3 | 21 | 20.03 | -0.97 |
| 4 | 46 | 47.01 | 1.01 |
| 5 | 70 | 67.82 | -2.18 |
| 6 | 72 | 74.44 | 2.44 |
| 7 | 75 | 79.80 | 4.8 |

Considering the null hypothesis: “The mean is not significantly different from zero” (H0: $\mu = 0$).

Alternative hypothesis: “The mean is significantly different from zero” (H1: $\mu \neq 0$).

The t value is calculated from equation 13 using the calculated sample mean (X) of the sample D as 0.03938ft, standard deviation ($\sigma$) as 3.49ft, sample size (n) of 414 vehicles and the population mean ($\mu = 0$).

$$t = \left(\frac{(X-\mu)}{\frac{\sigma}{\sqrt{n}}}\right) \tag{13}$$

The calculated t-statistic is 0.22958 (from Equation 13). The t critical value for 95% confidence interval ($\alpha = 0.5$) and df (degrees of freedom) of 413 is 1.973. As the t calculated is less than t critical (0.229<1.973), the null hypothesis cannot be rejected. Thus we can say that the mean of the sample is not significantly different from zero. Therefore there is no significant difference between CAVC length estimate and the ground truth vehicle length.

Finally the vehicle classification in to standard bins is presented in Table 10. From the table we can observe the comparison between CAVC model and ground truth data. Except for medium vehicle class and Bin1class where CAVC model predicts 2 and 3 vehicles less than ground truth data respectively, the rest of classification is accurate enough. We can thus conclude that the classification data from CAVC model for synchronized traffic is very accurate. And CAVC model produced accurate vehicle length estimates and vehicle classification results compared to existing model for synchronized traffic flow.

**Table 10: Vehicle classification in to bins from CAVC model (synchronized flow)**

| | | **Ground truth** | **CAVC model** | Misclassification (CAVC-Ground) (No. of vehicles) |
|---|---|---|---|---|
| | | No of vehicles | No.of vehicles | |
| **ODOT** | | | | |
| Small | <= 28ft | 328 | 329 | (329-328) = 1 |
| Medium | <=46ft | 15 | 13 | -2 |
| Large | >46 ft | 71 | 72 | 1 |
| **WSDOT** | | | | |
| Bin 1: | <= 26 ft | 330 | 327 | -3 |
| Bin 2: | <= 39 ft | 8 | 9 | 1 |
| Bin 3: | <= 65 ft | 18 | 19 | 1 |
| Bin 4: | > 65 ft | 58 | 59 | 1 |
| **Total** | | 414 | 414 | |

## 5.2 CAVC model under stop-and-go flow

CAVC model is then applied for stop-and-go flow data and evaluated similar to existing model (section 3.3). The vehicles in stop-and-go traffic flow travel at very low speeds (section 1.2.5.3) and also stop on loops for certain period of time. Same data sample is taken for evaluating CAVC model for stop-and-go flow as used in for existing model. A total of 130 vehicles are used to compare the results for this new model.

In Figure 27 the graph is plotted which presents the vehicle length data for both CAVC model and existing model comparing them with ground truth vehicle length data. This graph displays the accuracy of vehicle length prediction using both old and new models with reference to ground truth data. A linear straight line (y=x) is drawn to represent the accuracy of the results and we can see that most of the data points from new CAVC model fall close to the straight line.

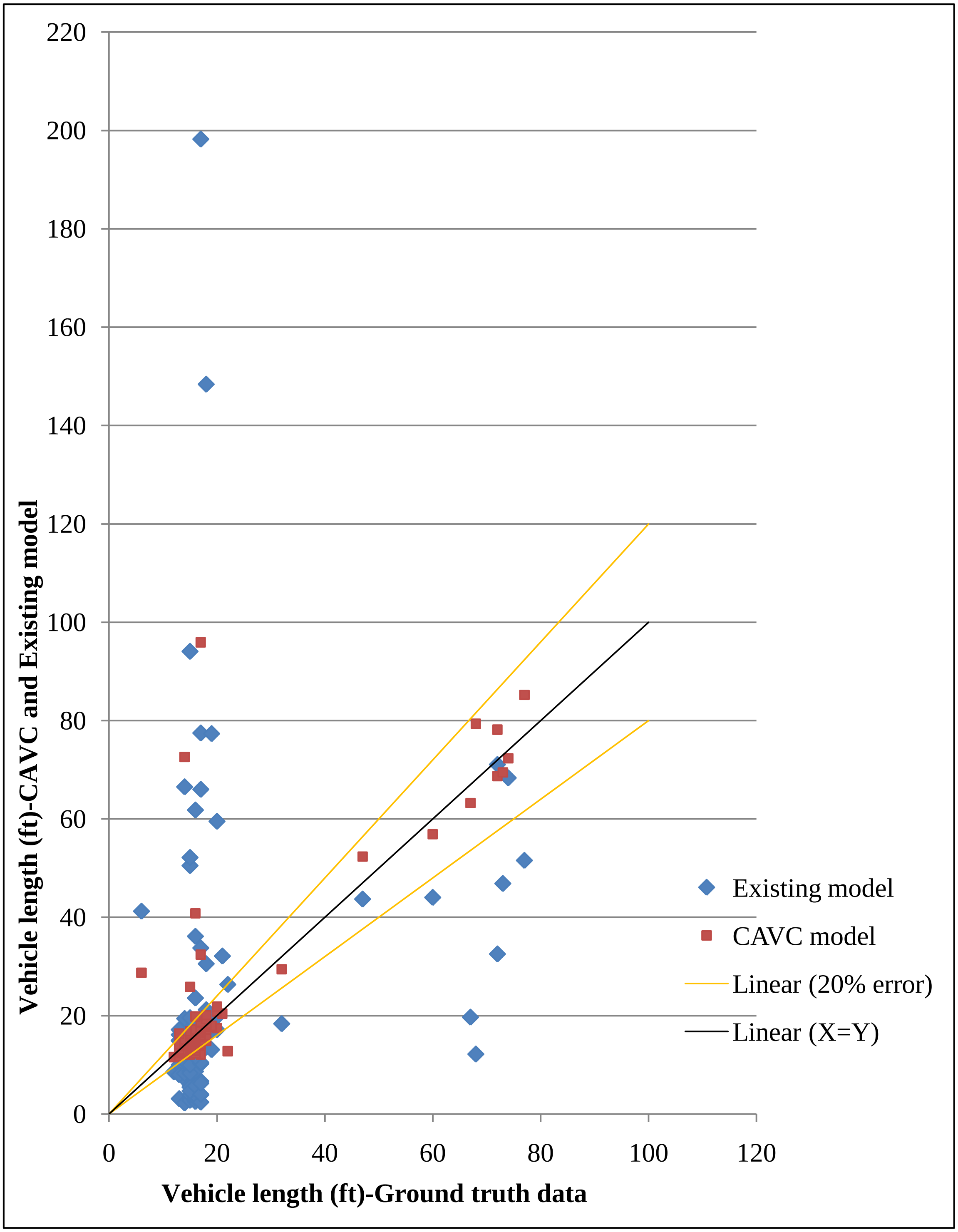


**Figure 27: CAVC and Existing model vs Ground truth data (stop-and-go flow)**

And we can also observe for few data points (CAVC model) fall on both sides of the straight line (y=x). And to represent the percentage error range (upper and lower boundary) about the true vehicle length value, two straight lines representing ±20% error ($y = \frac{6}{5}x$ and $y = \frac{4}{5}x$) are drawn. This is a significant amount of improvement as observed in (section 3.3) for existing model which resulted in error ranges of ±50% and still many data points were beyond the lines

We still can observe few outliers for example 17ft vehicle has been detected as 95ft, thus a small vehicle has been detected as large vehicle. This vehicle stopped on M loop and also on S loop for significant amount of time (18sec and 36sec respectively). This is a case where the vehicle did not observe continues motion (without any stops) on at least one loop. Thus the vehicle length estimated is larger than original vehicle length. Similarly in other examples 14ft, 16ft and 6ft vehicles have been detected as 72ft, 40ft and 28ft respectively. In these cases the vehicles stopped for longer periods of time thus observing multiple accelerating and decelerating behaviors. Vehicle entered M loop and started decelerating and finally stopped in such a way as to activate both loops simultaneously and a similar vehicle trajectory was observed on S loop resulting in overestimation of vehicle length.

Further statistical validation of results from CAVC model for stop-and-go flow is presented to prove that the difference (D) between the vehicle length estimate (CAVC model) and actual vehicle length (ground truth) is not significantly high. A t-test is conducted similar to synchronized flow data (section 5.1) on the sample D of size 130 stop-and-go vehicles. Theoretically if all the vehicle lengths have been exactly estimated the values in D should be equal to zero. The difference in length data (D) is statistically analyzed using t-test.

Null hypothesis: “The mean is not significantly different from zero” (H0: $\mu = 0$).

Alternative hypothesis: “The mean is significantly different from zero” (H1: $\mu \neq 0$).

The mean of the sample D is calculated as 1.3335ft, standard deviation as 9.46ft and sample size of 130 vehicles. The t value is calculated using equation 13 as 1.607. The t critical value for 95% confidence interval and df (degrees of freedom) of 129 is 1.979. Thus comparing t calculated by t critical (1.607<1.979), null hypothesis cannot be rejected. Thus we can say that mean of the sample is not significantly different from zero. Therefore there is no significant difference in vehicle length estimated (CAVC model) and the ground truth data.

**Table 11: Vehicle classification in to bins from CAVC model (stop-and-go data)**

| | | **Ground truth** | **CAVC model** | Misclassification (CAVC-Ground) (No. of vehicles) |
|---|---|---|---|---|
| | | No of vehicles | No.of vehicles | |
| **ODOT** | | | | |
| Small | <= 28ft | 121 | 116 | (115-120) = -5 |
| Medium | <=46ft | 1 | 4 | 3 |
| Large | >46 ft | 8 | 10 | 2 |
| **WSDOT** | | | | |
| Bin 1: | <= 26ft | 120 | 115 | -5 |
| Bin 2: | <= 39ft | 1 | 3 | 2 |
| Bin 3: | <= 65ft | 2 | 4 | 2 |
| Bin 4: | > 65ft | 7 | 8 | 1 |
| **Total** | | 130 | 130 | |

And finally vehicles are classified based on vehicle lengths estimated using CAVC model and compared with ground truth data as shown in Table 11. We can observe that CAVC model produced accurate vehicle classification compared to existing model. For instance in small vehicle bin existing model underestimated 13 vehicles whereas CAVC model reduced it to only 5 vehicles. Similarly for Bin3 vehicle classification underestimates only by 2 vehicles in CAVC model compared to 7 vehicles for existing model.

## 5.3 Conclusions

Existing model for free flow data produced accurate vehicle classification proving that the constant speed assumption is valid for free flow data. For synchronized vehicle data existing model observed error due to misclassification as high as -8 (negative represents underestimating number of vehicles) for small vehicle class and also an error 14 and -15 for Bin3 and Bin1vehicle class. The CAVC model on the other hand was successful in reducing the error to 1, 1, and -3 for small, Bin3 and Bin1 vehicle class respectively. This is significant improvement compared to the misclassification observed earlier in the existing model. On the other hand for stop-and-go vehicle data results show that existing model predicts most of the smaller vehicles as large vehicles. Thus existing model is over estimating most of small vehicles as either medium or large vehicles which is reduced by CAVC model. For example existing model produced an error due to misclassification of -14 (negative represents underestimating number of vehicles) for Bin1, -13 for small and 6 for large vehicle class. Whereas CAVC model was successful in reducing this to -5, -5, and -2 respectively for previously mentioned vehicle classes.

Traffic flow state identification methodology which is presented in this research works by analyzing the traffic flow parameters like on-time difference and speed. This methodology will allow us to switch back and forth between existing vehicle classification model and Constant Acceleration based Vehicle Classification model (CAVC). The vehicle length estimation using CAVC model is very accurate for synchronized flow compared to that in stop-and-go flow. This can be attributed to the stopping pattern of vehicles. It has been observed that CAVC model is accurate for vehicles observing constant motion (without stopping) and also for vehicle stopping once (either on M or S loop). But in other cases such as vehicles stopping on both loops simultaneously and also experiencing multiple stops in the loop detection area creates a complex scenario. More analysis of acceleration patterns observed during these stop conditions are required and their impact on vehicle length estimation has to be analyzed. Also more advanced data collection methods are recommended (GPS, etc.) to accurately study these acceleration patterns, stopping behaviors and headway distribution during stop-and-go traffic flow. Finally concluding that the proposed CAVC model can be used for improving vehicle length based classification.

## 5.4 Contribution to the area of research

This research has significant contribution to the area of vehicle classification data collection using dual loop detector data. As mentioned previously freeway vehicle classification data is collected mostly by using dual loop detectors as they are widely installed in the roadway network. This research unlike in any other previous studies concentrates in improving length based vehicle classification data during congestion, by introducing the acceleration factor for more accurate vehicle length estimation. Constant speed assumption for existing vehicle length based classification is a big source of error. This not only affects stop-and-go data but also synchronized vehicle classification data.

The main purpose of this research is not only quantifying the error due to existing model but also propose a new improved model (CAVC model) for accurate vehicle length estimate. This new model thus can be used for synchronized flow vehicles and also for non-stopping vehicles during stop-and-go flow. This new model considers application of acceleration factors on more vehicle by vehicle approach instead of specifying a range of acceleration values. Further the model incorporates a traffic identification methodology to determine whether the vehicle is observing free flow or synchronized flow or stop-and-go flow. This model uses two traffic flow parameters on-times difference and speed to identify the traffic state.

As CAVC model calculates vehicle acceleration for each individual vehicle depending on timestamps recorded it eliminates any error in acceleration prediction due to seasonal, location based and vehicle class based variance. Each vehicle is treated as a new vehicle and a distinct and more accurate acceleration value is calculated which yields more representative vehicle length based classification.

# REFERENCES

1. Bhagat, V., and Woods, D. (1997). "Loop Detector Crosstalk." *ITE Journal*, 67(2).
2. Bonsall, P., and Bell, M. (1987). *Information technology applications in transport.* VNU Science Press BV, The Netherlands.
3. Booth, C. J., and Kurpis, G. (1993). "The new IEEE standard dictionary of electrical and electronics terms [including abstracts of all current IEEE standards]." *New York: IEEE.*
4. Cheevarunothai, P., Wang, Y., and Nihan, N. L. (2006). "Identification and Correction of Dual-Loop Sensitivity Problems." *Transportation Research Record: Journal of the Transportation Research Board*, 1945(-1), 73-81.
5. Chen, L., and May, A. (1987). "Traffic detector errors and diagnostics." *Transportation Research Record*, 1132 82–93.
6. Coifman, B. (1999). "Using dual loop speed traps to identify detector errors." *Transportation Research Record: Journal of the Transportation Research Board*, 1683(-1), 47-58.
7. Coifman, B. (2001). "Improved velocity estimation using single loop detectors." *Transportation Research Part A: Policy and Practice*, 35(10), 863-880.
8. Coifman, B. (2003). "Identifying the onset of congestion rapidly with existing traffic detectors." *Transportation Research Part A: Policy and Practice*, 37(3), 277-291.
9. Coifman, B., and Yang, Y. (2004). "Estimating spatial measures of roadway network usage from remotely sensed data." *Transportation Research Record: Journal of the Transportation Research Board*, 1870(-1), 133-138.
10. Daganzo, C. F. (1997). *Fundamentals of transportation and traffic operations.* Pergamon.

11. Guide, T. M., and Counts, S. D. (2001). "TRAFFIC MONITORING GUIDE.".

12. Habib-Mattar, C., Polus, A., and Cohen, M. A. (2009). "Analysis of the Breakdown Process on Congested Freeways." *Transportation Research Record: Journal of the Transportation Research Board,* 2124(-1), 58-66.

13. Helbing, A. (2002). "Micro-and macro-simulation of freeway traffic." *Mathematical and Computer Modelling,* 35(5-6), 517-547.

14. Kates, R., Bogenberger, K., and Hoops, M. (1998). "Mesoscopic simulation with ANIMAL: Optimal utilization of downstream traffic detector data and the propagation of information." *Traffic and Granular Flow'97: Gerhard-Mercato-Universität Duisburg, Germany, 6-8 October 1997,* 453.

15. Kell, J. H., Fullerton, I. J., and Mills, M. (1990). "Traffic detector handbook." *Federal Highway Administration, Washington, DC*.

16. Kerner, B. S., and Konhäuser, P. (1994). "Structure and parameters of clusters in traffic flow." *Physical Review E,* 50(1), 54-83.

17. Kerner, B. S. (1998). "Experimental features of self-organization in traffic flow." *Phys.Rev.Lett.,* 81(17), 3797-3800.

18. Kerner, B. S. (1999). "The physics of traffic." *Physics World,* 8(99), 25.

19. Kerner, B. S. (2000). "Theory of breakdown phenomenon at highway bottlenecks." *Transportation Research Record: Journal of the Transportation Research Board,* 1710(-1), 136-144.

20. Kerner, B. S. (2001). "Complexity of synchronized flow and related problems for basic assumptions of traffic flow theories." *Networks and Spatial Economics,* 1(1), 35-76.

21. Kerner, B. S., and Klenov, S. L. (2005). "Deterministic microscopic three-phase traffic flow models." *Journal of Physics A: Mathematical and General,* 39 1775.

22. Klein, L. A. (2001). *Sensor technologies and data requirements for ITS.* Artech House Publishers.

23. Klein, L. A., Mills, M. K., and Gibson, D. R. P. (2006). *Traffic Detector Handbook: - Volume I.*

24. Kockelman, K. M. (2001). "Modeling traffic's flow-density relation: Accommodation of multiple flow regimes and traveler types." *Transportation,* 28(4), 363-374.

25. Lighthill, M., and Whitham, G. (1955). "On kinematic waves. II. A theory of traffic flow on long crowded roads." *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences,* 229(1178), 317-345.

26. May, A. D., Cayford, R., Coifman, B., and Merritt, G. (2003). "Loop detector data collection and travel time measurement in the Berkeley highway laboratory."

27. Nelson, P. (2000). "Synchronized traffic flow from a modified Lighthill-Whitman model." *Physical Review E,* 61(6), 6052-6055.

28. Neubert, L., Santen, L., Schadschneider, A., and Schreckenberg, M. (1999). "Single-vehicle data of highway traffic: A statistical analysis." *Physical Review E,* 60(6), 6480-6490.

29. Nihan, N. L., Zhang, X., Wang, Y., and Briglia, P. (2002). "Evaluation of dual-loop data accuracy using video ground truth data." *Washington, US.*

30. Nihan, N. L., Zhang, X., and Wang, Y. (2005). "Improved dual-loop detection system for collecting real-time truck data." *Transportation Research Record: Journal of the Transportation Research Board*, 1917(-1), 108-115.

31. Oman, R. M., and Oman, D. M. (1997). *How to solve physics problems.* Schaum's Outline Series.

32. Payne, H. J. (1971). "Models of freeway traffic and control." *Mathematical Models of Public Systems*, 1(1), 51-61.

33. Payne, H. J. (1979). "A critical review of a macroscopic freeway model." *Engineering Foundation Conference on Research Directions in Computer Control of Urban Traffic Systems*, 251-265.

34. Polus, A., and Pollatschek, M. A. (2002). "Stochastic nature of freeway capacity and its estimation." *Canadian Journal of Civil Engineering*, 29(6), 842-852.

35. Richards, P. I. (1956). "Shock waves on the highway." *Operations Research*, 42-51.

36. Sin, J., Lee, S. H., and Choi, K. (2002). "A single loop detector-based freeway speed estimation." *KSCE Journal of Civil Engineering*, 6(1), 1-9.

37. Viti, F., Verbeke, W., and Tampère, C. M. J. (2008). "Sensor Locations for Reliable Travel Time Prediction and Dynamic Management of Traffic Networks." *Transportation Research Record: Journal of the Transportation Research Board*, 2049(-1), 103-110.

38. Wang, Y., and Nihan, N. L. (2004). "Dynamic Estimation of Freeway Large-Truck Volumes Based on Single-Loop Measurements." *Journal of Intelligent Transportation Systems*, 8(3), 133-141.

39. Wei, H., Feng, C., Meyer, E., and Lee, J. (2005). "Video-capture-based approach to extract multiple vehicular trajectory data for traffic modeling." *Journal of Transportation Engineering.* 131 496.

40. Wei, H. (2008). "Characterize Dynamic Dilemma Zone and Minimize its Effect at Signalized Intersections." Ohio.

41. "Weigh-in-Motion (WIM) systems." accessed on 8/31/2010, http://training.ce.washington.edu/wsdot/Modules/04_design_parameters/wim.htm 2010).

42. Zhang, H. M., and Shen, W. (2009). "A Numerical Investigation of Stop-and-go Traffic Patterns Upstream of a Freeway Lane-drop." *88th Transportation Research Board Annual Meeting,* Washington, DC.