Microscale Vehicle Emission Modelling in Hong Kong

C.K.L. Wong[1], T.S. Lo[1], H.L.A. Wong [1], K.L. Lam[1], H. C. Frey[2], R. Smit[3], S. Hausberger[4], K. Weller[4], L. Ntziachristos[5]

[1] Mobile Source Group, Hong Kong Environmental Protection Department, Hong Kong SAR Government, China, carolw@epd.gov.hk

[2] Department of Civil, Construction and Environmental Engineering, North Carolina State University, Raleigh, North Carolina, 27695-7908, United States
[3] Centre of Transport Strategy, University of Queensland, Brisbane, Australia
[4] Institute of Internal Combustion Engines and Thermodynamics, Graz University of Technology, Ineffeldgasse 19, 8010 Graz, Austria
[5] Department of Mechanical Engineering, Aristotle University of Thessaloniki, Thessaloníki, Greece

Introduction

A vehicle emission model is a mathematical tool to estimate vehicular emissions according to vehicle fleet composition, activities and ambient conditions. Their applications include compilation of vehicle emission inventory, evaluating the environmental impact for various projects, and estimating the effectiveness of various emission reduction control policies. Most of the popular vehicle emission models, e.g., COPERT developed for EUJRC (COPERT, https://www.emisia.com/utilities/copert/), MOVES (MOVES, https://www.epa.gov/moves) developed by USEPA and EMFAC (EMFAC, https://www.arb.ca.gov/emfac/) developed by CARB are macroscale or mesoscale. That is, the temporal resolution is at most hourly and the spatial resolution is in area level.

Real-world application of vehicle emission models, however, is often in microscale. For example, in metropolis such as Hong Kong, the road segments are often of the scale around 200 m. The average distance of the road segments between junctions in the urban area is about 50 m. To correctly model the vehicle emissions, one has to consider much more details of the driving situations. The characteristics of stop-and-go driving, frequency of hard accelerations and road gradient may have significant impact to the emissions. These impacts, however, may have been largely ignored or averaged in the macroscale/mesoscale model. Microscale vehicle emission model would be more useful in these applications.

In this study, we strive at developing a microscale vehicle emission model. The targeting application is a short road segment in Hong Kong. We shall consider 3 models, namely Passenger car and Heavy duty Emission Model (PHEM), P∆P and Modal and Peak Model (MAPM), in this study. These models differ considerably in the modelling methodology. 1) PHEM, (Hausberger, 2016), (Luz, 2014), (Hausberger, 2012), (Hausberger, 2003), employs engine map to predict the second-by-second emissions. The model has been used to generate the vehicles emission factors in the HBEFA (HBEFA, http://www.hbefa.net/e/index.html). 2) P∆P is a microscale vehicle emission model used in Australia for light duty and heavy duty vehicles vehicles (Smit, 2013; 2014; Smit et al., 2017). The model involves multivariate regression of the emission to numerous and optimised variables. The model is mainly statistically based on power and its derivative. 3) MAPM uses power binning as the core methodology similar to MOVES. The fundamental variables for the prediction are the binned vehicle speed and vehicle specific power (VSP) for light duty vehicles and scaled traction power (STP) for heavy duty vehicles). Mappings from these variables to the emissions are then created for the prediction. This binning method has been mainly used for meso/macro-scale emission modelling in US, but its applicability could be extended to microscale if more details are added.

The targeting application area is streets in Hong Kong. The spatial resolution is around 200m, and the temporal resolution is about 10 seconds. The aforementioned three models are developed/calibrated for the Hong Kong specific situation. Real-world vehicle emission measurement data using Portable Emissions Measurement System (PEMS) are used for the model development and validation. Data from a total of 18 vehicles including double deck buses, coaches and taxis are used for model development and another 3 vehicles for model validation.

In this paper, we shall present in details: the methodology used, development of the three models and their validation results.

Measurement Methodology

PEMS data from 18 vehicles are used for the model development as follows:

Table 1. Vehicles used for the model development.

Vehicle Class	Fuel	Euro Standard	GVW	Vehicle Brand	No. of Vehicles
Double deck bus	Diesel	Euro V	24 t	ALEXANDER DENNIS	5
Coach	Diesel	Euro V	16 – 17 t	MAN	1
				DAEWOO	3
				VOLVO	1
				SCANIA	3
Taxi	LPG	Euro 5	< 3 t	TOYOTA	5

The testing procedure of PEMS follows the most stringent testing requirements stipulated in ISO 16183, US CFR 1065 Subpart J and PEMS requirements in Euro VI regulation. Zero checking/ calibrations, audit checking and span checking/ calibration of the gaseous analysers were performed every hour, 3 hours and twice a day, respectively. The PEMS tests were conducted in real-world driving. The drivers are professional drivers of the same type of vehicles. There are two types of measurements; 1) the testing vehicle follows the vehicles of the same type to simulate their driving characteristics; 2) the testing vehicle is driven on a pre-determined route which includes both urban and highway driving. For heavy duty vehicles, the payload would be 50-60% of the maximum, whereas the payload would be the weight of PEMS and its accessories which are >50% of the maximum for the light duty vehicles. 2-3 fuel samples from the fuel tank were analysed for each vehicle to ensure no abnormality in the fuel specification.

The analysers being used for CO_2, CO, NO, NO_2 and THC measurements are SEMTECH-DS and AVL GAS PEMS. For SCR vehicles, N_2O, CH_4, NH_3 & various HC species were also measured by A&D portable FTIR. Exhaust flow meter was used to measure the exhaust flow rate and the exhaust flow temperature. A speedometer was attached to the wheel of the vehicle to record the vehicle speed. A weather probe was installed at the top of the vehicle to record the ambient temperature and relative humidity. An engine speed sensor is used to record the engine speed of the vehicle. GPS with dead reckoning is used to record the position of the vehicle. Its result is combined with the measurement from barometer, and the survey data from the Lands Department to estimate the road gradient of the testing vehicle. The data frequency of all these measurement is 1Hz. The example installation of PEMS is shown in Figure 1.

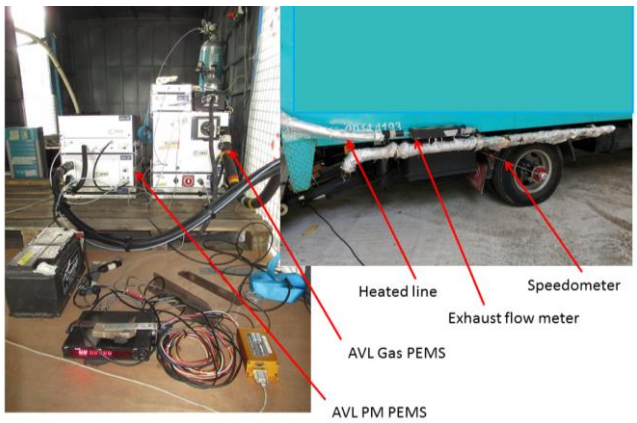


Figure 1: Example PEMS set-up for a heavy duty truck

PHEM

Model Development

The modelling methodology in PHEM is depicted in Figure 2. The heart of the model is the engine map (emissions vs. engine power and engine speed). To build the engine map using the PEMS data, the engine speed data from the engine speed sensors are used. The engine speed and the CO_2 emissions are used to determine the engine power through a generic CO_2 engine map.

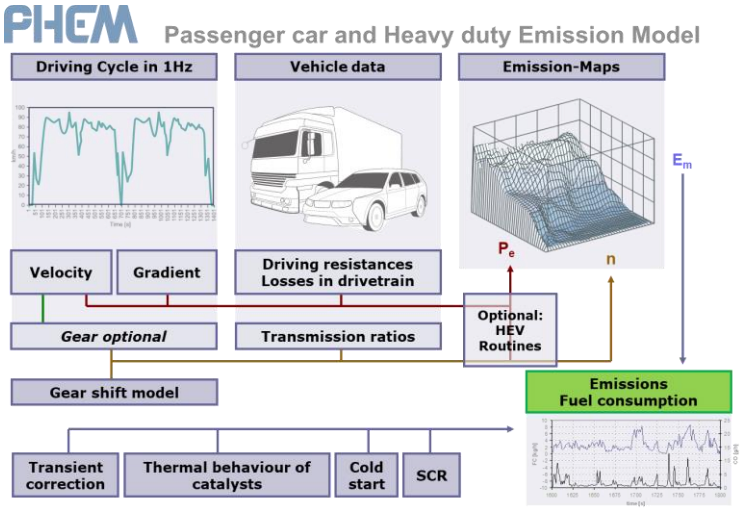


Figure 2: Schematic presentation of the methodology used in PHEM

The emission predictions from the engine maps can also be corrected by transient correction functions (Hausberger, 2016), (Luz, 2014), (Hausberger, 2012), (Hausberger, 2003). Due to time constraints, for the Hong Kong data set no transient correction was applied yet in PHEM.

For building the engine map, data from two road tests with at least 4,000 seconds for each vehicle are used for modelling. For this data set instantaneous engine speed and emission signals were accurately time aligned, which is a prerequisite for correct engine emission maps.

The rest of the data are used for the calibration. For calibration, the simulated and measured 100 second averaged emission results for every component are plotted over the 100 second averaged power values. Then linear equations are fitted for the simulated and the measured emissions. The difference between the two lines is added to the values in the emission map simply as function of the engine power value of the map point.

The SCR and the cold start corrections cannot be done in this study because no temperature upstream of SCR data is available from the measurements to set up the model.

PHEM Light, a simplified version of PHEM designed to be integrated into traffic models (e.g. SUMO), is also used. Characteristic emission lines (emission vs normalized engine power) are used instead of engine maps in PHEM Light. The model is independent of engine speed. No details on transmissions of the vehicle and gear shift model are needed. The transient corrections are not applied, but the model calibration is adjusted by the characteristic emission line generated by the rest of the PEMS data.

Model Application

Engine speed is estimated by a gear shift model of PHEM. Power is estimated by the second-by-second vehicle speed and road gradient (Hausberger, 2016), (Luz, 2014), (Hausberger, 2012), (Hausberger, 2003):

$P = P_R + P_L + P_a + P_s + P_{auxillaries}$,
where P_R is the power to overcome rolling resistance in [W], P_L is the power to overcome air resistance in [W], P_a is the acceleration power in [W], P_s is the power to overcome the road gradient in [W], $P_{Auxilliaries}$ is the power consumption of auxiliaries in [W].

PΔP

The model is designed to simulate the impacts of changing traffic and operational conditions, as well as a wide range or traffic measures, on air pollution and greenhouse gas emission levels *at fleet level* (Smit, 2013). These include, for instance different degrees of congestion, tunnel emissions, signals versus roundabouts, signal settings, eco-driving and dynamic speed limits.

Central to the PΔP is the estimation of vehicle power (P) and the change of power (ΔP). The following on-road instantaneous power is estimated with algorithms (Smit R., 2014) from the European ARTEMIS project (Rexeis et al., 2005) were adopted for the PΔP model:

- power required to overcome aerodynamic resistance
- power required to overcome tyre rolling resistance
- power required to overcome drive train/transmission resistance
- power required to overcome inertial resistance
- power required to overcome gravitational resistance
- power required to run auxiliaries

The power and the change of power are the explanatory variables. For power, both positive and negative engine powers and their transformations (log or square-root) are used. For the change of power, i) power difference: $\Delta P(\tau)_t = P_t - P_{t-\tau}$; ii) power oscillation: $\Delta P(\tau)_t = |P_t - P_{t-1}| + \ldots + |P_{t-\tau+1} - P_{t-\tau}|$ and iii) normalised power oscillation: $\Delta P(\tau)_t^* = \Delta P(\tau)_t / \Delta x(\tau)_t$ are used.

The preceding time period τ covered by $\Delta P(\tau)_t^*$ is variable, and takes the values of τ = 3, 6, 9, 12, 15, 30 and 60 seconds. Note that the last variable $\Delta P(\tau)_t^*$ (kW/m) represents the $\Delta P(\tau)_t^*$ variable normalised for distance over preceding time period τ.

In addition, the temperature of the exhaust is an important factor affecting the emissions especially for the taxis equipped with three way catalyst and vehicles equipped with SCR. The after-treatment device functions properly only when the temperature is above the light-off temperature, and emission control efficiency changes sharply around this temperature, which could be roughly described as a reverse logistic function $e_t = \beta_0 + \frac{\beta_1}{(1+\beta_2\, e^{T(\tau)t - T^*})} + \varepsilon$, where e_t is pollutant emission rate at t = t (g/s), $T(\tau)t$ is average exhaust temperature over the last τ seconds at time = t (°C) and T^* is light-off exhaust temperature (°C).

To linearize the effect of the exhaust temperature, the modelling team considers the transformation:

$$F_t = \left(b_0 + \frac{b_1}{(1 + \exp^{T(\tau)t - T^*})}\right) / (b_0 + b_1) \qquad \text{if } b_1 > 0$$

$$F_t = \left(b_0 + \frac{b_1}{(1 + \exp^{T(\tau)t - T^*})}\right) / \, b_0 \qquad \text{if } b_1 \leq 0$$

A method is required to estimate temperature as a function of variables for which input data are available. The model, $T_t = \beta_0 + \beta_1 \ln(1 + W(\Delta t)_t^+) + \beta_2 \ln(1 + W(\tau)_t^+) + \varepsilon$, is used to predict instantaneous exhaust temperature T_t as function of work. $W(\Delta t)_t^+$ is the accumulated positive work at time = t since the engine start and $W(\tau)_t^+$ is accumulated positive work over a period of τ seconds before time = t. Figure 3 shows two examples of the exhaust temperature predictions versus observed values.

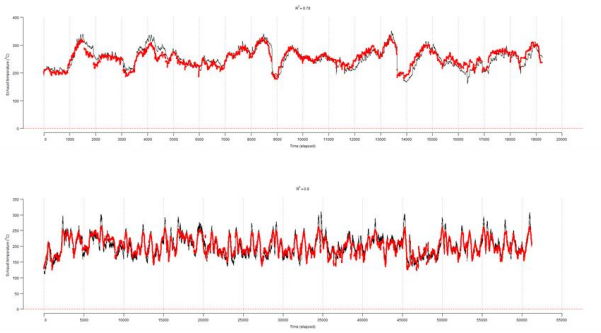


Figure 3: Time-series plots of temperature predictions for the coaches. Red dots: the exhaust temperature prediction; Black line: the measurement values after the tailpipe.

PΔP backcasts driving behaviour which computes accumulated work, temperature and other model variables, such as delta power since t = 0 (engine start).

Various PΔP model structures are possible with different levels of complexity and different expressions/definitions of power, delta power and temperature. For instance, the following structures could be used:

$$e_t = \beta_0 + \beta_1 P_t^+ + \varepsilon \ ,$$

$$e_t = \beta_0 + \beta_1 P_t^+ + \beta_2 \Delta P(\tau)_t^* + \beta_3 P_t^2 + \beta_4 T(\tau)_t' + \varepsilon, \text{ or}$$

$$e_t = \beta_0 + \beta_1 \ln(1 + P_t^+) + \beta_2 P_t^2 + \beta_3 T(\tau)_t' + \beta_4 P_t \Delta P(\tau)_t + \beta_5 P_t T_t + \varepsilon.$$

The best model structure is identified as possessing using i) Mallow's C_p criterion (minimized C_p value, with $C_p \leq p$ and low p value); ii) 'Variance inflation factors' (*VIFs*) are calculated to check for signifcant multicollinearity; and iii) 'Constancy of error variance' (*CEV*) was created to check for homoscedasticity.

Modal and Peak Models

MAPM is a model similar to MOVES (MOVES, https://www.epa.gov/moves) developed by U.S. EPA. The modelling team modified i) the definition of operating modes in MOVES to better represent Hong Kong driving situations, and ii) added two new modes, peak and fuel cut-off modes, to improve the prediction of emission peak and fuel cut-off driving. In this section, the modelling methodology is described.

It was demonstrated that vehicle energy use and emissions (EU&E) are proportional to vehicle specific power (VSP) (Frey et al., 2002) (Jimenez-Palacios, 1999). Emission rates of different pollutants have different relationships with VSP (Frey et al., 2010). In MOVES, VSP is defined as power over vehicle mass for light duty vehicles and Scaled Tractive Power (STP) for heavy duty vehicles power over scaling factor (f_{scale}).

MOVES cannot be directly applied to Hong Kong because 1) maximum speed limit for heavy duty vehicles in Hong Kong is 70 kph and so most of the MOVES Operating Modes in high speed are not applicable in Hong Kong; and 2) the model may "average out" some of the high frequency variability in specially low and high emissions that cannot be directly explained by speed, acceleration, grade, VSP, STP, or other parameters derived from these.

To tackle the issue 1) above, the modelling team tried to re-define the Operating modes such that each mode has sufficient data samples and CO_2 emissions for different modes are significantly different. To achieve this target, the modelling team has tried different definitions of operating modes to the best model.

To tackle the issue 2) above, the modelling team devised two additional modes: the fuel cut-off mode and the peak mode for predicting the specially low and high emissions, respectively. To build those modes, CO_2 emission analysis by CART tree was used. In the CART analysis, the

lagged VSP, i.e., VSP from 1 second ago, 2 seconds ago, etc., are also considered for modelling the possible “history” or “lag” effect of the emission.

MAPM for Taxi

Here we give an example to develop the model for taxis, the methodology of the developing the model for coaches and double deck buses are similar.

After a few iterations, the definition of the operating modes is proposed for taxi (Table 2). This definition has narrower speed range than MOVES. Braking is defined as being a deceleration of 3 km/h/s, or 1 km/h continuously for three seconds. Idle is defined as speed of 0 kph and acceleration of 0 kph/s.

Table 2: MAPM for an Average Hong Kong LPG Taxi Based on 20 kph Speed Ranges with Fuel Cut-Off and Peak Modes: Mode Definition, Time in Mode, Emission Rates and Total Emissions by Mode for CO_2, CO, HC, NO, and NO_2 (N= 401,251). (speed in kph)

HK Modes			Time		CO_2			CO			HC			NO			NO_2			Exhaust
Speed	VSP	HK Mode	Secs	Percent	Ave (g/s)	Total (g)	Percent of Total	Ave (g/s)	Total (g)	Percent of Total	Ave (g/s)	Total (g)	Percent of Total	Ave (g/s)	Total (g)	Percent of Total	Ave (g/s)	Total (g)	Percent of Total	Temperature (°C)
Fuel cut-off		-3	27939	7.0%	0.90	25092	2.6%	0.001	36	2.4%	4.80E-05	1.34	1.8%	1.93E-04	5.38	2.8%	1.10E-05	0.31	4.2%	182
-1		-2	27484	6.8%	1.19	32819	3.4%	0.002	42	2.8%	6.32E-05	1.74	2.4%	2.32E-04	6.38	3.3%	1.24E-05	0.34	4.7%	160
0		-1	84651	21.1%	1.09	92537	9.6%	0.001	64	4.3%	7.32E-05	6.19	8.4%	3.95E-05	3.34	1.7%	5.36E-06	0.45	6.3%	144
0-20	<0	0	15597	3.9%	1.22	19066	2.0%	0.001	20	1.3%	8.84E-05	1.38	1.9%	2.01E-04	3.14	1.6%	1.00E-05	0.16	2.2%	150
0-20	0-3	1	23004	5.7%	1.56	35811	3.7%	0.002	48	3.2%	1.96E-04	4.51	6.1%	3.22E-04	7.41	3.8%	1.16E-05	0.27	3.7%	151
0-20	3-6	2	8169	2.0%	2.75	22437	2.3%	0.004	35	2.3%	4.32E-04	3.53	4.8%	5.42E-04	4.43	2.3%	1.62E-05	0.13	1.8%	157
0-20	>6	3	3938	1.0%	3.77	14845	1.5%	0.006	22	1.5%	5.90E-04	2.32	3.1%	6.28E-04	2.47	1.3%	2.02E-05	0.08	1.1%	162
20-40	<0	20	18307	4.6%	1.64	29987	3.1%	0.002	36	2.4%	1.22E-04	2.24	3.0%	4.06E-04	7.44	3.8%	2.01E-05	0.37	5.1%	161
20-40	0-3	21	15406	3.8%	2.21	34100	3.5%	0.003	44	2.9%	1.96E-04	3.02	4.1%	7.69E-04	11.84	6.1%	3.46E-05	0.53	7.4%	164
20-40	3-6	22	13788	3.4%	3.03	41762	4.3%	0.004	53	3.6%	2.84E-04	3.91	5.3%	8.69E-04	11.98	6.2%	3.19E-05	0.44	6.1%	168
20-40	6-9	23	10997	2.7%	3.98	43792	4.5%	0.006	62	4.1%	4.55E-04	5.00	6.8%	8.02E-04	8.82	4.5%	2.76E-05	0.30	4.2%	174
20-40	9-12	24	6924	1.7%	4.94	34190	3.5%	0.007	50	3.3%	5.66E-04	3.92	5.3%	1.01E-03	6.99	3.6%	3.10E-05	0.21	3.0%	180
20-40	12-15	25	2984	0.7%	5.86	17501	1.8%	0.007	21	1.4%	6.08E-04	1.82	2.5%	1.19E-03	3.54	1.8%	3.54E-05	0.11	1.5%	184
20-40	>15	26	972	0.2%	7.08	6882	0.7%	0.010	10	0.7%	7.97E-04	0.78	1.1%	2.02E-03	1.96	1.0%	4.52E-05	0.04	0.6%	193
40-60	<0	40	13488	3.4%	1.93	25973	2.7%	0.003	35	2.4%	1.04E-04	1.40	1.9%	4.42E-04	5.97	3.1%	2.02E-05	0.27	3.8%	182
40-60	0-3	41	12210	3.0%	2.32	28269	2.9%	0.003	38	2.5%	1.24E-04	1.51	2.0%	5.91E-04	7.22	3.7%	2.44E-05	0.30	4.1%	186
40-60	3-6	42	13190	3.3%	3.02	39808	4.1%	0.004	54	3.6%	1.70E-04	2.24	3.0%	7.40E-04	9.77	5.0%	2.89E-05	0.38	5.3%	188
40-60	6-9	43	10903	2.7%	3.91	42637	4.4%	0.006	64	4.3%	2.48E-04	2.70	3.7%	7.21E-04	7.86	4.0%	2.70E-05	0.29	4.1%	192
40-60	9-12	44	8272	2.1%	4.92	40670	4.2%	0.008	64	4.2%	3.46E-04	2.86	3.9%	8.32E-04	6.88	3.5%	2.58E-05	0.21	2.9%	204
40-60	12-15	45	4597	1.1%	5.83	26812	2.8%	0.008	37	2.5%	3.75E-04	1.73	2.3%	1.02E-03	4.69	2.4%	3.10E-05	0.14	2.0%	207
40-60	>15	46	2520	0.6%	7.15	18009	1.9%	0.012	31	2.1%	5.31E-04	1.34	1.8%	1.73E-03	4.36	2.2%	3.91E-05	0.10	1.4%	207
>60	<0	60	10234	2.6%	2.35	24083	2.5%	0.005	52	3.5%	1.36E-04	1.39	1.9%	5.06E-04	5.17	2.7%	1.73E-05	0.18	2.4%	210
>60	0-3	61	10236	2.6%	2.55	26090	2.7%	0.005	46	3.1%	1.18E-04	1.20	1.6%	7.47E-04	7.65	3.9%	2.33E-05	0.24	3.3%	206
>60	3-6	62	14351	3.6%	3.05	43780	4.5%	0.005	71	4.8%	1.39E-04	1.99	2.7%	8.21E-04	11.78	6.0%	2.41E-05	0.35	4.8%	207
>60	6-9	63	14889	3.7%	3.74	55631	5.8%	0.006	96	6.4%	2.04E-04	3.03	4.1%	7.61E-04	11.32	5.8%	2.42E-05	0.36	5.0%	210
>60	9-12	64	12003	3.0%	4.58	54934	5.7%	0.010	116	7.7%	2.99E-04	3.59	4.9%	7.63E-04	9.15	4.7%	2.32E-05	0.28	3.8%	218
>60	12-15	65	7729	1.9%	5.43	41954	4.3%	0.012	92	6.1%	3.71E-04	2.87	3.9%	9.59E-04	7.41	3.8%	2.47E-05	0.19	2.6%	227
>60	15-18	66	3468	0.9%	6.18	21434	2.2%	0.016	55	3.7%	4.45E-04	1.54	2.1%	1.16E-03	4.01	2.1%	2.84E-05	0.10	1.4%	233
>60	18-21	67	1293	0.3%	6.82	8823	0.9%	0.020	26	1.7%	5.26E-04	0.68	0.9%	1.37E-03	1.77	0.9%	3.06E-05	0.04	0.5%	243
>60	>21	68	734	0.2%	7.97	5847	0.6%	0.043	32	2.1%	9.43E-04	0.69	0.9%	2.09E-03	1.54	0.8%	2.98E-05	0.02	0.3%	247
Peak		99	974	0.2%	10.54	10267	1.1%	0.046	45	3.0%	1.39E-03	1.36	1.8%	3.17E-03	3.09	1.6%	5.26E-05	0.05	0.7%	233

Fuel Cut-off Mode Development

CART analysis is used to identify conditions in which the fuel use rate of the measured data is lower than the lowest modal average rate. The CART tree is shown in Figure 4. In the figure, VSP_Lx represents the value of VSP in the previous x seconds. The fuel cut-off mode is defined as the lowest CO_2 emission leaf node (in red circle).

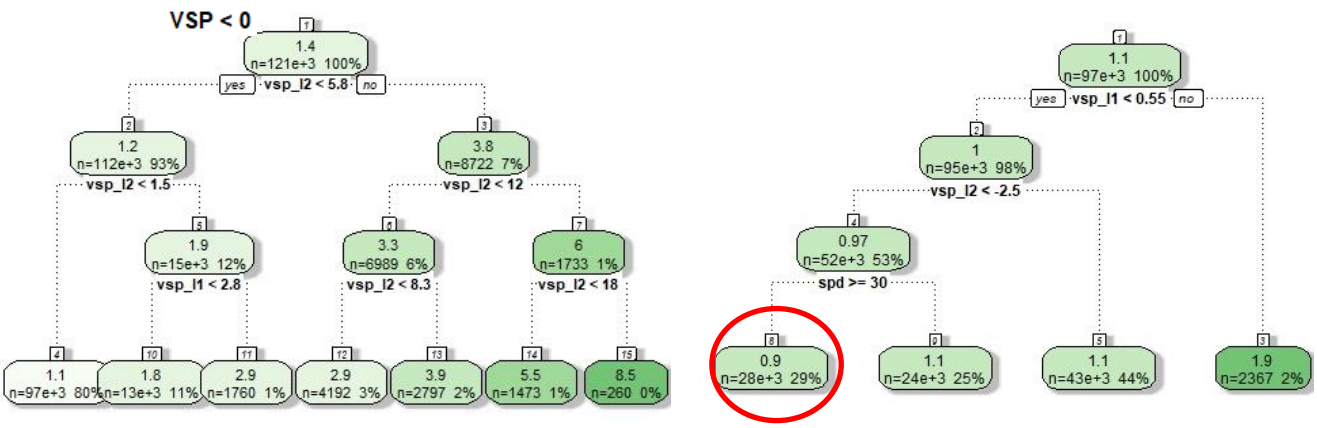


Figure 4: CART Regression Tree analysis for LPG Taxi Data for the definition of the fuel cut-off mode. Fuel cut-off mode is defined as the tree node in the red circle.

When CART-derived rule is applied to detect fuel cut-off in the measured data, it correctly predicts CO_2 emission rates that are less than the idle modal average rate for 22,102 data points. It also correctly identifies that 285,055 data points do not have fuel cut-off. It predicts, incorrectly, that 5,837 data points are fuel cut-off when they really are not (false positives), and it predicts that 88,257 data points are not fuel cut-off when they really are (false negatives).

Peak Mode Development

The peak mode is defined by another CART tree analysis (Figure 5). The peak is defined as VSP_L1>19, VSP_L2>19 and speed <85 km/h. Based on the simplified rule, the mean CO_2 emission rate is 10.23 g/sec from a total sample size of 1183, which is substantially higher than the highest modal average rate of 8.62 g/sec. This rule leads to 827 true positives and only 356 false positives, while it has 397,348 true negatives and 2,719 false negatives.

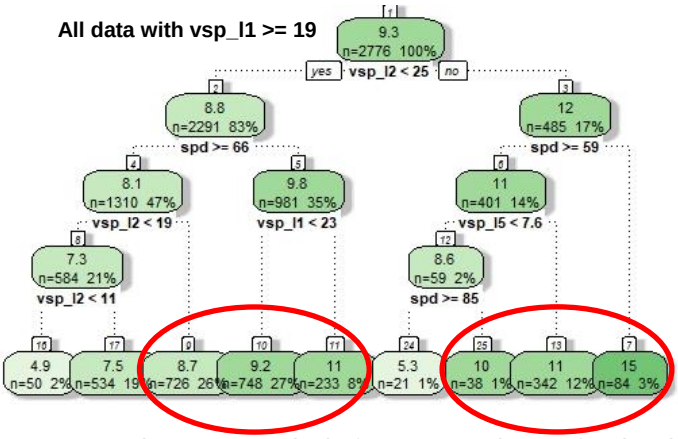


Figure 5: CART Regression Tree analysis for LPG Taxi Data for the definition of the peak mode. The rule for the peak mode is simplified as VSP_L1 ≥ 19, VSP_L2 ≥ 19 and speed < 85 (The CO_2 emissions are higher than 8.7 g/s)

Final MAPM for Taxi

The modal average emission rates for the final MAPM for LPG Taxis are shown in Figure 6. The modal average CO_2 emission rate clearly increases monotonically with positive VSP in each speed range. The CO emission rates are highest at high VSP in the two highest speed ranges. The HC emission rates are typically increasing monotonically with VSP in each speed range. The highest NO emission rate are observed at mode 26, 46 and 68, which is the highest range of VSP at the each speed range. The trend in NO_2 emission rates is somewhat counter-intuitive, in that it seems to either decrease with increasing positive VSP or to not have much of a trend with respect to VSP.

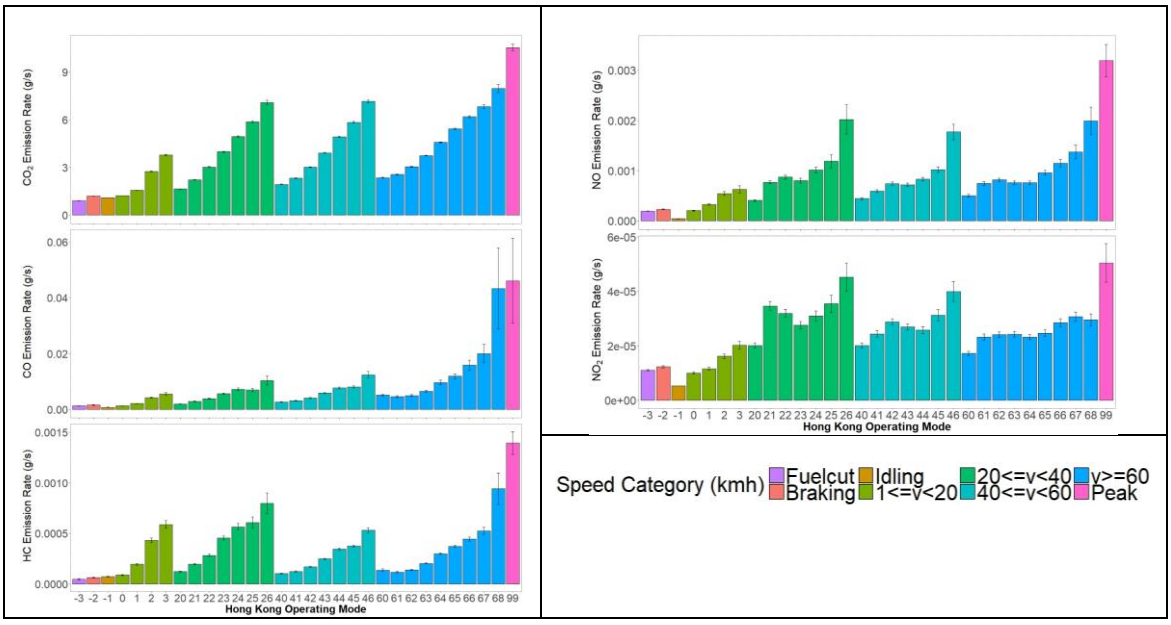


Figure 6: Modal average emission rates, with 95 percent confidence intervals on the mean, for the Modal and Peak Model for an Average Hong Kong LPG Taxi (N= 401,251)

Model Validation

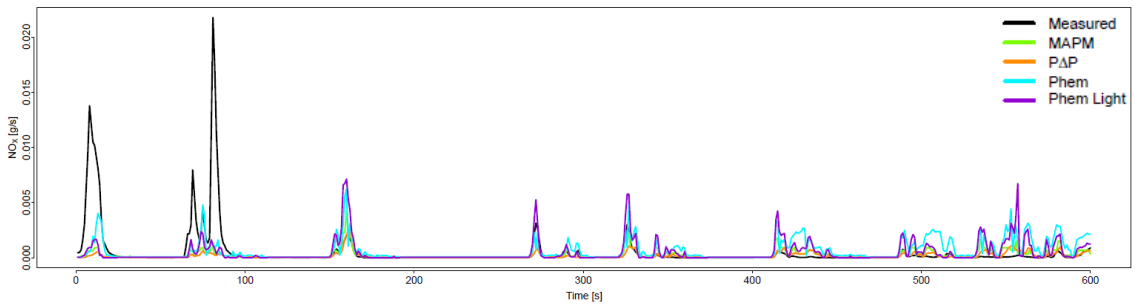
The models' predictions are validated by two PEMS datasets after the model development (Table 3), Dataset A and Dataset B.

For dataset A: while the vehicles of the validation data are included in the model development, the dataset itself is not. For dataset B: the vehicles are not included in the modelling. The NOx emissions of these 3 vehicles are close to the average emissions of the vehicles in that particular class.

Table 3: Datasets for model validation

Dataset A	**Dataset B**
PEMS data for 3 road networks for all the vehicles in the Modelling Dataset, each lasting for 10 minutes	PEMS data for 3 road networks for all the following vehicles which are not in the Modelling Dataset, each lasting for 10 minutes: ● 1 Euro 5 LPG taxi ● 1 Euro V diesel coach of 15-24 tonnes ● 1 Euro V diesel double deck bus of about 24 tonnes

Figures 7-9 show the exampled time series of the predictions and the measurement for NOx. It is observed that the predictions for double deck buses and coaches are better than those of taxis. There seems no significant difference between the prediction for dataset A and B. All models could predict the locations of the emission peaks quite accurately. However, none of the model could estimate the amplitude of the peak correctly, especially for taxis.



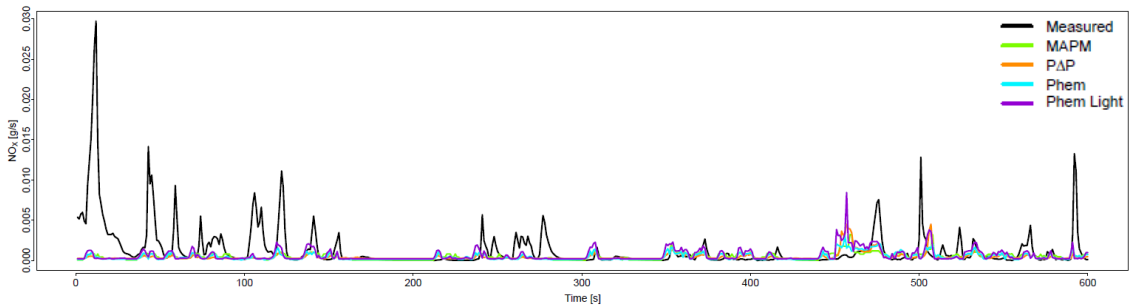


Figure 7: Time plot comparison of measured NOx emission rate (g/sec) for taxi. Above: validation of dataset A, and bottom: validation of dataset B.

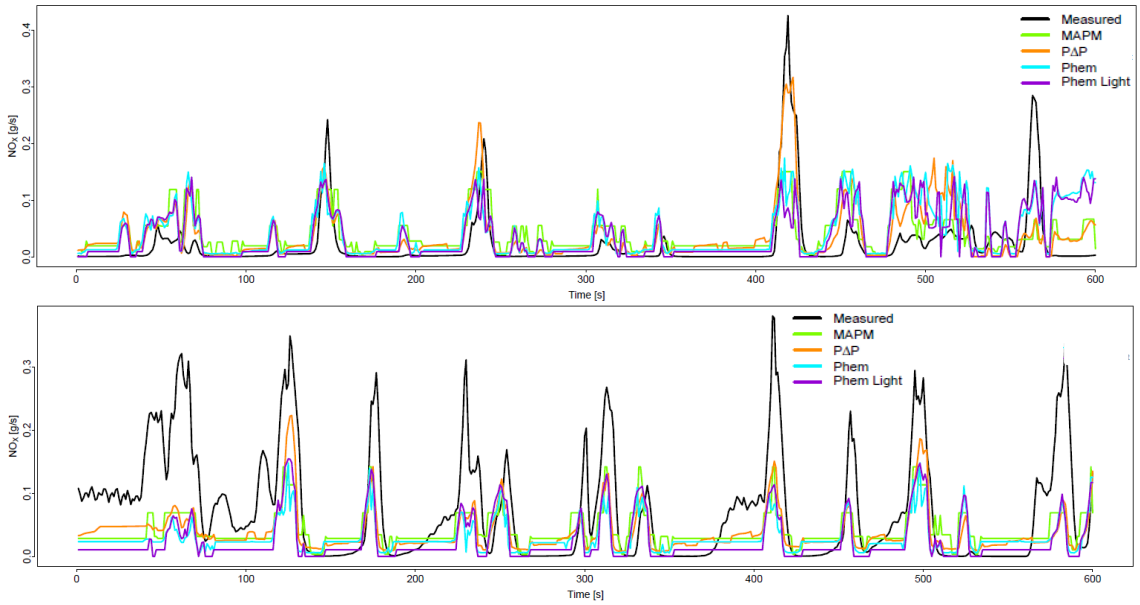


Figure 8: Time plot comparison of measured NOx emission rate (g/sec) for bus. Above: validation of dataset A, and bottom: validation of dataset B.

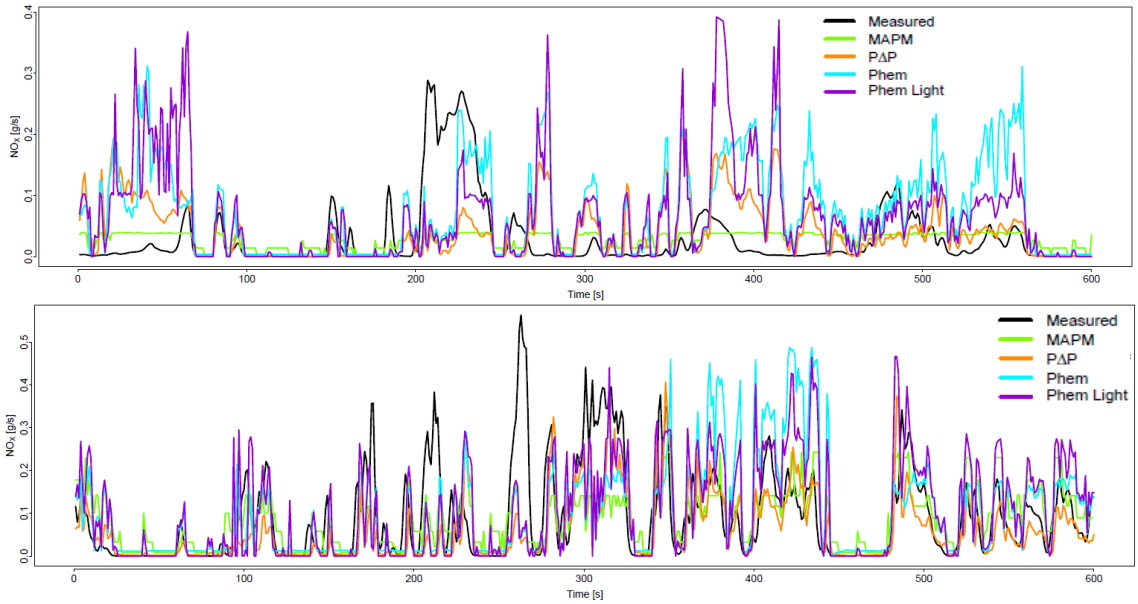


Figure 9: Time plot comparison of measured NOx emission rate (g/sec) for coach. Above: validation of dataset A, and bottom: validation of dataset B.

To further quantify the errors, three statistical metrics are used. Firstly, Mean absolute error (MAE)

(Hyndman and Koehler, 2006; Makridakis et al. 1998) is used to quantity the errors in second-by-second level.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|e_i|$$

where $e_i = y_i - \hat{y}_i$.

To eliminate scale dependency, we employ MAEm which was defined as the ratio [%] of the MAE to the respective average observed emission rate (g/s) in each case

$$MAEm = 100 * \frac{MAE}{\bar{y}}$$

The ideal value for both MAE and MAEm is 0. To quantify the error of the whole trip, Total emissions error (TEE)

$$TEE = 100 * \frac{\sum_{i=1}^{n}\hat{y}_i - \sum_{i=1}^{n}y_i}{\sum_{i=1}^{n}y_i}$$

is computed as a measure of the accuracy in predicting the total sum of emissions based on one-step-forecasts. The ideal value for TEE is 0. Pearson's r (R) (Bennett et al., 2013) is a measure of the linear correlation between the predicted and the observed values.

$$R = \frac{\sum_{i=1}^{n}(y_i - \hat{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{\hat{y}})^2}}$$

Pearson's R is insensitive to systematic overestimation/underestimation of the observed values. The ideal value for R in our case is 1.

All the statistical metrics above were calculated in 10-second time steps, i.e., the 10-second average emission was predicted and compared. And the values were then averaged over all vehicles in the vehicle class. The results are shown in table 4-5. The accuracies for CO_2 for all models are good, no matter in 10-second level (MAEm), and 30-minute level (TEE). Also, the prediction correlated very well with the measurement as indicated in the value of R. There is no significant difference between the results in dataset A and B, meaning that the models are no over-specified for the modelling dataset.

Table 4: Validation results of CO_2 for different vehicle classes and models

	Metric	Model	Dataset A			Dataset B		
			Bus	Coach	Taxi	Bus	Coach	Taxi
CO_2	MAEm	MAPM	20%	24%	15%	21%	23%	21%
		PΔP	19%	20%	12%	21%	19%	16%
		PHEM	20%	22%	15%	23%	21%	20%
		PHEM Light	29%	21%	16%	42%	21%	22%
	TEE	MAPM	5%	7%	2%	4%	6%	1%
		PΔP	6%	5%	2%	9%	7%	1%
		PHEM	6%	9%	6%	7%	2%	1%
		PHEM Light	13%	7%	3%	26%	3%	14%
	R	MAPM	94%	93%	94%	92%	92%	88%
		PΔP	95%	95%	96%	93%	92%	92%
		PHEM	93%	94%	94%	91%	91%	89%
		PHEM Light	95%	94%	93%	92%	92%	90%

Table 5: Validation results of NOx for different vehicle classes and models

	Metric	Model	Dataset A			Dataset B		
			Bus	Coach	Taxi	Bus	Coach	Taxi
NOx	MAEm	MAPM	83%	61%	132%	67%	47%	90%
		PΔP	82%	59%	123%	56%	51%	92%
		PHEM	106%	67%	223%	71%	54%	93%
		PHEM Light	94%	68%	188%	72%	44%	101%
	TEE	MAPM	19%	13%	53%	27%	15%	38%
		PΔP	14%	15%	26%	32%	43%	43%
		PHEM	38%	25%	153%	47%	25%	38%
		PHEM Light	35%	22%	112%	56%	15%	22%
	R	MAPM	58%	67%	41%	64%	74%	9%
		PΔP	49%	70%	24%	72%	76%	1%
		PHEM	42%	62%	30%	60%	72%	3%
		PHEM Light	40%	60%	35%	61%	80%	0%

Compared to CO_2 results, the NOx results are much worse. Regarding the 10-second level evaluation, the best predictions are for the Coach in which the MAEm is about 60% for all models. For buses and taxis, the values of MAEm rise to 82%-106% and 123%-223% for dataset A and 56%-72% and 90%-101% for dataset B. Roughly speaking, the MAPM and PΔP models perform better than PHEM and PHEM Light. Interestingly, the accuracy for dataset B is better than dataset A.

For the 30-minute evaluation, still the MAPM and PΔP models perform better than PHEM and PHEM Light. Especially for Taxi, the values of TEE of PHEM and PHEM Light could be double of that of MAPM and PΔP. The performance of PΔP drops when the evaluation data switches from dataset A to dataset B. It may mean that the model has some degree of over-specifying to the modelling dataset.

For the results of Pearson's correlation R, all models are basically the same, with MAPM and PΔP slightly better. However, the performances of all models are considerably better for buses and coaches, especially in dataset B.

The performance of the models for other pollutants CO, THC, NO, NO_2 and PM are more or less similar to NOx. If we weigh the importance of different vehicle classes by their share in total emissions of 2016 vehicle emission inventory in Hong Kong, and the pollutants by table 6 (Roughly representing their importance in the environment impacts in Hong Kong). The weighted averaged of all the statistical metrics are given in Table 7.

Table 6: Importance weighting of different pollutants by class

Weights	**FBDD**	**NFB**	**Taxi**
NO	27.3%	9.5%	3.2%
NO_2	17.1%	5.9%	2.0%
NOx	0.0%	0.0%	0.0%
CO	1.9%	0.3%	2.7%
CO_2	2.4%	0.5%	2.1%
THC	3.1%	0.5%	1.4%
PM_{10}	16.6%	3.4%	0.0%

Table 7: Averaged statistical metrics of all models

	Dataset A			**Dataset B**		
Model	**MAEm**	**TEE**	**R**	**MAEm**	**TEE**	**R**
MAPM	75%	22%	62%	69%	27%	60%
PΔP	85%	29%	50%	70%	36%	61%

PHEM	93%	37%	49%	82%	49%	54%
PHEM Light	82%	35%	50%	77%	43%	54%

The overall statistics indicate that the MAPM and PΔP models perform slightly better than PHEM and PHEM light in this validation exercise.

Discussion and Conclusion

This paper discusses the development/calibration of microscale models from the PEMS data in Hong Kong. Three models PHEM, PΔP and MAPM are involved in this study for the vehicle classes: double deck bus, coach and taxi. The models developed/calibrated were validated by independent datasets. The results indicate that the accuracy of all models are approximately the same, with MAPM and PΔP slightly better. The evaluation of 30-minute level resolution is better than that of the 10-second level. Also, the predictions for bus and coach are better than that for taxi.

One should note that besides accuracy, many aspects of the model are important. For example:

- The computational speed,
- Easiness to use,
- Whether the required input data are available, and
- Whether model parameters are easy to update or calibrate

are important factors affecting whether the model is suitable for the users. However, these factors are not considered in this study.

Acknowledgements

We thank Alberto Ayala, Robert Sawyer and Alex Wai for their review and discussion.

References

Bennett N.D., Croke B.F.W., Guariso G., Guillaume J.H.A., Hamilton S.H., Jakeman A.J., Andreassian V. (2013). Characterising performance of environmental models, *Environmental Modelling & Software*, *40*, 1–20.

COPERT, *https://www.emisia.com/utilities/copert/*

EMFAC, *https://www.arb.ca.gov/emfac/*

Frey H.C., Unal A., Chen J., Li S., Xuan C.(2002), Methodology for Developing Modal Emission Rates for EPA's Multi-Scale Motor Vehicle and Equipment Emission Estimation System, *EPA420-R-02-027, Prepared by North Carolina State University for the Office of Transportation and Air Quality, U.S. Environmental Protection Agency, Ann Arbor, MI, October 2002*

Frey, H.C., Zhang K., Rouphail N.M.(2010), Vehicle-Specific Emissions Modeling Based Upon On-Road Measurements, *Environmental Science and Technology,* 44(9):3594-3600.

Hausberger S, (2003). Simulation of Real World Vehicle Exhaust Emissions; *VKM-THD Mitteilungen; Heft/Volume 82; Verlag der Technischen Universitat;* ISBN 3-901351-74-4; Graz 2003.

Hausberger S., Rexeis M., Kies A., Schulte L., Verbeek R. et. al. (2012). Reduction and Testing of Greenhouse Gas. Emissions from Heavy Duty Vehicles. *LOT 2; Development and testing of a certification procedure for CO_2 emissions and fuel consumption of HDV;*DG CLIMA Contract N 070307/2009/548300/SER/C3; Graz 2012

Hausberger S., Rexeis M., Kies A., Weller K., Siberholz G. (2016). Future HD vehicle requirements due to legislation and effects on CO_2 and air quality, *3 INTERNATIONALER MOTORENKONGRESS 2016 ANTRIEBSTECHNIK 1M FAHRZEUG,* Baden Baden; 2016

HBEFA, *http://www.hbefa.net/e/index.html*

Hyndman R.J., Koehler A.B.(2006). Another look at measures of forecast accuracy. *International Journal of Forecasting,* 22(4), 679–688.

Jimenez-Palacios (1999), J. L. Ph.D. Dissertation*, Massachusetts Institute of Technology:* Cambridge, MA, 1999.

Luz R., Rexies M., Hausberger S., Schulte L., Hammer J., Steven H., Verbeek R. et. al. (2014). Development and validation of a methodology for monitoring and certification of greenhouse gas emissions from heavy duty vehicles through vehicle simulation, *Final report, Service contract CLIMA.C.2/SER/2012/0004;* Report No. | 07/14/Rex EM-1 2012/08 699 from 15.05.2014

Makridakis S.G., Wheelwright S.C., & Hyndman R.J. (1998). *Forecasting: methods and applications* (3rd ed). New York: John Wiley & Sons.

MOVES, *https://www.epa.gov/moves*

Rexeis M., Hausberger S., Riemersma I., Tartakovsky L., Zvirin Y., Van Poppel M., Cornelis E., (2005). ARTEMIS - Assessment and Reliability of Transport Emission Models and Inventory Systems, *WP400 - Heavy Duty Vehicle Emissions,* DGTREN Contract 1999-RD.10429.

Smit R.,(2013). Development and performance of a new vehicle emissions and fuel consumption software (PΔP) with a high resolution in time and space, *Atmospheric Pollution Research,* 4, 336-345.

Smit R., 2014. PΔP: A simulationtable tool for vehicle emissions and fuel consumption software with a high resolution in time and space, *Vehicle Technology Engineer,* SAE Australasia, July 2014, 17-21.

Smit, R., Kingston, P., Wainwright, D., Tooker, R. (2017). A tunnel study to validate motor vehicle emission prediction software in Australia, *Atmospheric Environment*, 151, 188-199.