8-31-2017

Estimating vehicles emissions at signalized intersections in the highway capacity manual

Yuanyuan Fan
New Jersey Institute of Technology

ABSTRACT

ESTIMATING VEHICLES EMISSIONS AT SIGNALIZED INTERSECTIONS IN THE HIGHWAY CAPACITY MANUAL

by
Yuanyuan Fan

Over the past decades, motor vehicle volumes have continued to increase at a high rate. As a result, engineers in the transportation field not only need more robust knowledge of traffic operation control and transportation planning, but more attention is also needed to understand and estimate the influences that this increasing volume of vehicles has on the environment, especially the influence on air quality. The EPA has stated that reducing carbon monoxide (CO) from vehicle emissions is the most significant way to control air pollution from the transportation sector.

The Highway Capacity Manual is a national and international resource that has become a guideline for evaluating the operation of roadway, transit and pedestrian facilities. The Highway Capacity Manual assesses the operation of a roadway based on the perception of its users. Performance measures are used to describe the traffic operation of the roadway. At present, no measures are provided to describe the operation of the roadway based on environmental impacts. The incorporation of air pollution estimation into the Highway Capacity Manual will allow the roadway's operation to be assessed both from an operational and environmental aspect, ultimately creating a sustainable development for both transportation and the environment.

The objective of this dissertation is to develop MOVES-like estimation models of vehicle emissions for pollutants at a signalized intersection that can be incorporated into the Highway Capacity Manual. "EPA's Motor Vehicle Emission Simulator (MOVES) is a state-of-the-art emission modeling system that estimates emissions for mobile sources at the national, county, and project level for criteria air pollutants, greenhouse gases, and air toxics." (EPA, 2014). A thorough understanding is needed about what parameters, and influence of these parameters on vehicle emissions. This dissertation develops two kinds of models in order to estimate emissions caused by

on-road vehicles. Two modeling approaches are used to estimate four kinds of emissions including CO, NO, NH3 and NOX separately. The following summarizes the work of this dissertation:

(1) Two modeling approaches are used to estimate vehicle emissions including: multiple linear regression and Artificial Neural Network (ANN). In the multiple linear regression modeling, two different models were developed including one model using operation modes as independent variables and another model using traffic related parameters as independent variables. Both model approaches and independent variables are used to estimate four types of pollutant emissions. Statistically, the emission models using traffic parameters as independent HCM related parameters are capable of providing a better emissions estimate based on the higher R square value. For CO, the variables found to be significant were volume to capacity ratio and grade with an R^2 of 61.56%. For NO, the variables found to be significant were volume to capacity ratio and grade with an R^2 of 99.47%. For NOx, the variables found to be significant were volume to capacity ratio and grade with an R^2 of 99.47%. For NH3, the variables found to be significant were volume to capacity ratio and grade with an R^2 of 99.25%. This study shows that volume to capacity dominate the emissions quality at a signalized intersection. The research found that for NOx, Idling and Moderate Speed Coasting were significant. For NH3, all variables were significant except Low Speed Coasting. For CO, Braking and Cruise/Acceleration were significant. It was also found that longer delay time reduces CO emissions, but it causes the other three pollutant emissions increase.

(2) The ANN modeling method using the Levenberg-Marquardt method was used to train the HCM related variables and MOVES emissions outputs. The parameters of volume to capacity ratio, and road grade are used to estimate emissions. The Validated R value of the obtained ANN model is found.

ESTIMATING VEHICLES EMISSIONS AT SIGNALIZED INTERSECTIONS IN THE HIGHWAY CAPACITY MANUAL

by
Yuanyuan Fan

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
In Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Transportation Engineering

John A. Reif, Jr. Department of Civil and Environmental Engineering

August 2017

.

APPROVAL PAGE

ESTIMATING VEHICLES' EMISSIONS AT SIGNALIZED INTERSECTION ANALYSIS OF THE HIGHWAY CAPACITY MANUAL

Yuanyuan Fan

Dr. Janice Daniel, Dissertation Advisor Date
Associate Professor of Civil and Environmental Engineering, NJIT

Dr. I Jy Steven Chien, Committee Member Date
Professor of Civil and Environmental Engineering, NJIT

Dr. Jo Young Lee, Committee Member Date
Assistant Professor of Civil and Environmental Engineering, NJIT

Dr. Bladikas, Athanassios, Committee Member Date
Professor of Industrial Manufactural Engineering, NJIT

Dr. Wen Zhang, Committee Member Date
Associate Professor of Civil and Environmental Engineering, NJ

BIOGRAPHICAL SKETCH

Author: Yuanyuan Fan

Degree: Doctor of Philosophy

Date: August 2017

Undergraduate and Graduate Education:

- Doctor of Philosophy in Transportation Engineering,
  New Jersey Institute of Technology, Newark, NJ, 2017

- Master of Science in Environmental Science,
  New Jersey Institute of Technology, Newark, NJ, 2012

- Bachelor of Science in Civil Engineering,
  Iowa State University, Ames, IA, 2010

Major: Transportation

Presentations and Publications:

Yuanyuan Fan and J. Daniel. Significance Analysis of Vehicle Operation Modes for Pollutant Emission at a Signalized Intersection. 5th International Conference on Sustainable Energy and Environment Engineering (CSEEE2016), Zhuhai, China 2016 Dec (pp511-515)

Yuanyuan Fan, Janice R. Daniel. Predicting Vehicles emissions at signalized intersection by Using Artificial Neural Network. The 9th International Workshop on Computational Transportation science,CTS 2017 July 14-15, 2017 Lanzhou, China

Jianjun Wang, Yuanyuan Fan, Yunlong Li, Song Miao, A Real-time Monitoring Method of Vehicle Pollutants' Emissions Based on LiDAR and Photogrammetry Technology Chinese Patent No. 2016111310729, Applied Date 12/09/2016

Yuanyuan Fan, Janice R. Daniel, Estimate CO Emissions at Signalized Intersection by Using Artificial Neural Network Method, Journal of Intelligent Transportation System (Submitted)

Yuanyuan Fan, Daniel Janice. Vehicle Emissions as a Performance Measure at Signalized Intersections in the Highway Capacity Manual. 97th Transportation Research Annual Meeting (Submitted)

Yuanyuan Fan, Janice R. Daniel. Validation of Vehicles Operation Modes from VISSIM and Modeling Development at a Signalized Intersection. (In preparation)

Yuanyuan Fan, Daniel Janice. Estimating Vehicle Emissions and Significance Analysis of impact factors in the Signalized Intersection. (In preparation)

To My Beloved God Jesus Christ – for His Great Mercy and Salvation

ACKNOWLEDGEMENT

I want to give thanks and gratitude to my dissertation advisor, Dr. Janice Daniel –for her intelligence, critical thinking and kindness. Her continuous support encourages me to pursue my Ph.D. degree. I would also like to thank my dissertation committee members: Dr. Steven Chien, Dr. Bladikas, Athanassios, Dr. Jo Young Lee and Dr. Wen Zhang–for their participation in my committee.

Thanks for the funding support from New Jersey Institute of Technology. They include project from Safety Bell Usage, ITS Resource Center, and Dynamic Message Signs. Thanks to my fellow graduate students who walked with me at this journey, they are Dr. Liuhui Zhao, Dr. Albert Ford, Dr. Wei Hao, Wanyi Fu, Liran Chen and Zijia Zhong.

I want to thanks my brothers and sisters in Newark Christian Fellowship–for their prayers for me in Christ. Thanks to my parents, Xiaocai Fan and Guifang Si, who always love me and support me in every situation. I thank my sister and brother, Baimei Si and Silong Fan –for their sibling love.

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

Chapter **Page**

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF TABLES
(Continued)

Table **Page**

LIST OF FIGURES

LIST OF FIGURES
(Continued)

CHAPTER 1

INTRODUCTION

1.1 Background and Problem Statement

The US Environmental Protection Agency (EPA) regulates six principal pollutants to ensure air quality. These pollutants include: carbon monoxide (CO), lead (Pb), nitrogen dioxide (NO_2), ozone (O_3), particulate matters (PM) and sulfur dioxide (SO2). Among these pollutants, about two-thirds of the total emissions of CO pollutant comes from the transportation sector. It has been shown that CO emissions from vehicles are as high as 90% of the total amount of total urban CO emission (EPA, 1993). The EPA has stated that reducing carbon monoxide (CO) from vehicle emissions is the most significant way to control air pollution from the transportation sector.

The Highway Capacity Manual (HCM) provides guidance that serves both traffic planners and traffic engineers in the planning, designing and operating of transportation facilities. Although the HCM is a tool recommended by the EPA to predict vehicles' speeds in the estimation of emissions (HCM, 2010), the HCM does not include air quality in determining the performance of transportation facilities. The HCM 2010 makes some references about air quality stating "vehicle emissions are a significant contributor to poor air quality", and referring to the Clean Air Act Amendments CAAA (HCM, 2010). Although the report "*Extent of Highway Capacity Manual Use in Planning*" (Dowling, 2012) expects air impact analysis to be included ultimately in the HCM, current HCM users do not have strategies in the HCM to estimate air quality based on the design or operation of the transportation facility.

Several vehicle emissions estimation tools have been developed in recent years. The second generation EPA vehicle emissions model is called Motor Vehicle Emission Models (MOVES). It was developed by EPA's Office of Transportation and Air Quality (OTAQ) in 2002 and first released to the public in 2010. "EPA's Motor Vehicle Emission Simulator (MOVES) is a state-of-the-art emission modeling system that estimates emissions for mobile sources at the national, county, and project level for criteria air pollutants, greenhouse gases, and air toxics." (EPA, 2014). MOVES serves both the state and local agencies and it is capable of estimating 59 pollutants from 13 vehicle classes, five source types and five road types.

Another emissions research tool that has been used in research is the Comprehensive Modal Emissions Model (CMEM). This emission tool was developed by the University of California-Riverside in 1995. CMEM is a module model that divides the whole estimation system into six parts, including power demand, engine speed, air to fuel ratio, fuel rate, engine-out emissions, and catalyst pass fraction.

Emission estimations tools such as MOVES and CMEM often need vehicle activity data, including information on the operation mode of vehicles to complete the emission procedure. Research has shown that emission results are sensitive to the operation modes of vehicles (LeBlanc, 1995; Barth et al. 1997; Frey, H. Christopher, et al. 2002; Ritner, Mark, et al, 2013). Critical to the accuracy of the emissions estimation is how the vehicle activity data is collected. The existing research on vehicle emissions models discusses three types of approaches for collecting traffic activity data for emissions study. These approaches include: using simulation software (Abou-Senna et al., 2013; Chen et at. 2016); using lab experiment data (Djoric et al. 2014); and using sensors to measure field data

(Jimenez-Palacios, 1998). There are distinct advantages and disadvantages to using different sources of data. On-board emission technology is capable of collecting field emission data, but the cost is relatively high and only limited data is obtained at one time. Remote sensing technology is also difficult to collect multiple vehicle activity data at one time. Individual vehicle activity data or limited vehicle activity data is difficult to use in describing air quality performance influenced by vehicles' operation modes. On the contrary, the cost of simulation-based approach is relatively low. In addition, the functionality of the simulation software enables generating simultaneous estimation of emissions by importing the simulated vehicle trajectory data. For this reason, vehicle activity data from simulation software becomes the first choice for data collection in this dissertation.

Current emission models use a variety of parameters to model emissions' output. For example, Abou-Senna and Radwan (2013) used vehicle speed as a variable to estimate emissions; Zhang et al. (2013) applied acceleration as a variable to find the emission outputs; Andrew et al. (2012) and Rouphail et al. (2000) used control delay percentage and time to predict emissions. Coelho, Farias and Rouphail (2005) used speed control signal as important predictors; Shabinhkhani and Gonzals (2013) utilized operation modes, time, stop and volume as predictors for emissions output study. Akcelik et al. (2003) used instantaneous speed and acceleration rate to estimate emissions. Li et al. (2011) used signal timing and delay as predictors to find the emissions outputs. These studies indicate that vehicle emissions can be estimated using various predictors. At the same time, traffic engineers who are unfamiliar with these air pollution estimation tools find it difficult to conduct environmental analyses due to the complex vehicle emissions models. These traffic engineers are familiar with the HCM and may be able to use the HCM to estimate vehicle emissions. The HCM has a variety of performance measures that can be used not only for estimating the operational performance of the roadway, but also can be used in estimating vehicle emissions. This dissertation focuses on providing traffic engineers with tools that can be used for estimating vehicle emissions.

Many agencies in United States use the Highway Capacity Manual (HCM) in design and operation of their transportation facilities. Absent from the HCM, however, is the ability to estimate the performance of a roadway based on its impact on air quality. Although much has been studied on vehicle emissions and several models exist, none of the estimation methods have been incorporated into the HCM. This may be due to the fact that current emission estimation tools are not easily utilized by traffic engineers. Performance measures developed through the HCM methodology can be used in the development of a simplified vehicle emissions model that can be incorporated into the HCM. A reliable emission estimation methodology using traffic engineering parameters as inputs would be useful in realizing the goal of incorporating vehicle emissions into the HCM.

Some of the current emissions models have problems of utilizing complicated input files, which can be time consuming to generate and then to use in the emissions process. To overcome this, simplified models can be developed which provide a baseline analysis of vehicle emissions for use in evaluating the emissions performance of the roadway facility.

Many vehicle emissions models use speed as an input variable. Speed profile is one of the most significant inputs for the project level analysis in MOVES. In fact, either average speed or vehicle second-by-second speed can be applied. The project level has three ways of importing the speed profile, they are: average speed, drive schedule and operation mode distribution. The operation mode is the description of the vehicle running states, such as Braking, Idling, Low speed coasting, Cruise/acceleration, and Moderate speed coasting. The operation modes distribution method is reported to be the most accurate input method for estimating emissions, but this type of input is a complicated input method (EPA MOVES, 2010) and requires a large quantity of vehicle activity data to be collected and formatted before being imported into MOVES. This has, therefore, decreased the use of this method. Instead, multiple agencies currently use average speed instead of the operation mode input, although this can cause biases in the emissions estimation and possibly impact the accuracy of the emissions estimation. Therefore, an easily used emission model which does not sacrifice prediction accuracy is pursued.

Previous research has been performed to develop simplified emission estimation methods. Ozbay et al. (2012) developed a simplified vehicle emissions model based on MOVES outputs. In this study vehicle speed is used as an input variable and the Fourier series is used to fit the model. However, the manual used for EPA MOVES claims that vehicle speed cannot fully capture the relationship between emissions and the influence factors of vehicles' running states (EPA MOVES, 2004). In addition, studies have shown that the vehicle's operation mode can have a significant impact on the emissions results (LeBlanc, 1995; Barth et al. 1997; Frey H. Christopher, et al. 2002; Ritner Mark, et al. 2013). For this reason, it is necessary to develop a model that includes these operation modes. Another study by Stanek et al. (2013) generated linear, polynomial, and logarithmic models as the relationship between emissions and volume of vehicles. Vehicle volume was the only variable in that research. All of these studies utilize a relationship between emission outputs and a single influencing factor, such as speed, acceleration, volume. In addition, there are other potential variables that can be used to estimate vehicle emissions, including operation mode. At the same time, estimation accuracy should also be considered and evaluated in these models.

1.2 Research Objectives

The primary objective of this research is to develop a MOVES-like estimation model for use in estimating vehicle emissions of pollutants at signalized intersections. These models can be used as a performance measure for assessing roadway performance in the HCM. Pollutant emissions from on-road vehicles will be used as the dependent variable of interest. Vehicle operation modes will be used as independent variables as well as other traffic related variables. The relationship between the independent and dependent variables will be analyzed, and the best fitting model will be found to predict pollutant emissions for four pollutants: CO, NO, NOx, and NH_3. The result will assist HCM users in the design and assessment of roadways using both operational performance measures, as well as through the use of vehicle emissions estimates. The following tasks are to be accomplished in order to achieve the objectives.

Task 1: Validate the reliability of VISSIM results used for estimating vehicle operation mode

Task 1-1: Measure real-world vehicle speed and acceleration data on a second-by-second basis at an intersection.

Task 1-2: Simulate the intersection in VISSIM under the same conditions for the real-world conditions,

Task 1-3: Compare the field data of operation modes with the simulation data.

Task 2: Develop a vehicle emissions model by using multiple linear regression method

Task 2-1: Optimize the signal timing for various demand volumes at the intersection. The purpose of this step is to minimize the control delay.

Task 2-2: Simulate the intersection in VISSIM and generate second-by-second activity data for different demand volume conditions.

Task 2-3: Use MOVES to estimate emissions for four pollutants.

Task 2-4: Develop a vehicle emissions model relating vehicles' operation models.

Task 2-5: Develop vehicle emissions prediction models with HCM parameters.

Task 3: Develop vehicle emissions prediction models by using artificial neural network model (ANN)

Task 3-1: Use artificial neural network model (ANN) to develop an emissions model.

Task 3-2: Develop ANN Graphic User Interface model (GUI) to users.

CHAPTER 2

LITERATURE REVIEW

This chapter provides an overview of previous studies on vehicle emissions. Section 2.1 introduces the methodology assessing the automobile mode in studying urban street segments in the HCM. Section 2.2 introduces the method to calculating delay at intersections in HCM 2010. Section 2.3 introduces the pollutants the EPA mainly regulates. Section 2.4 introduces tools currently used to estimate vehicle emissions. The last section of this chapter is a literature review of possible influencing factors and existing methodologies on tailpipe level emissions.

2.1 Automobile Mode in Urban Street Segments in HCM 2010

This subsection of the literature review describes the methodology used in HCM 2010 for assessing the performance of the automobile mode in urban street segments. Understanding the operational behavior of the automobile mode in urban street segments becomes an important step for generating vehicle emissions that will be used in the second stage of developing the emissions model. The section begins with a discussion of urban street segments including some vital inputs for key procedures. Second, a discussion is provided on the used to determine the delay.

Input data for the automobile methodology consists of traffic characteristics and geometric design. Specifically, the input data elements for the boundary intersection include: demand flow rate; number of lanes upstream intersection width, and turn bay length; through control delay, through stopped vehicles, 2nd and 3rd-term back-of-queue size, and capacity. A summary of input data and their further notes presents in Table 2.1.

A thorough review for a signalized intersection measurement methodology will be introduced in next subsection.

Table 2.1 Input Data for Boundary Intersection in Urban Street Segment

Traffic Characteristic	Geometric Design			Performance Measurement		
Demand flow	Number of lanes upstream	Upstream intersection width	Turn bay length	Through control delay	Through stopped vehicles,2nd and 3rd-term back-of-queue size	Capacity
Hourly flow rate (veh/hr)	Through lanes, shared lanes, Exclusive turn lanes	Distance between the stops (two way)	Length of the bay that queued vehicles stored	HCM2010 Chapter 18	Used to calculate stop rate at signalized intersections	Need vehicles that entering upstream and leaving the downstream

Source: HCM (2010)

2.2 Intersection Performance in HCM

The methodology used to determine the operational performance of a signalized intersection includes evaluating four modes of travel at the intersection. These modes are automobile mode, pedestrian mode, bicycle mode and transit mode. Performance measures cited in HCM includes four aspects. They are volume-to-capacity ratio, automobile delay, queue storage ratio, and delay. This research especially focuses on automobile mode and delay calculation. Delay, the primary performance measure, is found form step eight. Figure 2.1 exhibit these descriptions. Step one of the auto mode methodology is to determine the movement group and lane group. The two groups are different only when a shared lane is presented. The second step is to determine the movement of group flow rate.

The third step determines the lane group flow rate. The fourth step is to calculate the adjusted saturation flow rate. Step 4: Determine Adjusted Saturation Flow Rate.

$$s = s_0 f_w f_{HV} f_g f_p f_{bb} f_a f_{LU} f_{LT} f_{RT} f_{Lpb} f_{Rpb} \tag{2.1}$$

Where s=adjusted saturation flow rate (veh/h/ln)
S_0= base saturation flow rate (pc/h/ln)
f_w=adjusted factor for lane width,
f_{HV}=adjustment factor for heavy vehicles in traffic stream,
f_g=adjustment factor for approaching grade,
f_p=adjustment factor for existence of a parking lane and parking activity adjacent to lane group,
f_{bb}= adjustment factor for blocking effect of local buses that stop within intersection area,
f_a= adjustment factor for area type,
f_{LU}= adjustment factor for lane utilization,
f_{LT}= adjustment factor for left-turn vehicle presence in a lane group,
f_{Lpb}= pedestrian adjustment factor for left-turn group, and
f_{Rpb}= pedestrian-bicycle adjustment factor for right-turn groups.

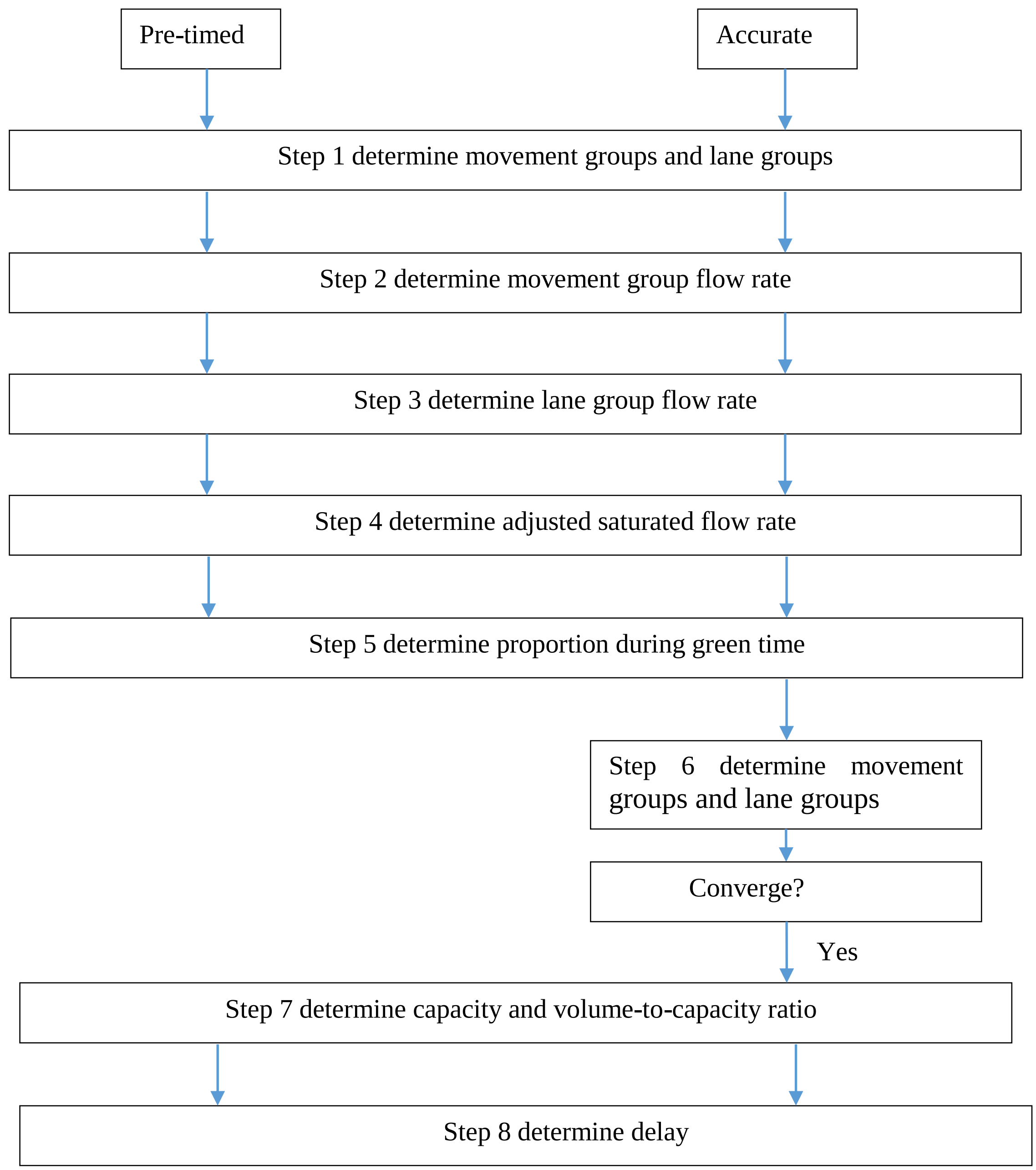


Figure 2.1 Methodology flowchart of delay (HCM 2010).

Step 5 is to determine the proportion of vehicles arriving during green. Step 6 of the HCM is to determine the phase duration. Step 7 is to determine capacity and volume to capacity ratio. The proportion of vehicles arriving during the green impacts the control delay and queue size. It is found that shorter delay and queue size when there is a larger proportion of vehicles arriving at the green time. The following equation represents proportion of vehicles arriving at each lane group.

$$P = R_p\left(\frac{g}{C}\right) \tag{2.2}$$

Step 6 is to determine signal phase duration, this procedure is suitable for those phase duration is not known. The time of an actuated phase is defined as:

$$D_p = l_1 + g_s + g_e + Y + R_c \tag{2.3}$$

Where

D_p is the phase duration (s)
l_1= start up lost time=2 secs.
g_s= queue service time,
g_e = green extension time,
Y= yellow change interval,
R_c = red clearance interval.

The effective green time for the phase is computed with the equation:

$$g = D_p - l_1 - l_2 = g_s + g_e + e \tag{2.4}$$

The effective green time in this type of duration unknown control is represented as:

$$G = D_p - L_1 - L_2 = g_s + g_e + e \tag{2.5}$$

Where

L_1 = Clearance lost time,

e= extensive lost time, 2 sec.

Step 7 is to determine capacity and volume-to-capacity ratio.

$$C = Ns\frac{g}{C} \tag{2.6}$$

Where C is cycle length (s).

Volume capacity ratio is defined as:

$$X = \frac{v}{c} \tag{2.7}$$

Where
v is demand flow rate (veh/hr),
c is capacity (veh/hr).
X = volume- to-capacity ratio.
And then find critical intersection Volume-to-Capacity Ratio

Step 8 is to determine the delay. The delay means vehicles' average delay during the study period. These delays include uniform delay, incremental delay and initial queue delay. The control delay equation is represented as:

$$d = d_1 + d_2 + d_3 \tag{2.8}$$

Where
d = control delay(s/veh),
d_1 =uniform delay (s/veh),
d_2 = incremental delay (s/veh), and
d_3 = initial queue delay (s/veh).

Queue size at the end of interval i represented as:

$$Q = Q_{i-1} - \left(s/3600 - q/N\right)\ t_{d,i} >= 0.0 \tag{2.9}$$

Where
Q=queue size at the end of interval i (veh)
q= arrival flow rate=v/3600 (s/veh).
$T_{d,i}$= duration of time interval ii during the arriving flow rate and saturation flow rate are constant (s).

To calculate uniform delay,

$$d_1 = \frac{0.5C[1 - g/\mathrm{C}]^2}{[1 - g/\mathrm{C}, \min(\mathrm{X}, 1.0)]} \tag{2.10}$$

To calculate the incremental delay,

$$d_2 = 900T[(X_A - 1) + \sqrt{(X_A - 1)^2 + \frac{8kIX_A}{c_A T}} \tag{2.11}$$

With

$$X_A = \frac{v}{c_A} \tag{2.12}$$

Where

C = cycle length

X= v/c ratio

k = incremental delay factor;

I = upstream filtering/metering adjustment factor;

T = duration of analysis period (h).

2.3 EPA Regulations and MOVES

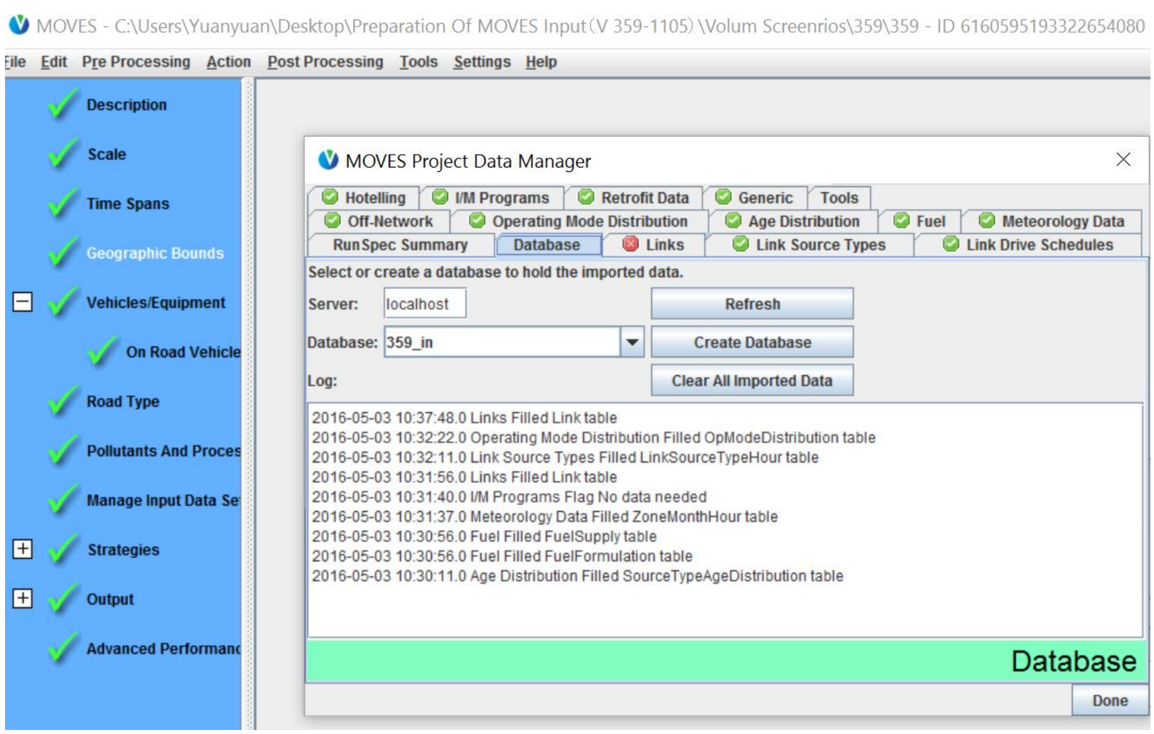
The Clean Air Act Amendment (CAAA) requires the Environmental Protection Agency (EPA) to set National Ambient Air Quality Standards (NAAQS). EPA addresses the most recent updated NAAQS for the six principal pollutants. The pollutants are carbon monoxide, lead, nitrogen dioxide, ozone, particle pollutants and sulfur dioxide. The pollutants are set as primary and secondary standards separately. Primary pollutants are mainly for protecting people's health and the secondary standard set for protecting public welfare. Pollutants mainly caused by the mobile source are listed in Table 2.2 and they include carbon monoxide, ozone, nitrogen dioxide, and particle pollution.

Table 2.2 Six Principal Ambient Air Pollutants Standards

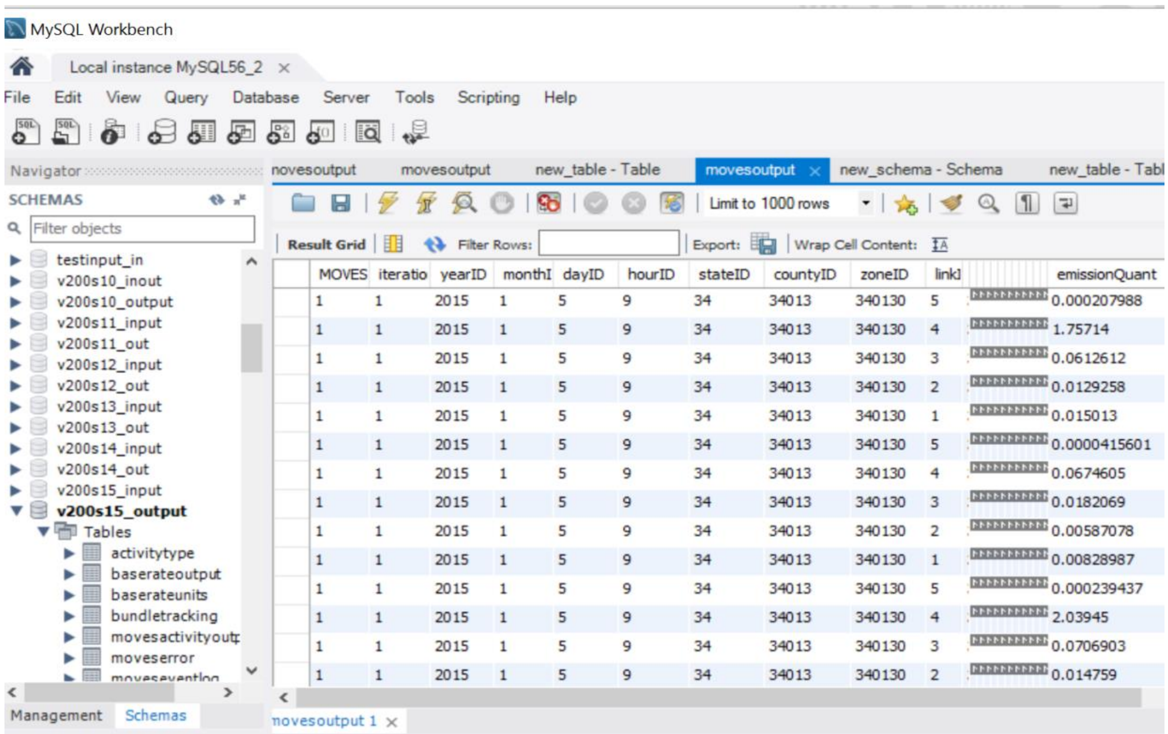
Pollutants		Primary/Secondary	Averaging Time	Level	From
Carbon Monoxide		Primary	8 hour	9ppm	Not to be exceeded more than per year
			1hour	35ppm	
		Primary and secondary	3 months Avg.	0.15 µg/m^3	Not to be exceeded
Nitrogen Dioxide		Primary	1 hour	100 ppb	98th percentile, average over three years
		Primary and secondary	Annual	53 ppb	Annual Mean
Ozone		Primary and secondary	8hour	0.075ppm	Average over 3 years
Particle Pollution	PM2.5	Primary	Annual	12µg/m^3	Annual Mean average 3 years
		Secondary	Annual	15 µg/m^3	Annual Mean average 3 years
		Primary and secondary	24 hour	35 µg/m^3	98th percentile, average over three years
	PM10	Primary and secondary	24 hour	150 µg/m^3	Not to be exceeded more than once per year

Source: EPA, 2012

Motor Vehicle Emission Simulator (MOVES) was released by EPA in 2010 to estimate on-road vehicle emissions. Project level analysis is used for roadway intersections, highways, transit projects, parking lots and intermodal terminals. MOVES input file includes “link, metrology, link source type, fuel supply, fuel formulation, source type age distribution, off-network link” which is required to import to Project Data Manager. Among all input database: metrology, I/M (inspection and maintenance), fuel, and fuel formulation can be exported using default data for the location and time that selected in study. Thirteen types of vehicles in Vehicle/Equipment are available to make combinations with fuel type for selected vehicles, such as gasoline-passenger car or diesel fuel-school bus. MOVES has five types of road types. These road types include off-network, rural restrained access, rural unrestricted access, urban restricted access, urban unrestricted access. “Urban Unrestricted Access” is selected from road types for this case study because this proposal is focus on signalized intersection emissions study in urban area. An output file name is then set up and units chosen for all outputs. The pollutants chosen to be studied are CO, NH_3. NO, and NO_X. These pollutants were chosen because these pollutants are MOVES’s outputs allowed to estimate in once. In addition, gram is selected for weight unit, joules for energy use and mile for distance. The following snapshot exhibits the output from MySQL Browser that MOVES used to provide emission outputs.



(a)



(b)

Figure 2.2 MOVES input and output interfaces. (a) MOVES graphic user interface (GUI); (b) MySQL output interface.

The methodology that MOVES uses to calculate emissions is based on the Vehicle Specific Power (VSP), which is a function dependent on grade, speed and acceleration. The function used in EPA MOVES is represented as:

$$VSP\ (kW/ton) = v * \left[1.1a + 9.81a\ \left(tan\left(sin\left(grade\right)\right)\right) + 0.132\right] + 0.000302\ v^3 \qquad (2.13)$$

Where:

a = Vehicle acceleration (mph/sec)
v = Vehicle speed (mph)

2.4 Comprehensive Modal Emissions Model (CMEM)

Comprehensive Modal Emissions Model is another widely used software that has been used to simulate vehicles emissions. This model was first introduced by University of California-Riverside in 1995. CMEM has been funded and supported by EPA from 1999. The model structure of CMEM is shown in Fig. 2.3.

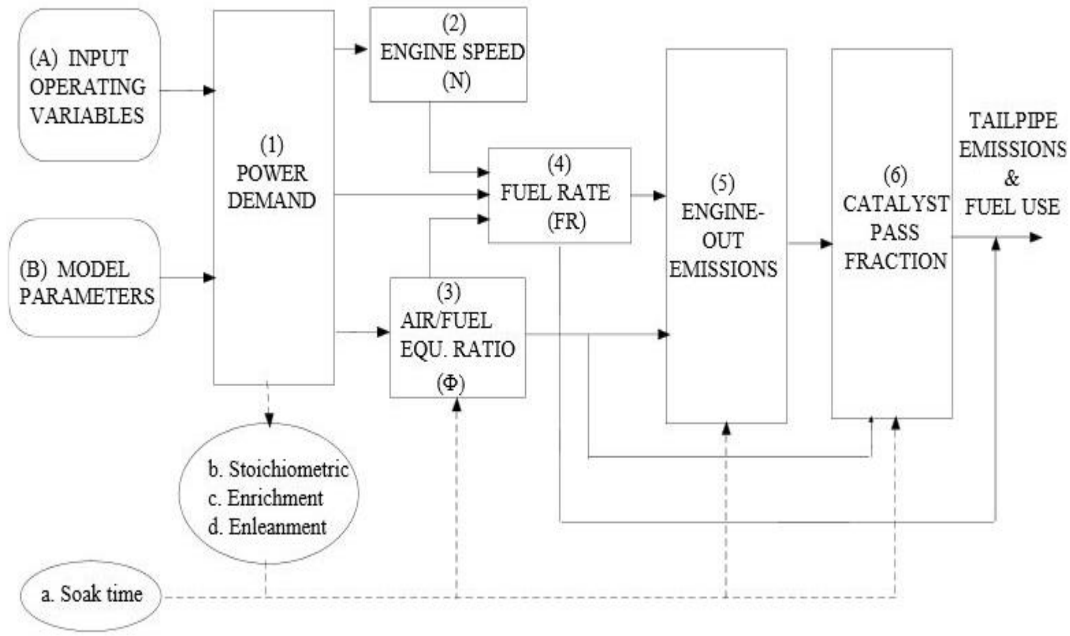


Figure 2.3 CMEM model structure (CMEM manual 2006)

From Figure 2.3 we can know, the tailpipe emission can be expressed by the following:

$$TailpipeEmissions = FR * \frac{g_{emission}}{g_{fuel}} * CPF \tag{2.14}$$

Where, FR is the fuel use rate in grams per second; CPF is the catalyst pass fraction. It is defined as ratio of tailpipe to engine emission. CPF is a function of fuel to air ratio and engine out emissions.

In Figure 2.3, there are total six modules. Each module has its model to be presented in mathematics form. In our literature review, it is necessary to mention some key concepts, as for the detailed description and illustration, they can be found from the CMEM manual. First, Engine Power Demand Module is presented as following:

$$p = (\mathrm{M}\cdot \mathrm{a} + M \cdot G \cdot \sin\theta) + \frac{1}{2} Cd \cdot A\rho v + M \cdot \mathrm{g} \cdot \mathrm{Cr} \cdot \cos\theta) \cdot \mathrm{v}/1000 \tag{2.15}$$

Where

M =mass,

V=speed m/sec
A=acceleration m/s^2,
g=gravity constant,
Θ= grade angel,
Cd is the coefficient of the rolling resistance
A is the frontal surface area (m^2)
ρ = air density (kg/m^3)
Cr =coefficient of the rolling resistance.

Engine speed module is shown as:

$$N(t) = \mathrm{S} \cdot \frac{R(\mathrm{L})}{R(\mathrm{Lg})} \cdot v(\mathrm{t}) \tag{2.16}$$

Where
$N(t)$ =engine speed at time t,
S= engine speed/ vehicle speed ratio in top gear Lg (rpm/mph)
$R(\mathrm{L})$ =gear ratio in L$^{\mathrm{th}}$ gear,
$v(\mathrm{t})$ = the vehicle speed at time t,
Fuel rate module is presented as:

$$\begin{aligned} &\mathrm{FR} \approx (\mathrm{k} \cdot N \cdot V + \frac{P}{\eta}) \frac{1}{43.2} \cdot (1 + b1 \cdot (\mathrm{N} - \mathrm{N}_0)^2) \\ &K = K_0 \cdot (1 + \mathrm{C} \cdot (\mathrm{N} - \mathrm{N}_0)) \\ &\mathrm{N}_0 \approx 30 \cdot \sqrt{\frac{3.0}{V}} \end{aligned} \tag{2.17}$$

Where
FR=fuel use rate in grams/sec,
P= engine power output in KW,
K = the engine fraction factor,
N = engine speed,
V= engine displacement (liter),
η = efficiency of engine which is 0.45,
$b1$ =10^{-4}
c $\approx$ 0.00125
And Engine-out Emission Module is represented as:

$$E_i = a_i \cdot FR + r_i \tag{2.18}$$

Where E_i is the engine-out emission rate in g/s,
a_i, r_i =emission index coefficients of pollutant i.

2.5 Previous Research

There are numerous studies estimating and predicting vehicle emissions at intersections in recent years. The most prominent factors associated with vehicles emissions in the intersection system can be classified into four categories. These categories include: traffic related, vehicles mode and year related, traffic control strategies related, and others factors such as season and weather. The data-compiling method used in these researches is also different to each other: some traffic data is acquired in a project-level study, and then emissions under the proposed traffic scenario can be calculated. The data used to analyze vehicle emission at an intersection have several collection sources in some studies the traffic data is obtained by collectors who stand at the intersection and use devices to capture the measurement. And in some studies the traffic data is collected through GPS (Zhang et al., 2013). In other studies the traffic data is obtained by using the traffic simulation software create virtual traffic data. Then the traffic data is imported into the emission simulation software such as MOVES to be further analyzed by a statistical method.

2.5.1 Traffic Condition in a Study Period

The existing research for the traffic conditions is mainly related to the following parameters, including: vehicle volume at the intersection, speed, number of stops, queue length and delay time.

Yu et al. (1998) have applied remote sensing to study vehicle emission produced per time duration in Huston area. They found a vehicle's emissions are related to vehicle's instantaneous activity profile. In other words, emissions output is a function of instantaneous speed and acceleration. The authors found that initial indicators of engine load are vehicle's velocity and operating mode. Li Jie et al. (2012) indicated that by optimizing traffic control can reduce the vehicles emission. They used the image processing method to obtain the real vehicles trajectory to calibrate the driving behavior parameters in VISSIM, to identify and adjust the most influential parameters, so as to ensure the correctness of the simulation outputs results.

Stanek and Breiland (2013) estimated vehicles emissions for two lanes roadway by using the Synchro and SimTraffic simulation software. They developed a multivariate regression model to estimate the vehicles pollutants emission per day. The authors considered daily volume, percentage of major volume, turning traffic and peak hour as the influencing factors. Statistical R Square reached to 0.856 and 0.921 for all way stop and roundabout respectively.

Abou-Senna and Radwan (2013) studied the relationship of vehicle emission and speed. They found that with the increasing of speed from 20mph to 80mph, CO_2 emission first decreased and then increased, and at the speed of 60mph, there is a minimum emission appeared. In their study, VISSIM and MOVES were utilized to generate the traffic data and the corresponding emissions, respectively. The results were then analyzed using stepwise regression. In this research, Abou-Senna and Radwan also estimated the CO_2 emission at the condition of 0% grade and non-truck condition. In addition, Abou-Senna and Radwan found the relationship of vehicle flow rate and CO_2 emission. Emission linearly increases as flow rate increase from 1000vph to 7000vph.

Zhang et al. (2013) established a model to evaluate emissions due to acceleration at an intersection. Field data were collected at an intersection during the peak hour. GPS data were also collected from passenger cars about 2500ft away from the intersection. Acceleration models were built to calculate the instantaneous speed.

Stanek and Breiland (2013) estimated Green House Gas (GHG) emission at two lane roadway intersections. The concept of this research is that fuel and emissions are convertible since carbon amounts can be measured both in polluted emissions and in fuel.

Thus, the research uses fuel consumption to present emission based on demand volume. The authors sought to find out how different control types would affect vehicle emission and the fuel consumption. The analysis assumes an intersection of 2500 feet with four legs. There are fifty-four groups of variables are designed as inputs for the simulation model to obtain emission outputs, of which each group of variables include intersection daily volume, volumes of major and minor streets, and turning traffic percentages. By the Synchro and SimTraffic, the experimental data are obtained. Based on the obtained data, three forms of regression models are used to best fit these data in order to find a quick estimation of the emissions. These regression forms include linear, polynomial, and logarithmic respectively. The others parameters assumed are Peak Hour Factor (PHF), conflicting pedestrians and bicyclists, heavy Vehicle percentage, passenger car distribution between light passenger car and SUV, speed limit, and single lane approaches. The experimental parameters include the split of major street volume, the direction volume, and the turning percentages. Multivariable regression method was used to analyze the relationships between fuel consumption and key variables. The study found that among these variables, the vehicle volume is the most significant variable for all the regressions models. It also shows major and minor street splits were not prominent for all models. Based on the R-square value, the polynomial and logarithmic equations shows a better fit because R square values are closer to 1. Polynomial regression contains volume square as an independent variable. The form of the polynomial model is presented as:

$$Y = b_o + b_1 x_1 + b_2 x_2 + \ldots b_n x_n \tag{2.19}$$

Where

b_o to b0 to b_n are regression coefficients;

x_1 to x_n are intersection volume properties.

The logarithm model contains Logarithm of fuel consumption value as the dependent variable and is presented as follows:

$$Log(y) = b_o + b_1 x_1 + b_2 x_2 + \ldots b_n x_n \qquad (2.20)$$

Andrew et al. (2012) studied emissions at congested and uncongested intersections using MOVES2010. In order to measure the amount of nitrogen oxides and PM emission at a signalized intersection, traffic intersection scenarios using levels of service from B to E were established to estimate the emission. The study found that emission is more sensitive to control delay during congested conditions. Cruising and acceleration can cause more than 80% of the total emissions and idling can cause 18% of emissions. In the second step, the methodology used to calculate the time spent in each mode was split into two parts. In the first part, percentage of control delay at an intersection was determined by considering queue length, number of lanes, vehicle spacing in queue, vehicle flow and cycle length. In the second part, the time spent at each activity was calculated by considering variables including percentage of control delay, vehicle flow, and the maximum speed. In third step, multiply emission factors (EFs) was determined and would change with the corresponding time in each mode to obtain the total emission amount as grams per hour. The methodology for studying EFs is based on MOVES. With MOVES each vehicle's amount of emission per second is obtained.

Similarly, Coelho, Farias, and Rouphail (2005) used the time-in-mode concept when they studied the impact of speed control signals on pollutant emissions. The fixed emission rate, which is the same as the EFs prepared for the total emission estimation model, and the differences between with and without control devices are expressed as

$$DE = EF_i t_i + EF_a t_a + EF_d t_d - EF_c t_c \qquad (2.21)$$

Where a is acceleration, i is idling, d means deceleration and c means cruising. EFi means emission factor in idling, t_i is the time duration in the idling. EF_a is the emission factor in the acceleration mode, t_a is time spent on acceleration, EF_d is emission factor during deceleration, td is the time spend on deceleration, EF_c is emission factor during cruising, and t_c is the time spent on cruising. The common elements of these papers are that they all used different vehicle deceleration, acceleration, idling time, and their emission rate to calculate the total emission.

Shabinhkhani and Gonzales (2013) presented an analytical model of vehicle emissions at a signalized intersection at a microscopic level. MOVES is used to obtain emission factors for different driving modes, and then the total intersection emission was estimated with an analytical traffic model. The model takes into consideration stops in a signal cycle, idling time, and cruising time, and this model is developed by kinematic wave theory. MOVES is used on a project level to study emission at the microscopic level. In that research, the author assumed some parameters were fixed in the input files.

Some studies use traffic simulation such as VISSIM, and Next Generation SIMulation (NGSIM) which provide second-by-second vehicle trajectory (reference). After the driving modes are obtained from the vehicles, emissions can then be calculated. In a research Shabinhkhani and Gonzales selected 126 data records from 1000 traffic trajectories ready to use by smoothing, filling the gaps and selecting the acceleration, as well as a decelerating event. Emissions are then calculated by MOVES, each vehicle’s stop is input as a link file. With regards to traffic, the authors designed different traffic volume and signal control time for intersections to obtain emission factors (EF) at intersections. To create an analytical model, kinematic wave theory was used in estimating the traffic state. The number of vehicles is calculated, then time of idling and cruising are modeled. Then, the emission model is calibrated in terms of speed and acceleration. The calibrated simulation is applied to validate the model. An isolated intersection is simulated depending on different demands. Stops per vehicle and the time spent on each mode are obtained from the simulation. Finally, the emission estimate is determined. This paper pointed out that decreasing the number of stops from approaching vehicles is significant to reducing the emissions at an intersection.

Akcelik et al. (2003) developed the emission model using SIDRA and MOTION. In that research, instantaneous speed and acceleration rate are generated by MOTION. SIDRA is used to simulate the drive cycle, which included cruising, acceleration, deceleration and idling. Fuel consumption in each mode are then added together to get the total emission amount. The model for estimating the emissions is represented as:

$$\Delta F = \begin{cases} \left\{ \alpha + \beta_1 R_T v + \left[\beta_2 M_v a^2 v / 1000 \right]_{a>0} \right\} \Delta t & R_T > 0 \\ \alpha \ \Delta t & R_T <= 0 \end{cases} \tag{2.22}$$

Where R_T=total tractive force, α, β_2 are parameters, v is velocity, a is acceleration rate. And F is emissions value.

Kyoungho et al. (2002) developed a method to evaluate environmental impact for transportation planning purposes network wide. The research involved the vehicles participating in the test are five light-duty vehicles and three light-duty trucks. The authors acquired the experimental data from Oak Ridge National Laboratory (ORNL) including fuel consumption and emissions rates. The experimental data includes second-by-second speed, instantaneous acceleration and measure of effectiveness (MOE) which is presented below. They are measured and collected from the selected vehicles and each of them is measured 1300 to 1600 times. With the data available, the proposed model is developed based on the relationship between tractive effort and other variables that Post et al. (1981)

first revealed. Many combinations of speed and acceleration are used in the derivation of models. The final model includes three forms of variables of speed and acceleration. As shown below:

$$MOEe=\sum_{i=0}^{3}\sum_{j=0}^{3}(Ki,j*si*aj) \tag{2.23}$$

$$ln(\,MOE)e=\sum_{i=0}^{3}\sum_{j=0}^{3}(Ki,j*si*aj) \tag{2.24}$$

$$ln(MOE)_e=\begin{cases}\sum_{i=0}^{3}\sum_{j=0}^{3}(Li,j*si*aj) & for\ a>=0\\ \sum_{i=0}^{3}\sum_{j=0}^{3}(Li,j*si*aj) & for\ a<=0\end{cases} \tag{2.25}$$

Where

(MOE)e= Instantaneous emission rate

Kei,j= Model regression coefficient for MOE "e" at speed power "i" and acceleration power "j"

Lei,j=Model regression coefficient for MOE "e" at speed power "i" and acceleration power "j" for positive accelerations

Mei,j=Model regression coefficient for MOE "e" at speed power "i" and acceleration power "j" for negative accelerations

s=Instantaneous Speed (km/h)

a=Instantaneous acceleration (m/s2)

The first equation shows a third degree polynomial combination to estimate instantaneous emission rate. The equation can include most ORNL data except a few negative dependent variables. Equation 2.24 is used by a data transformation technique to transfer results in the first equation. Equation 2.25 is used to express positive and negative acceleration separately in regression model. The real-world data is used to validate the proposed model. The field data included EPA measurement at automotive Testing laboratory in Ohio and EPA's national vehicle and fuels emission lab in Michigan in 1997. The result shows a good fit in the model compared with the field data. As a conclusion, the model developed by Kyoungho et al. (2002) revealed the relationship between speed and acceleration. Traffic and drive related factors are especially emphasized in the model and they are represented by the instantaneous speed and acceleration. But this model does not consider the heavy-duty vehicle emissions, which plays a significant constitution of road emission source.

Yao et al. (2013) has brought video camera to emissions estimation. The light of their study mainly shows to obtain vehicles activities data from the video camera. Then speed file are imported into MOVES to estimate the emissions. Even there maybe still exist error from camera calibration or image warping (Yao et al., 2013). This method compared to the many traditional methods has improved the vehicles' operation mode input so that accuracy of emissions results output is said to be improved.

2.5.2 Vehicles Conditions

The presence of trucks in the traffic flow can greatly affect CO2 emissions (Abou-Senna and Radwan, 2013). In addition, vehicle conditions including vehicle types, ages, weight, mode and the year the vehicle was produce can also impact tailpipe emissions. Poor maintenance can worsen emissions in all types of vehicles. When the vehicle speed is at 45 mph, CO2 emissions increase from 0.6 kg/veh-mi at 0% of truck to 1.2 kg/veh-mi at 15% of trucks involved. Ozguven el al. (2013) estimated emissions based on various vehicle types with the methodology based on MOVES. The study applies an approximated emission function to estimate the emission using MOVES output. Thirteen vehicle types were included in the analysis. These vehicles are assigned to run in MOVES with speeds ranging between2 mph to 80 mph, with an increasing quantity of 5 mph each time to obtain different emissions. The relationship between emission and speeds were then determined. An eight order Fourier series function was investigated to best fit the output emissions from

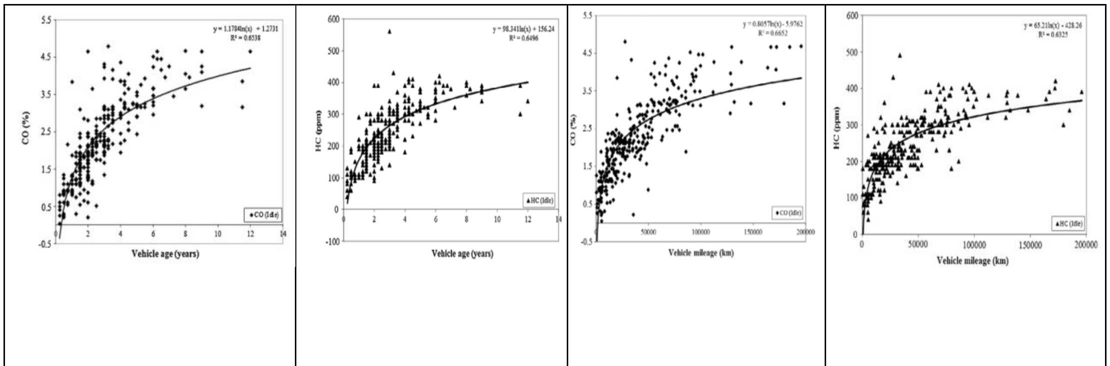
MOVES. The function is related to the roadway link speed, and also 16 coefficients are used to weight the velocity series, as shown in Equation 2.24. In addition to the vehicles in motion, idling is also considered in the development of the function. The vehicle type and pollutant level at a speed of zero is analyzed by using MOVES in project-level carbon monoxide analysis. Then the formulated Fourier function and idling estimation parameters are input into Assist-Me or Advanced Software for Statewide Integrated Sustainable Transportation System Monitoring and Evaluation. This software is used for visualizing and analyzing results of the transportation planning model. The Fourier function is represented as:

$$\begin{aligned} &Emission\ level = a0 + a1*cos(w*V) + b1*sin(w*V) + a2*cos(2w*V) \\ &+b2*sin(2w*V) + a3*cos(3w*V) + b3*sin(3w*V) + a4*cos(4w*V) \\ &+b4*sin(4w*V) + a5*cos(5w*V) + b5*sin(5w*V) + a3*cos(6w*V) \\ &+b6*sin(6w*V) + a7*cos(7w*V) + b7*sin(7w*V) + a8*cos(8w*V) \\ &+b8*sin(8w*V) \end{aligned} \qquad (2.26)$$

Where a, b are function coefficients; V is link speed, w is the basic angular frequency of emission fluctuation, and the unit of emission level is grams/veh-hour.

Rouphail et al. (2000) explored the field observed emissions at the signalized arterials. Data in this research is obtained from real-time through the use of portable, on-board Emission Measurement Unit. The rate of vehicle emissions was evaluated in each mode including deceleration, idling and acceleration. Then the relationship between control delay and vehicle emission were studied. Four different vehicle types are used in the data collection. These include 1996 Oldsmobile Cutlass sedan, 1998 Plymouth Breeze sedan, 1999 Ford Taurus sedan, and Ford Club Wagon 15 passenger. Total 72 hour testing experiment was carried out with a travel distance of 2000 vehicle miles. Then the data are analyzed. This analysis shows car type and mode can cause very different emission results.

Pandey et al. (2016) have described the influences of the petrol vehicle's age and its mileage on tailpipe emissions of HC and CO. The plot of the influence relationship is as shown in Figure 2.4, which illustrates that as vehicles mileage and age increases, Emissions increase correspondingly. Vehicles after four years or mileage beyond 50000 miles will lead to a relative steady increase.



(a) Vehicle age vs. CO emission (Idle).	(b) Vehicle age vs. HC emission (Idle).	(c) Vehicle mileage vs. CO emission (Idle)	(d) Vehicle mileage vs. HC emission (Idle)

Figure 2.4 Influences of the vehicle's mileage and age on HC and CO emission.

2.5.3 Traffic Control Strategies

The previous research which emphasized the impact of traffic control strategies on vehicle emissions often focused on pre-timed control or an adaptive traffic control strategy applied at the study intersection; also an isolated signal intersection or coordinated intersections. The cycle time is often 50-100 seconds long in some cases. And some authors working on this aspect designed different scenarios based on different green time and cycle length to find the relationship between emission and length of various time (cycle length and green time). Li et al. (2011) investigated the impact of signal timing on vehicle emissions at an intersection. The study was based on a typical case when experienced by a vehicle at an intersection which includes deceleration, a full stop, and acceleration. The methodology can be divided into as two parts, which is called “two-stage approach”. The first step is to develop an optimization model in order to find out a “trade-off” between stops and vehicle delays; the output from the first part then is used in the second part to model the vehicle emission. In the first stage, number of stops and delay are formulated based on the study of Webster in 1958. Then an index gives a soft approach to obtain the balance between number of stops and delay. A vehicle trajectory of stopped delay and control delay in the under-saturation condition is analyzed under pre-timed traffic control. Li (add reference) et al. (2011) also used a model from Webster (1958) about the average delay per vehicle and the stop rate of vehicles to calculate the designed cycle length. By using the Webster’s model, several traffic and pre-timed control factors are involved, which include: traffic demand, cycle length, green time, effective green spilt, saturation flow and degree of saturation. All the potential cycle lengths are selected between 50 seconds to 200 seconds, then the average delay and stop rate are calculated.

Frey et al. (2002) used the Vehicle Specific Power mode (VSP) to estimate the vehicle emissions. It is presented as:

$$VSP = v\left(1.1a + 9.81 sin\left(\theta\right) + 0.132\right) + 0.000302 \qquad (2.27)$$

Where, θ is road grade, v is vehicle speed, and a is acceleration. Individual vehicle emissions are calculated based on the range of VSP value. Based on Frey’s study, VSP can be calculated from -2 to 39, which is then classified into 14 bins. Each corresponding bin can be then used to estimate CO, CO_2, NO_X, and HC amounts as a unit of grams per second.

2.5.4 Weather and Other Characteristic

Temperature and road grad change can affect CO_2 emission (Abou-Senna and Radwan, 2013). Abou-Senna and Radwan designed an approach for developing a microscopic transportation emissions model, which can be used to predict CO2 emissions on limited access highways. Key parameters, such as traffic-related (volume, truck percentage, speed limits), geometry-related (road grade) and environment-related (temperature), are selected for detailed evaluation. The results demonstrated that temperature changes from 50F to 100F increase about 12% CO_2 emission.

Early in 1997, Enns et al, have found that increases grade can cause CO production increase. Honmark (2002) indicated a relationship between engine load due to operation on a grade and elevated emissions in Baltimol. The author found that the grade is statistically significant in modeling vehicles emissions' model. This finding is quantified by Cicero Fenamdez et al. in 1997. The authors found that there will be 0.04g/mile of HC and 3g/ mile of CO increase produced by vehicles when 1% grade increases. Then Zhang and Frey (2012) carried out sensitivity analysis study to evaluate the importance of road grade with respect to vehicle's specific power and emissions. Their finding found that VSP and emissions are different as road grade and speed changes. They pointed out that when grade changes from 0-6%, VSP will increase 20 wt/ton. Sentoff et al. (2014) compared VSP frequency when grade is account and without account in calculating of second-by-second VSP. They found it is true in various road types. Sentoff et al. (2014) reveals that up to 48% CO emissions difference can be caused with and without considering grade in emissions calculation. They proved that account in road grade also could cause difference in operation mode distribution in all road type.

2.6 Summary

The research of the existing literature is mainly focus on the study of the impact of several aspects on the vehicle emissions. The one is the traffic condition, including the vehicle volume at the intersection, speed, number of stops, queue length and delay time; the second is the vehicle conditions, including the vehicle types, ages, mileage, weight, mode and the year the vehicle; and the third is the traffic control strategy, such as pre-timed control or an adaptive traffic control strategy; the fourth is the weather condition, such as temperature and road grad change, and the road grade also influence vehicle's emissions too. But as for the impact of operation modes of on-road vehicles, although it is said very important according to the description of reference (LeBlanc, 1995; Barth et al. 1997, Frey, H. Christopher, et al 2002, Ritner, Mark, et al, 2013,), but at present, there is no intense research to quantitatively evaluate the relationship between the operation modes and emissions. In addition, in this research, we use MOVES to obtain the emissions output, with its most accurate input method, to get the emission result, which is also few applied in current study.

CHAPTER 3

METHODOLOGY

3.1 Data Generation and Processing Method

As previously stated, the objective of this dissertation is to develop a vehicle emissions model that is capable of estimating Carbon Monoxide (CO), Ammonia (NH_3), Nitric Oxide (NO), and Nitrogen Oxide (NO_X), from on-road vehicles. Two model approaches are used including: multiple linear regression and ANN. For the multiple linear regression modeling, two kinds of regression models are developed: one using the operation modes as independent variables and another regression model using traffic related parameters as independent variables. Both regression models use the four types of pollutant emissions as the dependent variables. To achieve this goal, experimental data being used in the development of the estimation models is first generated using transportation and vehicles emission simulation software. The detailed methodological approach for the development of the emissions models is illustrated in Figure 3.1. Each module of the flow chart is discussed separately.

The organization of this chapter includes: Section 3.1 reviews the tasks to be performed. A flowchart of the methodology illustrates and explains the approach being used. Section 3.2 identifies the road characteristics for the study intersections. Section 3.3 discusses how vehicle activity data are collected through the traffic simulation software of VISSIM. Section 3.4 discusses how emission quantities of on-road vehicles are estimated through emission simulation software of MOVES. Section 3.5 and Section 3.6 provide two statistical analysis methods to develop the vehicle emissions estimation models, including multiple linear regression method and ANN method.

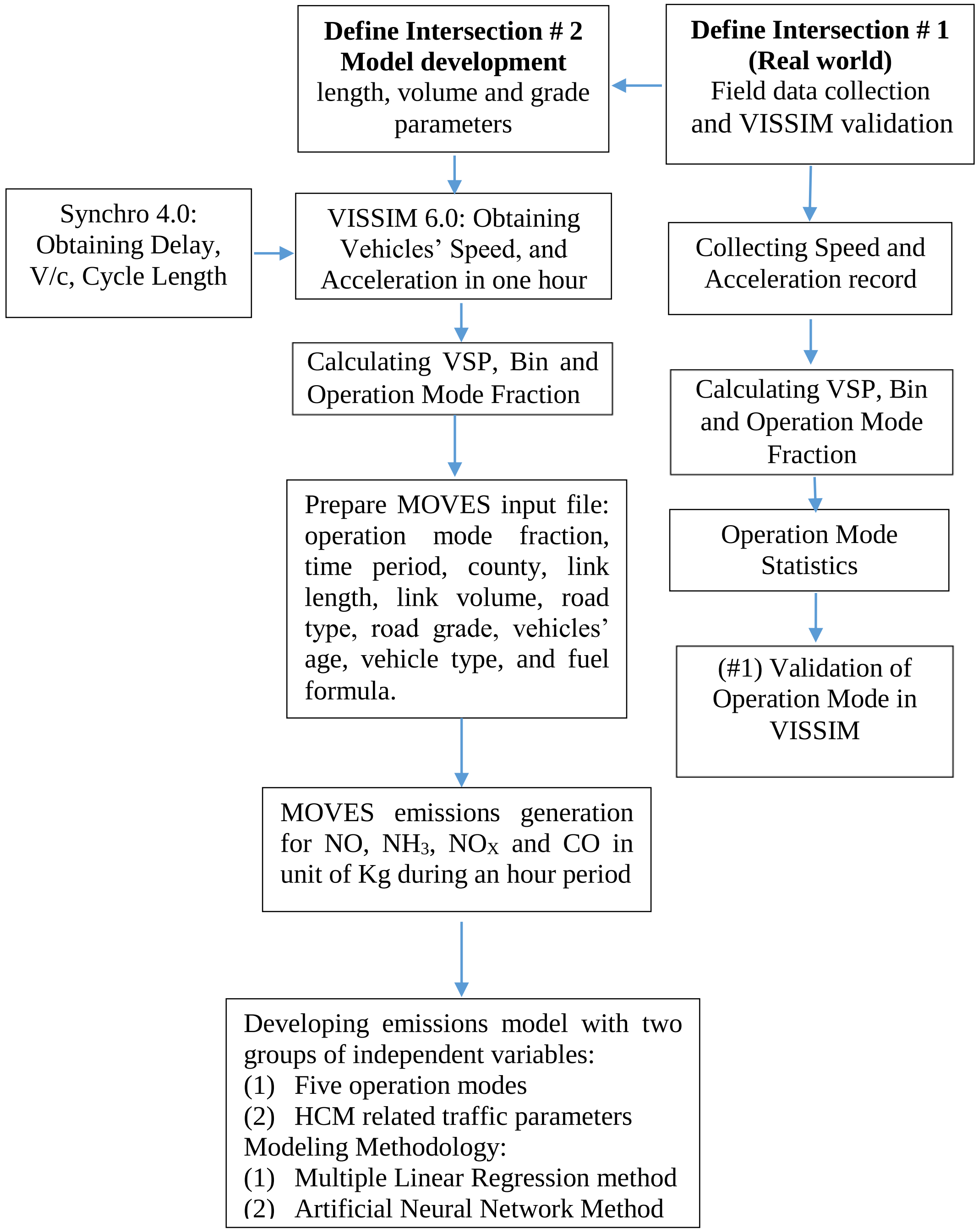


Figure 3.1 Flowchart of the methodology.

In this dissertation, the total emission at a signalized intersection within an hour period is interested. Figure 3.1 shows the flow chart of the methodological approach used in estimating the vehicle emissions estimation models. Firstly, we define two intersections. Intersection #1 is a real world signalized intersection, and intersection #2 is a virtual intersection. The purpose of the first intersection is to determine whether VISSIM can provide reliable operation mode fractions, which are the important parameters for generating vehicles emissions. Intersection #2 is established to generate vehicle emissions using a range of vehicular volume and grade as the key variables.

Figure 3.1 is the flow chart of methodology. The overall procedure can be divided into three parts. The first is to validate the operation mode fractions generated from VISSIM at a real world signalized intersection (Intersection #1). The second part is to generate operation mode fractions using VISSIM at a general signalized intersection (Intersection #2), and further to produce four types of emissions using MOVES. The third part is to develop the emission estimation models by using multiple linear regression (MLR) and Artificial Neural Network (ANN). In order to estimate the emission models using HCM related parameters, as many as possible variables, either traffic related parameters or road geographic parameters, are initially considered and collected. These variables include vehicular volume, road grade, delay, cycle length, Volume to capacity ratio (V/c) and green to cycle length ratio (g/c). Then from VISSIM, we can obtain every vehicle's second-by-second speed and acceleration at the intersection for a one-hour period. After getting the vehicle's second-by-second speed and acceleration from VISSIM, the vehicle's specific power (VSP) can be calculated and the operation mode percentage can be determined. Finally, using the operation mode fractions and other MOVES inputs, such as link source type, link, age distribution, metrology, fuel, impair and maintenance program (I/M), hoteling and retrofit data, vehicle emissions can be generated in unit of kg within one hour's period. In addition, a correlation analysis is carried out to determine correlation between the generated outputs and all possible influencing factors.

For the multiple linear regression models, we have two groups of variables to develop two kinds of multiple linear regression models. The first group of variables uses operation modes as the independent variables and aims to determine the contribution of each operation mode on each pollutant emission separately. The second group of variables uses HCM related variables such as volume and grade, to estimate the four types of emissions separately. This model aims to determine the contribution of each traffic related parameters on each pollutant emission separately.

VISSIM version 6 is used for generating the vehicle running data used in the development of the operation modes. The input parameters of VISSIM include: vehicle volume, signal timing, turning percentage, road characteristics, vehicle characteristics, driving behavior, and desired speed. Among them, the first two parameters are adjustable and the remaining variables use fixed values for each simulation run in this research. The output of VISSIM provides each vehicle's: speed, acceleration, delay time, in-queue, dwelling time, and number of stops. The advantage of using VISSIM is that we can obtain the speed and acceleration of every vehicle at every second within one-hour period. This data can be used to categorize the operation modes of the vehicles, which will be used as independent variables in the vehicle emission models.

The operation modes of vehicles are based on the second-by-second speed and acceleration information obtained from VISSIM. This data is used to categorize every car into five operation modes, i.e., Braking, Idling, Cruise/Acceleration, Low Speed Coasting, and Moderate Speed Coasting. Using the second-by-second speed and acceleration data obtained from VISSIM, the time fractions (or percentage) each vehicle spends in the five operation modes are calculated and imported into MOVES. Then the corresponding emissions of various vehicles pollutants are determined. Furthermore, through statistical analysis by using multiple linear regression method and ANN method, we can establish the relationship between vehicle emissions and the five operation modes, which can be used to ultimately estimate vehicles pollutant emissions.

3.2 Roadway Identification

The roadway used in this research generate microscopic data through VISSIM6. As shown in Figure 3.2, is a four-lane urban unrestricted access road with a signalized intersection. The length of the intersection is 1000 ft per leg. Details of the intersection layout and volume is zoomed in and shown on the right side of the figure to illustrate the intersection characteristics.

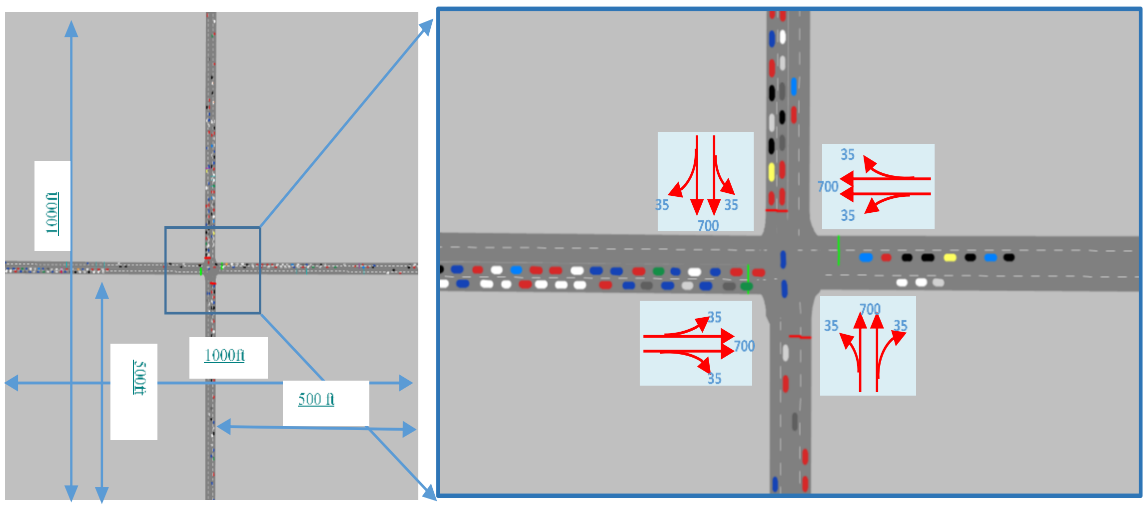


Figure 3.2 Roadway identification and simulation in VISSIM.

The road used in this dissertation has the following constraints (HCM 2010):

- 12 ft lanes in each direction
- No heavy vehicles
- Grades vary from -5% to 5%
- No parking
- No bus stop in the intersection
- Turning traffic 5%

From previous research on estimation of emissions, road grade has been found to have an important impact on vehicles' emissions. In this research, we will include road grade as one important influencing factor on vehicles emissions and determine its influence on vehicles emissions. Furthermore, we will explore the impact of volume changes on vehicle emissions.

3.3 Vehicle Activity Data Generation from VISSIM

VISSIM is used to generate data of on-road vehicles at the second-by-second level. The input to VISSIM includes operation time, number of vehicles, signal timing, random seed and geometric-related parameters, to generate the simulated vehicles runs. A simulation time of 3600 seconds is used for each simulation run. The volume of vehicles is to be simulated ranges from 200 veh/hr to 880 veh/hr per lane group approaching the intersection. The outcome of VISSIM includes the speed and acceleration of every vehicle for every second, which is called the activity data and will be used to categorize the operation modes of each vehicle into one of 23 mode IDs based on the EPA MOVES standards. In order to calculate the operation mode fractions, vehicle specific power (VSP) will first be calculated. VSP is a useful parameter that represents the amount of power required by one unit weight of vehicle to overcome rolling resistance, engine force and air force. The EPA introduced in VSP as an important influential factor of vehicle emissions when it was known that the speed alone could not fully reveal the relationship between vehicle emissions and the vehicle's performance. Generally, if the air resistance is assumed as a constant and the road grade is assumed to be zero, then the VSP only depends on the vehicle's instantaneous speed and acceleration/deceleration. Each vehicle's VSP and its instantaneous speed are combined to determine the vehicle's five operation modes, i.e., Braking, Idling, Cruise/Acceleration, Low Speed Coasting, and Moderate Speed Coasting.

Table 3.1 shows an example for determining the operation mode, mode ID and fraction for the studied intersection under the case when the intersection volume is selected to be 200 vph. The 200 vph includes the total number of vehicles at the intersection within an hour with 90% the vehicles through movements, 5% turning right and 5% turning left. Column 1 in Table 3.1 provides the 'link' number that corresponds to the operational mode of the vehicles. In MOVES, a 'link' is defined as the partial length of roadway that belongs to the same operation mode. For example, at an upstream location when the signal turns green, vehicles tend to accelerate. This is a link of acceleration, which is different from other operation modes such as a deceleration mode when approaching vehicles need to stop. There are five links based on the operation mode definition.

Table 3.1 Link, Link Length and Volume

Link	Operation Mode	Mode ID	Frequency(vph)	Frequency (vph) based on Link	Operation mode Fraction	Op Mode Fraction based on Link	Link Length	Link Volume
1	Braking	0	8467	8467	0.1207	0.1207	0.0549	97
2	Idling	1	8139	8139	0.1160	0.1160	0.0528	93
3	Low Speed Coasting	11	20366	20366	0.2903	0.2903	0.1321	232
5	Moderate Speed Coasting	21	11921	11921	0.1699	0.1699	0.0773	136
4	Cruise/Acceleration	12	4837	21264	0.0689	0.3031	0.1379	242
		13	1638		0.0233			
		14	1566		0.0223			
		15	11273		0.1607			
		16	731		0.0104			
		22	39		0.0006			
		23	25		0.0004			
		24	60		0.0009			
		25	135		0.0019			
		27	0		0.0000			
		28	70		0.0010			
		29	202		0.0029			
		30	688		0.0098			
		33	0		0.0000			
		35	0		0.0000			
		37	0		0.0000			
		38	0		0.0000			
		39	0		0.0000			
		40	0		0.0000			

In Table 3.1, the second column “Operation Mode” classifies vehicles’ operation activity into one of the previously stated five operation modes. Column 3 includes 23 kinds of ‘Mode ID’. Each vehicle’s activity is assigned a specific Mode ID based on the VSP and speed data. The list of mode IDs is provided in the Table 3.2. The frequency in column 4 represents the total number of seconds that vehicles at the intersection operate in the corresponding Mode ID. Column 5 shows the frequency based on the corresponding Link. Column 6, ‘Operation mode Fraction’, is the percentage of vehicles associated with the frequency of the Mode ID. In column 7, ‘Op-Mode Fraction based on Link’, is the number of vehicles associated with the frequency of the corresponding Link. The information provided in Table 3.1 also includes the link length and link volume parameters, which are used as inputs to MOVES for estimating vehicle emissions.

Table 3.2 EPA Defined VSP and Speed Range for Each Operation Mode

Mode ID	Operation Mode
0	Braking: Acceleration<-2 mph/s, or<-1 mph/s for 3 consecutive seconds
1	Idling: -1≤Speed<1
11	Low Speed Coasting: VSP<0; 1≤Speed<25
12	Cruise/Acceleration: 0≤VSP<3; 1≤Speed<25
13	Cruise/Acceleration: 3≤VSP<6; 1≤Speed<25
14	Cruise/Acceleration: 6≤VSP<9; 1≤Speed<25
15	Cruise/Acceleration: 9≤VSP<12; 1≤Speed<25
16	Cruise/Acceleration: 12≤VSP; 1≤Speed<25
21	Moderate Speed Coasting: VSP<0; 25≤Speed<50
22	Cruise/Acceleration: 0≤VSP<3; 25≤Speed<50
23	Cruise/Acceleration: 3≤VSP<6; 25≤Speed<50
24	Cruise/Acceleration: 6≤VSP<9; 25≤Speed<50
25	Cruise/Acceleration: 9≤VSP<12; 25≤Speed<50
27	Cruise/Acceleration: 12≤VSP<18; 25≤Speed<50
28	Cruise/Acceleration: 18≤VSP<24; 25≤Speed<50
29	Cruise/Acceleration: 24≤VSP<30; 25≤Speed<50
30	Cruise/Acceleration: 30≤VSP; 25≤Speed<50
33	Cruise/Acceleration: VSP<6; 50≤Speed
35	Cruise/Acceleration: 6≤VSP<12; 50≤Speed
37	Cruise/Acceleration: 12≤VSP<18; 50≤Speed
38	Cruise/Acceleration: 18≤VSP<24; 50≤Speed
39	Cruise/Acceleration: 24≤VSP<30; 50≤Speed
40	Cruise/Acceleration: 30≤VSP; 50≤Speed

3.4 Emissions Quantities Generated by MOVES

This section explains how emissions quantities are produced by MOVES. The operational mode of each vehicle is calculated based on a combination of the VSP and speed data ranges. The VSP calculation method can be found in Chapter 2 from EPA MOVES.

$$VSP\ (kW/ton) = v * \left[1.1a + 9.81a\ \left(tan\left(sin\left(grade\right)\right)\right) + 0.132\right] + 0.000302\ v^3 \quad (3.1)$$

Where, a is the instantaneous acceleration, v is the speed, and grade is the ratio of the road height over the road length. These ranges and the operation modes associated with speeds are presented in Table 3.2. For every vehicle and for each second of the simulation within an hour period, a vehicle's operational mode can be defined. Hence, the number of seconds that vehicles are in a particular operational mode can be determined. The EPA provides an emission rate for each Mode ID or operational mode, which is then used to convert the number of seconds for each operational mode to calculate the total emissions during an hour.

For most of the input files in MOVES, default values can be used including fuel, vehicle age distribution, impair and Maintenance program (I/M), which can be extracted from the EPA MOVES database. Files of operation mode fraction, link, volume and grade are significant in determining vehicles emissions' quantity of interested pollutants in MOVES and therefore default values are not used.

3.5 Emissions Model Development

3.5.1 Multiple Linear Regression Model Development

The vehicle emissions model to be developed in this research will be based on the data obtained from MOVES. The model will be developed using multiple linear regression to estimate the relationship between pollutant emissions and the five operation modes. A second model will estimate the relationship between pollutant emissions and HCM related parameters.

First, five operation modes are taken as the independent variables in the multiple linear regression process. These operation modes are denoted as vector $\mathbf{X} = [1, x_1, x_2, \cdots, x_m]$, where m is represented as the number of the independent variables, i.e., m=1 to 5. The four types of pollutant emission elements are the independent variables and they are denoted as vector $\mathbf{Y} = [y_1, y_2, y_3, y_4]^{\mathrm{T}}$. Let $\boldsymbol{\beta}$ to be the matrix coefficient, and $\boldsymbol{\beta} = [\beta_{0i}, \beta_{1i}, \cdots, \beta_{mi}]^{\mathrm{T}}, i = 1, 2, 3, 4$.

The second kind of multiple linear regression model involves the HCM related parameters as the independent vector. These HCM related parameter, grade and v/c, are denoted as vector $\mathbf{X} = [1, x_1, x_2, \cdots, x_m]$, where m represent the independent number, i.e., m=1 for grade and m=2 for v/c. The four types of pollutant emissions are the dependent variables, denoted as vector $\mathbf{Y} = [y_1, y_2, y_3, y_4]^{\mathrm{T}}$.

X and Y satisfy the following linear regression model (Kutner and Chris,2004):

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{3.2}$$

Where $\boldsymbol{\varepsilon}$ refers to the random noise vector and satisfies the following conditions of the mathematical expectation $\mathrm{E}(\boldsymbol{\varepsilon})$ and the variance:

$$\begin{cases} \mathrm{E}(\boldsymbol{\varepsilon}) = 0 \\ \mathrm{Cov}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}) = \sigma^2 \mathbf{I}_4 \end{cases}, \tag{3.3}$$

Where: $\mathbf{I}_4$ is a 4×4 identity matrix, and σ^2 is the variance of the random noises. A multiple linear regression model with more than two predictor variables can be presented as:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_2 + \ldots \beta_p x_{i,p} + \varepsilon_i \tag{3.4}$$

The coefficient $\boldsymbol{\beta}$ can be calculated by using the minimum least square method. The estimated values are denoted as

$$\hat{\boldsymbol{\beta}} = [\hat{\beta}_{0i}, \hat{\beta}_{1i}, \cdots, \hat{\beta}_{mi}]^{\mathrm{T}} \tag{3.5}$$

And fitting errors can be expressed as:

$$\mathbf{Q} = \sum_{i=1}^{n} (\mathrm{Y_i} - \beta_0 - \beta_1 x_{i1} - \ldots - \beta_{(p-1)} x_{i,p-1})^2 \tag{3.6}$$

So the multiple linear regression equation between $\mathbf{X}$ and $\mathbf{Y}$ can be expressed as:

$$\hat{y}_i = \hat{\beta}_{0i} + \hat{\beta}_{1i} x_1 + \hat{\beta}_{2i} x_2 + \cdots + \hat{\beta}_{mi} x_m, \qquad i = 1,2,3,4 \tag{3.7}$$

Where, the vector $\hat{\mathbf{Y}} = [\hat{y}_1, \hat{y}_2, \hat{y}_3, \hat{y}_4]^{\mathrm{T}}$ is the estimator of the four types of emissions and $\hat{\boldsymbol{\beta}}$ refers to the regression model estimated coefficients.

3.5.2 Significance Test of Multiple Linear Regression Equation

The F-distribution test is used to verify if the obtained linear regression equations are significant, as shown in the following:

$$\hat{y}_i(k) = \hat{\beta}_{0i} + \hat{\beta}_{1i} x(k) + \hat{\beta}_{2i} x(k) + \cdots + \hat{\beta}_{mi} x_m(k) \quad k=1\sim n,\ i=1, 2, 3, 4 \tag{3.8}$$

where k is the serial number of the sample data of $\mathbf{X}$, and n is the sample size of the experiment data.

Let

$$SSR = \sum_{k=1}^{n} [\hat{y}_i(k) - \overline{y}_i]^2, \mathrm{SST} = \sum_{k=1}^{n} [y_i(k) - \hat{y}_i(k)]^2 \tag{3.9}$$

Where $\hat{y}_i(k)$ is the estimated pollutants emissions value of type i in the *k*-th number of sampling using the predicted model of Equation (3.8). $\overline{y}_i$ is the mean value of the sample. *SSR* is the regression sum of squares, and *SST* means the residual sum of squares.

The F-test is used to verify the significance of the obtained multiple linear regression models. Suppose the Null hypothesis is stated as $H_{0i} : \hat{\beta}_{0i} = \hat{\beta}_{1i} = \cdots = \hat{\beta}_{mi} = 0$, and the alternative hypothesis is stated as H_1 : not all betas equal to zero. Based on Equation (3.9), we have:

$$F_i = \frac{SSR / m}{\mathrm{SST}/ (n - m - 1)} \sim \mathrm{F}(m, n - m - 1) \tag{3.10}$$

Where *m* is the number of the independent variables, and *n* is the number of the sample size of the independents and dependents. If we choose a significance level of α, such as 0.02 or 0.05, then we can find the corresponding critical value of $\mathrm{F}_{(1-\alpha)}(m, n - m - 1)$, as shown in Figure 3.3. If the calculated F_i is greater than $\mathrm{F}_{(1-\alpha)}(m, n - m - 1)$, which means the obtained regression model is a small probability event, hence, H_0 hypothesis should be rejected. The results can be used to verify whether the constructed multiple linear regression relationships between independents and dependents variables are significant or not.

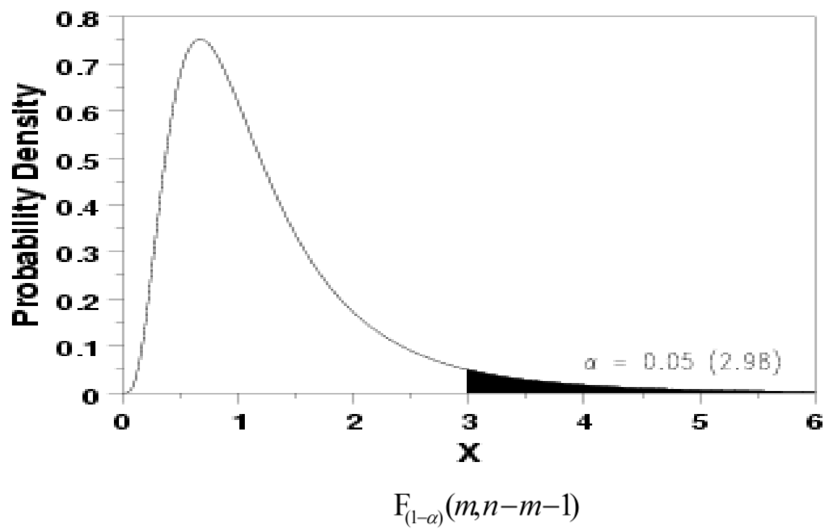


Figure 3.3 F distribution plot (William, 2006).

3.5.3 Calculation of Standard Regression Coefficients

The standardized regression coefficients $\hat{\beta}_i^*$ can be used to reveal the importance of an independent variable's influence on the dependent. The method of obtaining $\hat{\beta}_i^*$ is to transform the distribution of the independents into the standard normal distribution, so that we can transform non-standardized coefficients of $\hat{\beta}$ to be standardized values of $\hat{\beta}_i^*$. Standardized coefficients also are recommended for improving the calculation accuracy (Neter, et al. 1989). Standardized betas measure the impact of independent variables on dependent variable. The method of calculating the standardized regression coefficients is as following:

$$\hat{\beta}_i^* = \hat{\beta}_i \cdot \frac{s_{x_1}}{s_{yi}} \tag{3.11}$$

Where $\hat{\boldsymbol{\beta}}_{\mathbf{1i}}^{*}$ is the standardized regression coefficients, s_{x_1} and s_{yi} are standard deviation of the independent variables and the dependent variables, respectively. They presented as:

$$s_x = \sqrt{\frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \overline{x})^2} \tag{3.12}$$

$$s_y = \sqrt{\frac{1}{(n-1)}\sum_{i=1}^{n}(\mathrm{y_i} - \overline{\mathrm{y}})^2} \tag{3.13}$$

From Equation (3.8) we can obtain the following equation, where $\overline{y}$ refers to the mean value of the observed dependent and $\overline{x}$ the mean value of observed independent value.

$$\overline{y}_i = \hat{\beta}_{0i} + \hat{\beta}_{1i}\overline{x}_1 + \hat{\beta}_{2i}\overline{x}_2 + \cdots + \hat{\beta}_{mi}\overline{x}_m, \qquad i = 1,2,3,4 \tag{3.14}$$

Subtracting Equation (3.14) from (3.8), and dividing the difference by the standard deviation s_{y_i} , then:

$$y_i^* = \hat{\beta}_{1i} \cdot \frac{s_{x_1}}{s_{y_i}} \cdot x_1^* + \hat{\beta}_{2i} \cdot \frac{s_{x_2}}{s_{y_i}} \cdot x_2^* + \cdots + \hat{\beta}_{mi} \cdot \frac{s_{x_m}}{s_{y_i}} \cdot x_m^*, \qquad i = 1,2,3,4 \tag{3.15}$$

The larger the value of the standardized value of $\hat{\boldsymbol{\beta}}_{\mathrm{i}}^{*}$, the greater the influence of that independent variable on the emissions. Therefore, we can quantitatively evaluate the contribution of various operation modes on the different kinds of polluted emissions.

3.6 Emissions Predicting Using Artificial Neural Network

3.6.1 Introduction

Artificial Neural Networks is simulating the structure and function of the human neural network. It is a kind of computing method, which is based on the mechanism of human neurons. ANN consists of many Neurons. They are organized according to a certain topology, and connected with each other. ANN is a network, which has the function of parallel processing. Anderson (1992) provides a classical framework of neural network architecture selection. All the problems are categorized into one of the five major application of using neural network including: prediction, classification, data association, data conceptualization and data filtering. Prediction happens when a set of input data and target data is given and approximate function then generated after the network learning. The success of prediction depends on the quality of training data and architecture of the natural network. (Taylor, 2006) Artificial Neural Networks (ANN) method is applied in estimating four types of emissions through the HCM traffic related variables in this dissertation. Back-propagation and five other learning methods are commonly used in prediction purpose. In addition, for number of neurons, the following Equation 3.16 is suggested a good rule-of- thumb (Anderson, 1992)

$$\mathrm{Ni} = \frac{\text{Number of training Data Pairs}}{(\text{Number of input Neurons} + \text{Number of output Neurons}) * \alpha} \tag{3.16}$$

Where *Ni* is the number of neurons in the *i*-th layer. α is the factor from one to fifty, which is selected based on the noiseless of the data. Nearly noiseless data is suggested to use α value between one and five. Typically, noisy data is to use ten. It is noted that too many neurons within a hidden layer causes training set to be simply memorized (Taylor, 2006)

ANN has the following characteristics: (1) massively parallel processing; (2) robustness and fault tolerance; (3) self-learning ability; (4) large-scale nonlinear systems adaptive, collective operations. At present, ANN has been widely used in various fields including Economics, Industry, Transportation, and others (Taylor, 2006).

3.6.2 Model of Artificial Neurons

Fig. 3.4 shows the schematic of the model of an artificial neuron.

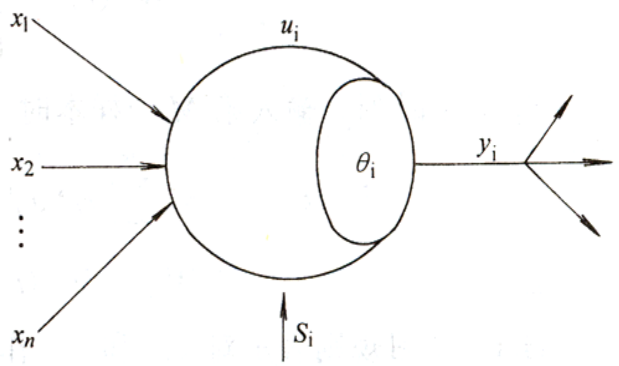


Figure 3.4 Model of artificial neuron.

In Figure 3.4, u_i refers to the inner status of the neuron, θ refers to the threshold value which refers to an inner status value of the artificial neuron. Only when the external stimulus value exceeds the threshold value, is the artificial neuron able to produce a response. x_j $(j=1,2,3,\ldots n)$ refers to the various variables input into the neuron. W_{ij} is the weight value connected with the neurons i and j. S_i refers to the external environmental influence, and y_i is the output value of the neuron. The model can be expressed as:

$$\begin{aligned} Net_i &= \sum W_{ij} x_j + s_i - \theta_i \\ \mu_i &= g(\mathrm{Net_i}) \end{aligned} \tag{3.17}$$

$$y(\mathrm{t}) = H[\mu_i(\mathrm{t})] \tag{3.18}$$

Where Net_i represents the net input of a neuron, g function means the active function. *H* refers to the output function. The most commonly used *H* functions are shown in Fig. 3.5

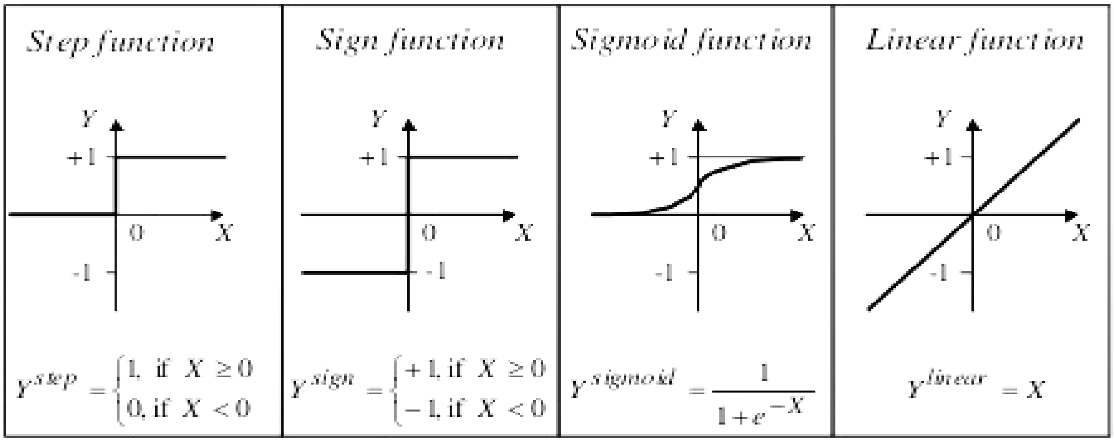


Figure 3.5 Activation function of an artificial neuron (Neural Network, 2017).

In Figure 3.5, the x-axis refers to the input of the *H* function, and the y-axis refers to the output of the *H* function. In neural network area, the sigmoid function is generally used. For the Sigmoid function, the x input is from -∞ to +∞, the y output is from 0 to 1 and suitable to neural networks with output values between 0 and 1. For the Linear function, the x input is from -∞ to +∞, the y output is also from -∞ to +∞ and suitable to neural networks with output values between -∞ and +∞. Therefore, according to the modelling needs, we can choose different activation functions for the artificial neuron modeling.

3.6.3 Neuron's Learning Rule (Adjustment of Weight Value)

Figure 3.6 shows the complex structure of the ANN and weight values.

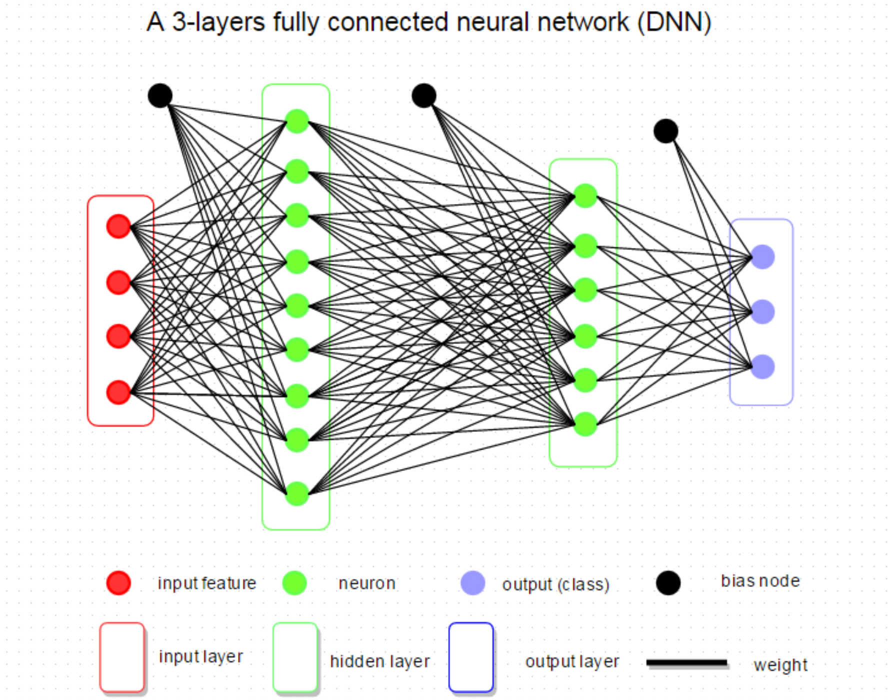


Figure 3.6 ANN structure.

In order to adjust the weights in Fig. 3.6, there are two main learning rules:

1）Hebb Learning Rule:

Assume x_j is the input of the j-th neuron. It may come from the other neuron's output, or an input from the exterior networks. Since the i-th neuron output y_i could be the next neuron's input, we denote it as x_i instead of y_i. Suppose W_{ij} is the weight value of input x_i to neuron *j*. Let x_i^s and x_j^s be the x_i and x_j values at sample S. When there are M-1 sample being input into the neuron, in terms of Hebb rule, the weights are expressed as:

$$W_{ij}(\mathrm{t}) = \begin{cases} \sum_{s=0}^{M-1} x_i^{(s)} & (\mathrm{i} \neq \mathrm{j}) \\ 0 & (\mathrm{i} = \mathrm{j}) \end{cases} \tag{3.19}$$

At the moment (t+1), inputting M sample, the weight values become:

$$W_{ij}(\mathrm{t}+1)=\sum_{x=0}^{M}x_i^{(\mathrm{s})}x_j^{(\mathrm{s})}=W_{i\,j}(\mathrm{t})+x_j^{(\mathrm{M})}x_j^{(\mathrm{M})} \qquad i\neq j \tag{3.20}$$

Then the weight change can be calculated by:

$$\Delta W_{i\,j}=W_{i\,j}(\mathrm{t}+1)-W_{i\,j}(\mathrm{t})=x_j^{(\mathrm{M})}x_j^{(\mathrm{M})} \tag{3.21}$$

Hebb rule shows that when $x_j^{(\mathrm{s})}$ and $x_i^{(\mathrm{s})}$ do calculation (not clear with what is "do calculation", and if they match, then the weight value $\Delta W_{i\,j}$ is to be strengthened, or the weight value will be weakened.

We take consideration of other neuron's inputs and exterior's inputs, and denote them as x_i , then the rule is:

$$Net_i=\sum_{j=1}W_{ij}x_j-\theta \tag{3.22}$$

Further, when $W_{i0}=-\theta$, and let x_0 =1, then we have

$$Net_i=\sum_{j=1}W_{ij}x_j \tag{3.23}$$

let the active function $g(\mathrm{Net_i})=\mathrm{Net_i}$, the model is simplified:

$$\mathrm{y_i}=\mathrm{f}(\mu_\mathrm{i})=f(\mathrm{Net_i})=(\sum_{j=1}W_{ij}x_j) \tag{3.24}$$

Or

$$\mathrm{y_i}(\mathrm{t}+1)=\mathrm{f}(\sum_{j=0}W_{ij}x_j(\mathrm{t})) \tag{3.25}$$

If it is represented as vector,

$$Y=f(\mathrm{WX^T}) \tag{3.26}$$

Or

$$Y_{(t+1)}=f(\mathrm{WX^T}(\mathrm{t})) \tag{3.27}$$

It shows the input vectors and output vectors have related with function of *f*, and this relationship can be either linear or non-linear.

2) Widrow-Hoff Rule

At time t, neuron i has a weight value $W_{ij}(\mathrm{t})$ to the jth input. At time t+1, the factor of j is xi, at this moment, the expectation of neuron i is d, and the real output is xi. In terms of Widrow-Hoff Rule,

$$W_{ij}(\mathrm{t}+1) = W_{ij}(\mathrm{t}) + \alpha(\mathrm{d_i} - \mathrm{x_i})\,\mathrm{x_j} \tag{3.28}$$

or

$$\Delta W_{ij} = W_{ij}(\mathrm{t}+1) - W_{ij}(\mathrm{t}) = \alpha(\mathrm{d_i} - \mathrm{x_i})\,\mathrm{x_j} \tag{3.29}$$

Where α is modifier factor, when α >0, δ is the difference value from the expected value and the real output $\delta = d_i - x_i$.

3.6.4 Error Back Propagation Method (BP)

BP network has been most widely used for data training, which represents about 80 to 90 percent among all other networks, as shown in Fig. 3.7. There are mainly three characteristics of a BP network. First, Pattern recognition and classification are especially used in language, words and graph identification. It is also used for categories in medical characteristics, and diagnosis. Second, BP network can be used for function approximation. It can be used for nonlinear curve fitting control, the trajectory of the robot, and industry control. It can also be used for data compression, storage and retrieval.

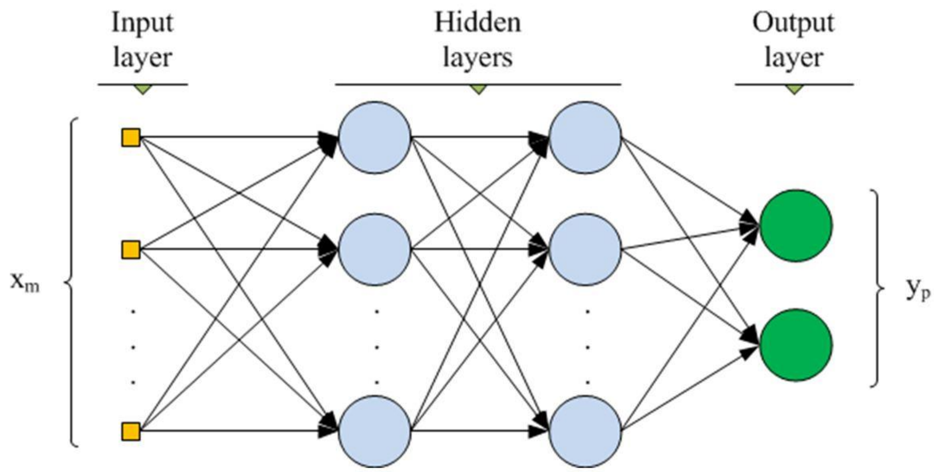


Figure 3.7 BP neural network structure.

Error Back Propagation Method shows the above structures in Figure 3.8. The neuron model that applied for nodes is represents as:

$$Y = f[\mathrm{WX}^{\mathrm{T}}] \tag{3.30}$$

Or

$$Y(\mathrm{t}-1) = f[\mathrm{WX}^{\mathrm{T}}(\mathrm{t})] \tag{3.31}$$

The characteristics of BP network starts from the input layer, experiences hidden layers and finishes as an output layer. There are no connections in the same layer, but only for the adjacent layers. For BP network, all transfer functions use Equation 3.31, which is also called sigmoid function.

$$f(\mathrm{x}) = \frac{1}{1+e^{-x}} \tag{3.32}$$

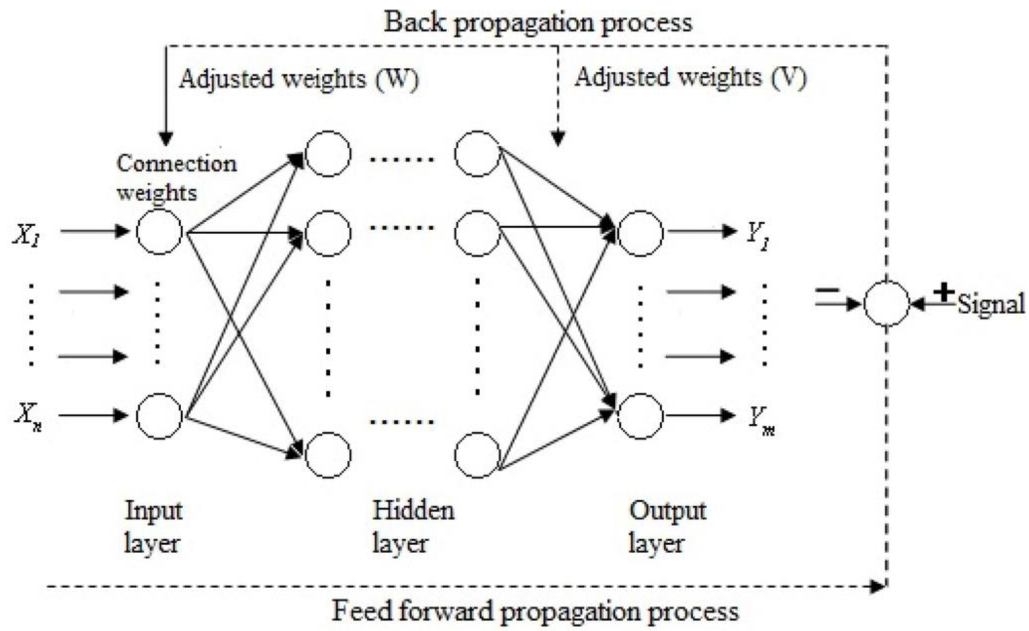


Figure 3.8 BP neural network calculation procedures.

The procedure of BP neuron network generally consists of two steps: one is the forward working to calculate the connection weight values; it starts from the nodes of the first layer to the last layer, until every node of every layer has been all calculated. The other step is called a learning step. In this procedure, the output of each node remains unchanged. From the output layer, backwards adjusting the weights of every node one layer by one layer until the input layer is reached. This procedure is then repeated until the error is limited within the settled error.

CHAPTER 4

EXPERIMENTAL DATA PROCESSING AND RESULTS ANALYSIS

4.1 Data Processing Flowchart

This chapter consists three parts. Section 4.2 describes the VISSIM outputs validation; section 4.3 describes the emission data generation; and section 4.4 describes the emission model development. As described in Chapter 3, VISSIM was used to generate operation mode fraction data. Chi-squared tests were then used to determine whether VISSIM simulation data and field collected data are close to each other. The operation mode data were then used in MOVES 2014a and four types of pollutant emissions were generated. These pollutants included CO, NO, NO_X and NH_3. Using the emissions results from MOVES, multiple linear regression models were estimated. In order to improve the estimation accuracy, an ANN model was developed based on the HCM related parameters. Finally, MOVES-like estimation models of vehicle emissions at a signalized intersection were developed Figure 4.1 shows a flowchart of the steps taken in developing the models.

It is noted that two groups of variables are prepared for developing the multiple linear regression (MLR) models: the operation modes and the HCM related variables. As mentioned in Chapter 1, vehicle's operation modes have a significant impact on the emissions results (LeBlanc, 1995; Barth et al. 1997; Frey H. Christopher, et al. 2002; Ritner Mark, et al. 2013). Hence, operation mode model helps one understand how each operation mode can influence the four types of emissions, respectively in various conditions. The HCM related variables considered in the development of the multiple linear regression model included delay, V/c, cycle length, road grade and g/c as the potential variables. The purpose of this model is to use the HCM related parameters to predict vehicles emissions. Correlation analysis is used to avoid multicollinearity in the MLR model.

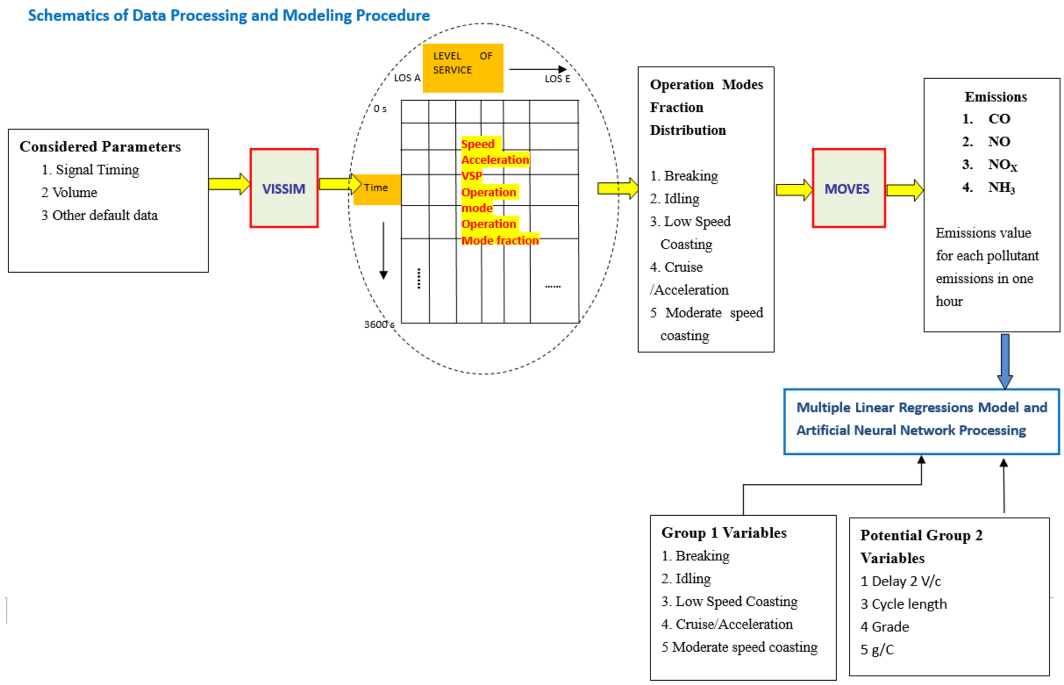


Figure 4.1 Data processing flowchart.

4.2 Validation of VISSIM Output Operation Mode Data by Using Real World Vehicle Running Data

There are many benefits to using simulation to study emissions. Some of these benefits include lower costs, lower risk and an easier way to capture a large quantity of vehicles' second-by-second data. Various traffic simulation models are used to process emissions study. VISSIM has been used in previous research in the study of vehicle emissions and the use of the simulation's output. Some research also tested the reliability of the data generated from VISSIM for emissions' study (Song et al. 2013). To verify the speed and acceleration results obtained from VISSIM in this dissertation, field data was collected and compared with the simulation data output from VISSIM. This validation has practical significance as it verifies the feasibility of using VISIMM's output results to study vehicles emissions. Therefore, validation was done to determine the adequacy of using VISSIM's speed and acceleration output data for using in MOVES. A Chi-square test was then performed to illustrate whether the VISSIM data were close to the field data at one intersection.

Figure 4.2 shows the validation procedure flowchart. Using VISSIM, a real world signalized intersection was simulated. Volume and signal timing at the selected intersection during 6-7 pm on weekdays was surveyed in the field and further used in the VISSIM simulation to generate the vehicle's running information. A total of 30 simulation runs with different random seeds were applied in VISSIM. Based on the second-by-second speed and acceleration output obtained from VISSIM simulation, the VSP, Mode ID, and then the average operation mode fraction was calculated.

Speed and acceleration data were also collected at the same intersection under similar operating conditions in the real world. Using this data, the operation mode fraction was calculated using a similar approach as was used for the simulation data. In the field measurement, vehicles running data were collected for 32 runs by driving from different approach directions. A chi-square test was performed where the calculated Chi-square value, which is the difference between the observed (field) data and expected (VISSIM) data, is 0.304147. Compared to the Chi-square table, which has a degree of freedom of 4, for a significance level of 5%, the Chi-square table value is 9.488. The calculated Chi-square value is smaller than the Chi-square critical value for a significance level of 0.05. Based on this comparison, the observed value and the simulation value are likely to be close to each other. Therefore, VISSIM output data is accepted for using in this emissions study.

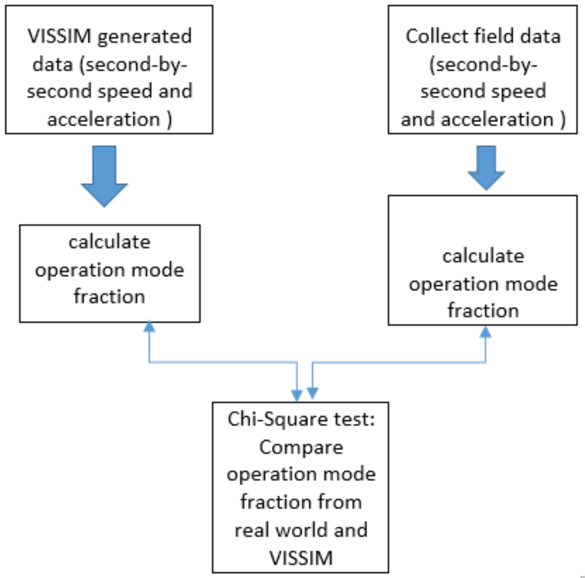


Figure 4.2 VISSIM validation processing flowchart.

4.2.1 Intersection Selection

Figure 4.3 shows the intersection used to perform the field test to validate VISSIM's speed and acceleration data. The data collection procedure included collecting speed and acceleration data by using an in-vehicle mounted device. Further details of this device is provided below. The device collected information about the running data including speed and acceleration experienced during the vehicles operation. The operation modes' fractions were then calculated using this data and the results compared to similar results obtained from VISSIM. The VISSIM simulation model used similar geometric, volume and speed data as existed in the location where the field data were gathered. Finally the field-measured operation modes were compared with the VISSIM simulated operation modes, to validate the effectiveness of the VISSIM simulation output.

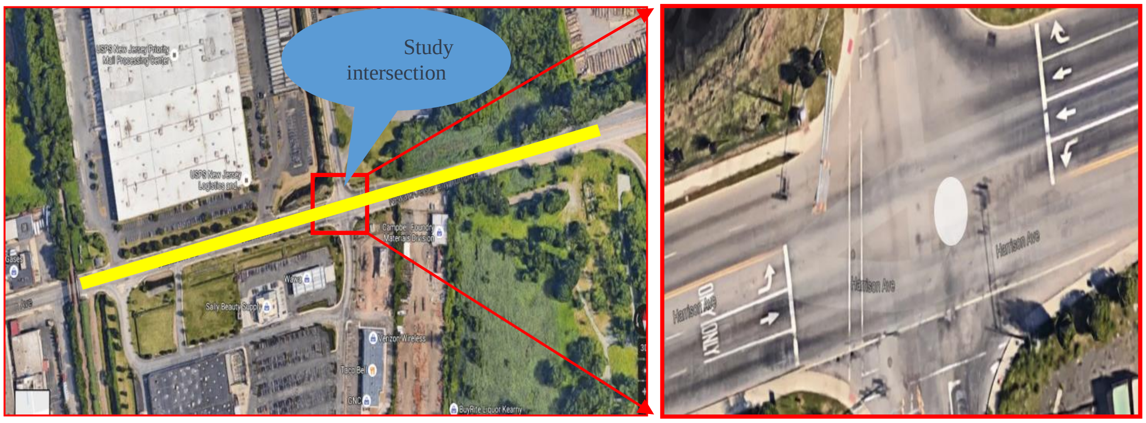


Figure 4.3 Intersection (1): snap shot of study location from Google map.

Figure 4.4 illustrates the average volume of the intersection during 6-7 pm on weekdays. This volume is used in VISSIM input to obtain the simulated vehicles running data such as second-by-second speed and acceleration, and further used to calculate the simulated operation modes fractions at this intersection. Speed and acceleration data was measured and this data is then used to calculate the field VSP and operation modes fractions. And finally, both the operation modes fractions from VISSIM simulation and the field data are compared.

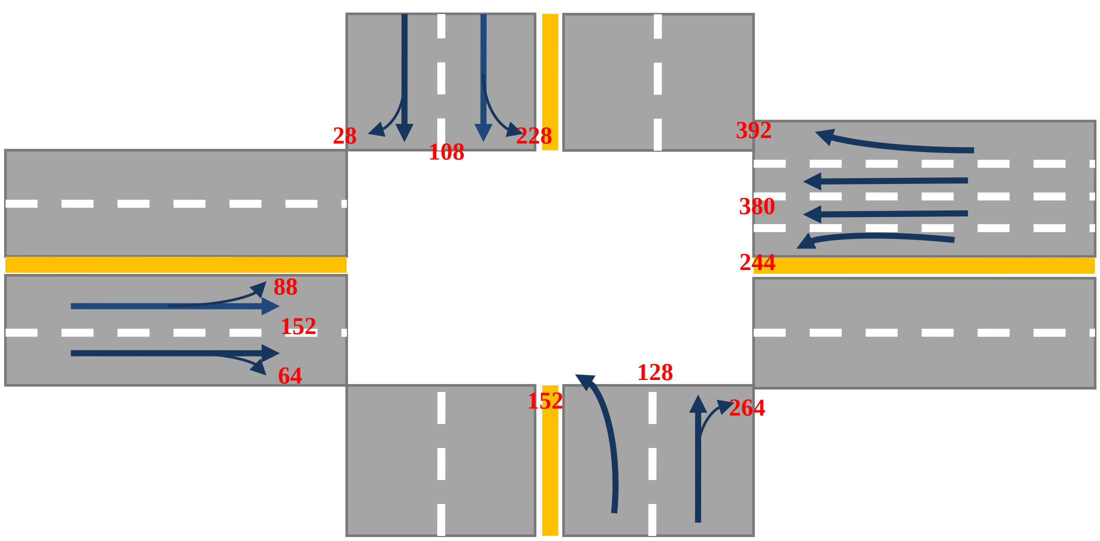


Figure 4.4 Vehicles average volume distribution at each direction between 6-7 pm on weekdays.

4.2.2 VISSIM Simulation Results

VISSIM simulation is designed for exactly the same study location, as shown in Figure 4.5.

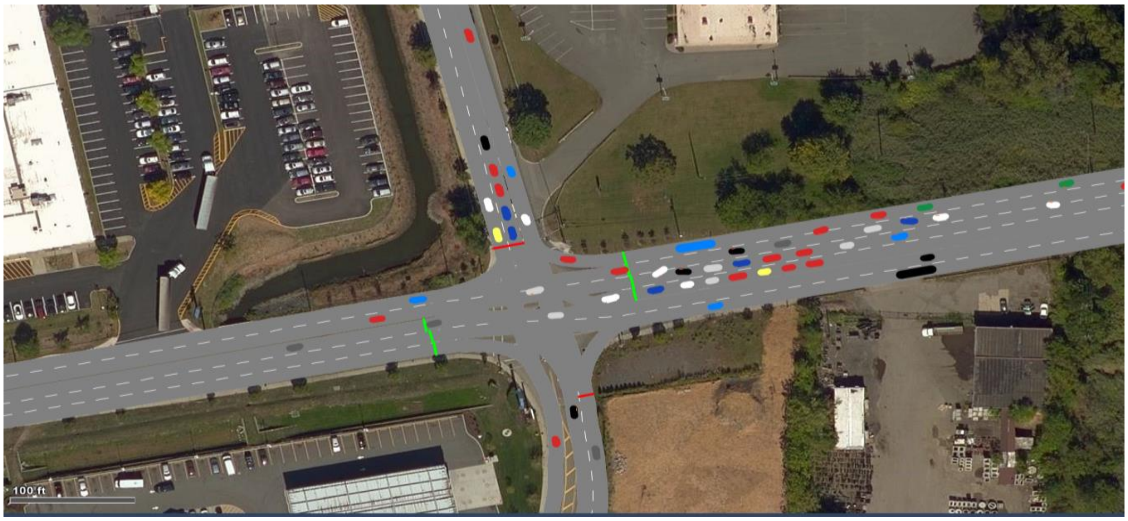


Figure 4.5 VISSIM simulation location for data collection during 6-7 pm on weekdays.

4.2.3 Field Measurement Results

Speed and acceleration were obtained using a measurement device named “Performance-box”. “Performance-box” is a tool that is installed in a vehicle and can measure the vehicle’s instantaneous speed and location and acceleration. The device is a professional car test tool that is accurate and easy to measure vehicles’ running data, such as location, distance, speed, acceleration, running time information, and other parameters. The device is installed on the window of the tested car as shown in Figure 4.6. Performance-Box is a GPS based performance meter that allows one to measure G-forces, speed, lap & split times, braking distance and many more. The new Predictive Lap Timing function (live comparison to best lap) provides instant feedback during one’s driving times. All parameters are logged to an SD memory card ten times per second for downloading and comparison. Performance-Box contains an integrated antenna, is compact in size and very easy to install. (Performance-Box Technical Specifications. Accessed on January 19, 2017)

The technical specifications of Performance-Box have the following characteristics:

(1) Velocity: accuracy of 0.1 Km/h (averaged over 4 samples); of 10 Hz; minimum velocity of 0.1 Km/h; maximum velocity of 1600 km/h; resolution of 0.01 Km/h.

(2) Distance: accuracy of 0.05 % (<50cm per Km); resolution of 1 cm.

(3) Acceleration: accuracy of 1 %; maximum of 4 G; resolution of 0.01 G.

(4) Heading: resolution of 0.01° (averaged over 4 samples): accuracy of 0.1°.

(5) Lap Timing: resolution of 0.01 s; accuracy of 0.01 s.

Figure 4.6 Performance-box installed on the window of tested car.
(Photo credit by Yuanyuan Fan)

Considering the VISSIM simulation uses car following algorithm, our testing also tries to follow its leading vehicle to collect the running data. At the experiment intersection, the test car follows other cars. Under this rule, several test runs were performed in the study area using the Performance-box. A similar field data collection method has been used in data collection of field car following trajectories in Beijing (Song et al. 2013).

Figure 4.7 shows the speed plot of one extracted data specifically for the study, which also shows the start and end point of the study intersection. The red plot is the speed record of the vehicle, the x-axis refers to the time in 0.1 second unit, and the y-axis refers to the speed value (km/h). When the speed is zero, this means the vehicle is idling. The plot on the right is the topographic map of the running route, in which the x-axis and the y-axis refer to the distance from the start point.

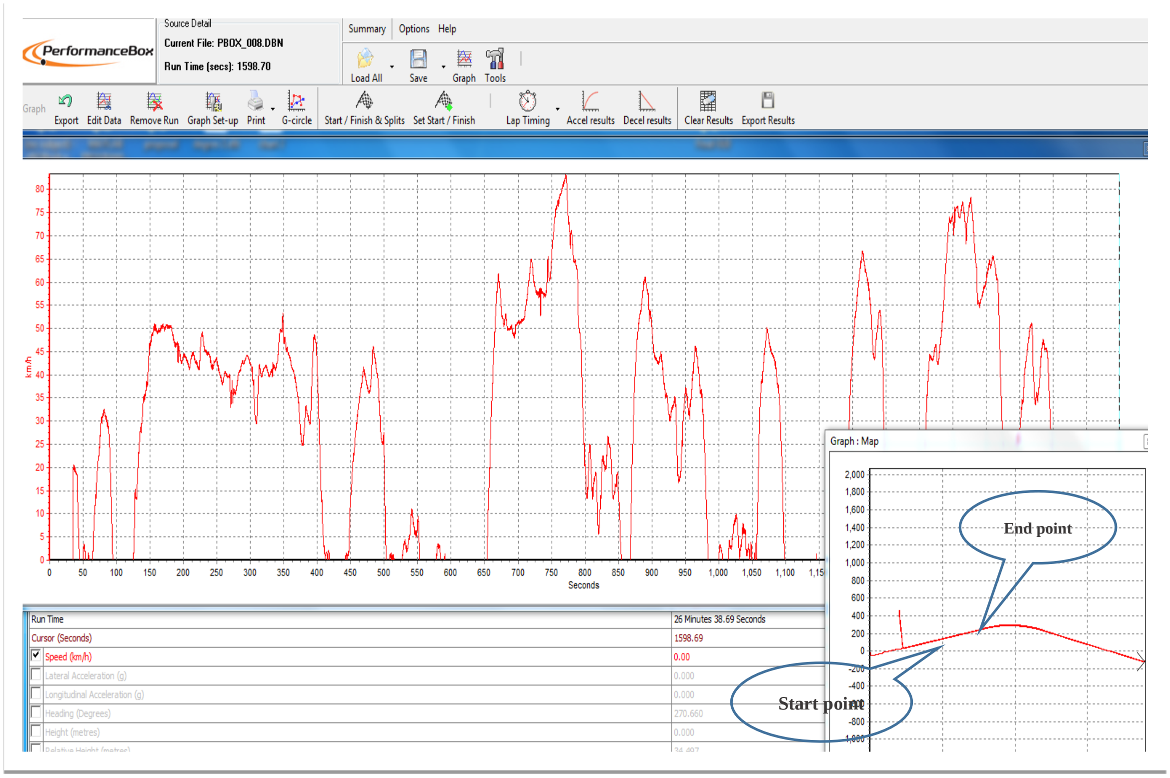


Figure 4.7 Performance-box measured vehicles speed and trajectory plot.

The procedure was used to obtain the field measurement of five operation modes. Approximately one hour of tests were performed for every approach to the intersection, such as from south to north, from north to south, from east to west, west to east, from south right turn to east, etc. The variables collected from the field including speed, location, and acceleration. Further data processing was performed by which we used the recorded speed and acceleration data of vehicles to determine operation modes fractions. Next, we averaged the collected data of the five operation modes fractions in every direction. Further, the averaged field measured operation modes fractions in each direction were weighted based on the volume to capacity ratio as shown in Figure 4.4. Finally, the weighted operation modes fractions in each direction were summed and the five operation modes fractions at the intersection in one-hour period were obtained.

4.2.4 Comparison of the Simulated and the Field Operation Modes

The comparison of the simulated operation modes fractions by VISSIM and the field measured operation modes fractions are compared and illustrated in Figure 4.8.

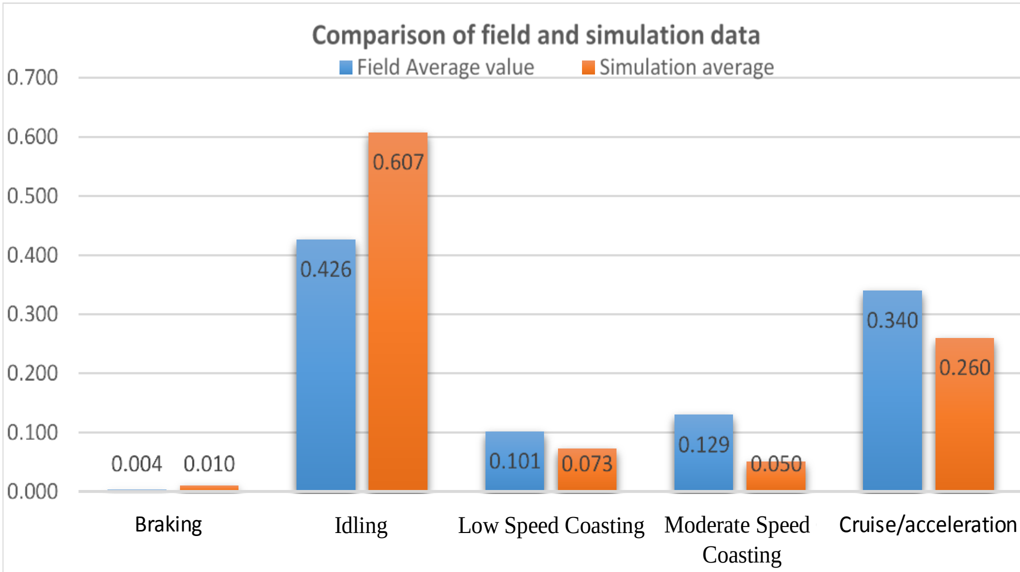


Figure 4.8 Comparison between VISSIM simulated and field measured operation modes fractions.

Further, in order to validate the similarity of the two kind's data, the Chi-Square Test and P-value test were applied. The Chi-Square Test can determine how closely the field collected operation mode as compared to the VISSIM generated data. The calculated Chi-square value between the observed (field) data and expected (VISSIM) data is 0.304147. Compared to the Chi-square table, which has a degree of freedom of 4, for a significance level of 5%, the Chi-square table value is 9.488. The calculated value is smaller than the Chi-square critical value for a significance level of 0.05. Therefore, the null hypothesis is not rejected. The field-collected data is statistically close to the VISSIM simulation data.

4.3 Data Generation and Correlation Analysis

At the data generation stage, the volumes used in VISSIM to simulate the studied intersection ranges from 200 to 880 vehicles per hour (Level of Service A to E). The detailed experimental data is listed in the Appendix. The entire volume range was used to generate the data. Traffic simulation data is obtained from VISSIM 6 and emissions outputs is obtained from MOVES 2014a. MOVES 2014a is the newest EPA authorized vehicles emission estimation tool. The output of VISSIM is the speed and acceleration of every car at every second during an hour period. Then, through calculation, we obtain the percentages of five operation modes, which is further input into MOVES. Additional information obtained from MOVES also includes: time period, county, link length, link volume, road type, road grade, temperature, humidity, vehicles' age, vehicle type, and fuel formula. Next we get the emissions of the four pollutants. Vehicle operation modes as well as other traffic-related variables will be used as independent variables respectively to predict the emissions of vehicles. The relationship between pollutant emissions, i.e., CO, NO, NOx, and independent variables are analyzed using the two types of models previously described.

4.3.1 Operation Mode Data as Variable Model

Two groups of independent parameters are used to develop vehicles emissions estimation models. Parameters in group one used the five operation modes to predict four types of emissions. The purpose of using operation mode to evaluate emissions is to determine how operation mode contributes to the four types of emissions under the different conditions. Figure 4.9 shows the change of operation mode under different grade and volume conditions. It is noted that as volume and grade increased, each operation mode contributes different fraction to the overall five operation modes. But the total operation modes in each condition is sum as 100%.

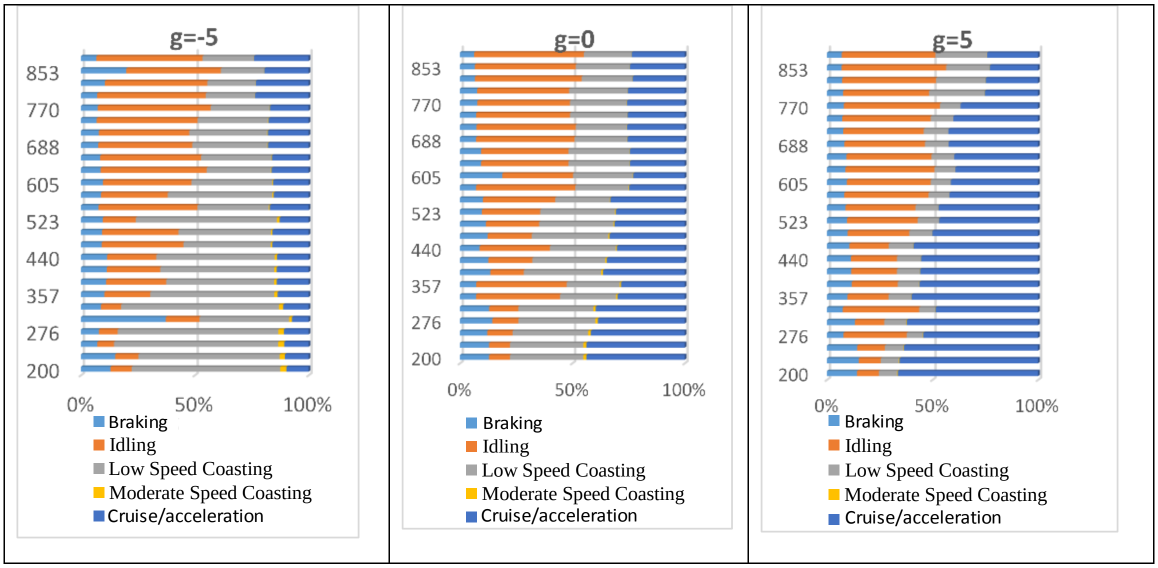


Figure 4.9 Fractions distribution of five operation modes under different vehicle volumes

The figure shows for a downgrade of 5 %, that the majority of vehicles operate under "Low speed coasting" with this percentage decreasing as the volume increases. As the volume increases the largest percentage of vehicles operate under "Idling". These results are intuitive as more vehicles added to the intersection will reduce speeds resulting in more vehicles idle vehicles. For the condition where grade is 0%, the largest percentage of vehicles is operating at cruise acceleration at low volume. As volume increases, the largest percentage of vehicles is operating at idling. At a grade of 5%, the largest percent of vehicles is operating at cruise acceleration. Similar to the other grades, as the grade increases, the largest percentage of vehicles is operating at idling.

4.3.2 HCM Variable and Correlation

Using HCM related parameters to evaluate emissions is aimed to develop a MOVES-like estimation model of vehicle emissions for pollutants at signalized intersections with HCM parameters. In HCM 2010, the parameters that used to estimate delay at signalized intersection includes several variables: volume to capacity ratio (v/c), green time to cycle length (g/c), number of lanes, lane width, heavy vehicles, approaching grade, adjustment of existence of parking adjacent to the lane group, area type, lane utilization, left turn lane characteristics, right turn lane group characteristics, presence of pedestrians, and for pedestrian and bicycle conflicts. Among all these variables, volume to capacity ratio, green time to cycle length, approaching grade, and delay are considered as potential variables. A correlation analysis was performed to develop emission models with variables that would not result in multicollinearity. Further discussion of the correlation analysis is provided below. At the completion of determining the correlation between the independent variables, a model containing volume to capacity together with grade had the highest R^2 for the four emissions models.

Four types of emissions are used as the dependent variables, with corresponding potential HCM related variables as mentioned above. Table 4.1 shows the range of variables used in generating the emissions models. MOVES generated CO is in the range of 0.4kg/hr-1.8kg/hr kg per hour. NOx emissions range is between 0.5kg/hr-2.1kg/hr. NH_3 is in the range of 0.2kg/hr-0.9kg/hr. NOX is between 0.5kg/hr-2.1kg/hr. Based on the delays at the intersection for the volume conditions, the intersection data are for level of service conditions between A and E. Volume to capacity ratio ranges from 0.28 to 1.02. Delay time from 7.7 sec/veh to 59.1 sec/veh, and approaching road grade is from -5 to 5. It is noted that default simulation speed is 30 mph, turning percentage is 5%, and no heavy vehicles are accounted in the vehicle flows.

Table 4.1 Description of HCM related variables and Emissions

	Description	Number
Dependent Variables	MOVES output	74 sets
NO	Tail pipe	0.4kg/hr-1.8kg/hr
NH3	Tail pipe	0.2kg/hr-0.9kg/hr
NOX	Tail pipe	0.5kg/hr-2.1kg/hr
CO	Tail pipe	8.2 kg/hr-79kg/hr
Independent Variables	HCM related	74 sets
Delay		[7.7sec/veh,59.1 sec/veh]
V/c	Volume to capacity ratio	[0.28,1.02]
g/C	green to Cycle	[0.43,0.47]
Road Grade		[-5, 5]
Cycle length		[40 sec to 100 sec]
Continuous Variable		
Desired speed		30mph
Turning percentage		5%
Truck Percentage		0%

4.3.3 Correlation Matrix Study

The correlation coefficient is a way to describe how closely two variables are related. It is an indicator reflecting the strength of the linear relationship between two physical characteristics. The correlation coefficient is a numerical measure of two variables' linear strength. The correlation coefficient is usually denoted as r. The correlation coefficient is always between -1 and 1. A positive value indicates that the least-square line has a positive slope, whereas a negative slope shows two variables have a negative relationship. A value of the correlation coefficient close to -1 or 1 indicates that a strong linear relationship exists between two variables. If the value of correlation coefficient close to 0, it indicates weak relationship between two physical characteristics. If r=0, two variables are said to not have any relationship. Table 4.2 to 4.5 illustrates the correlation matrix analysis for the variables under study. The correlation analysis is made between four types of pollutants, road grade, delay, v/c, g/c and cycle length. In Table 4.2, the correlation between NOx, grade, delay, v/c, cycle length and g/c is determined. The correlation table shows cycle length and volume to capacity ratio having a strong linear relationship. The g/c and cycle length have a strong relationship. The g/c also has a strong relationship with v/c. Road grade has very weak relationship with cycle length, v/c and g/c. Delay has a strong correlation with cycle length, v/c, and g/c, but has a very weak linear relationship with road grade. The last column of table 4.2 shows that NO_X emission has a strong correlation with cycle length, v/c, g/c and delay. In addition, v/c and NO_X exist the strongest linear relationship with each other. Therefore, as independent variables, grade together with v/c are available to build the model, which also is believed to provide the highest R^2 among these parameters.

Table 4.2 NO_X Correlation Matrix

	Cycle Length	V/c	g/C	Road Grade	Delay	NO_X
Cycle Length	1	.820	.923	.072	.976	.847
Volume to Capacity		1	.822	.054	.820	.996
Green to Cycle			1	.044	.883	.850
Road Grade				1	.092	.005
Delay					1	.843
NO_X						1

The second correlation analysis is between NH_3, grade, delay, v/c, cycle length and g/c. The correlation table shows that v/c and cycle length are strongly correlated. The g/c and cycle length, volume to capacity ratio also has strong correlation coefficients. Road grade is weakly correlated with cycle length, volume to capacity ratio, and green to cycle length. Delay is strongly correlated with cycle length, volume to capacity ratio, green to cycle length but does have a weak correlation with road grade. NH_3 shows a strong linear relationship with cycle length, volume to capacity ratio, g/c and delay. V/c and NH_3 also exist the strongest linear relationship with each other. Therefore, as independent variables, grade together with v/c are available to build the model, which is also believed to provide the highest R^2 among these parameters.

The third correlation analysis is between NO, grade, delay, v/c, cycle length and g/c. The Table 4.4 shows V/c and cycle length are strongly correlated. The g/c and cycle length, volume to capacity ratio also has strong correlation coefficients. Road grade is weakly correlated with cycle length, volume to capacity ratio, and green to cycle length. Delay is strongly correlated with cycle length, volume to capacity ratio, green to cycle length but has a weak correlation with road grade. NO shows a strong linear relationship with cycle length, volume to capacity ratio, g/c and delay. V/c and NH_3 also exist the strongest linear relationship with each other. Therefore, as independent variables, grade together with v/c are available to build the model, which is also believed to provide the highest R^2 among these parameters.

The fourth correlation analysis is between CO, grade, delay, v/c, cycle length and g/c. V/c and cycle length are strongly correlated. The g/c and cycle length, volume to capacity ratio also has strong correlation coefficients. Road grade is weakly correlated with cycle length, volume to capacity ratio, and green to cycle length. Delay is strongly correlated with cycle length, volume to capacity ratio, green to cycle length but does have a weak correlation with road grade. CO shows medium correlation with cycle length, volume to capacity ratio, g/c grade and delay. The order of the strength is road grade, v/c, g/c, cycle length and delay. It illustrates in Table 4.5.

Table 4.3 NH_3 Correlation Matrix

	Cycle Length	V/c	g/C	Road Grade	Delay	NH_3
Cycle Length	1	0.820	0.923	0.072	0.976	0.847
Volume to Capacity		1	0.822	0.054	0.820	0.996
Green to Cycle			1	0.044	0.883	0.849
Road Grade				1	0.092	0.095
Delay					1	0.846
NH_3						1

Table 4.4 NO Correlation Matrix

	Cycle Length	V/c	g/C	Road Grade	Delay	NO
Cycle Length	1	0.820	0.923	0.072	0.976	0.847
Volume to Capacity		1	0.822	0.054	0.820	0.996
Green to Cycle			1	0.044	0.883	0.850
Road Grade				1	0.092	0.005
Delay					1	0.843
NO						1

Table 4.5 CO Correlation Matrix

	Cycle Length	V/c	G/C	Road Grade	Delay	CO
Cycle Length	1	0.82	0.923	0.072	0.976	0.362
Volume to Capacity		1	0.822	0.054	0.82	0.494
Green to Cycle			1	0.044	0.883	0.38
Road Grade				1	0.092	0.64
Delay					1	0.35
CO						1

CHAPTER 5 Emission Estimation Model Results

The previous chapter describes the variables to be used in estimating vehicle emissions. This chapter presents the model development and their results. Three emissions models are estimated. Model 1 estimates vehicle emissions using operation mode variables using MLR. Model 2 estimates vehicle emissions based on HCM related variables by using MLR. Model 3 estimates vehicle emissions based on HCM related variables by using ANN modeling. The coefficients of Model 1 can be used to determine the contributing weight of each operation mode on vehicle's emission. Models 2 and 3 are intended to estimate and predict vehicles emissions. MOVES generated emissions are treated as the observed emissions quantities. Emissions estimated using Models 2 and 3 are treated as the expected values. The observed emissions from MOVES and the estimated emissions from Models 2 and 3 are tested by Chi-square test to determine how closely the models fit the observed data. The MLR models are verified by F test.

5.1 Multiple Linear Regression Models Development

5.1.1 Operation Modes and Traffic Related Parameters as Independent Variables

The model was fit using Minitab 17 Statistics software. Model 1, which estimated vehicle emissions using operation mode as the independent variables is shown in Table 5.1. Volume to capacity ratio and grade shows the influences on traffic related variables to pollutant emission's generation. In the other model, which illustrated in Table 5.2, we take the fractions of five operation modes as the independent variables, and also find their influences on four kinds of emissions including NO, NO_X, NH_3, and CO as the dependent variables.

The obtained coefficients of the independent variables in Models 1 and 2, where multiple linear regression is used in estimating the models, are used to explain the influence of vehicle operation modes on pollutant emissions. In this test, the P value is used to determine significant variables. "The P-value is defined as the probability under the assumption of hypothesis H_0 of obtaining a result equal to or more extreme than what was actually observed. The smaller the P-value, the larger the significance because it tells the investigator that the hypothesis under consideration may not adequately explain the observation. Then, based on the obtained P-value, we can determine whether the factors are significantly related to the emissions outputs or not. The hypothesis H_0 is rejected if any of these probabilities is less than or equal to a small, fixed but arbitrarily pre-defined threshold value α, which is referred to as the level of significance." (Bhattacharya and Bhaskar; Habtzghi 2002, Hung et al 1997). In this research, the level of significance is set to 0.05. A P-value of of less than 0.05 will result in H_0 being rejected. From Table 5.1, we can find that for NO, NO_X and CO, grade is not significant to these emissions. Volume has significant influence on each type of emissions.

Table 5.1 HCM Related Variables' Significance for NO, CO, NH3 and NO_X

Variable	Coef	T-Value	P-Value	Variable	Coef	T-Value	P-Value
		NO				CO	
Grade	-0.001	-0.670	0.504	Grade	-0.189	-0.790	0.431
V/c	0.002	121.600	0.000	V/c	0.033	9.960	0.000
		NH_3				NO_X	
Grade	0.009	10.850	0.000	Grade	-0.001	-0.670	0.504
V/c	0.001	83.490	0.000	V/c	0.002	121.600	0.000

In Table 5.2, it shows five operation mode including Braking, Idling, Low Speed Coasting, Moderate Speed and Cruise/acceleration are used to measure the weight of the drive mode contribution to the emissions. Not every operation mode has a significant influence on pollutant emissions. For NO and NO_X, Idling is not a significant parameter. Moderate Speed is not significant for CO, all operation mode have significant influence on NH_3.

Table 5.2 Operation Mode Variables Significance for NO, CO, NH3 and NO_X

Variable	Coef	T-Value	P-Value	Coef	T-Value	P-Value
		NO			CO	
Braking	-0.0739	-2.91	0.005	-31.48	-6.35	0
Idling	-0.0305	-1.22	0.227	-15.94	-3.27	0.002
Low Speed Coasting	0.0577	2.84	0.006	-14.63	-3.68	0
Moderate Speed Coasting	-0.061	-2.87	0.005	4.75	1.15	0.256
Cruise/acceleration	0.0811	3.16	0.002	32.58	6.49	0
		NH_3			NO_X	
Braking	0.0627	3.66	0	-0.0844	-2.91	0.005
Idling	-0.07	-4.16	0	-0.0348	-1.22	0.227
Low Speed Coasting	-0.0289	-2.11	0.039	0.0659	2.84	0.006
Moderate Speed Coasting	0.0353	2.46	0.016	-0.0696	-2.87	0.005
Cruise/acceleration	-0.0591	-3.41	0.001	0.0927	3.16	0.002

5.1.2 Standardized and Unstandardized Coefficients of Regression Equations for Four Pollutants Emissions

Table 5.1 provides a coefficient analysis that allows one to understand if variables are significant in the multiple linear regression model. Table 5.2 illustrates standardized and unstandardized regression coefficients analysis between pollutants and five operation modes. The non-standardized coefficients are not suitable to horizontally compare the impacts of independent variables. This is because in the non-standardized regression models the independent variables may have different units or distributions. After standardizing the non-standardized regression models, the independent variables have been all transformed into the standard normal distribution N (0, 1), then the obtained standardized coefficients can be used to be horizontally compared, which also means that the standardized coefficients $\hat{\beta}_i^*$ can be directly used to quantitatively evaluate the impact size of each operation mode on the emissions of each pollutant. The larger the absolute value of the standardized coefficients, the greater the impacts of the independent variables on the dependent variables.

The obtained unstandardized and standardized regression coefficients are listed in Table 5.3. By the standardized coefficients, we can evaluate and compare the significance size of each of the five operation modes on four vehicle missions including NO, CO, NH_3 and NO_X. Unstandardized coefficients can be used for emission estimation, but operation mode is not easy to measure in the real world. Therefore, for estimation purpose, we plan to use macroscopic parameters to develop the estimation models. HCM related parameters are used to develop the emissions model and will be discussed in later of this chapter.

Table 5.3 Standardized Coefficients and Unstandardized Coefficients for Operation Mode of MLR

	Unstandardized Coef.	Standardized Coef.	Unstandardized Coef.	Standardized Coef.
	NO		NOx	
Braking	-.063	-.991	-.071	-.991
Idling	/	/	/	/
Low Speed Coasting	-.071	.033	.081	.033
Moderate Speed	-.018	-.132	-.095	-.132
Cruise/acceleration	.102	1.087	.080	1.087
	NH3		CO	
Braking	-.099	-3.368	-11.576	-36.797
Idling	/		-4.510	-.966
Low Speed Coasting	-.071	-.072	-3.761	-.398
Moderate Speed	-.018	-.060	/	/
Cruise/acceleration	.102	3.429	30.630	11.720

The variable coefficients and the R^2 for the models provide regression results of the multiple linear regression models. In Multiple linear regressions, NO R^2 is 0.999, NH3 R^2 is 0.998, NO_X R^2 is 0.999, and CO R^2 is 0.853.

According to the standardized coefficients, for NO and NO_X emission, Cruise/acceleration have the most significant influence on emissions. Braking, Moderate speed, Low Speed Coasting, Moderate Speed Coasting and Braking are next in terms of their influence on emissions. For NH3 Cruise/acceleration has the most significant influence, then in turn Moderate speed, Low Speed Coasting and braking has the next significant influence. For CO emission, Cruise/acceleration has the most significant influence, then in turn idling, and braking has the next significant influence. The findings reveal that Cruise/acceleration has the most significant influences on vehicles emissions. This finding matches the results by Andrew et al. (2012). Their research found that more than 80% of total emission are generated from cruising and acceleration.

5.1.3 Multiple Linear Regression Results Using HCM Related Variables

The results of Model 2 development is shown below in Table 5.4. For NO, the R^2 is 99.47% with all variables having a p value smaller than 0.05. This indicates that there is a small chance that this variable is not a significant variable. The t-statistic calculates the difference between the dependent variable NO and the independent variables. The greater the magnitude of T (it can be either positive or negative), the greater the evidence against the null hypothesis that there is no significant difference between the measured emissions and the prediction variables. The closer T is to 0, the more likely there is not a significant difference. So for all t value in the test, the larger the better. VIF, Variance Inflation Factor, is 1. VIF measures how much the variance (the square of the estimate's standard deviation) of an estimated regression coefficient is increased because of collinearity. It is the reverse of tolerance; it means the level of reliable of variable as independents. VIF >5 is means tolerance is small.

For NH3, the model results are shown below with an R^2 of 99.25%. All variables have p value smaller than 0.05, and VIF is 1.

The model developed for NO is shown below with an R^2 of 99.25%. All variables have a p value smaller than 0.05 and the VIF is 1.

The model development for CO is shown below with an R^2 of 61.56%. All variables have a p value smaller than 0.05, and VIF is 1

Table 5.4 HCM Related Variables Model Development

Depended variables	R^2	Independent Variable	Estimated Coefficient	T-value	P-Value	VIF
NO	99.47%	Constant	-0.058	-3.690	0.001	
		Volume to capacity	1.749	80.200	0.000	1.000
		Grade	-0.004	-3.490	0.001	1.000
NH_3	99.25%	Constant	-0.023	-2.620	0.013	
		Volume to capacity	0.809	0.012	0.000	1.000
		Grade	0.002	3.090	0.004	1.000
NO_X	99.47%	Constant	-0.066	-3.690	0.001	
		Volume to capacity	1.997	80.200	0.000	1.000
		Grade	-0.0048	-3.49	0.001	1.000
CO	61.56%	Constant	9.04	5.49	0.000	
		Volume to capacity	10.49	4.58	0.000	1.000
		Grade	0.701	5.55	0.000	1.000

5.1.4 Analysis of Prediction Equations for HCM Related Variables by Non-Standardized Coefficients

1) For NO emissions we find that the emission of NO can be estimated by the volume to capacity ratio, and grade. These two variables have positive coefficients with the emission of NO, which indicates that these two variables have positive relationship to NO's emissions. It reveals that higher volume to capacity ratio will result in heavier NO emissions. On the other hand, road grade has a negative reaction for emission of NO, which shows that steeper grades will make the emission of NO decrease.

2) For emission of NH_3 two variables including volume to capacity ratio and grade, both variables have positive coefficients for the emission of NH_3, which indicates that these two variables all have a positive relationship to NH_3's emission. It reveals that greater volume, and higher road grade would result in heavier NH_3 emission.

3) For emission of NO_X, we can find that the emission of NO_X can be estimated by

volume to capacity ratio and grade. Among these three variables, volume to capacity ratio has positive coefficients for the emission of NO_X, which indicates that they have positive relationship with the emission of NO_X. On the other hand, road grade has a negative coefficient for emission of NO_X, which indicates that it has negative relationship with NO_X's emission. It reveals that, the grade change would not have an obvious positive increase on NOx generation. Furthermore, volume to capacity ratio has the biggest absolute coefficient, which indicates that volume has the most significant influence on the emission of NO_X.

4) For emission of CO， we find that the emission of CO can be estimated by grade and volume to capacity ratio. Both variables have positive coefficients for the emission of CO, which indicates that all the variables have positive co-relationship with CO's emission. When volume to capacity ratio and road grade increases, then the CO emission will increase.

5.1.5 Standardized Multiple Linear Regressions Experiments

The standardized regression coefficients can be used to reveal the importance of an independent variable's influence on the dependent variable. From Table 5.5, it can be seen that for NO, NH_3 and NO_X, volume to capacity ratio dominates the total quantity of these pollutants. Grade only plays a marginal influence on these three types of emissions. For CO, grade has more weight on measuring the generation of CO emissions with V/c also significant in CO emissions generation. Both variables show positive signs indicating that an increase of V/c and grade will cause the CO emission to be heavier.

Table 5.5 Standardized Coefficients of V/C and Grade

Standardized Coefficients	V/c	Grade
NO	0.999	-0.043
NH_3	0.993	0.046
NO_X	0.999	-0.043
CO	0.488	0.591

5.1.6 Verification of Regression Models by F test

The F statistical values are calculated and used to identify the significance of the above obtained linear regression equations. Table 5.6 shows the statistics of the MLR equations using operation mode as the independent variables.

Table 5.6, the test is used to determine if a null hypothesis that the obtained coefficients in the models are equal to 0. First we have obtained the F value from the SPSS 17.0. The F value is compared to F table using a significance level of 5% and 1%. The critical value of $\mathrm{F}_{(1-\alpha)}(m, n-m-1)$ can be obtained by checking the F-distribution table for a chosen α value. M refers to the number of random errors, which is 5 for the Five Operation Modes regression model and 2 for the traffic Parameters regression model. N is the sample size of the experimental data, which is 78 for operation mode model and 37 for HCM parameter model These values represent half of the data which was used to develop the models. The remaining half is used for model validation. When the value of the significance level α is set to 0.1, the critical value $\mathrm{F}_{(1-\alpha)}(m, n-m-1)$ is 1.93 for the operation mode regression model and 2.46 for the traffic related regression model. When the value of the significance level α is set to 0.05, the critical value of $\mathrm{F}_{(1-\alpha)}(m, n-m-1)$ is 2.5 for the operation mode regression model and 3.31 for the traffic related regression model. From Table 5.6 and Table 5.7, we know that the F statistical values for four pollutants' linear regression models are all higher than of F table value when either set α as 0.1 or 0.05.

Therefore, the models of regression equations are of significance and effective.

Table 5.6 F Statistics of Multiple Linear Regression Equations for NO, NH_3, NO_X, and CO as Operation Mode as Variables

Parameter	Linear regression equations for	SST	SSR	F	$F_{(1-0.1)}(5,72)$	$F_{(1-0.05)}(5,72)$
Five Operation Modes regression	NO	13.110	13.102	23358.251	1.93	2.5
	NH_3	2.820	2.809	3900.653	1.93	2.5
	NO_X	17.093	17.083	23353.438	1.93	2.5
	CO	2097.469	1789.7	87.76	1.93	2.5

Table 5.7 Statistics of Multiple Linear Regression Equations for NO, NH_3, NO_X, and CO as HCM Parameters as Variables

Parameter	Linear regression equations for	SST	SSR	F	$F_{(1-0.1)}(3,74)$	$F_{(1-0.05)}(3,72)$
Traffic parameters (V/c and grade) regression	NO	5.837	5.806	3216.580	2.46	3.31
	NH_3	1.263	1.253	2239.328	2.46	3.31
	NO_X	7.61	7.57	3217.054	2.46	3.31
	CO	879.308	541.277	27.221	2.46	3.31

5.2 ANN Precision Estimation Models

In order to increase the estimation capability of the pollutants emissions, ANN model method are used. The ANN models are developed from the traffic parameters by volume to capacity ratio and grade. By developed ANN model, we can compare the validation R, to find out the closeness of estimated model and experimental results.

5.2.1 Structure Design of ANN Model

The ANN model developed for emission study is designed to be of two and three layers, with the number of neurons in the hidden layer calculated using equation 3.16. The structure includes the input layer, hidden layers and the output layer. The input layer includes the volume to capacity ratio and grade, which refer to the two inputs of the ANN model. The first and second hidden layer transfer function uses a sigmoid function. Three commonly used algorithms include Levenberg-Marquardt, Bayesian regularization and scaled conjugate gradient are tested for the training data. They are denoted as 'trainlm', 'trainbr' and 'trainscg' in the MATALB software. We have 74 data pairs in total for inputs and outputs pair. It is noted that 70% of the data pair are randomly chosen to be in training (52 data pair), 15% for validation and 15% for testing. The validation R is used to compare the goodness of fit of the ANN model, based on the hidden layer, number of neurons and the algorithm. We also noted that training and validation results vary due to the different initial and training sample. Therefore, for 5 times each structure of the ANN model is trained and the average validation R value used for comparison. It is worthy to mention that based on Anderson's (1992) rule for identifying the number of neurons, as given in equation 3.16, α value is suggested to take one to five, ten and twenty based on the level of data's noise. In this study, we took all suggested number of neurons depending on the training size, and the number of inputs and outputs.

5.2.2 ANN Modeling Results Analysis

The results of the four ANN models for each of the four vehicle emissions are illustrated in Tables 5.8 and 5.9. CO shows large difference when compared to the models for NO,

NO_X and NH_3. For this reason, two ANN models are fitted for emission input and output data. One is for CO, and the other is for NO, NO_X and NH_3. For the CO results shown Table 5.8, the number of inputs is 2 and the output is CO only. Based on equation 3.16, the number of neurons in hidden layer can be one, two, three, four, six, nine, and seventeen. Algorithms including Levenberg-Marquardt (lm), Bayesian regularization (br) and scaled conjugate gradient (scg) are tested to train 70% of randomly selected data pairs. The validation R is used to check the effects of fitting the results compared to the training R. The number of hidden layer is tested for one and two. From the Table 5.8, we found using scg algorithm, with two hidden layers and six neurons achieves the best results. The validation R is equal to 0.920822.

Table 5.8 ANN Validation R Results for CO

CO Input and output (2,1)	Description	Algorithm				Algorithm		
		Lm	br	scg		Lm	br	scg
Ni								
1	No. of Neurons in hidden layer (1 hidden layer)	0.729244	0.741072	0.865533	No. of Neurons in hidden layer (2 hidden layer)	0.865636	0.790968	0.828442
2		0.789525	0.863229	0.798139		0.849978	0.811972	0.882938
3		0.855642	0.84032	0.836659		0.870412	0.820008	0.902746
4		0.819807	0.858684	0.880127		0.848866	0.886542	0.80915
6		0.853736	0.854618	0.912984		0.878188	0.87126	0.920822
9		0.866419	0.906302	0.893705		0.88321	0.874246	0.895124
17		0.82661	0.906648	0.809937		0.798552	0.908492	0.895236

The ANN fitting model is applied for NO, NO_X and NH_3 input-output data pair, for these three pollutant emissions have the similar tendency, therefore one model was used to fit all three emissions. The number of inputs are two and outputs are three. Based on equation 3.16, number of neurons in the hidden layer could be one, two, three, five and ten. The optimized fitting was found using the scg algorithm, two hidden layers and one neuron in each layer. The optimized fitting model has validation R of 0.99925 based on Bayesian regularization with hidden layer of two and each layer of two neurons.

Table 5.9 ANN R^2 Results for NO, NO_X and NH_3

NO, NOX,NH3 Input and output (2,3)			Algorithm Lm	br	scg		Algorithm Lm	br	scg
Ni	1	No. of Neurons in hidden layer (1 Hidden layer)	0.99907	0.99907	0.99790	No. of Neurons in hidden layer (2 hidden layer)	0.99915	0.99868	0.99491
	2		0.99605	0.99888	0.99801		0.99915	$\boxed{0.99925}$	0.89976
	3		0.99927	0.99906	0.99573		0.99857	0.99910	0.95932
	5		0.99919	0.99853	0.99559		0.99854	0.99870	0.99123
	10		0.99848	0.99876	0.99644		0.99917	0.99870	0.99262

5.2.3 GUI Design for the ANN Model from HCM Related Parameters

The results of the obtained ANN model show that the fitting accuracy is high. Hence, by this ANN model, if we import the measureable grade and calculated V/c, the emissions of the four pollutants can be precisely predicted.

Therefore, based on ANN model, a friendly graphic user's interface (GUI) is developed. The purpose is to provide traffic engineers an access to estimate environmental problems caused by traffic parameters. Without traffic simulation software such as VISSIM and MOVES, the developed GUI can still provide a quick and accurate emissions.

The experiment data used in development of the ANN model is based on traffic simulation data and the EPA MOVES 2014. The data base in MOVES is from lab testing and authorized through the nation and states. But this ANN model also has its limitation. Based on the factors of analyzing signalized intersection of HCM 2010, several factors were not included as variables in the analysis including: percentage of heavy vehicles in traffic stream; existence of a parking lane and parking activity adjacent to lane group; blocking effect of local buses that stop within intersection area; area type; lane utilization;

left and right-turn vehicle presence in a lane group. All these variables can be considered for inclusion in a vehicle emissions model in future work.

5.3 Validation and Comparison of MLR and ANN Models

There are a total of 74 groups of data involved in the model development and the validation. 70% of data are used to develop the ANN model and the other 30% used for validating and testing the model. The MOVES generated emissions are compared with the developed ANN model's results. Root-mean-square error (RMSE) is the measure to indicate the difference between the experimental emissions data and ANN model predicted emissions. The smaller RMSE indicates the better fitting effect. It shows the predicted values are closer to the experimental outputs.

Previously, we also use MLR to estimate four types of emissions. Therefore, it is useful to compare two types of model's RMSE value, so that one can understand which models are better tools for estimating emissions. Table 5.10 shows the RMSE values for each pollutant by using two different methods: MLR and ANN. It is found that the ANN model have lower RMSE compare to the MLR model. Therefore, the ANN model is more accurate for emissions estimation than MLR model.

Table 5.10 RMSE Values for Each Pollutant by Using MLR and ANN

	ANN		MLR	
	Validation R	RMSE	R^2	RMSE
CO	92.08%	1.6632E+00	61.56%	5.91770
NO	99.93%	1.2237E-02	99.25%	0.84292
NOX	99.93%	1.2237E-02	99.47%	4.76456
NH3	99.93%	1.2237E-02	99.25%	0.44941

CHAPTER 6

CONCLUSIONS

6.1 Conclusion

The motivation of this dissertation is to incorporate vehicles' emission estimation into the Highway Capacity Manual. The Highway Capacity Manual is a national and international resource that has become a guideline for evaluating the operation of roadway, transit and pedestrian facilities. The operation of the roadway is described based on a performance measure that describes the roadway's operation based on the perception of its users. The performance measures are used describe the traffic operation of the roadway. At present, no measures are provided to describe the operation of the roadway based on the environmental impacts. The incorporation of air pollution estimation into the Highway Capacity Manual is believed to create a sustainable development for both transportation and environment. In addition, The HCM is recommended by the EPA to predict vehicles' speeds in the estimation of emissions (HCM, 2010). The HCM 2010 makes some references about air quality that stating "vehicle emissions are a significant contributor to poor air quality", and referring to the Clean Air Act Amendments CAAA (HCM, 2010). Furthermore, the report *Extent of Highway Capacity Manual Use in Planning* (Dowling, 2012) expects air impact analysis to be ultimately included in the HCM.

The objective of this dissertation was to develop MOVES-like estimation models of vehicle emissions for pollutants at a signalized intersection. A thorough understanding is needed about what parameters, and the influence of these parameters on emission quantity. This dissertation develops two kinds of models in order to make estimation of emissions caused by on-road vehicles. Two modeling approaches are used to estimate four kinds of emissions including CO, NO, NH_3 and NO_X separately. The following points conclude the work of this dissertation: (please correct the following based on the corrections already made for the abstract)

(1) Two modeling approaches are used to estimate vehicle emissions including: multiple linear regression and Artificial Neural Network (ANN). In the multiple linear regression modeling, two different models were developed including one model using operation modes as independent variables and another model using traffic related parameters as independent variables. Both model approaches and independent variables are used to estimate four types of pollutant emissions. Statistically, the emission models using traffic parameters as independent HCM related parameters are capable of providing a better emissions estimate based on the higher R square value. For CO, the variables found to be significant were volume to capacity ratio and grade with an R^2 of 61.56%. For NO, the variables found to be significant were volume to capacity ratio and grade with an R^2 of 99.47%. For NOx, the variables found to be significant were volume to capacity ratio and grade with an R^2 of 99.47%. For NH3, the variables found to be significant were volume to capacity ratio and grade with an R^2 of 99.25%. This study shows that volume to capacity dominate the emissions quality at a signalized intersection. The research found that for NOx, Idling and Moderate Speed Coasting were significant. For NH3, all variables were significant except Low Speed Coasting. For CO, Braking and Cruise/Acceleration were significant. It was also found that longer delay time reduces CO emissions, but it causes the other three pollutant emissions increase.

(2) The ANN modeling method using the Levenberg-Marquardt method was used to train the HCM related variables and MOVES emissions outputs. The parameters of volume to capacity ratio, and road grade are used to estimate emissions. The Validated R value of the obtained ANN model is found.

6.2 Significance of Research

The research performed in this dissertation fills the gap and allows researchers and engineers to use HCM related variables to estimate vehicle emissions. The incorporation of vehicles emission into the HCM can benefit transportation planners and traffic engineers evaluating the performance of a roadway to capture environmental influences based on on-road traffic. There are also more detailed findings and significance of that are found in the following paragraphs.

The resulting models are based on a large dataset that included microscopic data under the various volume scenarios on second-by-second basis for each vehicle. The classification of operation modes are based on VSP and speed, the standard is EPA MOVES. It helps one to understand how each operation mode could have influence on emissions' generation.

The data source used to develop the emissions model is based on a rich data set. The models developed in this research are based on a significant dataset gathered through both simulation data and field studies. VISSIM was run for 25 volume conditions with these volumes ranging between 220 vph per lane group to 880 vph per lane group. The volume ranges were established to ensure that the intersection would be evaluated at all LOS between A and E. These volumes were then run under three road grade levels at g=-5%, g=0% and g=+5%, resulting in 75 scenarios of volume and grade that was run in VISSIM. This resulted in a total of 165,060 vehicles being simulated during these 75 scenarios.

Speed and acceleration data were captured for every vehicle on a second-by-second basis for 3600 seconds or for one hour. For a total of 5.94 million seconds of vehicle data. Using each vehicle-second speed and acceleration, the vehicle's operation mode was then determined. The use of this activity data to generate vehicle emissions model, rather than the use of average speed, results in emissions models that are based on a rich set of microscopic data.

As previously mentioned, the volumes used in the analysis were selected to ensure all levels of service were captured in the development of the emissions model. Prior to using VISSIM to generate speed and acceleration data for each of the volume scenarios, the intersection's signal timing was optimized for each volume condition using Synchro. This required running Synchro 75 times to gather the signal timing information under each volume condition. The result was the ability to run VISSIM using signal timings that were based on optimal signal timing condition. Without performing this step, the speed and acceleration data gathered from VISSIM may have been influenced by more than the volume conditions. If poor signal timing was used, the speed and acceleration data would have had some impact on these values.

Despite the widespread use of the HCM, it is limited in its ability to allow traffic engineers to be able to estimate the performance of the roadway using the same types of data used for performing an operational roadway analysis. This research fills this gap and makes the estimation of vehicle emissions for feasible for the types of users currently performing operation analysis. This is a significant accomplishment that has not before been done. In the past, environmental engineers worked on vehicle emissions and traffic engineers worked on roadway operation. With the inclusion of the models developed within this dissertation, traffic engineers are more readily able to estimate vehicle emissions using models that have been validated and based on a rich data set.

The models developed were based on operational modes, which has not before been performed. From the study, it is noted that as volume and grade increased, each operation mode contributes different fraction to the overall five operation modes. But the total operation modes in each condition are sum as 100% in each situation. It is found the cruise and acceleration mode greatly increased when the approaching road grade increases. This phenomenon is especially obvious when lane group volume is heavy. Andrew et al. (2013) found that emissions generated from cruise and acceleration takes up to 80% of the total.

The other finding of the research is that we have found the most significant variables to reflect the influence of on-road vehicles to the environmental are volume to capacity ratio and the approaching road grade, for a signalized intersection, with signal timing optimized, based on the various volume and default parameters.

An ANN model has seldom to be seen in the existing research for the purpose of estimating emissions at signalized intersections with traffic related parameters. This study also proved that ANN provides a more accurate emission prediction model compared with another model used in this study. It is especially obvious for CO and NOx. Grade as a significant variable, it was over simplified in the existing emissions study, but this research found emission is significant for a vehicle's emission generation. Grade is especially more significant in estimate CO, from the MLR coefficient can be seen this.

6.3 Future Work

The models in this dissertation are built by using both linear and non-linear regression models. Model 1 is based on the variables of five operation mode, and Model 2 is based on the HCM2010 related parameters for analyzing signalized intersection. For HCM traffic related parameters, several factors were not included as variables in the analysis including: percentage of heavy vehicles in traffic stream; existence of a parking lane and parking activity adjacent to lane group; blocking effect of local buses that stop within intersection area; area type; lane utilization; left and right-turn vehicle presence in a lane group. All these variables can be studied for inclusion in the models for estimating vehicle emissions in a future work. A simple plan for the future work is as follows:

- Bring heavy traffic into the vehicle flow, change the percentage to find the influence on emissions.
- Determine if existence of a parking lane and parking activity adjacent to lane group could have an effect on vehicles' speed and acceleration so that influence operation mode and emissions generation.
- Determine if the blocking of local buses that stop within intersection area could impact the speed and acceleration of vehicles, and in so doing influence the total emissions generation.
- Take into consideration of area type changes for total emissions study
- Determine if left and right-turn vehicle presence changes in a lane group would cause vehicles operation mode to change so that it influences the total emissions.
- Revise the default value of MOVES input file such as vehicles age distribution, to find out if the increase aged vehicles can increase the vehicles emissions.
- Change the desired speed in signalized intersection to find out if speed limit is a significant factor for estimation emissions.

Taking into consideration each parameter above to determine its impact on emissions analysis and modeling would better satisfy different kinds of estimation needs when estimating environmental influences caused by the traffic side.

APPENDIX A

EMISSISON DATA AND POTENTIAL VARIABLES

Table A.1 and Table A.2 show the experimental data. The first table is the emission data and traffic related data. The later one is the operation mode and vehicles emission data.

Table A.1 Experimental Data for Traffic Related Parameters and Vehicles Emissions

v	Cycle Length	g/C	v/c	Delay	G	(NO)	(NH3)	(NOx)	(CO)
220	40	0.43	0.29	7.7	5	0.463668	0.237903	0.529431	15.01989
220	40	0.43	0.28	7.7	-5	0.464009	0.210749	0.52982	8.239245
220	40	0.43	0.28	7.7	0	0.46817	0.212639	0.534571	8.827985
247	40	0.43	0.33	7.8	5	0.519541	0.265037	0.593228	16.23091
247	40	0.43	0.32	7.8	0	0.525255	0.238567	0.599753	9.530512
247	40	0.43	0.31	7.8	-5	0.538149	0.244423	0.614476	12.09101
276	40	0.43	0.36	8	5	0.576491	0.279666	0.658255	14.07557
276	40	0.43	0.35	7.9	0	0.581849	0.264272	0.664374	9.827095
303	40	0.43	0.38	8	-5	0.582976	0.264783	0.665661	8.420522
276	40	0.43	0.34	7.9	-5	0.598941	0.272034	0.68389	13.00839
303	40	0.43	0.4	8.1	5	0.638881	0.323195	0.729495	19.21658
303	40	0.43	0.38	8.1	0	0.64047	0.290896	0.73131	10.93906
330	40	0.43	0.42	8.2	0	0.681685	0.309616	0.778369	8.240266
330	40	0.43	0.43	8.3	5	0.682402	0.330248	0.779188	14.24361
330	40	0.43	0.41	8.2	-5	0.713284	0.323968	0.81445	15.12046
357	40	0.43	0.46	8.4	0	0.733377	0.333094	0.837392	8.438867
357	40	0.43	0.45	8.3	-5	0.752252	0.341666	0.858945	11.28802
357	40	0.43	0.47	8.5	5	0.754783	0.371317	0.861835	21.55297
385	40	0.43	0.48	8.5	-5	0.799976	0.363342	0.913437	10.11465
385	40	0.43	0.51	8.6	5	0.804719	0.402709	0.918853	21.1053
385	40	0.43	0.49	8.6	0	0.809494	0.367666	0.924307	12.91961
413	40	0.43	0.52	8.7	-5	0.861826	0.391435	0.98406	11.66861
413	40	0.43	0.53	8.7	0	0.863846	0.392352	0.986366	12.90131
413	40	0.43	0.54	8.8	5	0.864173	0.431566	0.986739	22.23572
440	40	0.43	0.56	8.9	0	0.913799	0.415041	1.043404	12.10729
440	40	0.43	0.55	8.8	-5	0.920804	0.418221	1.051402	13.10663
440	40	0.43	0.58	9	5	0.920917	0.45937	1.051531	23.67233
468	40	0.43	0.59	9	-5	0.962517	0.437167	1.099032	10.39211
468	40	0.43	0.6	9.1	0	0.980054	0.445133	1.119057	14.5261
468	40	0.43	0.62	9.2	5	0.987665	0.489508	1.127746	27.73532
495	40	0.43	0.62	9.2	-5	1.02179	0.464089	1.166713	11.55475
495	40	0.43	0.65	9.8	5	1.02808	0.506803	1.173894	23.33692

495	40	0.43	0.64	9.4	0	1.031747	0.468612	1.178081	14.40619
523	45	0.43	0.67	10.8	5	1.079074	0.53102	1.232119	22.20087
523	40	0.43	0.67	10.1	0	1.092673	0.496283	1.247648	15.34671
523	40	0.43	0.65	9.7	-5	1.115922	0.506843	1.274197	20.18813
550	40	0.43	0.69	10.7	-5	1.121224	0.50925	1.280248	11.13961
550	45	0.43	0.71	12.1	5	1.138293	0.556848	1.299738	23.78456
578	45	0.43	0.73	12.6	0	1.179247	0.535604	1.3465	12.98772
578	45	0.43	0.75	13.3	5	1.183981	0.576743	1.351906	21.18441
578	45	0.43	0.71	11.8	-5	1.204396	0.547027	1.375217	15.20547
605	45	0.43	0.76	13.8	0	1.209963	0.549556	1.381573	12.06755
605	55	0.455	0.76	14.9	5	1.23425	0.607431	1.409305	21.27041
605	45	0.43	0.74	13	-5	1.235039	0.560944	1.410205	12.47227
633	45	0.425	0.78	14.3	-5	1.278816	0.580828	1.460193	11.82888
633	55	0.455	0.79	16.2	5	1.288692	0.630874	1.471468	21.15264
633	55	0.445	0.77	15.3	0	1.293349	0.587428	1.476787	14.46487
660	55	0.455	0.79	15.8	-5	1.339804	0.608528	1.529831	12.88652
660	55	0.455	0.83	17.8	5	1.345992	0.661457	1.536893	22.8075
660	55	0.445	0.81	16.7	0	1.348384	0.612425	1.539626	15.07919
688	55	0.445	0.84	13.8	0	1.404474	0.637901	1.603672	16.39538
688	55	0.46	0.79	19	-5	1.408917	0.63992	1.608747	14.80764
688	55	0.445	0.86	19.8	5	1.413561	0.688784	1.614048	26.37711
715	55	0.445	0.88	20.6	0	1.457633	0.662045	1.664371	16.99346
715	55	0.47	0.85	18.8	-5	1.467063	0.666328	1.675139	15.6767
715	65	0.455	0.88	22.8	5	1.472072	0.71401	1.680858	27.89661
743	65	0.455	0.87	21.6	-5	1.516892	0.688961	1.732035	15.44825
743	70	0.46	0.91	26.4	5	1.521844	0.735585	1.73769	27.08645
743	70	0.46	0.89	23.7	0	1.522252	0.691394	1.738154	17.98499
770	70	0.46	0.9	24.5	-5	1.554034	0.705829	1.774444	14.59476
770	65	0.455	0.95	31	5	1.562209	0.760127	1.78378	24.86989
770	65	0.46	0.93	27.3	0	1.576211	0.715902	1.799768	16.16398
798	65	0.455	0.94	28.4	-5	1.617757	0.734772	1.847206	17.53451
798	70	0.455	0.95	32.4	0	1.634897	0.742556	1.866776	19.25528
798	75	0.46	0.97	37.2	5	1.635689	0.793093	1.867682	18.71535
825	75	0.46	0.95	33.6	-5	1.659786	0.753861	1.895195	17.47304
853	75	0.46	0.99	39.8	-5	1.66693	0.757106	1.903354	14.35942
825	75	0.46	0.98	38.4	0	1.675899	0.76118	1.913595	18.07855
825	75	0.46	1	44.1	5	1.682098	0.812779	1.920673	18.85347
853	100	0.47	1.01	51.4	5	1.72461	0.832168	1.969215	17.86352
853	100	0.47	1	46.2	0	1.741325	0.790896	1.988299	19.65795
880	100	0.48	1.02	52.6	0	1.785516	0.810968	2.03876	19.6899
880	100	0.47	1	46.6	-5	1.78982	0.812922	2.043673	19.52345
880	100	0.47	1.05	59.1	5	1.79583	0.86638	2.050535	19.71665

Table A.2 Experimental Data for Operation Modes and Vehicles Emissions

g	Volume	Breaking	Idling	Low Speed Coasting	Moderate Speed Coasting	cruis/acceleration	(NO)	(NH3)	(NOx)	(CO)
0	200	0.1298	0.0930	0.3265	0.0134	0.4374	0.4256	0.1933	0.4859	8.0239
0	220	0.1298	0.0930	0.3265	0.0134	0.4374	0.4682	0.2126	0.5346	8.8280
0	247	0.1216	0.1128	0.3360	0.0119	0.4177	0.5253	0.2386	0.5998	9.5305
0	276	0.1440	0.1162	0.3422	0.0113	0.3862	0.5818	0.2643	0.6644	9.8271
0	303	0.1302	0.1295	0.3342	0.0102	0.3959	0.6405	0.2909	0.7313	10.9391
0	330	0.0719	0.3736	0.2492	0.0066	0.2986	0.6817	0.3096	0.7784	8.2403
0	357	0.0728	0.4021	0.2360	0.0058	0.2832	0.7334	0.3331	0.8374	8.4389
0	385	0.1354	0.1487	0.3445	0.0080	0.3634	0.8095	0.3677	0.9243	12.9196
0	413	0.1264	0.1959	0.3245	0.0066	0.3466	0.8638	0.3924	0.9864	12.9013
0	440	0.0875	0.3129	0.2934	0.0045	0.3016	0.9138	0.4150	1.0434	12.1073
0	468	0.1231	0.1964	0.3405	0.0058	0.3343	0.9801	0.4451	1.1191	14.5261
0	495	0.1157	0.2367	0.3307	0.0044	0.3126	1.0317	0.4686	1.1781	14.4062
0	523	0.0977	0.2596	0.3323	0.0041	0.3063	1.0927	0.4963	1.2476	15.3467
0	550	0.1024	0.3218	0.2425	0.0025	0.3311	1.1346	0.5153	1.2955	27.2060
0	578	0.0716	0.4419	0.2368	0.0027	0.2470	1.1792	0.5356	1.3465	12.9877
0	605	0.1887	0.3146	0.2655	0.0020	0.2292	1.2100	0.5496	1.3816	12.0676
0	633	0.0943	0.3885	0.2716	0.0016	0.2439	1.2933	0.5874	1.4768	14.4649
0	660	0.0943	0.3885	0.2716	0.0016	0.2439	1.3484	0.6124	1.5396	15.0792
0	688	0.0719	0.4383	0.2336	0.0014	0.2549	1.4045	0.6379	1.6037	16.3954
0	715	0.0738	0.4417	0.2272	0.0012	0.2561	1.4576	0.6620	1.6644	16.9935
0	743	0.0724	0.4176	0.2542	0.0013	0.2545	1.5223	0.6914	1.7382	17.9850
0	770	0.0775	0.4121	0.2522	0.0014	0.2569	1.5762	0.7159	1.7998	16.1640
0	798	0.0767	0.4091	0.2600	0.0010	0.2531	1.6349	0.7426	1.8668	19.2553
0	825	0.0669	0.4748	0.2248	0.0013	0.2322	1.6759	0.7612	1.9136	18.0785
0	853	0.0665	0.4490	0.2400	0.0010	0.2436	1.7413	0.7909	1.9883	19.6579
0	880	0.0633	0.4882	0.2118	0.0008	0.2358	1.7855	0.8110	2.0388	19.6899
5	200	0.1423	0.1035	0.0866	0.0032	0.6643	0.4229	0.2155	0.4829	13.9911
5	220	0.1509	0.1039	0.0859	0.0030	0.6562	0.4637	0.2379	0.5294	15.0199

5	247	0.1430	0.1313	0.0874	0.0026	0.6357	0.5195	0.2650	0.5932	16.2309
5	276	0.0785	0.2993	0.0761	0.0020	0.5441	0.5765	0.2797	0.6583	14.0756
5	303	0.1325	0.1389	0.1032	0.0030	0.6224	0.6389	0.3232	0.7295	19.2166
5	330	0.0748	0.3621	0.0754	0.0016	0.4862	0.6824	0.3302	0.7792	14.2436
5	357	0.0970	0.1945	0.1070	0.0016	0.5999	0.7548	0.3713	0.8618	21.5530
5	385	0.1175	0.2181	0.1011	0.0016	0.5618	0.8047	0.4027	0.9189	21.1053
5	413	0.1150	0.2165	0.1080	0.0015	0.5590	0.8642	0.4316	0.9867	22.2357
5	440	0.1135	0.2179	0.1123	0.0014	0.5548	0.9209	0.4594	1.0515	23.6723
5	468	0.1063	0.1863	0.1157	0.0014	0.5903	0.9877	0.4895	1.1277	27.7353
5	495	0.0979	0.2919	0.1086	0.0007	0.5009	1.0281	0.5068	1.1739	23.3369
5	523	0.0951	0.3344	0.1008	0.0008	0.4690	1.0791	0.5310	1.2321	22.2009
5	550	0.0881	0.3309	0.1087	0.0006	0.4717	1.1383	0.5568	1.2997	23.7846
5	578	0.0820	0.3994	0.0980	0.0005	0.4201	1.1840	0.5767	1.3519	21.1844
5	605	0.0941	0.3972	0.0944	0.0005	0.4137	1.2342	0.6074	1.4093	21.2704
5	633	0.0875	0.4217	0.0984	0.0004	0.3920	1.2887	0.6309	1.4715	21.1526
5	660	0.0923	0.4027	0.1075	0.0002	0.3973	1.3460	0.6615	1.5369	22.8075
5	688	0.0826	0.3817	0.1114	0.0004	0.4239	1.4136	0.6888	1.6140	26.3771
5	715	0.0772	0.3813	0.1166	0.0003	0.4246	1.4721	0.7140	1.6809	27.8966
5	743	0.0726	0.4186	0.1072	0.0002	0.4014	1.5218	0.7356	1.7377	27.0864
5	770	0.0799	0.4556	0.0971	0.0002	0.3673	1.5622	0.7601	1.7838	24.8699
5	798	0.0764	0.4070	0.2630	0.0010	0.2526	1.6357	0.7931	1.8677	18.7154
5	825	0.0719	0.4448	0.2341	0.0009	0.2483	1.6821	0.8128	1.9207	18.8535
5	853	0.0696	0.4947	0.2055	0.0011	0.2291	1.7246	0.8322	1.9692	17.8635
5	880	0.0701	0.4436	0.2444	0.0008	0.2411	1.7958	0.8664	2.0505	19.7167
-5	200	8.1244	0.4256	0.1933	0.4859	8.1244	0.4256	0.1933	0.4859	8.1244
-5	220	8.2392	0.4640	0.2107	0.5298	8.2392	0.4640	0.2107	0.5298	8.2392
-5	247	12.0910	0.5381	0.2444	0.6145	12.0910	0.5381	0.2444	0.6145	12.0910
-5	276	13.0084	0.5989	0.2720	0.6839	13.0084	0.5989	0.2720	0.6839	13.0084
-5	303	8.4205	0.5830	0.2648	0.6657	8.4205	0.5830	0.2648	0.6657	8.4205
-5	330	15.1205	0.7133	0.3240	0.8144	15.1205	0.7133	0.3240	0.8144	15.1205

-5	357	11.2880	0.7523	0.3417	0.8589	11.2880	0.7523	0.3417	0.8589	11.2880
-5	385	10.1147	0.8000	0.3633	0.9134	10.1147	0.8000	0.3633	0.9134	10.1147
-5	413	11.6686	0.8618	0.3914	0.9841	11.6686	0.8618	0.3914	0.9841	11.6686
-5	440	13.1066	0.9208	0.4182	1.0514	13.1066	0.9208	0.4182	1.0514	13.1066
-5	468	10.3921	0.9625	0.4372	1.0990	10.3921	0.9625	0.4372	1.0990	10.3921
-5	495	11.5548	1.0218	0.4641	1.1667	11.5548	1.0218	0.4641	1.1667	11.5548
-5	523	20.1881	1.1159	0.5068	1.2742	20.1881	1.1159	0.5068	1.2742	20.1881
-5	550	11.1396	1.1212	0.5093	1.2802	11.1396	1.1212	0.5093	1.2802	11.1396
-5	578	15.2055	1.2044	0.5470	1.3752	15.2055	1.2044	0.5470	1.3752	15.2055
-5	605	12.4723	1.2350	0.5609	1.4102	12.4723	1.2350	0.5609	1.4102	12.4723
-5	633	11.8289	1.2788	0.5808	1.4602	11.8289	1.2788	0.5808	1.4602	11.8289
-5	660	12.8865	1.3398	0.6085	1.5298	12.8865	1.3398	0.6085	1.5298	12.8865
-5	688	14.8076	1.4089	0.6399	1.6087	14.8076	1.4089	0.6399	1.6087	14.8076
-5	715	15.6767	1.4671	0.6663	1.6751	15.6767	1.4671	0.6663	1.6751	15.6767
-5	743	15.4483	1.5169	0.6890	1.7320	15.4483	1.5169	0.6890	1.7320	15.4483
-5	770	14.5948	1.5540	0.7058	1.7744	14.5948	1.5540	0.7058	1.7744	14.5948
-5	798	17.5345	1.6178	0.7348	1.8472	17.5345	1.6178	0.7348	1.8472	17.5345
-5	825	17.4730	1.6598	0.7539	1.8952	17.4730	1.6598	0.7539	1.8952	17.4730
-5	853	14.3594	1.6669	0.7571	1.9034	14.3594	1.6669	0.7571	1.9034	14.3594
-5	880	19.5235	1.7898	0.8129	2.0437	19.5235	1.7898	0.8129	2.0437	19.5235

REFERENCES

Abou-Senna, Hatem, and Essam Radwan. "VISSIM/MOVES integration to investigate the effect of major key parameters on CO< sub> 2</sub> emissions." Transportation Research Part D: Transport and Environment 21 (2013): 39-46.

Aerial Vehicle, http://www2.ece.ohio-state.edu/~coifman/documents/UAV-traffic.pdf

Ahn, Kyoungho, Hesham Rakha, Antonio Trani, and Michel Van Aerde. "Estimating vehicle fuel consumption and emissions based on instantaneous speed and acceleration levels." Journal of Transportation Engineering 128.2 (2002): 182-190.

Akçelik, Rahmi, and Mark Besley. Acceleration and Deceleration Models.Presented at 23rd Conference of the Australian Institutes of Transport Research (CAITR 2001), Monash University, Melbourne, Australia,Dec. 2001, revised July 11, 2002.

Andrew Papson, Seth Hartley, and Kai-Ling Kuo. "Analysis of Emissions at Congested and Uncongested Intersections with Motor Vehicle Emission Simulation 2010." Transportation Research Record: Journal of the Transportation Research Board 2270.1 (2012): 124-131.Analysis Model for vehicle Emission at a Signalized Intersection: Integrating Traffic and Microscopic Emission Model

Benjamin Coifman, Mark McCord, Rabi G. Mishalani, Michael Iswalt, Yuxiong Ji, Roadway Traffic Monitoring from an Unmanned.

Bhattacharya, Bhaskar; Habtzghi, DeSale (2002). "Median of the p value under the alternative hypothesis". The American Statistician. American Statistical Association. 56 (3): 202–6

Choi, David, et al. "MOVES sensitivity analysis: the impacts of temperature and humidity on emissions." Retrieved April 25 (2010): 2013.

Coelho, Margarida C., Tiago L. Farias, and Nagui M. Rouphail. "Impact of speed control traffic signals on pollutant emissions." Transportation Research Part D: Transport and Environment 10.4 (2005): 323-340.

Dion, Francois, Hesham Rakha, and Youn-Soo Kang. "Comparison of delay estimates at under-saturated and over-saturated pre-timed signalized intersections." Transportation Research Part B: Methodological (2004): 99-122.

Environmental Protection Agency- Six Principal Pollutants, http://www.epa.gov/airtrends/, Accessed in 2013.

EPA, Methodology for developing model emission rates for EPA'S multy-scale motor vehicle and equipment emission system, October 2002

EPA 2014 , https://www3.epa.gov/otaq/models/moves/index.htm

Eric Saund, Christopher Paulson, Gregory Burton, Eric Peeters. System and method for detecting, tracking and estimating the speed of vehicles from a mobile platform, US9070289 B2, Publication date: Jun 30, 2015

Farzaneh, Reza Mohamadreza, et al. "Developing Texas-Specific Drive Cycles for Use with the MOVES Model." Transportation Research Board 94th Annual Meeting. No. 15-4919. 2015.

Frey, H. Christopher, Alper Unal, N. M. Rouphail, and J. D. Colyar. (2002, April). Use of on-board tailpipe emissions measurements for development of mobile source emission factors. In Proceedings of US Environmental Protection Agency Emission Inventory Conference (pp. 1-13).

Hallmark, Shauna L. Analysis and prediction of individual vehicle activity for microscopic traffic modeling. Diss. Georgia Institute of Technology, 1999.

Hatem Abou-Senna and Essam Radwan. Developing a Microscopic Transportation Emissions Model to Estimate Carbon Dioxide Emissions on Limited-Access Highways. Journal of the Transportation Research Board, No. 2428, Transportation Research Board of the National Academies, Washington,D.C., 2014, pp. 44–53. DOI: 10.3141/2428

Highway Capacity Manual. TRB, National Research Council, Washington D.C., 2010.

How To Measure the Speed of Your Drone, UAV, or Quadcopter With a Radar Gun, http://flyingcameradronereviews.blogspot.com/2014/08/how-to-measure-speed-of-your-drone-uav-quadcopter-with-radar-gun.html.

Huai, Tao, Thomas D. Durbin, Ted Younglove, George Scora, Matthew Barth, and Joseph M. Norbeck. "Vehicle specific power approach to estimating on-road NH3 emissions from light-duty vehicles." Environmental science & technology 39.24 (2005)

Hung, HM James, Robert T. O'Neill, Peter Bauer, and Karl Kohne. (1997). "The behavior of the p-value when the alternative hypothesis is true". Biometrics. International Biometric Society. 53 (1): 11–22.

Joumard, Robert. "Methods of estimation of atmospheric emissions from transport: European scientist network and scientific state-of-the-art." INRETS report LTE 9901 (1999).

Jimenez-Palacios, Jose Luis. Understanding and quantifying motor vehicle emissions with vehicle specific power and TILDAS remote sensing. Diss. Massachusetts Institute of Technology, 1998.

Drake, Joseph S., Joseph L. Schofer, and Adolf D. May Jr. “A statistical Analysis of Speed-Density Hypotheses,” Transportation Research Record 154, p.78, 1967

Koupal, J., et al. MOVES2004 energy and emission inputs (Draft report). Prepared for US Environmental Protection Agency. EPA-420-P-05-003, Washington, DC, 2005.

Kutner, Michael H., Chris Nachtsheim, and John Neter. Applied linear regression models. McGraw-Hill/Irwin, 2004.

LeBlanc, David C., et al. "Driving pattern variability and impacts on vehicle carbon monoxide emissions." Transportation Research Record 1472 (1995).

Li, Jing-Quan; Wu, Guoyuan; Zou, Ning. Investigation of the impacts of signal timing on vehicle emissions at an isolated intersection. Transportation Research Part D: Transport and Environment, Volume 16, Issue 5, 2011, pp 409-414

Manual, Highway Capacity. "Highway capacity manual." (2010).

Matthew Muresan, ,S M Kamal Hossain,Liping Fu ,Chaozhe Jiang. A Trajectory-Clustering Based Method for Integrating Microscopic Traffic Simulation and 2 Emission Estimation Models. the 95th Annual Meeting of the Transportation Research 28 Board, 2016.

Neural Network. https://leonardoaraujosantos.gitbooks.io/artificial-inteligence/content/neural_networks.html. (Accessed on Jan 19, 2016)

Ozbay, Kaan, et al. "ASSIST-ME." Transportation Research Record: Journal of the Transportation Research Board 2399.1 (2013): 63-73.

Ozguven, Eren Erman, Kaan Ozbay, and Shrisan Iyer. "A Simplified Emissions Estimation Methodology Based on MOVES to Estimate 1 Vehicle Emissions from Transportation Assignment and Simulation Models 2." Transportation Research Board 92nd AnnualMeeting. No. 13-4402. 2013.

Performance-Box technical specifications. http://racelogic.co.uk/_downloads/Techical_Specs/PERFORMANCE%20BOX%20-%20Technical%20Specs.pdf. (Accessed on January 19, 2017)

Priddy KL, Keller PE. Artificial neural networks: an introduction. SPIE press; 2005.

Rakha, Hesham, Youn-Soo Kang, and François Dion. "Estimating vehicle stops at undersaturated and oversaturated fixed-time signalized intersections." Transportation Research Record: Journal of the Transportation Research Board 1776.1 (2001): 128-137.

Rouphail, Nagui M., H. Christopher Frey, James D. Colyar, and Alper Unal. "Vehicle emissions and traffic measures: exploratory analysis of field observations at signalized arterials." In 80th Annual Meeting of the Transportation Research Board, Washington, DC. 2001.

Roger, P. R., Elena, S. P., & William, R. M. (2004). Traffic engineering. Pearson Education, Inc. Upper Saddle River, NJ, 7458, 5-6.

Ritner, Mark, Kurt K. Westerlund, C. David Cooper, and Michael Claggett. "Accounting for acceleration and deceleration emissions in intersection dispersion modeling using MOVES and CAL3QHC." Journal of the Air & Waste Management Association 63, no. 6 (2013): 724-736.

R. V. Hogg and A. T. Craig. Introduction to Mathematical Statistics (Fifth Edition), Higher Education Press, 2004, pp. 471–478Navidi, William Cyrus. Statistics for engineers and scientists. Vol. 1. New York: McGraw-Hill, 2006. pp. 806

Scora, George, and Matthew Barth. "Comprehensive modal emissions model (CMEM), version 3.01." User guide. Centre for Environmental Research and Technology. University of California, Riverside (2006).

Shabihkhani, Rooholamin, and Eric J. Gonzales. "Analytical Model for Vehicle Emissions at a Signalized Intersection: Integrating Traffic and Microscopic Emissions Models." Transportation Research Board 92nd Annual Meeting. No. 13-5208. 2013.

Slavin, Courtney; Figliozzi, Miguel A. Impact of Traffic Signal Timing on Sidewalk-Level Particulate Matter Concentrations. Transportation Research Record: Journal of the Transportation Research Board, Issue 2340, 2013, pp 29–37

Stanek, David, and Chris Breiland. "Quick Estimation Method for Greenhouse Gas Emissions at Intersections." Transportation Research Board 92nd Annual Meeting. No. 13-1428. 2013.

Traffic Detector Handbook, Chapter 4, May, 2006. http://www.fhwa.dot.gov/publications/research/operations/its/06108/04.cfm, accessed May 13, 2014

Taylor, Brian J., ed. Methods and procedures for the verification and validation of artificial neural networks. Springer Science & Business Media, 2006.

Using MOVES in Project-Level Carbon Monoxide Analysis, EPA, December 2010

Webster, F. V. Traffic Signal Settings. TRRL Technical Paper 39. U.K.Transport and Road Research Laboratory, Crowthorne, Berkshire, United Kingdom, 1958.