

Microscopic Emission and Fuel Consumption Modeling for Light-duty Vehicles Using Portable Emission Measurement System Data

Wei Lei, Hui Chen, and Lin Lu

Abstract—Microscopic emission and fuel consumption models have been widely recognized as an effective method to quantify real traffic emission and energy consumption when they are applied with microscopic traffic simulation models. This paper presents a framework for developing the Microscopic Emission (HC, CO, NO_x, and CO₂) and Fuel consumption (MEF) models for light-duty vehicles. The variable of composite acceleration is introduced into the MEF model with the purpose of capturing the effects of historical accelerations interacting with current speed on emission and fuel consumption. The MEF model is calibrated by multivariate least-squares method for two types of light-duty vehicle using on-board data collected in Beijing, China by a Portable Emission Measurement System (PEMS). The instantaneous validation results shows the MEF model performs better with lower Mean Absolute Percentage Error (MAPE) compared to other two models. Moreover, the aggregate validation results tells the MEF model produces reasonable estimations compared to actual measurements with prediction errors within 12%, 10%, 19%, and 9% for HC, CO, NO_x emissions and fuel consumption, respectively.

Keywords—Emission, Fuel consumption, Light-duty vehicle, Microscopic, Modeling.

I. INTRODUCTION

AT present, traffic-induced exhaust emission and energy consumption is becoming one of the severest challenges for urban transportation in many metropolitans, including Beijing, in the world. Accurately and effectively estimating and controlling urban traffic emissions and fuel consumption has increasingly attracted attentions from transportation professionals and decision makers. Emission and fuel consumption models have been widely recognized as one of the most effective approaches for quantifying the environmental impacts of traffic on air quality. However, some state-of-the-art macroscopic emission models, such as MOBILE6 [1], EMFAC2002 [2], and COPERT III [3], usually employ simplified mathematical formulas to calculate fuel and emission rates based on average drive speeds without much regard to the transient effects of vehicle operating conditions. Even some recently developed models, for example ARTEMIS [4] and MOVES [5], are also known as aggregate

emission models based on different traffic situations and road facilities. Oppositely, microscopic models developed based on instantaneous driving variables e.g. speed and acceleration can more accurately capture the effects of vehicle dynamic characteristics on emissions. In addition, microscopic emission and fuel consumption models can be conveniently combined with microscopic traffic simulation models for evaluating the environmental impacts of operational-level transportation projects.

On the other hand, methods of vehicle emission data collection also deserve much concern because they considerably influence the prediction accuracy of developed models. Typically, several measurement methods, such as laboratory measurement [6], remote sensing [7], tunnel study [8] [9], and on-board emission measurement [10] [11] have been widely used in this research field. However, development of microscopic models with satisfactory prediction accuracy requires abundant on-road emission data. As a new data collection technology, Portable Emission Measurement System (PEMS) provides researchers with possibilities to gather the real-world emission and fuel consumption data. In recent decades, PEMS is increasingly being used in various transportation research projects to analyze real-world vehicle emissions impacts [12] [13]. Moreover, recent studies also proved that PEMS is reliable of measuring real-world vehicle emissions with high accuracy compared with laboratory measurements [14].

This paper presents a Microscopic Emission and Fuel consumption (MEF) model for two categories of light-duty vehicles (LDV) developed using PEMS measurements collected in Beijing, China. The MEF model is a statistical model constructed based on instantaneous speed and acceleration adopting multivariate least-square regression method. This paper is organized as follows. Firstly, it present the collection, analysis and preprocessing for the PEMS data. Secondly, it provides descriptions of the general form of existed microscopic emission models and describes the structures of the MEF model and other two related models, VT-Micro [15] and POLY [16]. Thirdly, the calibration and validation results of the MEF model as well as VT-Micro model and POLY model compared to field measurements for two categories of light-duty vehicles are presented. Finally, the main findings of our study and identify the future research direction are summarized.

Wei Lei is with Intelligent Transportation System Centre (ITSC), Wuhan University of Technology (WUT), Wuhan, 430063, China (corresponding author, phone: 0046-0760833279; e-mail: leiwei.whut@gmail.com).

Hui Chen is with ITSC, WUT, Wuhan, 430063, China (e-mail: chenhui.whut@gmail.com).

Lin Lu is with Advanced Heat Engine Research Centre (AHERC), WUT, Wuhan, 430063, China (e-mail: lulin@mail.whut.edu.cn).

II. DATA SOURCE

The data that were utilized in this paper to develop the emission and fuel consumption models were collected in Beijing, China from 18th to 29th, November, 2007. A PEMS instrument, On Board Emission Measurement System (OBS-2200), was employed in this experiment to gather real-world data for 28 light-duty gasoline vehicles. OBS-2200 consists of vibration-proof gas analyzers, a laptop PC with software for system controlling and data logging, accessory sensors, and a tailpipe attachment with a Pitot tube. CO and CO₂ concentration can be measured by a NDIR analyzer, HC and NO_x concentration can be measured by a FID analyzer and a CLD analyzer. GPS data (speed, latitude, longitude, altitude, etc.) and other external signals (temperature, humidity, atmospheric pressure, etc.) can be saved into the laptop PC by the data logging software [17].

A test route was predesigned to cover different road types with appropriate proportions in Beijing in order to capture real-world driving behaviors of the typical driving population. Each of the 28 recruited vehicles in this experiment was tested twice on this route. Specifically, the first test was usually during the peak-hours in the morning (7:00-9:00) and the second test was during the off-peak hours (11:00-13:00). Each run of each vehicle typically lasted for about 30 ~ 40 minutes. The OBS-2200 system automatically recorded the real time emission and fuel consumption data and vehicle operation data second by second.

In total 76,415 raw data records were finally obtained and saved into the OBS-2200 system in our experiments. For the purpose of convenient query, a MySQL database was developed to preserve the measured data. Moreover, a MATLAB process program was developed to be integrated with this MySQL database, which could be easily used to query the measured data from MySQL database and process them by MATLAB. Each data record was comprised of vehicle speed (m/s), fuel consumption rate (g/s), emission rate (g/s), relative humidity, latitude, longitude, and altitude (m) etc. Typically, speed values range from 0 to 30 m/s, and acceleration values vary from -2.8 to 1.6 m/s². Fig. 1 illustrates an example of speed/acceleration distribution of driving data from one test vehicle. It shows that the majority of the speed and acceleration data takes place when the vehicles operate during stationary state (acceleration ranging between -0.4 and 0.4 m/s²).

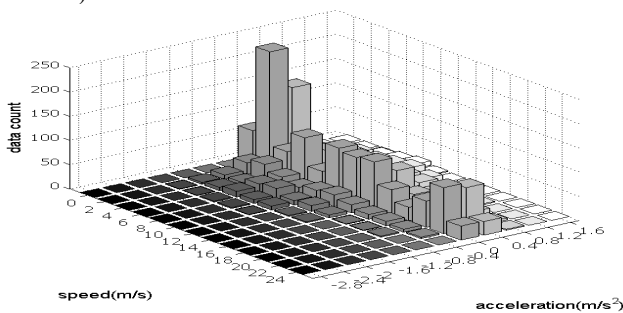


Fig. 1 An example of speed and acceleration distribution of driving data from one test vehicle

To ensure all the data used in our modeling are reliably valid, three process procedures on all raw data were

performed. First of all, synchronization between the independent variables (speed and acceleration) and the dependent variables (emission and fuel consumption rates) is of importance to modeling quality in our study. Hence, a time offset procedure was designed to eliminate the time lag between emission profile and speed profile, which is resulted from the transport of gases from engine to the analyzers in OBS system. Specifically, the two profiles were shifted according to the time axis in a way such that the first crest of the emission data well match the engine start (indicated by a sudden increase in vehicle speed) [18]. Secondly, all negative emission data were removed from the database because they make no sense. Finally, a data smoothing technique, namely moving average smoothing, was also adopted in this paper to reduce unwanted noise from raw data [19].

III. MODELING METHODOLOGY

Several methodologies for developing microscopic vehicle emission and fuel consumption models are described briefly in this second. And then we proposed the basic framework of the MEF model based on another two existing models.

A. Methodology review

Microscopic models generally take kinematic variables e.g. speed and acceleration as input to calculate second-by-second vehicle emission and fuel consumption rates (g/s). The effect of various driving modes (acceleration, deceleration, cruise, idle) on emission and energy consumption is taken into account in these models, which is in general demonstrated as:

$$e_i(t) = \sum_j e_i(c_j, x_j(t)) \quad (1)$$

Where: i is the species of pollutions (or the fuel consumption); c_j is the category of vehicle j ; $x_j(t)$ denotes instantaneous variables of vehicle j at time t ; $e_i(c_j, x_j(t))$ denotes the emission of species i (or the fuel consumption) for vehicle j at time t ; $e_i(t)$ denotes the total amount of emissions of species i (or the fuel consumption) generated at time t in a given area.

Microscopic models presented in this general form can usually be classified into four categories, namely emission maps, regression-based models, load-based models, and neural network models. Emission maps are actually look-up tables for querying data of emission and fuel consumption rates according to each combination of vehicle speed and acceleration or of engine speed and torque. This type of emission models can be too sensitive to the driving cycle to generate satisfactory results [20]. Regression models normally adopt mathematic functions of instantaneous speed and accelerations as explanatory to predict emission rates. Although they are usually lack of clear physical basis, they can produce estimates of emissions with high quality [15]. Load-based models provide another modeling method, which is expressed as functions of several causal variable variables with detailed physical basis of emission generating in an engine [6]. However, too many input parameters are typically desired to feed into this type of models, which limit their application to a certain extent. Neural network models were also studied by researches in the past decade [21]. Due to the requirement of large quantity of input data for training the networks, these models usually take much computational time

for a run when they are combined with traffic simulation models. Among these types of modes discussed above, developing regression-based models using multivariate least-square methods based on PEMS data is what we focus on in this paper.

B. Model development

Most of the existing regression-based microscopic emission and fuel consumption models take account the speed and acceleration at a time point to estimate the emission and fuel consumption rates at the same time point. For example, the VT-Micro model has been developed by researchers based on laboratory chassis dynamometer measurements using a combination of linear, quadratic, and cubic speed and acceleration terms, which is demonstrated below [15]:

$$e(t) = \begin{cases} \exp(\sum_{i=0}^3 \sum_{j=0}^3 L_{i,j} \times v(t)^i \times a(t)^j) & a \geq 0 \\ \exp(\sum_{i=0}^3 \sum_{j=0}^3 M_{i,j} \times v(t)^i \times a(t)^j) & a < 0 \end{cases} \quad (2)$$

where $e(t)$ is the emission rate (or fuel consumption) at time t (g/s); $L_{i,j}$ and $M_{i,j}$ are the regression model coefficients; $v(t)$ and $a(t)$ are instantaneous speed and acceleration at time t .

However, researchers have also developed the POLY models that identified the acceleration or deceleration of previous time periods takes more impact on emissions than that of current time does [16]. The basic form of POLY models can be described by:

$$e(t) = \beta_0 + \beta_1 v(t) + \beta_2 v^2(t) + \beta_3 v^3(t) + \beta_T T'(t) + \beta_{T''} T''(t) + \beta_{a_1} a(t) + \dots + \beta_{a_9} a(t-9) + \beta_w w(t) \quad (3)$$

where β_0 is a constant and β_x is the coefficient for variable x . $a(t-k)$ is acceleration at time $t-k$ where $k=1 \dots 9$; $T'(t)$ and $T''(t)$ are acceleration and deceleration time up to time t since its inception; $w(t)$ is specific power at time t , which is equal to the product of $v(t)$ and $a(t)$.

Based on the study on the VT-Micro and the POLY, we take not only the current speed and acceleration but also the history acceleration of previous nine seconds before current time point t into account in the MEF model. The framework of the MEF model is presented by:

$$e(t) = \begin{cases} \exp(\sum_{m=0}^3 \sum_{n=0}^3 \lambda_{m,n} \times v(t)^m \times \bar{a}(t)^n) & \bar{a} \geq 0 \\ \exp(\sum_{m=0}^3 \sum_{n=0}^3 \gamma_{m,n} \times v(t)^m \times \bar{a}(t)^n) & \bar{a} < 0 \end{cases} \quad (4)$$

where $\bar{a}(t)$ is the composite acceleration at time t derived from the current acceleration $a(t)$ and the nine historical accelerations $a(t-1), \dots, a(t-9)$, which can be expressed as:

$$\bar{a}(t) = \sum_{i=0}^9 \omega_i \cdot a(t-i) \quad (5)$$

where ω_i is the weight for $a(t-i)$, $0 \leq \omega_i \leq 1$, $\omega_0 + \omega_1 + \dots + \omega_9 = 1$; $\lambda_{m,n}$ and $\gamma_{m,n}$ are the model coefficients for $e(t)$ at speed power m and acceleration power n . Assuming these nine historical acceleration give the same degree of impact on emission and fuel consumption, we can transform equation (5) into:

$$\bar{a}(t) = \alpha \cdot a(t) + (1-\alpha) \cdot \sum_{i=1}^9 a(t-i) \quad (6)$$

where α is defined in this paper as Acceleration Impact Factors (AIF) for constructing composite acceleration, $0 \leq \alpha \leq 1$. The purpose of introducing the variable of composite acceleration into the MEF model is to capture the effects of historical

accelerations interacting with current speed, which is not taken account into POLY model, on emission and fuel consumption based on the framework of the VT-Micro model. Further, the AIF (α) can also be regarded as an indicator that indirectly reflects the magnitude of impact of current acceleration and historical acceleration on emission rate. The nature logarithm is adopted in the MEF model in order to transform the emission and fuel consumption rate mainly based on two concerns. Firstly, negative predictions generated by the model can be avoided. Secondly, the magnitude of disparity among the raw emission data can be lowered by this transformation technique in order to improve the quality of model calibration.

IV. MODEL CALIBRATION

In this section, the MEF model was calibrated for two categories of light-duty vehicles using the PEMS data collected in Beijing. From the perspective of comparative study, the VT-Micro model and POLY model were also calibrated with the same dataset that used for the MEF model. The comparison and analysis of calibration results for the MEF model are provided at the end of this section.

A. Vehicle classification

Generally, vehicles should be classified into several categories before developing emission and fuel consumption models because different types of vehicles have different emission properties, which considerably impacts the calibrated coefficients for one model framework. Therefore, it is reasonable and necessary to divide vehicles into different groups and calibrate the models for each vehicle group. In this study, the classification criteria set up in the CMEM model [6], a state-of-the-art load-based emission model, was adopted to classify the 28 test vehicles into two categories, LDVI and LDVII, according to four basic vehicle characteristics, namely fuel delivery system (carbureted or fuel injection), emission control technology (no catalyst, 2-way catalyst, or 3-way catalyst), accumulated mileage ($\geq 50,000$ miles, $\leq 50,000$ miles), and power-to-weight ratio (high or low). Table I illustrates the final vehicle classification results in this paper.

B. Model Calibration

Before calibrating the models, the experimental data from five vehicles of category LDVI and eight vehicles of category LDVII were selected purposefully to construct two composite datasets, which would be used in the following calibration.

TABLE I
 CLASSIFICATION RESULTS FOR ALL TEST VEHICLES

Vehicle category	Fuel delivery system	Emission Control Tech.	Mileage (mile)	Power-to-weight ratio	Number of test vehicles
LDVI	FI	3-way catalyst	<50,000	low (<0.039hp/lb)	11
LDVII	FI	3-way catalyst	>50,000	low (<0.039hp/lb)	17

For both composite datasets, vehicles selection was performed in such a way that driving mileages of the chosen vehicles cover the entire mileage scope within the critical values listed in Table I. Finally more than 12,000 and 18,000 data points in total are included in the two composite data samples respectively for category LDVI and LDVII.

For the MEF model, the calibration process should be a complex nonlinear programming problem that costs too much computation time to reach a solution. In order to simply it, we determined the value of AIF (α) in advance. Specifically, we set the value of α to be from 0 to 1 at increments of 0.1 and adopt multivariate least-squares regression method to calibrate the MEF model using the predesigned composite datasets for LDVI vehicle category. The calibration quality with different values of α can be compared by Correlation Coefficient (R), which can be represented as follows:

$$R = \frac{\sum_{t=1}^T [\hat{e}(t) - \bar{\hat{e}}] \cdot [e(t) - \bar{e}]}{\sqrt{\sum_{t=1}^T (\hat{e}(t) - \bar{\hat{e}})^2 \sum_{t=1}^T (e(t) - \bar{e})^2}} \quad (7)$$

where $e(t)$ are the field measurements and $\hat{e}(t)$ are the model predictions; \bar{e} and $\bar{\hat{e}}$ are the average of $e(t)$ and $\hat{e}(t)$, respectively; T denotes the total number of data points included in the dataset for calibration.

The R statistics of calibration results against α values for the MEF model are demonstrated in Fig. 2, from which it can be seen that the R values for submodels of HC, CO and NOx will increase along with the growth of α value before it reaches 0.3, and the corresponding critical value of α for submodels of CO₂ and FUEL is near 0.8. Finally we define $\alpha=0.5$ in the MEF model with an overall consideration such that all the five submodels can keep relatively high R values. Further, as the basis of the MEF model, the VT-Micro and the POLY model are also calibrated with the same datasets by multivariate least-squares regression method for the purpose of comparison study with the MEF model in this paper.

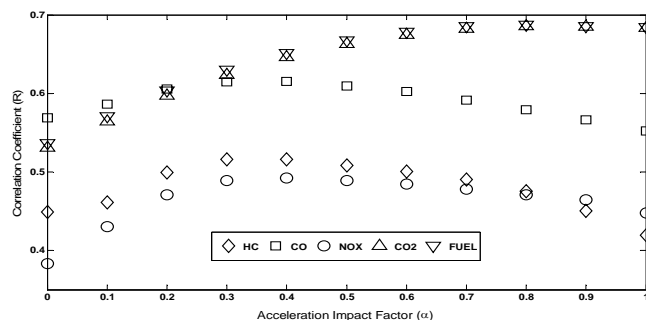


Fig. 2 Correlation Coefficient (R) against AIF (α) for calibration of the MEF model

The calibration results of HC, CO emission and FUEL consumption of the MEF model as well as the corresponding t statistic of each parameter for composite datasets of LDVI and LDVII vehicles are illustrated in Table II and Table III, respectively. In order to keep the same with the general framework of the VT-Micro model developed by Rakha [15], some parameters are still remained in our study though they are not statistically significant by the t statistics at the 95% confidence level. Another reason for keeping these variables is that removing them from the model framework would cause the reduction of model performance. The values of adjusted R^2 , which is statistically considered as an effective goodness-of-fit measurement in multiple linear regression models, for each submodel are also showed. It can be seen from Table II and Table III that the MEF model always shows lower adjusted R^2 in negative acceleration regime than in positive acceleration regime. In addition, we can preliminarily conclude that the CO submodel performs better than the other two, and this is simply because in linear regression the higher the adjusted R^2 is, the more proportion of variance in the dependent variables can be accounted for by the explanatory variables.

TABLE II
 CALIBRATED PARAMETERS OF THE MEF MODEL ($\alpha=0.5$) FOR LDVI VEHICLE CATEGORY

HC			CO			FUEL											
Value	t		Value	t		Value	t		Value	t		Value	t				
$\lambda_{0,0}$	-8.9680	-193.20	$\gamma_{0,0}$	-8.9184	-175.92	$\lambda_{0,0}$	-5.4401	-164.30	$\gamma_{0,0}$	-5.3600	-136.65	$\lambda_{0,0}$	-1.9116	-61.18	$\gamma_{0,0}$	-1.9428	-55.45
$\lambda_{0,1}$	5.0509	4.44	$\gamma_{0,1}$	0.1177	0.21	$\lambda_{0,1}$	6.1901	7.62	$\gamma_{0,1}$	-1.9663	-4.46	$\lambda_{0,1}$	2.7630	3.61	$\gamma_{0,1}$	-0.4093	-1.04
$\lambda_{0,2}$	-7.0006	-2.11	$\gamma_{0,2}$	-0.2350	-0.23	$\lambda_{0,2}$	-8.6441	-3.65	$\gamma_{0,2}$	-2.3950	-3.01	$\lambda_{0,2}$	-2.5418	-1.14	$\gamma_{0,2}$	-0.6861	-0.97
$\lambda_{0,3}$	3.4359	1.51	$\gamma_{0,3}$	-0.1795	-0.43	$\lambda_{0,3}$	3.7943	2.34	$\gamma_{0,3}$	-0.8173	-2.51	$\lambda_{0,3}$	0.6289	0.41	$\gamma_{0,3}$	-0.2422	-0.83
$\lambda_{1,0}$	0.1640	5.39	$\gamma_{1,0}$	0.2440	6.84	$\lambda_{1,0}$	0.3588	16.56	$\gamma_{1,0}$	0.3705	13.43	$\lambda_{1,0}$	0.1108	5.42	$\gamma_{1,0}$	0.1342	5.45
$\lambda_{2,0}$	-0.0081	-2.34	$\gamma_{2,0}$	-0.0178	-3.93	$\lambda_{2,0}$	-0.0221	-8.99	$\gamma_{2,0}$	-0.0260	-7.45	$\lambda_{2,0}$	-0.0008	-0.36	$\gamma_{2,0}$	-0.0052	-1.66
$\lambda_{3,0}$	0.0002	1.93	$\gamma_{3,0}$	0.0005	3.50	$\lambda_{3,0}$	0.0005	6.86	$\gamma_{3,0}$	0.0007	6.11	$\lambda_{3,0}$	-0.0005	-0.73	$\gamma_{3,0}$	0.0001	1.48
$\lambda_{1,1}$	-0.7278	-1.74	$\gamma_{1,1}$	0.3314	1.18	$\lambda_{1,1}$	-1.1653	-3.91	$\gamma_{1,1}$	0.8556	3.94	$\lambda_{1,1}$	-0.0850	-0.30	$\gamma_{1,1}$	0.2968	1.53
$\lambda_{2,1}$	0.0495	1.24	$\gamma_{2,1}$	-0.0249	-0.74	$\lambda_{2,1}$	0.0906	3.18	$\gamma_{2,1}$	-0.0657	-2.54	$\lambda_{2,1}$	-0.0198	-0.74	$\gamma_{2,1}$	0.0116	0.50
$\lambda_{3,1}$	-0.0008	-0.77	$\gamma_{3,1}$	0.0007	0.70	$\lambda_{3,1}$	-0.0018	-2.41	$\gamma_{3,1}$	0.0015	1.89	$\lambda_{3,1}$	0.0011	1.55	$\gamma_{3,1}$	-0.0007	-1.02
$\lambda_{1,2}$	1.9105	1.68	$\gamma_{1,2}$	0.4369	0.94	$\lambda_{1,2}$	2.2321	2.74	$\gamma_{1,2}$	0.8341	2.33	$\lambda_{1,2}$	0.2791	0.36	$\gamma_{1,2}$	0.3028	0.95
$\lambda_{2,2}$	-0.1149	-1.11	$\gamma_{2,2}$	-0.0370	-0.68	$\lambda_{2,2}$	-0.1549	-2.11	$\gamma_{2,2}$	-0.0609	-1.45	$\lambda_{2,2}$	0.0275	0.40	$\gamma_{2,2}$	0.0154	0.41
$\lambda_{3,2}$	0.0015	0.57	$\gamma_{3,2}$	0.0009	0.52	$\lambda_{3,2}$	0.0027	1.41	$\gamma_{3,2}$	0.0011	0.88	$\lambda_{3,2}$	-0.0019	-1.06	$\gamma_{3,2}$	-0.0011	-0.95
$\lambda_{1,3}$	-1.0900	-1.44	$\gamma_{1,3}$	0.1278	0.71	$\lambda_{1,3}$	-1.0589	-1.95	$\gamma_{1,3}$	0.2029	1.46	$\lambda_{1,3}$	-0.1128	-0.22	$\gamma_{1,3}$	0.0573	0.46
$\lambda_{2,3}$	0.0719	1.08	$\gamma_{2,3}$	-0.0097	-0.47	$\lambda_{2,3}$	0.0725	1.53	$\gamma_{2,3}$	-0.0111	-0.69	$\lambda_{2,3}$	-0.0111	-0.25	$\gamma_{2,3}$	0.0095	0.66
$\lambda_{3,3}$	-0.0012	-0.71	$\gamma_{3,3}$	0.0002	0.25	$\lambda_{3,3}$	-0.0013	-1.06	$\gamma_{3,3}$	0.0007	0.15	$\lambda_{3,3}$	0.0008	0.72	$\gamma_{3,3}$	-0.0005	-1.15
adjusted $R^2=0.456$			adjusted $R^2 = 0.322$			adjusted $R^2 = 0.582$			adjusted $R^2 = 0.424$			adjusted $R^2 = 0.495$			adjusted $R^2 = 0.394$		

TABLE III
CALIBRATED PARAMETERS OF THE MEF MODEL ($\alpha=0.5$) FOR LDVII VEHICLE CATEGORY

HC						CO						FUEL									
Value	t	Value	t	Value	t	Value	t	Value	t	Value	t	Value	t	Value	t						
$\lambda_{0,0}$	-6.3291	-151.20	$\gamma_{0,0}$	-6.35599	-143.81	$\lambda_{0,0}$	-4.37417	-162.19	$\gamma_{0,0}$	-4.34111	-131.26	$\lambda_{0,0}$	-2.02233	-65.40	$\gamma_{0,0}$	-1.98666	-57.13				
$\lambda_{0,1}$	1.4134	1.46	$\gamma_{0,1}$	0.7361	1.47	$\lambda_{0,1}$	3.9721	6.35	$\gamma_{0,1}$	-0.4493	-1.20	$\lambda_{0,1}$	3.0116	4.20	$\gamma_{0,1}$	-0.4183	-1.06				
$\lambda_{0,2}$	-1.98799	-0.65	$\gamma_{0,2}$	0.7722	0.82	$\lambda_{0,2}$	-7.5711	-3.82	$\gamma_{0,2}$	-0.3978	-0.57	$\lambda_{0,2}$	-3.6463	-1.60	$\gamma_{0,2}$	-1.0615	-1.44				
$\lambda_{0,3}$	2.2297	0.93	$\gamma_{0,3}$	0.0216	0.06	$\lambda_{0,3}$	5.4554	3.54	$\gamma_{0,3}$	-0.3386	-1.16	$\lambda_{0,3}$	2.0060	1.14	$\gamma_{0,3}$	-0.6250	-2.04				
$\lambda_{1,1}$	0.2626	8.79	$\gamma_{1,0}$	0.2009	5.98	$\lambda_{1,0}$	0.3387	17.59	$\gamma_{1,0}$	0.2620	10.43	$\lambda_{1,0}$	0.1334	6.05	$\gamma_{1,0}$	0.1305	4.94				
$\lambda_{2,1}$	-0.0226	-6.12	$\gamma_{2,0}$	-0.0134	-3.17	$\lambda_{2,0}$	-0.0279	-11.72	$\gamma_{2,0}$	-0.0180	-5.71	$\lambda_{2,0}$	-0.0060	-2.21	$\gamma_{2,0}$	-0.0061	-1.84				
$\lambda_{3,1}$	0.0005	4.44	$\gamma_{3,0}$	0.0002	1.84	$\lambda_{3,0}$	0.0007	10.43	$\gamma_{3,0}$	0.0005	4.89	$\lambda_{3,0}$	0.0002	2.22	$\gamma_{3,0}$	0.0002	1.83				
$\lambda_{1,2}$	0.0739	0.19	$\gamma_{1,1}$	-0.6077	-2.18	$\lambda_{1,1}$	-0.9961	-3.90	$\gamma_{1,1}$	-0.1856	-0.89	$\lambda_{1,1}$	-0.2541	-0.87	$\gamma_{1,1}$	0.1031	0.47				
$\lambda_{2,2}$	0.0227	0.51	$\gamma_{2,1}$	0.0890	2.52	$\lambda_{2,1}$	0.1308	4.54	$\gamma_{2,1}$	0.0567	2.14	$\lambda_{2,1}$	0.0359	1.08	$\gamma_{2,1}$	0.0299	1.07				
$\lambda_{3,2}$	-0.0010	-0.76	$\gamma_{3,1}$	-0.0024	-2.11	$\lambda_{3,1}$	-0.0043	-4.91	$\gamma_{3,1}$	-0.0018	-2.07	$\lambda_{3,1}$	-0.0013	-1.24	$\gamma_{3,1}$	-0.0013	-1.43				
$\lambda_{1,3}$	0.5227	0.44	$\gamma_{1,2}$	-0.6958	-1.37	$\lambda_{1,2}$	2.9568	3.87	$\gamma_{1,2}$	-0.4664	-1.24	$\lambda_{1,2}$	0.7478	0.85	$\gamma_{1,2}$	0.3768	0.95				
$\lambda_{2,3}$	-0.1231	-0.93	$\gamma_{2,2}$	0.0623	0.91	$\lambda_{2,2}$	-0.3705	-4.34	$\gamma_{2,2}$	0.0855	1.66	$\lambda_{2,2}$	-0.0932	-0.95	$\gamma_{2,2}$	-0.0099	-0.18				
$\lambda_{3,3}$	0.0054	1.27	$\gamma_{3,2}$	-0.0002	-0.08	$\lambda_{3,2}$	0.0129	4.74	$\gamma_{3,2}$	-0.0018	-0.97	$\lambda_{3,2}$	0.0032	1.03	$\gamma_{3,2}$	0.0001	0.08				
$\lambda_{1,4}$	-0.9017	-0.97	$\gamma_{1,3}$	-0.0524	-0.23	$\lambda_{1,3}$	-2.3323	-3.89	$\gamma_{1,3}$	-0.0877	-0.52	$\lambda_{1,3}$	-0.7446	-1.08	$\gamma_{1,3}$	0.2607	1.46				
$\lambda_{2,4}$	0.1326	1.25	$\gamma_{2,3}$	-0.0132	-0.38	$\lambda_{2,3}$	0.2871	4.19	$\gamma_{2,3}$	0.0144	0.56	$\lambda_{2,3}$	0.0935	1.19	$\gamma_{2,3}$	-0.0226	-0.83				
$\lambda_{3,4}$	-0.0054	-1.51	$\gamma_{3,3}$	0.0014	1.06	$\lambda_{3,3}$	-0.0103	-4.48	$\gamma_{3,3}$	-0.0002	0.18	$\lambda_{3,3}$	-0.0034	-1.28	$\gamma_{3,3}$	0.0008	0.73				
adjusted $R^2=0.487$		adjusted $R^2 = 0.362$				adjusted $R^2 = 0.594$				adjusted $R^2 = 0.458$				adjusted $R^2 = 0.465$				adjusted $R^2 = 0.381$			

V. MODEL VALIDATION

Model validation seems to be the most important step in the model building sequence to test the capability of models to generate reasonable predictions from new inputs that are different from those used in calibration. In order to examine and analyze the prediction performance of the MEF model, we implement two types of validation procedure, namely instantaneous validation and aggregate validation, using the out-of-sample datasets that are not adopted in the calibration procedure. The purpose of instantaneous validation is to examine the microscopic prediction ability of the MEF model in second-by-second resolution, whereas the aggregate validation is adopted to verify its macroscopic forecast accuracy of the average emission factors (g/km) during specific driving periods.

A. Instantaneous validation

In instantaneous validation, field measurements of six vehicles (not used in previous calibration) were used as the data samples. Three of them were prepared for LDVI category and the other three were for LDVII. Then the predicted second-by-second emission rates and fuel consumption rates generated from MEF model based on the real driving parameters (speed and acceleration) can be compared with the real world experimental data. In addition, the same data samples were also used to validate the VT-Micro model and POLY model, and the comparison and analysis of the validation results of these three models would be presented below. In this paper, two widely used statistics, the Mean Absolute Percentage Error (MAPE) and the Root Mean Square Error (RMSE) were adopted to verify the validation quality. MAPE and RMSE can be expressed as follows:

$$MAPE = \sum_i \left| \frac{\hat{e}(t) - e(t)}{e(t)} \right| \quad (8)$$

$$RMSE = \sqrt{\sum_{t=1}^T [\hat{e}(t) - e(t)]^2 / T} \quad (9)$$

the meaning of all the variables in the right hand side of equation (8) and (9) are the same as those in equation (7). The instantaneous validation results based on these two statistics are listed in Table IV and Table V, respectively. It can be seen from Table IV that most of *MAPEs* derived from the MEF model are smaller than those from VT-Micro model and POLY model. It seems that *MAPEs* for CO₂ emission and fuel consumption are comparatively lower than those for the other three pollutants. On the other hand, we can also say that the MEF model performs best by comparing *RMSEs* listed in table V because the MEF model generates the smallest *RMSE* values in most of time. However, POLY model can also give the lowest *RMSE* values in some cases. In addition, the *RMSE* of HC, CO and NO_x emission for LDVII vehicles are obviously larger than those for LDVI vehicles. But the *RMSE* of CO₂ emission and fuel consumption for both vehicle categories are always similar. This may imply that the LDVII vehicles can produce more emissions than LDVI vehicles do based on the same fuel consumption.

To illustrate the instantaneous estimation performance of the MEF model compared to the other two models, two second-by-second validation examples are graphically showed in Fig. 3 (LDVI) and 4 (LDVII). From Fig. 3, it is easy to tell that all of the three models can well capture the trend of real world emission trajectory and fuel consumption trajectory. During low emission periods, all of them have similar prediction performance. But during high emission periods, MEF model performs better than the other two models because it can give closer estimations to the peak values. However, all of them failed to reflect the peak measurements for HC emission. In Fig. 4, the emission and fuel consumption profiles generated by these three models can still follow the tendency of field measurements. We notice that the MEF model make more

accurate estimations at low emission areas while POLY models give overestimations at the same areas. However, all of these models still considerably underestimate HC emission at peak emission areas.

B. Aggregate validation

In our experiments, all of the recruited vehicles were tested on the same predesigned test route from the start to the end. So the aggregate emission factors (g/km) and fuel consumption (g/km) of each run for vehicles of the same category are expected to be similar or comparable. To validate the MEF model in an aggregate way, we randomly selected three data sequences for both categories as validation samples. Each data sample is composed of data of peak-hour run and off-peak hour run from the same vehicle. Then we calculated the real-world average emission factors and energy consumption rates for each data sample by dividing the total amounts of emissions and energy consumption by the distance of the test route. Afterwards, the second-by-second predictions generated by the MEF model were aggregated into a sum, which can be also divided by total distance to derive the average emission factors and energy consumption. The same procedure would be performed on VT-Micro model and POLY model similarly.

From the comparison point of view, ARTEMIS road emission model [4] was also adopted in this paper to estimate the average emission factors and fuel consumptions. The ARTEMIS road model was developed within the framework of the EU 5th FP project ARTEMIS (Assessment and Reliability of Transport Emission Models and Inventory Systems). It is composed of two types of emission models, namely Traffic Situation based models and Average Speed based models. The former are designed as non-continuous or discreet models, opposite to instantaneous or average speed models, based on different traffic situations, which can be described by typical speed-time curves. However, average speed based models are

also provided in ARTEMIS in order to keep continuity with previous models like COPERT. In our study, the average speed based models for gasoline passenger car in ARTEMIS model were used and the general mathematical form for these models can be demonstrated as follows:

$$EF = \frac{a + c \cdot v + e \cdot v^2}{1 + b \cdot v + d \cdot v^2} + \frac{f}{v} \quad (10)$$

where EF is emission factor or fuel consumption (g/km), v is average speed, a, b, c, d, e, f are model coefficients. However, models of CO_2 emission are not constructed in ARTEMIS. Hence only the emission factors of HC, CO, NO_x and fuel consumption are calculated in ARTEMIS model and compared with estimations of other models.

Fig. 5 illustrates the comparison of aggregate emission factors and fuel consumption derived by MEF, ARTEMIS, VT-Micro, and POLY models with actual measurements. It can be seen from Fig. 5(a) (LDVI category) that the MEF give the closest predictions to measurements compared to the other three models in most cases. The average absolute error between the MEF predictions and measurements are 12.0%, 9.9%, 11.6%, and 6.3% for HC, CO, NO_x , and fuel consumption, respectively. Generally VT-Micro model has similar prediction performance with MEF model while POLY model tends to generate considerable overestimation of emissions and fuel consumption. It can be also noted that ARTEMIS model greatly underestimate CO emission factors. On the other hand, we can still tell from Fig. 5(b) (LDVII category) that MEF model performs best among the four models with the average absolute errors of 11.4%, 5.4%, 18.7% and 8.1% for HC, CO, NO_x , and fuel consumption, respectively. However, ARTEMIS model still give prodigious underestimation of HC, CO, and NO_x emission factors, which may implies that these data collected in Beijing cannot be applied with ARTEMIS model because it was developed based on European measurements.

TABLE IV
 COMPARISON OF MODEL VALIDATION RESULTS BY MEAN ABSOLUTE PERCENTAGE ERROR (MAPE)

Vehicle number	Mean Absolute Percentage Error (%)														
	CO			CO ₂			HC			NO _x			FUEL		
	MEF	VT-Micro	POLY	MEF	VT-Micro	POLY	MEF	VT-Micro	POLY	MEF	VT-Micro	POLY	MEF	VT-Micro	POLY
LDVI-1	58	64	61	49	52	51	65	71	68	81	84	88	42	49	51
LDVI-2	53	58	58	56	54	60	73	77	73	79	82	89	50	54	60
LDVI-3	52	54	54	54	54	58	66	70	89	80	82	73	54	54	57
LDVII-1	52	52	57	66	69	74	52	60	62	63	67	65	47	46	54
LDVII-2	64	71	72	43	51	50	46	51	53	60	63	66	44	48	52
LDVII-3	56	57	59	54	55	56	50	51	53	58	60	60	54	54	56

TABLE V
 COMPARISON OF MODEL VALIDATION RESULTS BY ROOT MEAN SQUARE ERROR (RMSE)

Vehicle number	Root Mean Square Error (g/s)														
	CO			CO ₂			HC (10 ⁻²)			NO _x (10 ⁻²)			FUEL		
	MEF	VT-Micr	POLY	MEF	VT-Micr	POLY	MEF	VT-Micr	POLY	MEF	VT-Micr	POLY	MEF	VT-Micr	POLY
LDVI-1	0.066	0.073	0.063	1.428	1.514	1.602	0.14	0.18	0.16	0.59	0.62	0.56	0.482	0.508	0.536
LDVI-2	0.053	0.059	0.051	1.491	1.524	1.571	0.29	0.30	0.27	0.60	0.62	0.60	0.496	0.508	0.520
LDVI-3	0.049	0.054	0.047	1.218	1.226	1.236	0.28	0.30	0.28	0.90	0.92	0.78	0.406	0.414	0.417
LDVII-1	0.097	0.097	0.095	1.129	1.143	1.151	0.87	0.88	0.71	1.28	1.37	1.48	0.408	0.403	0.393
LDVII-2	0.117	0.129	0.139	1.314	1.406	1.419	0.46	0.49	0.55	1.40	1.57	1.66	0.529	0.567	0.580
LDVII-3	0.104	0.103	0.093	1.113	1.150	1.158	0.46	0.47	0.42	1.48	1.50	1.25	0.406	0.402	0.388

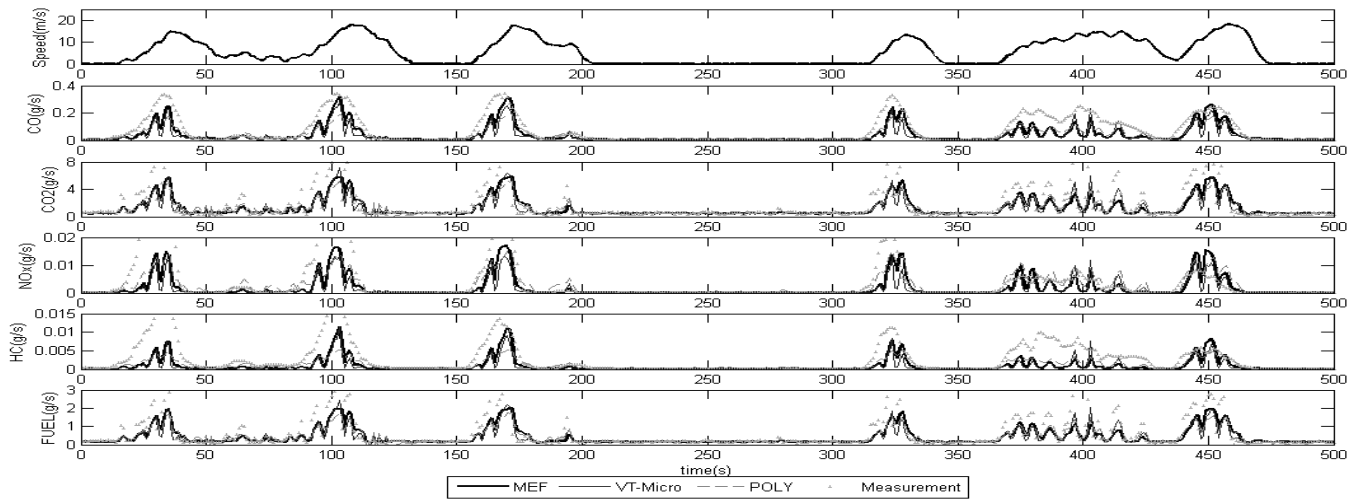


Fig. 3 An example of instantaneous validation of the MEF, VT-Micro, and POLY model for LDVI vehicles

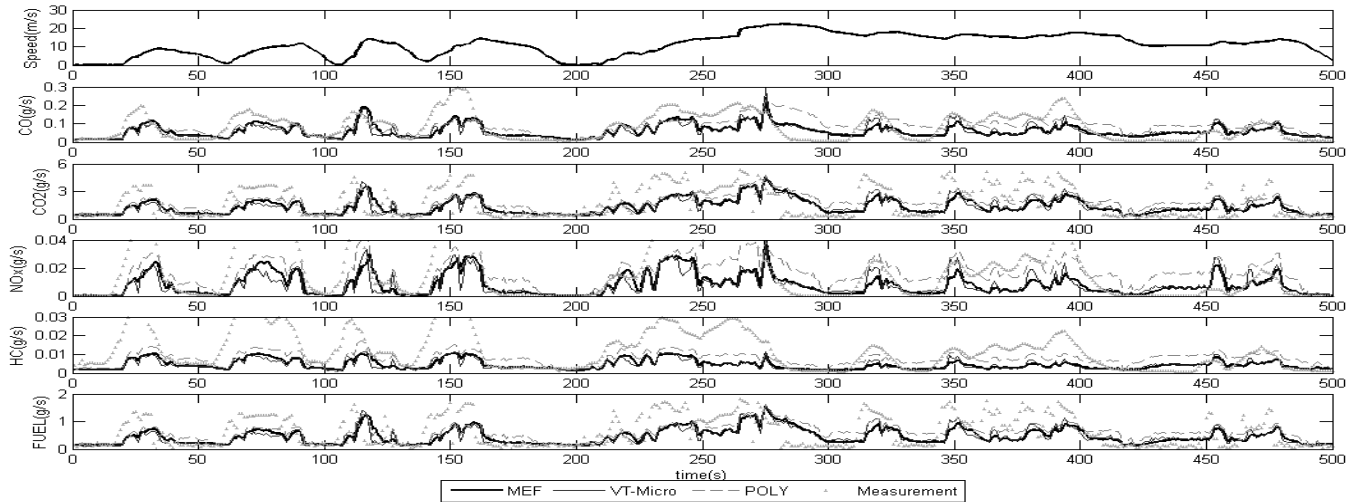


Fig. 4 An example of instantaneous validation of the MEF, VT-Micro, and POLY model for LDVII vehicles

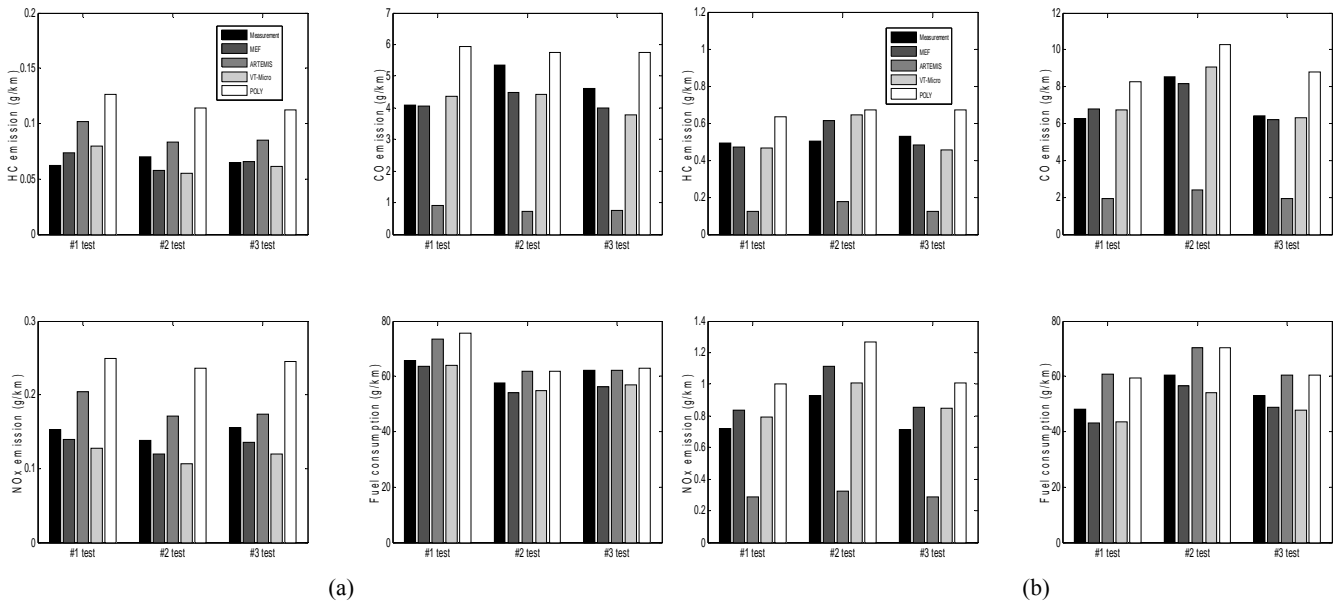


Fig. 5 Comparison of aggregate emission factors and fuel consumption derived from the MEF, ARTEMIS, VT-Micro, and POLY models with measurements for (a) LDVI vehicles; (b) LDVII vehicles.

VI. CONCLUSION

In this paper, we presented a framework for developing microscopic emission (HC, CO, NO_x, and CO₂) and fuel consumption models (MEF model) for light-duty vehicles. Based on the previous study on VT-Micro model and POLY model, we introduced composite acceleration, which is calculated as the weighted mean of acceleration of current time and accelerations of nine previous seconds, into MEF model with the purpose of capturing the effects of historical accelerations interacting with current speed on emission and fuel consumption. The data used in this paper were collected from 28 vehicles under real-world driving conditions in Beijing by a PEMS instrument (OBS-2200). All of the test vehicles were divided into two categories, LDVI category and LDVII category. The MEF model, as well as VT-Micro and POLY model, was separately calibrated for both categories using the collected PEMS data by multivariate least-squares regression method. Both instantaneous validation and aggregate validation were performed to verify the prediction accuracy of the MEF model. In instantaneous validation, the MEF model well captured the trend of real-world emission and fuel trajectories and showed best performance according to *MAPE* compared to the VT-Micro and the POLY model. In aggregate validation, the ARTEMIS model was also adopted in this paper for comparison with the MEF model, and the results showed that MEF model can give most accurate estimations compared to actual measurements.

In the future research, several issues related to the MEF model are supposed to be investigated further. First, the value of AIF (α) in the MEF model needs further analysis by using some numerical methods for model identification, such as Gaussian-Newton method and Levenberg-Marquardt (LM) method. Second, more PEMS data should be collected from other vehicle categories, such as trucks, bus, and high-duty vehicles, in order to broaden the application range of the MEF model. Third, the applicability of the MEF model, which is developed for Beijing, in other cities should be examined.

ACKNOWLEDGMENT

The authors would like to acknowledge the financial support of National Natural Science Foundation of China (Grant No. 50876078) and Self-determined and Innovative Research Funds of WUT.

REFERENCES

- [1] EPA, "User's Guide to MOBILE6.0. Mobile Source Emission Factor Model". EPA-420-R-02-001. United States Environmental Protection Agency, 2001.
- [2] CARB, "EMFAC User's Guide". U.S. California: California Air Resource Board, 2002.
- [3] C. Kouridis, L. Ntziachristos, and Z. Samaras, "COPERT III Computer programme to calculate emissions from road transport". European Environment Agency, 2000.
- [4] M. Keller, and N. Kljun, "ARTEMIS Road Emission Model User Guide". EU Commission, 2007.

- [5] EPA, "Draft Motor Vehicle Emission Simulator (MOVES) 2009". EPA-420-B-09-008. United States Environmental Protection Agency, 2009.
- [6] M. Barth, F. An, T. Younglove, G. Scora, C. Levine, M. Ross, and T. Wenzel, "Comprehensive Modal Emission Model (CMEM), version 2.0, User's Guide". University of California, Riverside, 2000.
- [7] H. Kuhns, C. Mazzoleni, H. Moosmuller, D. Nikolc, R. Keislar, P. Barber, L. Zheng, V. Etyemezian, and J. Watson, "Remote sensing of PM, NO, CO and HC emission factors for on-road gasoline and diesel engine vehicles in Las Vegas, NV". *Science of the Total Environment*, Vol. 322, No. 1-3, pp. 123-137, 2004.
- [8] N. Rogak, U. Pott, T. Dann, and D. Wang, "Gaseous emissions from vehicles in a traffic tunnel in Vancouver, British Columbia". *Air & Waste Manage. Assoc.*, Vol. 48, No.7, pp.604-615, 1998.
- [9] C. John, R. Friedrich, J. Staehelin, K. Schl pfer, and W.A. Stahel, "Comparison of emission factors for road traffic from a tunnel study (Gubrist tunnel, Switzerland) and from emission modeling". *Atmospheric Environment*, Vol. 33, No. 20, pp. 3367-3376, 1999.
- [10] F. Yang, and L. Yu, "A Microscopic Emission Model for the Light-Duty Vehicles Based on PEMS Data". In *Proceedings of International Conference of Transportation Engineering 2007 (ICTE 2007)*, Chengdu, China, 2007.
- [11] D.A.M.M. Elst, R.T.M. Smokers, and J.W. De Koning, "Evaluation of the Capabilities of On-Board Emission Measurement System for the Purpose of Generating Real-Life Emission Factors". TNO Report 04.OR.VM.039.1/DE, 2004.
- [12] G. Song, L. Yu, and X. Zhang, "Emission Analysis at Toll Station Area in Beijing with Portable Emission Measurement System". *Transportation Research Record: Journal of the Transportation Research Board*, 2058, pp.106-114, 2008.
- [13] X. Li, G. Li, S. Pang, X. Yang, and J. Tian, "Signal timing of intersections using integrated optimization of traffic quality, emissions and fuel consumption: a note". *Transportation Research Part D* 9, Vol. 9, pp. 401-407, 2004.
- [14] H. Liu, M. Barth, G. Scora, N. Davis, and J. Lents, "Using Portable Emission Measurement System for Transportation Emissions Studies: Comparison with Laboratory Methods". In *Proceedings of 89th Transportation Research Board Annual Meeting*. USA, Washington D.C., 2010.
- [15] H. Rakha, K. Ahn, and A. Trani, "Development of VT-Micro model for estimating hot stabilized light duty vehicle and truck emission". *Transportation Research Part D*, Vol.9, pp.49-74, 2004.
- [16] H. Teng, L. Yu, and Y. Qi, "Statistical micro-scale emission models incorporating acceleration and deceleration". In *Proceedings of 81th Transportation Research Board Annual Meeting*. USA, Washington D.C., 2002.
- [17] HORIBA, "On Board Emission Measurement System OBS-2200 Instruction Manual". HORIBA Ltd, 2005.
- [18] Y. Kamarianakis, and H.O. Gao, "Accounting for exhaust gas transport dynamics in instantaneous emission models via Smooth Transition Regression". In *Proceedings of 89th Transportation Research Board Annual Meeting*. USA Washington D.C., 2010.
- [19] R.K. Jr, "Data smoothing using a least squares fit C++ class". *ISA Transactions*, 37, pp.3-19, 1998.
- [20] M. Barth, F. An, J. Norbeck, and M. Ross, "Modal Emissions Modeling: A Physical Approach". *Transportation Research Record: Journal of the Transportation Research Board*, 1520, pp.81-88, 1996.
- [21] H. Guo, J. Zeng, and Y. Hu, "Neural Network Modeling of Vehicle Gross Emitter Prediction Based on Remote Sensing Data". In *Proceedings of the 2006 IEEE International Conference*. Taipei, Taiwan, 2006