Predicting vehicular emissions in high spatial resolution using pervasively measured transportation data and microscopic emissions model

Predicting vehicular emissions in high spatial resolution using pervasively measured transportation data and microscopic emissions model

Marguerite Nyhan[1,*], Stanislav Sobolevsky[2], Chaogui Kang[3], Prudence Robinson[1], Andrea Corti[4], Michael Szell[5], David Streets[6], Zifeng Lu[6], Rex Britter[1], Steven R.H. Barrett[7], Carlo Ratti[1]

[1]Massachusetts Institute of Technology, SENSEable City Laboratory, Cambridge, MA, United States.

[2]Centre for Urban Science and Progress, New York University, New York City, United States.

[3]Wuhan University, Wuhan, Hubei, China.

[4]Politecnico di Milano, 32 Piazza Leonardo da Vinci, Milano, Italy.

[5]Center for Complex Network Research, Department of Physics, Northeastern University, Boston, United States.

[6]Argonne National Laboratory, National Aeronautics and Space Administration (NASA), Lemont, Illinois, United States.

[7]Massachusetts Institute of Technology, Department of Aeronautics and Astronautics, Cambridge, MA, United States.

[*]Corresponding author.

Abstract

Air pollution related to traffic emissions pose an especially significant problem in cities; this is due to it's adverse impact on human health and well-being. Previous studies which have aimed to quantify emissions from the transportation sector have been limited by either simulated or coarsely resolved traffic volume data. Emissions inventories form the basis of urban pollution models, therefore in this study, Global Positioning System (GPS) trajectory data from a taxi fleet of over 15,000 vehicles were analyzed with the aim of predicting air pollution emissions for Singapore. This novel approach enabled the quantification of instantaneous drive cycle parameters in high spatio-temporal resolution, which provided the basis for a microscopic emissions model. Carbon dioxide (CO_2), nitrogen oxides (NO_x), volatile organic compounds (VOCs) and particulate matter (PM) emissions were thus estimated. Highly localized areas of elevated emissions levels were identified, with a spatio-temporal precision not possible with previously used methods for estimating emissions. Relatively higher emissions areas were mainly concentrated in a few districts that were the

Singapore Downtown Core area, to the north of the central urban region and to the east of it. Daily emissions quantified for the total motor vehicle population of Singapore were found to be comparable to another emissions dataset. Results demonstrated that high-resolution spatio-temporal vehicle traces detected using GPS in large taxi fleets could be used to infer highly localized areas of elevated acceleration and air pollution emissions in cities, and may become a complement to traditional emission estimates, especially in emerging cities and countries where reliable fine-grained urban air quality data is not easily available. This is the first study of its kind to investigate measured microscopic vehicle movement in tandem with microscopic emissions modeling for a substantial study domain.

Keywords: air quality, transportation, emissions, microscopic emissions model, microscopic vehicle movement.

1. Introduction

With mass urbanization happening at an unprecedented scale, urban air quality is becoming an issue of global concern (WHO, 2014). Growth in populations, traffic, industrialization and energy usage have led to increased air pollution levels and subsequent public health effects at the urban, regional and global scale (Akimoto, 2003; Molina et al., 2004; Gurhar et al., 2010) The World Health Organization estimates that ambient air pollution leads to approximately 3.7 million premature deaths annually worldwide, with South-East Asia and the Western Pacific Regions having the largest air pollution-related health burden (WHO, 2014).

The adverse impact of air pollution exposure on human health is well documented in the literature (WHO, 2014). Epidemiological studies have quantified the relationship between adverse health effects and both long- and short-term exposure to air pollution (Bell et al., 2004; Jerrett et al., 2005; Laden et al., 2006; Lewtas, 2007; Krewski et al., 2009; Nyhan et al., 2014a; Nyhan et al., 2014b). In assessing the impact of air pollution on mortality in the United States, Caiazzo et al. (2013) reported that the largest sector contributor of pollutant-related mortalities is road transportation, causing approximately 53,000 $PM_{2.5}$-related deaths and approximately 5000 ozone-related deaths per year. These figures corresponded to premature deaths from cardiovascular diseases and lung cancer due to long-term exposure to $PM_{2.5}$ (where $PM_{2.5}$ refers to the particulate matter fraction which is less than 2.5μm in aerodynamic diameter).

Traditional methods for monitoring urban air quality employ discrete measurement stations which sample atmospheric conditions at specific sites throughout a city. Networks vary both in size and scale. The London Air Quality Network has over 50 sites classified as roadside, background, suburban and industrial that are dispersed throughout the whole metropolitan area (Laxen et al., 2003). Singapore, which is the focus of this study, has 14 high-grade stations operated by the National Environment Agency, gathering data throughout the island (NEA, 2015). Traditional approaches to monitoring air quality have several limitations, including significant investment required to set up and maintain the measurement networks. Furthermore, as air quality can exhibit large variations over a relatively small scales (Britter and Hanna, 2003), sampling biases can be introduced which make the assessment of human exposure and the sources of pollutants difficult (Vardoukalis et al., 2005). As a result of this, municipal air quality monitoring is often supplemented by air quality models such as the AERMOD modeling system (USEPA, 2009) and the ADMS Urban model (CERC, 2015) to improve the spatial and temporal resolution of air pollution estimates. Sparsely located air quality monitors are limited in their usefulness for accurately determining the locations of air pollution sources. Therefore, air quality monitoring using distributed networks of sensors has gained traction as sensors are becoming smaller, less expensive yet more reliable (Chong et al., 2003; Burke et al., 2006; Cuff et al., 2008; Paulos et al., 2009; Kumar et al., 2015), providing a wealth of high spatial resolution air quality information.

The availability of large transportation and mobility datasets from sensors, Global Positioning System (GPS)-enabled devices, along with improvements in methods and computational facilities for analyzing these have led to advancements in the field of urban computing research in recent times. So-called opportunistic sensing which is the use of data that is collected for one purpose but can be reused for another one (Campbell et al., 2008), has proved useful in many research studies. Examples include using various anonymized or aggregated spatio-temporal datasets created by different aspects of human activity, such as cell phone data (Gonzales et al, 2008; Sobolevsky et al, 2013; Hoteit et al, 2014; Kung et al, 2014; Pei et al, 2014; Grauwin et al, 2014) or vehicle GPS traces (Kang et al, 2013). One such example of opportunistically utilizing vehicle GPS traces is a recent study by Santi et al., (2014) where the economic and environmental benefits of vehicle pooling in New York were quantified based on the analyses of a taxi GPS dataset consisting of 150 million trips.

Emissions from on-road motor vehicles constitute one of the largest contributions to air pollutants such as carbon monoxide, nitrogen dioxide, ozone, selected volatile organic compounds and fine particulates (Molina and Molina, 2004), and also represent a factor in the spatial variability of air quality in urban areas (Fecht et al., 2016). Vehicle emissions have typically been estimated with the use of either measured (through loop detectors or similar) or modeled (using a transport simulator) traffic data. Based on this information, emission factors are commonly used to convert traffic loads into emissions (NARSTO, 2005). Emission factors vary from location to location, and depend on the vehicle model and road conditions (Zhang and Morawska, 2002; North et al., 2006). The application of emission factors to traffic loads is unable to account for real driving conditions as they happen on the road (Samuel et al., 2002). Thus, as an alternative, different vehicles models with different load factors are often used as probes, whose emissions (and eventually the emission of nearby vehicles) are measured on the road (Canagaranta et al., 2004; Shorter et al., 2005). The aforementioned approaches do not allow the high resolution spatiotemporal mapping of emissions as they do not take into account the 'drive cycle' which is the description of a vehicle's velocity over time. The drive cycle allows the precise determination of consumption and hence emissions (Mantazeri et al., 2003; Int Panis et al., 2006). In the widely used MOBILE Model (USEPA, 2012), only 14 different drive cycles are used; however, these are only expressed as average speed. Many studies have examined the impact of different vehicle modes (idling, moving and accelerating) on the release of pollutants. In a study by Frey et al., (2003) average emissions were observed to be five times greater during periods of acceleration for hydrocarbons and carbon dioxide, and reached ten times as much for nitric oxide and carbon monoxide compared with levels found in an idling vehicle. Similarly, ultrafine particulates released whilst a vehicle is accelerating have also been shown to increase significantly (Fruin et al., 2008). Hence, there is a need for the use of more detailed drive cycles, including velocity and acceleration parameters resolved in high spatial and temporal resolution, in modeling emissions from transportation.

Many studies have led to the development of models that consider variations in speed and are appropriate for instantaneous emission modeling. These include the Comprehensive Modal Emissions Model developed at the University of California (An et al., 1997; Barth et al., 2006) and others (e.g. Rakha et al., 2004; Pelkmans et al., 2004; El-Sgawarby et al., 2005). Along with this, significant effort has been devoted to the use of micro-simulation methods for transportation modeling on road networks, for representing real-time, behavior-based policies (e.g. Ben-Akiva et al., 1997; Hu and Mahmassani, 1997; Liu et al., 2006). Individual driver behavior and individual vehicle's real-time space-time trajectories are explicitly represented through traffic micro-simulation models and these produce detailed vehicle operation, instantaneous speed and acceleration of vehicles that are necessary for microscopic emissions models. A review by Fontes et al., (2015) examined combining various micro-simulation tools for assessing the impacts of road traffic on the environment, and identified best practices which would aim to minimize errors in combining these. Int Panis et al., (2006) presented a methodology for making instantaneous emission modeling compatible with traffic micro-simulation models. In particular, the emissions caused by acceleration and deceleration of vehicles were modeled based on microscopic traffic simulation model integrated with an instantaneous emission model. The functions developed by Int Panis et al., (2006) were incorporated into a study addressing optimum mitigation strategies for urban transportation emissions by Osorio and Nanduri (2015) where a combination of macroscopic and microscopic traffic simulators and emissions models were employed.

Recent developments in the field of vehicle emissions have seen the uptake of cell phones and their built in sensors as on-board diagnostic systems - using the data gathered from the GPS and accelerometer to monitor the drive cycle and hence consumption and emissions (Thiagarajan et al., 2009). These approaches have been mostly confined to single or small numbers of vehicles. In this study, however, it is intended to extend an emissions model to a large vehicle fleet using GPS data collected. Intelligent Speed Adaption (ISA) systems are technologies which incorporate GPS navigation to apply speed limits to cars on specific road areas. Systems for monitoring and controlling vehicle velocities include ISA systems (Duynstee et al., 2001; Int Panis et al., 2006). These could also be used for reducing emissions and fuel consumption on road networks, but require fine-grained emissions predictions based on real-time GPS data.

The purpose of this study is to use data routinely captured by existing transportation networks and vehicle fleets to predict vehicular emissions in high spatial resolution. For this, GPS measurements gathered by a large taxi fleet in Singapore would be analyzed. Parameters representative of vehicle drive cycles would then be characterized in high spatial and temporal resolution at points throughout the road network. A microscopic emissions model would be implemented to predict the emissions of carbon dioxide (CO_2), nitrogen oxide (NO_x), volatile organic compounds (VOCs) and particulate matter (PM) throughout the study domain, where particulate matter here refers to total suspended particles. Highly localized areas of elevated emissions would thereby be identified, with a higher spatiotemporal precision than commonly used methods. This is the first study to implement a microscopic emissions model using measured microscopic vehicle trajectory data for an entire urban region.

2. Methodology

2.1. Overview of methodology

In order to develop an emissions inventory, GPS trajectory data from 15,236 taxis were analyzed. From this, the instantaneous parameters of velocity and accelleration were derived and used as inputs for a microscopic emissions model. Emissions of CO_2, NO_x, VOCs and PM were predicted across the road network of Singapore using this model. An analyses was completed which compares the taxi data used to the overall traffic on the road network in Singapore. Following this, emissions from the remainder of the total motor vehicle population of Singapore were also estimated. The results were compared to emissions estimates produced to those attained from the National Aeronautical and Space Agency (Streets and Lu, 2012).

2.2. Study domain and GPS data processing

The study domain included the island of Singapore, which covers approximately 718 km^2. Singapore has a population of 5,469,700 people (Singapore Department of Statistics, 2014), therefore has an average population density of 7,618 persons per km^2.

Our analysis used vehicle GPS traces collected over a period of one week from 15,236 taxis in Singapore. The raw data included the following parameters: identification number of the vehicle, a timestamp of when each location measurement was performed, the corresponding latitude and longitude defining the position of the vehicle. The data samples were collected at varying temporal intervals every few seconds. Our data was collected from an undisclosed vehicle fleet operator, which operates over the majority of the island of Singapore on a 24-hour basis. Each vehicle contained within the fleet transmits information including its identification number, location and status at various intervals to a central operations base. The dataset contained over 120 million vehicle-GPS samples measured from the 21st February 2011 to the 27th February 2011.

The GPS trace data was utilized to infer both the location of each vehicle, its velocity and its acceleration. In applying a data cleaning process to the dataset, erroneous GPS points which fell outside the boundary of Singapore or which have an unreasonable distance from its previous location at a given time interval (distance/time ≤ 150 km/h) were eliminated. The instantaneous velocities of vehicles were determined based on the time and distance between geo-referenced points. The data was filtered so as to only examine changes in velocity that occurred over short temporal ranges, where two consecutive data points were separated by no more than 5 seconds as intervals greater than this are unable to depict the microstructure of the acceleration profile. A secondary filtering process was applied to the data to remove errors attributed to GPS measurements, as these may be affected by the multi-path effect within urban canyons (Parkinson, 1996). An outlier filter was used that removed all the acceleration values that exceeded 10 ms^{-2} as these values are generally not attainable in an average car. The normative driving cycle, used to homologate vehicles emissions are characterized by a maximum acceleration of 1.5 ms^{-2} for FTP-72 and 4 ms^{-2} for LA92 (Guzella and Sciarreta, 2005; Metric Mind Corporation, 2012), therefore sampling points with an acceleration value between 0.5 and 10 m s^{-2} were used in this study.

2.3. Comparison of taxi fleet and total traffic

By applying the above filters, the distribution of the sampling intervals of the 15,236 taxis, indicate that only 7.71% of the logged data has a sampling interval of less than 5 seconds as well as a valid acceleration value. This indicates that the majority of vehicles demonstrate intermittent data logging at intervals greater than 5s. The spatial distribution of the valid samples was correlated with a co-efficient of determination of 0.75 to the spatial distribution of the raw vehicle-GPS points. In order to examine the spatial distributions of GPS points, the city was divided into road links. The valid accelerations of all the vehicles were then attributed to one of the road links based on their latitude and longitude data, and were projected onto a map of Singapore.

Aslam et al., (2012) demonstrated that vehicular GPS taxi network data can be used to infer general traffic patterns in Singapore. Aslam et al., (2012) used data from the same taxi fleet as used herein this study. Measured traffic data (i.e. counts of vehicles on road links per time intervals) were obtained through loop count data from the Land Transport Authority (LTA) of Singapore. By examining the fraction of road segments the taxi fleet covers during workdays, it was concluded that 700 taxis were sufficient to cover 70% of the roads for the majority of the day's 1-hour time windows, with the exception of those in the middle of the night when vehicle numbers are sparse. Further to this, Aslam et al., (2012) also observed that 2000 taxis were sufficient to cover 90% of the total loop detector locations during a period of 15 minutes in the morning (from 08:00~08:15) on all workdays. Similarly, we compared our taxi fleet data to measured traffic data obtained from loop detectors operated by Singapore's LTA for the same time period as our study. To achieve this, the taxi data was synchronized with the loop detector data, which was aggregated every 15 minutes. The time series of GPS points for taxis were first matched to road links and then segments on the road network of Singapore. The number of taxis on road segments where loop detectors are located, were counted every 15 minutes. These counts were then compared to the loop counts which were regarded as the ground truth for traffic conditions. Figure 1 shows the taxi and loop detector count data for 15 randomly selected Singapore road segments. The taxi distribution tended to underestimate the loop distribution and this underestimation was variable across road segments. On each road link, a bias was observed which varied throughout the day, however this bias was relatively consistent across days.

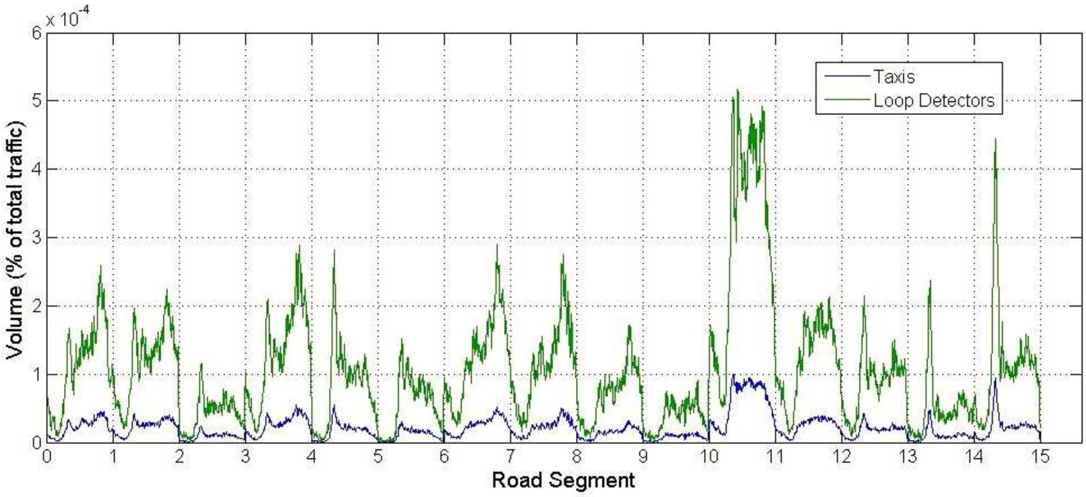


Figure 1. Distribution of traffic volumes (i.e. number of vehicles per road segment) on 15 randomly selected road segments for the 23rd February 2011. The x-axis includes 15 road segments including a point for every 15 minutes during the 24-hour day. The y-axis represents the percentage of traffic at that location and time. The taxi distribution (in blue) underestimates the loop distribution (in green) and the underestimation is variable.

For inferring general traffic patterns, an artificial neural network model was employed, as has been used in another study for predicting traffic volumes on road links (Moretti et al., 2015). The model utilized was a simple corrective model for inferring vehicle distribution as detected by loop detectors from vehicle distribution as determined by the taxi fleet. A 2-layer feed-forward network was implemented, with a tan-sigmoid transfer function in the hidden layer and linear transfer function in the output layer. The model was run for 500 road segments. In determining the performance of the model, a linear regression between modeled traffic volume and the corresponding targets of measured traffic volume was conducted. Figure 2 shows the results of learning for trained model for a sample of data points. As there is a strong association between the modeled and measured traffic volumes, this demonstrates that the taxi fleet data may be used to predict general traffic on specific road segments, and the results were similar across the road network of Singapore.

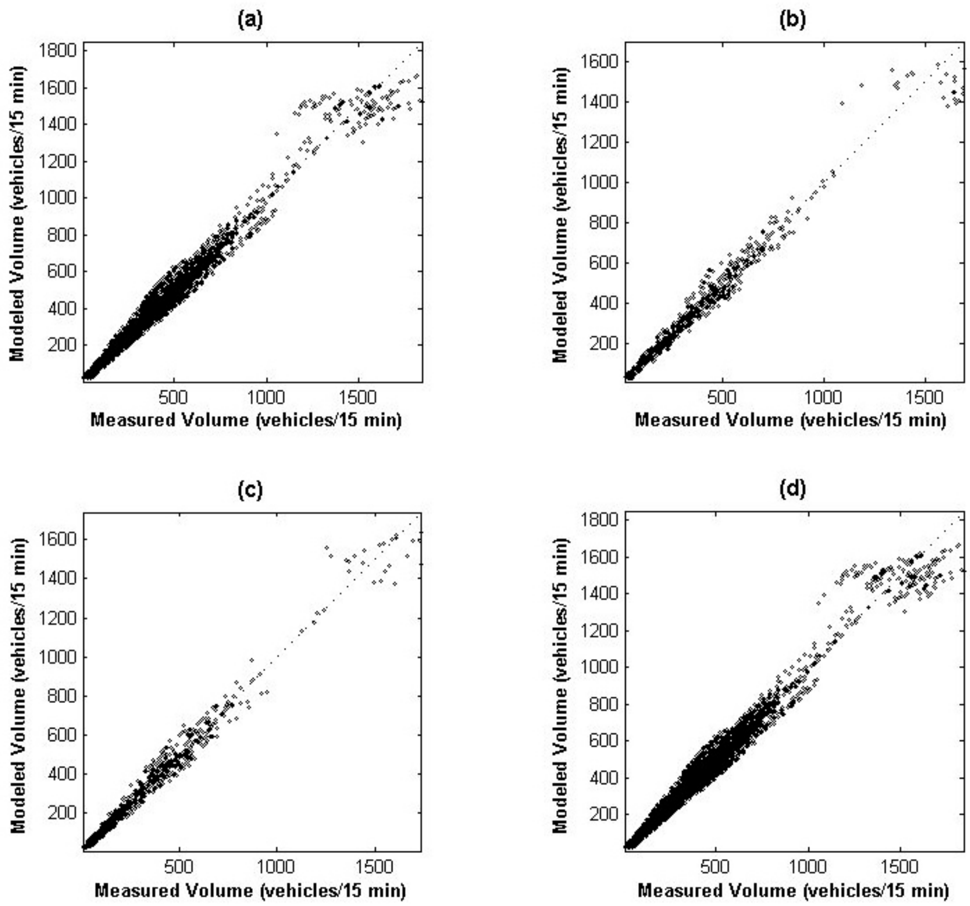


Figure 2. Results of the feed-forward artificial neural network model implementation. Regression plot of a partial set of modeled traffic volumes versus corresponding measured traffic volume for the (a) training phase (R^2=96%), (b) validation phase (R^2=93%), (c) tesing

phase (R^2=92%) and (d) overall model (R^2=94%). A sub-sample of points are presented for clarity.

2.4. Microscopic emissions model

A microscopic emissions model was implemented and this computed the instantaneous air pollution emissions associated with CO_2, NO_x, VOCs and PM. The emissions model was based on a model developed by Int Panis et al., (2006), and has been adopted by Osorio and Nanduri (2015). The model utilizes instantaneous velocity and accelerations derived from the GPS dataset to compute emissions. The emission rate at a given time-instant t is given in the following equation:

$$ER_n^k(t) = max[E_{0n}^k, f_{1n_1}^k + f_{2n}^k v_n(t) + f_{3n}^k v_n(t)^2 + f_{4n}^k a_n(t) + f_{5n}^k a_n(t)^2 + f_{6n}^k v_n(t) a_n(t)], \quad (1)$$

where k is the pollutant type, i.e. $k \in$ {CO_2, NO_x, VOC, PM}, $v_n(t)$ is the instantaneous speed of vehicle n at time t (in m/s), $ER_n^k(t)$ is the instantaneous emissions rate of pollutant k (in g/s), $a_n(t)$ is the instantaneous acceleration of vehicle n at time t (in m/s^2), E_{0n}^k is the lower limit of emission rate for each pollutant type (in g/s), and f_{1n}^k, f_{2n}^k, f_{3n}^k, f_{4n}^k, f_{5n}^k and f_{6n}^k are the emission rate constants specific to each vehicle and pollutant type. Equation (1) holds for CO_2 and PM emissions. For NO_x and VOC emissions, the emissions rate coefficients differ depending on whether the vehicle is in acceleration or deceleration mode. If $a_n(t) \geq -0.5\ m/s$, then

$$ER_n^k(t) = max[E_{0n}^k, f_{1n_1}^k + f_{2n}^k v_n(t) + f_{3n}^k v_n(t)^2 + f_{4(1)n}^k a_n(t) + f_{5(1)n}^k a_n(t)^2 + f_{6(1)n}^k v_n(t) a_n(t)], \quad (2)$$

Otherwise, if $a_n(t) < -0.5\ m/s$, then

$$ER_n^k(t) = max[E_{0n}^k, f_{1n_1}^k + f_{2n}^k v_n(t) + f_{3n}^k v_n(t)^2 + f_{4(2)n}^k a_n(t) + f_{5(2)n}^k a_n(t)^2 + f_{6(2)n}^k v_n(t) a_n(t)], \quad (3)$$

The lower limit of the emissions rate E_0 is fixed to zero for all pollutant types and vehicle types. The emission rate constants (e.g., f_1, f_2, etc.) are specified for each pollutant type and vehicle type, and were determined from emissions measurements of on-road instrumented vehicles. These were determined for the car, heavy duty vehicle (HDV, diesel) and bus (diesel) categories. A table describing these emission rate constants are described in Int Panis et al., (2006).

For each pollutant, the expected total emissions (in g) in the specified vehicle network during the simulation period were computed by:

$$E[TE^k] = \sum_{l \in L} E[TE_l^k], \quad (4)$$

where L is the set of all road links in the network, and $E\left[TE_l^k\right]$ denotes the total emissions (in g) of pollutant k on link l. The latter term in Equation (4) is approximated by:

$$E\left[TE_l^k\right] = E[ER^{k,l}]E[T_l]\lambda_l \Delta T, \quad (5)$$

where $E[ER^{k,l}]$ denotes the expected emissions rate (in g/s) for link l and pollutant type k, $E[T_l]$ is the travel time on link l, λ_l is the arrival rate of vehicles to link l and ΔT is the total simulation time. For a given link l and pollutant type k, the term $\lambda_l \Delta T$ approximated the expected total demand over the time period of interest, while $E[ER^{k,l}]E[T_l]$ approximated the expected emissions per vehicle. The emissions computed for each road link were projected onto a map of Singapore.

Emissions for the total motor vehicle population, represented by general traffic patterns, across the road network of Singapore were quantified. Emissions were estimated on a daily basis according to Equation (5). In this scenario however, the arrival rates of vehicles to each road link, λ_l, were predicted using the traffic model described in Section 2.3. Daily emissions were calculated for each of five days of data available, and the mean of these five days was then compared to mean daily emissions estimated by Streets and Lu, (2012).

2.5. Vehicle fleet composition

The emissions model took into consideration the estimated composition of the vehicle fleet of Singapore. This was based on information collected by the Land Transport Authority of Singapore (LTA, 2015). The data-set yielded counts of the various categories of motor vehicles within the overall transportation fleet i.e. Cars, Taxis, Motorcycles, Goods and Other Vehicles, and Buses, and these categories were further stratified by type of fuel used i.e. petrol, diesel, petrol-electric, petrol-CNG, CNG and electric for each of the respective categories of vehicle type. Data for the year 2011 were used as this corresponded to our vehicle data-set (see Table 1 for details).

Table 1. Motor vehicle population in Singapore by category and type of fuel used for the year 2011. Figures exclude tax exempted vehicles for off-the-road use (RU plates).

Cars	Petrol	596,947
	Diesel	346
	Petrol-Electric	3,786
	Petrol-CNG	2,642
	CNG	-
	Electric	2
	Total	603,723
Taxis	Petrol	279
	Diesel	23,880
	Petrol-Electric	56
	Petrol-CNG	2,836
	CNG	-
	Electric	-
	Total	27,051
Motorcycles	Petrol	145,672
	Electric	8
	Total	145,680
Goods & Other Vehicles	Petrol	9,058
	Diesel	136,076
	Petrol-Electric	1
	Petrol-CNG	14
	CNG	8
	Electric	1
	Diesel-Electric	-
	Total	145,158
Buses	Petrol	194
	Diesel	16,433
	Petrol-Electric	-
	Petrol-CNG	8
	CNG	14
	Electric	3
	Total	16,652

3. Results

3.1. Spatial distribution of accelerations and predicted emissions

Figure 3 shows counts of all valid acceleration data on each link on the road network. Higher counts of valid accelerations were concentrated in the Singapore Downtown Core area, at Changi International Airport and some parts of Jurong, Bishan and Yishun. As demonstrated in Section 3.3., the taxi data may be used to predict general traffic on road segments, therefore counts of valid accelerations were proportional to the distribution of vehicles in the city, and proportional to the number of accelerations of each road link. Valid accelerations on each road link were utilized for the emissions model. However, areas such as the Singapore Downtown Core area and the vicinity of Changi International Airport which were characterized by a relatively higher number of sample points of acceleration than other areas. This may indicate a bias in the dataset.

The spatial distributions of vehicle emissions computed for each road link in Singapore are shown in Figure 4. With regards emissions related to specific parameters, we can see that for all of CO_2, NO_x, VOC and PM, elevated levels were identified in a concentrated number of locations in the Singapore Downtown Core area, south of Newton and in Geylang. Elevated levels were also identified in the area surrounding Changi International Airport, Bishan and Jurong West.

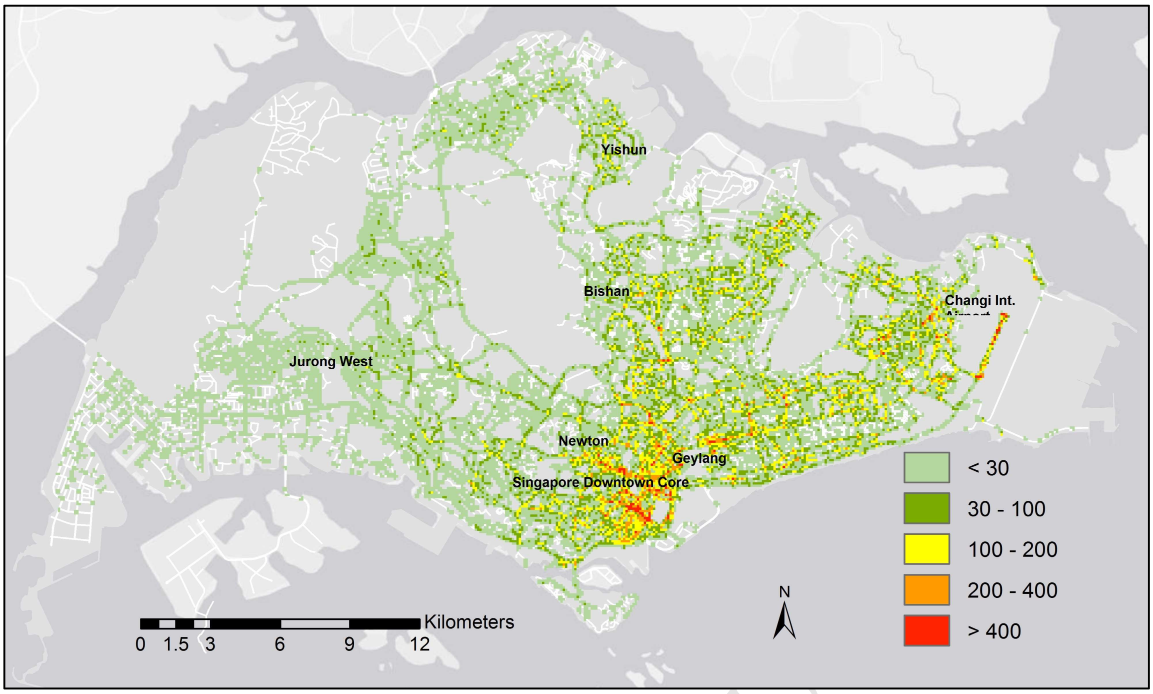
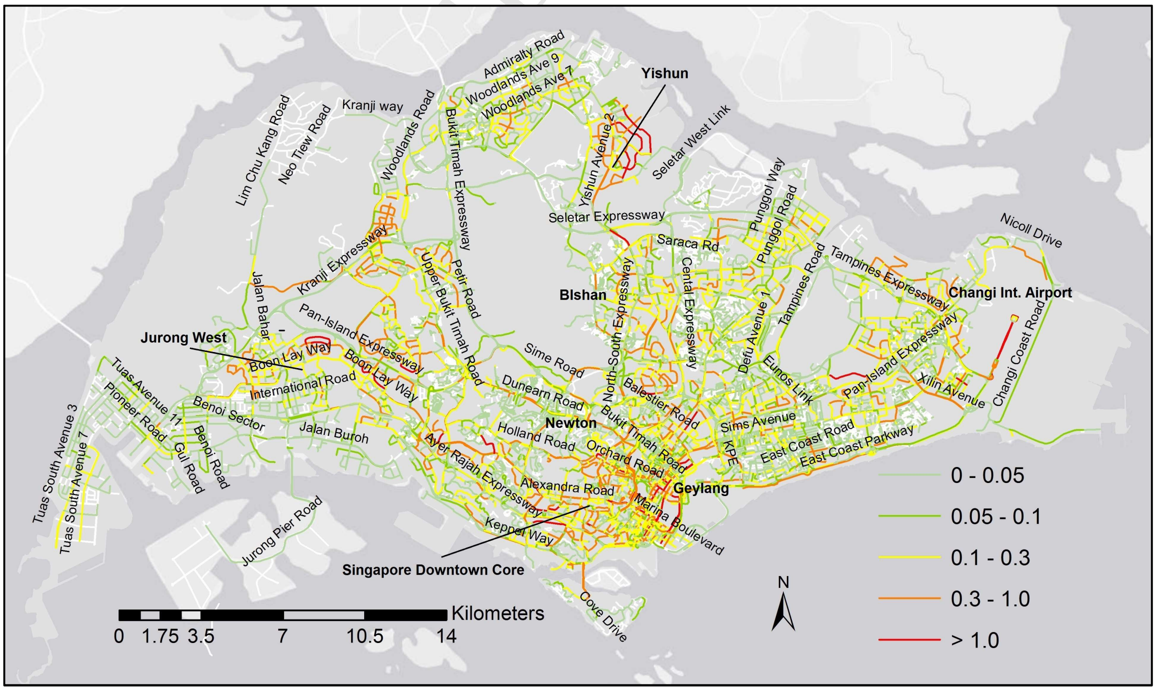
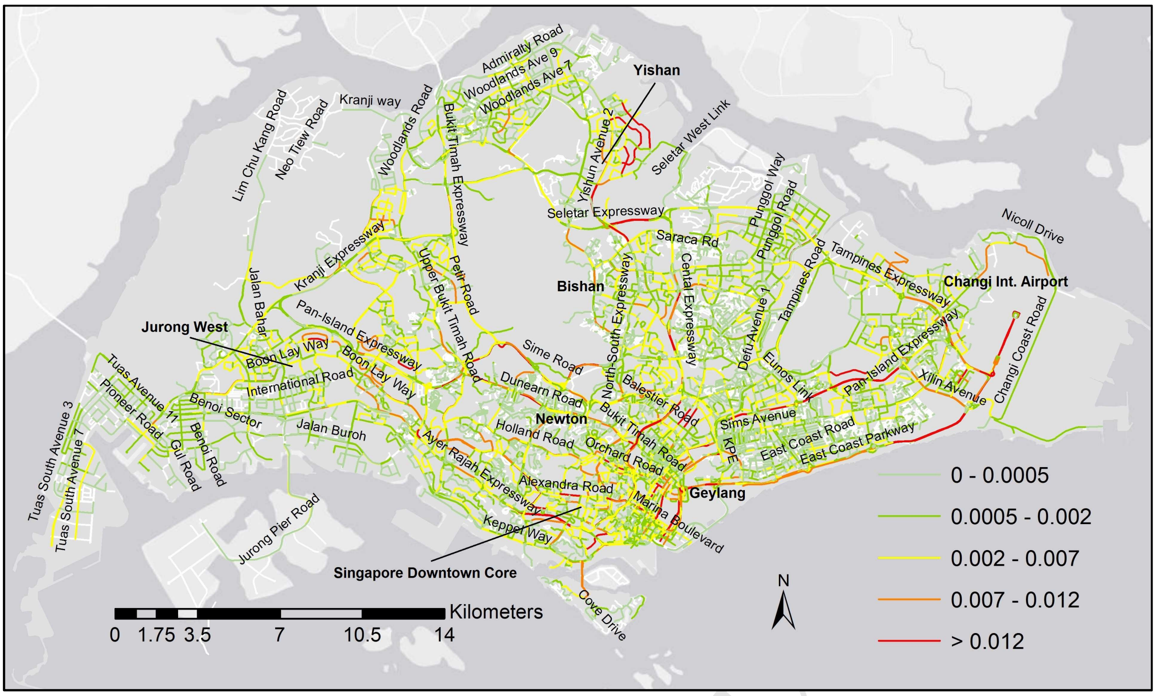


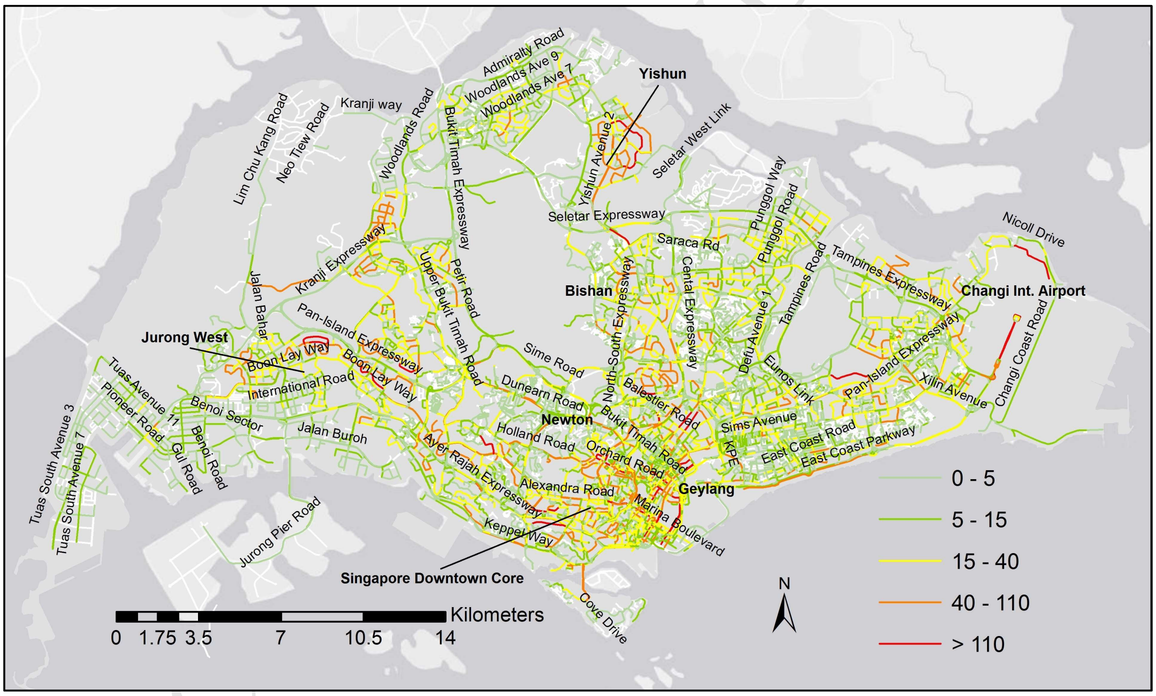
Figure 3. Spatial distribution of the number of valid accelerations in Singapore on the 23rd February 2011. Locations where relatively higher numbers of valid accelerations are observed in the vicinity of the Singapore Downtown Core area and the Changi International Airport in the east.



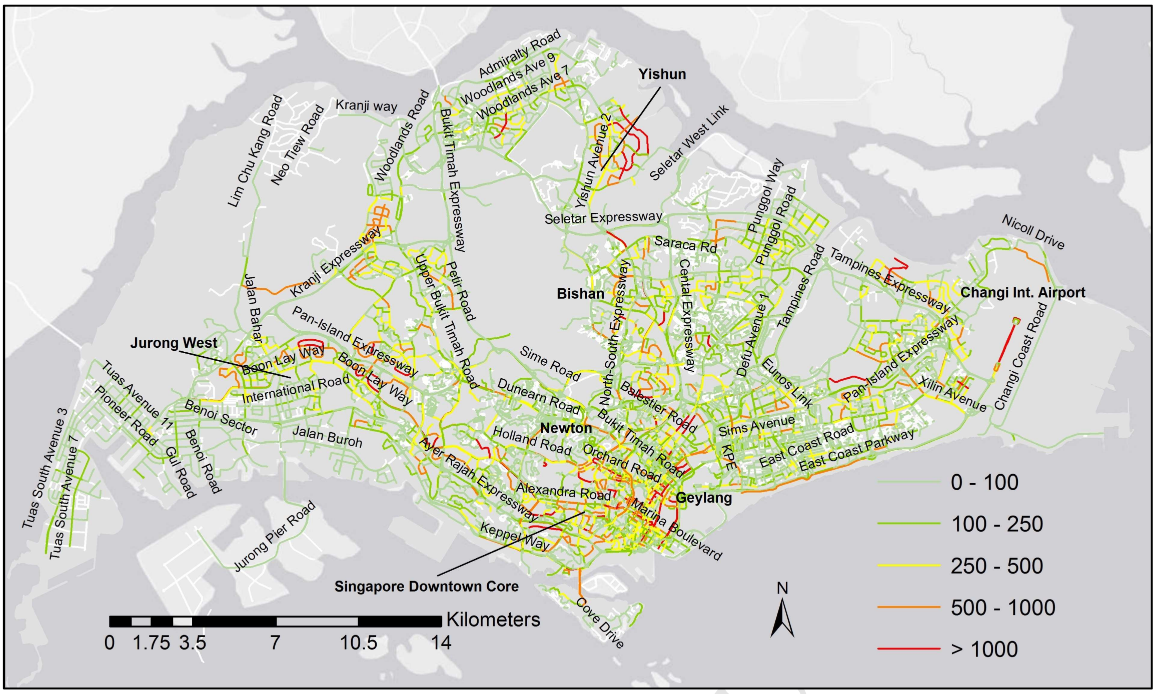
(a)


Admiralty Road
Woodlands Ave 9
Woodlands Ave 7
Yishan
Kranji way
Lim Chu Kang Road
Neo Tiew Road
Woodlands Road
Bukit Timah Expressway
Yishun Avenue 2
Seletar West Link
Seletar Expressway
Punggol Way
Punggol Road
Nicoll Drive
Saraca Rd
Tampines Expressway
Changi Int. Airport
Kranji Expressway
Upper Bukit Timah Road
Pelir Road
Bishan
Central Expressway
Defu Avenue 1
Tampines Road
Jalan Bahar
Jurong West
Pan-Island Expressway
Boon Lay Way
International Road
Sime Road
North-South Expressway
Eunos Link
Pan-Island Expressway
Changi Coast Road
Xilin Avenue
Tuas Avenue 11
Pioneer Road
Benoi Sector
Dunearn Road
Newton
Balestier Road
Bukit Timah Road
Sims Avenue
Jalan Buroh
Holland Road
Orchard Road
East Coast Road
East Coast Parkway
KPE
Tuas South Avenue 3
Tuas South Avenue 7
Gul Road
Benoi Road
Ayer Rajah Expressway
Alexandra Road
Geylang
Marina Boulevard
Jurong Pier Road
Keppel Way
Singapore Downtown Core
Cove Drive
0 1.75 3.5 7 10.5 14
Kilometers
N
0 - 0.0005
0.0005 - 0.002
0.002 - 0.007
0.007 - 0.012
> 0.012

(b)


Admiralty Road
Woodlands Ave 9
Woodlands Ave 7
Yishun
Kranji way
Lim Chu Kang Road
Neo Tiew Road
Woodlands Road
Bukit Timah Expressway
Yishun Avenue 2
Seletar West Link
Seletar Expressway
Punggol Way
Punggol Road
Nicoll Drive
Saraca Rd
Tampines Expressway
Changi Int. Airport
Kranji Expressway
Upper Bukit Timah Road
Pelir Road
Bishan
Cental Expressway
Defu Avenue 1
Tampines Road
Jalan Bahar
Jurong West
Pan-Island Expressway
Boon Lay Way
International Road
Sime Road
North-South Expressway
Eunos Link
Pan-Island Expressway
Changi Coast Road
Xilin Avenue
Tuas Avenue 11
Pioneer Road
Benoi Sector
Dunearn Road
Newton
Balestier Road
Bukit Timah Road
Sims Avenue
Jalan Buroh
Holland Road
Orchard Road
East Coast Road
East Coast Parkway
KPE
Tuas South Avenue 3
Tuas South Avenue 7
Gul Road
Benoi Road
Ayer Rajah Expressway
Alexandra Road
Geylang
Marina Boulevard
Jurong Pier Road
Keppel Way
Singapore Downtown Core
Cove Drive
0 1.75 3.5 7 10.5 14
Kilometers
N
0 - 5
5 - 15
15 - 40
40 - 110
> 110

(c)



(d)

Figure 4. Spatial distributions of predicted daily emissions from the vehicle fleet for each road link for the parameters of (a) CO_2 (tonnes/day), (b) NO_x (tonnes/day), (c) VOC (g/day), and (d) PM (g/day) in Singapore on the 23rd February 2011. Locations of relatively high-emissions, are observed in the Singapore Downtown Core area in the south-center of Singapore and in other locations throughout the island.

The locations where predicted emissions were relatively higher across Singapore can be identified for the four pollution parameters of CO_2, NO_x, VOC and PM. In terms of CO_2 emissions, the areas which were identified as having relatively higher CO_2 output from the vehicle fleet. Marina South and Raffles Place in the Downtown Core area, the Harbour Front area, Jurong East, Clementi, Sin Ming and an area close to the Seletar Reservoir in Yishun were identified. In the east of Singapore, the area between Tampines and Changi International Airport was identified as having relatively higher CO_2 emissions than other areas. The Bukit Timah Road - Whitley Road Intersection was selected as having relatively higher CO_2 emissions as were the busy areas Novena, Newton, Somerset, Dhoby Ghaut (north) and Farrer Park which are located north of the central region of Singapore.

Relatively higher levels of NO_x emissions were predicted in the Downtown Core Area such as in Chinatown, Outram Park, Clarke Quay and Raffles Place. The Chin Swee Tunnel - Havelock Road intersection area was also identified. North of Dhoby Ghaut, City Hall, on the

Central Expressway side of Fort Canning Park and Little India were areas where relatively higher NO_x emissions were predicted. Connected to these, Somerset and Orchard were areas with relatively higher NO_x emissions. The Moulmein Flyover, the Jalan Bukit Merah - Lower Delta Road Intersection (located west of the Downtown Core area), the Kallang-Paya Lebar Expressway (KPE) - Nicoll Highway Intersection (located east of the Downtown Core area), and further east, an area in the vicinity of Changi International Airport was also identified. For VOC emissions, the areas of elevated emissions were observed to be centrally located with a few areas scattered in other parts of Singapore. Located centrally were Orchard Road, the River Valley Road - Zion Road Intersection, Outram Park, Marina South, Suntec City and Little India. Moving east from the urban central region - Selegie, Lavender, Kallang, Geylang, and further east, the Layang Avenue - Pasir Ris Drive 1 Intersection and the Pan Island Expressway (PIE) - Tampines Expressway (TPE) Intersection near Changi International Airport were identified as hotspots for VOC emissions. North-east of the central region was Tao Payoh and further north was Sin Ming (Yishun area). Westwards from the Downtown Core areas were Bukit Merah, the Hollande Road - Farrer Road Intersection. Further west was Clementi, Jurong East and Bukit Batok. In the north-west of Singapore, Choa Chu Kang was observed to have relatively higher levels of VOC.

For PM emissions, all the areas of relatively highest predicted emissions were concentrated in the Downtown Core area with some areas identified to the east of it. The areas identified included the areas of Outram Park, Chinatown, Raffles Place and Clarke Quay. South of these the Tajang Pagar area near Keppel Road was chosen and slightly north of these, the Havelock Road - Outram Road Intersection. River Place near the Chin Swee Tunnel and the Central Expressway side of Fort Canning Park were identified. On the east of the Downtown Core area were Bugis, Beach Road and Geylang. On the west side of the Downtown Core area; the Jalan Bukit Merah - Lower Delta Road Intersection was included in the selection of areas determined to have increased PM emissions relative to the rest of the island. Other areas were the extent of Orchard Road, Farrer Park and Balestier.

3.2. Comparison of predicted emissions for the total motor vehicle population

Total emissions of each pollutant parameter for the vehicle fleet studied were computed for each day and the means determined are presented in Table 2. The mean daily CO_2 and NO_x

emissions determined were respectively representative of 7.9% (±3%) and 7.6% (±1.4%) of emissions estimates for the total motor vehicle population of Singapore (including the previously calculated vehicle fleet emissions). For VOC and PM, the proportions were smaller, whereas the total daily emissions computed were approximately 3.2% (±1.7%) and 3.5% (±1.6%) (respectively) of total motor vehicle population emissions (see Table 2 for details).

Daily emissions from the total motor vehicle population were then computed for one week and compared to other transportation emissions estimated by Streets and Lu (2012) (see Table 3). The overall emissions levels computed for the entire fleet were comparable to those attained from Streets and Lu (2012). Whereas our analyses predicted mean daily emissions from the entire motor vehicle population to be 27656 (±3049) tonnes for CO_2, Streets and Lu computed 24417 tonnes/day. Therefore, the relative difference in emissions was found to be 15% (±1.7%). For NO_x we determined total daily emissions to be 155 (±33.1) tonnes/day while Streets and Lu computed 121 tonnes/day, and this corresponded to a relative difference of 24% (±4.9%). A larger disparity was observed in the case of VOC. We predicted total emissions to be 9.7 (±2.6) tonnes/day whereas Streets and Lu determined a value of 21.6 tonnes/day. This is equivalent to a relative difference of -49% (±12.3%). Finally, for PM we computed 8.5 (±3.4) tonnes/day while Streets and Lu predicted 14.1 tonnes/day. Similar to VOC, we calculated a relatively lower value to Streets and Lu by 39% (±15.5%), but exhibiting a larger uncertainty.

Table 2. Modeled emissions for the taxi fleet and the proportion of modeled taxi emissions in the estimated total motor vehicle population emissions, for each of four air pollutant parameters.

	Modeled Emissions Taxi Fleet (tonnes/day)	Proportion of Modeled Taxi Emissions in the Total Motor Vehicle Population Emissions %
	Mean (SD)	**Mean (SD)**
CO_2	2176.6 (1023.5)	7.9 (3.0)
NO_x	11.9 (2.8)	7.6 (1.4)
VOC	0.3 (0.2)	3.2 (1.7)
PM	0.3 (0.2)	3.5 (1.6)

Table 3. Comparison of the mean daily emissions predicted for the total motor vehicle population of Singapore to estimated ground transportation emissions attained from Streets and Lu (2012).

	Predicted Total Motor Vehicle Population Emissions (tonnes/day)	Streets and Lu (2012) (tonnes/day)		
	Mean (SD)	**Mean**	**Range of Ratios**	**Average difference (SD) (%)**
CO_2	27656 (3049)	24417	(1.1-1.3)	15.1 (1.7)
NO_x	155.2 (33.1)	121	(1.0-1.6)	24.1 (4.9)
VOC	9.7 (2.6)	21.6	(0.4-0.7)	-49.3 (12.3)
PM	8.5 (3.4)	14.1	(0.4-0.8)	-38.7 (15.5)

4. Discussion

Recent advances in urban computing and the availability of large transportation GPS datasets have presented new opportunities for real-time transportation and emissions modeling. Transportation and emissions modeling conducted in previous studies have been limited by coarsely resolved predicted or measured traffic information. In this study, we analyzed GPS traces from a fleet of over 15,000 vehicles in Singapore with the aim of using this information to make predictions of emissions in high spatial resolution throughout the study domain. The instantaneous velocities and accelerations of vehicles, which were extracted in high spatial and temporal resolution, were inputted into a microscopic emissions model. The air pollution emissions of CO_2, NO_x, VOC and PM were thus quantified. The spatial distributions of the emissions were examined and this enabled highly localized areas of elevated emission levels to be identified. The study demonstrated how instantaneous drive cycles can be used to predict vehicular pollutant emissions and this forms an important component of the urban emissions inventory.

An analyses demonstrated that the taxi data could be used to predict overall traffic volumes on road segments throughout the road network. Emissions from the taxi fleet and then the total motor vehicle population were therefore predicted for the study domain of Singapore. The subsequent emissions levels computed for the entire motor vehicle population was comparable to those attained from Streets and Lu (2012). Whereas the modeled values are in the same order of magnitude for each pollutant parameter, the results likely varied due to the different emissions modeling methods employed. Further to this, in the case of Streets and Lu (2012) estimates of emissions from the transportation sector are from the year 2012, while our data are representative of one week of data for the year 2011. Predicted emissions computed for CO_2 and NO_x were higher than VOC and PM. The reason for this is that the emissions function parameters used are higher for CO_2 and NO_x. CO_2 and PM emission estimates are more sensitive to vehicle velocities than VOC and NO_x which are more sensitive to accellerations (Int Panis et al., 2006).

This paper presents a novel methodology for making instantaneous emission modeling compatible with microscopic traffic patterns (measured on a second by second basis). Previous studies have focused on the microscopic traffic simulation coupled with microscopic emissions modeling (Int Panis et al., 2006) or a combination of macroscopic and microscopic traffic simulation combined with microscopic emissions modeling (Osorio and Nanduri, 2015). However, to the authors knowledge, a study investigating measured microscopic vehicle movement (measured on a second by second basis using GPS) in tandem with microscopic emissions modeling has not been completed successfully for a substantially sized vehicle fleet and study domain, rather have been limited to small ad hoc deployments.

The methodology described in this study has the potential to inform environmental policy related to transportation in urban areas. With the framework proposed, where appropriate data is available, responsive and adaptive strategies could be implemented should the emissions model be applied using real-time GPS data. The methodology described demonstrated the potential for linking GPS measured vehicle movements directly with microscopic emissions models (based on the instantaneous driving speed and acceleration) for quantifying traffic emissions. Although the computation of emissions is clearly a useful application, it is in the implementation and evaluation of real-time, technology-based environmental policies related to transportation where its application would be most beneficial. Technologies for monitoring and controlling vehicle velocities include Intelligent Speed Adaption (ISA) systems (Duynstee et al., 2001; Int Panis et al., 2006). ISA systems are electronic systems installed in vehicles, which utilize GPS navigation to evaluate the vehicle location and apply appropriate speed limits on specific road segments ISA systems combined with an appropriate real-time emissions model could be utilized for minimizing emissions and fuel consumption in urban road networks in the future.

Environment related transportation policies such as restricting vehicles in a city-center zone or restricting odd/even number plates in urban regions have been adopted in a number of cities in recent years (Fensterer et al., 2014; Holman et al., 2015). Whereas these have helped in the reduction of congestion and pollution levels in urban centers, more beneficial approaches may be based on the detection of the specific, fixed positions where emissions take place, rather than in substantial urban regions. With the dynamic fine grain emissions inventory presented in this study, it may become feasible to target air pollution emissions mitigation efforts in a far more direct manner. The health and economic benefits of reducing air pollution emissions across various sectors including transportation, thereby improving air quality, has been quantified in many reports. For example, the US EPA computed the costs for the implementation of the 1990 Clean Air Act to be about 65 million dollars, with a potential benefit reaching 2 trillion dollars from 1990 to 2020, potentially avoiding approximately 230,000 premature deaths in 2020 (USEPA, 2011).

For the first time, the data collected allow us to see an emission inventory not as something static which only changes from one road segment to the other, but which has more detailed characteristics with spatiotemporal variation. This enables a better estimate of the impact of pollution on the urban population which also exhibits variable spatial and temporal distribution profiles over the course of the day (Nyhan et al., forthcoming). The advantage of the proposed method is that by interrogating and interpreting easily accessible data from existing fleets (such as vehicle or bus services), considerable information regarding air pollution emissions can be obtained at a low cost and minimal effort in cities. Such a system can be applied in other cities, perhaps through government encouragement to make transportation GPS data available. This information may be of considerable value in determining the most appropriate locations of where to take action to reducing emissions and subsequently air pollution concentration levels in cities. This type of data could also be used to compute fine-grained fuel consumption patterns from the transportation sector.

This approach we adopted for predicting emissions has some limitations. In the development of the emissions model functions, Int Panis et al., (2006) primarily used measurements made in urban traffic (with low speeds) for determining functional forms and the variables in the emissions equations used in this study. This is considered sufficient for the purposes of evaluating the effects of speed management in urban networks. It is possible that the emission functions for highway traffic (at higher speeds) differ for those of urban traffic and the traffic on highways was insufficiently represented in the functions used. The emissions model did not allow for the specific model or age of the vehicles to be considered in computations either. Some additional measures would also be needed to verify the quality of the acceleration data obtained from GPS traces. There are inherent inaccuracies associated with GPS measurements, which however are compensated by the large volume of data collected. There is a necessity to connect the movements of the subset of vehicles with the movement of all the vehicles in the city. For this, calibrations parameters could be applied based on the sampling of the available vehicles versus the total number of vehicles. Finally, additional work would be needed to link the emissions predicted for various parameters to local measured air pollution concentration levels. A future study by these authors will therefore examine the relationship between predicted emissions using the methodology described herein this study, and measured or modeled values of air pollution concentrations.

This methodology described in this paper may be replicated in a number of cities worldwide, as GPS traces from vehicles become increasingly available. Vehicle fleet operators can do a major public service by providing GPS data for research, in particular for predicting emissions and other information relevant to environmental health from it. This information may be used for designing air pollution intervention strategies (long-term, short-term, responsive and adaptive) for the protection of human health and well-being.

5. Conclusions

Through analyzing GPS data from a large transportation fleet in Singapore, fine grained emissions were estimated in high spatial resolution. The emissions model was based on the inputs of velocity and acceleration parameters extracted from the data. Air pollution emissions related to CO_2, NO_x, VOC and PM were thereby quantified. The spatial distributions of the emissions were investigated thereby enabling highly localized areas of relatively higher emissions levels to be identified. This study also shows how the instantaneous drive cycles can be applied in the estimation of the overall emissions from the transportation sector within the study area.

7. Acknowledgements

All the authors wish to thank the MIT SENSEable City Lab Consortium and the Singapore-MIT Alliance for Research & Technology program for supporting the research. M. Nyhan would like to thank Fulbright and the Irish Environmental Protection Agency. The authors would also like to acknowledge Dr. Luc Int. Panis for providing advice on some modeling aspects of the study.

8. References

Akimoto, H., 2003. Global Air Quality and Pollution. Science 302, 1716-1719.

An, F., Barth, M., Ross, M., Norbeck, J., 1997. The development of a comprehensive modal emission model: operating under hot-stabilize conditions. Transportation Research Record. 1587, 52-62.

Aslam, J., Lim, S., Pan, X., Rus, D., 2012. City-scale traffic estimation from a roving sensor network. SenSys' 12, November 6-9, 2012. ACM 978-1-1169-4.

Barth, M., Scora, G., 2006. Comprehensive Modal Emission Model (CMEM), version 3.01 User's Guide. Riverside: University of California.

Bell, M.L., Davis, D.L., Fletcher, T., 2004. A Retrospective Assessment of Mortality from the London Smog Episode of 1952: The Role of Influenza and Pollution; Environmental Health Perspectives. 112, 6-8.

Ben-Akiva, M., Koutsopoulos, H.N., Mishalani, R., Yang, Q., 1997. Simulation laboratory for evaluating dynamic traffic management systems. Journal of Transportation Engineering. 123(4), 283-289.

Britter, R. E., and Hanna, S. R., 2003. Flow and dispersion in urban areas. Annual Review of Fluid Mechanics. 35, 469–496.

Burke, J. A., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S., and Srivastava, M. B., 2006. Participatory Sensing. UC Los Angeles: Center for Embedded Network Sensing. Retrieved from <http://escholarship.org/uc/item/19h777qd>

Caiazzo, F., Ashok, A., Waitz, I.A., Yim, S.H.L., and Barrett, S.R.H., 2013. Air pollution and early deaths in the United States. Part I: Quantifying the impact of major sectors in 2005. Atmospheric Environment. 79, 198-208.

Campbell, A. T., Eisenman, S. B., Lane, N. D., Miluzzo, E., Peterson, R. A., Lu, H., Zheng, X., Musolesi, M., Fodor, K. and Ahn, G-S., 2008. The Rise of People-Centric Sensing. IEEE Internet Computing. 12, 12-21.

Canagaranta, M. R., Jayne, J. T., Ghertner, D. A., Herndon, S., Shi, Q., Jimenez, J. L., Silva, P. J., Williams, P., Lanni, T., Drewnick, F., Demerjian, K. L., Kolb, C. E., and Worsnop, D. R., 2004. Case studies of particulate emissions from in-use New York City vehicles. Aerosol Science and Technology. 38, 555-573.

CERC, 2015. ADMS Urban, Cambridge Environmental Research Consultants, Cambridge, United Kingdom.

Chong, C-Y. and Kumar, S. P., 2003. Sensor networks: evolution, opportunities and challenges. Proceedings of the IEEE. 91, 1247-1256.

Cuff, D., Hansen, M., and Kang, J., 2008. Urban sensing: out of the woods. Communications of the ACM. 51, 24-33.

Duynstee, L., Katteler, H., Martens, A. Intelligent speed adaptation: selected results of the practical trial. Proc. 8th world congress on intelligent transport systems, Sydney, Australia, 30 September-4 October, 2001.

El-Sgawarby, I., Kyoungho, A., Rakha, H., 2005. Comparitive field evaluation of vehicle cruise speed and acceleration levels impacts on hot stabilized emissions. Transportation Research. 10D, 13-30.

Fecht, D., Hansell, A.L., Morley, D., Dajnak, D., Vienneau, D., Beevers, S., Toledano, M.B., Kelly, F.J., Anderson, H.R., Gulliver, J., 2016. Spatial and temporal associations of road traffic noise and air pollution in London: Implications for epidemiological studies. Environment International. 88, 235-242.

Fensterer, V., Kuchenhoff, H., Maier, V., Wichmann, H.-E., Breitner, S., Peters, A., Gu, J., Cyrys, J., 2014. Evaluation of the impact of low emission zone and heavy traffic-ban in Munich (Germany) on the reduction of PM_{10} in ambient air. International Journal of Environment and Public Health. 11(5), 5094-5112.

Fontes, T., Pereira, S.R., Fernandes, P., Bandeira, J.M., Coelho, M.C., 2015. How to combine different micro-simulation tools to assess the environmental impacts of road traffic? Lessons and directions. Transportation Research Part D. 34, 293-306.

Frey, H. C., Unal, A., Rouphail, N. M. and Colyar, J. D., 2003. On-road measurement of vehicle tailpipe emissions using a portable instrument. Journal of the Air & Waste Management Association. 53, 992–1002.

Fruin, S., Westerdahl, D., Sax, T., Sioutas, C., Fine, P.M., 2008. Measurements and predictors of on-road ultrafine particle concentrations and associated pollutants in Los Angeles. Atmospheric Environment. 42(2), 207-219.

González, M.C., Hidalgo, C.A., Barabási, A.L., 2008. Understanding individual human mobility patterns. Nature. 453, 779–782.

Grauwin, S., Sobolevsky, S., Moritz, S., Godor, I., Ratti, C., 2014. Towards a comparative science of cities: using mobile traffic records in New York, London and Hong Kong. Computational Approaches for Urban Environments, Geo-technologies and the Environment. 13, 363–387.

Gurhar, B.R., Jain, A., Sharma, A., Agarwal, A., Gupta, P., Nagpure, A.S. Lelieveld, J., 2010. Human health risks in megacities due to air pollution. Atmospheric Environment. 44, 4606-4613.

Guzella, L. and Sciarreta, A., 2005. Vehicle propulsion systems: introduction to modeling and optimization. Springer Verlag.

Holman, C., Harrison, R., Querol, X., 2015. Review of the efficacy of low emission zones to improve urban air quality in European cities. Atmospheric Environment. 111, 161-169.

Hoteit, S., Secci, S., Sobolevsky, S., Ratti, C., Pujolle, G., 2014. Estimating human trajectories and hotspots through mobile phone data. Computer Networks. 64, 296-307.

Hu, T.-Y., and Mahmassani, H.S., 1997. Day-to-day evolution of network flows under real-time information and reactive signal control. Transportation Research. 5C(1), 51-69.

Int Panis, L., Broekx, S., Liu, R., 2006. Modeling instantaneous traffic emission and the influence of traffic speed limits. Science of the Total Environment. 371, 270-285.

Jerrett, M., Burnett, R.T., Ma, R., Pope, C.A., III, Krewski, D., Newbold, K.B., Thurston, G., Shi, Y., Finkelstein, N., Calle, E.E., Thun, M.J., 2005. Spatial Analysis of Air Pollution and Mortality in Los Angeles. Epidemiology. 16, 727-736.

Kang, C., Sobolevsky, S., Liu, Y., Ratti, C., 2013. Exploring human movements in Singapore: A comparative analysis based on mobile phone and taxicab usages. In Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing (p. 1). ACM.

Kumar, P., Morawska, L., Martani, C., Biskos, G., Neophytou, M., Di Sabatino, S., Bell, M., Norford, L., Britter, R., 2015. The rise of low-cost sensing for managing air pollution in cities. Environment International. 75, 199-205.

Kung, K.S., Greco, K., Sobolevsky, S., Ratti, C., 2014. Exploring universal patterns in human home-work commuting from mobile phone data. PLoS ONE. 9(6): e96180. doi:10.1371/journal.pone.0096180

Krewski, D., Jerrett, M., Burnett, R.T., Ma, R., Hughes, E., Shi, Y., Turner, M.C., Pope, C.A. 3rd, Thurston, G., Calle, E.E., Thun, M.J., Beckerman, B., DeLuca, P., Finkelstein, N., Ito, K., Moore, D.K., Newbold, K.B., Ramsay, T., Ross, Z., Shin, H., Tempalski, B., 2009. Extended follow-up and spatial analysis of the American Cancer Society study linking particulate matter and mortality. Research Report from the Health Effects Institute: 5-114; Discussion 115-136.

Laden, F., Schwartz, J., Speizer, F.E., Dockery, D.W., 2006. Reduction in fine particulate air pollution and mortality extended follow-up of the Harvard six cities study. American Journal of Respiratory and Critical Care Medicine. 173, 667-672.

Land Transport Authority of Singapore, 2015. Transportation fleet composition datasets available at <data.gov.sg>

Laxen, D., Wilson, P., Marner, B., Moorcroft, S., and Brown Y., 2003. Review of Air Quality Monitoring in London. Report for the Mayor, the Association of London Government (ALG) and the Department for Environment Food and Rural Affairs (DEFRA). Air Quality Consultants: Bristol.

Lewtas, J., 2007. Air pollution combustion emissions: Characterization of causative agents and mechanisms associated with cancer, reproductive, and cardiovascular effects. Mutation Research. 636, 95-133.

Liu, R., van Vliet, D., Watling, D., 2006. Micro-simulation models incorporating both demand and supply dynamics. Transportation Research. 40A(2), 125-50.

Mantazeri-Gh, M. and Naghizadeh, M., 2003. Development of car drive cycle for simulation of emissions and fuel economy. Proceedings of the 15th European Simulation Symposium. 1-6.

Metric Mind Corporation, 2012. Data drive cycles. Retrieved from <http://www.metricmind.com/data/cycles.pdf>

Molina, L. T., Molina, M. J., Slott, R. S., Kolb, C. E., Gbor, P. K., Meng, F., Singh, R. B., Galvez, O., Sloan, J. J., Anderson, W. P., Tang, X., Hu, M., Xie, S., Shao, M., Zhu, T., Zhang, Y. H., Gurjar, B. R., Artaxo, P. E., Oyola, P., Gramsch, E., Hidalgo, D., and Gertler, A. W.,

2004. Air Quality in Selected Megacities. Journal of the Air & Waste Management Association. 54(12), 1-73.

Molina, M. J. and Molina, L. T., 2004. Megacities and atmospheric pollution. Journal of Air and Waste Management. 54, 644-680.

Moretti, F., Pizzuti, S., Panzieri, S., Annunziato, M., 2015. Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling. Neurocomputing, 167, 3-7.

NARSTO, 2005. Improving emission inventories for effective air quality management across North America. NARSTO 05-001. Pasco, Washington, USA.

NEA (National Environment Agency), 2015. Available at <www.nea.gov.sg>

North, R. J., Noland, R. B., Ochieng, W. Y., and Polak, J. W., 2006. Modeling of particulate matter mass emissions from a light- duty diesel vehicle. Transportation Research Part D: Transport and Environment. 11, 344-357.

Nyhan, M., Misstear, B.D., McNabola, A., 2014a. Comparison of particulate matter dose and acute heart rate variability response in cyclists, pedestrians, bus and train passengers. Science of the Total Environment. 468-469, 821-831.

Nyhan, M., Misstear, B., McNabola, A., 2014b. Evaluating artificial neural networks for predicting minute ventilation and lung deposited dose in commuting cyclists. Journal of Transport & Health. 1(4), 305-315.

Nyhan, M., Britter, R., Grauwin, S., Laden, F., McNabola, A., Misstear, B., Ratti, C., forthcoming. Exposure Track – the impact of mobile device based mobility patterns on quantifying population exposure to air pollution.

Osorio, C., Nanduri, K., 2015. Urban transportation emissions mitigation: Coupling high-resolution vehicular emissions and traffic models for traffic signal optimization. Transportation Research Part B. Accepted and in press.

Paulos, E., Honicky, R.J., and Hooker, B., 2009. Citizen Science: Enabling Participatory Urbanism. In Foth, M. (Ed.) Urban Infomatics: The practice and promise of the real-time city. IGI Global, Hershey, PA.

Parkinson B.W., 1996. GPS error analysis, Global Positioning System: Theory and applications. Progress in Astronautics and Aeronautics. 163, 469-483.

Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.L., Li, T., Zhou, C., 2014. A new insight into land use classification based on aggregated mobile phone data. International Journal of Geographical Information Science. 28(9) p.1-20.

Pelkmans, L., Debal, P., Hood, T., Hauser, G., Delgado, M.R., 2004. Development of a simulation tool to calculate fuel consumption and emissions of vehicles operating in dynamic conditions. SAE 2004 spring fuels and lubricants, 2004-01(1873). Warrendale PA, (USA): SAE International (Society of Automotive Engineers).

Rakha, H., Kyoungha, A., Trani, A., 2004. Development of a VT-Micro model for estimating hot stabilized light-duty vehicle and truck emissions. Tranportation Research. 9D, 49-74.

Samuel, S., Austin, L. and Morrey, D., 2002. Automotive test drive cycles for emission measurement and real-world emission levels- a review. Proceedings of the Institution of Mechanical Engineers, Part D: Jouranl of Automobile Engineering. 216, 555-564.

Santi, P., Resta, G., Szell, M., Sobolevsky, S., Strogatz, S., Ratti, C., 2014. Quantifying the benefits of vehicle pooling with sharability networks. Proceedings of the National Academy of Sciences. 111(37), 13290-13294.

Singapore Department of Statistics, 2015. Yearbook of Statistics Singapore 2015. Available at <www.singstat.gov.sg>

Shorter, J., Herndon, S., Zahniser, M. S., Nelson, D. D., Womrhoudt, J., Demerjian, K. L. and Kolb, C. E., 2005. Real-time measurements of nitrogen oxide emissions from in-use New

York City transit buses using a chase vehicle. Environmental Science and Technology. 39, 7991- 8000.

Sobolevsky, S., Szell, M., Campari, R., Couronné, T., Smoreda, Z., Ratti, C., 2013. Delineating geographical regions with networks of human interactions in an extensive set of countries. PLoS ONE. 8(12), e81707 doi: 10.1371/journal.pone.0081707.

Streets, D., Lu., Z., 2012. Anthropogenic emissions inventory for 5 major sectors for South East Asia and China. Argonne National Laboratory for the National Aeronautics and Space Administration (NASA) SEAC4RS project.

Thiagarajan, A., Ravindranath, L., LaCurts, K., Madden, S., Balakrishnan, H, Toledo, S. and Eriksson, J., 2009. VTrack: Accurate, energy-aware road traffic delay estimation using mobile phones. SenSys'09, Berkley, CA, USA.

USEPA (United States Environmental Protection Agency), 2009. AERMOD Modeling System. Available at www.epa.gov

USEPA (United States Environmental Protection Agency), 2011. The Benefits and Costs of the Clean Air Act from 1990 to 2020. Final Report-Rev. A. US Environmental Protection Agency Office for Air and Radiation.

USEPA (United States Environmental Protection Agency), 2012. MOBILE Model documentation. Available at <www.epa.gov>

Vardoukalis, S., Gonzalez-Flesca, N., Fisher, B. E. A., and Pericleous, K., 2005. Spatial variability of air pollution in the vicinity of a permanent monitoring station in central Paris. Atmospheric Environment. 39, 2725-2736.

WHO (World Health Organization), 2014. Urban Population Growth, Global Health Observatory. Available at <www.who.int>

WHO (World Health Organization), 2014. Burden of disease from ambient air pollution for 2012. Available at <www.who.int>

Zhang, J. and Morawska, L., 2002. Combustion sources of particles: 2. Emission factors, and measurement methods. Chemosphere. 49, 1059-1074.

List of Tables

Table 1. Motor vehicle population in Singapore by category and type of fuel used for the year 2011. Figures exclude tax exempted vehicles for off-the-road use (RU plates).

Table 2. Modeled emissions for the taxi fleet and the proportion of modeled taxi emissions in the estimated total motor vehicle population emissions, for each of four air pollutant parameters.

Table 3. Comparison of the mean daily emissions predicted for the total motor vehicle population of Singapore to estimated ground transportation emissions attained from Streets and Lu (2012).

List of Figures

Figure 1. Distribution of traffic volumes (i.e. number of vehicles per road segment) on 15 randomly selected road segments for the 23rd February 2011. The x-axis includes 15 road segments including a point for every 15 minutes during the 24-hour day. The y-axis represents the percentage of traffic at that location and time. The taxi distribution (in blue) underestimates the loop distribution (in green) and the underestimation is variable.

Figure 2. Results of the feed-forward artificial neural network model implementation. Regression plot of a partial set of modeled traffic volumes versus corresponding measured traffic volume for the (a) training phase (R^2=96%), (b) validation phase (R^2=93%), (c) tesing phase (R^2=92%) and (d) overall model (R^2=94%). A sub-sample of points are presented for clarity.

Figure 3. Spatial distribution of the number of valid accelerations in Singapore on the 23rd February 2011. Locations where relatively higher numbers of valid accelerations are observed in the vicinity of the Singapore Downtown Core area and the Changi International Airport in the east.

Figure 4. Spatial distributions of predicted daily emissions from the vehicle fleet for each road link for the parameters of (a) CO_2 (tonnes/day), (b) NO_x (tonnes/day), (c) VOC (g/day), and (d) PM (g/day) in Singapore on the 23rd February 2011. Locations of relatively high-

emissions, are observed in the Singapore Downtown Core area in the south-center of Singapore and in other locations throughout the island.

Highlights

We present a novel method for predicting air pollution emissions using transport data

Study uses measured microscopic transport data and a microscopic emissions model

GPS data from over 15000 vehicles were analyzed to quantify speeds and accelerations

CO_2, NO_x, VOCs and PM were modeled in high spatio-temporal resolution

Highly localized areas of elevated emissions were identified