Environ Model Assess (2012) 17:375–387
DOI 10.1007/s10666-011-9296-9

An Evaluation of Microscopic Emission Models for Traffic Pollution Simulation Using On-board Measurement

Xiaoliang Ma · Wei Lei · Ingmar Andréasson · Hui Chen

Received: 6 October 2010 / Accepted: 27 September 2011 / Published online: 27 October 2011

Abstract As a result of the continuously increasing numbers of motor vehicles in metropolitan areas worldwide, road traffic emission levels have been recognized as a challenge during the planning and management of transportation. Experiments were conducted to collect on-road emission data using portable emission measurement systems in two Chinese cities in order to estimate real traffic emissions and energy consumption levels and to build computational models for operational transport environment projects. In total, dynamic pollutant emissions and fuel consumption levels from dozens of light duty vehicles, primarily from four different vehicle classes, were measured at a second-by-second level. Using the collected data, several microscopic emission models including CMEM, VT-Micro, EMIT, and POLY were evaluated and compared through calibration and validation procedures. Non-linear optimization methods are applied for the calibration of the CMEM and EMIT models. Numerical results show that the models can realize performance levels close to the CMEM model in most cases. The VT-Micro model shows advantages in its unanimous performance and ability to describe low emission profiles while the EMIT model has a clear physics basis and a simple model structure. Both of them can be applied when extensive emission computation is required in estimating environmental impacts resulting from dynamic road traffic.

Keywords Vehicle emission · Traffic pollution · On-board data collection · Microscopic emission models · Calibration and validation

X. Ma (✉)
Faculty of Traffic and Logistics, Department of Transportation Sciences, Royal Institute of Technology (KTH),
Teknikringen 72,
Stockholm 10044, Sweden
e-mail: liang@kth.se

X. Ma · I. Andréasson
Centre for Traffic Research (CTR),
Royal Institute of Technology (KTH),
Teknikringen 72,
Stockholm 10044, Sweden

W. Lei · H. Chen
Intelligent Transportation System Centre (ITSC),
Wuhan University of Technology (WHUT),
Wuhan 430063 Hubei, China

1 Introduction

The rapid increase in number of transport vessels, especially vehicles for road transport, threatens sustainable development of our society in terms of energy consumption and traffic-induced pollutant emissions. At present, the transportation sector accounts for more than 20% of global energy consumption and road transport produces about 25% of principal greenhouse gases (GHG), considered to be the major factor of global climate change. Meanwhile, road traffic emissions are the main source of local pollutant emissions for urban areas such as carbon monoxide (CO), nitrogen oxides (NO_x), volatile organic compounds, and particulate matters (PMs). All of these emissions result not only in environmental problems but also add to the deterioration of human health and social welfare. Therefore, management of transport environments has become one of the most essential and urgent aspects for sustainable urban planning and management.

In recent years, China has become one of the largest energy consumers and producers of GHG emissions in the World due to its rapid economic growth. The industrialization process has, however, led to serious environmental deterioration undermining public health, air and water quality, agriculture, and ultimately the entire Chinese economy. Among all factors, emissions from road traffic are, unsurprisingly, identified as a major source of air pollution. Although emission regulations have been devised according to the European Union (EU) standards, quantitative analysis of traffic emissions in urban areas is indispensible for the evaluation of urban planning policies and for devising road traffic management strategies.

This paper describes our recent efforts collecting on-road mobile emission data in two Chinese cities, Beijing and Wuhan, using portable emission measurement systems (PEMS). Using collected data, four different microscopic emission models have been evaluated and compared through calibration and validation procedures. The main objective was to evaluate the applicability of those models with on-board emission data collected on urban roads in Chinese cities and to calibrate the models so that they can be applied in operational transportation projects, the main purpose of which is analysis and management of local traffic impacts. The remainder of this paper is organized into four sections. In the second section, we briefly describe the microscopic emission models. In the third section, previous work on vehicle emission data collection and analysis is described. The essential model evaluation procedures and results are presented in the fourth section. Finally, we summarize our study and identify future work towards environmental impact management of road traffic.

2 Microscopic Emission Models

One of the most fundamental research questions for environmental impact analysis of road transport is how to model the levels of vehicle exhaust emissions (CO_x, NO_x, HC, PM, etc.) and fuel consumption accurately. The current vehicle emission/energy models can generally be classified into aggregate models and microscopic models. Aggregate models are mostly used for transport environment planning with static approaches to estimate total or average traffic emission and fuel consumption. The models often require input data such as vehicle fleet composition, average speed on road links, total vehicle distance of travel, etc. For example, the MOBILE6 [24], MOVES [25], COPERT III [10], and ARTEMIS [11] models, developed by US Environmental Protection Agency (EPA) and EU respectively, are well-known aggregated emission models which have been applied in many transport environment planning projects.

With increasing concerns on management of dynamic traffic emission and energy consumption in local traffic environment, microscopic emission models have become a research focus due to their capability of evaluating traffic impacts in operational transportation projects prior to field implementation. The most important aspect is that they can be used together with microscopic traffic models for management of emission and energy consumption through dynamic traffic control and operations, e.g., signal control [15], in-vehicle or dynamic traffic management [19], and traffic calming [1].

Mathematically, a microscopic model can be represented by:

$$E_{i,j}(t) = f\left(C_j, a, v, \ldots, t, \Phi_{i,j}\right) \tag{1}$$

where $E_{i,j}(t)$is the emission or fuel consumption of gas type i for the vehicle category C_j and $\Phi_{i,j}$ represents parameters for the model. The models estimate vehicle fuel consumption and emissions based on instantaneous states of each individual vehicle on roads, e.g., vehicle speed $v(t)$, acceleration $a(t)$, engine performance indices, etc. A number of microscopic emission models have gained popularity in the literature such as CMEM, VT-Micro, POLY, and EMIT.

CMEM [2] is a load-based model developed at UC, Riverside using chassis dynamometer data of the US NCHRP database. The database includes original emission data collected from 300 automobiles and light trucks in 26 vehicle categories [21]. Second-by-second tailpipe emissions are modeled as a product of fuel rate (FR), engine-out emission index ($g_{\text{emission}}/g_{\text{fuel}}$), and catalyst pass fraction (CPF):

$$E_{\text{tailpipe}} = \text{FR} \times (g_{\text{emission}}/g_{\text{fuel}}) \times \text{CPF} \tag{2}$$

The model decomposes the physical emission generation process into modular components corresponding to physical phenomena: engine power, engine speed, air-to-fuel rate, fuel use, engine-out emission, and catalyst pass fraction. It calculates emission levels under stoichiometric, cold-start, enrichment, and enleanment conditions.

The VT-Micro model [20] adopts a third-order polynomial regression of instantaneous vehicle speed and acceleration. Mathematically, the model can be described by

$$\text{MOE}_e = \begin{cases} \exp\left(\sum_{i=0}^{3}\sum_{j=0}^{3} L_{i,j}^{e} \cdot v^i \cdot a^j\right) & \text{for } a \geq 0 \\ \exp\left(\sum_{i=0}^{3}\sum_{j=0}^{3} M_{i,j}^{e} \cdot v^i \cdot (-a)^j\right) & \text{for } a < 0 \end{cases} \tag{3}$$

where MOE_e (grams per second) is the instantaneous fuel consumption or emission rate of type e, $L^e{}_{i,j}$ and $M^e{}_{i,j}$ are regression model coefficients, v^i is the power i of the instantaneous vehicle speed (kilometers per hour), and a^j is the power j of the instantaneous vehicle acceleration or deceleration (kilometers per hour per second). A logarithm transformation is applied to MOE_e by the exponential form of Eq. 3, indicating non-negative values for MOE_e. The model was established for positive and negative accelerations separately to ensure a better compliance with the measurement data over the full range of the vehicle-operation envelope. The raw data utilized to develop this model was collected at the Oak Ridge National Laboratory and represents emissions from six light duty vehicles (LDVs) and three light trucks.

EMIT [3] is another microscopic emission model jointly developed by MIT and Ford. Based on similar but simplified ideas as in CMEM, the model consists of two parts: engine-out (EO) and tailpipe (TP) emission modules. The first module calculates instantaneous fuel consumption and engine-out emission rates of the emission species i using instantaneous speed v and acceleration a, i.e.,

$$\mathrm{EO}_i = \begin{cases} \alpha_i + \beta_i \cdot v + \gamma_i \cdot v^2 + \delta_i \cdot v^3 + \lambda_i \cdot a \cdot v & \text{if } p > 0 \\ \alpha_i^{'} & \text{if } p = 0 \end{cases} \quad (4)$$

where a_i, β_i, γ_i, δ_i, λ_i, and $\alpha_i^{'}$are model coefficients and p is the tractive power defined by:

$$p = A \cdot v + B \cdot v^2 + C \cdot v^3 + M \cdot a \cdot v + M \cdot g \cdot \sin\theta \cdot v \quad (5)$$

where A is the rolling resistance coefficient (kilowatts per meter per second); B is the speed correction to rolling resistance coefficient and C is the air drag resistance coefficient; M is the vehicle mass (kilograms) and g is the gravitational constant (9.81 m/s^2); and finally, θ is the road grade. The EO fuel consumption and emission rates are used in the second module to estimate tailpipe emission

$$\mathrm{TP}_i = \mathrm{EO}_i \cdot \mathrm{CPF}_i \quad (6)$$

where CPF_i is the catalyst pass fraction for gas species i and is simply a multi-regime linear regression of EO_i. Note that the tailpipe emission of CO_2 is estimated directly from Eq. 4 since it is not so much different from the engine-out CO_2. But for HC, CO, and NO_x, the EO and CPF should be estimated first and then their product is the final tailpipe-out rates for each pollutant gas.

In the development of POLY [23], vehicles are categorized into 41 classes according to vehicle size i, model year j, and emitter type k. For each vehicle category, a certain type m of emission is estimated by a regression model. The model formulation extends from the principle of the tractive power formula in Eq. 5 and takes into account several factors: instantaneous speed, acceleration or deceleration, specific power, and the extent of acceleration and deceleration. Differencing from other approaches, the POLY model includes historical acceleration components with a mathematical form of

$$\begin{aligned} e_{i,j,k,m} = \beta_0 &+ \beta_V \cdot v(t) + \beta_{V^2} \cdot v^2(t) + \beta_{V^3} \cdot v^3(t) + \beta_{T'} \cdot T'(t) \\ &+ \beta_{T''} \cdot T''(t) + \beta_{a_t} \cdot a(t) + \cdots + \beta_{a_{t-9}} \cdot a(t-9) \\ &+ \beta_W \cdot W(t) + \varepsilon_{i,j,k,m} \end{aligned} \quad (7)$$

where β_0 is a constant and β_x is the coefficient for variable x. $v(t)$ is vehicle speed at time t. $a(t)$ and $a(t-k)$ are combined acceleration or deceleration at time t and $t-k$ where k= 1,… 9; $T'(t)$ and $T''(t)$ are the corresponding acceleration and deceleration times up to t since their inception; $W(t)$ is the specific power at time t, which is equal to the product of $v(t)$ and $a(t)$; and $\varepsilon_{i,j,k,m}$ is the error term.

3 Data Collection and Preparation

Data collection is vital for modeling vehicle emission levels. In general, several methods have been widely applied for direct emission measurement including chassis dynamometer tests [2], remote sensing [6], and on-board emission measurement [7]. In addition, air quality data can be used indirectly to estimate actual vehicle fleet emission [18, 22]. The chassis dynamometer test is the most classical approach where vehicles or engines are placed on a test platform and operated according to legislated vehicle driving cycles so that emission factors can be obtained under controlled conditions. Many studies [9], however, have shown that the pre-defined driving cycles are unable to capture real emission patterns in comparison with on-road tests.

PEMS is a relatively new technique where real-time emissions can be measured on-board in test runs under real traffic conditions. With PEMS, tailpipe exhausts are collected and transferred to an on-board measurement unit through sampling pipes and then a gas analyzer estimates the real emission rates. This method makes it possible to study dynamic characteristics of vehicle emissions and has gained wide attention among both traffic and environment planners. The current development of MOVES [12] by US EPA has adopted the measurement approach to collect data. In addition, the method has been widely accepted by car manufacturers as an important method to evaluate vehicular emission performance.

3.1 PEMS Units

The data collected in this study is measured by mainly two PEMS in two Chinese cities: OEM-2100 used in Wuhan and OBS-2200 in Beijing. OBS-2200 consists of vibration-proof gas analyzers, a laptop PC with software for system control and data logging, accessory sensors, and a tailpipe attachment with a Pitot tube. The gas analyzing unit can calculate CO, CO_2, THC, and NO_x concentrations. In addition, other external signals such as temperature, air pressure, humidity, etc. are measured and saved into PC by the logging software [8]. OEM-2100 is another widely used on-board system consisting of two five-gas analyzers, an on-board computer and an engine diagnostic scanner [5]. The system can detect real-time engine and vehicle-operation parameters both from the OBD link and vehicle sensors. The main difference between OEM-2100 and OBS-2200 system is that HC is measured by an NDIR analyzer rather than a CLD analyzer. In addition, the OEM-2100 system calculates exhaust mass flow using engine parameters (e.g., intake air mass flow, engine displacement, and engine speed), whereas OBS-2200 measures flow volume directly using a flow meter.

3.2 Data Collection

One essential objective of the study is to build an emission database for traffic-induced air pollution research based on PEMS measurement of different vehicle classes. The measurement in Wuhan was performed by an OEM-2100 system between November and December of 2006. Vehicle types, driver characteristics, test time, and routes were carefully considered for the design of on-road data collection procedures. The experiment targeted at several popular brands (see Table 1), but the majority is LDV4 vehicles (see Table 2) according to the CMEM classification criteria. The data collection activity using OBS-2200 systems was carried out in Beijing in cooperation with the Chinese Automotive Technology and Research Center (CATARC) in 2007. The equipment was well calibrated and validated by dynamometer test data at CATARC. In total, 43 LDVs were recruited for the two tests (see Table 2). The test in Beijing covers also several popular passenger car makes including FAW Red Flag, VM Jetta, VM Satana, and Hyundai Elantra, all of which were less than 5 years old at the time. Engine displacements of vehicles in the two tests range from 1.6–2.7 l and their odometers cover from 2,000–300,000 km. Due to limited resources, only LDVs were measured in the experiment.

Driver behavior affects vehicle emissions significantly, e.g., aggressive behavior may produce more emissions than during normal driving behavior. Therefore, test drivers are all required to have more than 5 years of driving experience. They were requested to follow traffic flows without abnormal behavior (e.g., frequent lane changing and overtaking) so that real traffic characteristics can be captured in the study. To reflect common traffic conditions, test runs were performed between 7:00 am and 8:00 pm in order to cover both rush hours and non-rush hour periods. In order to reflect different road conditions, the test route is set up consisting of highways, urban arterial roads, and collector (feed) roads. All experimental data were collected from test runs on the route. Figure 1 shows the test route in Wuhan in which different types of urban roads are included. In general, more than 80,000 data records were collected. Each record comprises vehicle speed (kilometers per hour), fuel consumption rate (grams per second), emission rate (grams per second) for four gases (HC, CO, NO_x, CO_2), air-to-fuel ratio, exhaust gas temperature (degrees Celsius), relative humidity, position, and altitude (meters). In the collected data, the vehicle speed values range from 0–110 km/h, and vehicle acceleration values vary from −3.5–3.5 m/s^2.

Table 1 Average emission factor of the popular makes of passenger vehicle samples in Wuhan

Vehicle category	NO_x (g/km)	HC (g/km)	CO (g/km)
Citroen Elysee	1.05	1.55	42.73
Citroen Fukang	1.89	1.92	30.34
VM Jetta	1.96	2.58	18.60
Hyundai Elantra	0.32	0.39	2.96
Euro II limits	0.225	0.275	2.2

3.3 Data Preparation

Measurement data were saved in an emission database created in a MySQL server. The database includes instantaneous emission levels and fuel measurements as well as vehicle information such as license, brand, category, mileage, engine displacement, production year, etc. In addition, a data processing program was developed in MATLAB scripts to connect with the MySQL server so that we can easily query and manipulate data for further analysis. Emission measurement using PEMS is a rather delicate procedure unavoidably involving noise and interruptions. Unreasonable data, such as negative measurements because of zero-drift problems in the test equipment, are removed from the samples. In addition, since emission and speed data are not synchronized due to the flow blockage of exhaust gas in the sampling pipe, a data post-processing procedure [13] was also performed to offset the time lag in order to ensure that vehicle real-time states (speed and acceleration) are associated with instantaneous emission and fuel consumption rates in a correct manner.

In order to evaluate the experiment in Wuhan, detailed data analysis was carried out by Lei [13]. In particular, passenger car samples of four most popular makes are evaluated and compared to the Euro II standards, with which they should comply (see Tables 1 and 2). Tested vehicles, however, emit much more pollutants than standard values probably because the increasing mileage (LDV4, corresponding to more than 50,000 miles) leads to fast deterioration of vehicle operating status. The relatively newer Elantra vehicles (LDV6) show better emission performance. In general, average emission levels of the LDV6 vehicles measured in Wuhan are comparable to those tested in Beijing, whereas the emission levels of

Table 2 Classification of tested light duty vehicles

Vehicle category	Mileage	Power-to-weight ratio	Number of tested vehicles
LDV4	≥50,000	low (<0.039 hp/lb)	25
LDV5	≥50,000	high (>0.039 hp/lb)	1
LDV6	<50,000	low (<0.039 hp/lb)	16
LDV7	<50,000	high (>0.039 hp/lb)	1

All vehicles in the categories use fuel injection and three-way emission control

LDV4 vehicles measured in Wuhan are higher than those tested in Beijing. This indicates that the inspection regulations might need to be strengthened in Wuhan and high-emitting vehicles should be strictly eliminated from roads for improving air quality.

4 Model Evaluation

In this section, we present our model evaluation study based on emission data collected in the previous experiment. Four models described in the second section are analyzed and compared in the model calibration and validation procedures.

4.1 Vehicle Classification

Vehicle classification is an important procedure [2] for emission modeling since different categories of vehicles have different emission properties. Aggregation of vehicles with similar emission characteristics into homogeneous classes is a common way of simplifying the final model structure. Several vehicle classification methods were proposed in previous research [2, 3, 20, 23], among which the classification criteria of the CMEM model are widely accepted and therefore adopted in the study. In general, the 43 test vehicles (all are gasoline LDV) are classified into four categories according to emission control technology, accumulated mileage and power-to-weight ratio. Table 2 illustrates the classification results based on the CMEM category standards [21].

4.2 Model Calibration

For model calibration, experimental data from ten LDV4 and eight LDV6 vehicles were selected to formulate composite data samples. For both categories, vehicle

Fig. 1 Test routes (represented by *thick lines* in the map) in Wuhan, Hubei Province

mileages in the composite data samples cover a wide scope of the critical mileage values. For the LDV5 and LDV7 classes, model parameters are also estimated although only data from a single vehicle are available. In total, about 22,000 and 20,000 data records were used to formulate composite data sets for LDV4 and LDV6 in calibration. For LDV5 and LDV7, we randomly chose a majority of the data samples for calibration whereas the rest were used for validation. Figure 2 illustrates the speed/acceleration histogram for composite data sets of LDV4. The majority of speed and acceleration data were collected during stationary state (acceleration ranging between −0.5 and 0.5 m/s^2).

The latest CMEM 3.01 calculation software was used in the study. In calibration, one essential challenge is that the software does not support any calibration against real data explicitly. Therefore, we have developed a two-stage procedure [14] to recalibrate the CMEM model. The basic idea is to adjust some fuel-related parameters in the CMEM model to fulfill least square error (LSE) concerning fuel rate. Then, by fixing fuel-related parameters, emission-related parameters are tuned to implement LSE in tailpipe emissions. In both stages, a non-linear optimization problem has to be solved to find proper parameters, i.e.,

$$\min_{\theta} \Phi(\theta) = \sum_t \left(\widehat{S}(t) - S_{\mathrm{CMEM}}(v(t), \theta)\right)^2 \quad \theta \in \Omega \qquad (8)$$

where θ is parameter vector and Ω defines a set of parameter constraints that make the CMEM model valid, and $\hat{S}$ and S_{CMEM} are the measurements and CMEM model predictions, respectively. We have applied a combination of grid search and non-linear simplex methods in solving the optimization of Eq. 8. That is, some initial candidates were found by evaluating mesh grid points in the valid parameter space of CMEM; then using the candidates as initial vector, the best parameter vector is found by applying the non-linear simplex algorithm.

For the VT-Micro and POLY models, multiple linear regression procedures were applied in MATLAB to get model coefficients in Eqs. 3 and 7. Some related statistics can be derived simultaneously given second-by-second speed and acceleration inputs and corresponding emission and fuel consumption data.

Calibration of the EMIT model poses another challenge as the engine-out emission data are not available in PEMS measurements. This means that the parameters in Eq. 4 cannot be estimated directly through linear regression. Therefore, we simplify the calculation of CPF with a single regime model. Using collected emission data as reference, we reformulated the model calibration problem as a non-linear LS optimization problem for each emission factor i as follows:

$$\min_{\varphi} \Psi_i(\varphi) \qquad (9)$$

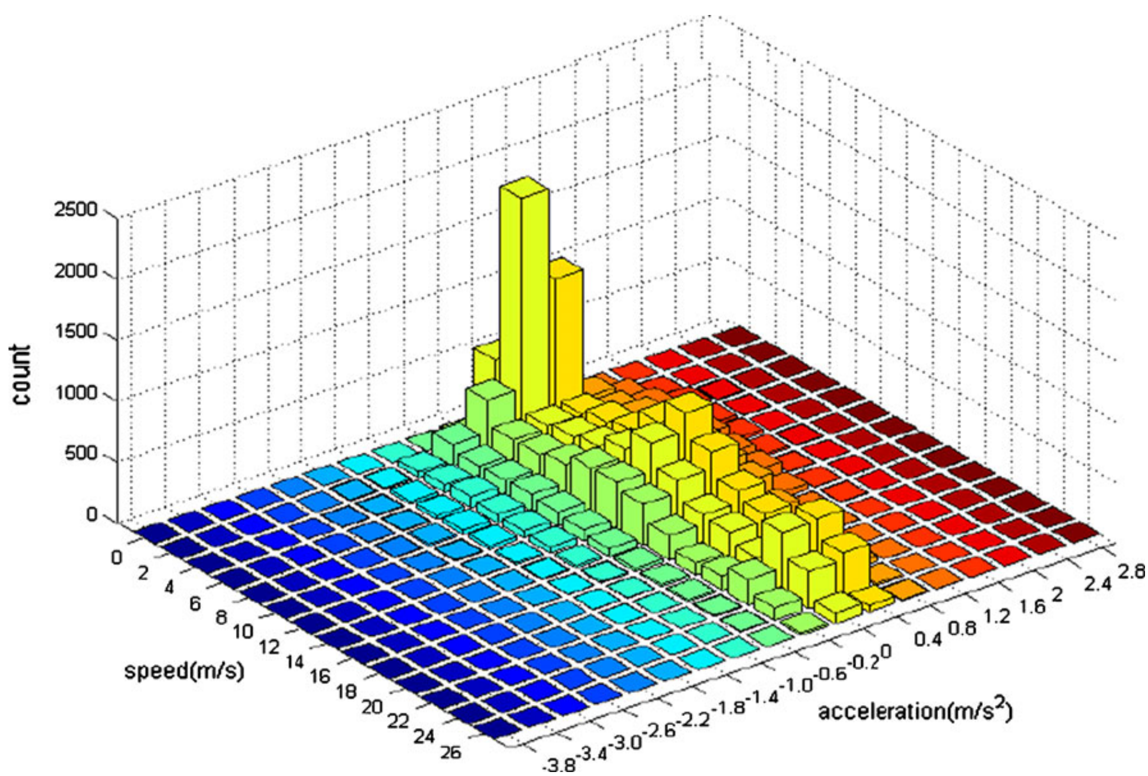

Fig. 2 Speed and acceleration histogram of the composite datasets for LDV4

where

$$\Psi_i(\varphi) = \mathbf{e}(\varphi)^T\mathbf{e}(\varphi) = \sum_{t\in\mathbf{T}} \left(\widehat{e}_i^{\mathrm{TP}}(t) - e_i^{\mathrm{EMIT}}(v(t), a(t), \varphi)\right)^2 \tag{10}$$

$$e_i^{\mathrm{EMIT}}(v(t), a(t), \varphi) = \mathrm{EO}_i(t)\cdot \mathrm{CPF}_i(t) \tag{11}$$

$$\mathrm{CPF}_i(t) = m_i \mathrm{EO}_i(t) + q_i \tag{12}$$

$$\mathrm{EO}_i(t) = \begin{cases} \alpha_i + \beta_i\cdot v(t) + \gamma_i\cdot v^2(t) + \delta_i\cdot v^3(t) + \lambda_i\cdot a(t)\cdot v(t) & \text{if } p(t) > 0 \\ \alpha_i' & \text{if } p(t) = 0 \end{cases} \tag{13}$$

s. t.

$$\mathrm{EO}_i(t) = a_i + \beta_i\cdot v(t) + \gamma_i\cdot v^2(t) + \delta_i\cdot v^3(t) + \lambda_i\cdot a\cdot v(t) > 0, \quad v(t) > 0 \quad t\in\mathbf{T} \tag{14}$$

$$\mathrm{EO}_i(t) = \alpha_i + \beta_i\cdot v(t) + \gamma_i\cdot v^2(t) + \delta_i\cdot v^3(t) + \lambda_i\cdot a(t)\cdot v(t) < \kappa_i, \quad v(t) > 0, \quad \kappa_i > 0 \quad t\in\mathbf{T} \tag{15}$$

where $\varphi = [m\; q\; \alpha\; \beta\; \gamma\; \delta\; \lambda]^T$ is a parameter vector, and $\widehat{e}_i^{\mathrm{TP}}(t)$is the tailpipe emission measurement and $e_i^{\mathrm{EMIT}}(v(t), a(t), \varphi)$is the model prediction value of tailpipe emission of factor i. In addition, since model estimation of Eqs. 9–13 can lead to unreasonable engine-out emission (i.e., negative or unexpected large emissions), numerical constraints are added to the engine-out outputs. κ_i is a numerical upper bound for engine-out emission of the factor i. Such a formulation significantly augments the complexity of the original least square (LS) optimization problem. For the typical non-linear programming without constraints, various numerical approaches can be applied [16]. For example, a widely used method is the Levenberg–Marquardt algorithm:

$$\varphi_{k+1} = \varphi_k - \frac{1}{2}\left(\mathbf{J}_k^T\mathbf{J}_k + \eta\cdot\mathbf{I}\right)^{-1}\mathbf{g}_k \tag{16}$$

where $\mathbf{J}_k = \partial\mathbf{e}(\varphi_k)/\partial\varphi_k$ is the Jacobian matrix of the error function whereas $\mathbf{g}_k = \nabla_\phi\Psi(\varphi_k)$is the gradient vector of $\Psi(\varphi_k)$; η is a small positive value; and $\mathbf{I}$ is the unit matrix. This approach has been implemented in software tools such as MATLAB. However, the method may suffer from frequent existences of local minima in real applications. Therefore, multiple initial points are normally assigned first and then numerical searches are conducted iteratively from each initial guess. The final solution is the one with the best fit,[1] i.e., realizes the smallest LS error. The inclusion of numerical constraints of Eq. 14 makes the problem more difficult and expensive to solve [16]. In our numerical studies, the Optimization Toolbox in MATLAB is applied in which efficient algorithms, e.g., interior point method, are implemented. Given the size of the data, the problem is computationally expensive (each run may take dozens of hours when doing calculations using a Pentium Dual-Core 2.8 GHz, 4 GB desktop computer with the Linux operating system), and especially it is necessary to fit the model for three different types of emission factors, i.e., HC, CO, and NO_x. To accelerate the computation, the parallel computing technique in MATLAB [17] has been adopted in the analysis, i.e., numerical searches from different random initial estimates towards final convergence can be implemented with parallelism. This can successfully reduce nearly half of the computation time being used for solving numerical optimization without parallelization, thanks to the dual-core processor. The computation in constrained optimization normally takes longer than that without constraints even if the same parallel computing technique is applied to reduce computation time.

Based on the corresponding approaches illustrated above, we have calibrated all four models with the formulated composite data sets. Table 3 summarizes the values of adjusted R^2 (statistically considered as an effective goodness-of-fit measure in multiple linear regression models) for two regression models: VT-Micro and POLY. Tables 4, 5 and 6 show the calibrated model parameters for various emission factors including CO_2, CO, and NOx, and the corresponding t statistic for each parameter of the LDV4 vehicle class. To keep the same formats as the original version of VT-Micro and POLY models presented by Rakha et al.[20] and Teng et al.[23], respectively, the parameters that are not statistically significant (comparing to the t statistics at the 5% significance level) still remain in the models. In addition, according to the results in the POLY model, all the terms related to the definition of vehicle traction power in the Eq. 5 are remarkably different in comparison to the

Table 3 Goodness of fit (adjusted R^2) of the models for the LDV4 vehicle category

Emission types	VT-Micro		POLY
	Acceleration	Deceleration	
CO_2	0.4325	0.3002	0.4469
CO	0.3514	0.2229	0.1656
NO_x	0.4640	0.2911	0.2203
HC	0.2544	0.1260	0.1751

[1] A sub-optimal point may be selected if it gives better physical meaning to the model.

Table 4 Calibrated parameters of CO_2 emission for the LDV4 vehicle category

VT-Micro				EMIT		POLY	
Acceleration		Deceleration					
Coefficient	Value	Coefficient	Value	Coefficient	Value	Coefficient	Value
$L_{0,0}$	−0.6700[a]	$M_{0,0}$	−0.6700[a]	α'	0.63284	β_0	0.71517[a]
$L_{0,1}$	0.2974[a]	$M_{0,1}$	0.0166	α	0.15797	β_v	0.06229[a]
$L_{0,2}$	−0.0170	$M_{0,2}$	−0.0086	β	−0.00925	$\beta_v{}^2$	−0.00165[a]
$L_{0,3}$	−0.0019	$M_{0,3}$	−0.0007	γ	0.00039	$\beta_v{}^3$	3.34e−05[a]
$L_{1,0}$	0.0121[a]	$M_{1,0}$	0.0258[a]	δ	0.17096	$\beta_T{}'$	0.02493[a]
$L_{2,0}$	0.0004[a]	$M_{2,0}$	−0.0002	λ	0.70556	$\beta_T{}'''$	0.00661[b]
$L_{3,0}$	−3.67e−06[a]	$M_{3,0}$	2.27e−06			$B_{a(t)}$	0.06461[a]
$L_{1,1}$	0.0283[a]	$M_{1,1}$	0.0043			$B_{a(t-1)}$	0.00368
$L_{2,1}$	−0.0009[a]	$M_{2,1}$	0.00043			$B_{a(t-2)}$	−0.00587
$L_{3,1}$	9.44e−06[a]	$M_{3,1}$	−4.62e−06			$B_{a(t-3)}$	0.01049
$L_{1,2}$	−0.0077	$M_{1,2}$	0.0007			$B_{a(t-4)}$	0.04845[a]
$L_{2,2}$	0.0003[b]	$M_{2,2}$	5.61e−05			$B_{a(t-5)}$	0.01399
$L_{3,2}$	−3.73e-06[a]	$M_{3,2}$	−5.11e−07			$B_{a(t-6)}$	0.01957
$L_{1,3}$	0.0007	$M_{1,3}$	3.39e−05			$B_{a(t-7)}$	−0.01045
$L_{2,3}$	−2.65e−05[b]	$M_{2,3}$	1.28e−06			$B_{a(t-8)}$	−0.00258
$L_{3,3}$	3.38e−07[a]	$M_{3,3}$	−7.61e−10			$B_{a(t-9)}$	0.00908
						β_W	0.02907[a]

[a] Levels of significance at 1% level

[b] Levels of significance at 5% level

Table 5 Calibrated parameters of CO emission for the LDV4 vehicle category

VT-Micro				EMIT		POLY	
Acceleration		Deceleration					
Coefficient	Value	Coefficient	Value	Coefficient	Value	Coefficient	Value
$L_{0,0}$	−4.7400[a]	$M_{0,0}$	−4.7400[a]	α'	0.0276	β_0	0.0285[a]
$L_{0,1}$	0.8933[a]	$M_{0,1}$	−0.2942[a]	α	12.4035	β_v	0.0077[a]
$L_{0,2}$	−0.2148[a]	$M_{0,2}$	−0.0454	β	0.0250	$\beta_v{}^2$	−0.0005[a]
$L_{0,3}$	0.0139[b]	$M_{0,3}$	−0.0024	γ	0.0037	$\beta_v{}^3$	9.94e−06[a]
$L_{1,0}$	0.0661[a]	$M_{1,0}$	0.1149[a]	δ	0.0008	$\beta_T{}'$	−0.0048[a]
$L_{2,0}$	−0.0011[a]	$M_{2,0}$	−0.0029[a]	λ	0.1538	$\beta_T{}'''$	0.0021[a]
$L_{3,0}$	8.07e−06[a]	$M_{3,0}$	2.47e−05[a]	m	0.0015	$B_{a(t)}$	−0.0090
$L_{1,1}$	−0.0266[b]	$M_{1,1}$	0.0653[a]	q	−0.0152	$B_{a(t-1)}$	−0.0007
$L_{2,1}$	0.0004	$M_{2,1}$	−0.0020[a]			$B_{a(t-2)}$	−0.0005
$L_{3,1}$	−2.43e−07	$M_{3,1}$	1.60e−05[a]			$B_{a(t-3)}$	0.0004
$L_{1,2}$	0.0095	$M_{1,2}$	0.0121[b]			$B_{a(t-4)}$	0.0007
$L_{2,2}$	−0.0001	$M_{2,2}$	−0.0004[b]			$B_{a(t-5)}$	0.0007
$L_{3,2}$	−5.37e−08	$M_{3,2}$	3.92e−06[b]			$B_{a(t-6)}$	0.0014
$L_{1,3}$	−0.0007	$M_{1,3}$	0.0007			$B_{a(t-7)}$	−0.0004
$L_{2,3}$	6.44e−06	$M_{2,3}$	−2.82e−05			$B_{a(t-8)}$	−0.0002
$L_{3,3}$	2.76e−08	$M_{3,3}$	3.00e−07			$B_{a(t-9)}$	−0.0043
						β_W	0.0024[a]

[a] Levels of significance at 1% level

[b] Levels of significance at 5% level

Table 6 Calibrated Parameters of NO_x emission for the LDV4 vehicle category

VT-Micro				EMIT		POLY	
Acceleration		Deceleration					
Coefficient	Value	Coefficient	Value	Coefficient	Value	Coefficient	Value
$L_{0,0}$	−9.0000[a]	$M_{0,0}$	−9.0000[a]	α'	0.0003	β_0	0.00052[b]
$L_{0,1}$	1.54553[a]	$M_{0,1}$	−0.5642[a]	α	0.0142	β_v	0.00073[a]
$L_{0,2}$	−0.2239[a]	$M_{0,2}$	−0.1352[a]	β	0.0387	β_v^2	−2.38e−05[a]
$L_{0,3}$	0.0026	$M_{0,3}$	−0.0068[b]	γ	−0.0024	β_v^3	3.09e−07[a]
$L_{1,0}$	0.13445[a]	$M_{1,0}$	0.1585[a]	δ	7.37e-05	β_T'	0.00023[a]
$L_{2,0}$	−0.00205[a]	$M_{2,0}$	−0.0032[a]	λ	0.0053	β_T'''	0.00013[a]
$L_{3,0}$	1.17e-05[a]	$M_{3,0}$	2.35e−05[a]	m	0.0088	$B_{a(t)}$	0.00013
$L_{1,1}$	−0.02569	$M_{1,1}$	0.0948[a]	q	0.0396	$B_{a(t-1)}$	1.54e−05
$L_{2,1}$	6.38e−05	$M_{2,1}$	−0.0020[a]			$B_{a(t-2)}$	4.42e−05
$L_{3,1}$	3.65e−06	$M_{3,1}$	1.36e−05[a]			$B_{a(t-3)}$	0.00032
$L_{1,2}$	−0.00325	$M_{1,2}$	0.02062[a]			$B_{a(t-4)}$	0.00042[b]
$L_{2,2}$	0.00035	$M_{2,2}$	−0.0005[b]			$B_{a(t-5)}$	0.00040[b]
$L_{3,2}$	−4.45e−06	$M_{3,2}$	4.15e−05[b]			$B_{a(t-6)}$	0.00038[b]
$L_{1,3}$	0.00132	$M_{1,3}$	0.0010[b]			$B_{a(t-7)}$	8.61e−05
$L_{2,3}$	−5.64e−05[b]	$M_{2,3}$	−2.96e−05			$B_{a(t-8)}$	−0.00022
$L_{3,3}$	5.72e−07[a]	$M_{3,3}$	2.67e−07			$B_{a(t-9)}$	0.00057[a]
						β_W	0.00024[a]

[a] Levels of significance at 1% level

[b] Levels of significance at 5% level

historical acceleration terms. In particular, the coefficients of the V^3 term are always positive which indicates that the final emission is proportional to the traction power. Meanwhile, the corresponding parameters of the term for the specific power W are also positive, proving that more kinetic power leads to higher level of emission pollutants. The pure regression-based formulation in VT-Micro makes it difficult to analyze the physical meaning of the calculated model coefficients. However, according to the adjusted R^2 value, the VT-Micro model achieves relatively better calibration results than the POLY model, although it does not give equally sound performance in the deceleration stage.

4.3 Model Validation

After calibration of different vehicle classes using composite data sets, it is important to analyze prediction performance of emission models using out-of-sample data sets meaning data not included in the calibration procedure. We have chosen continuous emission time series data sets of individual vehicles for validation. There are several statistical indices for evaluation of model performance, among which the Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE) are the most widely accepted measures. The RMSE is estimated by

$$\mathrm{RMSE} = \sqrt{\sum\nolimits_{i=1}^{N} \left(y_i - \widehat{y}_i\right)^2 / N} \tag{17}$$

where y_i is the raw data obtained from the on-board emission experiments, $\hat{y}_i$ is the emission model prediction (g/s), and N is the size of samples. The MAPE index is computed by:

$$\mathrm{MAPE} = \sum_{i=1}^{N} \left|y_i - \widehat{y}_i\right| / \Sigma_{i=1}^{N} y_i \tag{18}$$

Tables 7 and 8 show the model validation results using out-of-sample vehicle emission time series data sets and two statistical measures mentioned above. Because of the wide acceptance of the CMEM model in many transportation projects, it is used as an essential reference for comparison of model performance. It is not difficult to tell from the tables that RMSE and MAPE do not necessarily produce consistent validation results. With the RMSE measure, there is no clear winner (for instance VT-Micro is a good performer for predicting CO whereas POLY performs well in predicting NO_x whereas EMIT is good at predicting performance on average.

Table 7 Comparison of the validation results by RMSE

Vehicle type	Root mean square error (g/s)																			
	CO_2				CO				NO_x (10^{-2})				HC (10^{-2})				Fuel			
	VT- Micro	EMIT	POLY	CMEM	VT- Micro	EMIT	POLY	CMEM	VT- Micro	EMIT	POLY	CMEM	VT- Micro	EMIT	POLY	CMEM	VT- Micro	EMIT	POLY	CMEM
LDV4-1	1.190	1.185	1.179	1.228	0.093	0.116	0.099	0.059	2.798	2.802	2.405	2.403	1.461	1.314	1.310	1.217	0.422	0.419	0.418	0.426
LDV4-2	0.869	0.838	0.843	0.904	0.103	0.132	0.088	0.062	3.115	3.087	2.598	2.802	1.422	1.219	1.229	1.241	0.318	0.299	0.300	0.318
LDV5	3.222	3.099	3.137	3.354	1.375	1.252	1.247	1.325	1.118	1.033	0.974	0.734	0.336	0.344	0.336	0.376	1.575	1.487	1.502	1.653
LDV6-1	1.399	1.387	1.405	1.343	0.101	0.126	0.091	0.083	0.933	0.895	0.873	0.687	0.242	0.208	0.207	0.207	0.472	0.470	0.476	0.454
LDV6-2	1.546	1.539	1.554	1.621	0.130	0.140	0.124	0.104	1.282	1.232	1.210	0.938	0.174	0.174	0.178	0.164	0.521	0.518	0.522	0.551
LDV7	2.056	1.911	1.992	1.521	0.397	0.395	0.423	0.389	0.750	0.695	0.687	0.490	0.305	0.308	0.301	0.312	0.779	0.707	0.763	0.628

Table 8 Comparison of the validation results by MAPE

Vehicle type	Mean absolute percentage error (%)																			
	CO_2				CO				NO_x				HC				Fuel			
	VT- Micro	EMIT	POLY	CMEM	VT- Micro	EMIT	POLY	CMEM	VT- Micro	EMIT	POLY	CMEM	VT- Micro	EMIT	POLY	CMEM	VT- Micro	EMIT	POLY	CMEM
LDV4-1	63	71	71	73	72	89	90	53	89	85	79	82	89	76	77	84	64	72	71	72
LDV4-2	42	46	46	48	65	56	58	38	90	78	73	84	90	72	73	86	43	45	45	47
LDV5	45	45	47	43	90	79	80	96	79	87	81	72	68	70	68	89	53	50	52	52
LDV6-1	44	48	48	47	74	71	74	63	91	97	99	88	79	89	83	79	45	48	49	47
LDV6-2	71	75	76	75	78	79	81	73	94	93	92	86	84	138	134	103	70	74	75	75
LDV7	40	37	40	27	96	90	96	97	98	96	98	75	75	77	75	88	43	39	43	31

According to the validation results in terms of the MAPE index, the VT-Micro model generally gives the best prediction performance in modeling CO_2 emission and fuel consumption. Both the EMIT and POLY models can achieve performance similar to the calibrated CMEM model. For the NO_x emission factor, both EMIT and POLY perform equally well in comparison to the prediction of the CMEM model, and the POLY model shows a slight advantage in its performance when predicting the NO_x emission for the vehicle category LDV4. For the CO factor, all models perform almost equally acceptable though they are on average inferior compared with the CMEM model. Concerning HC, the VT-Micro realizes similar error percentages to the CMEM model. Both the EMIT and POLY models also perform well.

Figure 3 illustrates a validation example of emission predictions for the vehicle category LDV6 in a driving cycle. From the validation curves in the graphs, it is not difficult to tell that all models are able to capture the trends of CO_2 emissions and fuel consumption measurements although their individual performances show slight variations. On the other hand, none of the models are able to completely reflect the trends of the emissions of CO, NO_x, and HC, especially when the levels are highly fluctuating. When the emission level keeps a low profile such as NO_x, the VT-Micro model shows its relative advantage since both the POLY and EMIT models tend to overestimate the emission results. One major drawback of the POLY model is that it may produce negative emission values during predictions. On the other hand, although the EMIT model may also lead to negative emission outputs in its original development [4], the calibration method that we have applied by adding optimization constraints makes it less likely to obtain negative engine-out emissions as well as tailpipe emissions

5 Summary and Discussion

In order to implement local traffic environment analysis for sustainable urban development in China, two portable emission data collection systems, OEM-2100 and OBS-2200, were used in this study to collect second-by-second emission patterns in two Chinese cities. After initial reduction, emission data of about 20 vehicles were selected to formulate composite calibration data sets for vehicle categories LDV4, LDV5, LDV6, and LDV7 based on the vehicle classification proposed in the CMEM model. The main concern of this paper was to conduct model calibration and validation using

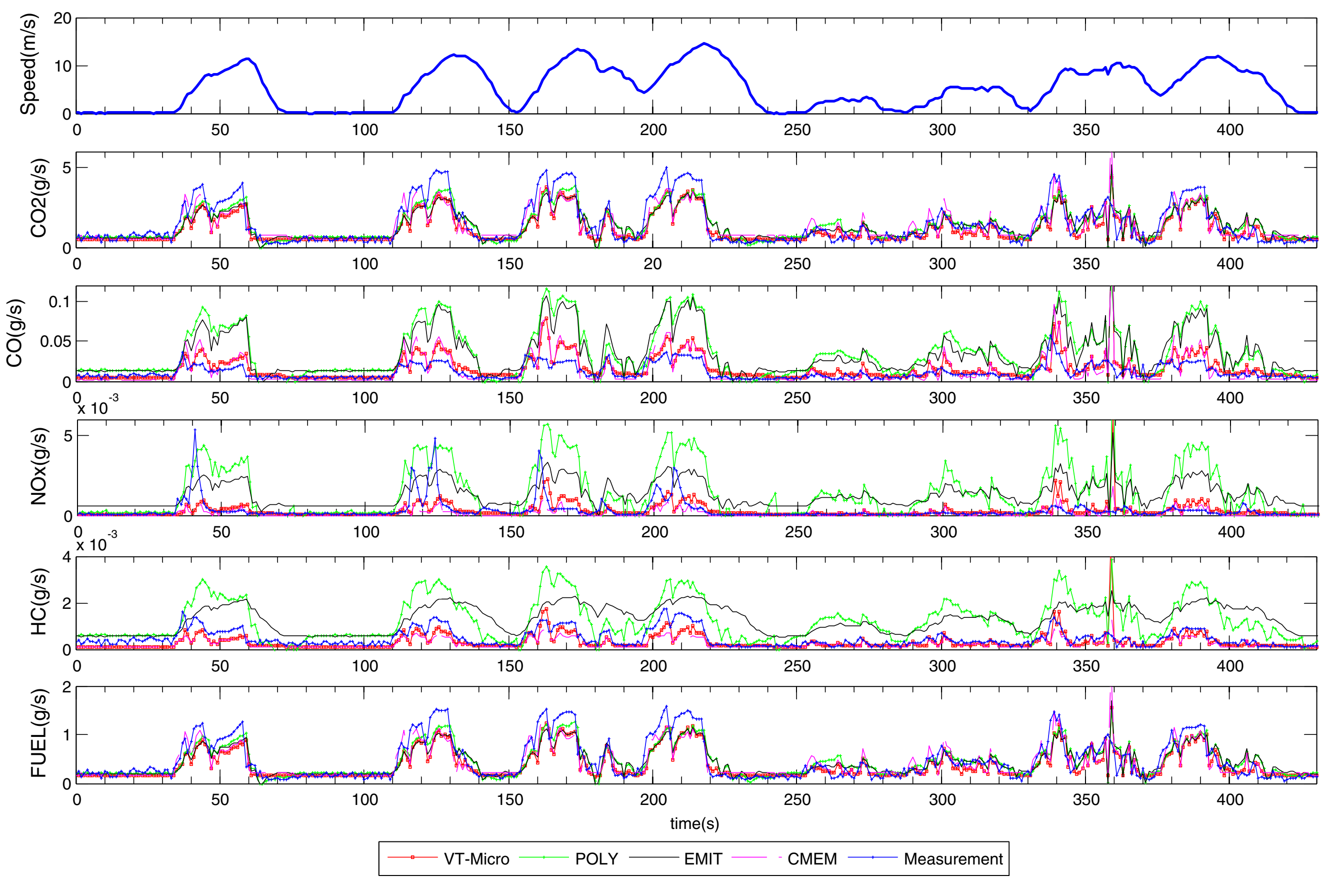

Fig. 3 An example of the instantaneous model validation for the LDV6 vehicle category

local data sets so that the existing models are able to be compared according to their prediction performances and moreover applied in later application projects on traffic impact analysis and management. Due to the lack of engine-out emission data, the EMIT model was calibrated using a non-linear least square approach and gradient-based numerical schemes help to find the final model parameters. In addition, parallel computing was applied to accelerate the numerical optimization procedure. The CMEM 3.01 software was also calibrated by a two-stage numerical tuning process in which selected parameters in the CMEM model were re-estimated according to fuel and emission measurements.

The models were finally compared according to the RMSE and MAPE validation indices using the CMEM model as a standard reference. Numerical results show that all models are better able to predict the CO_2 emissions and fuel consumption levels than the HC, CO, and NO_x emissions. Although it is difficult to completely reflect the trends for the emissions of HC, CO, and NO_x, especially with sudden and intensive fluctuations, all the models can achieve performance comparable to the calibrated CMEM model in most of the cases. The VT-Micro model shows advantages in its unanimous performance and ability of producing low emission profiles, while the EMIT model has a good general prediction capacity. Although the calibration of the EMIT model with the PEMS data is computationally expensive, the model is formulated according to the physics principle and has a much simpler structure than the load-based CMEM model. Therefore, both the VT-Micro and EMIT models have the potential to be applied for further environmental impact analysis of road traffic on the microscopic level despite representing different approaches.

Collection of emission data is a delicate and expensive procedure. In this study, however, we were only able of collecting PEMS emission patterns from limited vehicle classes. To accurately compute the instantaneous emissions of mixed traffic fleets and thoroughly combine the emission models with traffic simulation, it is still necessary to extend the PEMS data collection and modeling processes to other vehicle categories, especially large passenger vehicles and heavy duty trucks. Nevertheless, the calibration and validation methodology of this study will be able to be applied once more experimental data become available in the future.

Acknowledgments The authors would like to acknowledge the funding support of the WHUTLink research collaboration program (grant No. 348-2007-6398) through the Swedish Research Council (Vetenskapsrådet).

References

1. Ahn, K., & Rakha, H. (2009). A field evaluation case study of the environmental and energy impacts of traffic calming. *Transportation Research Part D: Transport and Environment, 14*(6), 411–424.
2. Barth, M., An, F., Younglove, T., Scora, G., Levine, C., Ross, M., et al. (2000). *Comprehensive Modal Emission Model (CMEM), version 2.0, User's Guide*. Riverside: University of California.
3. Cappiello, A., Chabini, I., Nam, E., Abou-Zeid, M., Lue, A., 2002. A statistical model of vehicle emissions and fuel consumption. Proceedings of the 5th IEEE International Conference on Intelligent Transportation Systems.
4. Cappiello, A., 2002. Modeling Traffic Flow Emissions. Master Degree Thesis, Massachusetts Institute of Technology.
5. CATI, 2003. OEM-2100 Montana System Operation Manual draft version 2.02. Clean Air technologies international INC.
6. Ekstrom, M., & Andreasson, K. (2004). Evaluation of the COPERT III emission model with on-road optical remote sensing measurements. *Atmospheric Environment, 38*(38), 6631–6641.
7. Elst, D.A.M.M., Smokers, R.T.M., De Koning, J.W., 2004. 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.
8. HORIBA Ltd, 2005. On-board emission measurement system OBS-2200 Instruction Manual.
9. Joumard, R., André, M., Vidon, R., Tassel, P., & Pruvost, C. (2000). Influence of driving cycles on unit emissions from passenger cars. *Atmospheric Environment, 42*(18), 4621–4628.
10. Kouridis, C., Ntziachristos, L., Samaras, Z., 2000. COPERT III Computer programme to calculate emissions from road transport. European Environment Agency
11. Keller, M., Kljun, N., 2007. ARTEMIS Road Emission Model User Guide. EU Commission
12. Koupal, J., Hart, C., Brzezinski, D., Giannelli. R., 2002. Draft Design and Implementation Plan for EPA's Multi-Scale Motor Vehicle and Equipment Emission System (MOVES). US EPA Office of Transportation and Air Quality, Report No. EPA420-P-02-008.
13. Lei, W., 2007. Study on the Driving Cycle and Emission Characteristics of Light-duty Vehicle in Wuhan. M.Sc. thesis, Wuhan University of Technology, China
14. Lei, W., Ma, X., Chen, H. (2010). Assessment of traffic environment using fine-tuned dynamic vehicle emission models. In: Proceeding of the 13th IEEE International Conference on Intelligent Transportation Systems, 2010.
15. Li, X., Li, G., Pang, S., Yang, X., & Tian, J. (2004). Signal timing of intersections using integrated optimization of traffic quality emissions and fuel consumption: a note. *Transportation Research Part D 9, 9*, 401–407.
16. Luenberger, D. G. (1984). *Linear and nonlinear programming* (2nd ed.). Massachusetts: Addison-Wesley Publishing Co.
17. MathWorks, 2009. Improving optimization performance with parallel computing. MATLAB Digest.
18. Palmgren, F., Berkowicz, R., Ziv, A., & Hertel, O. (1999). Actual car fleet emissions estimated from urban air quality measurements and street pollution models. *Science of the Total Environment, 235*, 101–109.
19. Panis, L. I., Broekx, S., & Liu, R. (2006). Modeling instantaneous traffic emission and the influence of traffic speed limits. *Science of the Total Environment, 371*(1–3), 270–285.
20. Rakha, H., Ahn, K., & Trani, A. (2004). Development of VT-Micro model for estimating hot stabilized light duty vehicle and truck emission. *Transportation Research Part D, 9*, 49–74.
21. Scora, G., Barth, M., 2006. Comprehensive Modal Emission Model (CMEM), version 3.01, User's Guide. University of California, Riverside, Center for Environmental Research and Technology.
22. Staehelin, J., Keller, C., Stahel, W., Schlpfer, K., & Wunderli, S. (1998). Emission factors from road traffic from a tunnel study (Gubrist Tunnel, Switzerland) Part III: results of organic com-

pounds, SO_2 and speciation of organic exhaust emissions. *Atmospheric Environment, 32*(6), 999–1009.
23. Teng, H., Yu, L., Qi, Y., 2002. Statistical micro-scale emission models incorporating acceleration and deceleration. Proceedings of 81th Transportation Research Board Annual Meeting, Washington DC, USA.
24. US EPA, 2001. User's guide to MOBILE6.0. Mobile source emission factor model. EPA-420-R-02-001. United States Environmental Protection Agency.
25. US EPA, 2009. Motor vehicle emission simulator (MOVES) 2009. EPA-420-B-09-008. United States Environmental Protection Agency.